2018

Cost-effectiveness Analysis of Preimplantation Genetic Screening

William Andrew Moye
Walden University

Walden University

College of Health Sciences

This is to certify that the doctoral dissertation by

William Moye

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Michael Furukawa, Committee Chairperson, Health Services Faculty
Dr. Lawrence Fulton, Committee Member, Health Services Faculty
Dr. Gudeta Fufaa, University Reviewer, Health Services Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2018

Abstract

Cost-effectiveness Analysis of Preimplantation Genetic Screening

by

William Andrew Moye

MBA, University of Florida, 2007

BS, University of Arizona, 1997

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Health Sciences

Walden University

February 2018

Abstract

In vitro fertilization (IVF) is used to help infertile couples achieve a live birth. Clinical studies have suggested that multiple, consecutive cycles of IVF can increase live birth rate significantly. Others have documented improved live birth rates from the use of new laboratory techniques for preimplantation genetic screening (PGS). This genetic screening technique seeks to determine the ploidy of the embryo prior to implantation into the woman. To date, no study has examined the cost-effectiveness of using IVF in conjunction with PGS compared to that of IVF alone for 3 consecutive cycles in achieving a live birth. This study compared the incremental cost-effectiveness ratios (ICER) from each intervention arm based on the clinical probabilities for each outcome and this study was grounded in the protection motivation theory. Costs were obtained from secondary sources, such as the literature and government databases. The model was constructed using a decision-analytical approach that allowed for *z* test statistical analysis of the outcomes, where the ICER is the dependent variable and the independent variables are the 2 interventions. The robustness of the model was tested through univariate and probabilistic sensitivity analysis and stratified by age groups. The results showed that PGS with IVF was cost-effective for women aged under 40 and women aged 40-42, but not for women over 42. Based on a willingness-to-pay threshold of $100,000, IVF with PGS was the most cost-effective strategy in all age groups. The positive social change implication of this study is such that understanding the costs associated with a new technology to achieve a live birth is significant for society to help guide clinical treatment of these patients.

Cost-effectiveness Analysis of Preimplantation Genetic Screening

by

William Andrew Moye

MBA, University of Florida, 2007

BS, University of Arizona, 1997

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Health Sciences

Walden University

February 2018

Dedication

To my beautiful family: Leanne, Devin and Zenia. Without their love and support I could not go pursue all the things I want to do–like get my PhD!! I love you!

Acknowledgments

Thanks to Dr. Mike Furukawa for his guidance during this process, I certainly appreciate it. Also many thanks to Dr. Lawrence Fulton for his help with the cost-effectiveness models. Thanks to everyone at Walden who supported me through this process.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

In vitro fertilization (IVF) is an assisted reproductive technology (ART) technique used by reproductive endocrinologists to help infertile couples conceive (Society for Assisted Reproductive Technologies, 2014). However, for most Americans, IVF is extremely expensive, and can cost on average $12,500 per IVF cycle, not including drug costs (American Society for Reproductive Medicine, 2014). The 2012 national average live birth rate (Singleton or more) per embryo transfer in the United States is 46.9% for women aged under 35, 37.8% for women aged 35 to 37, 28.4% for women aged 38 to 40, and less than 16.1% for women 41 and older (National Center for Chronic Disease Prevention and Health Promotion, 2014). Cumulatively, live birth rates increase with each IVF attempt, but each cycle is costly and success can vary depending on diagnosis of infertility. Because of costs and other factors, however, drop-out rates before the third cycle of IVF have been reported between 17% and 65% (Luke et al., 2012).

Preimplantation genetic screening (PGS) is a genetic technique used to analyze the complement of chromosomes from an individual embryo and determine if it is normal (*euploid*) or abnormal (*aneuploid*). This type of screening is now typically done using array comparative genomic hybridization (aCGH) and has been shown to greatly increase pregnancy rates and live birth rates, depending on the age of the participant (Hodes-Wertz, Grifo, Ghadir, Kaplan, & Laskin, 2012; Yang et al., 2012). The use of PGS in clinical care is controversial, and no clear recent consensus is available for patients or clinicians on when and where to use PGS in conjunction with IVF (Brezina & Kutteh, 2015).

Background of Problem

PGS is a tool that reproductive endocrinologists have been employing in recent years to select the best embryos for implantation during IVF cycles. PGS can be performed in a variety of ways, including Array Comparative Genomic Hybridization (aCGH) (Handyside, 2013). PGS involves looking for aneuploidy, or an abnormal number of chromosomes in the embryo, prior to implantation in the IVF cycle. The aim of PGS is to decrease miscarriage rates and increase live birth rates (Harper et al., 2012).

Microarray, or aCGH, is a tool that has been used to determine the etiology of products of conception for some time (Ford & Schust, 2009). Related to that, it has been established that roughly 50% of all miscarriages from women with recurrent pregnancy loss (RPL) are due to a chromosomal abnormality as detected by aCGH (Ford & Schust, 2009; Foyouzi, Cedars, & Huddlest, 2012). Researchers have speculated about the use of aCGH as a method to detect these chromosomal abnormalities in embryos prior to implantation in an IVF cycle in order to improve pregnancy rates and decrease miscarriage rates (Harper et al., 2012). Next generation sequencing (NGS) is a relatively new technique to perform ploidy analysis as well, but researchers have shown it to be just as effective at screening for euploid embryos as aCGH (Yang et al., 2015). Together, aCGH and NGS technologies represent PGS 2.0 or PGS #2, distinguishing them from prior techniques (Gleicher, Kushnir, & Barad, 2014). However, guidelines from professional societies that govern the use of PGS have stated that no evidence exists that the use of PGS increases the live-birth rate for women with RPL (American Society for Reproductive Medicine, 2008; Harper & SenGupta, 2012).

Researchers have performed multiple randomized control trials (RCTs) to understand the connection between PGS and implantation rates. Three studies underscore the current understanding of the use of PGS by CGH and implantation rates and miscarriage rates. The first, performed by Yang et al. (2012), was a pilot study of young, good prognosis patients. The results were that the pregnancy rate of the group that underwent PGS by CGH achieved a 69.1% pregnancy rate, compared to the control group at 41.7% (Yang et al., 2012). However, this study did not specifically look at women with RPL, nor did it look at live-birth rates.

A more recent study by Keltz (2013) looked at a number of patients that had PGS by CGH and compared them to a control group without PGS. This study looked at implantation rates, ongoing pregnancy rates, miscarriage rates, and multiple rates. The results were that the PGS group had implantation rates double that of the control group, and significantly decreased miscarriage and multiple rates (Keltz, 2013). However, this study did not focus specifically on women with RPL nor did it observe live-birth rates.

Lastly, the most recent analysis of CGH in PGS does look at RPL. Hodes-Wertz et al. (2012), looked at 2,282 embryos from couples with RPL and analyzed them through CGH. 35% were euploid (normal) and 60% were aneuploid. In 181 transfer cycles, the miscarriage rate was 6.9% compared to 33.5% in an RPL control population and 23.7% in a general infertile population. The researchers demonstrated a significant benefit to PGS in RPL couple for miscarriages, but did not look at live-birth rates (Hodes-Wertz et al., 2012).

Other methods for determining the best method for achieving a live birth include expectant management, or simply allowing a couple women with RPL to have a natural birth without IVF. Limited studies in determining the cost-effectiveness of PGS have been performed. Murugappan, Ohno, & Lathi (2015) examined the current literature to determine if a cycle of IVF with PGS would be more cost-effective than expectant management. The authors concluded that although IVF with PGS resulted in a lower clinical miscarriage rate than expectant management, it was a costlier technique with fewer clinical pregnancies.

Other researchers have pointed to cumulative IVF pregnancy rates as a means to increase live birth rate. In a study on cumulative birth rates, Luke et al. (2012) described the pregnancy rate for women who proceed to three cycles of IVF. The authors found that if a women proceeds to at least three cycles of IVF, the live-birth rate can be similar to natural fecundity rates for women. Other researchers have confirmed these findings and suggested that if women can proceed to three or more cycles, then a high (> 70%) live birth rate can be achieved (Smith, Tilling, Nelson, & Lawlor, 2015).

Cost-effectiveness analysis (CEA) for genetic tests and their use in healthcare have been described (Garber & Phelps, 1997; Grosse, Wordsworth, & Payne, 2008). Many authors have described genetic testing in the context of cost-utility and have determined that continued studies to understand the efficacy of genetic tests as they relate to economics must be performed, despite growing evidence that genetic tests provide improved outcomes and better health (Phillips, Sakowski, Trosman, Douglas, Liang, & Neumann, 2014).

Considering CEAs specifically of PGS and IVF, the results have been mixed concerning the use of PGS. Multiple researchers found no benefit to the use of PGS when compared with some other intervention. In a study examining the cost-effectiveness of two interventions, IVF combined with PGS and expectant management for patients with recurrent pregnancy loss, researchers found that expectant management was a far more cost-effective strategy that IVF in conjunction with PGS for patients with RPL, regardless of age group (Murugappan et al., 2015). Further, Salem, Ho, Bendikson, Chung, & Paulson (2016) concluded that PGS in a fresh transfer was not cost-effective compared with IVF alone in patients older than 39. The analysis was made using previous data concerning false positive rates in PGS (Salem et al., 2016). Other researchers looked at similar retrospective data and determined that PGS with IVF was not a cost-effective strategy compared with IVF alone (Kushnir, Darmon, Albertini, Barad, & Gleicher, 2016). However, in recent studies, researchers have looked at retrospective data to determine the cost per live birth with IVF and PGS compared to IVF with fresh transfers and subsequent frozen transfers in different cohorts of women. Hodes-Wertz, McCulloh, and Grifo (2015) conducted a recent single-center study by using direct costs for IVF and PGS and found that PGS was cost-effective in all age groups but women over 42. Adding to that is the conclusion of Resetkova, Tobler, Kearns, & Werner (2013) who found all age groups saw cost-effectiveness benefits from the use of PGS.

These contrary findings from recent CEA present a gap in the understanding of this technology in routine clinical practice. Further, no studies have been performed

comparing IVF in conjunction with PGS in one cycle compared to three consecutive cycles. Comparing the results of past studies for live birth rates in cumulative IVF cycles with live-birth studies from IVF combined with PGS will result in two end data points of the arms of a decision model and contribute to the current knowledge base of this nascent technology.

Problem Statement

There have been few health economic analyses for the use of PGS, and previous analyses were rooted in the technology and costs of the time, both of which have changed dramatically over the past few years (Gleicher et al., 2014). To date, no CEA from the perspective of the payer (in most IVF cases, the patient) have been done to determine if PGS should be used during the first cycle of IVF, versus going on to have three consecutive cycles of IVF without PGS.

The problem that I sought to solve with this research was to provide insight into the cost effectiveness of PGS testing in first cycle of IVF versus continued IVF cycles using a cost-effectiveness model for decision analytics. I examined the potential outcomes using previously published data about live birth rates, PGS live birth rates, and costs based on public information and private clinic data.

Purpose of the Study

The purpose of the study was to understand the cost-effectiveness of genetic screening in IVF and determine its utility compared to multiple cycles of IVF using secondary sources of data. For patients, IVF is expensive and is covered by health insurance in only a few states (Society for Assisted Reproductive Technologies, 2014).

