EPIDEMIOLOGICAL INVESTIGATION OF INFANT HEALTH AND ENVIRONMENTAL EXPOSURES

by

Shaina L. Stacy

B.S. Chemistry, Carnegie Mellon University, 2010

M.P.H. Environmental and Occupational Health, University of Pittsburgh, 2012

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Shaina L. Stacy

It was defended on

April 17, 2015

and approved by

Vincent C. Arena, Ph.D., Associate Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Aaron Barchowsky, Ph.D., Professor, Department of Environmental and Occupational Health, Graduate School of Public Health, University of Pittsburgh

LuAnn L. Brink, Ph.D., Adjunct Assistant Professor, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

Chair: Bruce R. Pitt, Ph.D., Professor and Chair, Department of Environmental and Occupational Health, Graduate School of Public Health, University of Pittsburgh

Dissertation Advisor/Co-Chair: Evelyn O. Talbott, Dr.PH., Professor, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

Bruce R. Pitt, Ph.D.
Evelyn O. Talbott, Dr.PH.

EPIDEMIOLOGICAL INVESTIGATION OF INFANT HEALTH AND ENVIRONMENTAL EXPOSURES

Shaina L. Stacy, PhD

University of Pittsburgh, 2015

ABSTRACT

The impact of "place" on health is a classical and key element of epidemiology. Recent advances in geographic information systems have facilitated the use of spatial methods to investigate public health issues. Such approaches are particularly helpful when a public health phenomenon is relatively new and adequate environmental exposure information is lacking. The overarching objective of the present epidemiological investigation is to use spatial methods to explore relationships between several infant and children's health outcomes and potential environmental exposures. The public health significance of this work is to identify possible sources of harmful exposures that may motivate further research, primary prevention efforts, and eventually policies to further limit exposures in these sensitive populations. It is well known that the embryo/fetus is particularly sensitive to the effects of environmental agents. Early life exposures are of public health significance since they may harm infant health and also have further adverse consequences in childhood and adulthood. The present work encompasses two relatively new but growing areas of interest related to fetal, infant, and children's health: 1) unconventional natural gas development and adverse birth outcomes, and 2) sources of air toxics and childhood autism spectrum disorder (ASD). Geographic information systems are used to spatially link health outcomes, including birth weight, small for gestational age, preterm birth, and ASD, with nearby

sources or with aggregated (e.g., census tract-level) estimates of exposure. Logistic regression is conducted to determine associations between risk for each of the above health effects and the exposures of interest, adjusting for other sociodemographic and personal risk factors. Overall, results indicate that environmental factors have a small but important role to play in the health of infants and children, even after accounting for other potentially confounding factors. Since spatial surrogates for exposure are the primary focus of this investigation, future work will benefit from improved individual exposure assessment and a prospective study design to confirm and further explain these associations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Funding

The Heinz Endowments, C1697, Perinatal Outcomes and Marcellus Shale (PI: Dr. Bruce Pitt)

The Heinz Endowments, C1627, Case-Control Study of Personal and Environmental Risk Factors for Childhood Autism in Southwestern Pennsylvania (PI: Dr. Evelyn Talbott)

Acknowledgments

I would like to thank my dissertation committee, Dr. Vincent Arena, Dr. Aaron Barchowsky, Dr. LuAnn Brink, Dr. Bruce Pitt, and especially my advisor Dr. Evelyn Talbott, for their continued guidance and support. Their knowledge and valuable insights have enabled my success in the program. I would like to extend a huge thank you to Dr. Talbott for her mentorship over the last several years. Under her advisement, I was able to pursue my passion for environmental health issues and to grow as a scientist. Dr. Pitt has also been instrumental in guiding me through the PhD program in Environmental and Occupational Health and for helping me to find a research home with Dr. Talbott and Dr. Brink.

I would also like to thank my family for their continued support throughout this journey. Thank you to my mother, Mrs. Helen Stacy, for instilling in me the importance of education and encouraging me to pursue my dreams (and for always sending me back to Pittsburgh with a lot of delicious, home-cooked food!). Thank you to my grandmother, Mrs. Helen Gruber, who sat with

me and taught me to tell time and memorize the state capitals when I was little and also instilled in me a love of learning. I know you and Pappy are watching from above and would be proud. Thanks also to Mrs. Patricia Conrad for being an amazing mentor and friend.

I would like to extend my thanks and gratitude toward many others who have helped shape this journey: Dr. James Peterson, my MPH advisor; Dr. Felicia Wu, my master's thesis advisor and a valuable mentor early in the program; the Student Affairs staff, particularly Joan, Caitlin, and Robin for their continued support and friendship; Anna, Juley, and Judy for all of their help and for making 507 Parran a fun place to work; and the faculty and staff of the Department of Environmental and Occupational Health for their assistance and support.

Last but not least, I would like to thank my friends for letting me vent, making me laugh, and always being there. I feel so blessed to have gone on this journey with a group of such amazing, intelligent, and passionate people who also happen to like Harry Potter as much as I do.

List of abbreviations

Health effects: ASD (autism spectrum disorder), IUGR (intrauterine growth restriction), LBW (low birth weight), SGA (small for gestational age)

Exposure: AADT (annual average daily traffic), HAP (hazardous air pollutant), NATA (National Air Toxics Assessment), NMR (nearest major road), PERC (perchloroethylene), PM (particulate matter), TRI (Toxics Release Inventory), UGD (unconventional natural gas development), VOC (volatile organic compound)

Statistics: CI (confidence interval), EFA (exploratory factor analysis), IDW (inverse distance weighting), IQR (interquartile range), LUR (land use regression), OR (odds ratio)

1.0 INTRODUCTION

The impact of “place” on health is a classical and key element of epidemiology [1]. There has been renewed interest in place over the last several decades, likely due to improvements in personal computing and the accessibility of geographic information systems (GIS) software [1]. When a public health phenomenon is new, aspects of place, such as proximity to a potential source of exposure, can serve as a useful surrogate. The following epidemiological investigation explores several infant and children’s health issues from a place perspective.

The embryo/fetus is uniquely vulnerable to toxic chemicals in the environment [2; 3]. Major windows of developmental vulnerability exist prenatally and during infancy and early childhood [3]. Early life exposures are not only detrimental to the infant herself, but can impair health further down the road in childhood and even adulthood [2; 4]. For these reasons, these sensitive populations are of particular interest in studies of environmental health effects, and such investigations have covered numerous potential exposures, pathways, and outcomes. The present work focuses on two relatively new but growing areas of interest in this field: 1) unconventional natural gas development (UGD) and adverse birth outcomes, and 2) ambient air toxics and childhood autism spectrum disorder (ASD). The first investigation helps to address the uncertainties surrounding the potential health effects of novel natural gas extraction methods, while the second seeks to clarify potential environmental contributions to ASD, many of which

are still unknown. Both topics have also received their share of media attention and public scrutiny.

Due to the current lack of adequate exposure data to link with these health outcomes, spatial approaches will be used to explore the relationships between UGD and adverse birth outcomes (Specific Aim 1), and between sources of air toxics and ASD (Specific Aims 2a and b). These spatial methods, such as proximity and aggregation, will serve as surrogates or estimates of exposure. This investigation will address the following specific aims:

- **Specific Aim 1:** To assess the impact of unconventional natural gas development on infant health in southwestern Pennsylvania using well density as a surrogate for exposure.

  Hypothesis: The risk for adverse birth outcomes will be greater for those infants born to mothers living in more densely drilled areas.
- **Specific Aim 2a:** To explore associations between ASD risk and groups of air toxics using exploratory factor analysis. This type of factor analysis is primarily hypothesis-generating and is not driven by *a priori* expectations.
- **Specific Aim 2b:** To explore associations between ASD risk and proximity to major sources of environmental styrene exposure, i.e. industrial and traffic.

  Hypothesis: Increased ASD risk will be associated with living near major industrial sources of styrene and near major roadways, after taking into account maternal risk factors for ASD.

The remainder of this section provides further background on these topics, broken down by specific aim, to lay the foundation of the three chapters that follow.

1.1 ADVERSE BIRTH OUTCOMES AND UNCONVENTIONAL NATURAL GAS DEVELOPMENT

There is a vast literature regarding environmental risk factors for adverse birth outcomes such as preterm birth, low birth weight (LBW), and small for gestational age (SGA). Each of these outcomes is also related to a variety of genetic, behavioral, and sociodemographic risk factors, so elucidating environmental contributions is a difficult task. Multiple factors may interact to cause or increase risk for the outcome. Nevertheless, clarifying associations between these infant health effects and suspected environmental agents is an important step, especially in the context of new or increased exposures to toxic compounds, such as those that might occur in areas where UGD activities are increasing. The following section first defines each outcome of interest (preterm birth, LBW, and SGA) and summarizes the genetic, environmental, and other contributions to their development. Next, potential exposures related to UGD activities and their relationships to these health effects are described.

1.1.1 Prematurity.

An infant is considered preterm or premature if born before 37 weeks of gestation [5]. In the United States, preterm birth affected 1 out of 9 infants in 2012, and 35% of infant deaths in 2010 were due to causes related to preterm birth [6]. Preterm birth complications are the primary cause of death in children under five years of age worldwide [5]. Premature infants are at an increased risk of a variety of further complications, including respiratory, gastrointestinal, and neurodevelopmental [7].

There are three “types” of premature birth: those due to induced labor or prelabor caesarean section, which compose about 30-35% of premature births, and spontaneous preterm labor with and without intact membranes, making up 40-45% and 25-30% of preterm births, respectively [7]. In the United States, the rate of preterm births increased from 9.5% in 1981 to 12.8% in 2006 [8], mostly due to early obstetric intervention and increases in preterm multiple births related to assisted reproductive technologies [7]. The rate has since been on the decline to 11.4% in 2013, for both multiple and singleton births [8]. This decline has also been observed in singleton births [8].

Goldenberg et al. (2008) have reviewed a variety of factors associated preterm birth, which include everything from biological (e.g., chemokines and cytokines) [9] and genetic (polymorphisms in the *TNFα* and *IL6* genes) markers [10; 11], to bacterial intrauterine infection [12], to a host of maternal demographic and behavioral factors that increase the risk for premature birth, among them low socioeconomic and educational status [13; 14], very young or old maternal age [15], race [16; 17], psychosocial stress [17], smoking [18], and low pre-pregnancy BMI [19].

Many studies have also found an environmental component to the risk for premature birth, even after accounting for the sociodemographic and behavioral factors listed above. In epidemiology studies, researchers have found associations between premature birth and ambient levels of air pollutants, such as carbon monoxide (CO) [20; 21], sulfur dioxide (SO_2) [20; 22; 23], particulate matter (PM) [22-28], polycyclic aromatic hydrocarbons (PAHs) [27; 29], and mixtures of pollutants [30]. Other studies have found no effect for CO, nitrogen dioxide (NO_2), or ozone (O_3) [25]. Relatively fewer studies have included preterm birth as an endpoint (as compared to, for example, birth weight), and there have been inconsistencies in the roles of

individual pollutants as well as timing of exposures, leading one review to conclude that there is not yet enough evidence to support a causal role of any one air pollutant for preterm birth [31]. As reviewed in Stillerman et al. [32], associations between prematurity and other environmental exposures, namely polychlorinated biphenyls (PCBs) [33] and certain pesticides and herbicides [34; 35], have also been observed, but these are outside the scope of the present work.

1.1.2 Low birth weight and small for gestational age.

Low birth weight (LBW) infants weigh less than 2,500 g (5 pounds 8 ounces), regardless of gestational age; term low birth weight (TLBW) infants weigh less than this threshold at term (37-42 weeks) [36]. The rise in preterm births described in the previous section in part contributed to a 20% increase in the rate of low birth weight in the United States from 1990-2006 [8]. The LBW rate has trended slightly downward in recent years and was 8.0% in 2013 [8]. LBW is one of the main risk factors for infant morbidity and mortality and the second major cause of death in the perinatal period [36].

Small for gestational age (SGA) is another measure used to describe unusually small infants. It is often incorrectly used interchangeably with the term intrauterine growth restriction (IUGR). An infant is considered to be SGA if he or she weighs below the 10th percentile expected for their sex and gestational age [37]. IUGR, on the other hand, is a condition in which the fetus does not reach its genetic growth potential [36]. An IUGR infant would not necessarily also be classified as SGA, and vice versa [36]. Restricted fetal growth has been linked to poor chronic health outcomes in adulthood (the Barker Hypothesis), including coronary heart disease, stroke, diabetes, and hypertension [4; 38].

It is thought that heredity contributes 40% to birth weight while environmental factors make up the remaining 60% [36]. The genetic, maternal, sociodemographic, and behavioral factors linked to low birth weight have been reviewed by Bernabé et al. (2004). Fetal growth restriction has been associated with chromosomal duplications, deletions, translocations, and other anomalies, such as Turner's syndrome and trisomies 13, 18, and 21 [36]. Male infants tend to weigh 150 to 200 g more than females [36]. Pre-pregnancy weight is also an important predictor; smaller mothers tend to have smaller babies [39]. Many of the risk factors associated with premature birth are also related to LBW, including mother's age (very young and older mothers) [15], marital status [40], socioeconomic status [41], race [16], education [42], and inadequate prenatal care [43]. Low birth weight and IUGR are also associated with maternal cigarette smoking [18; 44], alcohol consumption and drug abuse [44-46], as well as caffeine intake [47] and psychosocial stress [44; 48]. These sociodemographic and behavioral factors are in turn related to each other, so sometimes it is difficult to discern whether, for example, adolescent mothers tend to have LBW infants due to an underlying biological factor or due to related issues like living in poverty or poor access to prenatal care [36].

A variety of other factors, from concomitant disease or infection to characteristics of the pregnancy itself, can also affect birth weight. For example, gestational diabetes may lead to an infant having excessive birth weight or, conversely, premature birth [49; 50]. Reduced fetal growth may result from intrauterine infections caused by microorganisms crossing the placenta [36]. Several studies have found that short intervals between births increase the risk for having an LBW or preterm baby [51; 52]. Lastly, multiple pregnancies may lead to IUGR [53; 54]. Since twins, triplets, etc. are often at a greater risk for restricted fetal growth and preterm birth

than singleton births, they are typically excluded from studies of environmental effects on adverse birth outcomes [8].

An overwhelming number of epidemiology studies have found evidence linking exposure to ambient air pollution and LBW, SGA, and other measures of reduced fetal growth. As with premature birth, associations have been observed for gaseous pollutants like CO [55-58], SO_2 [22; 55; 56; 59-61], and nitrogen oxides [20; 55; 59], in addition to particulates [22; 28; 55; 61-68], PAHs [29; 63; 69; 70], and mixtures of air pollutants [71; 72]. A few studies have observed no effect on the risk of LBW for nitrogen oxides, ozone, or PM_{10} (particulate matter with an aerodynamic diameter less than 10 μm) [22; 57]. In their review of air pollution and adverse pregnancy outcomes, Šrám et al. (2005) concluded that the combined evidence suggests causality, although further studies are needed to confirm this. Other environmental and occupational exposures that may increase the risk for LBW or SGA are PCBs [33], certain herbicides and insecticides [32; 34; 35], heavy metals [73-75], and benzene [76].

1.1.3 Biological mechanisms.

Several biological mechanisms for exposure to ambient air pollution leading to preterm birth and LBW have been proposed, although they are still not well understood. There are five major possible biological mechanisms for the effect of air pollution on cardiopulmonary outcomes that could also reasonably explain effects on LBW: oxidative stress, inflammation, coagulation, endothelial function, and hemodynamic responses [77].

Particulate matter is composed of a variety of compounds, including combustion products, PAHs, and transition metals, that could be responsible for its adverse effects on the developing fetus [78-81]. PAHs may result in LBW or IUGR [78; 80; 81] through increased

DNA adducts and damage that lead to the activation of apoptotic pathways [82]. In utero exposures to transition metals may produce oxidative stress that damages DNA, interrupting DNA transcription and also increasing placental DNA adducts [77]. PM may also act through an inflammatory pathway associated with inadequate placental perfusion that disrupts oxygenation of maternal blood, nutrient exchange between the mother and the fetus, or another process [77].

The remaining possible biological mechanisms for the effects of PM on the fetus—blood coagulation, endothelial function, and hemodynamic (blood flow and circulation) responses—have been studied less extensively in direct relation to adverse birth outcomes, but the pathways are biologically plausible and future work may elucidate them [77]. PM may induce the release of mediators that increase blood coagulability and thus viscosity, leading to adverse effects on placental functions [67; 83]. Studies of environmental tobacco smoke, which is composed of PM and other toxics, have found that constriction of the blood vessels leads to increased plasma endothelin levels and endothelial dysfunction [77; 84]. Finally, exposure to PM could indirectly influence IUGR and preterm birth through increases in blood pressure leading to gestational hypertension [77]. The biological mechanism leading to these outcomes may vary according to the timing of the exposure during pregnancy, and the component of PM responsible for them may differ between geographic regions, since the chemical composition of PM depends on the major sources of air pollution in the area [67]. These mechanisms may also be further modified by other factors, namely the mother's nutritional status [77].

1.1.4 UGD exposures.

Studies regarding the environmental and human health effects of the relatively recent unconventional natural gas development (UGD) boom in Pennsylvania have been limited. The

current lack of exposure data related to UGD activities complicates our ability to assess possible health outcomes. Here, the small but growing body of literature regarding possible UGD exposures is reviewed, and their potential for impacting infant health is described.

The use of novel gas extraction methods has flourished over the last decade to exploit the resources of the Marcellus Shale, a rock formation underlying much of Pennsylvania and parts of several other states. The Marcellus Shale reaches depths of up to about 8,000 feet and is believed to hold trillions of cubic feet of natural gas [85]. Horizontal drilling and hydraulic fracturing have facilitated the extraction of natural gas at these previously unattainable depths [86]. Hydraulic fracturing involves the pumping of large amounts of water mixed with sand and other chemicals under high pressure to fracture the shale around a well, allowing the natural gas to flow freely [86]. The chemicals required serve a variety of functions, including but not limited to biocides, corrosion inhibitors, gelling agents, surfactants, and pH adjusters [86]. Pollutants relevant to public health may be released to air and ground or surface waters at multiple points along the natural gas production chain. These releases are described in the next two sections.

1.1.4.1 Releases to air.

Air pollutants, including NO_x, $PM_{2.5}$, and VOCs, are emitted by: 1) diesel-powered drill rigs and hydraulic fracturing pumps during well drilling, 2) natural-gas-fired compressors used to maintain gas pressure, and 3) trucks transporting materials to and from drilling sites [87]. After the well has been drilled, gases are allowed to escape from the wellbore (completion venting), another significant source of VOCs [87]. Burning off or flaring the gases reduces VOC emissions but instead releases NO_x and CO [87]. As reviewed earlier, associations have been found between this suite of air pollutants and adverse birth outcomes like preterm birth and low birth weight.

Roy et al. developed an air emission inventory for UGD activities in the Marcellus Shale region, estimating emissions of NO_x, $PM_{2.5}$, and VOCs for a base year (2009) and projecting out to 2020. Their models took into account all major UGD sources of these pollutants (drilling, hydraulic fracturing, completion venting, compressors, and truck traffic) as well as the use of emission-control technologies [87]. The authors anticipate that overall Marcellus-wide emissions will increase substantially from 2009 to 2020 [87]. According to their models, UGD activities will contribute 12% (6-18%) of regional NO_x and 12% (7%-28%) of regional VOC emissions in 2020 [87]. UGD contributed negligibly to future regional $PM_{2.5}$ emissions, although they did find that certain PM components, namely elemental carbon, could contribute 14% (2-36%) of regional emissions. These predictions took into account recently implemented regulations for off-road diesel engines and the EPA's Oil and Gas Rule.

1.1.4.2 Releases to water.

Risks of UGD activities to otherwise potable groundwater and surface water include contamination from land disturbances, leaks, spills, and disposal of inadequately treated wastewater, as well as accumulation of contaminants in subsurface soil or stream sediments [88]. The greatest potential for water contamination perhaps accompanies the challenge of UGD flowback water storage, treatment, and disposal. When the pressure in the gas well is released, allowing the natural gas to flow freely back up the well, it brings with it flowback or processed water. Flowback fluid contains oil or natural gas hydrocarbon products, the chemical additives required to fracture the shale, and naturally-occurring contaminants from the shale layer itself. Unlike in the western US, where flowback fluids are usually disposed of in deep underground injection wells, UGD wastewaters in the Marcellus Shale used to be processed predominantly by municipal sewage treatment plants and industrial wastewater treatment plants. The effluent was

then discharged into surface waters, leading to possible contamination if inadequately treated [89; 90].

Several studies have noted elevated levels of metals, salts, and radioactive isotopes in UGD produced waters [89; 91; 92]. In Pennsylvania, one study found concentrations of barium, magnesium, manganese, strontium, total dissolved solids, chlorides, bromides, sulfate, and benzene elevated above water quality criteria at three wastewater treatment plant discharge sites that processed UGD wastewater [89]. Since flowback fluids contain hydrocarbons, they can also be a source of air pollution. One investigation found that workers involved with UGD wastewaters could be exposed to levels of benzene, which has been associated with birth defects [93], above allowable occupational health levels [94]. UGD has also resulted in methane contaminated drinking water sources [95].

Mitigating the effects of UGD on drinking water sources may continue to prove challenging in the Marcellus Shale region. In addition to the contamination of potable water by chemical additives, metals, salts, etc., the acquisition of the large volumes of water needed for fracking fluids from nearby groundwater and surface water sources may lead to drinking water shortages [88]. In Pennsylvania, a moratorium on treating flowback water at publically owned treatment works has since been implemented. Alternatives to this have included transporting wastewater to Ohio for injection to deep underground wells as well as impoundments and recycling, but these in turn introduce additional opportunities for accidents, contamination, and human exposure [88; 89]. Acquiring good baseline data prior to drilling and using best practices to ensure well integrity and minimize accidental leaks and spills from drilling operations will all reduce the risk of ground- and surface water contamination [88]. Further, Marcellus produced

waters have a distinct strontium isotopic signature that may aid in the identification of sources and mechanisms of contamination [91].

1.1.4.3 Linkage to health effects.

To date, only a few published studies have attempted to link UGD exposures to human health endpoints. Many of these studies have used surrogates for exposure, such as proximity to gas wells, in the absence of adequate exposure data. Others have used actual measurements of pollutants associated with UGD and compared them to human health standards. These investigations are summarized below, with an emphasis on those relating to infant health.

McKenzie et al. conducted a human health risk assessment of air emissions from UGD in Garfield County, Colorado: a rural area where agriculture is the only other major industry. The authors measured ambient levels of a suite of hydrocarbons to estimate subchronic and chronic hazard indices as well as cancer risks [96]. Risk estimates were calculated for communities "near" (≤ 0.5 miles) and "far" (> 0.5 miles) from UGD. One study limitation might be the use of this 0.5-mile threshold, which was selected based on odor complaints attributed to UGD during only one summer. Nevertheless, residents living closer to UGD had greater total subchronic hazard indices and cumulative cancer risks than those living further away [96]. For residents living within 0.5 mile, both the chronic and subchronic non-cancer health indices were driven mostly by exposures to trimethylbenzenes, aliphatic hydrocarbons, and xylenes [96]. Benzene was the predominant contributor to cumulative cancer risk for both groups of residents [96]. Alternatively, a study conducted in the Barnett Shale region of north-central Texas, which compared maximum hourly and maximum 24-hour concentrations of 105 VOCs to applicable federal and state odor- and health-based standards, found very few instances in which VOC levels exceeded these standards [97]. Unlike the McKenzie et al. study, these authors only

compared ambient concentrations of VOCs to relevant health standards and did not compare potential health risks between those living close to and far from UGD.

In Pennsylvania, where UGD is still relatively new compared to Colorado and Texas, Rabinowitz et al. compared survey-based reported health symptoms between households living <1 km, 1-2 km, and >2 km from the nearest gas well in Washington County. The number of reported health symptoms per person, reported skin conditions, and upper respiratory symptoms were significantly greater in households <1 km compared to >2 km from UGD [98]. Other factors considered in their models included age, gender, household education, and smoking [98]. No associations were found between well proximity and neurological, cardiovascular, or gastrointestinal symptoms [98]. Another study of self-reported health impacts conducted in Pennsylvania found that participants attributed 59 unique health impacts and 13 stressors to Marcellus Shale development; stress was reported the most frequently [99].

Two studies (only one published) and one abstract have described work related specifically to perinatal health outcomes and the expansion of UGD operations [100-102]. These studies differed somewhat in their outcomes of interest and methodology, but, since much is still unknown about actual levels of pollutants associated with drilling activities, all used mother's residential proximity to natural gas development as a surrogate for exposure.

For her PhD thesis work, Hill (2012) explored associations between proximity to UGD and low birth weight (LBW), prematurity, and 5 minute APGAR score in all of Pennsylvania from 2003 to 2010. She used a difference-in-differences model to test an array of distances between 1 and 10.5 km, inclusive, to define living "near" UGD; any mother that had no wells within the selected distance was considered a control. Mothers living within a certain distance (5, 10 or 15 km) from permitted and drilled wells were also compared to mothers living near

permitted wells that had not been drilled within that timeframe. Overall, she found evidence for increased LBW and reduced APGAR scores for infants born to mothers living closer to UGD compared to control mothers [101]. Although the sample size was quite large (1,069,699 births), all Pennsylvania births over the study period were considered, including those in urban areas like Pittsburgh and Philadelphia. It would have been more prudent to exclude urban births due to differing maternal demographics and to reduce confounding by urban sources of pollution. Since Hill tested multiple distance thresholds between 1 and 10.5 km, her work also might suffer from a multiple comparisons problem; i.e., the more comparisons one makes, the more likely a significant result will be found simply due to chance. Alternatively, a buffer region could have been established *a priori*.

To date, McKenzie et al. (2014) has the only published study regarding perinatal outcomes and UGD. The authors performed a retrospective cohort study of 124,842 singleton births in rural areas of Colorado between 1996 and 2009. These years were chosen to capture the growth of natural gas development, which expanded rapidly in Colorado starting around 2000. Associations between maternal residential proximity to UGD and the following outcomes were examined: congenital heart defects (CHDs), neural tube defects (NTDs), oral clefts, preterm birth, and term low birth weight (TLBW). An inverse distance weighted (IDW) well count metric, which took into account both proximity to and density of wells around the mother's residence, was used as a surrogate for exposure. Each mother was assigned an IDW well count by summing the inverse distance between the mother's residence and each existing natural gas well within a 10-mile radius. Mothers were then divided into exposure tertiles (low: 1 to 3.62 wells per mile, medium: 3.63 to 125 wells per mile, and high: 126 to 1400 wells per mile), which were compared to a referent group of mothers that had no wells within a 10-mile radius, i.e. an

IDW well count of zero. Models were adjusted for maternal age, education, tobacco use, ethnicity, alcohol use, parity at time of pregnancy, infant gender, gestational age (for term birth weight), and elevation of maternal residence [102].

McKenzie and colleagues found increased prevalence of certain birth defects associated with natural gas development. They observed a monotonic increase in the prevalence of CHDs with increasing exposure to natural gas development in both unadjusted and adjusted models ($p<0.0001$). Infants born to the most exposed mothers had 30% greater prevalence of CHDs than those born to mothers with no wells within a 10-mile radius (OR=1.3, 95% CI=1.2-1.5). When the authors further probed different types of CHDs, they found increased prevalence of pulmonary artery and valve defects (60%), ventricular septal defects (50%), and tricuspid valve defects (400%) in the most exposed group compared to the control. Prevalence of NTDs was positively associated with increased IDW well count for only the most exposed group compared to the referent (OR=2.0, 95% CI=1.0-3.9). No statistically significant associations were found for oral clefts. In contrast to the results for CHDs and NTDs, decreasing trends were found for preterm birth and TLBW and increasing exposure to UGD ($p<0.0001$). These relationships remained virtually unchanged when the authors reduced the buffer around the mother's residence to five and then two miles and confined the analysis to years of the most rapid UGD expansion (2000 to 2009) [102].

The McKenzie et al. study was the first published work to investigate associations between UGD and perinatal health outcomes. Their use of an IDW well count metric is a well-established epidemiological approach to estimating exposure from multiple fixed point locations [103; 104]. Urban areas of Colorado, where pollution from other major sources may have confounded the analysis, were excluded from the study. They identified several possible "dose-

response" relationships between increased prevalence of two types of birth defects and increasing exposure. As with any study using proximity as a surrogate for exposure, the investigation was primarily hypothesis-generating. Other limitations include those inherent in the available birth and gas well data. For example, birth defects are most likely undercounted, especially in rural areas, and not all are confirmed by medical record review [102].

The next section reviews the literature concerning autism spectrum disorder and sources of ambient air toxics, including a summary of the potential neurotoxic effects of one particular pollutant of interest, styrene.

1.2 AUTISM SPECTRUM DISORDER AND AMBIENT AIR TOXICS

This section provides background on the second specific aim concerning environmental exposures and autism spectrum disorder (ASD). First, the diagnostic features and trends in prevalence of ASD will be described. Following will be a summary of the genetic, sociodemographic, and environmental factors thought to be associated with ASD. The section ends with a brief introduction to styrene, a potential neurotoxicant that may be associated with increased ASD risk.

