Determining the Relationship Between Male Waist Circumference and Male Fertility Markers

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37365621">https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37365621</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Determining the Relationship Between Male Waist Circumference
and Male Fertility Markers

David Bauer

A Thesis in the Field of Biology
For the Degree of Master of Liberal Arts in Extension Studies

Harvard University

November 2020
Abstract

Rates of obesity and overweight have been increasing in developing nations for decades and at the same time infertility rates have also been climbing. In the United States obesity among men above 20 years of age has more than tripled since the 1970s to over 33% in 2016 while at the same time, certain key markers of male fertility have decreased.

A great deal of research has been directed towards finding links between obesity and adverse reproductive outcomes in women, such as infertility and common pregnancy complications. However the same amount of investigation has not been undertaken for men’s fertility or reproductive health more generally. Furthermore the research that has been done for men’s fertility in relation to obesity or overweight has not yielded clear results.

The most commonly used clinical measurement of adiposity is body mass index (BMI), derived by dividing an individual’s mass in kg by their height in meters squared (kg/m²). Using a BMI chart or calculator, a clinician can categorize patient’s weight as “normal” (18.5 to < 25), “underweight” (<18.5) “overweight” (25 to < 30) or “obese” (≥30). This makes for a useful clinical measure, both easy to obtain and well known to be positively correlated with patient fat mass. However BMI is also positively correlated to lean mass, and suffers from an inability to factor in weight distribution. Another clinical measurement of adiposity, waist circumference (WC), a measurement that has been more closely linked to some adverse health trends than measures of mass or BMI may be a better proxy for central adiposity. Research that examines the link between central obesity as measured by WC and male fertility is limited.

To examine the relationship between WC and markers of male infertility I analyzed WC and fertility markers in 682 semen samples from 276 men undergoing assisted reproductive treatment for
infertility. By analyzing anthropometric data and creating generalized linear models to estimate semen quality parameters based on WC tertiles, I was able to determine that in couples seeking reproductive therapy, WC is inversely related to sperm concentration and total sperm count.
Acknowledgments

This work would not have been possible if not for the many people who supported me. First, I’d like to provide my deepest and sincerest thanks to my thesis director Dr. Jorge Chavarro, my research advisor Dr. James Morris and Dr. Lidia Mínguez-Alarcón, all of whom provided invaluable support, insight and answers at every stage of the thesis writing process. All of your kindness, and patience throughout this project is deeply appreciated.

I would like to thank the administrative staff and IT teams from the Harvard School of Public Health, whose assistance has been invaluable. In addition, I would like to extend special thanks to the Massachusetts General Hospital Fertility Center staff, members of the EARTH Study Team and to all participants of the EARTH study for making my research possible.

Lastly I would like to thank my informal research assistants, Katharine, Madeline, Maxwell and Cooper for your unwavering and exuberant support throughout the entire project.

I could not have completed this without all of your help.
Table of Contents

ACKNOWLEDGMENTS .......................................................................................................................... V

LIST OF TABLES .............................................................................................................................. VIII

LIST OF FIGURES ............................................................................................................................ IX

CHAPTER I: INTRODUCTION ........................................................................................................... 1
  Overview of the Problem ................................................................................................................... 1
  The Burden of Obesity ...................................................................................................................... 2
  The Decline in Male Fertility ........................................................................................................... 3
  Research Goal .................................................................................................................................. 5
  Definition of Terms ........................................................................................................................... 6
  Background ....................................................................................................................................... 9
  Question and Hypothesis .................................................................................................................. 12
  Selection of Confounders ................................................................................................................. 12

CHAPTER II: RESEARCH METHODS .............................................................................................. 26
  Dataset Cleaning and Preparation .................................................................................................. 26
  Modification and Generation of New Variables ................................................................................. 27
  Analysis of Demographic and Reproductive Characteristics ............................................................. 28
  Semen Quality Parameters Selection and Analysis ............................................................................. 29
  Creation of Mixed Models ............................................................................................................... 30

CHAPTER III: RESULTS .................................................................................................................... 32
  Demographic and Reproductive Characteristics .............................................................................. 32
List of Tables

Table 1. Demographic and reproductive characteristics [median (IQR) or counts (%)] by tertiles of waist circumference among 276 men in the EARTH Study. ................................................................. 39

Table 2. Distribution of semen quality parameters and abstinence time among 276 men contributing 682 semen samples in the EARTH Study. ................................................................. 42

Table 3. Semen quality parameters (adjusted mean, 95% CI) by tertiles of waist circumference among 276 men contributing 682 semen samples in the EARTH Study ................................................................. 46
List of Figures

Figure 1: Mixed Procedure Residual Diagnostics .............................................................. 31
Figure 2: Height Distribution ......................................................................................... 34
Figure 3: BMI Distribution ......................................................................................... 35
Chapter I: Introduction

Overview of the Problem

Since the 1970s obesity has steadily risen (WHO, 2018) paralleling a trend of decreasing fertility rates (Levine, 2017). While not fully understood, a significant amount of research has been completed regarding obesity and its negative link to fertility and adverse pregnancy outcomes in women (Bhattacharya, 2007; Wang, 2002; Funabiki, 2011). For men there has unfortunately been less fertility research completed and its findings are less clear (Sermondade, 2012; Chavarro, 2010; Guzick, 2001; Capelouto, 2018).

Obesity is generally defined and clinically assessed using the body mass index (BMI): body mass in kilograms divided by the square of the body height in meters. BMI has been used almost universally to make determinations of adiposity in clinical settings; by taking a patients height and weight and utilizing a BMI chart or calculator a clinician can categorize patients’ BMIs as “normal weight” (18.5 to < 25), “underweight” (<18.5) “overweight” (25 to < 30) or “obese” (≥ 30). While BMI is easy to obtain and known to be positively related to fat mass, it is also positively correlated to lean mass and does not capture information on weight distribution (WHO, 2018). Currently, research examining a link between male overall obesity and central adiposity, as measured by waist circumference, and male fertility is scarce.
The Burden of Obesity

As of the WHO’s most recent data on the subject, in 2016 more than 1.9 billion adults worldwide were overweight. 650 million, or about 13% of the world’s adult population (11% of men and 15% of women) were obese. During the same time period, in the United States, the Centers for Disease Control and Prevention (CDC) reported the prevalence of obesity as 93.3 million adults or 39.8% of the adult population (CDC, 2016).

Worldwide, the prevalence of overweight and obesity among children and adolescents aged 5-19 has risen dramatically from just 4% in 1975 to just over 18% in 2016 (WHO, 2018). Previously, the obesity epidemic was primarily a problem of high-income countries, however, overweight and obesity rates are climbing in urban areas of low and middle-income countries. Two areas with explosive growth rates are Africa, which is expecting a fifty percent increases in childhood overweight and obesity between 2010 and 2020 and Asia which now has over half (340 million) of the overweight and obese children under 5 years of age worldwide (Gupta, 2012).

Increased BMI is a major risk factor for several noncommunicable diseases including diabetes and cardiovascular diseases (both heart disease and stroke). BMI is positively correlated with the risk for these diseases as well as various forms of cancer, including endometrial, breast, ovarian, prostate, liver, gallbladder, kidney, and colon. Obesity during childhood is associated with a number of childhood health issues such as hypertension, cardiovascular disease and insulin resistance and it is also associated with premature death and disability in adulthood. Obesity is estimated to directly cost $209.7 billion annually in the United States, accounting for 20% of all annual health spending. Indirect costs of obesity are estimated to add an additional
$66 billion per year, yielding a total cost for obesity in the United States of $275 billion annually (Spiker, 2016).

The Decline in Male Fertility

Infertility is a disease of the reproductive system, defined by failure to achieve clinical pregnancy after 12 months or more of regular unprotected sexual intercourse (Zegers-Hochschild, 2009). While reliable figures for global prevalence of infertility are lacking, the WHO estimates that globally 72.4 million couples experience fertility problems (Mascarenhas, 2012; Boivin, 2007). Infertility rates are increasing within the United States, with 12% of men and 15.5% of women meeting the definition of infertile (Louis, 2013; Thoma, 2013).

“Male infertility” refers to a male's inability to achieve pregnancy with a fertile female after 1 year of regular coitus (Jequier, 2011). Infertility in men is commonly due to a decrease in semen quality; semen quality is frequently the surrogate used to measure male fertility (Cooper, 2010). “Male factor” infertility is the term used to describe this decrease in sperm concentration and/or degradation in motility and/or morphology. Male factor infertility affects 7% of men and is the cause of approximately half of infertility cases. Oligospermia (low sperm concentration) is the primary cause of male factor infertility, accounting for almost 90% of cases, it is also positively associated with abnormal semen parameters, most notably asthenospermia (poor sperm motility), and teratospermia (abnormal sperm morphology) (Brugh, 2004; Hirsh, 2003).

Over the last four decades, there has been a decline in semen quality. In 2017, Levine et al. performed a meta-regression analysis in their systematic review of temporal trends in sperm counts and found a significant decline in both sperm concentrations (SC) and total sperm count (TSC) in semen samples collected from 1973 to 2011. The analysis, which included data from
men in North America, Europe, Australia and New Zealand, suggested that this trend was not slowing and may continue. When restricted to samples collected in years 1996-2011, their analysis did not show any sign of “leveling off” (Levine, 2017).

While a number of studies have examined this downward trend in semen quality, few have examined the underlying causes of this trend, instead suggesting that poorly defined “environmental factors” may be to blame (Swan, 1997; Levine, 2017; Carlsen, 1992). To date, the only study to formally test this hypothesis analyzed data from men recruited between 2000 and 2017 to investigate the secular trends in semen parameters, and evaluate whether demographic, reproductive, nutritional and environmental factors previously related to semen quality contributed to the observed secular trends in semen quality. With the exception of urinary concentrations of phthalates, the researchers did not find any environmental, lifestyle, or nutritional factors which attenuated the downward trend of semen parameters. Relevantly, the researchers were unable to determine if obesity negatively correlated with semen quality in the men with evaluable demographic data because the prevalence of overweight and obesity remained stable, paralleling national trends over the study period (total cohort median 27.1 BMI, kg/m²) (Mínguez-Alarcón, 2018).

Researchers frequently focus on women’s role in reproduction, but it is also important to note that men play an essential role in creating healthy live births through their contribution of healthy gametes. It has been established that obesity is a risk factor for pregnancy chances, infertility, anovulation and poor pregnancy outcomes (Steeg, 2007; Dunn, 2015) and in addition some research has been done to examine BMI’s effect on the male role in decline in markers of clinical reproductive outcomes. In 2012 Sermondade et al. performed a review of the available literature examining the association of BMI and sperm count. In their research they obtained
primary data from 20 studies (n = 13,453) to determine that overweight and obesity were associated with an increased prevalence of azoospermia or oligozoospermia (Sermondade, 2012). While the exact mechanisms are not fully understood; male overweight and obesity has been found to affect sperm production and function through a number of different mechanisms including lowered circulating testosterone due to an increase in aromatization of androgens in adipose tissue, decreased amplitude and pulsatility of GnRH in the pituitary, increase scrotal temperature, oxidative stress and accumulation of toxic substances and liposoluble endocrine disruptors in fat tissue (Craig, 2017; Chavarro 2010; Bellastella 2019).