For those reasons, it is important to know what interventions provide the best clinical outcomes at the lowest costs. If the intervention of a genetic test adds cost in the short-run, but saves costs in the long-run because of the clinical information it provides, then that intervention needs to be considered as a first-line therapy (Grosse et al., 2008). Because the health benefit is achieving a live birth, it is a more set endpoint than using a quality-adjusted life year (QALY), which is subjective.

The importance of providing the best care while using the least resources is a critical part of reducing health-care expenditures in the United States while increasing quality of care. Through CEA of the two interventions, I aim to shed light on which technique might be preferred in the clinic to help couples achieve a live birth (Garber & Phelps, 1997).

Research Question and Hypotheses

Research Question

RQ1-Quantitative: What is the cost-effectiveness of IVF paired with PGS in one cycle compared to three consecutive cycles of IVF in achieving a live birth?

RQ2-Quantitative: What are the statistical differences in means between cost-effectiveness among women aged under 40, women aged 40-42, and women aged over 40 of women in each treatment arm?

Hypotheses

Null Hypothesis (H_0): There was no statistically significant difference in the cost-effectiveness of proceeding to three cycles of IVF without PGS in achieving a live birth than one cycle of IVF with PGS.

Alternative Hypothesis (H_a): There was a statistically significant difference in cost-effectiveness in utilizing one cycle of IVF with PGS in achieving a live birth than three cycles of IVF.

Variables

The independent variable in this research problem was the method of treatment (Mt), either one cycle of IVF with PGS (IVF-PGS) or three cycles of IVF (3xIVF).

The dependent variable was cost, or cost-effectiveness as determined via cost-effectiveness ratio of costs/live birth.

Theoretical Framework

The theoretical framework for this study was Rogers (1975) protection motivation theory (PMT) as it applies to genetic screening. PMT describes factors that motivate people to do certain things when it concerns their health. It focuses on cognitive processes and has three main components, including how the perception of severity of a health threat, the person's vulnerability to that threat, and the efficacy of the proposed protective mechanism. In this study, the health threat is pregnancy loss combined with economic threat of overpaying for an alternate intervention. Other researchers have described he literature describes the vulnerability, and the efficacy of the treatment arm is the clinical outcome. PMT was used as a model for genetic screening, including BRCA testing (Helmes, 2002).

Nature of the Study

The research design was a cross-sectional study based on economic modelling, using secondary data from the current literature and public data sets to determine

incremental cost effectiveness ratios (ICER). I utilized CEA and determined the results based on decision analytic models of two interventions to achieve a similar outcome. CEA is an economic tool used to compare costs and health outcomes of two interventions. Utilizing decision analytic models, I compared two interventions to aid in decisions about treatment course. Statistical analysis included sensitivity analyses to determine what key parameters provided the best outcome utilizing the least resources (Noyes & Holloway, 2004). Utilizing a cost-effectiveness decision model, the costs associated with each intervention—either IVF-PGS or three cycles of IVF (3xIVF)—were compared.

Many CEA models utilize health benefit endpoints that may be arbitrary, like QALY. Such endpoints are subject to interpretation about the costs associated with those benefits (Weinstein, Siegel, Gold, Kamlet, & Russell, 1997). Because the costs associated with getting to a live birth are known with some certainty, CEA should provide the best method for understanding the decisions needed for that treatment.

Definitions

Multiple clinical terms and their acronyms were used throughout the study and are defined in detail in Chapters 2 and 3. A comprehensive list of all acronyms used in this study can be found in Appendix A. Definitions of the most common terms are here:

Assisted reproductive technology (ART): Any technology used to assist pregnancy. It typically involves the manual handling of sperm or eggs, or both.

Clinical miscarriage: Any pregnancy loss less than 20 weeks gestation.

Cost-effectiveness analysis (CEA): A type of economic analysis that assumes that one wants to optimize some outcome given some resource restraint. Typically, this analysis results in a cost-effectiveness ratio, where the numerator is the net expenditure of health resources (costs) and denominator is the net improvement in health, or other health-related outcome, in this case, one live birth (Weinstein et al., 1997).

Cycle or IVF cycle: Any number of embryo transfers from a single given oocyte retrieval event.

In vitro fertilization (IVF): A form of ART where oocytes are retrieved surgically and then fertilized in vitro. The fertilized embryo(s) are then implanted in to the women to facilitate pregnancy.

Live birth: A live birth is defined as any pregnancy that terminates in a live birth after 24 weeks of gestation, where the child survived at least 30 days out of utero.

Preimplantation genetic diagnosis (PGD): A type of genetic testing used to look for a specific disease in a fertilized embryo, such as Tay Sachs or Huntington's Disease.

Preimplantation genetic screening (PGS): A type of genetic screening used to determine if a fertilized embryo has the correct complement of chromosomes.

Assumptions

As the analysis is done on a theoretical model, many of the assumptions were made based on the design of the model itself. I assumed certain specific states for each intervention and the outcomes of those interventions, including technical assumptions that defined each cycle and the number of embryos transferred. Further, the model assumed only three outcomes for any cycle: live birth, clinical miscarriage, or no

pregnancy. I also assumed that no couples dropped out of IVF treatment between cycles, and I used clinical probabilities from the literature to account for that assumption. No other outcomes were considered. These methodological assumptions are detailed in Chapter 3.

Scope and Delimitations

The scope of the study was limited to those patients who participated in the studies of the original datasets, and I looked only at the age of the participants as a factor in any differences between interventions. Although the number of patients in each intervention arm is significant, each clinical case is different, so each outcome is not exactly the same. This study's scope included three consecutive cycles of IVF or one cycle of IVF with PGS. These are not the only options available to patients wishing to conceive through ART.

There were multiple threats to validity. They included protocol-driven resource use, comparators that are uncommon or not recommended in clinical practice, nonrepresentative recruiting, restrictive inclusion and exclusion criteria, and artificially enhanced compliance (Ramsey et al., 2015). These threats were mitigated by using secondary data that were not protocol specific and ensuring that each intervention arm was one used in clinical practice.

Concerning specific delimitations of intrinsic limitations of the data or primary studies, I made every effort to normalize much of the data by using multiple sources and not relying on only one source for all data in the model. Further, statistical tests were used to determine the confidence intervals of each cost-effectiveness ratio (CER), as well

as the sensitivity of any one or multiple variables in the model. These statistical tests are recommended as best practice for modelling of CEAs (Briggs et al., 2012).

Limitations

As in any CEA, the main limitation of the model was the input of the model itself (Ramsey et al., 2005). Because clinical probability data for each intervention was drawn from both the literature and from public databases, there were intrinsic limitations about the design of the studies and the accuracy of the data reported by the CDC. Further, any limitations of the original studies used to populate the data in this model also existed in this study.

Significance of the Study

The significance of the study was in providing empirical data to support how a patient might go about getting pregnant through the use of IVF. As healthcare costs continue to rise, it is important to know if genetic screening tests can determine cost-effectiveness for certain patients in order to achieve certain health outcomes (Grosse et al., 2008). Understanding the different costs associated with multiple interventions to achieve a live birth is significant for society and also helps clinicians and the healthcare industry understand how new technologies benefit society. No studies on this have been done to date, so the research was novel.

Summary

IVF is a relatively expensive procedure and most patients in the United States must pay it out-of-pocket. Genetic screening for embryos has been shown to be an effective way to increase live birth rates and improve the effectiveness of IVF (Keltz et

al., 2013; Yang et al., 2015). Further, studies have shown that undergoing multiple, consecutive cycles of IVF also results in a high live birth rate (Smith et al., 2015). There have been multiple CEAs comparing the use of PGS in clinical practice compared to other interventions. The conclusions from those studies have been mixed, with some authors stating that PGS is not cost-effective, some saying it is cost-effective for only some age groups, and others stating that it is cost-effective for all age groups (Hodes-Wertz et al., 2015; Mersereau, Plunkett, & Cedars, 2008; Resetkova et al., 2013). Because there is no clear consensus between industry experts and clinicians, this presents a significant gap in the understanding of genetic screening in the use of routine clinical practice. To date, no cost-effectiveness studies have been performed that compare the cost-effectiveness of having IVF combined with PGS in one cycle, compared to three consecutive cycles of IVF.

The methodology I used is a decision-analytic model comparing two interventions using secondary sources of data. The two interventions were IVF paired with PGS in one cycle and IVF without PGS for three consecutive cycles. The study was a CEA, and I examined the cost-effectiveness ratios of each intervention as costs per live birth. The ratios were compared to provide an incremental cost-effectiveness ratio and statistical methods including sensitivity analyses were used to determine the robustness of the model. The model assumes certain clinical states for each intervention and was limited by the intrinsic limitations of the original data.

Chapter 2 provides details of the current base of knowledge in the field of IVF and genetic screening. It will also include discussion of the theoretical foundation of the

protection motivation theory. Lastly, I will detail the current literature in cost-effectiveness for genetic screening in general and in conjunction with PGS and IVF procedures and techniques.

Chapter 2: Literature Review

Introduction

IVF is a type of ART used to facilitate pregnancy in a woman and eventually lead to the clinical outcome of a healthy, live birth (Van Voorhis, 2007). The clinical effectiveness of IVF as a technique to achieve pregnancies and live births is well documented, and clinics that perform IVF are federally required to report their outcomes to the Centers for Disease Control (CDC) and voluntarily to the Society for Assisted Reproductive Technology (SART; Sunderam et al., 2015). In 2014, the most recent year for which data are available, there were 208,604 cycles of IVF performed in the United States, with a combined national average of 30% of cycles leading to a live birth for women aged 35-37, 19.5% for women aged 38-40, and 9.7% for women aged 41-42 (CDC, 2016).

Other data has suggested that cumulative IVF pregnancy rates increase as a woman goes through more than one cycle of IVF. In a study on cumulative birth rates, Luke et al. (2012) found that if a women proceeds to at least three cycles of IVF, the live-birth rate can be similar to natural fecundity numbers for women. Other studies have confirmed these findings and suggested that if women can proceed to three or more cycles, then a high (> 70%) live birth rate can be achieved (Smith et al., 2015).

PGS, also known as comprehensive chromosomal screening (CCS), is a genetic technique used to analyze the complement of chromosomes from an individual embryo and determine if it is normal (euploid) or abnormal (aneuploid). This type of screening is now typically done using (aCGH) or NGS and has been shown to greatly increase

pregnancy rates and live birth rates, depending on the age of the population studied (Hodes-Wertz et al., 2012; Yang et al., 2012). Recent studies have continued to highlight higher live birth rates by using some form of PGS versus just IVF without genetic screening (Feichtinger et al., 2015). The use of PGS in clinical care is controversial, and no clear recent consensus is available for patients or clinicians on when and where to use PGS in conjunction with IVF (Brezina & Kutteh, 2015).

CEAs are economic models based on data that provide economic endpoints per some health benefit, also known as a cost utility ratio, or ICER. These ratios usually have a quality adjusted life year (QALY) as the denominator, but many CEAs in fertility studies use pregnancy rate or live birth as the health benefit (Grosse, Wordsworth, & Payne, 2008). A more concrete outcome measure may be easier to effectively quantify, like live births.

CEA in IVF is usually performed from the perspective of the payer, which is almost always the patient. A few cost-effectiveness studies have been performed to examine different methods for IVF, live births, and outcomes, as well as to utilize decision models to understand the differences in interventions (Crawford et al., 2016; Fiddelers et al., 2009).