1.2.1 Autism spectrum disorder.

Autism spectrum disorder (ASD) is a range of neurodevelopmental disorders characterized by impaired social interaction and communication and by restricted and repetitive behaviors [105]. Until recently, the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV)

classified these disorders into three subtypes of varying degrees of severity: autistic disorder, Asperger's syndrome, and pervasive development disorder/PDD not otherwise specified [105]. Under the DSM-V, all are diagnosed as autism spectrum disorder, or ASD [106]. The prevalence of ASD has increased markedly over the last several decades [107; 108]. Although greater public awareness and changes in diagnostic criteria have contributed to this increase, it has been estimated that approximately 46% could be due to unknown factors [109]. The CDC recently estimated that ASD affects one in every 68 children [108].

Genetics play an important role in the development of ASD [110]. Evidence for a genetic component of autism has been gleaned from studies of families and twins. A child is at a 25-fold greater risk for being diagnosed with autism if he or she has an autistic sibling [111]. Further, concordance rates for monozygotic twins (70-90%) are higher than the rates for dizygotic twins (0-25%) [110; 112]. A clear "autism gene" has not been implicated, but current research suggests that the interaction of multiple genes leads to disease development [110]. The most common chromosomal abnormalities associated with ASD are inherited duplications found in the chromosomal region 15q11-15q13 (Angelman syndrome), which codes for the gene for ubiquitin protein ligase, among others [113]. Besides Angelman syndrome, ASD has been associated with a number of other disorders, including Rett syndrome and fragile X syndrome, which has lent further insight into the potential genetic risk factors for ASD [114]. None of the syndromes or molecules associated with ASD are unique to it, but instead lead to a spectrum of similar disorders that vary in severity, such as autism and Asperger syndrome [110]. Together, genetic syndromes, mutations, and copy number variants (CNV) account for approximately 10-20% of ASD cases, although each factor alone only represents 1-2% of cases [110].

Despite the previously estimated 46% contribution of "unknown factors" to ASD development [109], a recent twin study from the United Kingdom reported that genetic factors may explain most of ASD [115]. Colvert et al. obtained data from the population-based cohort Twins Early Development Study, which included all twin pairs born in England and Wales from 1994-1996. For all five of their ASD assessment methods, correlations among monozygotic twins (0.77-0.99) were significantly higher than those for dizygotic twins (0.22-0.65). The authors concluded that heritability could explain 56% to 95% of ASD. Environmental influences had a smaller but still significant effect (30%, 95% CI 8%-47%) for one of the five ASD measures [115].

The etiology of ASD is likely heterogeneous and complex, with a variety of possible genetic, sociodemographic, behavioral, and environmental factors interacting and contributing to disease development. The risk for developing ASD is higher for male than female children (4:1), although this effect is not X-chromosome driven [116]. Other factors associated with an increased risk of ASD include advanced parental age [117], greater parental educational attainment [118], low birth weight and preterm birth [119; 120], pregnancy complications [121], and maternal smoking [122].

1.2.2 Environmental exposures.

Recent studies have begun to explore associations between occupational and environmental exposures and ASD, including heavy metals [123; 124], solvents [125], PCBs and flame retardants [126], phthalates and phenols used in plastic products [127], and pesticides [128; 129]. A handful of studies have utilized databases of modeled estimates of ambient air pollutants [130-132], namely the United States Environmental Protection Agency's National Air Toxics

Assessment program (USEPA NATA). Others have focused on exposures to road traffic and related pollutants, such as nitric oxides, sulfur dioxide, ozone, PM_{10}, and/or $PM_{2.5}$ [133-137]. The major results of these air pollution and ASD studies are summarized next. Although the underlying biological mechanisms of these effects on ASD development remain unclear, the literature to date suggests that air pollution-induced oxidative stress, neuroinflammation, cerebrovascular dysfunction, microglial activation, and alterations in the blood-brain barrier contribute to the pathology of diseases of the central nervous system [138].

1.2.2.1 Ambient air toxics.

Windham et al. (2006) conducted their study of ambient air toxics and ASD for the San Francisco Bay area, using data from the California Centers for Autism and Developmental Disabilities Research and Epidemiology (CADDRE) to identify 284 ASD cases born in 1994. Controls (n=657) were randomly selected from California birth certificate records and matched to cases by sex and month of birth. Exposures to 19 potentially neurotoxic air pollutants were assigned to all subjects by census tract of the birth residence using the 1996 NATA assessment. The authors calculated summary index scores to explore associations between ASD risk and structural groups of air toxics (i.e., metals, aromatic solvents, and chlorinated solvents). Subjects were divided into exposure quartiles, with the first and second quartiles combined as the referent or comparison group. After adjustment for important sociodemographic factors, associations were found for metals (Adjusted Odds Ratio (AOR)=1.50, 95% Confidence Interval (CI)=1.05-2.12) and for chlorinated solvents (AOR=1.55, 95% CI=1.08-2.23), comparing the fourth quartile to the referent. Adjusting simultaneously for the three structural groups led to decreased risk for solvents and increased risk for metals (AOR=1.74, 95% CI=1.01-3.01) [132].

Kalkbrenner et al. (2010) carried out a similar investigation for North Carolina and West Virginia. Cases were 383 children with ASD and controls were 2,829 children with speech and language impairment, both recruited through the Autism and Developmental Disabilities Monitoring Network. The 1996 NATA assessment was again used to estimate exposures to ambient levels of metals, particulates, and VOCs by census tract of the child's residence. Individuals residing in census tracts with "high" estimates (80^{th} percentile) of pollutant concentrations were compared to those in census tracts with lower exposures (20^{th} percentile). No statistically significant results were found; adjusted odds ratios were elevated for methylene chloride (AOR=1.4, 95% CI=0.7-2.5), quinoline (OR=1.4, 95% CI=1.0-2.2), and styrene (OR=1.8, 95% CI=1.0-3.1) [130].

In a large, nationwide case-control study, Roberts et al. (2013) identified 325 children with ASD and 22,098 controls born to mothers who had participated in the Nurses' Health Study II between 1987 and 2002. Unlike the previous two studies, ASD diagnosis was not confirmed for all cases, although the authors did administer a diagnostic test by telephone to 50 (15%) randomly selected cases. Study subjects were assigned estimated pollutant levels from the NATA assessment closest to their year of birth, and individuals in higher quintiles of exposure were compared to those in the lowest. Increased risk for ASD was significantly associated with exposure to overall metals (AOR=1.6, 95% CI=1.1-2.4), lead (AOR=1.6, 95% CI=1.1-2.3), manganese (AOR=1.5, 95% CI=1.1-2.2), and nickel (AOR=1.7, 95% CI=1.1-2.5). The odds ratio for methylene chloride was elevated, but did not quite reach statistical significance (AOR=1.5, 95% CI=1.0-2.1) [131].

The most recent ASD investigation utilizing NATA estimates was a population-based case-control study conducted for southwestern Pennsylvania (Talbott et al., in review). Case and

control children were born in six counties (Allegheny, Armstrong, Beaver, Butler, Washington, and Westmoreland) from January 1, 2005 to December 31, 2009. ASD cases were recruited from autism clinics, treatment centers, local physicians, and through the PA school system. A child was diagnosed with ASD if he or she had: 1) a score of 15 or above on the Social Communication Questionnaire (SCQ), a positive screen for the presence of autistic features, and 2) written documentation, including test results, of an ASD diagnosis from a child psychologist or diagnostic center. Controls were randomly selected from the Pennsylvania Department of Health (PA DOH) state birth registry files and frequency-matched to cases on year of birth, gender, and race. All participants were administered an interview to obtain information on residential history, workplace and residential exposures, and personal risk factors. In the end, 217 cases and 226 controls were recruited and interviewed.

As in the previous three studies, exposures to air toxics were estimated using NATA. The 2005 assessment was used since it was closest to the time period under study (2005 to 2009). Census tract-level estimates were obtained for 30 hazardous air pollutants (HAPs) with established or potential neurotoxic, developmental, and/or endocrine-disrupting effects. A computer algorithm calculated location- and time-specific exposure estimates for each mother/child pair, taking into account changes in the mother's residence and time spent at each residence for each critical time period (i.e., three months before pregnancy, trimesters 1-3, child's first and second birthday). Participants were then divided into quartiles of exposure for each pollutant, with those in the first quartile as the referent or comparison group.

To explore associations between exposure level and ASD risk, logistic regression analyses were conducted, unadjusted and adjusted for mother's age, education, race, and smoking. Out of the 30 HAPs, styrene was found to be significantly associated with increased

odds of ASD for the 4th compared to the 1st quartile (AOR=1.97, 95% CI=1.13-3.43, p=0.02). In a second analysis using a random sample of 5,007 controls generated from PA DOH birth certificates, styrene had a borderline significant relationship with ASD (AOR=1.46, 95% CI=0.97-2.19, p=0.07) while arsenic, chromium, methanol, methylene chloride, and PAHs had significant associations with ASD ($p<0.05$). The point odds ratios for these additional five compounds were all elevated in the interviewed case-control analysis, but they did not reach statistical significance, possibly due to the lower sample size.

1.2.2.2 Traffic.

The remaining investigations focused on ASD risk and traffic-related pollutants, using proximity to major roads as a surrogate for traffic-related exposures [136] or levels of ambient air pollutants derived either from regulatory monitors or spatiotemporal models, such as land use regression (LUR) [133-135; 137]. Volk et al. (2011, 2013) conducted two population-based, case-control investigations in California as part of the Childhood Autism Risks from Genetics and the Environment (CHARGE) study. In the 2011 study, 304 cases of ASD and 259 typically developing controls were recruited using California birth records from 2003 to 2009. Controls were matched to cases on age, sex, and general geographical area. A detailed residential history from three months before conception through the current address was obtained by personal interview. Distances from each residence to the nearest freeway (state or interstate highway) and to the nearest major road (state, interstate, or major arterial) were calculated. Exposure cut points were based on the closest 10% (<309 m), the next 15% (309-647 m), and the next 25% (647-1,419 m) compared to the remaining 50% (>1,419 m). ASD was associated with residential proximity to a freeway during the third trimester (AOR=1.96, 95% CI=1.01-3.93), suggesting a late-pregnancy effect [136].

In a subsequent study, Volk et al. (2013) expanded their investigation of ASD and traffic to include estimates of traffic-related air pollutants. Exposures were assigned using both actual measurements from USEPA Air Quality System (AQS) monitors and model-based estimates for $PM_{2.5}$, PM_{10}, ozone, and nitrogen dioxide. Increased risk for ASD development was found for participants living within the highest quartile of exposure to traffic-related air pollution during gestation (AOR=1.98, 95% CI=1.20-3.31) and during the first year of life (AOR=3.10, 95% CI=1.76-5.57). ASD risk increased per an interquartile range (IQR) increase of 8.7 $\mu g/m^3$of $PM_{2.5}$ during gestation (AOR=2.08, 95% CI=1.93-2.25) and the first year of life (AOR=2.12, 95% CI=1.45-3.10). PM_{10} and nitrogen dioxide were also associated with ASD during gestation and the first year of life [137].

Becerra et al. (2013) conducted another California-based study of traffic-related pollution and ASD risk for children born from 1995 to 2006. Cases of ASD (n=7,603) were identified through the California Department of Developmental Services and were matched to 75,782 controls on birth year, sex, and gestational age. Exposures to carbon monoxide, nitrogen dioxide, nitric oxide, ozone, and particulate matter were assigned using the USEPA air monitor closest to the birth residence. LUR models were also used to estimate exposures to nitrogen dioxide and nitric oxide. Adjustments were made for other important sociodemographic and behavioral risk factors, except for maternal smoking. Increased risk for ASD was found per IQR increase of ozone (AOR=1.12, 95% CI=1.06-1.19 per 11.54 ppb increase) and $PM_{2.5}$ (AOR=1.15, 95% CI=1.06-1.24 per 4.68 $\mu g/m^3$ increase). For LUR-based estimates of nitrogen dioxide and nitric oxide, ASD risk increased 3–9% per IQR increase [133].

The two most recent investigations focused on ASD and exposure to particulate matter only [134; 135]. Raz et al (2014) conducted a nested case-control study of participants in the

Nurses' Health Study II (NHS II) of 245 ASD cases frequency matched to 1,522 controls born between 1990 and 2002. A spatiotemporal model for the continental US was used to estimate monthly averages of $PM_{2.5}$ and $PM_{10-2.5}$ at residential addresses. The authors found that higher maternal exposure to $PM_{2.5}$ during pregnancy (AOR=1.57, 95% CI=1.22-2.03 per IQR increase) and during the first trimester (OR=1.42, 95% CI=1.09-1.86 per IQR increase) were associated with increased odds of ASD [135]. Kalkbrenner et al. (2015) used geostatistical interpolation to assign daily levels of PM_{10} from regulatory air monitors to 979 ASD cases and 14,666 controls in North Carolina and California born between 1994 and 2000. Similar to Raz et al. (2014), an association was found between ASD risk and exposure to PM_{10} during the third trimester (AOR=1.36, 95% CI=1.13-1.63), even in models adjusting for all three trimesters at the same time (AOR=1.38, 95% CI=1.03-1.84) [134].

1.2.2.3 Multiple pollutants.

Using proximity to major road [136] or another surrogate for traffic-related pollutants (e.g., traffic density) is one approach to investigating the association of a source itself with the disease outcome—in this case, traffic and ASD—rather than with an individual air toxic. "Traffic" represents a variety of pollutants emitted from gasoline and diesel vehicles, including but not limited to particulate matter, hydrocarbons, gases like NO_x and SO_2, and other air toxics. Several of the ASD studies have also investigated mechanistic and structural groups of compounds and their associations with ASD [131; 132]. Since people are typically exposed to a complex mixture of air contaminants, a multi-pollutant approach to studying the effects of air toxics on ASD and other health outcomes may prove beneficial [139]. Further, a multi-pollutant framework would better aid air quality policies and permit the identification of specific sources of harmful pollution [139].

In addition to looking at individual air toxics, one investigation to date has attempted to explore associations between ASD and pollution sources [140]. The authors used a dimension reduction technique called factor analysis to examine the correlation structure among 24 neurotoxicants. This method reduced their set of environmental variables from 24 to 4 main "factors" of interest. Pollutants loading on the same factor are correlated with each other and may share a common characteristic, such as belonging to the same structural family or emitting from the same source. The air toxic most significantly related to ASD risk in each factor was then chosen to represent that factor in subsequent multi-pollutant logistic regression models.

Von Ehrenstein et al. found that autism risks increased per interquartile range increase of individual pollutants loading on factor 1 (1,3-butadiene, meta/para-xylene, other aromatic solvents, lead, and perchloroethylene), factor 3 (formaldehyde), and not loading on any factor (trichloroethylene) after adjustment. Associations with 1,3-butadiene, xylenes, and lead tended to weaken in 2- and 3-pollutant models, while associations with formaldehyde and trichloroethylene were unaffected. In general, pollutants for which the authors observed the strongest associations with ASD are related to road traffic (e.g., 1,3-butadiene and aromatic hydrocarbons) and point sources like dry cleaners and industrial stationary cleaning or degreasing operations (perchloroethylene and trichloroethylene) [140]. One limitation of their study is the selection of an "indicator" pollutant from each factor to represent that factor rather than incorporating the factor itself into their models.

1.2.2.4 Styrene.

Two previous investigations, including the recent case-control study by Talbott et al. (in review), have suggested an association between ASD risk and exposure to modeled ambient levels of styrene, an aromatic hydrocarbon [130]. These findings motivated a sub-analysis of styrene

sources as part of the present work regarding ASD and sources of air toxics. Styrene (ethenylbenzene) is a volatile, colorless liquid with a characteristic sweet, sharp odor [141]. It is primarily used in the manufacture of plastics and resins (e.g. polystyrene, styrene-butadiene rubber, and unsaturated polyester resins) [141]. The sources, exposure pathways, and health effects of styrene are summarized here.

Occupational exposure to styrene is greatest in the reinforced plastics and boat building industries [141-143]. In occupational settings, styrene enters the body primarily through absorption in the lungs, although a small amount of dermal absorption also occurs [141]. Emissions from major point sources, such as styrene production plants and the industries mentioned previously, also contribute to ambient levels of styrene relevant to the exposure of the general population [144]. Other non-occupational sources of exposure include exhaust from gasoline- or diesel-powered vehicles [145], cigarette smoke [146], off-gassing of residual styrene from products in the home (carpet glues, flooring materials, etc.) [147], and food, due to either natural processes or the leaching of styrene from packaging [148; 149].

The acute symptoms of occupational styrene exposure have been well documented. Workers exposed to styrene vapor have reported the following symptoms: acute eye and throat irritation, nausea, vomiting, weakness, dizziness, headache, and loss of appetite ("styrene sickness") [142; 150; 151]. Styrene is also an ototoxicant, causing damage to the outer hair cells of the inner ear, which leads to eventual hearing loss [152]. In a recent review of the styrene profile in the National Toxicology Program 12th Report on Carcinogens, it was reaffirmed that styrene's classification as "reasonably anticipated to be a human carcinogen" was appropriate due to sufficient evidence of carcinogenicity from experimental animal studies and limited but credible evidence from human studies [153]. Styrene-7,8-oxide, a major metabolite of styrene in

both experimental animals and humans, has also been classified as a probable human carcinogen [153].

Both animal studies and human occupational studies have suggested that styrene has neurotoxic properties. Styrene is lipophilic and therefore may be easily absorbed by the lipid-rich nervous system [154]. Further, styrene partially partitions in the brain following inhalation exposure in animal models. In one study, rats exposed via inhalation to 2,800 ppm of styrene had an average concentration of 25 mg/100 cm^3 in brain tissue, although this decreased to 8.6 mg/100 cm^3 one hour later [155]. Another investigation exposed rats to 300 ppm of styrene over 11 weeks; at four weeks, the concentration of styrene in the brain peaked at 47 nmol/g [156]. Styrene exposure may also elevate levels of specific markers of nervous system damage. Rosengren and Haglid exposed 32 male Sprague Dawley rats to 320 ppm styrene via inhalation chambers for three months. They observed significantly elevated levels of glial fibrillary acidic protein (GFA), suggestive of cell proliferation in response to central nervous system (CNS) damage [157].

Occupational studies have suggested that chronic exposure to styrene hinders neurobehavioral performance. A cross sectional study of former workers of a polyester boat building plant found that exposed workers performed worse than control workers in several tests of neurobehavioral performance (e.g. symbol-digit substitution and digit span forwards) [142]. Less than 10 years of exposure to an average styrene level of 155 mg/m^3 appeared to lead to persistent neurotoxic effects in this particular study [142]. A meta-analysis of the human neurobehavioral effects of chronic styrene exposure revealed that cumulative styrene exposure was significantly associated with increased choice reaction time (CRT) and increased color confusion index (CCI) [158]. Eight work-years of exposure to 20 ppm styrene, a contemporary

limit for occupational exposure at the time of the study, resulted in a 6.5% increase in CRT and 2.2% increase in CCI [158]. One study found that workers in the reinforced plastics industry were at increased risk for mortality from diseases of the CNS [143]. To date, no experimental animal or human studies, except for the two epidemiologic studies summarized previously, have investigated exposure to styrene directly in relation to ASD.

As has been described, a variety of sociodemographic, behavioral, and environmental factors have been associated with increased risk for adverse birth outcomes (preterm birth, LBW, and SGA) and ASD. In the next three sections, associations between these health outcomes and exposure sources of interest are explored more fully. Specific Aim 1 (UGD and adverse birth outcomes) is addressed in the first paper, while Specific Aims 2a and b, both concerning ASD and sources of air toxics, are discussed in the second and third papers.

2.0 PERINATAL OUTCOMES AND UNCONVENTIONAL NATURAL GAS OPERATIONS IN SOUTHWEST PENNSYLVANIA

The data presented in this chapter appears in: Stacy SL, Brink LL, Larkin JC, Sadovsky Y, Goldstein BD, Pitt BR, Talbott EO. Perinatal Outcomes and Unconventional Natural Gas Operations in Southwest Pennsylvania. *PLOS ONE*, accepted for publication.

2.1 ABSTRACT

Unconventional gas drilling (UGD) has enabled extraordinarily rapid growth in the extraction of natural gas. Despite frequently expressed public concern, human health studies have not kept pace. We investigated the association of proximity to UGD in the Marcellus Shale formation and perinatal outcomes in a retrospective cohort study of 15,451 live births in Southwest Pennsylvania from 2007-2010. Mothers were categorized into exposure quartiles based on inverse distance weighted (IDW) well count; least exposed mothers (first quartile) had an IDW well count less than 0.87 wells per mile, while the most exposed (fourth quartile) had 6.00 wells or greater per mile. Multivariate linear (birth weight) or logistical (small for gestational age (SGA) and prematurity) regression analyses, accounting for differences in maternal and child risk factors, were performed. There was no significant association of proximity and density of UGD with prematurity. Comparison of the most to least exposed, however, revealed lower birth

weight (3323 +/- 558 vs 3344 +/- 544 g) and a higher incidence of SGA (6.5 vs 4.8%, respectively; odds ratio: 1.34; 95% confidence interval: 1.10-1.63). While the clinical significance of the differences in birth weight among the exposure groups is unclear, the present findings further emphasize the need for larger studies, in regio-specific fashion, with more precise characterization of exposure over an extended period of time to evaluate the potential public health significance of UGD.

2.2 INTRODUCTION

Unconventional gas development (UGD), characterized by advances in engineering, including horizontal drilling and high volume hydraulic fracturing, enables extraction of large amounts of fossil fuel from shale deposits at depths that were previously unapproachable [86]. In Pennsylvania, UGD in the Marcellus shale formation has rapidly advanced from only 44 such wells known to be drilled before 2007 to 2,864 wells drilled during the 2007-2010 period of our study, and with continued rapid expansion to as many as 80,000 forecasted [159].

Several recent reviews summarizing the evolving UGD process describe the potential for adverse health effects and delineate challenges that have contributed to as yet minimal understanding of public health impact [86; 88; 160]. UGD is a dynamic process encompassing preparation of the site, well development and production, the removal of wastes and the downstream distribution of gas [86]. The well is drilled vertically into a shale layer often 1.5 km underground and then turned laterally within the shale layer for another 2-3 km before holes are blown at intervals in the pipe. This is followed by the high-pressure injection of approximately 5 million gallons of water to hydraulically fracture the shale layer, allowing the release of gas

tightly bound to the shale. Added to this water is a complex mixture, including approximately 15% of a physical agent (usually silica) to prop open the fractures and about 0.5-2.0% of an evolving mixture of about 6-10 chemicals (e.g., surfactants, biocides, metal chelators, and others), that enhance release and flow of the gas. Return or flowback fluids include mixtures of the hydrofracturing agents, hydrocarbon products (methane and other volatile organic hydrocarbons including benzene) and, of particular toxicological significance, naturally occurring agents dissolved from the shale bed (e.g., brine, radionuclides, arsenic, barium, strontium and other metals) [161; 162]. Over a thousand diesel truck trips are usually required for site preparation, bringing hydrofracturing fluids and disposing of the approximately 1-2 million gallons of fluid that flows back from the well. In the western US, flowback fluids are generally rapidly disposed of in deep underground injection wells. Such wells are uncommon in Pennsylvania. UGD operators first discharged to publically owned treatment works, which treated the wastewater and discharged to the regional rivers until it was determined that this practice was associated with increasing concentrations of bromine and other contaminants in drinking water pulled from the rivers [163; 164]. Next, the flowback waters were transported to deep underground injection wells in Ohio. However, the resultant mild earthquakes in Ohio have led to a variety of attempted solutions to deal with these flowback fluids on the surface, including impoundments and recycling, thereby increasing the opportunity for human exposure [89]. This continues to be the current situation in Pennsylvania. As flowback fluids also contain hydrocarbon product, they can be a source of air pollution. Esswein et al recently reported that workers involved with waste fluids could be exposed to levels of benzene above allowable occupational health levels [94]. This is pertinent as benzene in air has been associated with adverse birth outcomes [93].

Wells can be hydrofractured intermittently on multiple occasions to stimulate product flow. A more continuous process of product development occurs in region-specific patterns. This includes condensate tanks and glycol dehydrators to separate dry (methane) and wet (higher carbons such as ethane) gas components of product and diesel fuel operated compressors (to liquefy gas for shipping via pipelines) [90]. As such, concern about air pollution is both direct (flaring of methane gas at well heads, controlled burning of natural gas and release of VOCs including benzene, toluene, ethylbenzene and xylene) and indirect (traffic, diesel operated compressors).

Major challenges in assessing and quantifying environmental, ecological and human health related effects (existing and potential) of UGD exist largely due to the dynamic and complex nature of the evolving UGD process itself as well as differences in geology between site locations, UGD technique and community demography. Together, these factors make it difficult to compare experiences, historically and concomitantly, within and between regional efforts. Several recent studies have provided measurements of likely pollutants, focusing on hydrocarbons found in air [66] or on thermogenic methane found in shallow drinking water sources [90; 95; 165]. A study in Colorado revealed that those living within 0.5 miles of a well were exposed to air pollutant levels, including benzene, that significantly increased non-cancer risk [96]. However, there is still a lack of information linking potential exposures with public health risks, which led the State of New York to the following declaration: “Until the science provides sufficient information to determine the level of risk to public health from HVHF and whether the risks can be adequately managed, HVHF should not proceed in New York State” [166].

The embryo/fetus is particularly sensitive to the effects of environmental agents [3]. A host of environmental and behavioral risk factors have been identified and linked to low birth weight and prematurity. They include most notably cigarette smoking [167; 168], maternal occupational exposures to metals [73; 74], and recently $PM_{2.5}$ and ozone [62; 66; 67]. The mechanism is thought to be one involving oxidative stress or inflammation [77]. Xu et al. have noted a relationship in southwestern Pennsylvania of low birth weight and $PM_{2.5}$ [67]. The strength of using birth outcomes is the availability of data and the ability to capture the critical time of exposure and linkage to outcomes within the nine month period [169]. McKenzie et al used a retrospective cohort design and exposure estimates from an inverse distance weighted (IDW) approach to explore associations between maternal residential proximity to hydraulic fracturing sites in Colorado and birth outcomes [102]. They found an increase in the prevalence of congenital heart defects and, to a lesser extent, neural tube defects with increasing exposure to natural gas extraction. They also found an increase in birth weight associated with well density [102].

We adapted the epidemiological and geographic information systems (GIS) approaches of McKenzie et al to explore the potential effects of UGD on infants born to mothers living in Southwestern PA where unconventional drilling of the Marcellus Shale has been rapidly expanding. The objective of the present study is to use readily available data on birth outcomes for Southwestern Pennsylvania to investigate the relationship of proximity to UGD and perinatal outcomes for 2007-2010.

2.3 METHODS

Natural gas well and birth data were collected for Butler, Washington and Westmoreland counties in PA for the years 2007 to 2010. The UGD locations were obtained from the Pennsylvania Department of Environmental Protection (PADEP), that defines UGD as wells having both a lateral component and hydraulic fracturing, a process relatively new to Pennsylvania until 2005 [159]. The PADEP dataset also includes information on drilling commencement dates, known as the spud date, and well status (active, abandoned, etc.) [159]. Birth data for these counties were obtained using information from birth certificates, which had also been geocoded by the Pennsylvania Department of Health (PA DOH) Bureau of Vital Statistics. This study was approved by the University of Pittsburgh Institutional Review Board (IRB number PRO12060174). Individual data on these births was accessed through a password protected application with the PADOH. Information was abstracted regarding maternal risk factors (age, education, cigarette smoking history, use of Women, Infant and Children/WIC assistance, gestational diabetes, prenatal visits, pre-pregnancy weight, and birth parity) as well as gestational age and gender of child at birth [170]. Multiple births, records without a valid geocode (X, Y coordinate), and those with missing birth outcome and demographic information were excluded from the analysis. Exact point distances between singleton-birth residences with complete information and natural gas wells were calculated using ArcMap (version 10.1; ESRI Inc., Redlands, CA).

We calculated an inverse distance weighted (IDW) well count for each mother living within 10-miles of UGD to account for both the number of unconventional wells within this buffer as well as distance of each well from the mother's residence [102]. This metric, shown

below in Equation 1, gives greater weight to unconventional wells closest to the mother's residence:

Equation 1. Inverse distance weighted well count

$$\text{IDW well count} = \sum_{i=1}^{n} \frac{1}{di}$$

where the IDW well count is the inverse distance weighted count of unconventional wells within a 10-mile radius of maternal residence in the birth year, n is the number of existing unconventional wells within a 10-mile radius of maternal residence in the birth year, and d_i is the distance of the i^{th} individual well from the mother's residence. For example, a mother's residence that has two wells, both 0.5 mile away, would have an IDW well count of 4. Mothers were categorized into exposure quartiles according to their IDW well counts:

Group 1: IDW Well Count >0 but <0.87

Group 2: IDW Well Count ≥0.87 but <2.60

Group 3: IDW Well Count ≥2.60 but <6.00

Group 4: IDW Well Count ≥6.00

Three indicator variables were created, using the first quartile (Group 1) as the referent group. The 10% of births that did not live within 10 miles of UGD were eliminated from the analysis due to notable sociodemographic differences; these mothers were more African American (7% compared to 3%), smoked more during pregnancy (25% versus 20%), and had a higher proportion receiving WIC assistance (41% versus 32%).