Subsequently, couples undergoing assisted reproductive treatment (ART) in which the male is overweight and obese are less likely to be successful. In their review of paternal obesity’s effect on reproductive potential, Campbell et al. found that in obese men undergoing ART, there was a statistically significant decrease in live birth rate and that the odds of a nonviable pregnancy are significantly greater for couples where the male partner is obese: OR 2.87 (95% CI 1.34–6.13) (Campbell, 2015).

Research Goal

While the increase in obesity and the decrease in male fertility have been occurring in parallel, the existing research on the subject still leaves the role of central adiposity on men’s fertility uncertain. Using a dataset from couples seeking ART, I will use the men’s semen quality parameters as a proxy for male fertility and WC as a measurement for central obesity to investigate an association between the two, and whether this relation is independent of BMI.
Definition of Terms

“Assisted reproductive technologies (ART)”: All fertility treatments in which both gametes and embryos are handled; which generally consist of eggs being surgically removed from a woman’s ovary, combined with sperm in the laboratory and placing the resulting embryo(s) in a uterus.

“Azoospermia”: Ejaculate lacking sperm.

“Body Mass Index (BMI)”: A measure of body obesity calculated by dividing weight in kilograms (kg) by the square of their height in meters (m²).

“Central adiposity”: The accumulation of fat in the lower torso around the abdominal area including both subcutaneous fat and visceral fat.

“General Linear Mixed Models (GLMM)”: a linear regression model which the linear predictor contains random effects to account for correlations between observations, in addition to the usual fixed effects.

“Institutional Review Board (IRB)”: an administrative body established to protect the rights and welfare of human research subjects recruited to participate in research activities conducted under the auspices of the institution with which it is affiliated.

“In-vitro fertilization (IVF)”: A procedure in which eggs are removed from a woman’s ovary
and combined with sperm outside the body to form embryos which are then grown in the
laboratory for several days before being placed in a uterus or in storage.

“Male factor infertility”: Infertility caused primarily by male factors encompassing abnormal
semen parameters or function; anatomical, endocrine, genetic, functional or immunological
abnormalities of the reproductive system; chronic illness; and sexual conditions incompatible
with the ability to deposit semen in the vagina.

“Male infertility”: The inability of a male to achieve conception with a fertile female after 12
months of regular coitus.

“Oligozoospermia”: Semen with a low concentration of sperm.

“SAS”: A statistical software suite developed by SAS Institute for data management that can
mine, alter, manage and retrieve data from a variety of sources and perform statistical analysis on
it.

“Sexually Transmitted Infection (STI)”: Bacteria, viruses or parasites that cause sexually
transmitted diseases that generally spread from person to person in blood, semen, or vaginal and
other bodily fluids during sexual contact.

“Time to Pregnancy (TTP)”: period of unprotected intercourse leading to a clinically detected
pregnancy.
“Waist Circumference (WC)”: A measurement taken around the abdomen at the level of the umbilicus to measure central adiposity.
Background

Several studies have examined and found a link between obesity and fertility using BMI and semen quality parameters. However, these studies have all suffered from shortcomings, including generally being of a small sample size (Magnusdottir, 2005; Hanafy, 2007; Zorn, 2007; Pauli, 2008).

At this time only three large epidemiological studies have allowed rigorous examination of obesity’s relation to male fertility. The first, *The Agricultural Health Study* was initiated to study the potential link between handling pesticides and various health issues. A study of secondary data found that in 1,329 evaluable couples there was a positive correlation between male infertility and obesity. However, participant BMI was based on self-reported height and weight which was used to calculate BMI. In addition, fertility was based on the participant’s reported ability to achieve pregnancy after 12 continuous months of attempts (Sallmén, 2006).

The *Danish National Birth Cohort* (DNBC) is a prospective cohort study established to investigate the causal link between early life exposures and later disease states. 49,957 couples enrolled in the DNBC were evaluated for links between male obesity and fertility. Subfertility (a time to conceive of 12 months or greater) was found in couples where the male partner was overweight and obese (Sermondade, 2015). However, the DNBC only evaluated data from couples who attained a birth, without considering couples unable to conceive or who had miscarriages. More importantly, the DNBC’s data collection is limited to telephone interviews and maternal blood samples, no semen samples were collected, and no anthropometric data
were obtained directly by study staff.

Similarly, a secondary analysis of The Norwegian Mother and Child Study, (MoBa), a prospective population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health (NIPH) which focused on women in the second trimester of pregnancy, defined infertility as a couple requiring greater than 12 months to conceive. The analysis found that in couples where the man was overweight or obese, the infertility rates were increased significantly. Unfortunately, the MoBa, similarly to the DNBC used telephone interviews to obtain self-reported BMI values and the study was conducted on couples who had already conceived. Also, as with the DNBC the MoBa’s biological sample collection consisted only of maternal blood samples making analysis of semen quality parameters impossible (Paltiel, 2014).

Clearly there is a current lack of studies examining the link between obesity and fertility that have a sufficiently large participant population, semen quality parameter measurements, and anthropometric data collected by a clinician or research study staff.

In addition to the previously described insufficiencies, the marker for obesity in all three large scale studies as well as the smaller studies is one that could be improved. While potentially only comprising a very small portion of overall body fat, abdominal fat has an outsized effect on overall health. Waist circumference is a useful way to measure abdominal fat because it includes subcutaneous as well as intra-abdominal adipose tissue. Furthermore in men WC more uniformly predicts the distribution of adipose tissue among several fat compartments in the abdominal region than BMI or similar clinical measurements (Chan, 2003).

Increased levels of abdominal fat have been found to place individuals at a greater risk for developing obesity-related conditions, such as Type 2 Diabetes, high blood pressure, and coronary artery disease (Klein, 2007). Moreover, abdominal obesity has been found to be a
better predictor of associated obesity-related health risks than BMI (Janssen, 2004) including reproductive risks (Derby, 2006).

The Environment and Reproductive Health (EARTH) Study is an ongoing prospective cohort that began in 2010 to investigate the impact of environmental, nutritional and lifestyle factors on fertility among couples undergoing infertility treatment with ART at the Massachusetts General Hospital (MGH) Fertility Center (Boston, MA). The study was approved by the Institutional Review Boards (IRBs) of MGH and the Harvard T.H. Chan School of Public Health. All participants had the study’s procedures explained, and study questions answered by study staff before being provided a written informed consent form (ICF) for their signature.

Consented study participants are given questionnaires to collect demographic as well as lifestyle and medical history data and are asked if researchers can access their medical records. Participants then have their anthropometric data collected by study staff, including waist circumference in centimeter rounded to the nearest millimeter. Participants received follow up during subsequent treatment visits up through clinical outcome or until treatment was discontinued.

Unlike the Agricultural Health or the MoBa and DNBC studies, the EARTH study contains anthropometric measurements collected by study staff during the first visit including height, weight and waist circumference. Compared to the MoBa and DNBC studies the EARTH study contains much more robust fertility information as it includes data from men who have conceived as well as those that have been unable to. Importantly, the EARTH study contains data from semen analyses from the male partner.

Using these lifestyle and anthropometric data from the EARTH study dataset I propose to test the hypothesis that male central obesity, as determined using waist circumference values, is
inversely related to markers of spermatogenesis independently of overall adiposity as measured by BMI.

Question and Hypothesis

The aim of my investigation is to test the hypothesis that central obesity, measured using waist circumference, is linked to male infertility as reflected in semen quality parameters. To test this hypothesis, I propose analyzing the most recent semen, lifestyle and anthropometric data of male participants enrolled in the EARTH study. This hypothesis is based on the observation that obesity is related to a number of health risks, and that there is evidence that for some obesity linked health risks, waist circumference is a more accurate measurement than other commonly used clinical measurements such as BMI (Seo, 2017; Lofgren, 2004; Shen, 2012).

Selection of Confounders

“Outcome” refers to the characteristic that is being predicted. For this paper, the primary outcome measure will be clinical reproductive outcome as represented by the surrogate measurement of semen quality parameters. An “exposure” is the primary explanatory variable, and other variables that are believed to explain or predict a study outcome. Exposures include confounders or effect modifiers which also must be addressed in the analysis of the primary outcome to determine if they bias the estimate of the effect of exposure on the outcome (Kestenbaum, 2019).

To select confounders for the analyses a tentative list was created while reviewing prior
literature on the relation between different factors and semen quality and looking at the association of these and other lifestyle factors with WC. After comparing with available data in the dataset the following ten variables were selected to be included in the analyses: abstinence time, age, BMI, education level, height, infertility diagnosis, obesity, physical activity, race, recreational drug use, presence of sexual transmitted infection, smoking status and varicocele. A brief discussion of the rationale for each confounder follows.

**Abstinence Time**

One of the primary considerations during sperm donation is the duration of abstinence between production of the sperm sample and previous ejaculation. During in vitro fertilization (IVF), standard practice is to collect semen samples after 2 – 7 days of sexual abstinence per WHO guidelines (WHO, 2010). The basis for the 2 – 7 day interval is that duration of abstinence is positively correlated with increased sperm concentration, volume, and sperm normal forms (De Jonge, 2004; Comar, 2017). Not all samples collected in the EARTH study adhered to the 2 – 7 day guidelines and given the correlation between abstinence time and semen quality parameters, it is a necessary covariate to consider.

**Age**

The variable that is most widely perceived as being related to male fertility and semen quality is the age of the sperm donor. Several groups have examined the effects of male age on semen quality and have reported general declines with advancing age. A literature review performed in 2001 by Kidd, et al. reviewed 70 studies evaluating age and semen quality
parameters including concentration, morphology, motility and semen volume. The findings of
the review showed decreases in morphology, motility and semen volume corresponding to
increase in age. Concentration, however, was not consistently found to be affected by age (Kidd,
2001). In Johnson et al. (2015), the researchers compared study datasets for over 90,000 subjects
and similar to Kidd’s findings, Johnson found a statistically significant decline in total sperm
count, morphology, motility and semen volume with increasing age and again were unable to
find a consistent link between age and semen concentration.

Body Mass Index

As discussed previously, BMI has been the subject of studies as a risk factor for male
infertility. Like WC, BMI is a measure of adiposity; however BMI suffers from inaccuracy
measuring total amount of adiposity of participants with higher amounts of lean mass. BMI is
also not as useful to measure central adiposity as WC and central adiposity is directly linked to a
number of adverse health risks (Ross, 2020). The relationship between BMI and semen quality
has been studied in the EARTH study previously but the findings were somewhat complicated.
While male BMI was found to have a slightly negative effect on semen quality, the results of
male BMI was positive with regard to fertilization and live-birth rates (Colaci, 2012). The
inclusion of BMI as a covariate will allow me to evaluate if central adiposity (as measured using
WC) is related to semen quality (as measured by male fertility markers) independent of markers
of overall adiposity reflected in BMI.
Education

Men who have a college or graduate degree are respectively 3 and 5 times more likely to seek infertility treatment as compared to men without higher education (Hotaling, 2012). This raises an important question of whether men with higher education levels seek infertility treatment because they are—for some undefined reason—less fertile than the general population. Since this is an open question, education level will be evaluated as a potential confounding variable. The EARTH dataset is comprised only of patients seeking IVF and because participants with higher education levels are more likely to seek infertility treatment, the number of participants without college degrees is relatively small. For this reason, while the EARTH dataset contains several categories of educational level (including categories for some high school, high school graduate, some college, etc.) these will be simplified into a new variable with only two values denoting whether the participant has completed a post-secondary degree or not.