The purpose of this quantitative study is to assess the cost-effectiveness of two different interventions involved with advanced reproductive technology in achieving a live birth. I will also examine the variation of cost-effectiveness for each intervention by age group. Only a few researchers have looked at cost-effectiveness associated with PGS and IVF, and none have examined cumulative IVF compared with one cycle of IVF with

PGS. In a study on the cost effectiveness of IVF alone compared to IVF paired with PGS in women aged 38-40 and in women over 40, Mersereau et al. (2008) found that for women aged 38-40, IVF without PGS was more cost effective than with PGS, but for women over 40, the cost effectiveness was similar. The study did not take into account the effects of multiple cycles on cost and percentage of having a live birth, nor did it take into account the most recent technological advances associated with PGS, like NGS.

More recently, Murugappan et al. (2015) examined the cost-effectiveness for women with recurrent pregnancy loss with either expectant management or with IVF paired with PGS. The authors concluded that expectant management was far more cost-effective than IVF with PGS. However, Murugappan et al. did not look at multiple cycles of IVF compared to one cycle with PGS. The lack of research in the literature comparing these two typical interventions provides a unique opportunity to understand the cost effectiveness of PGS testing in first cycle of IVF versus continued IVF cycles using a cost-effectiveness model for decision analytics.

Lastly, a recent abstract presented at the American Society of Reproductive Medicine looked at false positive rates for PGS and determined that because of false positives, women over 39 would have a decreased cost effectiveness to use IVF in conjunction with PGS (Salem et al., 2016). The additional factor of false positives to the CEA has not been examined before.

In this literature review, I will describe the literature search strategy I used followed by a review of the theoretical foundation of the study. I will also provide technical background on infertility, assisted reproductive techniques, in vitro fertilization,

and PGS. Lastly, I will review the current literature about cost-effectiveness of PGS as it relates to IVF and live birth rates.

Literature Search Strategy

For the literature search, I targeted multiple keywords followed by a snowball search approach to subsequent literature. Because some of these technologies are new and rapidly evolving, I limited the technical and clinical searches for PGS and IVF to the past ten years, and often only reviewed articles published in 2012 or later. Search terms were used individually and together in Boolean search terms of *AND* or *OR*. Key search terms were the following :

- *Assisted reproductive technology (ART)*
- *Centers for Disease Control (CDC)*
- *Comprehensive chromosomal screening (CCS)*
- *Cost-effectiveness analysis (CEA)*
- *Cumulative IVF*
- *Economic analysis*
- *Genetic screening*
- *In vitro fertilization (IVF)*
- *Live birth rates*
- *Preimplantation genetic screening (PGS)*
- *Protection motivation theory (PMT)*
- *Regulation of IVF*

Theoretical Foundation

The theoretical framework for this study was Rogers's (1975) protection motivation theory (PMT) as it applies to genetic screening. PMT is used to describe factors that motivate people to do certain things when it concerns their health. It focuses on cognitive processes and has three main components, including how the perception of severity of a health threat, the person's vulnerability to that threat, and the efficacy of the proposed protective mechanism. In this study, , the health threat is pregnancy loss combined with economic threat of overpaying for an alternate intervention. The literature describes the vulnerability, and the efficacy of the treatment arm is the clinical outcome. PMT is used as a model for genetic screening, including BRCA testing (Helmes, 2002). Rogers (1975) introduced PMT to describe how people deal with a perceived threat and how the level of fear of a certain outcome will change behavior and can be used to predict how much a threat has to be feared in order for someone to change his or her behavior (Boer & Seydel, 2005). Further, PMT can be used to explain the decisions and communication factors that apply when people decide to undergo a potentially prophylactic procedure, intervention, or test.

Genetic screening is just such an intervention and PMT is applied to understand the level of apprehension about a potential adverse outcome in compelling someone to take on the risk of screening. Further, Rogers (1975) describes two constructs that lead to a certain protective behavior. The first is the *threat appraisal*, which is the aggregation of all the factors that increase the likelihood of a protective response, like the perceived vulnerability to and severity of the health threat (in this case, an adverse IVF outcome

like an inability to implant or miscarriage), minus those factors that decrease the probability of a protective response, like costs or risks of the procedures. The second is the *coping appraisal*, formed by taking one's appraisals of self-efficacy (can I do it?) and response efficacy (will it work?) minus the financial aspect with the behavior (Ralph et al., 2014).

For genetic testing, PMT has been applied to describe the factors that contribute to the decision to undergo genetic protective testing. Helmes (2002) described the components of PMT that affected a woman's choice to have BRCA testing for breast cancer risk. While the author did not find every component of the PMT applied, the main factors such as fear, response efficacy, and costs played into the decision-making.

Only one study included the determinants of election to undergo PGS in the clinical setting. Gebhart, Hines, Penman, & Holland (2016) found that several factors that correspond to the PMT contribute to the patient-perceived determinants when choosing to accept or decline PGS, including "cost, religious and ethical beliefs and values, social and family support, provider influences, and the past reproductive experience of the patient" (Gebhart, Hines, Penman, & Holland, 2016).

For this study, I utilize PMT as the framework for the decision-making process involved with genetic screening. As this is an economic evaluation, I examined only the cost factor of PGS as a component of the PMT, although response efficacy (clinical outcomes) plays into the costs associated with PGS.

Literature Related to Key Variables and Concepts

Infertility

Infertility is clinically defined as the inability to conceive after one year of unprotected sexual intercourse, and affects approximately 10% of the population (Van Voorhis, 2007). Common reasons for female infertility include polycystic ovarian syndrome, a condition where a woman has a hormone imbalance issue, or primary ovarian insufficiency, a condition that is equivalent to early-onset menopause. Male infertility occurs less often, but can also contribute to a couple's infertility. Overall, a woman's age is the primary predictor of ability to conceive and have a healthy live birth (U.S. Department of Health and Human Services, Office on Women's Health, 2009).

Infertility may be treated through consultation with a medical professional that specializes in infertility. These specialists are reproductive endocrinologists and infertility physicians, (REI), and are board-certified obstetricians and gynecologists that have advanced training in the form of two or three year surgical fellowships (Accreditation Council for Graduate Medical Education, 2016). The American Society for Reproductive Medicine has defined other conditions separately from infertility but related to it, such a recurrent pregnancy loss (RPL). RPL is defined as a disease defined by two or more failed pregnancies of unknown etiology ("Definitions of infertility and recurrent pregnancy loss," 2013). Both infertility and recurrent pregnancy loss may be treated with the use of assisted reproductive technologies.

Assisted Reproductive Technology

The CDC (2014) defined assisted reproductive technology (ART) as any technology that involves handling both the sperm and eggs of human for the purposes of assisting reproduction. ART was introduced in the United States in 1981 by the birth of the first infant using techniques where eggs are surgically harvested, fertilized, and implanted in a woman (Sunderam et al., 2015). ART can include techniques such as IVF, gamete intrafallopian transfer (GIFT), pronuclear stage tubal transfer (PROST), tubal embryo transfer (TET), and zygote intrafallopian transfer (ZIFT) (American Society of Reproductive Medicine, 2015).

Other treatments for infertility exist, including intrauterine insemination (IUI), which involves injecting sperm into the uterus at the exact time of ovulation. REI may also choose to perform other surgical techniques to repair anatomical issues affecting infertility, or other drug therapies to aid in either male of female infertility problems. While many options exist for the treatment of infertility and RPL, IVF will be the focus of this study.

In Vitro Fertilization

IVF, combined with embryo transfer (ET), is a type of ART where a physician, typically a reproductive endocrinologist, uses hormonal stimulation drugs to induce a woman to ovulate multiple eggs. The eggs are then surgically harvested from the woman and put into a petri dish alongside sperm collected from a male partner. The sperm fertilizes the eggs in vitro, either naturally, or through an intervention known as intracytoplasmic sperm injection (ICSI). ICSI is a technique where an embryologist

collects sperm in a fine needle and injects one sperm into an egg, or oocyte, to commence fertilization. Once fertilization occurs, the resultant embryo is cultured for several days and then implanted into the uterus of the woman (Van Voorhis, 2007).

IVF has gone through a series of developments in technology over the years since Georgeanna and Howard Jones opened the first IVF clinic in the United States in 1980. However, the main tenets of IVF remain the same as when introduced. The process starts when a physician uses human menopausal gonadotropins (hMG) to stimulate the woman's ovary to ovulate multiple oocytes for retrieval and then retrieve them surgically. These gonadotropins can include follicle-stimulating hormone (FSH) and human chorionic gonadotropin (hCG), as well as other formulations of gonadotropins under various trade names (Jungheim, Meyer, & Broughton, 2015).

The oocytes are retrieved surgically by the reproductive endocrinologist, and the goal will be to remove as many oocytes as possible during the ovulation cycle. The reason for retrieving as many as possible is to attempt to create as many embryos as possible to have multiple options for implanting a healthy embryo back into the mother. This practice has led to an increase in multiple births over the years and has also led to a wave of concern about implantation of multiple embryos (Kulkarni et al., 2013).

Once the oocytes are retrieved, an embryologist, a scientist trained in the handling and evaluation of embryos as well as ART, will commence the fertilization process starting with fertilization of the oocyte on Day Zero (0) (Niakan, Han, Pedersen, Simon, & Pera, 2012). This process happens either through IVF, sometimes called natural fertilization, or through ICSI. Clinicians will utilize either approach depending on the

clinical presentation and screening prior to ART. Some studies suggest that in non-male factor infertility, ICSI does not provide any benefit over natural IVF (Eftekhar, Mohammadian, Yousefnejad, Molaei, & Aflatoonian, 2012).

Regardless of how fertilization occurs, an embryo will form and go through different stages of development. See Figure 1 for an illustration of this process. In general, egg and sperm will form a zygote and then start to divide in stages of one-cell, to two-cells, to four-cells, and then 8-cells by day 3. At this point, the embryo has developed into a morula, and then will turn into a blastocyst. This process generally takes 5-6 days from day 0 when fertilization occurs (Niakan et al., 2012).

As the embryo develops, the embryologist will grade the embryos daily through the use of morphology. Morphology is the process of examining the embryos under a microscope. There are different methods and criteria for grading and evaluating embryos, and no firm standard exists in the community for morphology grading, despite global efforts to standardize these non-invasive assessments (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011).

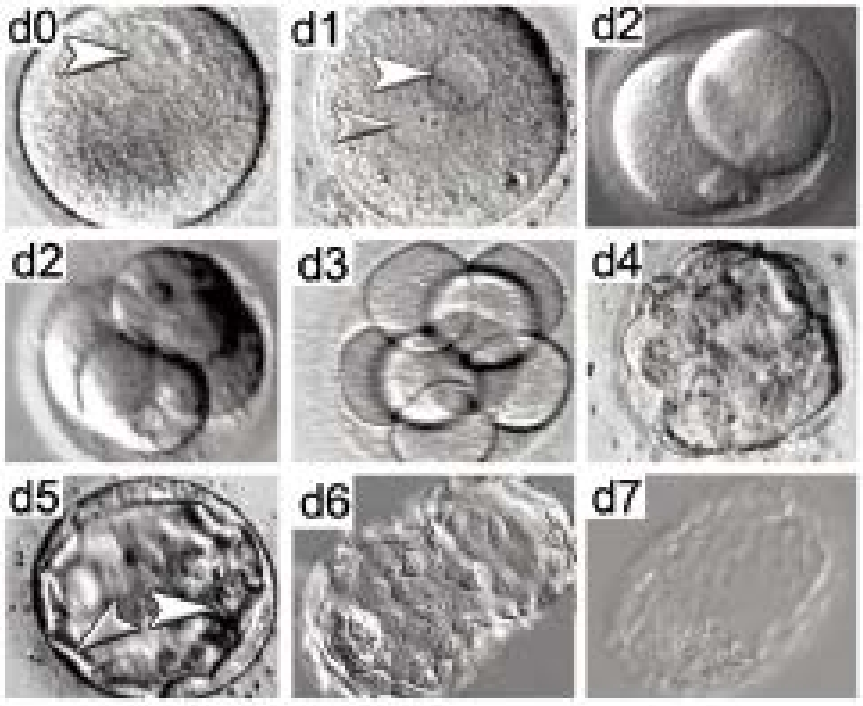

Figure 1. Stages of human pre-implantation embryo development. Reprinted from Niakan et al., 2012.