The outcomes assessed were continuous birth weight, small for gestational age (SGA), and prematurity (gestational age <37 weeks). To identify SGA births, birth weights were normalized to gestational age and estimates of SGA were deduced from nomograms identifying elements of fetal growth (SGA <10% of predicted weight for a given gestational age and

gender) [37]. Mean birth weights in each group were compared using analysis of variance (ANOVA), and proportions of SGA and premature infants were compared using chi-square tests. Outcomes were modeled using multivariate linear regression (continuous birth weight) or logistic regression (SGA and prematurity). All models were adjusted for gender of the child and mother's age, education (8^{th} grade or less; 9^{th}-12^{th} grade, no diploma; high school graduate or GED completed; some college credit, but not a degree; associate degree; bachelor's degree; master's degree; doctorate or professional degree), pre-pregnancy weight, prenatal care (1 if at least 1 visit; 0 otherwise), smoking (1 if smoked at all during pregnancy; 0 otherwise), gestational diabetes (1 if present; 0 otherwise), WIC (1 if received; 0 otherwise); African American (1 if yes; 0 otherwise) and parity (first child; second child; third child; fourth child or greater). The model for continuous birth weight was also adjusted for gestational age to account for the downward shift in birth weights accompanying shorter gestational ages due to earlier obstetric intervention observed in our dataset from the PADOH as well as nationally [171]. All statistical tests were performed using IBM SPSS Statistics 21 and assessed at a significance level of $\alpha = 0.05$.

2.4 RESULTS

2.4.1 Descriptive statistics.

This analysis included 509 active unconventional natural gas wells in Butler, Washington and Westmoreland counties from 2007 to 2010, representing 18% of the state-wide total of 2,864

[159]. Figure 1 shows the steps used to eliminate unavailable and missing birth certificate data, leading to the final sample of births with complete information.

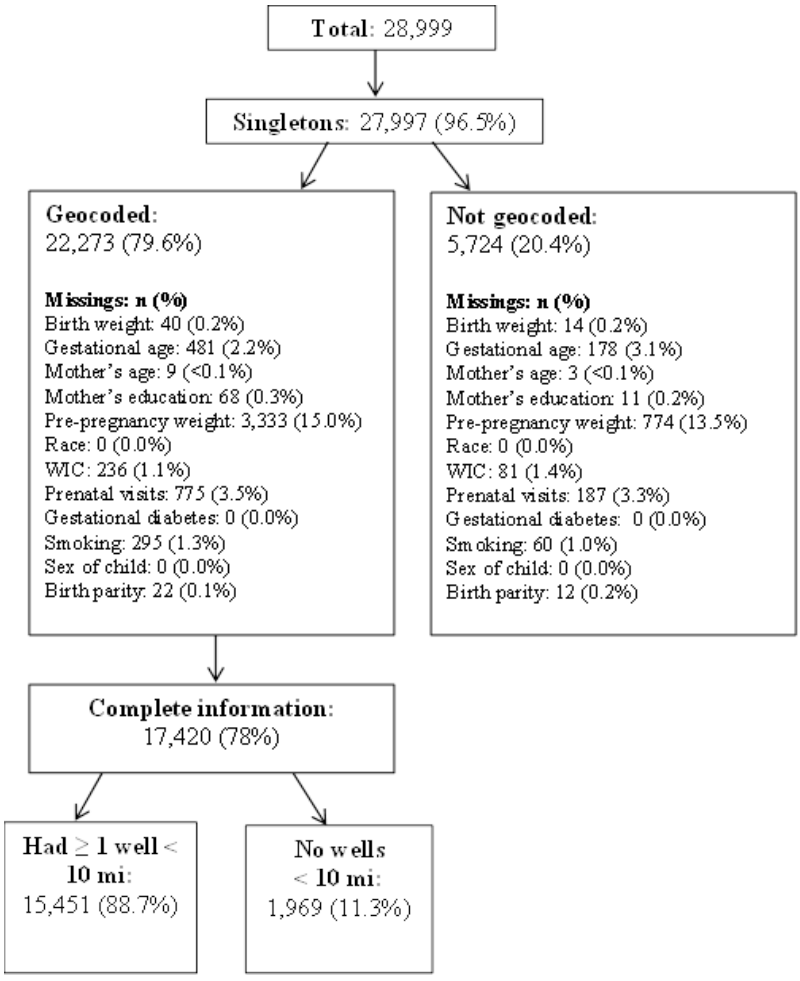


Figure 1. Flowchart of sample sizes and missing data for births in Butler, Washington, and Westmoreland Counties 2007-2010

There were 28,999 total births in these three counties from 2007 to 2010, and 27,997 (97%) of these were singleton live births. Out of the singleton birth residences, 5,724 (20%) were not geocoded to an X,Y coordinate and, since the dataset did not include an address or zip code for the mother's residence, were excluded from the analysis. This left 22,273 singleton births available for further analysis in ArcGIS. Birth weight was missing for 0.2% of these geocoded

singleton births, and gestational age was missing for 2.2%. Mother's age, mother's education, and birth order were missing for less than 1% of births. Pre-pregnancy weight was missing for 15% of mothers, WIC assistance for 1.1%, the number of prenatal visits for 3.5%, and information on smoking for 1.4%. The remaining 17,420 births had complete geographical and birth certificate information. Of these, 15,451 (89%) had at least one well within 10-miles of the mothers residence.

Table 1 shows the demographics of these 15,451 infant-mother pairs by quartile (the referent group (first quartile) and three exposure quartiles) as well as the proportions of SGA and premature infants in each group. Mother's education and parity were categorized into 8 and 4 groups, respectively; results are presented for percentage that completed high school/GED and first child. There were no significant differences in prenatal care, gestational diabetes, child gender, or parity between the referent and exposure quartiles. Differences in gestational ages and mother's ages between the four groups were small but statistically significant. Mother's education, pre-pregnancy weight, race, WIC assistance, and smoking were also statistically different between the four groups. Chi-square analyses showed statistically significant differences in the proportions of SGA and preterm births. All proportions of SGA were significantly less than the 10% expected for the population [31] but were similar to the general population (regardless of proximity to well) in various counties in our study.

Table 1. Maternal and child risk factors

Factor	Total N=15,451	Referent (First Quartile)[a] N=3,604	Second Quartile[a] N=3,905	Third Quartile[a] N=3,791	Fourth Quartile[a] N=4,151
Mother's age (years)[b]	28.6 ± 5.8	28.8 ± 5.8	28.7 ± 5.8	28.6 ± 5.7	28.3 ± 5.8
Mother's Education (% high school graduate/GED)[b]	22.7%	22.1%	22.5%	22.6%	23.6%
Pre-Pregnancy Weight (lbs)[b]	153.8 ± 39.1	152.6 ± 38.2	152.9 ± 38.2	155.2 ± 40.2	154.7 ± 39.9
Race (% African American)[b]	3.0%	2.6%	2.0%	3.4%	4.1%
WIC (% assistance)[b]	32.1%	29.6%	31.0%	33.6%	34.1%
Prenatal care (% at least one visit)	99.5%	99.5%	99.5%	99.5%	99.3%
Presence of gestational diabetes	4.1%	4.7%	3.7%	4.3%	3.9%
Cigarette smoking during pregnancy[b]	20.0%	19.6%	18.8%	19.9%	21.7%
Gestational age (weeks)[b]	38.7 ± 1.9	38.6 ± 1.9	38.8 ± 1.8	38.7 ± 1.9	38.7 ± 1.9
Birth weight (g)[b]	3345.8 ± 549.2	3343.9 ± 543.9	3370.4 ± 540.5	3345.4 ± 553.5	3323.1 ± 558.2
Small for gestational age[b]	5.5%	4.8%	5.2%	5.6%	6.5%
Premature[b]	7.7%	8.0%	6.7%	8.4%	7.9%
Congenital anomalies[b]	0.5%	0.3%	0.7%	0.4%	0.5%
Percent female	48.5%	48.7%	48.3%	48.6%	48.5%
Birth parity (first)	42.7%	42.8%	41.7%	42.2%	44.1%

[a]Referent (First quartile), <0.87 wells per mile; Second quartile, 0.87 to 2.59 wells per mile; Third quartile, 2.60 to 5.99 wells per mile; Fourth quartile, ≥6.00 wells per mile

[b]Difference between quartiles is significant (p-value <0.05)

2.4.2 Model results.

Table 2 shows the multivariate linear regression results for birth weight, adjusted for mother's age, education, pre-pregnancy weight, gestational age, child gender, prenatal visits, smoking, gestational diabetes, WIC, race, and birth order.

Table 2. Multivariate linear regression of birth weight and proximity

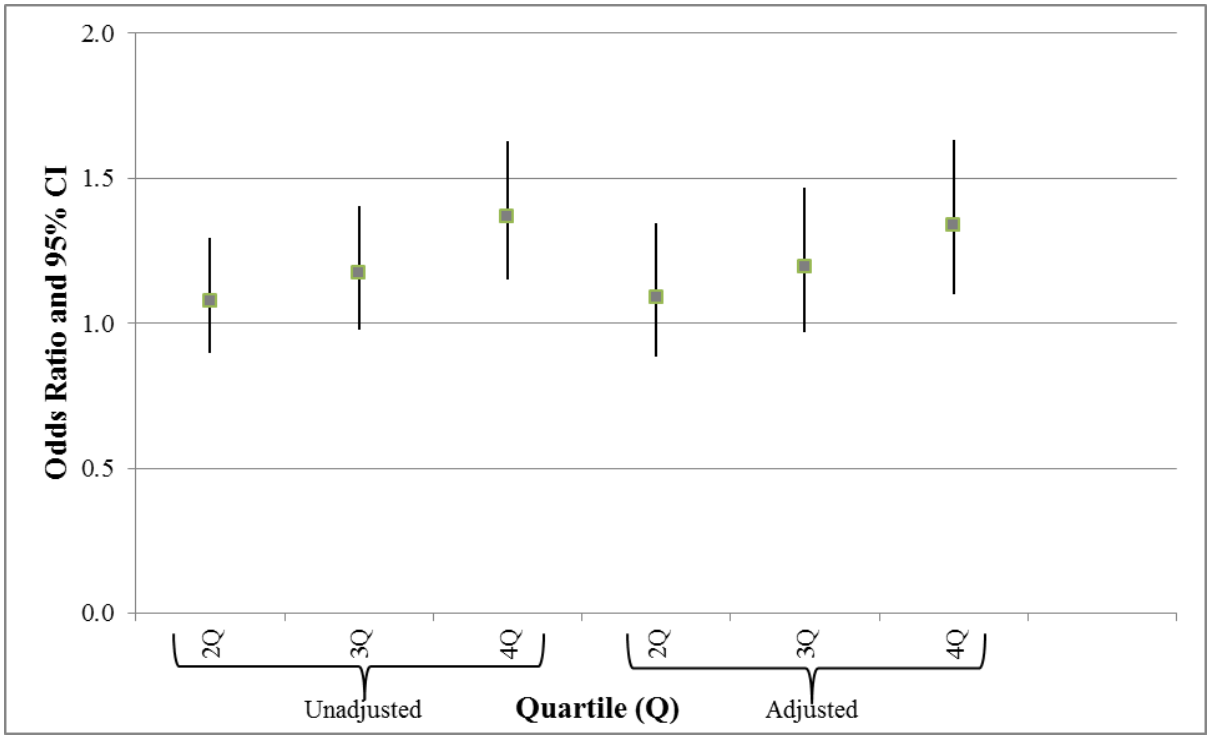
Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance (P)
	B	*Standard Error*	*Beta*		
Constant	-3711.86	93.06	-39.88		<0.01
Mother's Age	-2.95	0.77	-0.03	-3.82	<0.01
Mother's Education	17.88	2.72	0.05	6.58	<0.01
Pre-Pregnancy Weight	2.01	0.09	0.15	23.37	<0.01
Gestational Age	172.64	1.97	0.56	87.51	<0.01
Female	-133.90	6.63	-0.12	-20.19	<0.01
Prenatal Care	127.07	51.53	0.02	2.47	0.01
Smoking During Pregnancy	-184.69	9.07	-0.14	-20.37	<0.01
Gestational Diabetes	33.57	16.82	0.01	2.00	0.05
WIC	-27.44	8.62	-0.02	-3.18	<0.01
Race	-146.22	19.88	-0.05	-7.36	<0.01
Birth parity	65.89	4.01	0.12	16.41	<0.01
Low[a]	10.55	9.52	0.01	1.11	0.27
Medium[a]	-0.48	9.59	0.00	-0.05	0.96
High[a]	-21.83	9.39	-0.02	-2.32	0.02

[a]Low, Second quartile to referent; Medium, Third quartile to referent; High, Fourth quartile to referent

After accounting for these factors, we found that infants in the highest (fourth) exposure quartile tended to have lower birth weights than those in the referent group ($p = 0.02$). There were no significant differences in birth weight between the other exposure quartiles and the referent group. In accord with our current understanding, higher birth weights were associated with

mothers that were younger, more educated, had higher pre-pregnancy weights, had more prenatal care, did not smoke during pregnancy, had gestational diabetes, did not receive WIC, were Caucasian, and had previous children [36]. Higher birth weights were also associated with longer gestational ages and being male.

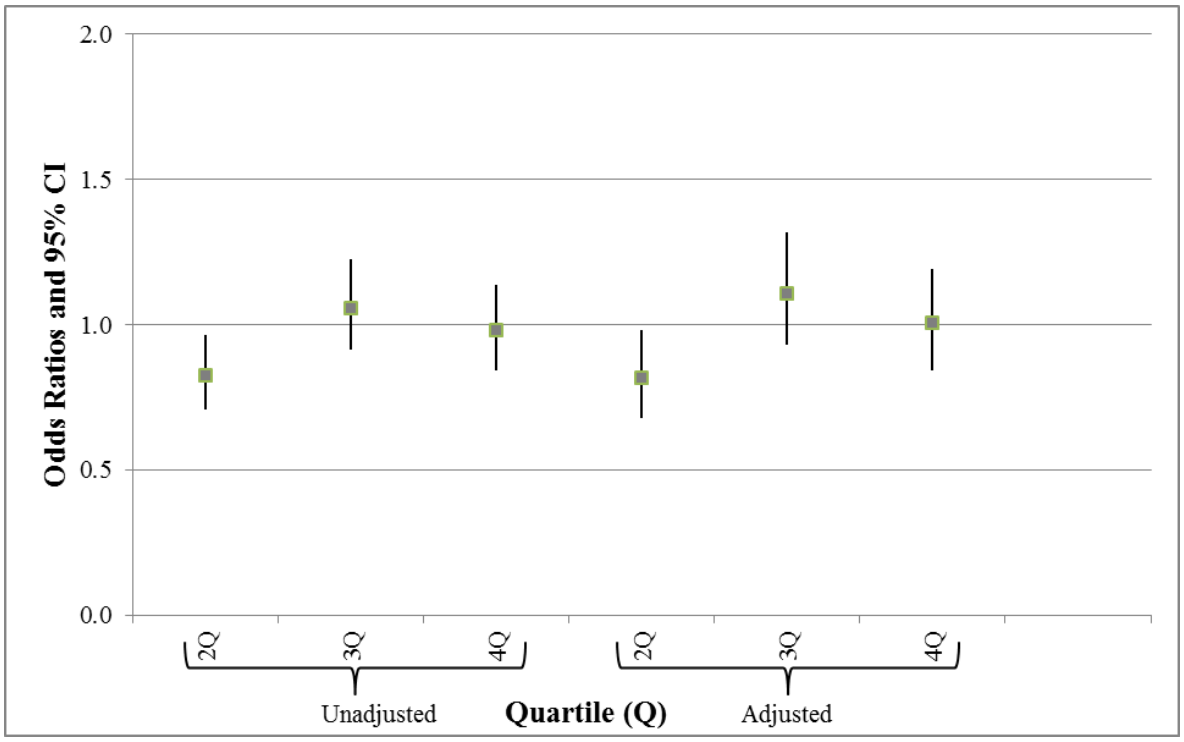
Figure 2 shows the unadjusted and adjusted odds ratios (OR) and 95% confidence intervals (CI) for SGA.



Key: Referent (First quartile), <0.87 wells per mile; Second quartile (2Q), 0.87 to 2.59 wells per mile; Third quartile (3Q), 2.60 to 5.99 wells per mile; Fourth quartile (4Q), ≥6.00 wells per mile
Figure 2. Unadjusted and adjusted odds ratios and 95% confidence intervals (CI) for SGA

The steady increase in SGA across quartiles (Table 1) resulted in a progressive increase in odds ratio for SGA (adjusted or unadjusted), suggestive of a dose-response relationship. In the adjusted model, the highest exposure group compared to the referent reached significance (OR=1.34, 95% CI=1.10-1.63).

Figure 3 shows the unadjusted and adjusted odds ratios and 95% confidence intervals for prematurity.



Key: Referent (First quartile), <0.87 wells per mile; Second quartile (2Q), 0.87 to 2.59 wells per mile; Third quartile (3Q), 2.60 to 5.99 wells per mile; Fourth quartile (4Q), ≥6.00 wells per mile
Figure 3. Unadjusted and adjusted odds ratios and 95% confidence intervals (CI) for prematurity

Prematurity was associated with mothers that were older, less educated, had no prenatal care, smoked, had gestational diabetes and had no previous births. Male babies were also more likely to be premature than females. There was no significant effect of well density on prematurity except for a slightly lower proportion of premature infants born to mothers in the second exposure quartile compared to the referent (adjusted OR=0.82, 95% CI=0.68-0.98).

2.5 DISCUSSION

We accessed public records of UGD and birth and used a geographic information system that enabled proximity and density of nearby UGD to be used as a surrogate for exposure. Based on this latter estimate, we identified four groups of mothers of comparable size that gave birth in the study period (2007-2010) in three counties in Southwest Pennsylvania with high levels of UGD activities. These four groups were relatively similar in various determinants of maternal and child risks for perinatal outcomes but had different levels of exposure (i.e. IDW well count) (Table 1). The information was readily compatible for multivariate linear and logistic regression analysis in which covariates of risk could be accounted for (at least within limits of available birth certificate data in Pennsylvania) and contribution of exposure could be assessed. Even when the SGA births were removed, a small but significant decrement in mean birth weight by quartile of exposure remained ($p<0.05$). McKenzie et al. were able to explore subsets of congenital anomalies and neural tube defects [102], but our dataset had insufficient power to explore such birth defects.

2.5.1 Comparison of existing studies on UGD and perinatal outcomes.

This analysis adds to possible health impact concerns recently described by McKenzie et al. in which there was an increase in birth defects associated with proximity to UGD in rural Colorado [102]. In contrast to the McKenzie et al. study, our observation of a decrement in birth weight in the highest exposure group is similar to preliminary reports of two other studies, including the original thesis work of Elaine Hill and a recent abstract [100; 101]. The differences in these studies on effects of UGD on birth weight from Colorado (where proximity and density were

associated with a protective effect) underscore the importance of assessing health impacts in a region-specific fashion.

Geological differences are known to account for differences in flowback water composition in different shale gas areas [172]. A notable regional difference between Colorado and Pennsylvania is that the disposal of flowback fluids is far more likely to lead to human exposure in Pennsylvania where deep underground injection has not been feasible [161]. Surface disposal sites are not readily available for geolocating, and thus could not be used in our IDW model. However, impoundments and other sites to which the flowback water is piped or trucked are likely to be near drilling sites, particularly when there are multiple sites in the area, and impoundments have been demonstrated to leak [161; 164]. Therefore, the IDW model is still likely to be representative of exposure risk. There are also important regional differences within Pennsylvania that may be pertinent to a comparison of our findings with those of other studies. Southwestern Pennsylvania is a “wet gas” area, which contains far higher levels of benzene and other relatively higher weight shale gas components than do the “dry gas” areas of the rest of the Marcellus shale regions of the state. The management of flowback fluids presents a risk of air pollution as well as water pollution. Studies with cooperating industries have shown very wide variation from site to site in methane emissions, and in worker benzene exposures [93; 173].

McKenzie et al. established criteria to restrict their analysis to rural areas, thereby minimizing the contributions of other industries, traffic, congestion and other confounding influences of a more urban environment [102]. Although UGD in Southwestern PA does not include the most dense areas of Allegheny County, the population density in the counties we studied surrounding Pittsburgh are greater than rural Colorado [174]; thus, our assessment of exposure likely included different contributing sources of confounding pollution and other

variables. McKenzie et al. also included impact of altitude that is important in Colorado but can be overlooked in the comparatively modest elevations in Southwestern PA. Non-white mothers were excluded in their analysis (as it was too small a group within existing cohorts) and their referent group was individuals >10 miles from UGD [102]. This group of mothers (those >10 miles) in the present study was composed of a somewhat different demographic of women than those living within 10 miles of UGD and were therefore excluded from the analysis; most notably, these mothers were more African American (7% compared to 3%), smoked more during pregnancy (25% versus 20%), and had a higher proportion receiving WIC assistance (41% versus 32%) (see Table 3, next page). In our study, 20% of mothers reported smoking during pregnancy (see Table 1) and, although slightly higher than the overall prevalence for the state of Pennsylvania (15%), it is similar to other reports of smoking in pregnancy for the counties and the time period under study [175]. According to the Pennsylvania Department of Health, the percent of mothers that smoked during pregnancy from 2010 to 2012 was 15% in Butler, 22% in Washington, and 20% in Westmoreland [175]. In a random sample of 5,007 birth certificates from 2005 to 2009 we obtained from the PADOH for a separate study, the proportions of mothers that smoked prior to and during pregnancy were also higher than the state: 20% for Butler, 32% for Washington, and 29% for Westmoreland.

Table 3. Maternal and child risk factors for geocoded versus not geocoded residences and those with versus without at least one well within 10 miles

Factor	Geocoded N=22,273	Not geocoded N=5,724	<10-miles N=15,451	≥10-miles N=1,969
Mother's age (years)	28.5 ± 5.8	28.1 ± 6.0	28.6 ± 5.8	27.5 ± 5.9
Mother's Education (% high school graduate/GED)	23.3%	25.6%	22.7%	27.4%
Pre-Pregnancy Weight (lbs)	154.1 ± 39.4	153.6 ± 39.4	153.8 ± 39.1	156.5 ± 41.9
Race (% African American)	3.5%	3.4%	3.0%	7.2%
WIC (% assistance)	33.2%	36.1%	32.1%	41.3%
Prenatal care (% at least one visit)	99.4%	99.1%	99.5%	99.4%
Presence of gestational diabetes	4.2%	4.4%	4.1%	4.4%
Cigarette smoking during pregnancy	20.9%	22.1%	20.0%	25.7%
Gestational age (weeks)	38.7 ± 1.9	38.7 ± 2.0	38.7 ± 1.9	38.5 ± 2.2
Birth weight (g)	3343.0 ± 553.9	3333.6 ± 558.9	3345.8 ± 549.2	3319.8 ± 594.8
Percent female	48.5%	50.0%	48.5%	48.5%
Birth parity (first)	42.6%	43.2%	42.7%	42.0%

Like McKenzie et al., we were persuaded that previous experience with multiple fixed sources of pollution and birth outcomes suggests that inverse density is the best surrogate for maternal exposure [103; 104]. Further, when we repeated the analyses using IDW well count as a continuous measure, the associations between increased exposure and smaller birth weights and increased odds of SGA (OR=1.009, 95% CI 1.003-1.015) remained significant ($p<0.01$). A

sensitivity analysis of 2010, the year with the most UGD activity in our study period, also showed an association between increased exposure and decreasing birth weights (p=0.03). A reanalysis (data not shown) adding county (categorically) to the adjusted linear regression led to similar conclusions regarding: a) association of lower birth weight and increased well density for the fourth quartile (p=0.02); and b) increased odds of SGA for the highest exposure group (OR=1.34, 95% CI=1.10-1.63, p=0.004).

Two other concomitant studies have findings similar to ours concerning birth weight. The PhD thesis of Elaine Hill at Cornell University compared birth outcomes for mothers who resided in regions in Pennsylvania in proximity to wells as a function of time (before and after permit and spud) [101]. Their model employed a difference-in-differences approach to compare groups that lived near permitted wells versus groups near permitted wells that underwent further development. An increase in prevalence of low birth weight at gestation and reduced 5 minute APGAR scores was reported while no impact on premature birth was detected for offspring of mothers living 1.5 miles or less from gas development [101]. In an abstract presented at a recent Annual Meeting of the American Economic Association, Currie et al. noted that proximity (within 1.5 miles) to a well increased low birth weight at term as measured in a multi-state sample [100]. Our study is the only one that is specifically limited to counties with intensive shale gas activities in Southwestern PA, thereby minimizing the heterogeneity of demography, geology, climate and other confounding variables.

It is only in recent years that drilling technology has rapidly advanced to be able to obtain substantial levels of natural gas tightly bound to deep underground shale layers. This continually evolving technology greatly differs from the past in using perhaps 5 million, rather than 50,000 gallons of hydrofracturing fluid under much higher pressures for each well; in having an

evolving suite of hydrofracturing chemicals, with over 500 having been used; in laterally bending the well within the shale layers for greater than a kilometer; in drilling in multiple directions from the same well head from larger drill pads for sequential periods of six months or longer; and in many other technological advances. Recent reviews of shale gas issues in the United States, Canada and Europe have been consistent in commenting on the lack of health-related information [86; 160].

2.5.2 Limitations.

This investigation is semi-ecological in nature. We had individual data on birth outcomes and risk factors; however, the final analysis grouped mothers into exposure categories to provide a clearer picture of possible dose-response relationships. In addition, there may be a number of unknown factors that led to our conclusion that well density was associated with lower birth weight and greater odds of SGA. As in any epidemiological study, these associations do not imply causation and are hypothesis generating only. The observed associations could be due to a contaminant related to UGD, an unknown confounding factor we were unable to account for in our analyses, or chance. Moreover, we assumed that the residence on the birth certificate was synonymous with exposure during the entire pregnancy, as we have no ability to evaluate transient occupancy of the pregnant mother. However, the counties under study have relatively stable populations. US Census data (2008-2012) for living in the same house one year and over for Butler, Washington and Westmoreland Counties shows 88.6%, 88.1% and 91.0% respectively as compared to 84.8% for the US and 87.8% for Pennsylvania [174].

Proximity is a primitive surrogate for exposure itself and is uninformative of route (water, air) or etiologic agent. Our observations were based on data deduced from the Department of

Environmental Protection (DEP) of Pennsylvania and assignments of longitude and latitude only from birth certificate data. Twenty percent of the birth certificate records did not have a corresponding geocode and, since no further information on address or zip code was available, these births were excluded from the analysis. However, the sociodemographic characteristics of this group were similar to those that were geocoded (Table 3). Up until recently, pertinent information from DEP was limited to date of permit request and drilling (SPUD) and status (active, plugged or abandoned). The available well permit number provides information on production and waste data [159]. Longitude and latitude defined proximity in our analyses, and we did not probe more complex issues of geology, climate or meteorological conditions; thus, the transmigration of potential pollutants in water or air remains unclear.

Other limitations in the birth dataset included the lack of a birth month and day; we were therefore only able to identify those wells drilled during the birth year of the infant. Active drilling of a well occurs over a period of only a few months, so incorporating more specific timings of exposure will be critical in future work as further data become available as to the time period during which air or water exposures are most likely. Birth weight data are reasonably precise as derived from birth certificates, but such certificates appear less reliable for gestational age [176], so derived information such as SGA may be spuriously affected. We also relied on birth certificates to incorporate non-exposure relative risks for mother and child. Although it is encouraging that in multivariate analyses, many of these contributing factors affected outcomes in a predictable fashion [36], incomplete information on many of these factors may have affected our conclusions in Table 2 and Figures 2 and 3. For example, socioeconomic status was inferred by use of assistance via WIC; smoking was neither quantitatively assessed nor confirmed beyond

self-reporting; the details of prenatal care, co-morbidities and nutritional status are not on birth certificates. As such, larger studies that include medical records will be helpful.

The relative monotonic increase in SGA (Table 1) and odds ratio for SGA (Figure 2) lends credence to the possibility that this association is indeed related to increased exposure to aspects of UGD. Similarly, a significant decrease in birth weight, after adjusting for covariates, was discernable only in the highest exposure quartile (Table 2). In contrast, changes in odds ratios for prematurity were not significant, except for a very small protective effect in the second quartile (Figure 3).

If the association of lower birth weight and proximity to well is indeed secondary to environmental exposure, then identifying the route of exposure and the agents, alone or in combination, is a critical and challenging next step. In the preliminary study of Currie et al., no differences between mothers with access to public or well water was found, suggesting that exposures may not be water derived [100]. Air pollution is well known to affect perinatal outcomes [62; 66; 67; 177], and a meta-analysis of 62 studies recently pointed to particulate matter, carbon monoxide and nitrogen dioxide [178]. Potential UGD derived air pollutants that are known to be associated with low birth weight include diesel exhaust [178], heavy metals [73-75], benzene [76] and other volatile organic compounds [179].