Height

There have been a number of concerns raised with regard to waist circumference as a measurement due to different anthropometric measurements among certain demographics. For instance, in a cross-sectional study of Taiwanese subjects, waist-to-height ratio (WHtR) was found to be a better predictor of cardiometabolic risk than BMI or WC (Li, 2013). And waist circumference-to-height ratio (WCHt) has been proposed as a marker of adiposity-related morbidity due to its association with cardiometabolic risk factors and visceral fat in adolescents (Brambilla, 2013). To examine this potential effect on the outcome, height will be included as a confounder.
Infertility Diagnosis

When participants sought fertility treatment and enrolled in the EARTH study, their chart was reviewed to capture data regarding the medical reason for seeking fertility treatment. This information is captured in the EARTH dataset, as the variable infdx1, which provides categorical information on a couple’s infertility diagnosis. The infertility diagnosis variable gives a numeric response describing whether there was a diagnosis of male factor or female factor infertility, or if the cause of subfertility is unexplained.

In the EARTH study the participants and their partners’ infertility diagnosis status were obtained from the patients’ medical records, or through questionnaires during their initial visit. The responses were coded in the dataset as male factor infertility, female factor infertility or unexplained (1, 2 and 3 respectively). Of the 276 evaluable men, 72 (26.1%) were diagnosed as infertile. 83 (30.1%) of men’s partners were diagnosed infertile and for 121 (43.8%), no specific infertility diagnosis was made for either partner. Based on the high number of female factor, relatively low number of male factor (Leung, 2018; CDC, 2006) and no option for both partners to have been diagnosed with infertility it is likely that the diagnoses were made in a way that is insufficient for use in this paper and that a new variable for male factor infertility must be generated and used instead.

Infertility diagnoses are made and evaluated by clinicians in a sequential fashion throughout the patients’ visits to receive treatment. Based on the high unexplained value and no diagnoses of both male and female infertility as well as the fact that they were collected during their initial visit for the study, these diagnoses are likely to be predominantly determined during preliminary visits. In the United States, women are more likely to visit infertility clinics than men (13.4% versus 7.5%) and men are more likely to go in response to their partner, yet are the
cause of approximately 50% of infertility diagnoses (Anderson, 2009; Leung, 2018). Assuming that these are diagnoses made in preliminary visits, this may partially explain why the infdx1 values represent higher rates of female infertility than male. In addition, other common but non-male factor based causes of infertility such as varicocele may have been included in the “male factor diagnosis” category. Therefore a new male factor infertility variable was generated using the existing values for semen analysis data in the EARTH database. The variable, MFI (for Male Factor Infertility), will be a numeric variable that will be 0 if sample volume, concentration, count, motility percent and normal morphology count are all equal to or above WHO 5th edition lower reference limits for fertile men; if any of the variables are lower than the reference limit the MFI will be 1 (see Appendix 1, lines 14 to 34). While much more time consuming and cumbersome than using the existing variables in the dataset, this will ensure that we can exclusively examine infertility as defined by the clinical criteria for male factor infertility.

Obesity

There have been multiple studies that have established a negative correlation between obesity and male factor infertility (Du Plessis, 2010). In a 2008 literature review Hammoud discussed recent articles examining obesity and male factor infertility. In the review Hammoud explained that within the literature, the marked changes in hormonal profile noted in obese subjects may suggest endocrine dysregulation resulting in decreased semen parameter quality and subsequently male factor infertility (Hammoud, 2008). A specific variable was created in the EARTH dataset to denote if the participant meets the clinical criteria for obesity ($\geq 30$ kg/m$^2$). Given obesity’s well documented effect on semen parameters it is worthwhile to include clinically determined obesity as a potential confounder.
Acute and chronic exercise in males has been known to be associated with suppression of endocrine functions at the hypothalamic and testicular levels axes. More specifically, in males engaging in vigorous exercise gonadotropin-releasing hormone (GnRH) and serum testosterone are suppressed (Vaamonde, 2017). In 2012, Palmer performed a study to determine if abnormal sperm physiology and function could be reversed in obese mice through diet and exercise. Palmer gave the mice one of two diets: a low fat control diet or a high fat diet to increase adiposity and allowed both groups 30 minutes of swimming three times per week to simulate light exercise. Some of the mice from both diet groups were then selected to receive a moderate training program, in which they were allowed to swim freely for 60 minutes, five times per week for the following 18 weeks. At the end of the study, sperm was collected postmortem from the mice and sperm count, motility, and morphology were assessed. Palmer found that mice who had the high fat diet (and resulting increased adiposity) had decreased motility and increased abnormal morphology compared to the mice on the control diet. Palmer also found that those mice that were on the high fat diet who began the moderate training program’s semen parameters were significantly improved across all parameters compared to the mice that only performed light exercise (Palmer, 2012).

While some of the underlying mechanisms have been understood on the topic of physical activity and male fertility, and demonstrated in animal model, in many cases preliminary clinical research has led to paradoxical or contradictory results (Wise, 2011; Hayden, 2018). In a 2018 review article, Hayden identified eleven studies that focused on assessing physiology of exercise response, including work by Vaamonde and Wise, and determined that either a sedentary
lifestyle or excessive exercise are both detrimental to sperm parameters and concluded that light and moderate exercise are therefore beneficial to sperm health (Hayden, 2018). One of the few studies Hayden reviewed to explicitly support the conclusion was a 2014 paper by Gaskins et al., which used mixed models on the EARTH dataset to attempt to find a relationship between exercise and clinical fertility outcomes. Gaskins et al. were unable to find a relationship between clinical live births and amounts of exercise, however subgroup analyses of different types of physical activity were found to be related to sperm concentration (Gaskins, 2014). While the relation between exercise and sperm parameters seems complex and nuanced, it does seem to exist and therefore total minutes of physical activity (including occupational physical activity) will be included as a potential confounding variable.

Race

Few studies have examined race and male factor infertility in-depth; however infertility studies that have examined racial data have yielded interesting but often contradictory findings. In a 2001 study Green and her team did a retrospective chart review of 756 patients seeking infertility treatment in Cincinnati, Ohio between 1998 and 1999. When broken down into groups by race, the patient groups were fairly representative of the demographics of the city, 644 patients or 85.4% were white (compared to 87.2% in the population), 77 patients or 10.2% were black (compared to 11.7% in the population), and 33 or 4.4% of patients were categorized as “other” (1.1% in the population). While not statistically significant, Green did find that the infertility diagnoses differed between racial groups. Male factor infertility in whites comprised 24.5% of all infertility diagnoses, in the “other” group, male factor infertility was 21.2%. However, male factor infertility only comprised 11.5% among black infertility diagnoses (Green,
In 2003 a questionnaire was mailed to 1,500 women who were seeking fertility care at Brigham and Women’s Hospital, a Harvard University teaching hospital located in Boston, Massachusetts which is in many ways comparable to Massachusetts General Hospital where the EARTH study is based. The Brigham and Women’s questionnaire asked a number of demographic questions and 561 women responded to it with evaluable information. In 2006 using these responses, an analysis of the socioeconomic and racial disparities was performed. The analysis included the following racial subgroups: Caucasian (n = 454), African American (n = 25), Hispanic (n = 22), Chinese (n = 24) and Other Asian (n = 33). Of these subgroups the rates of diagnosis due to male factor infertility were markedly different; in Caucasians it was 21.6%, African Americans 24.0 Hispanics, only 13.6%, Chinese 16.7% and Other Asians 6.1% (Jain, 2006). Unfortunately, no further subgroup analyses were performed on these results, and the results are not consistent with Green’s study’s findings.

In a 2001 retrospective study, Costabile analyzed demographic data of 700 consecutive patient records at a military infertility clinic to try to characterize patients seeking infertility treatment in an environment where all patients have equal access to no-cost medical care. In their findings Costabile noted that race was not a significant factor influencing the prevalence of male factor infertility. Costabile also found no racial predominance and that the study population’s racial demographics mirrored those of the Department of Defense (DOD) healthcare beneficiary population at large (Costabile, 2001).

Given the contradictory nature of the existing research on the topic, and the potential for effect on the outcome, race will be examined in this paper as a potential confounding variable.
Recreational Drug Use

The most widely used recreational drug in the world is cannabis (frequently referred to in the United States as “marijuana”) with an estimated 188 million users in 2017, or approximately 3.5% of the world population. In North America however, use is much higher, at an estimated 13.8% (UNODC, 2019). The principal psychoactive constituent of marijuana is the cannabinoid tetrahydrocannabinol (THC). THC binds to CB1 and CB2, the so called “brain” and “spleen” cannabinoid receptors, which despite their names are found in various central and peripheral tissues. While there are numerous hormonal changes that occur from the binding of THC to CB1 and CB2, including changes to LH, TSH and testosterone levels, the two most relevant to male fertility markers are its effect on sperm and in Sertoli cells. In sperm, mitochondrial activity is reduced when CB1 receptors in the sperm are activated by THC, this reduction in mitochondrial activity results in reduced sperm motility (Battista, 2008). Sertoli cells provide nutrients and the hormonal signals that allow nascent germ cells to develop into sperm. When the CB2 receptors on the surface of Sertoli cells bind active THC, the Sertoli cells induces programmed cell death (Battista, 2008). Notably, in a 2019 study using data from the EARTH dataset, the researchers surprisingly found that men who had ever smoked marijuana had higher sperm concentration and total sperm count than men who had never smoked marijuana and found no differences between current and past marijuana smokers. This finding is inconsistent with the known adverse effects of marijuana on testicular function (Nassan, 2019).

Despite a decline in usage starting in the early 2000’s, North America has and maintains some of the highest levels of cocaine usage in the world, with 2.1% of the population having used or using cocaine compared to 0.4% of the global population (Peacock, 2018). Cocaine is a stimulant and potent vasoconstrictor which causes anesthetic effects. Increased use of cocaine
among subjects has been linked to lower sperm counts and motility and increased morphological abnormalities (Samplaski, 2015). In animal models it has been demonstrated that the vasoconstrictive effect of cocaine use causes seminiferous tubule degeneration. Animal models have also demonstrated that cocaine negatively impacts spermatogenesis, potentially by increasing serum prolactin, while decreasing free testosterone (Sharma, 2013).

With such a high prevalence of use, and potential for unpredictable or paradoxical impact on semen quality parameters, cocaine and marijuana use are potential confounding variables to be evaluated in this paper.