The physician and the mother may elect to perform fresh embryo transfer (ET) when the embryo is in the Day 3, or cleavage, phase, or a frozen embryo transfer (FET) when the embryo is in the blastocyst stage at Day 5/6. Studies have shown that blastocyst transfers can result in higher implantation rates, but do not significantly alter the live birth outcomes if a birth is achieved (Maxwell, Melzer-Ross, McCulloh, & Grifo, 2015).

The American Society for Reproductive Medicine (2013) has provided guidance on blastocyst transfer at the discretion of the physician and only in good prognosis women. Systemic reviews have determined that FET, when used in good prognosis patients, increases the chances of achieving a live birth and decreases miscarriage rates (Roque et al., 2013). However, there is evidence that fresh blastocyst-stage transfers do

not significantly impact live birth rates compared to fresh cleavage stage transfers (Glujovsky & Farquhar, 2016).

Whether a woman elects to do a frozen or fresh embryo transfer, she must also decide to implant one embryo, known elective single embryo transfer (eSET), or multiple embryos. Multiple births is a side-effect of ART over the last thirty years and multiple births (twins, triplets, etc.) have increased dramatically specifically because of ART (Kulkarni et al., 2013). One way to decrease the rate of multiple births is to only implant one healthy embryo. eSET, combined with CCS or PGS, has shown promise in decreasing costs and reducing live birth rates (Crawford et al., 2016).

While the American Society for Reproductive Medicine maintains certain opinions about best practices for IVF and embryo transfer, there are no standard ways of doing things form ART clinic to ART clinic, and multiple factors go into how an REI treats infertility (Van Voorhis, Thomas, Surrey, & Sparks, 2010).

Preimplantation Genetic Screening

PGS, also known as comprehensive chromosomal screening, is a genetic test that uses a number of different technologies in order to determine the ploidy (number of chromosomes) of an embryo while it is still in vitro. The goal of PGS is to determine which embryos are euploid (normal number of pairs of chromosomes, 23) or aneuploidy (abnormal number of pairs). Aneuploidy is a major factor in women having miscarriages, either with the use of ART or not (Liu et al., 2012). Euploid embryos are selected for implantation in the uterus as part of IVF. This should be differentiated from other preimplantation genetic tests, such as preimplantation genetic diagnosis (PGD),

which looks at a specific disease that is known to be carried by one or both parents (Brezina & Kutteh, 2015).

An embryologist performs PGS by removing one or multiple cells from a growing embryo after fertilization in vitro. This occurs either at the cleavage stage (Day 3, or 8-cell stage) for a removal of one cell for analysis, or at the blastocyst stage, where tens or hundreds of cells are removed from the trophectoderm at Day 5 or 6. Generally, when a Day 5 or 6 embryo biopsy is performed, it is always in conjunction with vitrification (freezing) of the embryos for FET (Brezina & Kutteh, 2015). While both types of embryo transfers are done in the United States, there is a growing body of clinical evidence that frozen transfers are the preferred technique in order to get the most optimal clinical outcomes (Maxwell et al., 2015).

Multiple technologies exist that can be used to establish the ploidy of an embryo. They include fluorescent in-situ hybridization (FISH), quantitative polymerase chain reaction (qPCR, or quant PCR), aCGH, and NGS, also known as high-throughput genomic sequencing. Many in the field refer to PGS performed by FISH or qPCR as "PGS#1", and NGS and aCGH technology as "PGS#2". This is to indicate the extreme clinical differences between the two sets of technologies (Gleicher et al., 2014). Generally, PGS performed before 2011 in the United States is considered PGS#1, and PGS performed after 2011 is considered PGS#2 (Yang et al., 2012).

Each platform has its advantages and disadvantages, but aCGH has the most relevant data associated with higher implantation rates and live births than any of the other technologies (Capalbo et al., 2015; Handyside, 2013). NGS is a relatively new

technique to perform ploidy analysis, but has shown itself to be just as effective at screening for euploid embryos as aCGH (Yang et al., 2015). Table 1 describes the different technologies as they relate to PGS. This study will focus on live birth rates from PGS#2 technologies, and will imply PGS#2 when the term "PGS" is used. It will also assume that there is no difference in NGS v. aCGH.

Table 1

Different Technologies for PGS

Method	Turn-around time	Costs to run	Resolution	PGS #1 or #2
FISH	24 hours	Low	Low	1
qPCR	4 hours	Low	Low	1
aCGH	12-24 hours	Medium	Medium	2
NGS	16-72 hours	Medium-low	Low	2

Note. Adapted from Handyside, 2013, p. 598.

Concerning the clinical efficacy of PGS, multiple investigators have reported their findings and have concluded that the use of PGS in conjunction with IVF/ICSI can yield higher implantation rates, on-going pregnancy rates, and live birth rates, regardless of the age of the woman (Dahdouh, Balayla, & García-Velasco, 2015; Keltz et al., 2013; Munné, 2012; Yang et al., 2012). Other studies have looked specifically at certain age groups to determine clinical efficacy. One such study examined the outcomes of live birth for women aged 40-43. The study was a retrospective cohort study that examined 450 cycles that did not have PGS, and 170 where PGS was employed along with FET. The authors found significantly higher live birth rates for FET with PGS than without PGS, 45.9% compared to 19%, respectively (Lee et al, 2015).

In a recent retrospective study, researchers examined the data from the CDC for the years 2011 and 2012 to determine live birth rates by age when PGS was used. The authors found that there were 5,467 PGS cycles compared to 97,069 non-PGD/PGS cycles during that timeframe. The authors concluded that of the PGS cohort, PGS did not increase live birth rates for women less than 35 years old, but did improve live birth rates among women over 37 years old (Chang, Boulet, Jeng, Flowers, & Kissin, 2016).

Multiple systemic reviews of the literature on PGS were conducted in the past six years. The first, conducted in 2011, showed no benefit to PGS with IVF for any age group based on the literature to date (Mastenbroek, Twisk, van der Veen, & Repping, 2011). In fact, that review concluded that the use of PGS lowered live birth rates in women aged over 40 compared to IVF alone, and that no significant benefit was seen in any age group. However, the literature at the time was limited to studies using PGS#1 only, and did not take into account any studies using aCGH or NGS.

Dahdouh et al. (2015) found PGS did improve embryo selection and implantation rates based on three RCTs and eight observational studies reviewed. While the authors did see live birth rates in a few of the studies, they did not report on the statistical significance of those outcomes. They also did not resolve PGS by age group. Further, the authors discuss the needs for further examination of PGS with IVF, especially in conjunction with cumulative IVF.

Lee et al. (2015) conducted a systematic review that included three RCTs and 16 observational studies. The authors concluded that the majority of the studies and RCTs showed that PGS provided improved implantation rates and ongoing pregnancy rates.

They also concluded that not enough evidence was provided to make meaningful conclusions concerning age group stratification.

Other researchers attempted to examine the efficacy of PGS via systematic review for women of advanced maternal age (AMA). The authors searched the databases in 2010 and found six RCTs that met their inclusion criteria. The authors found that PGS in women with AMA showed significantly lower live birth rates than without PGS. However, the authors noted that the methodologies, techniques, and technologies used for all of the reviewed papers varied widely (Noble et al., 2010). Like some of the other earlier papers on this issue, PGS #1 technologies such as FISH were predominantly seen in these studies.

Lastly, Chen, Wei, Hu, and Quan (2015) conducted a meta-analysis of PGS studies to date. For live birth rates, the four studies examined that discuss this outcome show an average live birth rate using PGS of 65.6%. The studies examined all had different patient populations and looked at differing embryo transfer methods, such as SET or DET. Forman et al. (2012) looked at only SET, possible resulting in the lower live birth rates than the other studies. Concerning age, the one RCT did not discriminate based on age, and the three other studies examined women of different age groups, either older than 35 or younger. There was no conclusion drawn by the authors as to efficacy by age group. Table 2 is adapted from that study and compares the most relevant studies where live birth rates are the outcome.

Table 2

Studies examining live birth rate for PGS and non-PGS

Study	Type of Study	PGS live births / total	PGS Live Birth Rate	Non-PGS (Control group) live births / total	Non-PGS Live Birth Rate
Scott et al., 2013	RCT	61/72	84.7%	56/83	67.5%
Forman et al., 2012	Cohort	49/140	35%	63/182	34.6%
Greco et al., 2014	Cohort	59/88	67%	7/33	21%
Schoolcraft et al., 2010	Cohort	34/45	75.6%	78/113	69%

Note. Adapted from Chen et al., 2015, p. 10/21.

Despite the evidence, PGS remains controversial and is not recommended by any of the medical groups that govern the space, such as the American College of Obstetricians and Gynecologists (ACOG) or the American Society for Reproductive Medicine (ASRM) ("ACOG Committee Opinion No. 430," 2009). Many continue to doubt the clinical usefulness of PGS, and have suggested that non-PGS embryo selection would lead to higher live birth rates than performing PGS (Gleicher et al., 2014; Orvieto, 2016).

Trends in PGS Adoption

Despite the lack of consensus on routine clinical use of PGS, it continues to see increasing adoption among individual practices in the United States. The CDC has reported PGD rates for years but only for 2014 has the term "PGS" been added to clarify it is included with this rate. The CDC does not break out PGS versus PGD, however, so the exact trends are difficult to ascertain (Sunderam et al., 2015). Despite this, the overall

trend of all clinics reporting in the United States has ranged between 3 and 5% for the past five years (CDC, 2016).

Others have reported increased clinical usage overall in the United States from 2014 to 2016, years that are not covered by CDC statistics. A recent paper highlighted the increasing trend in PGS use in the U.S., mainly due to the advanced technologies of aCGH and NGS and better understanding of FETs. Some have estimated a 30% increase in PGS utilization per year since 2014. Some have also suggested a more accurate PGS utilization rate of 8% of all cycles (Mastenbroek & Repping, 2014).

Based on the available data, it is clear that each clinic views the use of PGS differently. The CDC data for 2014 shows an overall rate of PGD/PGS of 4%, however, there are clinics that report 0% PGD/PGS rates and others that report 86% of all cycles utilize PGD/PGS (Centers for Disease Control and Prevention, 2016). Some clinics advocate only frozen transfer and recommend PGS for every patient regardless of age ("IVF Success Rates with PGS | Arizona IVF Clinic," n.d.). The fact that every clinic has a different view of PGS makes this one of the most controversial aspects of this extremely important clinical decision. Mastenbroek and Repping (2014) state:

> In our view evaluating the (cost-)effectiveness of medical treatments is by far the greatest challenge in current day medicine, especially in an era where health care costs continue to increase to the extent where they are the number one item of expense for many governments across the globe.