In conclusion, a small but significant association between proximity to UGD and decreased birth weight was noted after accounting for a large number of contributing factors available from birth certificate data in southwest Pennsylvania. Although the medical and public health significance of this is unclear, it was noteworthy that there was a significant increase in incidence of SGA in the highest exposed group. Along with the first published study on the association of increased incidence of birth defects and proximity and density of nearby wells in

Colorado [102], there is a clear need for more complete studies including larger populations, better estimates of exposure and covariates and more refined medical records. The difference in outcomes as they relate to birth weight between our study and Colorado (but similar findings to ours in the original work of Hill [101] and preliminary results of Currie et al. [100]) underscores the importance of region-specific assessment of UGD impacts on public health. Although neither the route (water, air or soil) of exposure nor etiologic agents could be addressed, this study is among the first to report a human health effect associated with hydrofracturing. The embryo/fetus is particularly sensitive to the effects of environmental agents, which can have significant lifetime consequences [3]; therefore, further investigation appears warranted.

3.0 INVESTIGATING PRENATAL EXPOSURE TO GROUPS OF AIR TOXICS AND AUTISM SPECTRUM DISORDER USING EXPLORATORY FACTOR ANALYSIS

3.1 ABSTRACT

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by impaired social interaction and rigid behaviors and routines. The prevalence of ASD has increased markedly over the last several decades, motivating investigations into possible causes and risk factors. One area of interest has been the role of the environment, particularly ambient air pollution, in the development of childhood ASD. However, few studies have explored groups of air toxics, such as those emitted from a common source, in relation to ASD risk. In the present study, estimates of 30 ambient air toxics from the 2005 National Air Toxics Assessment (NATA), modeled at the census tract level, were linked to 217 cases of ASD and 224 controls born in southwestern Pennsylvania from 2005 to 2009. An exploratory factor analysis (varimax rotation) was conducted to reduce these 30 pollutants to a set of key predictors (factors). Factor scores were calculated using two methods: index scores based on sums of quartiles of exposure and linear regression. These scores informed two sets of logistic regression models to determine which factors were associated with increased ASD risk, unadjusted and adjusted for mother's age, race, education, and smoking. The results of each method for calculating factor scores were compared. A Spearman correlation matrix revealed that many of the 30 NATA air toxics were

highly correlated with each other. The air toxics loaded onto 7 main factors. Regardless of the method used to calculate the scores, the factors most associated with increased ASD risk appeared to represent traffic, other combustion sources, and certain types of manufacturing (plastics, rubbers, and adhesives). These sources should be targeted in future investigations of ASD risk and air pollution.

3.2 INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by impaired social interaction and communication and by restricted and repetitive behaviors [105]. The prevalence of ASD has increased markedly over the last several decades [107; 108] and, although part of this increase is attributed to greater awareness and changes in diagnostic criteria, it is estimated that approximately 46% is due to unknown factors [109]. The Centers for Disease Control (CDC) has estimated that ASD currently affects one in every 68 children [108]. The underlying genetic causes of ASD are complex and multifactorial. A definitive "autism gene" has not been identified; rather, a number of concomitant disorders and genetic mutations have been associated with the development of ASD [110]. Other factors found to be associated with an increased risk of ASD include: male sex [116], advanced maternal or paternal age [117], greater parental educational attainment [118], adverse birth outcomes such as low birth weight and preterm birth [119; 120], pregnancy complications [121], and maternal smoking [122].

Recent investigations in this field have sought to elucidate the environment's role in the development of ASD. Several studies have utilized the USEPA's National Air Toxics Assessment (NATA) database to explore associations between NATA estimated levels of

hazardous air pollutants (HAPs) and ASD risk [130-132]. Others have looked at traffic-related pollutants, including particulate matter with an aerodynamic diameter less than 10 (PM_{10}) or less than 2.5 μm ($PM_{2.5}$) [133-137], and industrial mercury emissions [123; 124].

The effects of individual air toxics on ASD risk have been the major focus of many of these studies [123; 124; 130-132], although several have also considered mechanistic groups of toxicants (e.g. metals or aromatic solvents) [132]. An association found between ASD and an individual air pollutant could be either a true association or a surrogate for a pollution source or something else that is the true etiologic agent. Since people are usually simultaneously exposed to a complex mixture of air pollutants, there has been a push in recent years toward a multi-pollutant approach to investigating air quality and human health [139]. A multi-pollutant approach would better address this complexity, aid air quality policies, and allow for the identification of specific sources of harmful pollution [139]. For these reasons, looking at groups of pollutants that share a certain characteristic, such as a common source, may be beneficial to advancing our knowledge of the environmental risk factors for ASD.

Dimension reduction techniques, such as factor analysis, reduce a larger set of *j* variables to a smaller set of *k* latent constructs or "factors" [180]. Ideally, the *k* factors explain much of the variance of the original *j x j* correlation matrix, and the factors can therefore be used to represent the original set of variables [180]. In an exploratory factor analysis (EFA), the results are driven by the mathematics of the method and do not require strong *a priori* expectations, although investigators may have some idea of what factors will emerge from the dataset [180]. As part of their study of prenatal exposure to toxic air pollutants and risk of childhood ASD, von Ehrenstein et al. (2014) used EFA to examine the correlation structure among 24 air toxics of interest, which loaded on 4 main factors. The air toxic most significantly related to ASD risk in each factor was

chosen to represent that factor in subsequent multi-pollutant logistic regression models [140]. Since pollutants can have multiple sources, such an "indicator" pollutant may not be a truly unique identifier for that factor or source [139].

The present work emerged from a population based case-control study conducted in southwestern Pennsylvania from 2005 to 2009 (Talbott et al., in review), which utilized the 2005 NATA assessment to explore relationships between census tract-level estimates of air toxics and ASD risk. EFA was used to reduce this set of highly correlated air toxics to a smaller set of constructs or factors. Rather than using an indicator pollutant from each factor, this work sought to incorporate the factors themselves into logistic regression models of ASD risk while also accounting for other sociodemographic and behavioral risk factors for ASD.

3.3 METHODS

3.3.1 Study population.

Case ascertainment and the study population have been described elsewhere (Talbott et al., in review) and will be summarized here. The study was approved by the University of Pittsburgh Institutional Review Board (IRB number PRO10010240). ASD cases were children born between January 1, 2005 and December 31, 2009 in a six-county area (Allegheny, Armstrong, Beaver, Butler, Washington, and Westmoreland) of southwestern Pennsylvania and were recruited from specialty autism clinics, treatment centers, or through the Pennsylvania School System. A case was required to have a documented diagnosis of ASD and a score of 15 or above on the Social Communication Questionnaire (SCQ), a positive screen for the presence of autistic

features. Controls were recruited from a random selection of births from the same six-county area and time period (2005-2009) using the Pennsylvania Department of Health (PA DOH) state birth registry. Children with an SCQ score of 15 or above were excluded as controls. Mothers of cases and controls were administered a personal interview by a trained interviewer using a structured questionnaire adapted from the CDC's Study to Explore Early Development. Residential history and all sociodemographic and behavioral (e.g. smoking) information were obtained from the interview.

3.3.2 Exposure assessment and statistical analysis.

Exposure to ambient air toxics was estimated using census tract-level, modeled data from the 2005 NATA assessment (http://www.epa.gov/ttn/atw/nata2005/tables.html). Out of the 177 air toxics available through NATA, 30 compounds, all of which had possible neurological, developmental, or endocrine-disrupting effects as well as diverse distributions in the study area, were used in the analysis. Details regarding the geocoding of residential addresses and linkage to dose estimates for each of the 30 air toxics are described in Talbott, et al. (in review). In the end, all mother-child pairs were assigned a geospatial and time-specific dose estimate for each of the 30 NATA air toxics and key developmental time periods (3 month period before last menstrual period/LMP, 1^{st} trimester, 2^{nd} trimester, 3^{rd} trimester, full pregnancy, 1^{st} year of life, and 2^{nd} year of life). The present analysis focuses on the full pregnancy period.

Spearman correlations were conducted to determine which of the 30 air toxics were correlated with each other. Since the concentrations of many of the NATA compounds were highly correlated, EFA was used to further examine their correlation structure and to reduce the set of 30 air toxics to a smaller set of latent factors. Factors were extracted using Principal

Component Analysis (PCA) and rotated using varimax rotation. The eigenvalue >1 rule was used to determine which factors to retain [140; 180]. An individual air toxic was identified as a component of a factor if it loaded greater than 0.5 on that factor. Factors defined by only one variable were excluded from further analysis [180].

Two different approaches were used to calculate factor scores. In "Method 1," NATA concentrations during the pregnancy period were first divided into quartiles based on the distribution of the controls. Index scores were calculated for each of the factors by summing the quartile ranks for each compound that loaded on the factor [132]. For example, if mercury and cadmium both loaded onto hypothetical Factor X, the quartile ranks (1, 2, 3, or 4) for these two metals were added together. Mothers living in census tracts within the fourth quartile of both mercury and cadmium would therefore have an index score of 8 for Factor X and would rank higher than those in the first quartile for both metals (index score of 2). In "Method 2," each factor score was calculated using the linear combination of air toxics loading on the factor. Therefore, every participant was assigned two scores for each factor, one using each of Methods 1 and 2.

Logistic regression analysis was performed to investigate associations between the factors identified above and ASD risk, unadjusted and adjusted for a set of *a priori* maternal risk factors: age, race, education, and smoking. Factor scores were analyzed as continuous variables. Since some of the mothers moved one or more times during the course of the pregnancy, a sensitivity analysis was conducted considering only those participants that did not move. The logistic regression results using Methods 1 and 2 to calculate factor scores were compared, and consistencies regarding which factors were positively associated (odds ratio >1) with ASD were

noted. All statistical analyses were performed in IBM SPSS Statistics 21 (IBM Corporation, Armonk, New York).

3.4 RESULTS

A total of 217 cases and 226 controls were consented and interviewed for the study. NATA data was unavailable for two controls that lived outside of the United States, and they were excluded from further analysis, leaving a final group of 224 controls.

3.4.1 Selection of factors.

The Spearman correlation matrix for the 30 NATA air toxics is available in Appendix A. Many of the metals (arsenic, chromium, cadmium, lead, etc.) were positively correlated with each other. The aromatic hydrocarbons (benzene, ethyl benzene, toluene, styrene, and xylenes) were also highly correlated with each other and with some of the metals. Allyl chloride was negatively correlated with several of the air toxics, including cadmium, mercury, toluene, and vinyl chloride, among others. Since the majority of the air toxics were correlated with each other, further examination using exploratory factor analysis was warranted.

Table 4 shows the initial eigenvalues and the percent of the variance explained both before and after rotation (varimax) of the first ten factors. Using the eigenvalue >1 rule, 9 factors were retained that together explained 81% of the variance. Factor 10 was the first factor not retained using this rule of thumb. Factor 1 alone explained almost 30% of the variance, while Factors 2 and 3 explained the next highest (9% each, after rotation).

Table 4. Initial eigenvalues and variance explained for the first ten factors (varimax rotation)

Factor	Initial Eigenvalues	Before Rotation		After Rotation	
		Percent of Variance	*Cumulative Percent*	*Percent of Variance*	*Cumulative Percent*
1	9.63	32	32	29	29
2	3.16	11	43	9	39
3	2.88	10	52	9	48
4	2.15	7	59	7	56
5	1.65	6	65	6	62
6	1.39	5	70	6	68
7	1.27	4	74	5	73
8	1.03	3	77	4	77
9	1.03	3	81	4	81
10	0.84	3	83		

Table 5 shows the rotated factor matrix used to determine which air toxics loaded (>0.5) on each of the 9 factors. A diverse array of pollutants loaded onto Factor 1 for a total of 13 pollutants: arsenic, benzene, ethyl benzene, toluene, xylenes, methylene chloride, perchloroethylene (PERC), diesel particulate matter (PM), cresol, cyanide, hexane, 1,1,1-trichloroethane, and methanol. The remaining factors had smaller groupings of pollutants. Manganese, lead, chromium, and nickel loaded on Factor 2, while the remaining metals (mercury, cadmium, and selenium) loaded on Factor 4. Dinitrotoluene, allyl chloride, and carbon disulfide loaded on Factor 3; trichloroethylene, nickel, and chromium on Factor 5; PAHs, cyanide, and cresol on Factor 6; and styrene and cumene on Factor 7. Since only one pollutant loaded onto each of Factors 8 (hydrazine) and 9 (vinyl chloride), these factors were excluded from further analysis, leaving 7 main factors that explained almost 75% of the variance (see Table 5) [180].

Table 5. Factor matrix rotated to the varimax criterion (coefficients >0.5 are in gray)

	Factor								
Pollutant	1	2	3	4	5	6	7	8	9
Arsenic	0.693	-0.044	0.018	0.466	0.330	-0.008	-0.037	0.161	-0.127
Cadmium	0.317	-0.006	-0.030	0.819	0.209	-0.050	-0.001	0.068	-0.116
Chromium	-0.073	0.755	0.012	-0.007	0.516	0.015	0.305	-0.062	-0.046
Mercury	-0.083	-0.025	-0.063	0.858	0.019	-0.105	0.020	-0.182	-0.022
Manganese	0.086	0.923	0.004	-0.004	-0.087	-0.008	-0.050	0.015	-0.001
Nickel	-0.094	0.719	-0.007	-0.062	0.608	0.010	0.129	0.015	0.038
Lead	0.097	0.920	-0.008	0.039	-0.054	-0.021	-0.111	0.041	-0.046
Selenium	-0.089	0.039	-0.034	0.685	-0.137	0.140	0.065	0.081	0.149
Benzene	0.949	0.045	0.011	-0.035	0.009	0.116	0.060	0.031	0.058
Ethyl benzene	0.688	-0.028	0.037	-0.051	0.067	-0.089	0.205	0.337	-0.055
Styrene	0.285	0.064	0.069	0.114	0.149	0.040	0.771	-0.145	-0.123
Toluene	0.965	0.034	0.010	-0.021	0.039	0.029	0.043	0.029	0.058
Xylenes	0.952	-0.004	0.096	0.022	0.141	0.010	0.098	-0.001	-0.026
Methylene chloride	0.548	0.089	0.321	0.110	0.037	0.086	0.228	0.161	0.148
PERC	0.725	0.032	0.025	-0.085	-0.049	0.046	-0.164	-0.037	0.200
Trichloroethylene	0.330	0.060	0.005	0.057	0.883	0.044	0.047	-0.024	0.074
Vinyl chloride	0.099	-0.045	0.002	0.023	0.042	-0.033	0.035	0.000	0.904
Hydrazine	0.128	0.032	0.010	-0.013	-0.034	0.001	-0.111	0.931	-0.001
PAHs	0.033	-0.011	0.026	-0.021	-0.011	0.935	0.029	0.022	-0.060
Diesel PM	0.679	0.083	-0.048	-0.021	-0.070	0.125	0.080	-0.031	-0.040
Allyl chloride	0.009	0.005	0.983	-0.010	0.019	-0.068	0.003	-0.033	0.000
Carbon disulfide	0.130	-0.026	0.860	-0.134	-0.027	0.297	0.065	0.089	-0.002
Cresol	0.570	0.030	0.098	0.001	0.007	0.598	0.030	-0.107	0.253
Cumene	0.382	-0.087	0.042	0.015	0.036	0.066	0.509	-0.037	0.235
Cyanide	0.665	-0.045	0.053	0.061	0.194	0.629	0.083	0.015	-0.124
Dinitrotoluene	0.011	0.005	0.983	-0.013	0.019	-0.068	0.003	-0.034	0.000
Ethylene oxide	0.472	-0.020	0.065	0.098	0.328	0.001	-0.435	-0.147	-0.136
Hexane	0.932	0.052	-0.027	-0.001	0.002	-0.031	0.186	-0.048	0.013
Trichloroethane	0.885	-0.028	0.035	0.128	0.171	0.052	0.043	0.092	0.093
Methanol	0.795	-0.068	0.071	0.147	0.296	0.292	0.085	0.054	-0.101

3.4.2 Descriptives of factor scores.

Table 6 shows the means, standard deviations, and ranges for the factor scores calculated using Methods 1 and 2. Method 1 calculated an index score for each factor by summing quartiles of exposure across all air toxics that loaded >0.5 on that factor [132]. Index scores for Factor 1 (n=13 air toxics) ranged from 13 (all pollutants in the first quartile) to 52 (all pollutants in the fourth quartile), while the range for Factor 7, which represented only 2 air toxics, was more narrow (2 to 8). In Method 2, scores were calculated by SPSS using the linear combination of pollutant variables multiplied by their coefficients (see Appendix A for the factor score coefficient matrix). As evident from Table 6, each factor's score distribution using Method 2 is standardized to a mean of 0 and a standard deviation of 1. Therefore, participants with more negative scores are relatively less exposed to a particular factor than those with more positive scores. For example, a participant assigned the minimum score for Factor 1 (-3) is less exposed to Factor 1 relative to someone assigned the maximum value (+4).

Table 6. Mean, standard deviation (S.D.), minimum, and maximum factor scores calculated using Methods 1 and 2

Factor	Method 1				Method 2			
	Mean	*S.D.*	*Minimum*	*Maximum*	*Mean*	*S.D.*	*Minimum*	*Maximum*
1	33	13	13	52	0	1	-3	4
2	10	4	4	16	0	1	-1	13
3	8	3	3	12	0	1	-1	19
4	7	3	3	12	0	1	-2	7
5	8	3	3	12	0	1	-2	16
6	8	3	3	12	0	1	-1	19
7	5	2	2	8	0	1	-4	8

3.4.3 Associations with ASD.

Figure 4 presents the unadjusted odds ratios (OR) and 95% confidence intervals (CI) for factor scores calculated using Method 1 (summing quartiles) and autism spectrum disorder. Associations with ASD were the most elevated for Factor 3 (OR=1.06, 95% CI=0.99-1.13), Factor 6 (OR=1.06, 95% CI=0.99-1.12), and Factor 7 (OR=1.11, 95% CI=1.01-1.22). Point odds ratios were also slightly elevated above 1.00 for Factors 2 (OR=1.02, 95% CI=0.97-1.08) and 5 (OR=1.04, 95% CI=0.97-1.11). After adjustment for mother's age, education, race, and smoking (Figure 5), the odds ratios were slightly attenuated, particularly for Factors 2 and 5. Factors 3, 6, and 7 remained similarly elevated (Factor 3: OR=1.04, 95% CI=0.97-1.11; Factor 6: OR=1.03, 95% CI=0.96-1.10; Factor 7: OR=1.08, 95% CI=0.98-1.18).

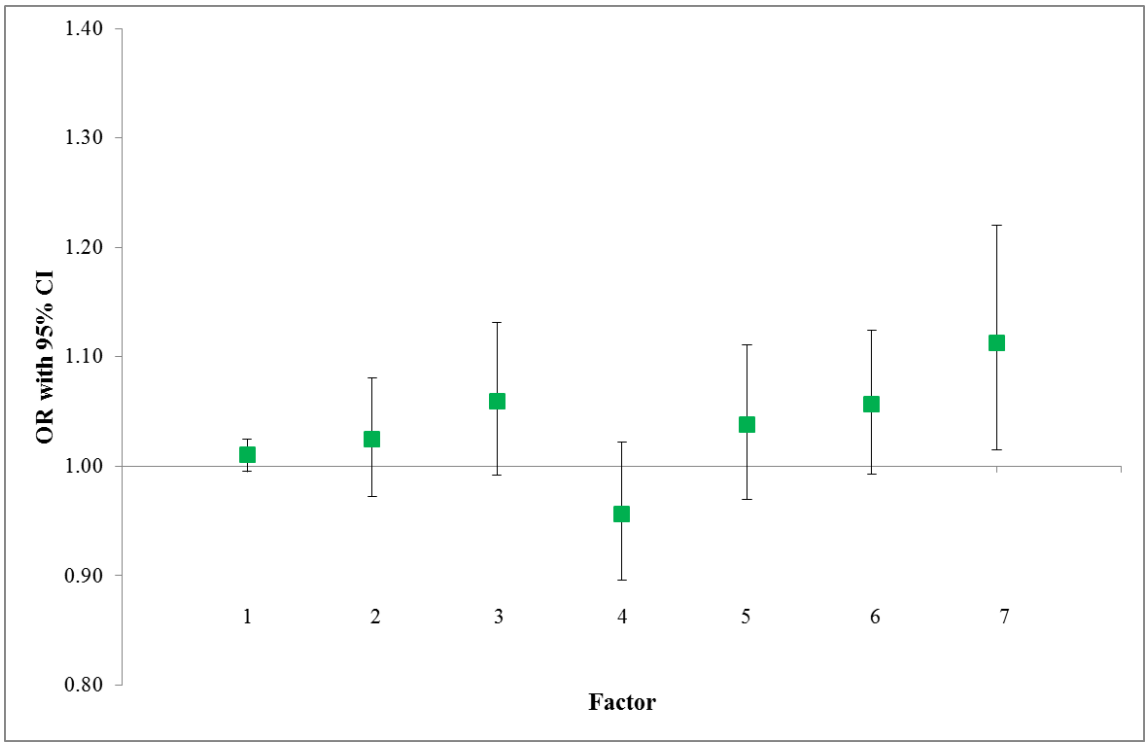


Figure 4. Unadjusted OR and 95% CI for 7 factors (Method 1); 217 cases, 224 controls

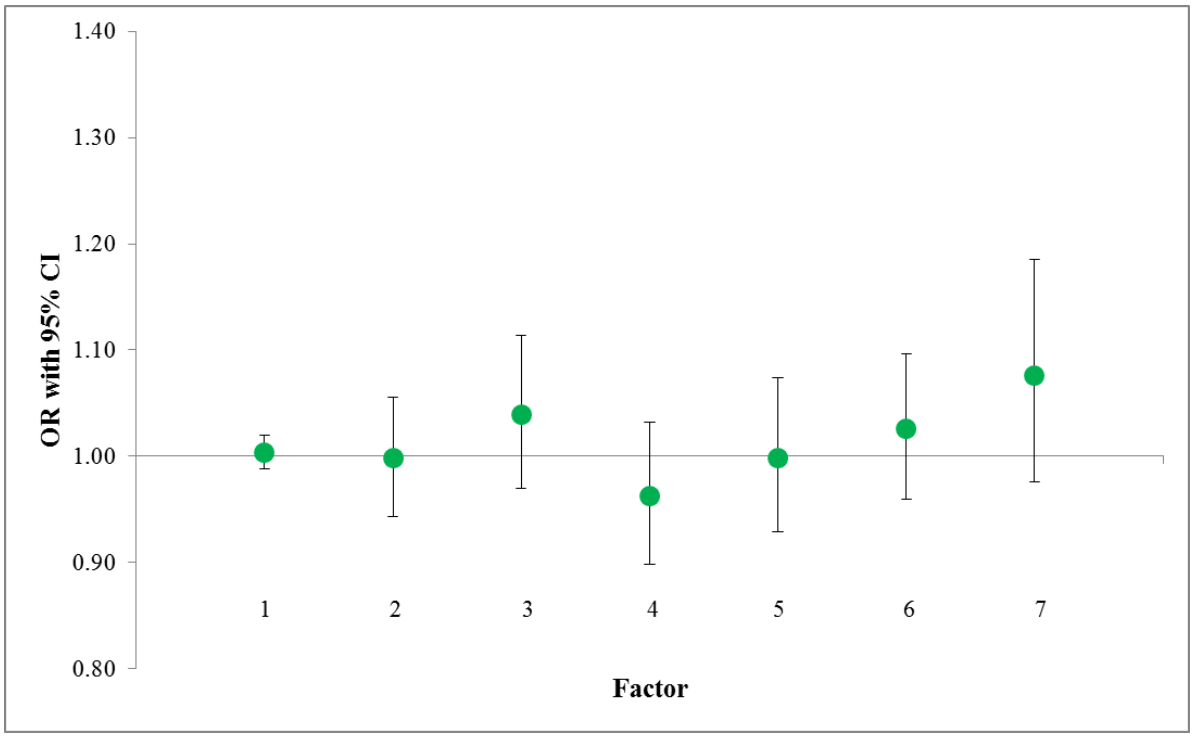


Figure 5. Adjusted OR and 95% CI for 7 factors (Method 1); 217 cases, 224 controls

The unadjusted and adjusted logistic regression results for ASD and factor scores calculated using Method 2 (linear combination) are shown in Figures 6 and 7, respectively. Unadjusted odds ratios were elevated for Factor 1 (OR=1.14, 95% CI=0.94-1.37), Factor 3 (OR=1.15, 95% CI=0.85-1.56), and Factor 6 (OR=1.32, 95% CI=0.89-1.97). After adjustment, the associations were again attenuated slightly, although the odds ratios and widths of the 95% confidence intervals were still similar in magnitude for Factor 1 (OR=1.04, 95% CI=0.85-1.28), Factor 3 (OR=1.09, 95% CI=0.80-1.48), and Factor 6 (OR=1.18, 95% CI=0.83-1.68). Conversely, the association with Factor 7 was more elevated after adjustment (OR=1.07, 0.88-1.30).

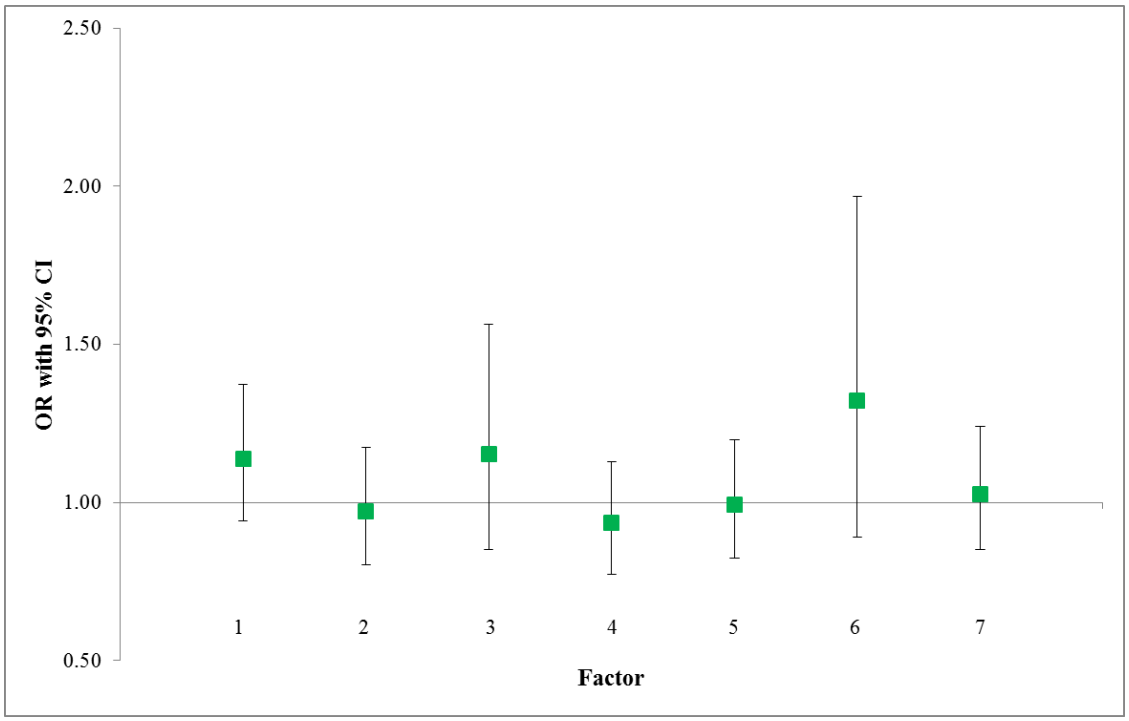


Figure 6. Unadjusted OR and 95% CI for 7 factors (Method 2); 217 cases, 224 controls

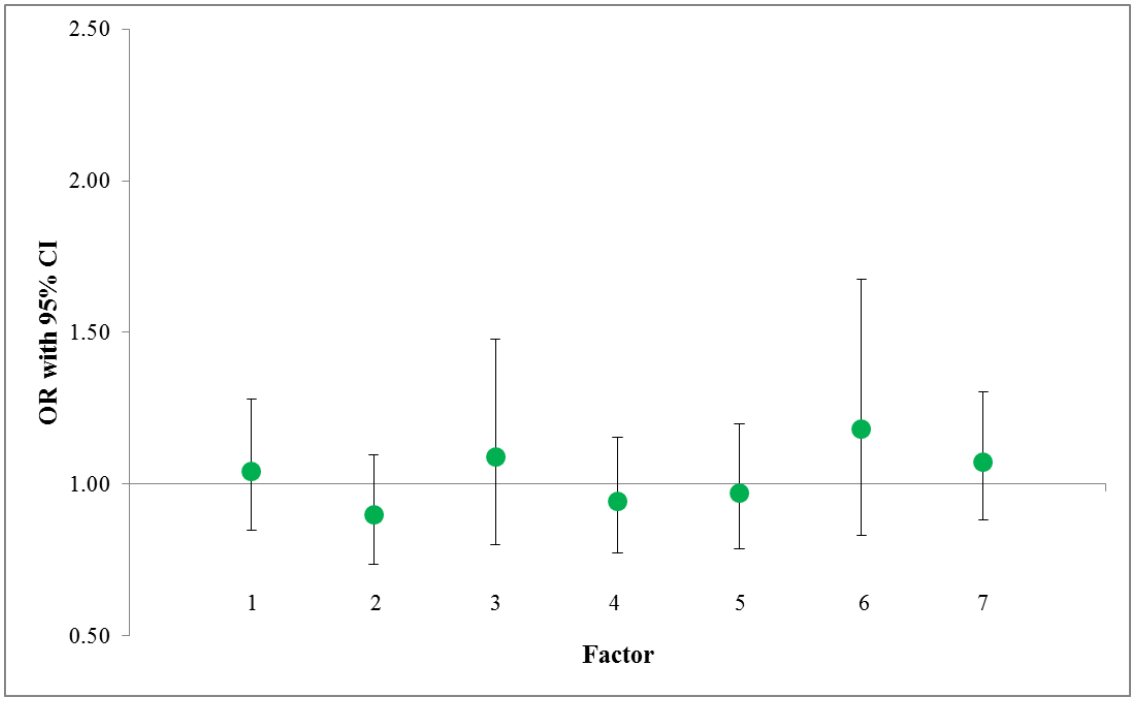


Figure 7. Adjusted OR and 95% CI for 7 factors (Method 2); 217 cases, 224 controls

A sensitivity analysis was conducted including only the 377 mothers (86% of the total sample) that did not move at all from the three months prior to pregnancy through the child's second year of life. Of these, 182 (48%) were cases and 195 (52%) were controls. The unadjusted and adjusted logistic regressions were repeated for the "non-movers," and the results using either method for calculating the factor scores were similar to those using all cases and controls (Figures 8-11).