Sexually Transmitted Infections

It is commonly known that chlamydia and gonorrhea are important and preventable causes of infertility in women. Testing of sexually transmitted infections (STI) in men is at times framed by the concern of transmission to women and prevention of female infertility, yet, STIs pose a serious risk to male fertility as well.

In a 2007 study, PCR assays were performed on 241 semen samples from infertility patients showing no clinical manifestations of STI. The assays were designed to determine the prevalence and quantity of DNA from common viral STIs in the donor’s semen. And, the presence of STI DNA was analyzed to determine association with several reproductive function related measures (genital tract inflammation, impaired accessory gland, etc.) as well as semen quality. 18.7% of the samples were found to have STI DNA present and almost all STI subgroups were found to have reduced sperm concentration, total sperm count and motility. The reduction in sperm parameters only met the criteria for significance for concentration and motility in the herpes simplex virus (HSV) group and total sperm count for the human
papillomavirus virus (HPV) group, leading the author to conclude that screening for subclinical STI should be increased to prevent male infertility (Bezold, 2007). With this relationship in mind, STIs are being included in my analysis as a potential confounding variable.

Smoking Status

According to the World Health Organization, approximately one third of the world’s adult population smokes tobacco. Smoking has been linked with a number of adverse reproductive health outcomes, including a decrease in semen quality. While the mechanism causing smoking’s effect on semen is not understood fully, it is believed that cotinine—a water soluble metabolite in tobacco penetrates the blood-testis barrier (Künzle, 2003).

A 1996 cross section study by Vine attempted to find a relationship between smoking and semen parameters, men recruited into the study had their smoking history recorded and were assigned into one of three groups according to their current smoking habits. The nonsmoking group were men who had smoked fewer than 100 cigarettes in their lifetime. The “light smokers” group were men who on average smoked 1 – 19 cigarettes per day, and the remainder of the recruits who smoked 20 or more cigarettes were assigned to the “heavy smokers” group. After participants’ semen samples were analyzed it was determined that both current smoking status (cigarettes smoked per day) and smoking history (years smoked) had a negative association with semen density, total sperm count and sperm motility (Vine, 1996).

Building on those findings, in a more recent cohort study, semen was collected from 655 smokers and 1,131 non-smokers. While sperm quality variables for smokers were within the lab normal range, a statistically significant negative trend was found for sperm density, total sperm count, total number of motile sperm, and normal forms (Künzle, 2003). This is consistent with a
recent review in which smoking, caffeine and alcohol usage were examined with respect to clinical reproductive outcomes; in Mínguez-Alarcón et al. (2018), the author noted that while smoking is negatively associated with semen quality that does not imply the same negative association effect on a couple’s ART outcome. Vanegas et al. examined smoking patterns and clinical reproductive outcomes of couples in the EARTH study but did not find any significant associations, and only found a possible beneficial effect of smoking cessation among men (Vanegas, 2017). Given the prevalence of smoking in the adult population and its potential effect, smoking is included in this paper as a confounding variable.

Varicocele

Varicocele is a condition clinically defined as abnormally dilated and tortuous veins in the pampiniform plexus of the spermatic cord and it is also the most common correctable cause of male infertility. Varicocele are correlated with decreased semen quality, though the exact mechanism or mechanisms are not known. Suggested pathophysiological mechanisms of varicocele causing decreased semen parameters include formation of anti-sperm antibodies, scrotal hyperthermia, hypoxia, oxidative stress and hormonal imbalances (Jensen, 2017). Despite the mechanism not being fully known, the risks are: for a couple, a history of varicocele multiplies the risk of infertility by a factor of 28 (Thonneau, 1992). In addition to being a significant risk factor, varicocele is independent of obesity measurement; in a study examining lifestyle factors of 816 men seeking infertility treatment, it was found that 32% of the men had varicocele and that among those with varicocele their BMI was not significantly different than those without, making varicocele a reasonable choice as a potentially confounding variable (Shafi, 2014).
Chapter II: Research Methods

Dataset Cleaning and Preparation

To test the hypothesis that central obesity, measured using waist circumference, is linked to male infertility, as reflected in semen quality parameters, I analyzed the most recent data from male participants of the EARTH study using SAS version 9.4 (2016 by SAS Institute Inc., Cary, NC, US) a copy of the final SAS code annotated with line numbers for easy reference will be included as Appendix 1. Preliminary code used to generate the finalized dataset and the SAS output file is not included in this paper due to length considerations; relevant tables and figures from the output are included in the paper as needed. The final SAS code sorted and analyzed the dataset as follows:

First, three macros were run (see Appendix 1, lines 2 through 4). The first three macros formatted variables as part of a process step (PROC FORMAT) to change some numeric variables in the dataset into character variables. Next, to format these data and create data subsets for subsequent analyses with participants broken into categories according to WHO waist circumference guideline cutoffs, the “TILE.sas” macro was used. The purpose of TILE.sas is to assign records to a given percentile (tertile, quartile, quintile, etc.) and obtain the median within those percentiles. In this instance, the macro was run to assign all records to tertiles based on the WC variable (mbq_waistcirc_m). These macros were pre-written and while the TILE.sas macro was edited for the analysis to specify WC tertiles, it and the format macros are not included along with the SAS code in the appendices due to space considerations. It is also necessary to note that in the most recent version of the SAS code in Appendix 1, the Tile.sas macro was not
invoked; it had been used in a previous iteration of the code to generate the most recent version of the dataset and did not need to be run again, and was therefore intentionally left out of the final version of the SAS code.

To determine which participants are evaluable, I excluded participants lacking waist circumference or all sperm parameter data. All observations were sorted by ID number and then by date these data were produced using the sort procedure (PROC SORT) (see Appendix 1, lines 40 through 45).

Modification and Generation of New Variables

In the EARTH dataset, participants’ former and current use of marijuana and cocaine were collected. For the purposes of evaluating if marijuana or cocaine are confounding variables all of the separate marijuana and cocaine usage variables were converted into a single recreational drug use variable (rec_drug) which was scored as 0 for never having used marijuana or cocaine, and 1 if the participant had used marijuana or cocaine at any point in their life (see Appendix 1, lines 9 through 11).

The EARTH dataset includes participant information for specific STIs: Chlamydia, trichomonas, HSV, HPV, syphilis, gonorrhea, Lymphogranuloma venereum, Mycoplasma genitalium as well as STI streptococcus and a variable for all other STIs not listed. Similar to the recreational drug use variable, these individual STI variables were combined to create a new variable "any_sti" which combined all responses to the previously listed STIs, and if the participant had ever had an STI they were coded as 1, and if they had not, they were coded as 0. These statements were run in a previous iteration of the SAS code to prepare the dataset for the final analyses and are subsequently not included in the final code included in Appendix 1, and are instead included as Appendix 2.
Analysis of Demographic and Reproductive Characteristics

To create Table 1, I generated descriptive statistics, including manual count, determination of minimum and maximum values and calculation of the median, first and third quartile values for all variables. These data for the table were generated in two distinct ways:

First for the total cohort data, the means process (PROC MEANS) was invoked for all continuous variables. For categorical variables, the frequency process (PROQ FREQ) was run (see Appendix 1, lines 50 through 60). Each variable in each WC ranked tertile had one-way analysis of variance (ANOVA) as well as a test for location differences (Wilcoxon) performed. Given the limited sample sizes of some subgroups and that the distribution of the outcomes cannot be assumed to be approximately normally distributed, nonparametric tests were used. The one-way nonparametric test for location and scale differences procedure (NPAR1WAY) were performed. The NPAR1WAY procedure also performs an analysis of variance on the variables, empirical distribution function statistics, pairwise multiple comparison analysis, and stratified analysis. The Wilcoxon specified as part of the NPAR1WAY procedure is the Wilcoxon rank-sum test (also referred to as the Mann–Whitney U test or Wilcoxon–Mann–Whitney test), a nonparametric test to compare two unmatched groups. This test is performed by default in the NPAR1WAY procedure, however by specifying it as an “option” in SAS, an analysis of Wilcoxon scores were performed. For multiple classification levels invoking this option produces the Kruskal-Wallis test which can compare more than two independent samples and will determine if samples originate from the same distribution.

Kruskal-Wallis tests of continuous variables and chi-squared or Fishers’ exact tests for categorical variables were performed to determine any association of participant demographic variables with the waist circumference groups (see Appendix 1, lines 54 through 82).
After the preliminary run of these processes, I reviewed the variable data and observed distribution and outliers, checked data integrity and performed necessary data transformations. Absolute count semen parameter variables were log transformed prior to their use in generating Tables 2 and 3. Any variables found to be erroneous, were manually replaced by the mean prior to generating the final datasets and running the final code to generate the final output.

Distribution data generated in the final output as figures are included for continuous variables at or near significance at the 0.05 alpha level. The values generated for all variables are provided and contextualized in the results section and the summary statistics were transcribed into Table 1.

Semen Quality Parameters Selection and Analysis

To evaluate the semen parameters I first selected the relevant, unique variables used in the semen quality dataset. Examining the available semen parameters in the dataset, compared against the WHO 5th Edition guidelines for semen analyses, I chose to include total ejaculate volume, concentration (expressed in millions of sperm per mL), total sperm count (expressed in millions of sperm), motility (expressed as the sum total of the percent of progressive and non-progressive sperm), and morphology which will be expressed as the percent morphologically normal sperm (total and normal count). In addition, abstinence time was included.

Table 2 was created by using the PROC MEANS procedure to generate descriptive statistics including those used in Table 1 as well as standard deviation (see Appendix 1, lines 87 to 94).
Creation of Mixed Models

In order to make inferences about the WC dataset which is comprised of non-independent longitudinal participant data, linear mixed models are used. Mixed modeling is modeling the longitudinal data by introducing participants-specific random effects. For normally distributed data, the MIXED process (PROC MIXED) in SAS will generate linear mixed models (GLMM) to fit to the WC dataset. As with standard linear modeling, GLMM generate fitted models to the dataset variable’s means. In addition to this, mixed models create models for the variable variances and the covariances. These GLMM were used to evaluate the relations of waist circumference with semen quality parameters. Any association found between waist circumference and semen quality were tested for effect modification by known and presumed predictors of semen quality.

The MODEL statement in PROC MIXED determines the specifications of the intended GLMM. In the MODEL statement one dependent variable and the fixed effects, determine the GLMM’s matrix. In my models the dependent variable is the semen parameters and the fixed effects for my models are groups of the previously discussed potential confounders, for the first model these include abstinence time, age and BMI. The second GLMM uses the confounding variables age, abstinence time and BMI and also include education, race and recreational drug use (see Appendix 1, lines 106 to 239 and 245 to 378). These adjusted semen parameter results, listed by WC tertile from these two sets of GLMMs were used to create Table 3. RESIDUAL and SOLUTION were included in the MODEL statement for each GLMM; RESIDUAL requests plots of both conditional and marginal residuals. SOLUTION requests the predicted values for
random effects (also called best linear unbiased prediction or BLUPs). Panels of residual diagnostics were requested in SAS using the ODS GRAPHICS statement. The panels generated both marginal and conditional residuals calculated in four distinct ways: raw, Studentized, Pearson and Scaled. Studentized residuals are divided by their own standard deviation to simulate standardization, Pearson are divided by the estimated standard deviation $Y_i$ and Scaled residuals have zero mean and are approximately uncorrelated (Schabenberger, 2005). The equations describing the differences between the diagnostics are presented in Figure 1.