This is why the CEA proposed in this study fills a gap in the literature and why it is adds to the growing body of evidence for the adoption of this technology.

Regulation of IVF

IVF has been a useful tool for several decades now and has proven to be a successful option for women and men with infertility. Since the inception of ART, there have been concerns about the regulation of this nascent technology and how it affects consumers and patients. In 1992, Congress enacted the Fertility Clinic Success Rate and Certification Act of 1992 (FCSRCA) (Public Law 102–493). This act mandated that all clinics performing these advanced reproductive techniques report their success rates via various metrics to the government, specifically, the CDC, in a standardized manner (CDC, 2016).

The FCSRCA requires clinics performing ART to report certain metrics, and they include live birth rates and PGD/ PGS rates. Data is collected via a web-based tool known as the National ART Surveillance System (NASS), and in 2013 over 94% of all U.S. clinics performing ART reported their results (Sunderam et al., 2015). These results are meant to guide prospective patients to understand each reporting clinic's success rate for performing ART, and assist with the selection of a potential IVF clinic. Further, the CDC aggregates this data year-by-year for analysis of trends in the space (CDC, 2016).

Other regulatory considerations of IVF include payment and insurance coverage. Most states do not require health insurance companies to provide coverage for IVF and /or ICSI, and no states mandate coverage of PGS (National Conference of State Legislatures, 2014). Of the few states that do mandate coverage, studies have found that birth rates increase due to the elimination of the financial aspect of doing multiple cycles

of IVF (Boulet et al., 2015). Further analyses of this kind are needed to dictate future policy for insurance coverage of IVF.

Cumulative IVF

Achieving implantation, on-going pregnancy, and ultimately, a live birth, using ART involves many factors. One such factor is the number of attempts made by using IVF, or the cumulative success rate of multiple cycles of IVF. While the data reported by the CDC provides details of success rates in cycles, it does not take into account any cumulative or multiple cycles, thus limiting our understanding of the actual success rates per cycle (Sunderam et al., 2015). To date, there is limited data to clarify how linking IVF cycles contributes to live birth rate success rates. However, two studies shed some light on the issue.

The first study was an analysis of CDC data to estimate live birth rates. Of 471,208 cycles from 2004 to 2009, there were 140,859 live birth rates. Overall, the authors estimate that by the third linked cumulative cycle of IVF, the conservative live birth rate estimate for all women is 52.4%. The authors conclude that the rate of natural fecundity can be achieved through linked IVF cycles. They also find that maternal age is the major factor in all live birth rate trends (Luke et al., 2012).

The second study is more recent and is a prospective study that examined the live birth rate for 156,947 in the UK that underwent multiple cycles of IVF from 2004 to 2010. Results showed that the conservative estimate of cumulative live birth rates for women undergoing at least three cycles of IVF was 44.6%. The authors were also able to stratify these results by age group and for women using their own oocytes, and the results

showed the conservative estimates by cycle #3 for ages <40, 41-42, and > than 42 to be 48.6%, 18.5%, and 5.4%, respectively (Smith et al., 2015).

Both studies make assumptions about patients that drop out after each successive cycle of IVF. Drop-out rates from IVF treatment have been reported between 17-65%, depending on the age of the woman, treatment schedule, and which number of cycles of treatment (Luke et al., 2012; Verberg et al., 2008). Calculating the cumulative birth rates, therefore, require certain assumptions for those patients that dropped out. In the Smith et al.'s (2015) study, their conservative estimates assumed that all of the patients that dropped out in a previous cycle would have achieved a live birth rate of 0% had they stayed in treatment. Reasons for dropping out include mental stress, financial issues, and other psychological factors (Verberg et al., 2008).

To date, there have been no studies that compare the cost-effectiveness of the use of PGS in IVF/ICSI with multiple cycles of IVF without PGS. The ability for a patient to assess their options in IVF treatment and to make decisions based on all the factors associated with likely outcomes is critical to societal change. CEAs aid in understanding the economic impact of making decisions and is a factor of decision-making as it relates to the protection motivation theory.

Cost-effectiveness Analysis

CEA is a kind of economic analysis that attempts to quantify the overall costs of something and compare them to some outcome. In a CEA for medical applications, the idea is to provide insight about the most cost-effective option between two or more interventions, considering health outcomes. It can also be defined as an analysis where

one wishes to maximize some objective (health outcomes) given a limited resource (money). It is a form of economic analysis and can be contrasted with cost-benefit analysis (CBA), where a CBA expresses all outcomes as a monetary unit, and a CEA will attempt to quantify the effectiveness compared to a relatively qualitative amount such as health outcomes (Grosse et al 2008). This unique idea gave rise to the idea of the ICER and the quality-adjusted life year (QALY).

An incremental cost-effective ratio is the basis for all CEAs for medical applications. It consists of a numerator that consists of the monetary amount of the net expenditure of health resources divided by a denominator of the incremental benefit of the health improvement, a non-monetary measure (Weinstein et al., 1997). An area of debate in this field remains what should be included in both numerator and denominator to best capture the overall cost-effectiveness of some intervention over another.

The numerator should consist of all the factors that express the net economic benefit of some intervention. This can include direct costs of the intervention, either fixed or variable, long-run opportunity costs, labor costs, or some variation of all of these. It can also include costs saved, or economic benefits gained. Many economists have agreed on the need to define clearly what costs, savings, and/or benefits are being measured in a CEA in order to understand the net economic improvement (Weinstein, et al, 1997).

The elements of the denominator are an even more complicated issue. Many attribute the introduction of the QALY to Weinstein and Stason (1977) in their groundbreaking work concerning hypertension. A QALY attempts to quantify the

extension of life (or the minimization of pain) through quantification from zero to one, where zero represents death and one represents perfect health. The QALY framework is designed to be applied in any disease-state because of the flexibility of the quantification of the health outcome via some instrument such as the EuroQol (Meltzer, 2001). These instruments seek to capture the level of health outcome from zero to one, and can be highly subjective.

CEAs have been used in recent years for both genetic screening techniques in many fields as well as in IVF for a multitude of purposes. These CEA provided insight into best practices to use for both genetics screening and IVF and can be applied to this study.

CEA in Genetic Testing and Screening

Genetic testing, such as PGS, provides a unique set of challenges for economists. While such testing holds promise to provide better diagnoses, eliminate unnecessary procedures, and predict future disease risk, the impact of such testing is still very new and therefore the timeline and costs associated with this testing is difficult to capture. An analysis of economic evaluations of genomic technologies showed that measuring outcomes is problematic and that the costs captured in the numerator of an ICER vary from study to study by costs captured, timeline, and scope (Buchanan, Wordsworth, & Schuh, 2013).

A number of studies examine genetic testing and their cost-effectiveness, and as described, vary in application, scope, and outcomes. One study looked at the CEA of *OncotypeDX* (Genomic Health, Redwood City, CA), a commercially available genetic

test that discriminates patients based on genetic markers for chemotherapy intervention. The study utilized a Markov decision model with 146 patients who either decided to use the test or not, and incremental cost ratio was overall costs for either intervention divided by a QALY-based score that encompassed longevity, recurrence, and other factors. The results were in a better cost-effectiveness ratio with the test than with standard care (Holt et al., 2013).

Another similar study sought to understand the CEA of a commercially available test for non-small cell lung cancer (NSCLC), known as *VeriStrat* (Biodesix, Boulder, Co.). The study looked at two outcomes from the perspective of the U.S. payer, cost per QALY gained and total direct lifetime costs. Again, the authors used a Markov decision model to look at two interventions – with the test and without the test. They made assumptions about the timeline of the costs captured through either intervention and used published data to compare clinical effectiveness of the test. The authors found that use of the test was more cost-effective than standard treatment without the test (Hornberger, Hirsch, Li, & Page, 2015).

In both studies, similar analytical and statistical methods were used, including Markov decision analytics, one-way sensitivity analyses, and probabilistic sensitivity analyses (PSA). Because genetic testing is rapidly evolving, CEAs of this kind of testing are a challenging field. Some have recommended methods that go beyond traditional CEA for this kind of intervention, including using cost-benefit and cost-utility assessments to better understand the decisions for consumers to use such tests (Grosse et al, 2008). Decision-making associated with genetic screening is part of the protection

motivation theory, where economic impact is one factor in making the decision to have the intervention.

CEA in IVF

There are multiple articles describing the costs and cost effectiveness of IVF, IVF with ICSI, eSET, and PGS, although each one varies in the scope, timeframe, and costs included as previously discussed. Further, studies examining the outcomes and reporting for economic evaluations involving fertility find that the time horizon, costs included, and outcome endpoints differ from study to study and that the intervention being studied may be biased (Goldhaber-Fiebert & Brandeau, 2015). These findings are consistent with other analyses of economic evaluations in this space.

Fiddelers et al., (2009) examined the cost-effectiveness of seven different IVF strategies that range from eSET with standard treatment to double embryo transfer (DET) with other conditions. The authors utilized Markov decision models combined with probabilistic sensitivity analysis as well as Monte Carlo simulations to understand the gain in effectiveness of any one particular strategy. The authors concluded that combining treatment strategies is not cost-effective, and that cumulative (up to three) cycles of IVF with eSET or DEP is equally cost-effective.

Cumulative IVF outcomes have also been examined through the lens of CEA. A recent study sought to understand the cost-effectiveness of DEP versus cumulative IVF attempts with eSET. The study was a retrospective analysis of previous data reported through NASS via the CDC. The authors also utilized sensitivity analysis to analyze the

data to determine total costs of eSET versus DET. They concluded that cumulative eSET for certain patients is a more cost-effective option than DET (Crawford et al., 2016).

Economic evaluations of ART and IVF are useful in estimating costs and obtaining meaningful cost estimates. A study examined the overall costs and elasticity of ART for multiple first-world regions/countries, including Canada, USA, UK, Scandinavia, Japan, and Australia. The review of overall costs for ART procedures in multiple countries revealed excellent cost data per live birth, where the United States had a cost per live birth with ART of $41,132 compared to $33,183 in Canada (2006 dollars) (Chambers, Sullivan, Ishihara, Chapman, & Adamson, 2009).

CEA in IVF and PGS

There are multiple CEA studies that specifically involve PGS and IVF and the results and conclusions are mixed. The first such study examined the cost-effectiveness of PGS #1 (FISH) in infertile women aged 38-40 and >40 undergoing one cycle of IVF with subsequent frozen cycles available to achieve a live birth. The authors utilized a decision-analysis model to examine the live birth rates with each intervention. They employed sensitivity analysis to determine which intervention provided the best costs per live birth. The authors concluded that IVF alone was a far more cost-effective strategy than PGS with IVF in women aged 38-40. For women aged 40 or older, they conclude that CEA is roughly equal between strategies (Mersereau et al., 2008). The main limitation of this study is that they did not look at PGS#2 technologies where the clinical probabilities of a live birth have increased dramatically, and that costs over the past eight years have changed dramatically.