The same factors had elevated associations with ASD in the sensitivity analysis of non-movers as in the analysis with the total sample. The odds ratios for Factors 3, 6, and 7 remained the most elevated in the unadjusted logistic regression using Method 1 (Figure 8). Factor 2 was less elevated initially than it was in the first analysis. Factors 3, 6, and 7 were similarly elevated after adjustment for mother's age, education, race and smoking (Figure 9).

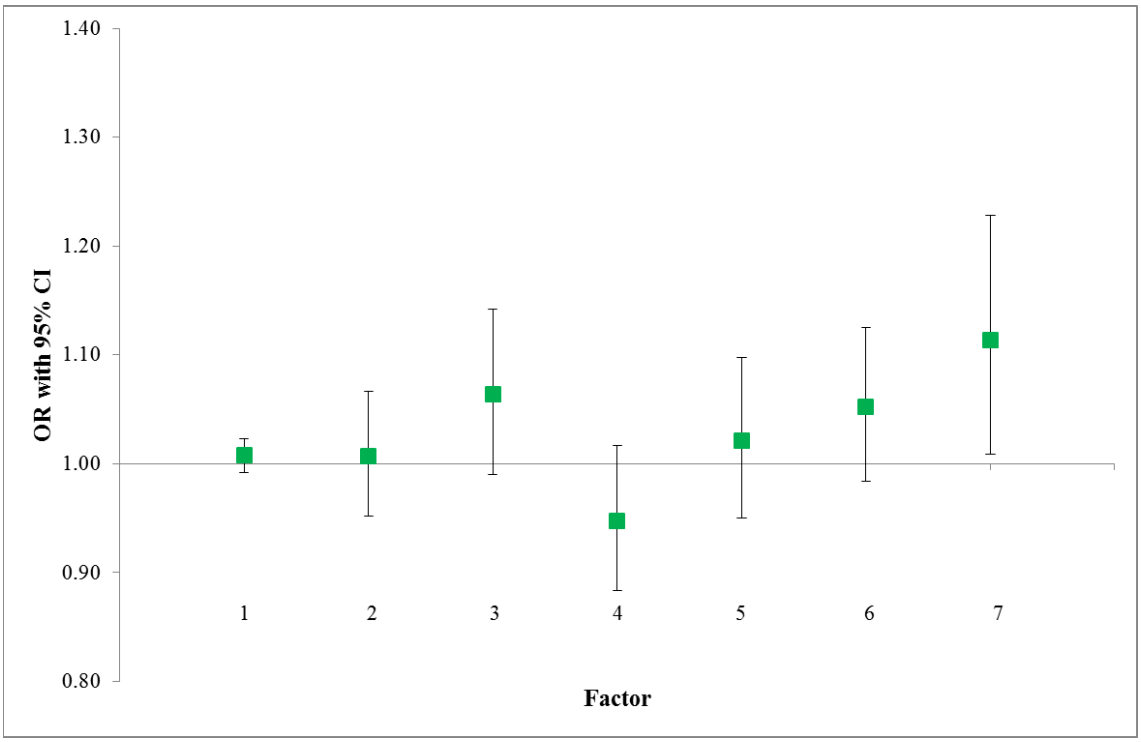


Figure 8. Unadjusted OR and 95% CI for 7 factors (Method 1); non-movers (182 cases, 195 controls)

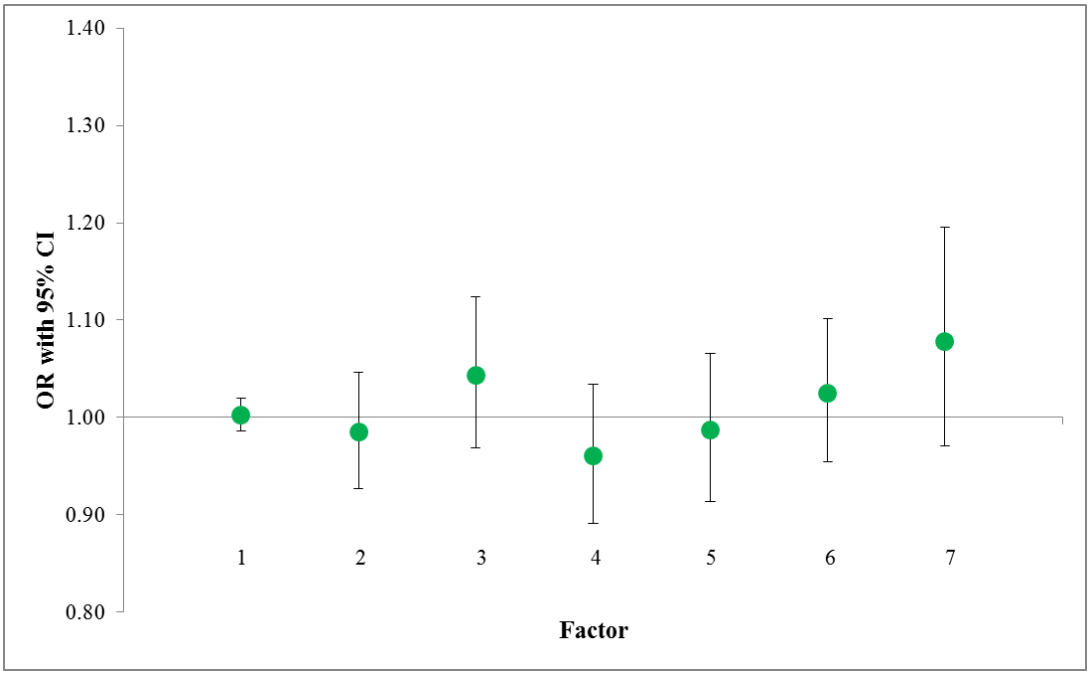


Figure 9. Adjusted OR and 95% CI for 7 factors (Method 1); non-movers (182 cases, 195 controls)

Using Method 2 to calculate the factor scores, the unadjusted logistic regression results for non-movers again showed elevated odds ratios for Factor 1 (OR=1.09, 95% CI=0.89-1.34), Factor 3 (OR=1.52, 95% CI=0.84-2.74), and Factor 6 (OR=1.37, 95% CI=0.88-2.13). The 95% confidence intervals were wider for Factors 3 and 6 compared to the analysis including everyone. After adjustment, the odds ratio for Factor 1 was attenuated to just above 1.00. The odds ratios and confidence intervals for Factors 3 and 6 were shifted slightly but still in the same direction.

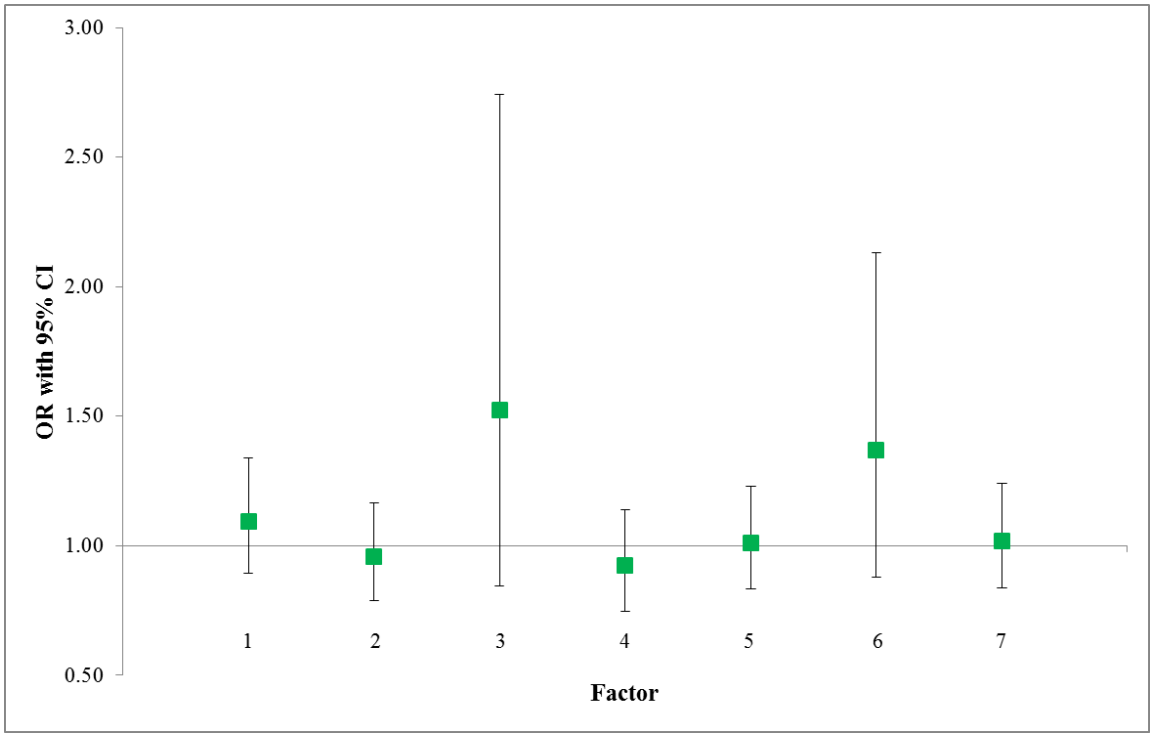


Figure 10. Unadjusted OR and 95% CI for 7 factors (Method 2); 182 cases, 195 controls

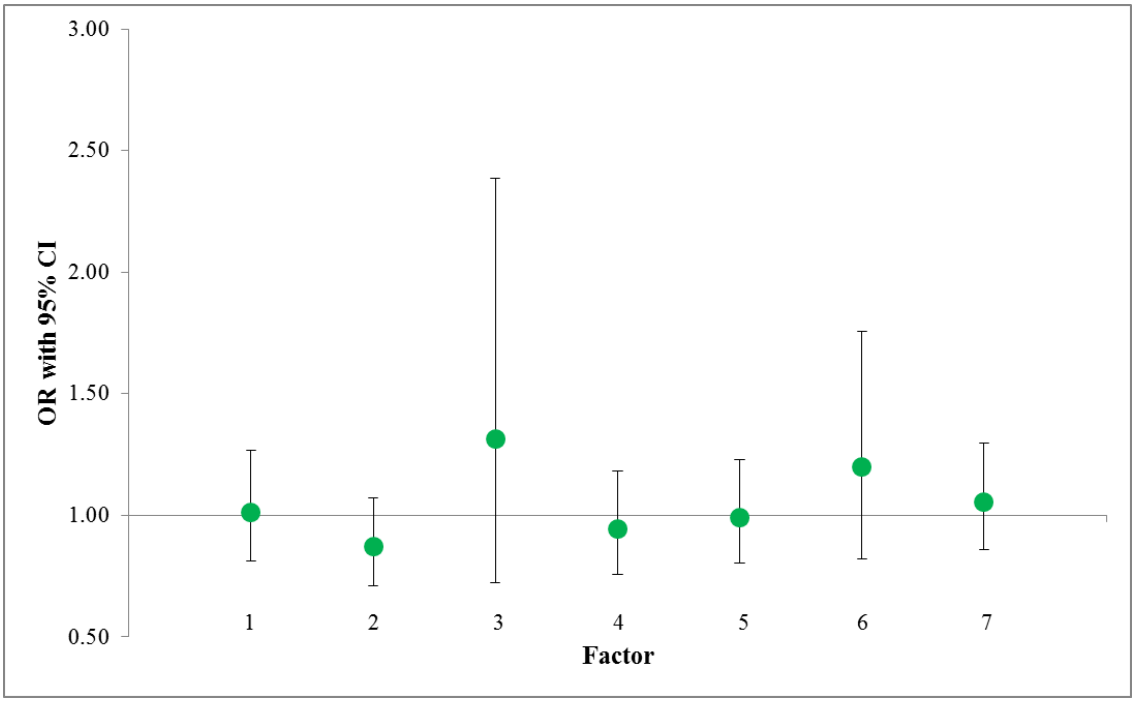


Figure 11. Adjusted OR and 95% CI for 7 factors (Method 2); 182 cases, 195 controls

3.5 DISCUSSION

3.5.1 Potential sources of exposure.

In the present study, exploratory factor analysis was employed to investigate ASD risk from a multi-pollutant standpoint and to identify potential sources of harmful exposure. Increased risk for ASD was primarily associated with traffic- and industry-related exposures. Calculating factor scores using Methods 1 and 2 both resulted in elevated associations between ASD and Factors 3 and 6. Factor 3 represented dinitrolutene, allyl chloride, and carbon disulfide. Although these three compounds have many varied uses, common sources include the production of plastics, rubbers, and adhesives [181]. PAHs, cyanide, and cresol loaded on Factor 6 and likely represent vehicle exhaust or other combustion sources, such as municipal trash incinerators, the burning of fossil fuels, and even cigarette smoke [181].

Using Method 2, a potential association was observed between ASD and Factor 1, which also represents many air toxics found in vehicle exhaust [136; 140; 145]. A relationship with Factor 7 (styrene and cumene) was found only using Method 1. In industry, styrene is primarily used in the manufacture of plastics, rubber (for tires and other automobile parts), and resins used in boats, tubs/shower stalls, liners, putty, and other products [144]. Cumene is used as a thinner for lacquers, paints, and enamels and is a constituent of crude oil and finished fuels [182]. Styrene and cumene are also constituents of automobile exhaust and cigarette smoke [144; 182].

3.5.2 Comparison with existing studies.

Prior studies have also observed associations between increased ASD risk and proximity to roads or traffic-related pollutants [136; 137; 140], such as those represented by Factors 1, 6, and 7 in our study. One investigation of in utero exposure to toxic air pollutants and ASD risk found that autism risks increased per interquartile range increase of individual pollutants loading on the same factor (1,3-butadiene, meta/para-xylene, other aromatic solvents, lead, PERC, and methylene chloride), many of which are found in vehicle exhaust [140]. Several of these pollutants were also found to covary in our study (benzene, ethylbenzene, toluene, xylenes, PERC, and methylene chloride). Rather than calculating factor scores and incorporating these into logistic regression models, von Ehrenstein et al. selected individual pollutants most associated with ASD from each factor and mutually adjusted for these in 2- and 3-pollutant models. As a result, although significant individually, associations with 1,3-butadiene, xylenes, and lead tended to weaken in 2- and 3-pollutant models, while associations with pollutants loading on other factors (formaldehyde) or no factor (trichloroethylene) were unaffected. In our study, trichloroethylene loaded on Factor 5, along with chromium and nickel, but there did not appear to an association with ASD, except for an elevated odds ratio in the unadjusted analysis using Method 1. While NATA modeled levels of air toxics were used to estimate exposure in our study, von Ehrenstein et al. utilized measurements from nearby (<5 km of the mother's residence during pregnancy) air monitoring stations [140].

Alternatively, Factor 7 in our study could represent industrial sources of styrene and cumene exposure. A possible association between styrene and increased ASD risk has been noted in two other studies. In an investigation of 1996 NATA estimates of air toxics and ASD risk in North Carolina and West Virginia, Kalkbrenner et al. (2010) found elevated odds of ASD

for children exposed to the highest quartile of styrene (OR=1.8, 95% CI=1.0-3.1). Single-pollutant analyses in our case-control study of southwestern Pennsylvania revealed significant associations between exposure to the highest quartile of styrene and increased ASD risk across key developmental milestones from pregnancy through the child's second year of life (Talbott et al., in review). Styrene was excluded from further analysis in the von Ehrenstein et al. study because more than 30% of its monitored levels were missing.

To date, cumene nor any of the air toxics loading on Factor 3 (dinitrotoluene, allyl chloride, carbon disulfide) have been examined as possible risk factors for ASD in any of the major epidemiologic studies. Further, we did not find elevated associations between ASD risk and groups of metals (Factors 2 and 4). The crude odds ratio for Factor 2 (manganese, lead, chromium, and nickel) was elevated, but the association was attenuated after adjustment for mother's age, education, race, and smoking (see Figures 4 and 5). An association between ASD and an overall measure of metals has been noted in previous studies [131; 132].

3.5.3 Limitations and conclusions.

Limitations of this study include the semi-ecological design. Personal risk factors were obtained at the individual level, while exposure to air toxics was assessed at the census tract (i.e. group) level. Further, NATA modeled estimates of air toxics for the most recent assessment year, 2005, were applied to all study years (2005 to 2009). Factor analysis and other dimension reduction techniques have their own set of limitations. Source-based approaches substitute one complex mixture, the air, with another complex mixture, the source, and results from one area may not be generalizable to other areas because of the location-specific nature of source signals [136; 137; 139; 140]. In addition, the results of an exploratory factor analysis are primarily driven by the

mathematical structure of the dataset rather than by the *a priori* expectations of the investigators [180]. Thus, the analyses described in this paper are intended to be hypothesis-generating. Since association does not necessarily equal causation, future work should determine: 1) which air toxics are the potentially responsible etiologic agents, 2) if two or more chemicals are interacting synergistically to increase ASD risk, or 3) if another agent associated with these exposure sources (but not currently accounted for in our analyses) is the true etiologic agent [139]. Utilizing databases such as the USEPA's Toxic Release Inventory (TRI) could aid in the identification of major industrial point sources of interest (i.e., those emitting air toxics loading on Factors 1, 3, 6, and 7).

This investigation adds to the small but growing body of literature regarding ambient air pollution and autism spectrum disorder. Risk for the disorder appears to be elevated for exposures related to traffic, other combustion sources, as well as certain types of manufacturing. Attenuation of many of the odds ratios after adjustment points to the importance of other confounding elements related to ASD risk, including sociodemographic characteristics (age, race, education) and behavioral factors, such as smoking. Nevertheless, these sources could be targeted in future investigations of ASD risk and air pollution. Further confirmation of traffic and industrial sources of exposure may inform future air quality regulations and policy.

4.0 AUTISM SPECTRUM DISORDER IN SOUTHWESTERN PENNSYLVANIA: A SUBANALYSIS OF EXPOSURE TO MULTIPLE SOURCES OF STYRENE

4.1 ABSTRACT

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by impaired social interaction and rigid behaviors and routines. The prevalence of ASD has increased markedly over the last several decades, motivating investigations into possible causes and risk factors. One particular area of interest has been the role of harmful ambient air toxics in the development of childhood ASD. A recent population based case-control study of 217 ASD cases and 226 controls conducted in southwestern Pennsylvania found an association between modeled estimates of ambient styrene exposure and increased ASD risk, taking into account relevant maternal risk factors. The present study is an extension of that work and investigates associations between ASD risk and exposure to multiple sources of environmental styrene (industrial and traffic), using proximity to source as a surrogate for exposure. Information on major industrial sources of styrene emissions was obtained from the Environmental Protection Agency's Toxics Release Inventory. Roadway and traffic data were available from the Pennsylvania Department of Transportation. Distances were calculated between mothers' residences and 1) TRI facilities and 2) major roads using ArcMap 10.1. Sociodemographic and exposure characteristics were compared for two subsets of the overall study population living within 3.2 km and 1.6 km of a

styrene-emitting TRI site. Logistic regression was conducted separately for exposure to styrene-emitting TRI facilities and major roadways as well as mutually adjusting for both exposures. Elevated ASD risk was associated with living within 3.2 km (2 miles) of a styrene TRI facility (adjusted OR=2.26, 95% CI=1.01-5.05) and within 300 meters of the nearest major road (adjusted OR=1.35, 95% CI=0.86-2.12). These associations remained stable in models adjusting for both.

4.2 INTRODUCTION

4.2.1 Autism Spectrum Disorder.

Over the last several decades, there has been a notable increase in the prevalence of autism spectrum disorder (ASD), a neurodevelopmental disorder characterized by impaired social interaction and communication and by restricted and repetitive behaviors [6; 105; 107; 108]. Although greater awareness and changes in diagnostic criteria have contributed to part of this increase, it has been estimated that as much as 46% is due to unknown factors [109]. ASD currently affects one in every 68 children [108].

A variety of genetic, sociodemographic, and environmental factors have been found to be associated with ASD. A child is at a greater risk for an ASD diagnosis if he or she has a sibling [111] or identical twin [110; 112; 183] with ASD. There are also a number of genetic mutations and disorders associated with ASD development, such as Angelman, Rett, and fragile X syndromes [110; 113; 114]. ASD risk is higher for boys than girls (4:1), although the effect does not appear to be X-linked [116]. Other factors related to increased ASD risk include advanced

maternal or paternal age [117], greater parental educational attainment [118], maternal smoking [122], adverse birth outcomes such as low birth weight and preterm birth [119; 120], and pregnancy complications [121].

Much of the work regarding ASD and environmental factors has focused on exposures to ambient air pollution due to industrial point, non-point, and mobile (e.g. traffic) sources [123; 124; 130-137; 140]. Several of these studies have estimated exposures to hazardous air pollutants (HAPs) using the USEPA's National Air Toxics Assessment (NATA) database [130-132]. NATA has modeled concentrations of HAPs by county and by census tract for the entire US approximately every three years from 1996 through 2005 [184]. These studies have noted associations between increased ASD risk and exposure to certain types of metals, chlorinated solvents, and aromatic solvents [130-132]. In a previous investigation using the study population described in this paper, we found a significant association between ASD and living in the highest quartile of styrene, an aromatic hydrocarbon (AOR=1.97, 95% CI=1.13-3.43, p=0.02). Styrene remained significant in multiple logistic regressions adjusting for air toxics significantly associated with ASD as well as mother's age, race, education, and smoking. Kalkbrenner also noted an elevated association between ASD and exposure to styrene as modeled by NATA (AOR=1.8, 95% CI=1.0-3.1) [130]. The properties, uses, and human health effects of styrene are described below.

4.2.2 Styrene.

Styrene (ethenylbenzene) is a volatile, colorless liquid with a characteristic sweet, sharp odor [141]. It is primarily used in the manufacture of plastics and resins (e.g. polystyrene, styrene-butadiene rubber, and unsaturated polyester resins) [141]. Occupational exposures occur most

notably in the reinforced plastics and boat building industries [141-143]. In occupational settings, styrene enters the body primarily through absorption in the lungs, although in its liquid phase a miniscule amount can be absorbed through the skin [141]. Workers exposed to styrene vapor have reported acute eye and throat irritation as well as symptoms associated with "styrene sickness," i.e. nausea, vomiting, weakness, dizziness, headache, and loss of appetite [142; 150; 151]. A well-established ototoxicant, styrene damages outer hair cells in the inner ear, which leads to eventual hearing loss [152]. Styrene has also been classified as "reasonably anticipated to be a human carcinogen" [153].

Experimental animal studies and human occupational studies suggest that styrene is neurotoxic. Styrene is lipophilic and therefore may be easily absorbed by the lipid-rich nervous system [154]. Animal studies have shown that styrene partially partitions in the brain following inhalation exposure [155; 156]. Styrene exposure may also elevate levels of specific markers of nervous system damage. When 32 male Sprague Dawley rats were exposed to 320 ppm styrene via inhalation chambers for three months, significantly elevated levels of glial fibrillary acidic protein (GFA) were observed, indicative of cell proliferation in response to central nervous system (CNS) damage [157].

Occupational studies in humans further support styrene's potential neurotoxicity. A cross sectional study of former workers of a polyester boat building plant found that exposed workers performed worse than control workers in several tests of neurobehavioral performance (e.g. symbol-digit substitution and digit span forwards) [142]. The authors concluded that less than 10 years of exposure to an average styrene level of 155 mg/m^3 may lead to persistent neurotoxic effects [142]. A meta-analysis of the human neurobehavioral effects of chronic styrene exposure revealed that cumulative styrene exposure was significantly associated with increased choice

reaction time (CRT) and increased color confusion index (CCI) [158]. Eight work-years of exposure to 20 ppm styrene, a contemporary limit for occupational exposure at the time of the study, resulted in a 6.5% increase in CRT and 2.2% increase in CCI [158]. Workers in the reinforced plastics industry are also at increased risk for mortality from diseases of the CNS [143].

The general population may be exposed to styrene from the following sources: 1) major point sources, such as styrene production plants and the industries mentioned previously [144], 2) exhaust from gasoline- or diesel-powered vehicles [145], 3) cigarette smoke [146], 4) off-gassing of residual styrene from household products (carpet glues, flooring materials, etc.) [147], and 5) food, due to either natural processes or the leaching of styrene from packaging [148; 149]. Due to these possible exposure sources and the above epidemiological and toxicological evidence to date, the following investigation of ASD and exposure to major sources of styrene was conducted. The objective was to explore associations between ASD risk and exposure to multiple sources of environmental styrene (industrial and traffic), using proximity to source as a surrogate for exposure. The hypothesis was that increased ASD risk would be associated with closer proximity to industrial facilities and roadways.

4.3 METHODS

4.3.1 Study population.

Case ascertainment and the study population have been described elsewhere (Talbott et al., in review) and will be summarized here. The study was approved by the University of Pittsburgh

Institutional Review Board (IRB number PRO10010240). ASD cases were children born between January 1, 2005 and December 31, 2009 in a six-county area (Allegheny, Armstrong, Beaver, Butler, Washington, and Westmoreland) of southwestern Pennsylvania and were recruited from specialty autism clinics, treatment centers, or through the Pennsylvania School System. A case was required to have a documented diagnosis of ASD and a score of 15 or above on the Social Communication Questionnaire (SCQ), a positive screen for the presence of autistic features. Controls were recruited from a random selection of births from the same six-county area and time period (2005-2009) using the Pennsylvania Department of Health (PA DOH) state birth registry. Children with an SCQ score of 15 or above were excluded as controls. Mothers of cases and controls were administered a personal interview by a trained interviewer using a structured questionnaire adapted from the CDC's Study to Explore Early Development. A complete residential history as well as demographic and behavioral characteristics of the parents, including age, race, education, smoking habits, and occupation, were obtained from the interview.

This study focused on two major sources of potential styrene exposure: industrial facilities and traffic. The United States Environmental Protection Agency's (US EPA) Toxics Release Inventory (TRI) provided information on major styrene-emitting facilities. Road and traffic data were obtained from the Pennsylvania Department of Transportation (PADOT). These datasets are described in the next two sections.

4.3.2 TRI facilities.

Geographic locations (latitude and longitude) and other facility information for industrial point sources of styrene were obtained from TRI facility reports [185]. The TRI Program compiles

information regarding the management of toxic chemicals that may pose a threat to human health and the environment. The program requires US facilities in certain industry sectors to report amounts of chemicals released to the environment. Such releases include emissions of a particular chemical to the air, water, or land as well as quantities of waste recycled, treated, or otherwise disposed. The TRI was established in 1986 as part of the Emergency Planning and Community Right-to-Know Act (EPCRA) to make information about industrial management of toxic chemicals available to the general public. Currently, there are over 650 chemicals covered by the program. Data is available for the years 1988 through the most recent assessment, 2013 [186].

For the present study, facility reports were downloaded for all styrene-reporting facilities in Pennsylvania for the years 2005 through 2009. The geographic coordinates and releases of styrene in pounds to air, water, and land were available for each facility. Although there were several facilities in counties neighboring the study area (in Clarion, Lawrence, and Marion), these were located relatively far (>16 km) from any of the birth residences in our study population and were therefore excluded from the analysis. Facilities within the six county area with styrene emissions totaling 50 or fewer pounds were also excluded. Point distances were calculated between each mother's residence (at her child's birth) and all styrene point sources within the six-county area during the birth year using ArcMap (version 10.1; ESRI Inc., Redlands, CA). Two "exposure" buffers were considered: living <3.2 km (2 miles) and <1.6 km (1 mile) from the nearest styrene-emitting facility.