Figure 1: Mixed Procedure Residual Diagnostics

<table>
<thead>
<tr>
<th>Type of Residual</th>
<th>Conditional</th>
<th>Marginal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>$r_{ci} = r_{mi} - z_i^T \hat{\gamma}$</td>
<td>$r_{mi} = Y_i - x_i^T \hat{\beta}$</td>
</tr>
<tr>
<td>Scaled</td>
<td>$r_{ci} = \bar{C}^{-1} r_m$</td>
<td>$r_{mi} = \bar{C}^{-1} r_m$</td>
</tr>
<tr>
<td>Studentized</td>
<td>$r_{student} = \frac{r_{ci}}{\sqrt{\text{var}[r_{ci}]}}$</td>
<td>$r_{student} = \frac{r_{mi}}{\sqrt{\text{var}[r_{mi}]}}$</td>
</tr>
<tr>
<td>Pearson</td>
<td>$r_{pearson} = \frac{r_{ci}}{\sqrt{\text{var}[Y_i</td>
<td>\gamma]}}$</td>
</tr>
</tbody>
</table>

The panels include a plot of residuals versus predicted values, a histogram with normal density overlaid, a Q-Q plot, and summary residual and fit statistics for each model. The residual diagnostic panels will be reviewed to ensure normality, spot outliers and compare against predicted values.
Chapter III: Results

Demographic and Reproductive Characteristics

Participants were asked to abstain from ejaculating for between 2 and 7 days, per WHO guidelines (WHO, 2010). Participants were asked to provide an estimated number of hours that they had remained abstinent prior to producing their semen sample. Of the 276 evaluable participants, the vast majority were within the recommended guidelines the median length of abstinence being 2.2 days. The lower quartile was 2.08 days and the upper quartile was 2.58 days. While outside the scope of this paper, it is worth briefly discussing the minimum value for abstinence time, 0.08 days (2 hours). This value is of note however because while it is far outside the recommended abstinence time per WHO guidelines, the semen parameters for the sample were (with the exception of morphology) above the group averages and all were within the established clinical normal ranges. This outlier is reminder that abstinence time has a very complex relationship with sperm quality (Hanson, 2018). The median values for abstinence time was almost identical across tertiles; 2.2, 2.3 and 2.2 days; no correlation exists between abstinence time and WC tertile at the alpha level of 0.05 (p=0.24)

Of the evaluable participants who had supplied age values, age ranged from a minimum of 25.23 up to 66.59 years. While this is a significant span of ages, most men in the study tended to be in their mid to upper 30s. The group had a median of 36.0 and the lower and upper quartiles were 32.79 and 39.54 years of age respectively. While the means increased slightly across tertiles (35.7, 35.8 and 37.5) this change did not reach statistical significance (p=0.22).

The EARTH dataset contained participants’ height and weight data allowing researchers
to calculate BMI, which was also captured as a distinct variable in the dataset. The CDC
Anthropometric Reference Data for Children and Adults gives the average height of males in the
US over age 20 as 175.7 cm, and the average height of white (non-Hispanic) males in the US
aged 20–39 years as 178.2 cm (Fryar, 2016). This is similar to the median height for this dataset,
which was 180 cm, with a range of 117 to 200 cm. One participant in the dataset had an
erroneous value listed for their height of 68 cm. This participant only had a single visit, so no
comparative data was available to compare and potentially correct the error. The participant’s
height for that visit was changed to the dataset’s mean height value of 180 cm. Mean height and
corresponding IQRs increased between the first and third WC tertiles and the Kruskal-Wallis test
determined significance at an alpha level of 0.05 (p=0.01). Distribution of Wilcoxon scores for
height data is shown as a box-and-whisker plot for all three tertiles in Figure 2.
The CDC ranges for BMI are less than 18.5 being underweight, 18.5 to 24.9 being normal, 25.0–29.9 kg/m² is overweight; 30.0 kg/m² is obese; and severe obesity is BMI at or above 40.0 kg/m². The values of the evaluated participants from the EARTH dataset ranged from 18.55 to 73.24, with the median being 26.82. This is slightly lower but still comparable to the national average; as of 2016 the mean BMI for men in the United States aged 20-39 was 28.7 kg/m² (Fryar, 2018). Predictably, BMI increased by mean and IQR across WC tertiles (means=24.1, 26.5 and 31.1 kg/m²), showing significance (p<.0001). Distribution of Wilcoxon Scores for BMI data is shown as a box-and-whisker plot for all three tertiles in Figure 3.
In the EARTH dataset education is coded from 1-6, with the values corresponding to highest level of education attained as follows: 1=did not graduate from high school, 2=high school graduate, 3=1 or 2 years of college, 4=3 or 4 years of college, 5=college graduate, 6=graduate degree. Given the disproportionately high number of study participants having college education or above, the number of patients for each of the 1-4 values were comparatively very small. The responses in the dataset were instead consolidated into a new numbering system of 1-3, with EARTH dataset values 1, 2, 3 and 4 being re-coded as 1, 5 re-coded as 2 and 6 re-coded as 3. By recoding participant responses in this way it allowed for meaningful analysis of participants based on educational level. Even after combining the non-college graduate categories, the college graduate and advanced degree holders were much more prevalent. Of the 276 total men in the cohort, the newly reassigned Group 1 was the smallest of the three new subgroups, making up only 11% (n=29) of patients. Group 2 (college graduates) were 28%
(n=76) of the cohort and group 3 (graduate degree) were 62% (n=171). The p-trend for all three groups across all tertiles was p=0.03.

Though not directly relevant to this paper it is notable that the ratio of participants with graduate degrees decreased across tertiles while those of participants with college degrees or less education had the exact opposite trend, with a smaller ratio in T1 increasing across tertiles.

For the Male Factor Infertility variable generated, 276 men were evaluated, and 162 (59%) met the criteria for infertility. Across the three WC tertiles there was an increase in both count and percentage, from 50 (53%) in T1 to 55 (61%) in T2 up to 57 (63%) in T3. These are similar to those given in a literature review in 2015, where it was noted that globally in couples suffering from infertility, 40-50% of cases were due to male factor infertility (Kumar, 2015). Kumar also noted that rates were found to be higher in high income countries such as the United States; this estimate of approximately 50% has also been confirmed by other researchers (Dupree, 2016). The Kruskal-Wallis test did not demonstrate evidence of an association between WC tertile and MFI (p=0.33).

Of the 276 men evaluated 70 (25.36%) met the clinical criteria for obesity (BMI ≥ 30 kg/m²). These numbers are lower than those of the general population, of which 42.4% of adults are obese, however this number is consistent with the obesity levels of residents of the Boston metropolitan area, where the EARTH study was conducted. In the Boston metropolitan area only 22% of adults are obese (CDC, 2020). For the greater New England area, the obesity rate overall is 25.4% (95% CI 24.7-26.1), with New England metropolitan areas being slightly lower (25.0% overall) and rural areas being slightly higher (28.7% overall). Participation in the study is not necessarily restricted to residents of Boston and surrounding urban communities, and so it is reasonable to assume that the recruitment area for the study includes some rural areas, making
25.36% a very reasonable representative value.

Unsurprisingly, the goodness of fit test indicated that there was a strong correlation between obesity and waist circumference tertile (p <.0001), with the lowest tertile only having 1.05% (n=1) of subjects meeting obesity criteria, 6.67% (n=6) in the second tertile and 69.23% (n=63) of those in the third tertile being clinically obese.

The mean physical activity for patients was 255.5 minutes/week ranging from 0 to 3,540 minutes of activity. The lower quartile were physically active 16 minutes per week and the upper 591. The median for the first and second quartiles were similar, 284, and 305 with the 3rd tertile being lower, with 180 though this change did not reach significance (p=0.297). Prior to running this analysis, two highly implausible values, 7,296 and 6,060 minutes from two different subjects were removed from the dataset and replaced with the mean.

For 275 of the men, race was evaluated; 238 (86.55%) of which were white and 37 (13%) were categorized as Black, Asian or “other”. These numbers are markedly different from local geographic population data, 2019 census data for the city of Boston shows the population is only 52.6% percent white. However, 87% is similar in composition to the responses to the mailed questionnaire for former Brigham and Women’s ART patients in the 2006 Jain study, in which 80.9% of respondents (454 of the 561 total respondents) were white (Caucasian). This is similar to the overall state demographics, which according to the latest United States census is 80.8% “white” (Jain, 2006; U.S. Census Bureau, 2019).

Of the 276 men evaluated, 126 (45.7%) use or at have at some point used recreational drugs. The values across the WC tertiles were all similar, the chi-squared test for WC tertile and recreational drug use did find some association (p=0.1033). Similarly, for all 276 men who had previously or currently have a STI (n=27, 9.8%) there was no association between WC and STI
history (p=0.4499). The same was also true for smokers, who made up 88 (32%) of the 276 men, (p=0.39) as well as the 20 (7%) men of the 275 men evaluated who had been diagnosed with varicocele (p=0.3514).

Of all covariates examined, BMI, obesity and height were strongly associated with outcome at the alpha 0.05 level. The only non-anthropometric exposure to show some level of association was recreational drug use, with p=0.10. The results discussed thus far are summarized in Table 1.
Table 1. Demographic and reproductive characteristics [median (IQR) or counts (%)] by tertiles of waist circumference among 276 men in the EARTH Study.