A more recent economic study examined the cost-effectiveness of two interventions, IVF combined with PGS, and expectant management for patients with recurrent pregnancy loss. Expectant management is a clinical strategy that allows the infertile couple to continue to attempt spontaneous natural conception without the use of ART. The study used a CEA decision tree to examine the two strategies and used the clinical effectiveness of each strategy obtained from the literature, specifically the clinical live birth rate obtained from Hodes-Wertz et al. (2012) where PGS was performed on 287 cycles of IVF for RPL patients. The clinical live birth rate for expectant management was obtained from Brigham et al (1999) that looked at live birth rates from 325 patients with RPL that were treated with expectant management. The study found that expectant management was a far more cost-effective strategy that IVF in conjunction with PGS for patients with RPL, regardless of age group (Murugappan et al., 2015).

A recent abstract from the American Society of Reproductive Medicine meeting concluded that PGS in a fresh transfer was not cost-effective compared to IVF alone in patients older than 39. The analysis was made using previous data concerning false positive rates in PGS (Salem et al., 2016). This was the first study to look at the influence of false positive rates on the cost-effectiveness of PGS and IVF.

Other studies conflict in their conclusions about PGS and IVF. Recent studies have looked at retrospective data to determine the cost per live birth with IVF and PGS compared to IVF with fresh transfers and subsequent frozen transfers in different cohorts of women. A single-center recent study by Hodes-Wertz, McCulloh, and Grifo (2015) examined retrospective data for patients using PGS and without PGS with routine IVF

from 2011-2013. There were 1,910 patients total, with 1,133 undergoing routine IVF with subsequent FET and 777 patients undergoing one cycle of PGS with IVF/ICSI. The overall live birth rate without PGS was 24%, and the live birth rate with PGS was 52.5%. The authors used direct costs for IVF and PGS and found that PGS was cost-effective in all age groups but women over 42.

The conclusions of that study have been seen in other CEA analyses. One recent abstract examined the cost-effectiveness of IVF with aCGH PGS compared to IVF alone for patients with RPL. The authors concluded that it cost $19,416 to achieve a live birth with PGS, versus $23,184 with IVF alone. These results were combined and the authors concluded that all age groups saw similar cost benefits (Resetkova et al., 2013).

However, others have looked at similar retrospective data and determined that there is a bias in the selection of women with PGS, and that explains any increase in live birth rate. They conclude that PGS with IVF is not a cost-effective strategy compared to IVF alone (Kushnir et al., 2016). In almost all cases, the authors recommend further studies to understand the true cost-effectiveness of PGS in any IVF treatment strategy.

Of note is the wide variation in cost estimates to achieve a live birth from study to study. Murugappan et al. (2015) reported a cost of $45,300 to achieve a live birth using PGS, whereas Resetkova et al (2013) reported a cost of $19,416 to achieve a live birth with PGS. Hodes-Wertz, McCulloh, and Grifo (2015) found that costs per live birth varied by age group, with costs for women under 35 totaling $65,278, women aged 35-39 of $65,841, women aged 40-42 of $89,350, and women aged over 42 of $291,907. Further, Chambers et al. (2009) reported a general cost per live birth of $41,132 in the

United States, where PGS or PGD costs were not considered. These variations in estimates lead to the large gaps in understanding the true cost-effectiveness for PGS.

Summary and Conclusions

The current literature associated with PGS and its impact on achieving a live birth via IVF reveals mixed conclusions where no consensus about its clinical utility has been reached (Orvieto, 2016). Summarizing what is known and not known for PGS, the main issue is the fierce debate between researchers on both the clinical utility of PGS as well as the cost-effectiveness. Concerning the clinical efficacy, the literature that does not include PGS#2 technology shows that PGS with IVF has a lower birth rate than with PGS, regardless of age. However, the literature points to increased live birth rates, decreased miscarriage rates, and increased pregnancy rates for all age groups with the use of PGS, leading more and more clinicians to adopt routine PGS use in their clinical practice (Mastenbroek & Repping, 2014). Further, most research agreed cumulative IVF attempts result in increased live birth rate in certain populations of patients.

The cost-effectiveness of using PGS with IVF is also not clear from the current literature. It is also not clear if using PGS is effective in all age groups, or if age makes a difference in whether PGS is cost-effective or not. Many studies show the benefit of using PGS in certain populations of patients, including all women under age 42 (Hodes-Wertz et al., 2015; Resetkova et al., 2013). Others have shown some cost-effectiveness in women but only in women aged 38-40, and not for older women (Mersereau et al., 2008). Other studies suggest that PGS is not cost-effective at all, regardless of age group (Kushnir et al., 2016). And still most CEA studies involving PGS do not stratify the

results by age group, making it difficult to assess the impact age makes on the results (Murugappan et al., 2015).

Dahdouh et al. (2015) stated this about the gaps in the literature concerning CEA and PGS:

> The cost-effectiveness of CCS can be resolved only following cumulative live birth rates. NGS may represent the next frontier of PGS and has the potential to decrease costs. Any future study evaluating the cost-effectiveness of CCS versus standard care must take into account all available procedures and related cost. These aforementioned points need to be considered, but until such data become available, the debate continues on this crucial aspect of PGS. (p. 1509)

Comparing cumulative live birth rates and IVF with PGS, along with understanding the variance associated with age group, is the next step to understanding the usefulness of PGS in the clinic, and it is what this study hopes to achieve.

This study seeks to continue to enhance the community's understanding of whether PGS with IVF in one cycle is cost-effective compared to going through multiple, cumulative cycles of IVF without PGS, and whether age makes a difference in that cost-effectiveness. While there are some recent studies that examined the cost-effectiveness of PGS based on retrospective data, but none to date have examined IVF with PGS in one cycle compared to achieving a live birth if a patient goes through three consecutive cycles of IVF without PGS.

Chapter 3 will include a description of the methodology of this study. Further, it will include the rationale for the design of the study and the methods used to analyze the

results. It will describe the model used for this study and the analytics used to determine the results of the study including any threats to validity. Lastly I will examine what costs will be included for each intervention and their source.

Chapter 3: Research Method

The purpose of this study was to compare the cost effectiveness of two interventions using secondary data to determine the best economic outcome based on limited resources. The two interventions are IVF utilizing PGS and cumulative, consecutive cycles of IVF without PGS. The clinical outcome is a live birth (defined as a singleton or multiple birth after 24 weeks of gestation that survived at least one month) and was stratified by age group. In this chapter, I will discuss the research design and methodology and as well as the rationale for the approach. The components involved in the costs of each intervention are also discussed, along with any threats to validity. Lastly, I will explore any ethical considerations with this research design.

Research Design and Rationale

The research design of this study was a cross-sectional, quantitative, economic modelling design using secondary sources. The independent variables in this study were the method of treatment, either one cycle of IVF with PGS (IVF-PGS) or three consecutive, cumulative cycles of IVF (3xIVF). The dependent variable was cost, or cost-effectiveness as determined by an incremental cost-effectiveness ratio (ICER) of costs per live birth. In order to study the effect of each intervention on cost-effectiveness, a theoretical decision-analytical model—in this case, a cost-effectiveness health model—was created using TreeAge Pro 2017 (www.treeage.com). The conceptual strategy was for the cost-effectiveness of two interventions to be compared using the model.

The rationale for using a cost-effectiveness model was that the vast majority of CEA in healthcare utilized a decision-analysis tool to determine the best path forward

between two or more interventions. Prior CEAs studies that have involved PGS and IVF have been conducted using a decision-analytical model or Markov model to understand how each intervention compares to another (Hodes-Wertz, McCulloh, & Grifo, 2015; Murugappan et al., 2015).

A cost-effectiveness model is theoretical, and does not limit the researcher based on any time or resource constraint other than collecting the relevant data, building the model, and analyzing the results. Statistical analysis for this research design included sensitivity analyses, probabilistic sensitivity analysis (PSA), and Monte Carlo simulations. More details on each statistical method and interpretation are in the data analysis section.

To explore the research questions, I compared each intervention via a decision-analytical model. First, the model allowed for direct comparison of cost-effectiveness for each intervention. Second, by stratifying the analysis by age group, I clearly illuminated any differences due to the age of the patient. I explain the methodology and assumptions made in the model in the next section.

Methodology

Population and Sampling

The population for the model was any infertile couple seeking treatment from an IVF clinic and data was from secondary sources. Many of the seminal works for PGS and clinical effectiveness in delivering a live birth have been prospective trials in which researchers looked at embryos compared to live births (Hodes-Wertz, Grifo, Ghadir, Kaplan, & Laskin, 2012). Researchers who conducted one cost-effectiveness study on

PGS in RPL patients used studies from the literature in order to determine the live-birth rates for each arm of the decision matrix (Murugappan et al., 2015). In that particular study, researchers compared the IVF with PGS of 287 cycles of IVF results with the expectant management results from 325 patients from a longitudinal study by Brigham et al (1999). Many CEAs are not prospective trials and used the literature to determine the outcomes of the decision arms (Garber & Phelps, 1997).

For this study, sample size needed to be determined for each arm. Three factors influence the sample size: the Effect Size, the Power, and the Alpha Level (or significance level). Knowing these three factors will allow a researcher to calculate the fourth. Effect Size is the importance of the treatment against the background of noise in the measurement. Power refers to the probability that you will observe an effect of treatment when there really is one. Lastly, the Alpha Level is the probability that the observed result is due to chance (Trochim, 2006).

Ideally, an equal number of participants should be in both arms, and because retrospective data collected from the CDC and other studies is being used, there will be some finite number of participants. Generally, a large power variable is preferable, and a smaller alpha variable when the chances of making a Type I error are too great. Effect size was based on the treatments available for both arms. In this study, because I was studying the effects out to a similar outcome (live birth) for both arms, then the Effect Size needed to be .5, indicating that I believed the treatment to result in a live birth for both arms.

However, because the samples from the literature are set, they were more likely to have to calculate the effect size from the sample set. Given the samples from other CEA studies of 250 patients, an effect size of .5, and an alpha of .05, a G*Power analysis provided a power of .99, which indicated that at least 250 patients in each arm was more than statistically significant enough to avoid Type I errors. Both interventions had more than 250 participants. Understanding the inputs for each intervention strategy determined the number of patients examined in each arm.

Intervention Strategies

The basic cost-effectiveness decision model is shown in Figure 2. For the consecutive cycle strategy (3xIVF), the model assumed that each patient underwent an IVF cycle without using PGS to screen for ploidy. In this strategy, a cycle is defined as any number of embryo transfers from a single given oocyte retrieval event. The model assumed that each cycle included one fresh and one FET. The model also assumed that all patients used their own oocytes, and no donor eggs were used. The possible outcomes of the first cycle were no pregnancy, clinical miscarriage, or live birth. If the patient did not have a live birth, they would move to the second theoretical cycle, where the possible outcomes were the same as the first cycle. Again, if no live birth, then the patient would move to a third cycle, where the possible outcomes were the same as the other two cycles. The model assumed that no patients dropped out between cycles, despite evidence that large number of patients do dropout between cycles (Luke et al., 2012). The decision tree ends at that point.

For the IVF-PGS strategy, the model assumed that every patient had at least one euploid embryo. At that point, each patient underwent IVF followed by a FET (no fresh transfer). The possible outcomes were no pregnancy, clinical miscarriage, or a live birth. The model did not make assumptions about the number of embryos transferred in this one cycle. For costs, both interventions assume that ICSI was employed for each cycle, with or without PGS.