4.3.3 Traffic.

The shapefile of Pennsylvania roadways was downloaded from the Pennsylvania Spatial Data Acccess (PASDA) [187] for 2008, the earliest year on the study period (2005-2008) for which data was available. The dataset includes locations of roadways (lines or "polylines" in ArcMap 10.1), type of route (PA, US, Interstate, etc.), street names, and estimates of traffic volume (annual average daily traffic, or AADT, in cars per day). About 6,300 raw traffic counts are collected by PADOT each year and used to develop AADT estimates by applying traffic expansion factors to traffic counts [188]. The distance between each residence and the nearest major road (PA, US, or I highway) within the six-county study area was calculated in ArcMap 10.1, which calculates the shortest distance (usually the perpendicular) from a point feature to the closest line segment of a polyline. Two exposure thresholds were examined: living <1000 m and <300 m from the nearest major road [136].

4.3.4 Statistical analysis.

The number of cases and controls living <3.2 km and <1.6 km of at least one styrene-emitting facility was noted. Simple t-tests and chi-square tests were conducted to compare the demographics (mother's age, education, race, and smoking) and exposure characteristics (total styrene emissions from the nearest TRI, distance to the nearest major road, AADT of the nearest major road, and census tract-level styrene estimate) of the cases and controls living within each buffer region. The occupations of both parents and the potential for additional chemical exposures through work were assessed for each subset.

Logistic regression was conducted to examine associations between ASD and proximity to styrene TRI facilities and traffic separately and in a final model mutually adjusting for both exposures. For proximity to TRI, participants living <3.2 km from the nearest facility were compared to those living ≥3.2 km. For proximity to traffic, participants living <1000 m versus ≥1000 m and those living <300 m versus ≥300 m from the nearest major road were compared. The final model mutually adjusted for proximity to TRI as categorized above and the highest exposure to traffic (<300 m versus ≥300 m). All models were further adjusted for mother's age, education, race, and smoking. A series of continuous metrics—distance to the nearest styrene facility, total styrene emissions (in pounds) of the nearest facility, distance to the nearest major road, and AADT of the nearest major road—were also examined as potential surrogates for exposure. Associations were considered significant at $p<0.05$, although borderline effects at $p<0.20$ were also considered.

4.4 RESULTS

4.4.1 Descriptives of TRI sites.

Table 7 shows descriptive statistics for the TRI facilities in the six county area that reported styrene emissions over the study period (2005 to 2009). The number of facilities reporting more than 50 pounds of total styrene emissions remained relatively constant at 15 or 16 sites from 2005 to 2008, except for a minimum of 13 sites in 2009. There was considerable variation in the amount of total styrene emissions within each year, with some sites emitting less than 100

pounds and one site emitting almost 75,000 pounds in 2008. There was little difference between the mean pounds of total styrene emissions across the study years (one-way ANOVA p>>0.10).

Table 7. Descriptive statistics for all TRI styrene-emitting facilities across the study period (2005 to 2009)

Year	Number of sites reporting any styrene	Number of sites reporting >50 pounds	Styrene descriptives (pounds)*			
			Mean	S.D.	Min	Max
2005	17	16	8,659	8,695	260	25,377
2006	16	15	8,539	10,227	76	38,519
2007	18	16	8,114	9,601	78	33,121
2008	19	15	9,840	18,674	60	74,500
2009	17	13	8,504	12,656	183	47,180

*Based on sites reporting >50 pounds of total styrene emissions

4.4.2 Demographics of study population.

Table 8 shows the demographic composition of the total study population (217 cases and 226 controls) as well as the subsets living within 3.2 km or 1.6 km of a styrene facility. For the total sample, mothers of controls were slightly older compared to mothers of cases (32 versus 30 years) and had a greater percentage that attended college (79% v. 55%). Conversely, there were more non-white mothers of cases than of controls (11% v. 3%) and more case mothers who reported smoking during pregnancy (25% v. 11%). About the same proportion of case and control children were male (78% and 77%, respectively).

Table 8. Demographic composition of the total study population and for subsets living within 3.2 km and 1.6 km of the closest TRI facility during the birth year (2005-2009)

Characteristic	Total		Within 3.2 km		Within 1.6 km	
	Cases (n=217)	Controls (n=226)	Cases (n=23)	Controls (n=10)	Cases (n=4)	Control (n=1)
Mother's age (mean ± SD)	30.4±5.4*	31.8 ± 4.7*	29.1± 4.9	31.1 ± 5.3	25.8 ± 4.3	40.0
Mother's education (% college)	55%*	79%*	44%*	90%*	50%	100%
Mother's race (% non-white)	11%*	3%*	13%	0%	0%	0%
Smoking during Pregnancy (% yes)	25%*	11%*	48%*	0%*	75%	0%
Child's sex (% male)	78%	77%	91%	90%	75%	100%

*t-test or chi-square test significant at $p<0.05$

There were 33 children whose mothers lived within 3.2 miles of the closest styrene TRI facility during the birth year (23 cases and 10 controls). Mothers of controls were again slightly older than those of cases (31 v. 29 years of age) and had higher educational attainment (90% v. 44% with a college education). Thirteen percent of the case mothers were nonwhite and 48% smoked during pregnancy, while none of the control mothers were nonwhite or smoked. There was little difference between the percentage of male children between the cases and controls (91% and 90%, respectively) in this subset, although these proportions were higher than those in the total sample. A very small subset of the total sample (4 cases and 1 control) lived within 1.6 km of a styrene facility. Case mothers within this group had a mean age of about 26 and were somewhat younger than mothers in the 3.2 km and overall samples, and most of these mothers (75%) smoked.

4.4.3 Occupational exposures.

The proportions of case and control parents in each group (total sample, 3.2 km, and 1.6 km) who worked outside of the home and who reported occupational exposures to chemicals and other substances are displayed in Table 9 (next page). Percentages are reported for the time period from 3 months prior to pregnancy until the child's second birthday. In the total study population, 74% of the 217 case mothers and 79% of the 226 control mothers worked at some point during the period 3 months prior to pregnancy until the child's second birthday. Among those that worked, 8% of case mothers and 7% of control mothers reported exposure to any of the chemicals/substances listed on the interview, such as alcohols, diesel fumes, oil based paints, and pesticides, among others (see Table 9 footnote). None of the mothers reported occupational exposure to styrene or any other specific aromatic hydrocarbon, except for xylenes (1 case, <1%). However, a few control mothers (<2%) reported exposures to adhesives and diesel fumes; styrene is used in the manufacture of some adhesives [141; 144] and is one of the many toxics in diesel exhaust [145].

Table 9. Occupational exposures of the total study population and for subsets living within 3.2 km and 1.6 km of the closest TRI facility during the birth year (2005-2009)

Worked?		Total		Within 3.2 km		Within 1.6 km	
		Cases (n=217)	Controls (n=226)	Cases (n=23)	Controls (n=10)	Cases (n=4)	Control (n=1)
Mother	*Yes*	74%	79%	65%	90%	50%	100%
	with chemicals[*]	8%[**]	7%[**]	20%	11%	25%	0%
Father[§]	*Yes*	93%	98%	87%	100%	50%	100%
	with chemicals[*]	36%	22%	20%	10%	0%	0%

[*] Percentage calculated out of those that worked

[**] Includes occupational exposures to the following chemicals and substances: adhesives, alcohols, anesthetic gases, automotive fluids, diesel fumes, oil based paints/paint strippers/thinners, perchloroetheylene, pesticides/herbicides/insecticides, pharmaceuticals/drugs, x-ray/radioactive, xylenes, other solvents/degreasers, hair dyes, or others

[§] Worked missing 1 case and 2 controls; worked with chemicals missing 21 cases and 11 controls

The majority of case and control fathers worked (93% and 98%, respectively), and, out of these, 36% of case fathers and 22% of control fathers reported occupational exposures to chemicals and other substances during the three month period prior to pregnancy. Sixteen fathers had gasoline- or diesel-associated exposures, while a few others reported working in industries or with substances in which styrene exposure may have occurred, e.g. tire production, flooring, plastics manufacturing, and adhesives [141; 144].

Among the subset of mothers living within 3.2 km of a styrene facility, 20% of case mothers and 11% of control mothers who worked reported occupational exposures to chemicals, including hospital-based exposures (medicines, IVs, xrays), cleaning products, calcium chloride, and citric acid. Among fathers in this subset who worked, 20% of case fathers and 10% of control fathers reported occupational exposure to chemicals. As is evident in Table 9, most of the working parents living the closest to styrene-emitting facilities (<1.6 km) did not work with chemicals at their jobs, except for one case mother.

4.4.4 Environmental exposures.

As the two primary styrene sources of interest in this analysis are industrial sites and traffic, the following proxies for exposure were considered: distance to the nearest styrene facility, total pounds of styrene emitted from the nearest facility, distance to the nearest major road (NMR), and average annual daily traffic (AADT) of the NMR. Table 10 compares these metrics between ASD cases and controls within the total study population and for those living in close proximity to styrene TRI facilities. Table 10 also shows the mean styrene level for each group as estimated by the National Air Toxics Assessment Program (NATA).

Table 10. Environmental exposures of the total study population and for subsets living within 3.2 km and 1.6 km of the closest TRI facility during the birth year (2005-2009)

Exposure metric	Total		Within 3.2 km		Within 1.6 km	
	Cases (n=217)	Controls (n=226)	Cases (n=23)	Controls (n=10)	Cases (n=4)	Control (n=1)
Mean distance to nearest TRI (km)	12.7 ± 8.7	12.6 ± 7.7	2.4 ± 0.60	2.1 ± 0.47	1.4 ± 0.23	0.37
Mean total styrene emissions from nearest TRI (pounds)	8,495 ± 6,651	9,567 ± 8,438	8,760 ± 5,922	8,939 ± 7,101	12,506 ± 7,935	6,967
Mean distance to NMR (m)	756 ± 663*	944 ± 940*	569 ± 409	852 ± 994	327 ± 149	3,490
Mean AADT of NMR (cars/day)	12,403 ± 7,505	12,352 ± 7,697	10,615 ± 8,153	9,731 ± 2,789	6,109 ± 1,823	5,224
Mean styrene level (ng/m^3)	38.8±20.4	37.3 ± 40.4	58.8±31.1	65.8± 37.3	70.3 ±26.2	26.8

Comparing cases and controls within each group, there were no striking differences between mean distances to the nearest TRI site nor the mean total styrene emissions from the nearest site. Mean total styrene emissions were highest for the four ASD cases living within 1.6 km of a facility (12,506 ± 7,935 pounds). The NMR tended to be closer for cases than controls; however, mean AADT of the nearest major road was not notably different between cases and controls in each group. Overall, 71% and 29% of cases lived within 1000-m and 300-m of the NMR, respectively. Lastly, mean styrene levels (ng/m^3) as estimated at the census tract level by NATA tended to increase as proximity to TRI sources of styrene increased, as was expected since NATA includes TRI emissions in its modeled air pollution estimates.

The Spearman correlation matrix including all of the exposure metrics in Table 10 (distance to nearest TRI, pounds of styrene emitted from nearest TRI, distance to NMR, AADT of NMR, and NATA-estimated styrene level) is presented in Table 11.

Table 11. Spearman correlation matrix for exposure metrics

	Nearest TRI	Total emissions	NATA styrene	NMR	AADT
Nearest TRI	1.000	-0.084	-.255**	0.031	-.251**
Total emissions		1.000	-0.050	0.046	.137**
NATA styrene			1.000	-.320**	.162**
NMR				1.000	0.039
AADT					1.000

**Correlation is significant at the 0.01 level (2-tailed)

Distance to the nearest TRI was not significantly correlated with total styrene emissions from the nearest TRI; however, it was negatively correlated with NATA styrene concentrations (i.e. concentrations decreased as distance increased). NATA-estimated styrene concentrations were negatively correlated with distance to NMR and positively correlated with AADT; these correlations would be expected since traffic is another component of NATA's ambient air pollution models [184]. Distance to NMR was not correlated with distance to the nearest TRI site. AADT was negatively correlated with distance to the nearest TRI site and positively correlated with styrene emissions in pounds. This suggests that areas in close proximity to major industrial sources of styrene are also near highly traveled major roadways.

4.4.5 Associations with ASD.

Figure 12 presents the odds ratios (OR) and 95% confidence intervals (CI) for ASD and proximity to styrene TRI facility (<3.2 km versus ≥3.2 km), unadjusted and adjusted for mother's age, race, education, and smoking. Odds of ASD were significantly associated with proximity to TRI in both the unadjusted (OR=2.56, 95% CI=1.19-5.52, p=0.016) and adjusted (OR=2.26, 95% CI=1.01-5.05, p=0.048) logistic regression models. Distance to the nearest

styrene facility as a continuous measure was not significantly associated with odds of ASD, nor was pounds of total styrene emitted from the nearest facility (data not shown).

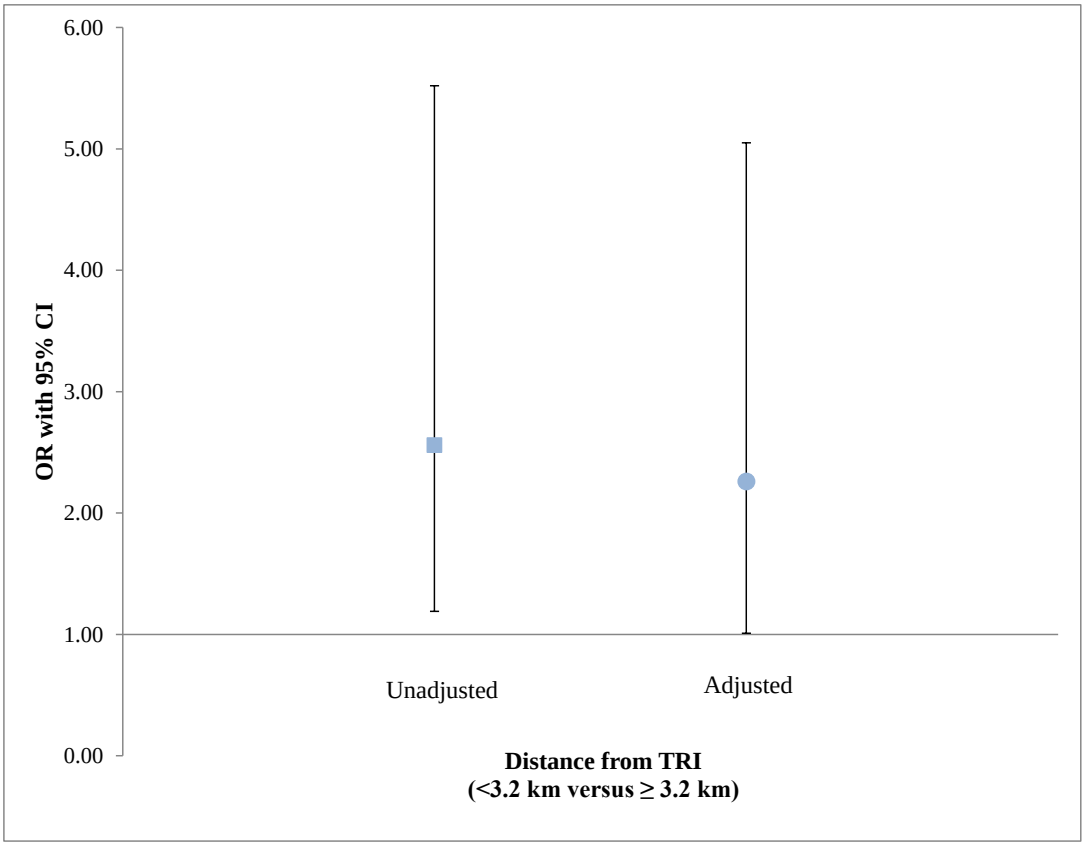


Figure 12. OR and 95% CI for proximity to styrene TRI sites (1 if <3.2 km, 0 if ≥3.2 km)

The logistic regression results for ASD and proximity to road are shown in Figure 13. Results are presented for both exposure buffers (<1000 m versus ≥1000 m and <300 m versus ≥300 m), unadjusted and adjusted for maternal risk factors. In the unadjusted model, ASD was positively associated with living <1000 m from a major road, although the relationship was not statistically significant (OR=1.22, 95% CI=0.82-1.83, p=0.342). Including mother's age, education, and race in the model did not substantially alter the relationship (OR=1.28, 95% CI=0.84-1.97, p=0.255). The association between ASD and proximity was somewhat stronger using the 300-m buffer and remained similar in magnitude after adjustment: unadjusted

OR=1.41, 95% CI=0.92-2.17, p=0.119; adjusted OR=1.35, 95% CI=0.86-2.12, p=0.187. Distance to the nearest major road as a continuous measure was significantly associated with odds of ASD (p<0.05), although AADT of the nearest major road was not (data not shown).

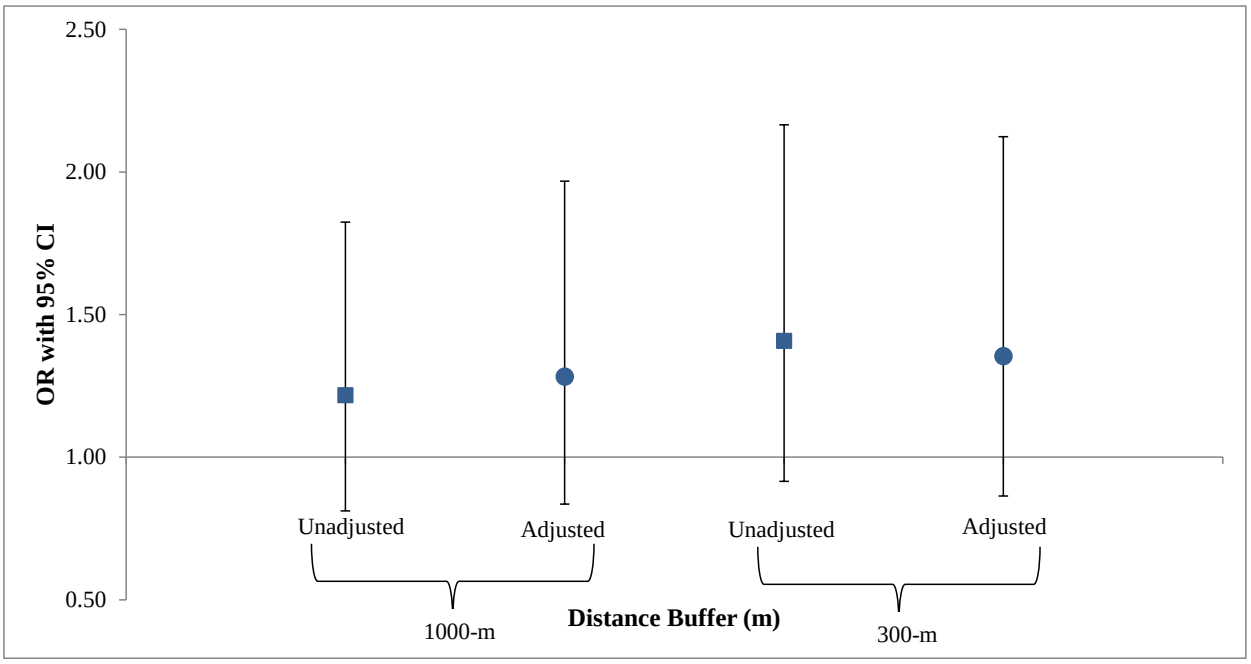


Figure 13. OR and 95% CI for proximity to NMR using two buffers, 1000-m and 300-m

Since the association of ASD and living within 300-m of the nearest major road appeared to be the stronger metric out of the two buffers (p<0.20), it was included in a final model mutually adjusting for proximity to TRI and road (Figure 14). Proximity to TRI (OR=2.58, 95% CI=1.20-5.58, p=0.016) and road (OR=1.42, 95% CI=0.92-2.19, p=0.113) were both positively associated with ASD before the inclusion of other maternal risk factors. Adding mother's age, education, etc. to the model attenuated the odds ratios somewhat, although they remained similar in magnitude and direction (TRI: OR=2.26, 95% CI=1.01-5.07, p=0.047; road: OR=1.36, 95% CI=0.86-2.13, p=0.186). Further, the associations between ASD and proximity to TRI and road

in models adjusted for both exposures were comparable to those observed in the individual models.

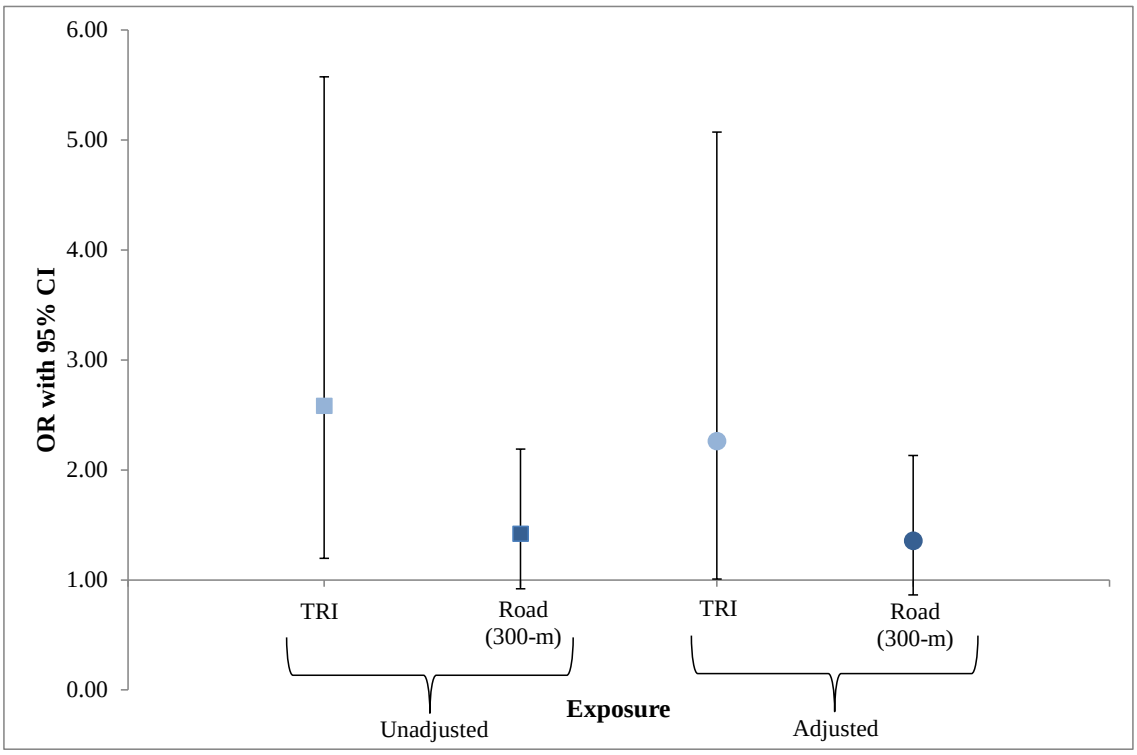


Figure 14. OR and 95% CI for models adjusted for proximity to TRI (<3.2 km) and NMR (<300-m)

4.5 DISCUSSION

4.5.1 Summary of findings and comparison with existing studies.

There appears to be an association of increased risk for ASD and proximity to industrial point sources of styrene (adjusted OR=2.26, 95% CI=1.01-5.05, p=0.048). This relationship was robust to inclusion of maternal risk factors (mother's age, education, and race) as well as other

sources of styrene (maternal smoking, proximity to major road). However, the association was only resolved when a simple dichotomous exposure metric was used (<3.2 km versus ≥3.2 km). There were 33 study participants (23 cases and 10 controls) who lived <3.2 km from a styrene-emitting TRI facility. There were several noteworthy differences between these 23 cases and 10 controls; namely, the cases tended to have less educational attainment (44% v. 90% had a college education) and included more non-white mothers (13% v. 0%) and smokers (48% v. 0%). The small size of the sample (4 cases and only 1 control) living <1.6 km precluded further analysis on this subset. Although pounds of emissions has proved a useful surrogate in studies of autism rates and exposure to mercury TRI facilities [123; 124], pounds of total styrene emitted from the nearest TRI site was not associated with increased ASD risk in our study.

An elevated association was also found between ASD risk and proximity to road. This relationship was strongest using the <300-m threshold (OR=1.35, 95% CI=0.86-2.12, p=0.187) and was not substantially influenced by including other maternal risk factors or proximity to TRI in the model. These results agree with previous studies that have found associations between ASD risk and proximity to road or traffic-related exposures [133-137; 140]. Volk et al. noted an association between ASD and residential proximity to a freeway (<309 m) during the third trimester (AOR=1.96, 95% CI=1.01-3.93) [136]. The results of the exploratory factor analysis in the previous chapter also revealed an elevated association between traffic-related pollutants and ASD.

4.5.2 Strengths, limitations, and conclusions.

The primary strength of this study was the incorporation of two major sources of environmental styrene. ASD risk was associated with proximity to TRI and proximity to roads individually.

These relationships remained in models mutually adjusting for both sources. Personal interviews conducted with cases and controls provided information on personal risk factors, which were considered in all models. The responses regarding occupation and associated chemical exposures revealed that confounding by occupational exposure to styrene was likely negligible since very few parents reported working in relevant industries. Logistic regressions were also adjusted for maternal smoking, another significant source of styrene exposure [146].

Despite these strengths, our study has several limitations characteristic of this type of investigation. The major limitation is that proximity to source was used as a surrogate for exposure. Further investigation is required to determine if exposure to styrene, particularly from industrial emissions, is the true etiologic agent responsible for the observed associations. Since styrene is just one of the many pollutants found in gasoline and diesel exhaust [133-137; 140; 145], identifying the agent (or agents) responsible for the individual association observed for proximity to road and ASD is an even more complex task. Alternatively, since disadvantaged groups have been found to live disproportionately close to pollution sites [189], these associations could instead be due to an unaccounted for factor related to low socioeconomic status. As seen in Table 8, there were some sociodemographic disparities between the 23 cases and 10 controls living <3.2 km from a styrene facility.

Although the current study was able to account for three important sources of styrene exposure (industrial, traffic, and smoking), future studies should probe other exposure sources in the home [147] as well as dietary intake [148; 149]. Mandelic acid and phenylglyoxylic acid, the two main urinary metabolites of styrene in humans, have been used as biomarkers of exposure in occupational studies [141] and would also be useful indicators of exposure in future work. Next steps should therefore incorporate a more refined exposure assessment, including biomonitoring,

as well as a prospective study design to elucidate styrene's role, if any, in the development of ASD.

5.0 CONCLUSIONS AND PUBLIC HEALTH SIGNIFICANCE

The public health significance of this work was to identify possible sources of harmful exposures that may motivate further research, primary prevention efforts, and eventually policies to further limit these exposures in infants and children. To achieve this, three specific aims were created:

- **Specific Aim 1:** To assess the impact of unconventional natural gas development on infant health in southwestern Pennsylvania using well density as a surrogate for exposure.

  Hypothesis: The risk for adverse birth outcomes will be greater for those infants born to mothers living in more densely drilled areas.

  Conclusions: Infants born to mothers living in the most densely drilled areas (i.e., fourth quartile) had lower birth weights and were at increased risk for SGA but not preterm birth.

- **Specific Aim 2a:** To explore associations between ASD risk and groups of air toxics using exploratory factor analysis.

  Conclusions: Factor analysis is a dimension reduction technique that may be utilized to study the contribution of multiple pollutants to ASD risk. In the current analysis, the two methods used to calculate factor scores produced fairly consistent results. Possible sources of air toxics that may be worthy of further study include traffic, combustion sources, and plastics, rubber, and adhesives manufacturing.

- **Specific Aim 2b:** To explore associations between ASD risk and proximity to major sources of environmental styrene exposure, i.e. industrial and traffic.

  Hypothesis: Increased ASD risk will be associated with living near major industrial sources of styrene and near major roadways, after taking into account maternal risk factors for ASD.

  Conclusions: ASD risk was elevated for those living close to styrene-emitting TRI facilities (<3.2 km) and to major roads (<300 m). These exposures remained elevated in logistic regression models mutually adjusting for both sources. Although personal risk factors were taken into account, results may be confounded by other variables related to living in disadvantaged areas close to major pollution sources.

Taken together, the results of these studies indicate that environmental exposures, such as pollutants associated with UGD, industrial air toxics, and traffic, may play an important role in the health of infants and children, even after accounting for other potential confounders. Since major windows of developmental vulnerability exist in utero and during infancy and early childhood [3], and early life exposures can influence health later in childhood and even adulthood [2; 4], further investigation of these issues appears warranted. As proximity and aggregation methods were used to represent exposure, future work will benefit from improved individual exposure assessment and a prospective study design to confirm and further elucidate these associations.