<table>
<thead>
<tr>
<th></th>
<th>Tertiles of waist circumference (cm)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total cohort</td>
<td>T1</td>
<td>T2</td>
<td>T3</td>
<td>p-value</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>N=96</td>
<td>N=89</td>
<td>N=91</td>
<td>across</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>tertiles</td>
<td></td>
</tr>
<tr>
<td>WC median (range)</td>
<td>-</td>
<td>67.5 - 91.5</td>
<td>92.0 - 101</td>
<td>102 - 200</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Abstinence Time (Days)</td>
<td>2.2 (2.08 – 2.58)</td>
<td>2.2 (2.1 – 2.5)</td>
<td>2.3 (2.2 – 2.7)</td>
<td>2.2 (2.0 – 2.7)</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td>36.0 (32.8 – 39.5)</td>
<td>35.7 (32.6 – 38.2)</td>
<td>35.8 (33.0 – 40.2)</td>
<td>37.5 (33.0 – 40.2)</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>Body Mass Index, kg/m²</td>
<td>26.8 (24.4 – 30.0)</td>
<td>24.1 (22.7 – 25.4)</td>
<td>26.5 (25.1 – 27.9)</td>
<td>31.1 (29.4 – 34.4)</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Education, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Some high school/college</td>
<td>29 (11)</td>
<td>8 (8)</td>
<td>5 (6)</td>
<td>16 (18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>College graduate</td>
<td>76 (28)</td>
<td>22 (23)</td>
<td>25 (28)</td>
<td>29 (32)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate degree</td>
<td>171 (62)</td>
<td>65 (68)</td>
<td>60 (67)</td>
<td>46 (51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height, cm</td>
<td>180 (175 – 184)</td>
<td>178 (173 – 183)</td>
<td>180 (175 – 185)</td>
<td>180 (175 – 185)</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Male Factor Infertility</td>
<td>162 (59)</td>
<td>50 (53)</td>
<td>55 (61)</td>
<td>57 (63)</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Obese, n (%)</td>
<td>70 (25)</td>
<td>1 (1)</td>
<td>6 (7)</td>
<td>63 (69)</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Physical Activity, min/week</td>
<td>256 (16-591)</td>
<td>284 (47 – 582)</td>
<td>305 (0 – 689)</td>
<td>180.0 (0 – 441)</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>White/Caucasian</td>
<td>238 (87)</td>
<td>75 (80)</td>
<td>81 (90)</td>
<td>82 (90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black/Asian/Other</td>
<td>37 (13)</td>
<td>19 (20)</td>
<td>9 (10)</td>
<td>9 (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recreational Drug Use</td>
<td>126 (46)</td>
<td>49 (52)</td>
<td>33 (37)</td>
<td>44 (48)</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Sexual Transmitted Infection</td>
<td>27 (10)</td>
<td>10 (11)</td>
<td>6 (7)</td>
<td>11 (12)</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>Ever Smoker, n (%)</td>
<td>88 (32)</td>
<td>36 (38)</td>
<td>22 (24)</td>
<td>30 (33)</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Varicocele, n (%)</td>
<td>20 (7)</td>
<td>10 (11)</td>
<td>5 (6)</td>
<td>5 (5)</td>
<td>0.35</td>
<td></td>
</tr>
</tbody>
</table>
Semen Quality Parameters

To evaluate the semen quality parameters and abstinence time, data for all 682 available semen samples were used, with the exception of normal morphology. Both the normal morphology count and percent morphology variables had 42 missing values, so only 640 values were included in their analyses. The means process was performed to obtain mean, standard deviation, quartile values and the values of how many samples fell below the WHO reference values. Note that total motile count, normal morphology count and abstinence time were also measured, even though they lack corresponding WHO reference values.

The WHO reference values for semen analysis were obtained by Cooper et al. by testing the characteristics of semen of men who had a time to pregnancy (TTP) of less than or equal to 12 months. The parameters of these semen samples from fertile men were used to generate one-sided lower reference limits using the fifth centiles and 95% confidence intervals for each characteristic (Cooper, 2010). The measurements from the 12 month or fewer TTP group were much higher than those of the general unscreened population, so Cooper included data from the general population for reference and comparison. Using the established semen analysis reference values and the values provided in Cooper for semen parameters from the general population, we can make comparisons with the findings from the EARTH study dataset.

The WHO reference value for semen volume is 1.5 mL (1.4-1.7), in the EARTH dataset 48 samples fell below that limit (17%). The mean value of the samples evaluated in the EARTH dataset was 2.80 mL, (SD 1.38), median 2.70 and the IQR was 1.70 – 3.70. This is significantly lower than the values from the general population. The semen volume IQR from the general
population was 2.2 – 4.2 with a median of 3.2 mL.

The WHO reference value for sperm concentration is 15 (12-16) million per mL, of the 682 evaluated samples 41 (15%) fell below that limit. The mean for sperm concentration is 63.69 mil/mL (SD 60.51) with an IQR of 23.20-88.0 and a median of 47.60 mil/mL, lower than the general population whose median is 64, and IQR are 36-100 mil/mL.

Total sperm count for the 682 samples had a mean of 158.24 million per ejaculate. 41 (15%) of the samples fell below the reference limit of 39 mil. (33-46) per ejaculate. The IQR and median were below those of the general population median 116.03 (compared to 196) and IQR 56.05-225.20 (101-336) mil. per ejaculate.

The largest percentage of samples below the reference limit were due to issues with motility. The reference limits for total motility by percent are 40% (38–42), and 300 (44%) of the 682 samples fell below that limit. The mean for motility was 43.83% (SD 23.25). In the general population, the IQR is 55-70%, but for the samples it was only 25-63%.

The 640 normal morphology values were evaluated with a mean of 6.23% (SD 3.32). The 25th, 50th and 75th centiles were 4.00, 6.00 and 8.00% respectively. These values are lower than those of the general population, whose first through third quartile values are 10.5, 14 and 16% respectively. 61 (22%) of the 640 values were below the minimum reference limit of 4.0% (3.0-4.0).

The semen parameter values discussed above as well as total count for motility, normal morphology and abstinence time (in days) are presented in Table 2, including all mean, quartile data and standard deviations. Table 2 also provides the number and percentage of semen samples that fell below the WHO reference limits.
Table 2. Distribution of semen quality parameters and abstinence time among 276 men contributing 682 semen samples in the EARTH Study.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>SD</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>N (%) below WHO 2010 reference limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ejaculate volume (mL)</td>
<td>2.80</td>
<td>1.38</td>
<td>1.70</td>
<td>2.70</td>
<td>3.70</td>
<td>120 (18)</td>
</tr>
<tr>
<td>Sperm concentration (mil/mL)</td>
<td>63.7</td>
<td>60.5</td>
<td>23.2</td>
<td>47.6</td>
<td>88.0</td>
<td>90 (13)</td>
</tr>
<tr>
<td>Total sperm count (mil)</td>
<td>158</td>
<td>143</td>
<td>56.1</td>
<td>116</td>
<td>225</td>
<td>106 (16)</td>
</tr>
<tr>
<td>Total motility (%)</td>
<td>43.8</td>
<td>23.2</td>
<td>25.0</td>
<td>44.0</td>
<td>63.0</td>
<td>300 (44)</td>
</tr>
<tr>
<td>Total motile count (mil/ejaculate)</td>
<td>88.1</td>
<td>102</td>
<td>15.8</td>
<td>49.9</td>
<td>131</td>
<td>-</td>
</tr>
<tr>
<td>Normal morphology (%)</td>
<td>6.23</td>
<td>3.32</td>
<td>4.00</td>
<td>6.00</td>
<td>8.00</td>
<td>95 (15)</td>
</tr>
<tr>
<td>Normal morphology count* (mil/ejaculate)</td>
<td>11.6</td>
<td>13.8</td>
<td>2.65</td>
<td>7.24</td>
<td>16.1</td>
<td>-</td>
</tr>
<tr>
<td>Abstinence Time (Days)</td>
<td>2.64</td>
<td>2.82</td>
<td>1.83</td>
<td>2.20</td>
<td>2.71</td>
<td>-</td>
</tr>
</tbody>
</table>

* A total of 42 semen samples (6%) had missing data for normal sperm morphology and thus total normal morphology count, resulting in N=640 semen samples.
Adjusted Semen Quality Parameters

The mixed process was run to generate GLMM for each semen parameter adjusted for abstinence time, age and BMI. Then GLMM were generated for each semen parameter adjusted for the covariates used in the first model but also including the non-anthropometric covariates that were found to have an association: race, education level and recreational drug use. Of the 682 samples evaluated from 276 men, morphology data (normal morphology count and as a percentage) was only available for 640 of the semen samples. For all other parameters 682 samples were used. To test for linear trends across the WC tertiles the median WC was used as a continuous variable for each tertile in the models.

The first analysis modeled ejaculate (sample) volume as the dependent variable. The mean when adjusted for abstinence time, age and BMI across tertiles lacked significance, (p-trend = 0.68). The adjusted mean volumes increased from T1 to T3 from 2.71 to 2.78 mL, but T2 was much higher than both, with a mean of 2.95 mL. When adjusted to include race, education level and recreational drug use, the mean ejaculate volumes across the three tertiles again lacked significance (p-trend= 0.75). The mean values were identical, and the confidence intervals were only slight different, with the most notable change being T3, 2.45-3.12 mL in the first model narrowing to 2.51-3.01 mL in the second model.

In the first analysis mean sperm concentration adjusted for abstinence time, age and BMI decreased across tertiles; in T1 the concentration was 46.8, T2 37.3 down to 19.4 mil/mL in T3. This change showed significance at the alpha =0.05 level. The p-trend was 0.03, however T3 when compared with T1 was slightly higher, with p=0.05. The second analysis, additionally
adjusting for race education level and recreational drug use showed a similar decrease in mean sperm concentration from 44.6 in T1, 39.6 in T2 to 20.2 mil/mL in T3. The p-trend for the second analysis was also significant (p=0.03), showing an inverse relationship to waist circumference.

Mean total sperm count was inversely related to waist circumference as well. Sperm count decreased across tertiles; T1 was 110 mil, T2, 94.7 mil and T3, 47.9 mil. Confidence interval decreased consistently from 76.0-159 in T1 to 24.8-92.5 in T3. When adding the adjustments for race, education and drug usage, they had slightly closer values (106, 100 and 49.7 mil) from T1 through T3. Both first and second analyses were significant, with p-trend = 0.05.

While total motility (%) slightly decreased from 45.4% at T1 to 42.5% at T2 to 41.5% at T3, it did not rise to the level of significance (p-trend = 0.35) in the first analysis. The second analysis also showed a decrease from T1 to T3 which was slightly greater than the initial analysis (T1=45.7%, T2=42.9%, T3=41.2%), however it was still not sufficient to rise to the level of significance (p-trend = 0.28).

Total motile count also decreased but failed to show a strong correlation. In the first analysis T1-T3 decreased from 49.6 mil/ejaculate to 38.7 for T2 and 34.1 for T3. The p-trend was 0.14, which narrowed to 0.10 when the third tertile was compared to the first. Likewise with the added covariates in the second analysis the motile count decreased from T1-T3 (51.8, 38.4, 33.2 mil/ejaculate), and as in the first, when the third tertile was compared to the first, the p-value was 0.10; in this analysis however the correlation between WC tertile and motility was much stronger with an overall p-trend of 0.07.

Normal morphology (%) was not shown to be correlated to WC tertile. In the first
analysis the patients in the first WC tertile had 6.38% normal semen morphology, in T2 5.90 and in T3 6.32 and the p-trend was 0.99. In the second analysis the values from T1 through T3 were 6.36%, 5.88% and 6.36% respectively and the p-trend was 0.92.

Only 640 samples were analyzed for the two measurements of morphology as 42 of the 682 samples were missing data. As a percentage, normal morphology was not correlated with WC in the first or second analyses. The p-trend for the first analysis was 0.99 and the p-trend of the mean in the second analysis was 0.92. The normal morphology count however had a much stronger correlation.

For T1 mean normal morphology count in the first analysis was 6.98 mil/ejaculate, which decreases to 5.55 in T2 and to 5.20 in T3. This resulted in a 0.08 p-trend of means across tertiles (p=0.10 when comparing T3 to T1).