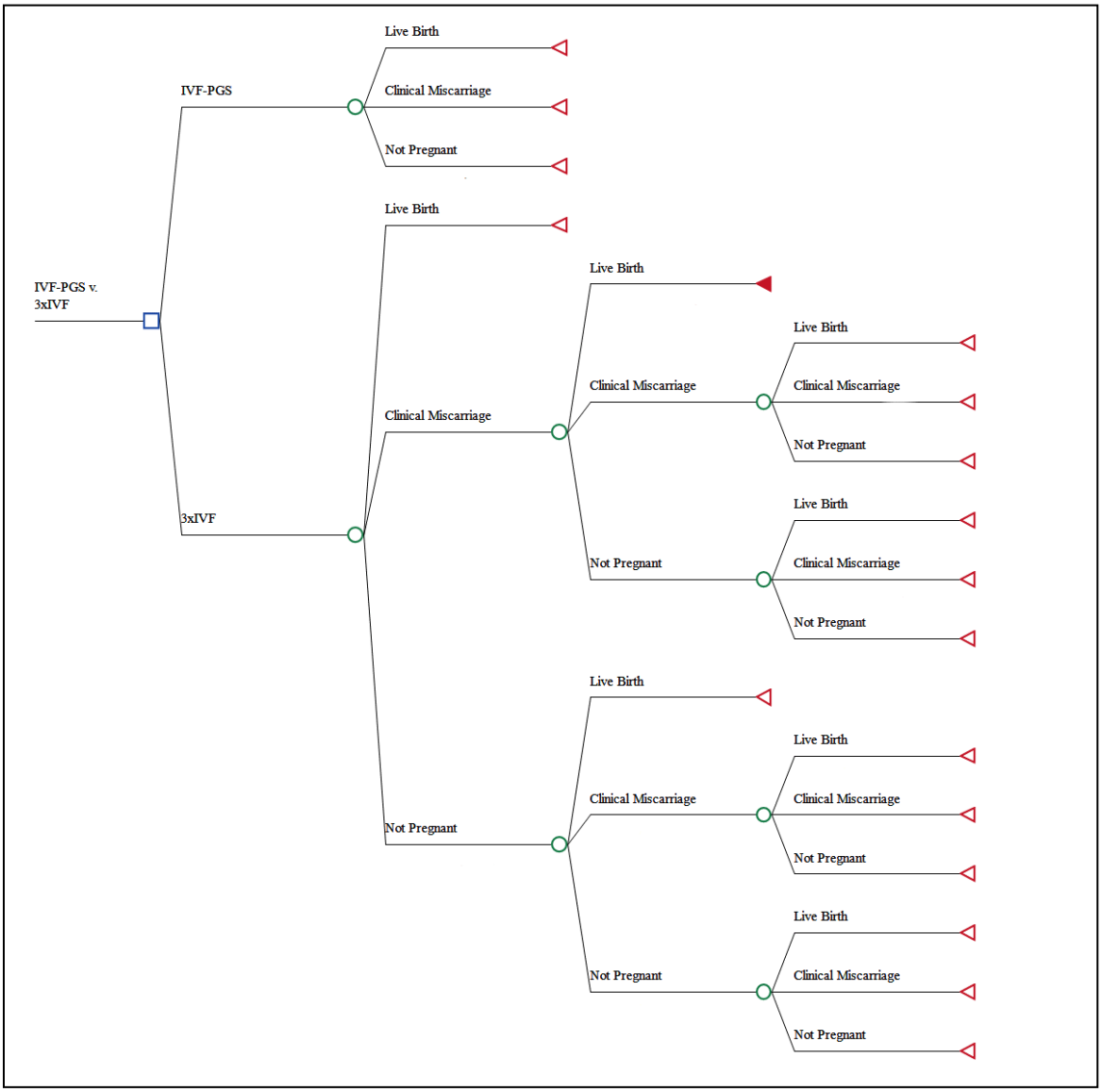

Figure 2. Simplified decision tree showing the two interventions, IVF with PGS and three consecutive cycles of IVF.

Clinical Probabilities

For each node of the model, there was an assumption based on the clinical probability of each possible outcome. For these statistical rates, I utilized published literature and recent studies to determine the outcome.

Cumulative (3xIVF) strategy. For the consecutive, cumulative IVF strategy (3xIVF), the probability rates were taken from the groundbreaking study on cumulative IVF and live birth rates by (Smith et al., 2015). In this study, the authors examined 257,655 cycles of IVF that were performed on 157,475 women in the United Kingdom between the years of 2003 and 2012. They excluded any women who had previously had a live birth from ART, or had started IVF prior to 2003. They also excluded any cycles that were not started exclusively for the purpose of achieving a live birth. This is the largest retrospective study to date that links multiple cycles of IVF with live birth rates (Smith et al., 2015).

Smith et al. (2015) were able to determine live birth rate in each individual cycle and infer cumulative live birth rates up to nine cycles of IVF. Further, the study was able to stratify these findings by age group for women under the age of 40, aged 40-42, and older than 43. The authors utilized three assumption strategies to deal with women who dropped out between each consecutive cycle. First, they calculated cumulative live birth rates and assumed that every woman that dropped out would have had the same rate of live birth as those that stayed in, known as the “optimal estimate”. Second, they assumed that 30% of the women that dropped out would have not achieved a live birth, known as the “prognostic-adjusted estimate”. Lastly, they assumed that no woman that dropped out would have achieved a live birth, known as the “conservative estimate”. For the purposes of this study, we assumed no participants dropped out between cycles and used the optimal estimate.

Another critical element to determine costs associated with each intervention is pregnancy rate and miscarriage rate. For the 3xIVF strategy, pregnancy rates from the 2014 CDC survey data was used for each age group. Because the CDC stratifies ages for each two year block as well as by frozen or fresh transfer, I averaged the pregnancy rates for all women under the age of 40 regardless of transfer method. I then calculated miscarriage rates by subtracting the 1st cycle live birth rate obtained from the Smith et al. (2015) study from the pregnancy rate published by the CDC ("ART Success Rates (ART) Data | Reproductive Health | CDC," n.d.). If the outcome was not a miscarriage or a live birth, then it was assumed to be a no pregnancy. See Chapter 4 for values.

Live birth rates for IVF-PGS. The clinical success rates from the literature for utilizing PGS in one cycle of IVF have ranged from 35% to 85%, depending on the study and the methods used **(**Chen et al., 2015). In their CEA analysis of IVF with PGS compared to expectant management for patients with recurrent pregnancy loss, Murugappan et al. (2015) utilized the clinical rates established by Hodes-Wertz, et al (2012) where PGS was performed on 287 cycles of IVF for RPL patients. That live birth rate in that study was 40% with a calculated miscarriage rate of 7%. However, these rates were not stratified by age group and so are difficult to apply to this study.

A more recent study stratified live birth rates for IVF with PGS by age group, and fell within the range of live birth rates established in other studies, but not does not reveal pregnancy rates or miscarriage rates (Hodes-Wertz et al., 2015). Another recent study by Salem, et al (2016) provides clinical miscarriage rates for patients in the same age groups, and are similar to the calculated miscarriage rates used by Murugappan et al (2015). This

study utilized these rates and for clinical probabilities used in the model, and are shown in Chapter 4.

Costs

The most important aspect of any CEA is what costs are associated with the analysis (Meltzer, 2001). For infertility treatment studies, accurate and consistent values for costs remain a difficult variable to ascertain and remain one of the fields most challenging topics (ESHRE Capri Workshop Group, 2015). Each CEA that involves PGS has used slightly different costs, but most recent analyses utilize direct costs to the patient associated with IVF, PGS, miscarriage, and births.

Recent studies involving PGS have identified the minimal costs that need to be associated with any PGS-related cost-effectiveness study. From Dahdouh (2015),

> Any future study evaluating the cost-effectiveness of CCS versus standard care must take into account all available procedures and related costs. These should include, at minimum, the costs of the following techniques: the IVF cycle, the CCS analysis, the first fresh embryo transfer and all subsequent frozen transfers, the obstetrical care, and any additional prenatal care provided to ongoing aneuploid gestations, from pregnancy time until delivery. (p. 1509)

Hodes-Wertz et al. (2015) identified the costs associated with their recent PGS cost-effectiveness, and it included costs for

> …routine IVF cycle with and without embryo transfer, FET of unscreened and euploid embryos, IVF with PGS (including shipping of specimen, outside laboratory cost to run specimen, and cryopreservation of embryos) with and

without transfer, dilation and curettage performed on early pregnancy failure, termination of pregnancy, laparoscopic salpingectomy per-formed for ectopic pregnancy, livebirth of singleton, twins, and triplets. (p.p505)

This study will include the direct costs for both strategies as defined above by Hodes-Wertz et al. (2015). However, because both strategies have an outcome of cost per live birth, we can assume that obstetrical care (costs for live birth obstetrical care) are equal in both strategies, and therefore can be excluded. The costs used in this study are detailed here:

- Procedure Costs (PC), including:
  - Routine IVF Cycle
  - Oocyte retrieval
  - Medications costs
  - ICSI
- Fresh Embryo Transfer (FT)
- Frozen Embryo Transfer (FET)
- PGS Costs (PGS) (includes procedure, shipping, and reporting)
- Clinical Dilation and Curettage (D&C) associated with miscarriage

The other issue facing researchers is the source of the costs associated with these procedures. This study used the perspective of the payer, in this case, the patient, to estimate costs. There are multiple studies that highlight the costs used for each kind of procedure. To attempt to normalize the variation in costs and methods from study to study, we averaged multiple studies' cost estimates, along with costs obtained with

permission from one established IVF clinics in the United States. See the section on Archival Data for more information on the sources for data used in this study.

I assumed that for the 3xIVF arm, there was one fresh embryo transfer and one FET per cycle. For the IVF-PGS arm, I assumed only one FET, as I assumed that all PGS procedures would be trophectoderm biopsies followed by a "freeze all" strategy (Maxwell et al., 2015). Further, I assumed that patients incurred the full cost of three cycles of IVF in 3xIVF strategy, and that costs for multiple cycles will be calculated per cycle. In other words, each cycle of IVF was assumed to incur the same costs as any other cycle (more discussion about the advent of discounted multi-cycle plans in the IVF community can be found in the Chapter 5).

Archival Data

I used secondary data from a combination of public domain and from individual clinics for clinical probabilities and costs. The sources of this data included data available from the literature, clinic websites, by permission from one private clinic, and from public databases maintained by the CDC. Data obtained from the literature are cited and any additional data obtained from the authors were retrieved by permission from the authors. Clinical data was retrieved from the CDC database. Costs were obtained from the literature as well as from one private clinic. The private clinic is located in New York City and performs over 2,300 cycles of IVF per year ("ART Success Rates (ART) Data | Reproductive Health | CDC," n.d.). The data for this clinic was obtained via permission and all data was de-identified of all personal health information (PHI). See Appendix B for the data use agreement letter.

The CDC maintains a database of data obtained from all clinics in the United States that perform ART. This data is updated annually and is typically two years aged from the current year due to the length of time for pregnancy. The data can be downloaded without permission at https://www.cdc.gov/art/artdata/index.html and results are delivered as an excel spreadsheet. They are also reported in a comprehensive report funded by the CDC (Sunderam et al., 2015). Specific costs, their sources, and the mean costs are detailed in Chapter 4.