APPENDIX: EXPLORATORY FACTOR ANALYSIS SUPPLEMENTAL TABLES

Table 12. Spearman correlation matrix for 30 NATA air toxics (pregnancy average, ng/m^3)

	Arsenic	Cadmium	Chromium	Mercury	Manganese	Nickel	Lead	Selenium	Benzene
Arsenic	1.000								
Cadmium	.717**	1.000							
Chromium	.454**	.443**	1.000						
Mercury	.235**	.644**	.223**	1.000					
Manganese	.494**	.394**	.395**	.156**	1.000				
Nickel	.494**	.438**	.492**	.110*	.604**	1.000			
Lead	.657**	.501**	.491**	0.069	.552**	.583**	1.000		
Selenium	.188**	.532**	0.042	.656**	0.049	.110*	0.046	1.000	
Benzene	.748**	.576**	.377**	.154**	.470**	.511**	.555**	.216**	1.000
Ethyl benzene	.768**	.537**	.402**	0.091	.460**	.477**	.567**	0.090	.903**
Styrene	.619**	.494**	.550**	.149**	.230**	.234**	.349**	0.053	.565**
Toluene	.754**	.580**	.360**	.159**	.476**	.506**	.558**	.189**	.991**
Xylenes	.810**	.592**	.434**	.126**	.463**	.482**	.586**	0.092	.942**
Methylene chloride	.686**	.502**	.358**	.096*	.415**	.392**	.471**	0.061	.679**
PERC	.625**	.463**	.210**	.115*	.416**	.410**	.454**	.227**	.870**
Trichloroethylene	.770**	.645**	.551**	.174**	.343**	.462**	.513**	0.065	.713**
Vinyl chloride	.314**	.433**	.104*	.301**	.419**	.209**	.182**	.393**	.461**
Hydrazine	.305**	.188**	-0.043	0.066	.410**	.328**	.211**	.340**	.295**
PAHs	.624**	.394**	.379**	-.101*	.262**	.361**	.526**	.131**	.768**
Diesel PM	.715**	.516**	.388**	0.083	.401**	.499**	.549**	.168**	.929**
Allyl chloride	0.068	-.167**	0.087	-.221**	-0.001	-.303**	-.152**	-.372**	-0.073
Carbon disulfide	.408**	-0.016	.210**	-.480**	.104*	.112*	.290**	-.154**	.491**
Cresol	.552**	.408**	.269**	0.090	.287**	.257**	.233**	.280**	.770**
Cumene	.633**	.423**	.413**	0.041	.269**	.270**	.371**	.151**	.694**
Cyanide	.761**	.534**	.498**	0.018	.319**	.406**	.531**	0.065	.778**
Dinitrotoluene	0.070	-.180**	0.086	-.240**	-0.007	-.302**	-.149**	-.385**	-0.062
Ethylene oxide	.653**	.475**	.243**	.135**	.447**	.410**	.483**	.186**	.794**
Hexane	.768**	.597**	.407**	.211**	.458**	.466**	.568**	.167**	.934**
Trichloroethane	.809**	.658**	.360**	.241**	.500**	.483**	.553**	.261**	.889**
Methanol	.831**	.642**	.478**	.117*	.415**	.468**	.558**	.118*	.823**

Table 12. (continued)

	Ethyl benzene	Styrene	Toluene	Xylenes	Methylene chloride	PERC	Trichloro-ethylene	Vinyl chloride	Hydrazine	PAHs
Ethyl benzene	1.000									
Styrene	.669**	1.000								
Toluene	.918**	.564**	1.000							
Xylenes	.950**	.709**	.959**	1.000						
Methylene chloride	.667**	.585**	.698**	.752**	1.000					
PERC	.760**	.319**	.872**	.778**	.568**	1.000				
Trichloroethylene	.700**	.656**	.716**	.773**	.705**	.572**	1.000			
Vinyl chloride	.400**	.274**	.475**	.430**	.300**	.452**	.272**	1.000		
Hydrazine	.303**	-0.085	.275**	.185**	.113*	.319**	-0.007	.230**	1.000	
PAHs	.711**	.596**	.738**	.749**	.627**	.662**	.660**	.326**	.124**	1.000
Diesel PM	.901**	.618**	.931**	.928**	.651**	.778**	.671**	.414**	.239**	.744**
Allyl chloride	0.036	.333**	-0.047	0.092	.244**	-.110*	.117*	-.111*	-.248**	-0.035
Carbon disulfide	.488**	.425**	.479**	.529**	.515**	.449**	.421**	.126**	.098*	.725**
Cresol	.645**	.493**	.764**	.729**	.553**	.733**	.563**	.522**	.136**	.666**
Cumene	.706**	.712**	.680**	.720**	.594**	.575**	.669**	.376**	.165**	.793**
Cyanide	.760**	.743**	.767**	.829**	.774**	.641**	.819**	.260**	0.032	.894**
Dinitrotoluene	0.041	.330**	-0.037	.099*	.246**	-.099*	.123**	-.119*	-.252**	-0.021
Ethylene oxide	.680**	.373**	.798**	.752**	.629**	.741**	.627**	.354**	.199**	.623**
Hexane	.913**	.662**	.948**	.956**	.690**	.769**	.711**	.474**	.212**	.712**
Trichloroethane	.853**	.565**	.906**	.897**	.744**	.786**	.743**	.549**	.297**	.708**
Methanol	.837**	.734**	.827**	.884**	.812**	.683**	.863**	.339**	.154**	.818**

Table 12. (continued)

	Diesel PM	Allyl chloride	Carbon disulfide	Cresol	Cumene	Cyanide	Dinitro-toluene	Ethylene oxide	Hexane	Trichloro-ethane	Methanol
Diesel PM	1.000										
Allyl chloride	-0.019	1.000									
Carbon disulfide	.503**	.352**	1.000								
Cresol	.740**	.154**	.540**	1.000							
Cumene	.702**	.144**	.630**	.654**	1.000						
Cyanide	.768**	.117*	.643**	.662**	.805**	1.000					
Dinitrotoluene	-0.010	.999**	.369**	.162**	.152**	.125**	1.000				
Ethylene oxide	.722**	0.070	.467**	.740**	.514**	.655**	0.078	1.000			
Hexane	.910**	0.032	.430**	.702**	.717**	.765**	0.038	.727**	1.000		
Trichloroethane	.841**	0.005	.448**	.743**	.684**	.759**	0.010	.762**	.871**	1.000	
Methanol	.804**	.114*	.551**	.654**	.786**	.938**	.120*	.666**	.828**	.846**	1.000

**Correlation is significant at the 0.01 level (2-tailed)

*Correlation is significant at the 0.05 level (2-tailed)

Table 13. Factor score coefficient matrix (Method 2)

	Factor								
Pollutant	1	2	3	4	5	6	7	8	9
Arsenic	0.057	-0.045	0.007	0.170	0.130	-0.050	-0.083	0.112	-0.096
Cadmium	0.006	-0.007	0.013	0.362	0.039	-0.033	-0.035	0.054	-0.070
Chromium	-0.059	0.204	-0.006	-0.028	0.210	0.008	0.177	-0.014	-0.021
Mercury	-0.022	0.015	0.016	0.409	-0.063	-0.034	-0.002	-0.148	0.012
Manganese	0.041	0.379	0.008	0.033	-0.205	0.003	-0.073	-0.026	0.003
Nickel	-0.067	0.179	-0.011	-0.057	0.310	0.016	0.032	0.050	0.073
Lead	0.043	0.375	0.007	0.049	-0.183	-0.002	-0.121	-0.007	-0.029
Selenium	-0.052	0.055	0.006	0.357	-0.142	0.116	0.047	0.089	0.154
Benzene	0.132	0.031	-0.022	-0.044	-0.086	-0.017	-0.017	-0.033	0.013
Ethyl benzene	0.076	-0.029	-0.012	-0.060	-0.010	-0.118	0.155	0.269	-0.082
Styrene	-0.008	-0.016	-0.004	0.018	-0.025	-0.038	0.583	-0.081	-0.163
Toluene	0.140	0.023	-0.019	-0.043	-0.066	-0.069	-0.031	-0.038	0.015
Xylenes	0.129	-0.009	0.013	-0.034	-0.005	-0.087	0.008	-0.056	-0.057
Methylene chloride	0.036	0.034	0.097	0.052	-0.052	-0.008	0.131	0.124	0.106
PERC	0.123	0.041	-0.003	-0.049	-0.075	-0.032	-0.192	-0.100	0.157
Trichloroethylene	-0.039	-0.105	-0.014	-0.055	0.555	-0.006	-0.066	0.012	0.110
Vinyl chloride	-0.024	-0.014	-0.006	0.036	0.079	-0.019	-0.041	0.017	0.808
Hydrazine	-0.038	0.001	-0.010	0.004	0.025	0.020	-0.018	0.814	0.019
PAHs	-0.078	0.008	-0.028	0.016	-0.016	0.554	-0.010	0.044	-0.047
Diesel PM	0.105	0.054	-0.037	-0.026	-0.133	0.013	0.024	-0.074	-0.071
Allyl chloride	-0.010	0.004	0.364	0.028	0.001	-0.080	-0.036	-0.047	-0.007
Carbon disulfide	-0.026	-0.004	0.294	-0.027	-0.024	0.130	0.014	0.067	-0.018
Cresol	0.031	0.034	0.002	0.011	-0.049	0.303	-0.066	-0.115	0.208
Cumene	0.011	-0.048	-0.014	-0.009	-0.028	-0.016	0.366	-0.008	0.162
Cyanide	0.026	-0.030	-0.019	0.003	0.055	0.311	-0.005	-0.001	-0.120
Dinitrotoluene	-0.010	0.004	0.364	0.026	0.001	-0.080	-0.036	-0.049	-0.007
Ethylene oxide	0.081	-0.027	0.035	0.006	0.190	-0.030	-0.413	-0.186	-0.088
Hexane	0.143	0.030	-0.032	-0.037	-0.110	-0.109	0.090	-0.098	-0.038
Trichloroethane	0.101	-0.020	-0.006	0.022	0.034	-0.045	-0.036	0.036	0.065
Methanol	0.059	-0.055	-0.001	0.022	0.113	0.101	-0.006	0.023	-0.099

BIBLIOGRAPHY

1. Auchincloss, A. H., Gebreab, S. Y., Mair, C., & Roux, A. V. D. (2012). A Review of Spatial Methods in Epidemiology, 2000-2010. *Annual Review of Public Health, 33*, 107-122.

2. Landrigan, P. J., & Goldman, L. R. (2011). Children's vulnerability to toxic chemicals: a challenge and opportunity to strengthen health and environmental policy. *Health Aff (Millwood), 30*(5), 842-850. doi: 10.1377/hlthaff.2011.0151

3. Sever, L. E. (1995). Epidemiologic Aspects of Environmental Hazards to Reproduction. In E. O. Talbott & G. F. Craun (Eds.), *Introduction to Environmental Epidemiology* (pp. 81-98). Boca Raton: CRC Press, Inc.

4. Barker, D. J. (2007). The origins of the developmental origins theory. *J Intern Med, 261*(5), 412-417. doi: 10.1111/j.1365-2796.2007.01809.x

5. World Health Organization. (2014). Preterm Birth. Retrieved February 7, 2015, from http://www.who.int/mediacentre/factsheets/fs363/en/

6. Centers for Disease Control and Prevention. (2014). Reproductive Health: Preterm Birth. Retrieved February 7, 2015, from http://www.cdc.gov/reproductivehealth/maternalinfanthealth/pretermbirth.htm

7. Goldenberg, R. L., Culhane, J. F., Iams, J. D., & Romero, R. (2008). Epidemiology and causes of preterm birth. *Lancet, 371*(9606), 75-84. doi: 10.1016/s0140-6736(08)60074-4

8. Martin, J. A., Hamilton, B. E., Osterman, M. J. K., Curtin, S. C., & Mathews, T. J. (2015). Births: final data for 2013. *National vital statistics reports, 64*(1). http://www.cdc.gov/nchs/data/nvsr/nvsr64/nvsr64_01.pdf

9. Goldenberg, R. L., Goepfert, A. R., & Ramsey, P. S. (2005). Biochemical markers for the prediction of preterm birth. *Am J Obstet Gynecol, 192*(5 Suppl), S36-46. doi: 10.1016/j.ajog.2005.02.015

10. Engel, S. A., Erichsen, H. C., Savitz, D. A., Thorp, J., Chanock, S. J., & Olshan, A. F. (2005). Risk of spontaneous preterm birth is associated with common proinflammatory cytokine polymorphisms. *Epidemiology, 16*(4), 469-477.

11. Macones, G. A., Parry, S., Elkousy, M., Clothier, B., Ural, S. H., & Strauss, J. F., 3rd. (2004). A polymorphism in the promoter region of TNF and bacterial vaginosis: preliminary evidence of gene-environment interaction in the etiology of spontaneous preterm birth. *Am J Obstet Gynecol, 190*(6), 1504-1508; discussion 1503A. doi: 10.1016/j.ajog.2004.01.001

12. Goldenberg, R. L., Hauth, J. C., & Andrews, W. W. (2000). Intrauterine infection and preterm delivery. *N Engl J Med, 342*(20), 1500-1507. doi: 10.1056/nejm200005183422007

13. Smith, L. K., Draper, E. S., Manktelow, B. N., Dorling, J. S., & Field, D. J. (2007). Socioeconomic inequalities in very preterm birth rates. *Arch Dis Child Fetal Neonatal Ed, 92*(1), F11-14. doi: 10.1136/adc.2005.090308

14. Thompson, J. M., Irgens, L. M., Rasmussen, S., & Daltveit, A. K. (2006). Secular trends in socio-economic status and the implications for preterm birth. *Paediatr Perinat Epidemiol, 20*(3), 182-187. doi: 10.1111/j.1365-3016.2006.00711.x

15. Weng, Y. H., Yang, C. Y., & Chiu, Y. W. (2014). Risk Assessment of Adverse Birth Outcomes in Relation to Maternal Age. *PLoS One, 9*(12), e114843. doi: 10.1371/journal.pone.0114843

16. Bryant, A. S., Worjoloh, A., Caughey, A. B., & Washington, A. E. (2010). Racial/ethnic disparities in obstetric outcomes and care: prevalence and determinants. *Am J Obstet Gynecol, 202*(4), 335-343. doi: 10.1016/j.ajog.2009.10.864

17. Copper, R. L., Goldenberg, R. L., Das, A., Elder, N., Swain, M., Norman, G., et al. (1996). The preterm prediction study: maternal stress is associated with spontaneous preterm birth at less than thirty-five weeks' gestation. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. *Am J Obstet Gynecol, 175*(5), 1286-1292.

18. Cnattingius, S. (2004). The epidemiology of smoking during pregnancy: smoking prevalence, maternal characteristics, and pregnancy outcomes. *Nicotine Tob Res, 6 Suppl 2*, S125-140. doi: 10.1080/14622200410001669187

19. Hendler, I., Goldenberg, R. L., Mercer, B. M., Iams, J. D., Meis, P. J., Moawad, A. H., et al. (2005). The Preterm Prediction Study: association between maternal body mass index and spontaneous and indicated preterm birth. *Am J Obstet Gynecol, 192*(3), 882-886. doi: 10.1016/j.ajog.2004.09.021

20. Liu, S., Krewski, D., Shi, Y., Chen, Y., & Burnett, R. T. (2003). Association between gaseous ambient air pollutants and adverse pregnancy outcomes in Vancouver, Canada. *Environ Health Perspect, 111*(14), 1773-1778.

21. Wilhelm, M., & Ritz, B. (2003). Residential proximity to traffic and adverse birth outcomes in Los Angeles county, California, 1994-1996. *Environ Health Perspect, 111*(2), 207-216.

22. Bobak, M. (2000). Outdoor air pollution, low birth weight, and prematurity. *Environ Health Perspect, 108*(2), 173-176.

23. Xu, X., Ding, H., & Wang, X. (1995). Acute effects of total suspended particles and sulfur dioxides on preterm delivery: a community-based cohort study. *Arch Environ Health, 50*(6), 407-415. doi: 10.1080/00039896.1995.9935976

24. Huynh, M., Woodruff, T. J., Parker, J. D., & Schoendorf, K. C. (2006). Relationships between air pollution and preterm birth in California. *Paediatr Perinat Epidemiol, 20*(6), 454-461. doi: 10.1111/j.1365-3016.2006.00759.x

25. Ritz, B., Yu, F., Chapa, G., & Fruin, S. (2000). Effect of air pollution on preterm birth among children born in Southern California between 1989 and 1993. *Epidemiology, 11*(5), 502-511.

26. Sagiv, S. K., Mendola, P., Loomis, D., Herring, A. H., Neas, L. M., Savitz, D. A., & Poole, C. (2005). A time-series analysis of air pollution and preterm birth in Pennsylvania, 1997-2001. *Environ Health Perspect, 113*(5), 602-606.

27. Wilhelm, M., Ghosh, J. K., Su, J., Cockburn, M., Jerrett, M., & Ritz, B. (2011). Traffic-related air toxics and preterm birth: a population-based case-control study in Los Angeles County, California. *Environ Health, 10*, 89. doi: 10.1186/1476-069x-10-89

28. Wilhelm, M., & Ritz, B. (2005). Local variations in CO and particulate air pollution and adverse birth outcomes in Los Angeles County, California, USA. *Environ Health Perspect, 113*(9), 1212-1221.

29. Perera, F. P., Tang, D., Rauh, V., Lester, K., Tsai, W. Y., Tu, Y. H., et al. (2005). Relationships among polycyclic aromatic hydrocarbon-DNA adducts, proximity to the World Trade Center, and effects on fetal growth. *Environ Health Perspect, 113*(8), 1062-1067.

30. Lin, M. C., Chiu, H. F., Yu, H. S., Tsai, S. S., Cheng, B. H., Wu, T. N., et al. (2001). Increased risk of preterm delivery in areas with air pollution from a petroleum refinery plant in Taiwan. *J Toxicol Environ Health A, 64*(8), 637-644. doi: 10.1080/152873901753246232

31. Sram, R. J., Binkova, B., Dejmek, J., & Bobak, M. (2005). Ambient air pollution and pregnancy outcomes: a review of the literature. *Environ Health Perspect, 113*(4), 375-382.

32. Stillerman, K. P., Mattison, D. R., Giudice, L. C., & Woodruff, T. J. (2008). Environmental exposures and adverse pregnancy outcomes: a review of the science. *Reprod Sci, 15*(7), 631-650. doi: 10.1177/1933719108322436

33. Longnecker, M. P., Klebanoff, M. A., Brock, J. W., & Guo, X. (2005). Maternal levels of polychlorinated biphenyls in relation to preterm and small-for-gestational-age birth. *Epidemiology, 16*(5), 641-647.

34. Dabrowski, S., Hanke, W., Polanska, K., Makowiec-Dabrowska, T., & Sobala, W. (2003). Pesticide exposure and birthweight: an epidemiological study in Central Poland. *Int J Occup Med Environ Health, 16*(1), 31-39.

35. Villanueva, C. M., Durand, G., Coutte, M. B., Chevrier, C., & Cordier, S. (2005). Atrazine in municipal drinking water and risk of low birth weight, preterm delivery, and small-for-gestational-age status. *Occup Environ Med, 62*(6), 400-405. doi: 10.1136/oem.2004.016469

36. Valero De Bernabe, J., Soriano, T., Albaladejo, R., Juarranz, M., Calle, M. E., Martinez, D., & Dominguez-Rojas, V. (2004). Risk factors for low birth weight: a review. *Eur J Obstet Gynecol Reprod Biol, 116*(1), 3-15. doi: 10.1016/j.ejogrb.2004.03.007

37. Olsen, I. E., Groveman, S. A., Lawson, M. L., Clark, R. H., & Zemel, B. S. (2010). New intrauterine growth curves based on United States data. *Pediatrics, 125*(2), e214-224. doi: 10.1542/peds.2009-0913

38. Osmond, C., & Barker, D. J. (2000). Fetal, infant, and childhood growth are predictors of coronary heart disease, diabetes, and hypertension in adult men and women. *Environ Health Perspect, 108 Suppl 3*, 545-553.

39. Neggers, Y., & Goldenberg, R. L. (2003). Some thoughts on body mass index, micronutrient intakes and pregnancy outcome. *J Nutr, 133*(5 Suppl 2), 1737S-1740S.

40. Holt, V. L., Danoff, N. L., Mueller, B. A., & Swanson, M. W. (1997). The association of change in maternal marital status between births and adverse pregnancy outcomes in the second birth. *Paediatr Perinat Epidemiol, 11 Suppl 1*, 31-40.

41. Lund, R., Modvig, J., Hilden, J., Rosdahl, N., Kure, L., & Schmidt, K. (1999). Risk of low birthweight in social districts of Copenhagen. *Scand J Public Health, 27*(2), 89-93.

42. Millar, W. J., & Chen, J. (1998). Maternal education and risk factors for small-for-gestational-age births. *Health Rep, 10*(2), 43-51 (Eng); 47-56 (Fre).

43. Desjardins, E., & Hardwick, D. (1999). How many visits by health professionals are needed to make a difference in low birthweight? A dose-response study of the Toronto Healthiest Babies Possible program. *Can J Public Health, 90*(4), 224-228.

44. McFarlane, J., Parker, B., & Soeken, K. (1996). Physical abuse, smoking, and substance use during pregnancy: prevalence, interrelationships, and effects on birth weight. *J Obstet Gynecol Neonatal Nurs, 25*(4), 313-320.

45. Lazzaroni, F., Bonassi, S., Magnani, M., Calvi, A., Repetto, E., Serra, F., et al. (1993). Moderate maternal drinking and outcome of pregnancy. *Eur J Epidemiol, 9*(6), 599-606.

46. Plessinger, M. A., & Woods, J. R., Jr. (1998). Cocaine in pregnancy. Recent data on maternal and fetal risks. *Obstet Gynecol Clin North Am, 25*(1), 99-118.

47. Christian, M. S., & Brent, R. L. (2001). Teratogen update: evaluation of the reproductive and developmental risks of caffeine. *Teratology, 64*(1), 51-78. doi: 10.1002/tera.1047

48. Wergeland, E., & Strand, K. (1998). Work pace control and pregnancy health in a population-based sample of employed women in Norway. *Scand J Work Environ Health, 24*(3), 206-212.

49. Ardawi, M. S., Nasrat, H. A., Jamal, H. S., Al-Sagaaf, H. M., & Mustafa, B. E. (2000). Screening for gestational diabetes mellitus in pregnant females. *Saudi Med J, 21*(2), 155-160.

50. Scholl, T. O., Sowers, M., Chen, X., & Lenders, C. (2001). Maternal glucose concentration influences fetal growth, gestation, and pregnancy complications. *Am J Epidemiol, 154*(6), 514-520.

51. Dafopoulos, K. C., Galazios, G. C., Tsikouras, P. N., Koutlaki, N. G., Liberis, V. A., & Anastasiadis, P. G. (2002). Interpregnancy interval and the risk of preterm birth in Thrace, Greece. *Eur J Obstet Gynecol Reprod Biol, 103*(1), 14-17.

52. Zhu, B. P., Rolfs, R. T., Nangle, B. E., & Horan, J. M. (1999). Effect of the interval between pregnancies on perinatal outcomes. *N Engl J Med, 340*(8), 589-594. doi: 10.1056/nejm199902253400801

53. Murray, S. R., & Norman, J. E. (2014). Multiple pregnancies following assisted reproductive technologies--a happy consequence or double trouble? *Semin Fetal Neonatal Med, 19*(4), 222-227. doi: 10.1016/j.siny.2014.03.001

54. Powers, W. F., & Wampler, N. S. (1996). Further defining the risks confronting twins. *Am J Obstet Gynecol, 175*(6), 1522-1528.

55. Ha, E. H., Hong, Y. C., Lee, B. E., Woo, B. H., Schwartz, J., & Christiani, D. C. (2001). Is air pollution a risk factor for low birth weight in Seoul? *Epidemiology, 12*(6), 643-648.

56. Maisonet, M., Bush, T. J., Correa, A., & Jaakkola, J. J. (2001). Relation between ambient air pollution and low birth weight in the Northeastern United States. *Environ Health Perspect, 109 Suppl 3*, 351-356.

57. Ritz, B., & Yu, F. (1999). The effect of ambient carbon monoxide on low birth weight among children born in southern California between 1989 and 1993. *Environ Health Perspect, 107*(1), 17-25.

58. Salam, M. T., Millstein, J., Li, Y. F., Lurmann, F. W., Margolis, H. G., & Gilliland, F. D. (2005). Birth outcomes and prenatal exposure to ozone, carbon monoxide, and particulate matter: results from the Children's Health Study. *Environ Health Perspect, 113*(11), 1638-1644.

59. Bobak, M., & Leon, D. A. (1999). Pregnancy outcomes and outdoor air pollution: an ecological study in districts of the Czech Republic 1986-8. *Occup Environ Med, 56*(8), 539-543.

60. Lin, C. M., Li, C. Y., Yang, G. Y., & Mao, I. F. (2004). Association between maternal exposure to elevated ambient sulfur dioxide during pregnancy and term low birth weight. *Environ Res, 96*(1), 41-50. doi: 10.1016/j.envres.2004.03.005

61. Wang, X., Ding, H., Ryan, L., & Xu, X. (1997). Association between air pollution and low birth weight: a community-based study. *Environ Health Perspect, 105*(5), 514-520.

62. Dadvand, P., Parker, J., Bell, M. L., Bonzini, M., Brauer, M., Darrow, L. A., et al. (2013). Maternal exposure to particulate air pollution and term birth weight: a multi-country evaluation of effect and heterogeneity. *Environ Health Perspect, 121*(3), 267-373. doi: 10.1289/ehp.1205575

63. Dejmek, J., Solansky, I., Benes, I., Lenicek, J., & Sram, R. J. (2000). The impact of polycyclic aromatic hydrocarbons and fine particles on pregnancy outcome. *Environ Health Perspect, 108*(12), 1159-1164.

64. Dugandzic, R., Dodds, L., Stieb, D., & Smith-Doiron, M. (2006). The association between low level exposures to ambient air pollution and term low birth weight: a retrospective cohort study. *Environ Health, 5*, 3. doi: 10.1186/1476-069x-5-3

65. Parker, J. D., Woodruff, T. J., Basu, R., & Schoendorf, K. C. (2005). Air pollution and birth weight among term infants in California. *Pediatrics, 115*(1), 121-128. doi: 10.1542/peds.2004-0889

66. Sapkota, A., Chelikowsky, A. P., Nachman, K. E., Cohen, A. J., & Ritz, B. (2012). Exposure to particulate matter and adverse birth outcomes: a comprehensive review and meta-analysis. *Air Quality, Atmosphere & Health, 5*(4), 369-381. doi: 10.1007/s11869-010-0106-3

67. Xu, X., Sharma, R. K., Talbott, E. O., Zborowski, J. V., Rager, J., Arena, V. C., & Volz, C. D. (2011). PM10 air pollution exposure during pregnancy and term low birth weight in

Allegheny County, PA, 1994-2000. *Int Arch Occup Environ Health, 84*(3), 251-257. doi: 10.1007/s00420-010-0545-z

68. Lee, P.-C., Talbott, E. O., Roberts, J. M., Catov, J. M., Sharma, R. K., & Ritz, B. (2011). Particulate Air Pollution Exposure and C-reactive Protein During Early Pregnancy. *Epidemiology, 22*(4), 524-531.

69. Choi, H., Rauh, V., Garfinkel, R., Tu, Y., & Perera, F. P. (2008). Prenatal exposure to airborne polycyclic aromatic hydrocarbons and risk of intrauterine growth restriction. *Environ Health Perspect, 116*(5), 658-665. doi: 10.1289/ehp.10958

70. Vassilev, Z. P., Robson, M. G., & Klotz, J. B. (2001). Associations of polycyclic organic matter in outdoor air with decreased birth weight: a pilot cross-sectional analysis. *J Toxicol Environ Health A, 64*(8), 595-605. doi: 10.1080/152873901753246205

71. Lin, M. C., Yu, H. S., Tsai, S. S., Cheng, B. H., Hsu, T. Y., Wu, T. N., & Yang, C. Y. (2001). Adverse pregnancy outcome in a petrochemical polluted area in Taiwan. *J Toxicol Environ Health A, 63*(8), 565-574. doi: 10.1080/152873901316857743

72. Rogers, J. F., Thompson, S. J., Addy, C. L., McKeown, R. E., Cowen, D. J., & Decoufle, P. (2000). Association of very low birth weight with exposures to environmental sulfur dioxide and total suspended particulates. *Am J Epidemiol, 151*(6), 602-613.