The adjusted semen parameter and abstinence time results are shown below in Table 3.
Table 3. Semen quality parameters (adjusted mean, 95% CI) by tertiles of waist circumference among 276 men contributing 682 semen samples in the EARTH Study

<table>
<thead>
<tr>
<th>Tertile</th>
<th>Ejaculate volume (mL)</th>
<th>Sperm concentration (mil/mL)</th>
<th>Total sperm count (mil)</th>
<th>Total motility (%)</th>
<th>Total motile count (mil/ejaculate)</th>
<th>Normal morphology (%)</th>
<th>Normal morphology count (mil/ejaculate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 (67.5 - 91.5)</td>
<td>2.71, 2.44-2.98</td>
<td>46.8, 32.9-66.6</td>
<td>110, 76.0-159</td>
<td>45.4, 40.4-50.5</td>
<td>49.6, 36.5-67.3</td>
<td>6.38, 5.62-7.06</td>
<td>6.98, 5.29-9.21</td>
</tr>
<tr>
<td>T2 (92.0 - 101)</td>
<td>2.95, 2.68-3.22</td>
<td>37.3, 26.3-52.8</td>
<td>94.7, 67.7-133</td>
<td>42.5, 37.9-47.1</td>
<td>42.9, 28.5-52.3</td>
<td>5.90, 5.25-6.54</td>
<td>5.55, 4.31-7.14</td>
</tr>
<tr>
<td>T3 (102 - 200)</td>
<td>2.78, 2.45-3.12</td>
<td>19.4, 10.1-37.1*</td>
<td>47.9, 20.2</td>
<td>41.5, 24.8-42.5</td>
<td>34.1, 24.8-42.5</td>
<td>6.32, 5.25-6.54</td>
<td>5.20, 4.31-7.14</td>
</tr>
<tr>
<td>p-trend</td>
<td>0.68, 0.03</td>
<td>0.05</td>
<td>0.35</td>
<td>0.14</td>
<td>0.09</td>
<td>0.08</td>
<td></td>
</tr>
</tbody>
</table>

Adjusted for Abstinence time, Age and BMI

| T1 (67.5 - 91.5) | 2.71, 2.43-2.98        | 44.6, 30.8-64.8             | 106, 71.9-155         | 45.7, 40.5-50.9   | 51.8, 38.0-70.6                    | 6.36, 5.62-7.09       | 7.20, 5.42-9.56                         |
| T2 (92.0 - 101)  | 2.95, 2.67-3.23        | 39.6, 28.4-55.2             | 100, 72.8-139         | 42.9, 38.4        | 38.4, 28.5-51.9                    | 5.88, 5.22-6.53       | 5.52, 4.28-7.12                         |
| T3 (102 - 200)   | 2.78, 2.51-3.01        | 20.2, 11.1-36.8             | 49.7, 26.9-91.9*      | 41.2, 36.8-46.5   | 41.2, 23.5-46.9*                   | 6.36, 5.62-7.08       | 5.08, 3.64-7.10                         |
| p-trend         | 0.75, 0.03             | 0.05                        | 0.28                  | 0.07              | 0.92                               | 0.14                  |                                        |

Adjusted for Abstinence time, Age, BMI, Race, Education Level, and Recreational Drug Use

*p-value <0.05 when compared that tertile with the lowest tertile. ip-value <0.10 when compared that tertile with the lowest tertile

*aA total of 42 semen samples (6%) had missing data for normal sperm morphology and thus total normal morphology count, resulting in N=640 semen samples.
Chapter IV: Discussion

Comments

Since the 1970s, there has been a steady increase in the relative prevalence of obesity and severe obesity in the United States. In 1971, 41.7% of men were overweight and 12.1% were obese. By 2016, the percentage of overweight had decreased to 35.8% but obesity experienced a threefold increase to 38.3%. During the same time frame, severe obesity (BMI ≥ 40) which included 0.6% of the population, increased almost ten-fold to 5.9% in 2016 (Fryar, 2018). During the same timeframe there has been a significant decline in semen quality markers (Levine, 2017; Carlsen, 1992) and it is estimated that 30 million men worldwide are infertile (Agarwal, 2015). With the twin global crises of obesity and male infertility it is increasingly important to understand and identify the links between obesity and male infertility.

In this paper I reviewed data and created analyses and models to measure association of WC with semen parameters as a proxy for male infertility for the cohort of men enrolled in the EARTH study. The models used indicate that of the semen quality parameters examined, WC was negatively associated with sperm concentration (%) and total sperm count. These findings were significant and independent of BMI as well as the related confounders age, abstinence time, education, race and recreational drug use. In addition, while lacking significance at the alpha =0.05 level, it was also determined that there is an association total motility count and normal morphology count.

These results show some consistency with established research on obesity and fertility. In 2010 Hofny et. al examined semen parameters and hormone profiles of oligozoospermic obese men in order
to examine the relationship between obesity and male factor infertility. Of the 122 obese males studied, Hofny determined that obesity (as measured by BMI) had significant positive correlation with abnormal sperm morphology and a significant negative correlation with sperm concentration and motility (Hofny, 2010). The review article by Du Pleiss published the same year, which included Hofny’s findings, determined that across the available studies examining obesity and male infertility, there was a consistent relationship between obesity as measured by BMI and decreased total sperm count, sperm concentration and total motility (Du Pleiss, 2010).

Limitations

While the dataset used provided a rich source of diverse data, and while attempts were made to ensure all analyses were designed to adequately and definitively answer aspects of the research question, there are some inherent limitations in this research paper that should be addressed.

One notable limitation of this paper is that in the EARTH study anthropometric and age data was collected at enrollment, and thus any changes in BMI, age and WC during the duration of the participant’s involvement with the EARTH study were not captured and were therefore not able to be included in this paper’s analyses. This also had an additional minor effect, in that it caused some data to have to be removed, as in the case of the participant whose height was entered as 63 cm. Without comparative data, each outlier data point had to be scrutinized and compared to national averages and a decision had to be made if it was an extreme outlier or an erroneous value.

The sample size of the study population was sufficiently large for most analyses; however the study population was largely homogenous in some measures, notably race and education where white graduate students were noticeably overrepresented. This caused some subgroups to be too small to
effectively evaluate, an issue that should be addressed by future cohort studies.
Conclusions

In this paper I have presented an association between increasing WC and a decrease in semen parameters, namely concentration, total sperm count, total motile count and normal morphology count in men who are part of a couple undergoing ART infertility treatment. The association was demonstrated to be independent of BMI and other confounding variables and the associations between WC and semen parameters were relatively consistent between WC tertiles.

As with previous researchers, my attempt to directly link male fertility with obesity has in some ways only provided partial answers. This establishment of an association does not suggest that central adiposity (the primary measurement of WC) is any more or less important for male fertility than other overall adiposity measures. Furthermore, the association was not consistent across all semen quality parameters, notably ejaculate volume. This suggests a deeper and more complex interaction between adiposity and male fertility. However, it is a first and necessary step that will inform further research.
References


Centers for Disease Control and Prevention, Division of Reproductive Health and American Society for Reproductive Medicine Society for Assisted Reproductive Technology (2006) *Assisted Reproductive Technology Success Rates, National Summary of Fertility Clinic Reports*.


Appendix 1: Primary SAS Code (without Macros)
libname DAT 'C:\Users\djb895\Desktop';
#include "c:\users\djb895\Desktop\Ruth EQ Formats.txt";
#include "c:\users\djb895\Desktop\Ruth male EQ Formats.txt'';
#include "c:\users\djb895\Desktop\Ruth moreEQ Formats.txt";

data allobservations;
set data.wc sq 050320;
if age_m= then age_m=36;
IF MFQ_MARIJUANAEVER=1 OR MFQ_MARIJCURR=1 OR MFQ_COCAINEVER=1 OR MFQ_COCAINCURR=1 THEN rec_drug=1;
ELSE rec_drug=0;

IF SAMPLEVOLUME >= 1.5 THEN badSAMPLEVOLUME=0;
ELSE badSAMPLEVOLUME=1;
IF SATOTCONC >= 15 THEN badSATOTCONC=0;
ELSE badSATOTCONC=1;
IF count >= 39 THEN bad_count=0;
ELSE bad_count=1;
IF SAMOTPC >= 40 THEN badSAMOTPC=0;
ELSE badSAMOTPC=1;
IF lnnormorrcount >= 1.386 THEN badlnnormorrcount=0;
ELSE badlnnormorrcount=1;
IF badSAMPLEVOLUME=1 or badSATOTCONC=1 or bad_count=1 or badSAMOTPC=1
or badlnnormorrcount=1 then mfi=1;
ELSE mfi=0;

IF bmi_m > 30 then MaleObesity=1;
Else MaleObesity=0;

run;*682 semen;
proc freq; tables rec_drug;run;

proc sort data=allobservations; by id dateproduced; run;
data numbermen;
set allobservations;
by id dateproduced;
if first.id;
run;*276men;

;*/-------------------------------- TABLE 1--------------------------------*/
**total;
proc means data=numbermen n median Q1 Q3 min max maxdec=2;
var age_m bmi_m mbq_heightcm_m totalFamin ;
run;
proc freq data=numbermen; tables races m/chisq;run;
proc freq data=numbermen; tables obese/chisq;run;
proc freq data=numbermen; tables smokstat m/chisq;run;
proc freq data=numbermen; tables educ_m/chisq;run;
** **

**tertiles;**

PROC MEANS DATA=numbermen N median Q1 Q3 maxdec=1;
CLASS rk3_mbq_waistcirc_m;
VAR Age_m bmi_m mbq_heightcm_m totalPAm
;
RUN;

PROC NPARLWAY DATA = numbermen Wilcoxon; CLASS rk3_mbq_waistcirc_m; VAR
Age_m bmi_m run;
PROC NPARLWAY DATA = numbermen Wilcoxon; CLASS rk3_mbq_waistcirc_m; VAR
Bmi_m run;
PROC NPARLWAY DATA = numbermen Wilcoxon; CLASS mbq_heightcm_m; VAR bmi_m;
RUN;
PROC NPARLWAY DATA = numbermen Wilcoxon; CLASS totalPAm; VAR bmi_m; run;
PROC FREQ DATA=numbermen; TABLES rk3_mbq_waistcirc_m*white/chisq; run;
PROC FREQ DATA=numbermen; TABLES rk3_mbq_waistcirc_m*obese/chisq; run;
PROC FREQ DATA=numbermen; TABLES rk3_mbq_waistcirc_m*smokstat_m/chisq;run;
PROC FREQ DATA=numbermen; TABLES rk3_mbq_waistcirc_m*educ_m/exact; run;
PROC FREQ DATA=numbermen; TABLES rk3_mbq_waistcirc_m*infdk1_m/exact; run;
PROC FREQ DATA=numbermen; TABLES rk3_mbq_waistcirc_m*any_sti/exact; run;
PROC FREQ DATA=numbermen; TABLES rk3_mbq_waistcirc_m*mbq_varicocele/exact; run;

//------------------------------------------------------------- TABLE 2--------------------------*/

PROC MEANS DATA=allobservations n mean std p25 p50 p75 maxdec=2; var
samplevolume
satotconc
count
smotpc totalmotcount
mornorm totalnormmorcount
abstinencehrs
;
RUN;

DATA volume; SET allobservations; IF samplevolume<1.5; run;
DATA concentration; SET allobservations; IF satotconc<15; run;
DATA count; SET allobservations; IF count<39; run;
DATA motility; SET allobservations; IF smotpc<40; run;
DATA morphology; SET allobservations; IF mornorm<4 and mornorm ne .; run;

//------------------------------------------------------------- TABLE 3--------------------------*/

PROC MIXED DATA=allobservations empirical method=ml;
CLASS id semencounter rk3_mbq_waistcirc_m (ref='1');
MODEL samplevolume= rk3_mbq_waistcirc_m lastejacdays age_m bmi_m / RESIDUAL
SOLUTION ALPHA=0.05;
RANDOM intercept/subject=ID; repeated /TYPE=UN r rcorr;
LSMEANS rk3_mbq_waistcirc_m /CL diff;
ODS OUTPUT LSMEANS = myout2;
RUN;
PROC MIXED DATA=allobservations empirical method=ml;
class ID semencounter;
model samplevolume= rk3_mbq_waistcirc_m lastejacdays age_m bmi_m / RESIDUAL
  SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
run;

proc mixed data=allobservations empirical method=ml;
class ID semencounter rk3_mbq_waistcirc_m (ref='1');
model lnconc= rk3_mbq_waistcirc_m lastejacdays age_m bmi_m / RESIDUAL
  SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans rk3_mbq_waistcirc_m /cl diff;
ods output LSMeans = myout2;
run;
data myout2_exp_fds;
retain Estimate Lower Upper; set myout2;
expest=exp(estimate); est=ROUND(expest,0.1); explow=exp(lower);
lower=ROUND(explow,0.1); expup=exp(upper); upper=ROUND(expun,0.1);
run;
proc print data=myout2_exp_fds; var est lower upper; run;
proc mixed data=allobservations empirical method=ml;
class ID semencounter;
model lnconc= rk3_mbq_waistcirc_m lastejacdays age_m bmi_m / RESIDUAL
  SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
run;
data myout2_exp_fds;
retain Estimate Lower Upper; set myout2;
expest=exp(estimate); est=ROUND(expest,0.1); explow=exp(lower);
lower=ROUND(explow,0.1); expup=exp(upper); upper=ROUND(expun,0.1);
run;
proc print data=myout2_exp_fds; var est lower upper; run;
proc mixed data=allobservations empirical method=ml;
class ID semencounter;
model lnconc= rk3_mbq_waistcirc_m lastejacdays age_m bmi_m / RESIDUAL
  SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
run;
proc mixed data=allobservations empirical method=ml;
class ID semencounter rk3_mbq_waistcirc_m (ref='1');
model sampctpc= rk3_mbq_waistcirc_m lastejacdays age_m bmi_m / RESIDUAL
  SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans rk3_mbq_waistcirc_m /cl diff;
run;
proc mixed data=allobservations empirical method=ml;
class ID semencounter;
model samctpc= r1 mmbq Waistcirtc m lastejacdays age_m bmi_m / RESIDUAL
SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
run;

proc mixed data=allobservations empirical method=ml;
where lnmount ne .;
class ID semencounter r1 mmbq Waistcirtc m (ref='1');
model lnmount= r1 mmbq Waistcirtc m lastejacdays age_m bmi_m / RESIDUAL
SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans r1 mmbq Waistcirtc m /cl diff;
ods output LSMeans = myout2;
run;

data myout2_exp_fds;
retain Estimate Lower Upper; set myout2;
exp=exp(Estimate); est=ROUND(exp,.1); explow=exp(Lower);
lower=ROUND(explow,.1); expup=exp(Upper); upper=ROUND(expup,.1); run;
proc print data=myout2 exp fds; var est lower upper; run;
proc mixed data=allobservations empirical method=ml;
where lnmount ne .;
class ID semencounter;
model lnmount= r1 mmbq Waistcirtc m lastejacdays age_m bmi_m / RESIDUAL
SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
run;

proc mixed data=allobservations empirical method=ml;
class ID semencounter r1 mmbq Waistcirtc m (ref='1');
where mnorm ne .;
model mnorm= r1 mmbq Waistcirtc m lastejacdays age_m bmi_m / RESIDUAL
SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans r1 mmbq Waistcirtc m /cl diff;
run;

proc mixed data=allobservations empirical method=ml;
class ID semencounter r1 mmbq Waistcirtc m (ref='1');
where lnmorfcount ne .;
model lnmorfcount= r1 mmbq Waistcirtc m lastejacdays age_m bmi_m / RESIDUAL
SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans r1 mmbq Waistcirtc m /cl diff;
ods output LSMeans = myout2;
run;

data myout2_exp_fds;
retain Estimate Lower Upper; set myout2;
expest=exp(Estimate); est=ROUND(expest,0.01); expholow=exp(Lower);
lower=ROUND(expholow,0.01); expup=exp(Upper); upper=ROUND(expup,0.01); run;
proc print data=myout2 exp fds; var est lower upper; run;
proc Mixed data=allobservations empirical method=m1;
    where lnncmorforcount ne .;
class ID semencounter;
model lnncmorforcount= rk3 mbg waistcirc m lastedjacdays age m bmi m / RESIDUAL
    SOLUTION ALPHA=0.05;
    random intercept/subject-ID; repeated /TYPE=UN r rcorr;
run;

/*---------------------------------------- TABLE 3 model 2----------------------------------------*/
proc Mixed data=allobservations empirical method=m1;
class ID semencounter rk3 mbg waistcirc m (ref='1');
model samplevolume= rk3 mbg waistcirc m lastedjacdays age_m bmi_m white ed3
    rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
    random intercept/subject-ID; repeated /TYPE=UN r rcorr;
    lsmeans rk3 mbg waistcirc m /cl diff;
    ods output lsmeans = myout2;
run;
proc Mixed data=allobservations empirical method=m1;
class ID semencounter;
model samplevolume= rk3 mbg waistcirc m lastedjacdays age_m bmi_m white ed3
    rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
    random intercept/subject-ID; repeated /TYPE=UN r rcorr;
run;
proc Mixed data=allobservations empirical method=m1;
class ID semencounter rk3 mbg waistcirc m (ref='1');
model lnncconc= rk3 mbg waistcirc m lastedjacdays age_m bmi_m white ed3
    rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
    random intercept/subject-ID; repeated /TYPE=UN r rcorr;
    lsmeans rk3 mbg waistcirc m /cl diff;
    ods output lsmeans = myout2;
run;
data myout2Exp_fds;
retain Estimate Lower Upper; set myout2;
expest=exp(Estimate); est=ROUND(expest,0.1); expholow=exp(Lower);
lower=ROUND(expholow,0.1); expup=exp(Upper); upper=ROUND(expup,0.1); run;
proc print data=myout2 exp fds; var est lower upper; run;
proc Mixed data=allobservations empirical method=m1;
class ID semencounter;
model lnncconc= rk3 mbg waistcirc m lastedjacdays age_m bmi_m white ed3
    rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
    random intercept/subject-ID; repeated /TYPE=UN r rcorr;
run;
proc Mixed data=allobservations empirical method=m1;
class ID semencounter rk3 mbg waistcirc m (ref='1');
model lnncornt= rk3 mbg waistcirc m lastedjacdays age m bmi m white ed3
    rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans rk3_mbq_waistcirc_m /cl diff;
ods output LSMeans = myout2;
run;
data myout2 exp_fds;
retain Estimate Lower Upper; set myout2;
exp_est=exp(Estimate); est=ROUND(exp_est,0.1); explow=exp(Lower);
lower=ROUND(explow,0.1); expup=exp(Upper); upper=ROUND(expup,0.1); run;
proc print data=myout2 exp fds; var est lower upper; run;
proc mixed data=allobservations empirical method=ml;
class ID semencounter;
model lncount= rk3_mbq_waistcirc_m lastejacdays age_m bmi_m white ed3 rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
run;
proc mixed data=allobservations empirical method=ml;
class ID semencounter;
model samctpc= rk3_mbq_waistcirc_m lastejacdays age_m bmi_m white ed3 rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans rk3_mbq_waistcirc_m /cl diff;
run;
proc mixed data=allobservations empirical method=ml;
class ID semencounter;
model lnmcount= rk3_mbq_waistcirc_m lastejacdays age_m bmi_m white ed3 rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans rk3_mbq_waistcirc_m /cl diff;
ods output LSMeans = myout2;
run;
data myout2 exp_fds;
retain Estimate Lower Upper; set myout2;
exp_est=exp(Estimate); est=ROUND(exp_est,0.1); explow=exp(Lower);
lower=ROUND(explow,0.1); expup=exp(Upper); upper=ROUND(expup,0.1); run;
proc print data=myout2_exp_fds; var est lower upper; run;
proc mixed data=allobservations empirical method=ml;
where lnmcount ne .;
class ID semencounter rk3_mbq_waistcirc_m (ref='1');
model lnmcount= rk3_mbq_waistcirc_m lastejacdays age_m bmi_m white ed3 rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans rk3_mbq_waistcirc_m /cl diff;
ods output LSMeans = myout2;
run;
where mornorm ne .;
model mornorm= rk3 mbq waistcirc m lastejacdays age_m bmi_m white ed3
rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans rk3 mbq_waistcirc m /cl diff;
run;

proc mixed data=allobservations empirical method=ml;
  where mornorm ne .;
class ID semencounter ;
model mornorm= rk3 mbq waistcirc m lastejacdays age_m bmi_m white ed3
  rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
run;

proc mixed data=allobservations empirical method=ml;
class ID semencounter rk3 mbq_waistcirc m (ref='1');
where llnormnormcount ne .;
model llnormnormcount= rk3 mbq waistcirc m lastejacdays age_m bmi_m white ed3
  rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
lsmeans rk3 mbq waistcirc m /cl diff;
ods output LSMeans = myout2;
run;
retain Estimate Lower Upper; set myout2;
expest=exp(Estimate); est=ROUND(expest,0.01); explow=exp(Lower);
lower=ROUND(explow,0.01); expup=exp(Upper); upper=ROUND(expup,0.01); run;
proc print data=myout2_exp_fds; var est lower upper; run;
proc mixed data=allobservations empirical method=ml;
class ID semencounter ;
model llnormnormcount= rk3 mbq waistcirc m lastejacdays age_m bmi_m white ed3
  rec_drug / RESIDUAL SOLUTION ALPHA=0.05;
random intercept/subject=ID; repeated /TYPE=UN r rcorr;
run;
Appendix 2: SAS Code Generating the Any_STI variable

```sas
any_sti = 0;

IF mfq_chlamydia=1 OR mfq_trichomonas=1 OR mfq_herpes=1 OR mfq_hp=1
OR mfq_syphilis=1 OR mfq_gonorrhea=1 OR mfq_lgv=1 OR mfq_mycoplasma=1
OR mfq_strep=1 OR mfq_othercomdis=1 THEN any_sti=1;
```