73. Li, X., Sundquist, J., & Sundquist, K. (2010). Parental occupation and risk of small-for-gestational-age births: a nationwide epidemiological study in Sweden. *Hum Reprod, 25*(4), 1044-1050. doi: 10.1093/humrep/deq004

74. Quansah, R., & Jaakkola, J. J. (2009). Paternal and maternal exposure to welding fumes and metal dusts or fumes and adverse pregnancy outcomes. *Int Arch Occup Environ Health, 82*(4), 529-537. doi: 10.1007/s00420-008-0349-6

75. Shirai, S., Suzuki, Y., Yoshinaga, J., & Mizumoto, Y. (2010). Maternal exposure to low-level heavy metals during pregnancy and birth size. *J Environ Sci Health A Tox Hazard Subst Environ Eng, 45*(11), 1468-1474. doi: 10.1080/10934529.2010.500942

76. Zahran, S., Weiler, S., Mielke, H. W., & Pena, A. A. (2012). Maternal benzene exposure and low birth weight risk in the United States: a natural experiment in gasoline reformulation. *Environ Res, 112*, 139-146. doi: 10.1016/j.envres.2011.11.008

77. Kannan, S., Misra, D. P., Dvonch, J. T., & Krishnakumar, A. (2006). Exposures to airborne particulate matter and adverse perinatal outcomes: a biologically plausible mechanistic framework for exploring potential effect modification by nutrition. *Environ Health Perspect, 114*(11), 1636-1642.

78. Dejmek, J., Selevan, S. G., Benes, I., Solansky, I., & Sram, R. J. (1999). Fetal growth and maternal exposure to particulate matter during pregnancy. *Environ Health Perspect, 107*(6), 475-480.

79. Mohorovic, L. (2004). First two months of pregnancy--critical time for preterm delivery and low birthweight caused by adverse effects of coal combustion toxics. *Early Hum Dev, 80*(2), 115-123. doi: 10.1016/j.earlhumdev.2004.06.001

80. Perera, F. P., Jedrychowski, W., Rauh, V., & Whyatt, R. M. (1999). Molecular epidemiologic research on the effects of environmental pollutants on the fetus. *Environ Health Perspect, 107 Suppl 3*, 451-460.

81. Perera, F. P., Whyatt, R. M., Jedrychowski, W., Rauh, V., Manchester, D., Santella, R. M., & Ottman, R. (1998). Recent developments in molecular epidemiology: A study of the effects of environmental polycyclic aromatic hydrocarbons on birth outcomes in Poland. *Am J Epidemiol, 147*(3), 309-314.

82. Nicol, C. J., Harrison, M. L., Laposa, R. R., Gimelshtein, I. L., & Wells, P. G. (1995). A teratologic suppressor role for p53 in benzo[a]pyrene-treated transgenic p53-deficient mice. *Nat Genet, 10*(2), 181-187. doi: 10.1038/ng0695-181

83. Maisonet, M., Correa, A., Misra, D., & Jaakkola, J. J. (2004). A review of the literature on the effects of ambient air pollution on fetal growth. *Environ Res, 95*(1), 106-115. doi: 10.1016/j.envres.2004.01.001

84. Goerre, S., Staehli, C., Shaw, S., & Luscher, T. F. (1995). Effect of cigarette smoking and nicotine on plasma endothelin-1 levels. *J Cardiovasc Pharmacol, 26 Suppl 3*, S236-238.

85. Pennsylvania Department of Environmental Protection. (2015). Marcellus Shale: Frequently Asked Questions. Retrieved March 6, 2015

86. Department of Energy. (2009). Modern Shale Gas Development in the United States: A Primer. *2014*. http://energy.gov/fe/downloads/modern-shale-gas-development-united-states-primer

87. Roy, A. A., Adams, P. J., & Robinson, A. L. (2014). Air pollutant emissions from the development, production, and processing of Marcellus Shale natural gas. *Journal of the Air & Waste Management Association, 64*(1), 19-37.

88. Small, M. J., Stern, P. C., Bomberg, E., Christopherson, S. M., Goldstein, B. D., Israel, A. L., et al. (2014). Risks and risk governance in unconventional shale gas development. *Environ Sci Technol, 48*(15), 8289-8297. doi: 10.1021/es502111u

89. Ferrar, K. J., Michanowicz, D. R., Christen, C. L., Mulcahy, N., Malone, S. L., & Sharma, R. K. (2013). Assessment of effluent contaminants from three facilities discharging

Marcellus Shale wastewater to surface waters in Pennsylvania. *Environ Sci Technol, 47*(7), 3472-3481. doi: 10.1021/es301411q

90. Vidic, R. D., Brantley, S. L., Vandenbossche, J. M., Yoxtheimer, D., & Abad, J. D. (2013). Impact of shale gas development on regional water quality. *Science, 340*(6134), 1235009. doi: 10.1126/science.1235009

91. Chapman, E. C., Capo, R. C., Stewart, B. W., Kirby, C. S., & Hammack, R. W. (2012). Geochemical and Strontium Isotope Characterization of Produced Waters from Marcellus Shale Natural Gas Extraction. *Environ Sci Technol, 46*(6), 3545-3553.

92. Warner, N. R., Christie, C. A., Jackson, R. B., & Vengosh, A. (2013). Impacts of Shale Gas Wastewater Disposal on Water Quality in Western Pennsylvania. *Environmental Science and Technology, 47*(20), 11849-11857.

93. Lupo, P. J., Symanski, E., Waller, D. K., Chan, W., Langlois, P. H., Canfield, M. A., & Mitchell, L. E. (2011). Maternal exposure to ambient levels of benzene and neural tube defects among offspring: Texas, 1999-2004. *Environ Health Perspect, 119*(3), 397-402. doi: 10.1289/ehp.1002212

94. Esswein, E. J. (2013). NIOSH field effort to assess chemical exposures in oil and gas workers: Health hazards in hydraulic fracturing. In I. o. Medicine (Ed.), *Health impact assessment of shale gas extraction: Workshop summary*. Washington, D.C.: The National Academic Press.

95. Osborn, S. G., Vengosh, A., Warner, N. R., & Jackson, R. B. (2011). Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing. *Proc Natl Acad Sci U S A, 108*(20), 8172-8176. doi: 10.1073/pnas.1100682108

96. McKenzie, L. M., Witter, R. Z., Newman, L. S., & Adgate, J. L. (2012). Human health risk assessment of air emissions from development of unconventional natural gas resources. *Sci Total Environ, 424*, 79-87. doi: 10.1016/j.scitotenv.2012.02.018

97. Bunch, A. G., Perry, C. S., Abraham, L., Wikoff, D. S., Tachovsky, J. A., Hixon, J. G., et al. (2014). Evaluation of impact of shale gas operations in the Barnett Shale region on volatile organic compounds in air and potential human health risks. *Sci Total Environ, 468-469*, 832-842. doi: 10.1016/j.scitotenv.2013.08.080

98. Rabinowitz, P. M., Slizovskiy, I. B., Lamers, V., Trufan, S. J., Holford, T. R., Dziura, J. D., et al. (2015). Proximity to natural gas wells and reported health status: results of a household survey in Washington County, Pennsylvania. *Environ Health Perspect, 123*(1), 21-26. doi: 10.1289/ehp.1307732

99. Ferrar, K. J., Kriesky, J., Christen, C. L., Marshall, L. P., Malone, S. L., Sharma, R. K., et al. (2013). Assessment and longitudinal analysis of health impacts and stressors perceived to

result from unconventional shale gas development in the Marcellus Shale region. *Int J Occup Environ Health, 19*(2), 104-112. doi: 10.1179/2049396713y.0000000024

100. Currie, J. M., Deutch, J., & Meckel, K. H. (2014). *The Impact of the Fracking Boom on Infant Health: Evidence from Detailed Location Data on Wells and Infants [abstract]*. American Economic Association Annual Meeting.

101. Hill, E. L. (2013). *Unconventional natural gas development and infant health: evidence from Pennsylvania*. Charles Dyson School Applied Economics and Management. Cornell University. Ithaca, NY. Retrieved from http//: dyson.cornell.edu/research/researchpdf/wp/2012/Cornell-Dyson-wp1212.pdf

102. McKenzie, L. M., Guo, R., Witter, R. Z., Savitz, D. A., Newman, L. S., & Adgate, J. L. (2014). Birth outcomes and maternal residential proximity to natural gas development in rural Colorado. *Environ Health Perspect, 122*(4), 412-417. doi: 10.1289/ehp.1306722

103. Brauer, M., Lencar, C., Tamburic, L., Koehoorn, M., Demers, P., & Karr, C. (2008). A cohort study of traffic-related air pollution impacts on birth outcomes. *Environ Health Perspect, 116*(5), 680-686. doi: 10.1289/ehp.10952

104. Ghosh, J. K., Wilhelm, M., Su, J., Goldberg, D., Cockburn, M., Jerrett, M., & Ritz, B. (2012). Assessing the influence of traffic-related air pollution on risk of term low birth weight on the basis of land-use-based regression models and measures of air toxics. *Am J Epidemiol, 175*(12), 1262-1274. doi: 10.1093/aje/kwr469

105. American Psychological Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR*. Washington, D.C.

106. American Psychological Association. (2013). *Diagnostic and statistical manual of mental disorders* (5 ed.). Arlington, VA.

107. Centers for Disease Control and Prevention. (2007). Prevalence of autism spectrum disorders—autism and developmental disabilities monitoring network, 14 sites, United States, 2002. *MMWR Surveillaince Summaries, 56*(1), 12-28.

108. Centers for Disease Control and Prevention. (2014). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010. *MMWR Surveillance Summaries, 63*(SS02), 1-21.

109. King, M., & Bearman, P. (2009). Diagnostic change and the increased prevalence of autism. *Int J Epidemiol, 38*(5), 1224-1234. doi: 10.1093/ije/dyp261

110. Abrahams, B. S., & Geschwind, D. H. (2008). Advances in autism genetics: on the threshold of a new neurobiology. *Nat Rev Genet, 9*(5), 341-355. doi: 10.1038/nrg2346

111. Jorde, L. B., Hasstedt, S. J., Ritvo, E. R., Mason-Brothers, A., Freeman, B. J., Pingree, C., et al. (1991). Complex segregation analysis of autism. *Am J Hum Genet, 49*(5), 932-938.

112. Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med, 25*(1), 63-77.

113. Peters, S. U., Beaudet, A. L., Madduri, N., & Bacino, C. A. (2004). Autism in Angelman syndrome: implications for autism research. *Clin Genet, 66*(6), 530-536. doi: 10.1111/j.1399-0004.2004.00362.x

114. Zoghbi, H. Y. (2003). Postnatal neurodevelopmental disorders: meeting at the synapse? *Science, 302*(5646), 826-830. doi: 10.1126/science.1089071

115. Colvert, E., Tick, B., McEwen, F., Stewart, C., Curran, S. R., Woodhouse, E., et al. (2015). Heritability of Autism Spectrum Disorder in a UK Population-Based Twin Sample. *JAMA Psychiatry*. doi: 10.1001/jamapsychiatry.2014.3028

116. Lai, M., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/Gender Differences and Autism: Setting the Scene for Future Research. *J Am Acad Child Adolesc Psychiatry, 54*(1), 11-24. doi: 10.1016/j.jaac.2014.10.003

117. Croen, L. A., Najjar, D. V., Fireman, B., & Grether, J. K. (2007). Maternal and paternal age and risk of autism spectrum disorders. *Arch Pediatr Adolesc Med, 161*(4), 334-340. doi: 10.1001/archpedi.161.4.334

118. Maenner, M. J., Arneson, C. L., & Durkin, M. S. (2009). Socioeconomic disparity in the prevalence of autism spectrum disorder in Wisconsin. *WMJ, 108*(5), 253-255.

119. Lampi, K. M., Lehtonen, L., Tran, P. L., Suominen, A., Lehti, V., Banerjee, P. N., et al. (2012). Risk of autism spectrum disorders in low birth weight and small for gestational age infants. *J Pediatr, 161*(5), 830-836. doi: 10.1016/j.jpeds.2012.04.058

120. Singh, G. K., Kenney, M. K., Ghandour, R. M., Kogan, M. D., & Lu, M. C. (2013). Mental Health Outcomes in US Children and Adolescents Born Prematurely or with Low Birthweight. *Depress Res Treat, 2013*, 570743. doi: 10.1155/2013/570743

121. Lyall, K., Pauls, D. L., Spiegelman, D., Ascherio, A., & Santangelo, S. L. (2012). Pregnancy complications and obstetric suboptimality in association with autism spectrum disorders in children of the Nurses' Health Study II. *Autism Res, 5*(1), 21-30. doi: 10.1002/aur.228

122. Kalkbrenner, A. E., Braun, J. M., Durkin, M. S., Maenner, M. J., Cunniff, C., Lee, L. C., et al. (2012). Maternal smoking during pregnancy and the prevalence of autism spectrum disorders, using data from the autism and developmental disabilities monitoring network. *Environ Health Perspect, 120*(7), 1042-1048. doi: 10.1289/ehp.1104556

123. Palmer, R. F., Blanchard, S., Stein, Z., Mandell, D., & Miller, C. (2006). Environmental mercury release, special education rates, and autism disorder: an ecological study of Texas. *Health Place, 12*(2), 203-209. doi: 10.1016/j.healthplace.2004.11.005

124. Palmer, R. F., Blanchard, S., & Wood, R. (2009). Proximity to point sources of environmental mercury release as a predictor of autism prevalence. *Health Place, 15*(1), 18-24. doi: 10.1016/j.healthplace.2008.02.001

125. McCanlies, E. C., Fekedulegn, D., Mnatsakanova, A., Burchfiel, C. M., Sanderson, W. T., Charles, L. E., & Hertz-Picciotto, I. (2012). Parental occupational exposures and autism spectrum disorder. *J Autism Dev Disord, 42*(11), 2323-2334. doi: 10.1007/s10803-012-1468-1

126. Braun, J. M., Kalkbrenner, A. E., Just, A. C., Yolton, K., Calafat, A. M., Sjodin, A., et al. (2014). Gestational exposure to endocrine-disrupting chemicals and reciprocal social, repetitive, and stereotypic behaviors in 4- and 5-year-old children: the HOME study. *Environ Health Perspect, 122*(5), 513-520. doi: 10.1289/ehp.1307261

127. Stein, T. P., Schluter, M. D., Steer, R. A., Guo, L., & Ming, X. (2015). Bisphenol A Exposure in Children With Autism Spectrum Disorders. *Autism Res*. doi: 10.1002/aur.1444

128. Roberts, E. M., English, P. B., Grether, J. K., Windham, G. C., Somberg, L., & Wolff, C. (2007). Maternal residence near agricultural pesticide applications and autism spectrum disorders among children in the California Central Valley. *Environ Health Perspect, 115*(10), 1482-1489. doi: 10.1289/ehp.10168

129. Shelton, J. F., Geraghty, E. M., Tancredi, D. J., Delwiche, L. D., Schmidt, R. J., Ritz, B., et al. (2014). Neurodevelopmental disorders and prenatal residential proximity to agricultural pesticides: the CHARGE study. *Environ Health Perspect, 122*(10), 1103-1109. doi: 10.1289/ehp.1307044

130. Kalkbrenner, A. E., Daniels, J. L., Chen, J. C., Poole, C., Emch, M., & Morrissey, J. (2010). Perinatal exposure to hazardous air pollutants and autism spectrum disorders at age 8. *Epidemiology, 21*(5), 631-641. doi: 10.1097/EDE.0b013e3181e65d76

131. Roberts, A. L., Lyall, K., Hart, J. E., Laden, F., Just, A. C., Bobb, J. F., et al. (2013). Perinatal air pollutant exposures and autism spectrum disorder in the children of Nurses' Health Study II participants. *Environ Health Perspect, 121*(8), 978-984. doi: 10.1289/ehp.1206187

132. Windham, G. C., Zhang, L., Gunier, R., Croen, L. A., & Grether, J. K. (2006). Autism spectrum disorders in relation to distribution of hazardous air pollutants in the san francisco bay area. *Environ Health Perspect, 114*(9), 1438-1444.

133. Becerra, T. A., Wilhelm, M., Olsen, J., Cockburn, M., & Ritz, B. (2013). Ambient air pollution and autism in Los Angeles county, California. *Environ Health Perspect, 121*(3), 380-386. doi: 10.1289/ehp.1205827

134. Kalkbrenner, A. E., Windham, G. C., Serre, M. L., Akita, Y., Wang, X., Hoffman, K., et al. (2015). Particulate matter exposure, prenatal and postnatal windows of susceptibility, and autism spectrum disorders. *Epidemiology, 26*(1), 30-42. doi: 10.1097/ede.0000000000000173

135. Raz, R., Roberts, A. L., Lyall, K., Hart, J. E., Just, A. C., Laden, F., & Weisskopf, M. G. (2014). Autism Spectrum Disorder and Particulate Matter Air Pollution before, during, and after Pregnancy: A Nested Case-Control Analysis within the Nurses' Health Study II Cohort. *Environ Health Perspect*. doi: 10.1289/ehp.1408133

136. Volk, H. E., Hertz-Picciotto, I., Delwiche, L., Lurmann, F., & McConnell, R. (2011). Residential proximity to freeways and autism in the CHARGE study. *Environ Health Perspect, 119*(6), 873-877. doi: 10.1289/ehp.1002835

137. Volk, H. E., Lurmann, F., Penfold, B., Hertz-Picciotto, I., & McConnell, R. (2013). Traffic-related air pollution, particulate matter, and autism. *JAMA Psychiatry, 70*(1), 71-77. doi: 10.1001/jamapsychiatry.2013.266

138. Genc, S., Zadeoglulari, Z., Fuss, S. H., & Genc, K. (2012). The Adverse Effects of Air Pollution on the Nervous System. *Journal of Toxicology*. doi: http://dx.doi.org/10.1155/2012/782462

139. Dominici, F., Peng, R. D., Barr, C. D., & Bell, M. L. (2010). Protecting human health from air pollution: shifting from a single-pollutant to a multipollutant approach. *Epidemiology, 21*(2), 187-194. doi: 10.1097/EDE.0b013e3181cc86e8

140. von Ehrenstein, O. S., Aralis, H., Cockburn, M., & Ritz, B. (2014). In utero exposure to toxic air pollutants and risk of childhood autism. *Epidemiology, 25*(6), 851-858. doi: 10.1097/ede.0000000000000150

141. Harkonen, H. (1978). Styrene, its experimental and clinical toxicology. A review. *Scand J Work Environ Health, 4 Suppl 2*, 104-113.

142. Viaene, M. K., Pauwels, W., Veulemans, H., Roels, H. A., & Masschelein, R. (2001). Neurobehavioural changes and persistence of complaints in workers exposed to styrene in a polyester boat building plant: influence of exposure characteristics and microsomal epoxide hydrolase phenotype. *Occup Environ Med, 58*(2), 103-112.

143. Welp, E., Kogevinas, M., Andersen, A., Bellander, T., Biocca, M., Coggon, D., et al. (1996). Exposure to styrene and mortality from nervous system diseases and mental disorders. *Am J Epidemiol, 144*(7), 623-633.

144. National Toxicology Program. (2006). *NTP-CERHR Monograph on the Potential Human Reproductive and Developmental Effects of Styrene*. National Institutes of Health.

145. Wang, J., Jin, L., Gao, J., Shi, J., Zhao, Y., Liu, S., et al. (2013). Investigation of speciated VOC in vehicular exhaust under ECE and EUDC test cycles. *Science of the Total Environment, 445-446*, 110-116.

146. Cohen, J. T., Carlson, G., Charnley, G., Coggon, D., Delzell, E., Graham, J. D., et al. (2002). A comprehensive evaluation of the potential health risks associated with occupational and environmental exposure to styrene. *J Toxicol Environ Health B Crit Rev, 5*(1-2), 1-265. doi: 10.1080/10937400252972162

147. Franck, U., Weller, A., Roder, S. W., Herberth, G., Junge, K. M., Kohajda, T., et al. (2014). Prenatal VOC exposure and redecoration are related to wheezing in early infancy. *Environ Int, 73*, 393-401. doi: 10.1016/j.envint.2014.08.013

148. Ahmad, M., & Bajahlan, A. S. (2007). Leaching of styrene and other aromatic compounds in drinking water from PS bottles. *J Environ Sci (China), 19*(4), 421-426.

149. Genualdi, S., Nyman, P., & Begley, T. (2014). Updated evaluation of the migration of styrene monomer and oligomers from polystyrene food contact materials to foods and food simulants. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess, 31*(4), 723-733. doi: 10.1080/19440049.2013.878040

150. Rogers, J. (1955). Industrial hygiene problems in the field of plastics. *Arch Ind Health, 12*, 470-471.

151. Wilson, R. (1944). Health hazards encountered in the manufacture of synthetic rubber. *Journal of the American Medical Association, 124*, 701-703.

152. Chen, G.-D., Chi, L.-H., Kostyniak, P. J., & Henderson, D. (2007). Styrene Induced Alterations in Biomarkers of Exposure and Effects in the Cochlea: Mechanisms of Hearing Loss. *Toxicological Sciences, 98*(1), 167-177.

153. Committee to Review the Styrene Assessment in the National Toxicology Program 12th Report on Carcinogens; Board on Environmental Studies and Toxicology; Division on Earth and Life Sciences; National Research Council. (2014). *Review of the Styrene Assessment in the National Toxicology Program 12th Report on Carcinogens: Workshop Summary*. Washington DC: National Academies Press.

154. Morata, T. C., & Campo, P. (2002). Ototoxic effects of styrene alone or in concert with other agents: A review. *Noise and Health, 4*(14), 15-24.

155. Shugaev, B. (1969). Concentrations of hydrocarbons in tissues as a measure of toxicity. *Archives of Environmental Health, 18*, 878-882.

156. Savolainen, H., & Vainio, H. (1977). Organ distribution and nervous system binding of styrene and styrene oxide. *Toxicology, 8*, 135-141.

157. Rosengren, L., & Haglid, K. (1989). Long term neurotoxicity of styrene: A quantitative study of glial fibrillary acidic protein (GFA) and S-100. *British Journal of Industrial Medicine, 46*(316-320).

158. Benignus, V. A., Geller, A. M., Boyes, W. K., & Bushnell, P. J. (2005). Human Neurobehavioral Effects of Long-Term Exposure to Styrene: A Meta-Analysis. *Environmental Health Perspectives, 113*(5), 532-538.

159. Pennsylvania Department of Protection. (2015). Oil and Gas Reports. Retrieved December 15, 2014, from http://www.portal.state.pa.us/portal/server.pt/community/oil_and_gas_reports/20297

160. Council of Canadian Academics. (2014). Environmental impacts of shale gas extraction in Canada. Retrieved September 18, 2014, from http://www.assembly.gov.nt.ca/sites/default/files/td_105-175.pdf

161. Adgate, J. L., Goldstein, B. D., & McKenzie, L. M. (2014). Potential public health hazards, exposures and health effects from unconventional natural gas development. *Environ Sci Technol, 48*(15), 8307-8320. doi: 10.1021/es404621d

162. Goldstein, B. D., Brooks, B. W., Cohen, S. D., Gates, A. E., Honeycutt, M. E., Morris, J. B., et al. (2014). The role of toxicological science in meeting the challenges and opportunities of hydraulic fracturing. *Toxicol Sci, 139*(2), 271-283. doi: 10.1093/toxsci/kfu061

163. States, S., Cyprych, G., Stoner, M., Wydra, F., Kuchta, J., Monnell, J., & Casson, L. (2013). Marcellus Shale drilling and brominated THMs in Pittsburgh, Pa., drinking water. *Journal-American Water Works Association, 105*(8), E432-E448. doi: http://dx.doi.org/10.5942/jawwa.2013.105.0093

164. Wilson, J. M., & VanBriesen, J. M. (2012). Oil and gas produced water management and surface drinking water sources in Pennsylvania. *Environmental Practice, 14*(4), 288-300. doi: http://dx.doi.org/10.1017/S1466046612000427

165. Entrekin, S., Evans-White, M., Johnson, B., & Hagenbuch, E. (2011). Rapid expansion of natural gas development poses a threat to surface waters. *Frontiers in Ecology and the Environment, 9*(9), 503-511. doi: http://dx.doi.org/10.1890/110053

166. New York State Department of Health. (2014). A Public Health Review of High Volume Hydraulic Fracturing for Shale Gas Development Executive Summary. http://www.health.ny.gov/press/reports/docs/high_volume_hydraulic_fracturing.pdf

167. Adriaanse, H. P., Knottnerus, J. A., Delgado, L. R., Cox, H. H., & Essed, G. G. (1996). Smoking in Dutch pregnant women and birth weight. *Patient Educ Couns, 28*(1), 25-30.

168. Aronson, R. A., Uttech, S., & Soref, M. (1993). The effect of maternal cigarette smoking on low birth weight and preterm birth in Wisconsin, 1991. *Wis Med J, 92*(11), 613-617.

169. Landrigan, P. J., Kimmel, C. A., Correa, A., & Eskenazi, B. (2004). Children's health and the environment: public health issues and challenges for risk assessment. *Environ Health Perspect, 112*(2), 257-265.

170. Pennsylvania Department of Health. (2015). Birth Statistics: State/County Statistics Birth Weight. Retrieved August 18, 2014, from http://www.portal.state.pa.us/portal/server.pt?open=514&objID=809799&mode=2

171. Centers for Disease Control and Prevention Prevention. (2008). Quickstats: Percentage of Small for Gestational Age Births, by Race and Hispanic Ethnicity—United States, 2005. Retrieved January 15, 2015, from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5750a5.htm

172. Bibby, K. J., Brantley, S. L., Reible, D. D., Linden, K. G., Mouser, P. J., Gregory, K. B., et al. (2013). Suggested reporting parameters for investigations of wastewater from unconventional shale gas extraction. *Environ Sci Technol, 47*(23), 13220-13221. doi: 10.1021/es404960z

173. Allen, D. T., Torres, V. M., Thomas, J., Sullivan, D. W., Harrison, M., Hendler, A., et al. (2013). Measurements of methane emissions at natural gas production sites in the United States. *Proc Natl Acad Sci U S A, 110*(44), 17768-17773. doi: 10.1073/pnas.1304880110

174. United States Census Bureau. (2014). Map of 2010 population density by county or county equivalent. Retrieved August 18, 2014, from https://www.census.gov/geo/maps-data/maps/thematic.html

175. Pennsylvania Department of Health. (2015). Epidemiologic Query and Mapping System. Retrieved March 19, 2015, from https://apps.health.pa.gov/EpiQMS/asp/ChooseDataset.asp

176. DiGiuseppe, D. L., Aron, D. C., Ranbom, L., Harper, D. L., & Rosenthal, G. E. (2002). Reliability of birth certificate data: a multi-hospital comparison to medical records information. *Matern Child Health J, 6*(3), 169-179.

177. Backes, C. H., Nelin, T., Gorr, M. W., & Wold, L. E. (2013). Early life exposure to air pollution: how bad is it? *Toxicol Lett, 216*(1), 47-53. doi: 10.1016/j.toxlet.2012.11.007

178. Stieb, D. M., Chen, L., Eshoul, M., & Judek, S. (2012). Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis. *Environ Res, 117*, 100-111. doi: 10.1016/j.envres.2012.05.007

179. Forand, S. P., Lewis-Michl, E. L., & Gomez, M. I. (2012). Adverse birth outcomes and maternal exposure to trichloroethylene and tetrachloroethylene through soil vapor intrusion in New York State. *Environ Health Perspect, 120*(4), 616-621. doi: 10.1289/ehp.1103884

180. Henson, R. K., & Roberts, J. K. (2006). Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice. *Educational and Psychological Measurement, 66*(3), 393-416.

181. United States Environmental Protection Agency. (2013). Health Effects Notebook for Hazardous Air Pollutants: List of Health Effects Fact Sheets. Retrieved April 1, 2015, from http://www.epa.gov/ttn/atw/hlthef/hapindex.html

182. United States Environmental Protection Agency. (2013). Cumene Hazard Summary. Retrieved February 2, 2015, from http://www.epa.gov/airtoxics/hlthef/cumene.html

183. Colvert, E., Tick, B., McEwen, F., Stewart, C., Curran, S. R., Woodhouse, E., et al. (2015). Heritability of Autism Spectrum Disorder in a UK Population-Based Twin Sample. *JAMA Psychiatry, 4*(10).

184. United States Environmental Protection Agency. (2011). Technology Transfer Network Air Toxics 2005 National-Scale Air Toxics Assessment. Retrieved April 1, 2015, from http://www.epa.gov/ttn/atw/nata2005/tables.html

185. United States Environmental Protection Agency. (2015). TRI Explorer Release Reports. Retrieved April 1, 2015, from http://iaspub.epa.gov/triexplorer/tri_release.facility

186. United States Environmental Protection Agency. (2015). Learn about the Toxics Release Inventory. Retrieved April 1, 2015, from http://www2.epa.gov/toxics-release-inventory-tri-program/learn-about-toxics-release-inventory

187. Pennsylvania Spatial Data Access. (2013). Pennsylvania state roads shapefile. Retrieved February 6, 2015, from ftp://www.pasda.psu.edu/pub/pasda/

188. Pennsylvania Department of Transportation. (2015). Traffic Information. Retrieved April 1, 2015, from http://www.dot.state.pa.us/Internet/Bureaus/pdPlanRes.nsf/HomePageTransportationPlanning?OpenForm

189. Braubach, M., & Fairburn, J. (2010). Social inequities in environmental risks associated with housing and residential location--a review of evidence. *The European Journal of Public Health, 20*(1), 36-42. doi: http://dx.doi.org/10.1093/eurpub/ckp221