Data Analysis Plan

All statistical analysis was performed in TreeAge Pro 17 and Microsoft Excel. Statistical practices include confidence interval determination via bootstrapping, *z* tests, one-way sensitivity analyses, probabilistic sensitivity analysis (PSA) via Monte Carlo simulations. A description of each statistical practice and the data analysis plan for each research question is listed below.

Confidence Intervals / z-test for ICER. Determining the confidence interval for an ICER is not trivial, as the ICER is a ratio and the uncertainly around the point estimate can vary based on either the change in costs or change in effects. Many methods for determining the confidence intervals from the ICER have been discussed in the literature. I used a bootstrapping method where 1000 resamples of the analysis are performed and then the 97.5% and 2.5% brackets are translated to the 95% confidence intervals. This technique has held up as a good way to estimate the ICER for a CEA (Ming-Yu & Xiao-Hua, 2007).

The analysis of bootstrapping also provides standard deviations. Assuming that the distribution for the ICER is normal, and with a known standard deviation, a z-score and p-value for significance can be used to test the hypotheses (Altman & Bland, 2011).

One-way sensitivity analysis. One-way sensitivity analysis, or univariate sensitivity analysis, is a form of deterministic sensitivity analysis and is a way to determine how the uncertainty of the model inputs affects the model outputs by changing one variable at a time and examining the results compared to the base case. This kind of analysis is useful to determine the impact of any one variable on the outcome of the study (Andronis, Barton, & Bryan, 2009).

Probabilistic sensitivity analysis. Probabilistic sensitivity analysis (PSA) is a way to look at multiple variables in the model all at once by assigning a distribution to each parameter and drawing a random value over multiple simulations. PSA is useful to test the robustness of a model over a wide range of possibilities (Andronis et al., 2009).

Monte Carlo simulations. Monte Carlo simulations are useful for normalizing variation by looking at many factors whose outcomes are uncertain, and may be run in parallel with PSA. These mathematical, computer-driven simulations are used frequently when groups of patients are being compared to each other versus individual patients (Halpern, Weinstein, Hunink, & Gazelle, 2000). Using this kind of analysis is consistent with current literature on the topic, and will allow for comparison of this study to others to help advance the knowledgebase of this area or research.

Research Question 1. What is the cost-effectiveness of IVF paired with PGS in one cycle compared to three consecutive cycles of IVF in achieving a similar live birth rate?

Null Hypothesis ($H_0$1): There was no statistically significant difference in the cost-effectiveness of proceeding to three cycles of IVF without PGS in achieving a live birth than one cycle of IVF with PGS. This hypothesis will be tested using a z-test for comparison. If $p<.05$ for the z-test for the ICER, then the null hypothesis will be rejected and the alternate hypothesis will be accepted.

Alternative Hypothesis ($H_1$1): There was a statistically significant difference in cost-effectiveness in utilizing one cycle of IVF with PGS in achieving a live birth than three cycles of IVF.

The main outcome of the data analysis was the cost-effectiveness ratio (CER) of each intervention, which is the cost per live birth. In order to make this calculation, the total costs involved with one cycle of IVF were determined:

Total Single Cycle Costs (TSCC) = Procedure Costs (PC) + Fresh Embryo Transfer Costs (FT) + Frozen Embryo Transfer Costs (FET) + Dilation and Curettage Costs (D&C) + Preimplantation Genetic Screening Costs (PGS)

D&C costs are only applied if the outcome is a clinical miscarriage. For the IVF-PGS intervention, the CER of costs per live birth was calculated as:

CER = TSCC / Clinical probability of live birth using PGS (LB_CP_PGS)

For the 3xIVF intervention, the CER was calculated as:

CER = (3x(TSCC-PGS)) / Clinical probability of live birth at 3rd cycle (LB_CP_IVF3)

The research question was answered by comparing the two cost-effectiveness ratios and determining which intervention provided the same clinical outcome for the least cost, also known as the Incremental Cost Effectiveness Ratio (ICER). This was calculated as the difference in total costs between strategies divided by the difference in live birth rates between the two strategies. For the ICER, 95% confidence intervals and variability around the ICER were reported from the simulations as per best modelling practices (Briggs et al., 2012). Z-statistics were calculated from the confidence intervals along with p-values for significance. The formula for this is:

$ICER = C_2 – C_1 / E_2 - E_1$

Where C_2 = Total costs of IVF-PGS

C_1 = Total costs of IVFx3

E_2 = Clinical probability of a live birth for IVF-PGS

E_1 = Clinical probability of a live birth for IVFx3

A series of univariate sensitivity analyses was performed to test the impact of different variables. First, the clinical probabilities for live birth for each intervention were modelled with a value of ±25% of each branch of baseline value (women under age 40). Secondly, costs were also varied from ±25% of baseline value for each branch. Results were reported in tabular form and presented as a Tornado graph.

To resolve parameter uncertainty, PSA was performed using a distribution for clinical probabilities and costs. Distribution for each variable was determined to be

Dirichlet distribution for clinical probabilities and a lognormal distribution for costs. Further, 1000 Monte Carlo simulations were used to validate the analysis over multiple scenarios.

Research Question 2. What are the statistical differences between cost-effectiveness among three different age groups of women in each treatment arm?

Null Hypothesis ($H_0$2): There was no statistically significant difference based on z-test in the cost-effectiveness of between three different age groups of women using either intervention.

Alternative Hypothesis ($H_1$2): There was a statistically significant difference based on a z-test in cost-effectiveness in one or more age groups using one or the other intervention.

Similar to the first research question, we examined the CER for each intervention stratified by age group. We then applied confidence intervals for each group's ICER and determined the *z* test and *p* value for significance. If the *p* value for any one of the age group's ICER is < .05, we can accept the alternate hypothesis and reject the null. The age groups were

- Women aged less than 40
- Women aged 40-42
- Women aged older than 42

The CER for each age group and each intervention provided six unique CERs, which were calculated in the same manner as described for research question #1. We then compared CERs by age group and intervention to determine the incremental cost

effectiveness ratios. We applied the same statistical tests used for the first research question, including univariate sensitivity analysis adjusting for live birth probability and costs. We also ran Monte Carlo simulations with PSA over 1000 iterations to determine the robustness of the model.

Other secondary analysis points include a willingness-to-pay (WTP) threshold defined as $100,000 per QALY, or in this case, per live birth. A WTP threshold is a benchmark used in CEA to determine a society's willingness to pay for improvements in health. It was previously set as $50,000 per QALY, based on years-old analysis of one aspect of healthcare. However, many have argued that $100,000 per QALY is an appropriate benchmark for today's CEA and that benchmark has been used in recent CEA involving IVF and PGS (Murugappan et al., 2015; Neumann, Cohen, & Weinstein, 2014). By defining the benchmark, the most effective intervention were compared to others using a similar yardstick.

Threats to Validity

The validity of cost-effectiveness models in intervention-based studies has been examined. O'Brien (1996) identified seven major threats to validity in CEA studies involving multiple interventions or head-to-head comparisons. These are still applicable today, and are relevant to this study. Table 3 describes each threat and the mitigation of those threats in this study.

Table 3

Threats and Mitigation

Threat[a]	Threat description[a]	Mitigation

Choice of comparison therapy	Ensuring that all available therapies are evaluated	The two interventions were determined by the gap in the literature and need for the analysis
Gold standard measurement of outcomes.	Some clinical trial do not allow for real-world clinical care outcomes	For IVF studies, live birth as the main clinical outcome ensures this is not a threat.
Intermediate rather than final health outcomes	Outcomes evaluated at a mid-point in the care continuum	As live birth is the main clinical outcome, this mitigates the threat.
Inadequate patient follow-up or sample size	Some clinical trials do not follow the patient through to the final outcome and miss certain economic factors	Sample size is adequate, as described elsewhere in this chapter. Further, costs associated with adverse events, such as miscarriage, were taken into account.
Protocol-driven costs and outcomes	Clinical trials where protocol specific costs are incurred	All costs associated in this study are from the regular routine treatment of patients, and not some specific protocol.
Geographical transferability of trial evidence	Costs and clinical care differ in various parts of the country and need to be assessed	Geographic differences in prices were mitigated through using cost data from multiple sources in different locations across the country.
Selected patient and provider populations	Hand-picking patient populations and thereby bias the study	By utilizing the literature and multiple sites for cost data, patient and provider populations were adequately managed.

[a.] From "Economic evaluation of pharmaceuticals. Frankenstein's monster or vampire of trials? " O'Brien, B. 1996. *Medical Care*, 34(12 Suppl), DS99-108.

Further threats to validity are detailed in the recent International Society For Pharmacoeconomics and Outcomes Research Task Force update to recommendations (Ramsey et al., 2005), originally published in 2005. The task force relayed five major threats to validity when using data driven by clinical trials or other studies. They include

protocol-driven resource use, comparators that are uncommon or not recommended in clinical practice, non-representative recruiting, restrictive inclusion and exclusion criteria, and artificially enhanced compliance (Ramsey et al., 2015).

Of these threats, this study is particularly susceptible to recruiting issues between interventions. The cumulative (3xIVF) arm uses over 150,000 women in determining the live birth rate, whereas the IVF-PGS arm used just over 1,900 women in the studies used to determine clinical effectiveness. However, because the sample size calculations are statistically significant, this threat is generally very low.

Ethical Considerations

Ethical considerations in any research study are essential to ensure no harm comes from the research itself. In this case, no human participants were directly used to determine any portion of the analysis. Further, all cost data from private clinics were obtained by permission using Data Use Agreements (DUA) that strictly define the data use and sources. The DUA ensure that no Personal Health Information (PHI) was used or revealed and also ensures that the research collaboration between the sources and users of the data do not breach any part of the Health Information Portability and Accountability Act (HIPAA) statutes. The study was reviewed and approved by Walden university Institutional Review Board – approval number 07 26 17 044 33 96.

Summary

The study design was a CEA using a decision-analytical tool to model a theoretical interventional comparison using secondary data sources. The two interventions studied are patients that undergo IVF in conjunction with PGS for one cycle

versus patients that undergo three consecutive cycles of IVF without PGS. For each intervention, costs and clinical probabilities at each node were used. For costs, multiple sources from the recent literature and two independent IVF clinics were used to determine direct costs. For clinical probabilities, the literature was used to determine the most relevant clinical probabilities of each intervention. The sample size from the literature and costs set were robust enough to avoid sampling errors. A benchmark of $100,000 per live birth was used as the willingness-to-pay threshold for comparison purposes.

All CEAs are vulnerable to threats to validity, mainly from the sources of the data and the variables examined. Most of these threats were mitigated through careful examination of the material and using multiple sources of data. Because the model is theoretical, no human participants were used and therefore ethical considerations were at a minimum. All actual data from clinics were de-identified of all personal health information before used in the research.

Chapter 4 will involve a description of the findings of the study. It will include the statistical results and describe the data that was included in the analysis. It will also present information related to each arm of the decision model.

Chapter 4: Results

Introduction

The purpose of the study was to examine two interventions in IVF and determine the cost-effectiveness of each in achieving a live birth. The two interventions were IVF using PGS in one cycle, and the other was three, consecutive cycles of IVF. The first research question was whether there was a statistical difference in the cost-effectiveness ratios of each intervention. The second was whether there were statistical differences by age group, specifically women under age 40, women aged 40-42, and women aged over 42. I hypothesized that there was a statistical difference overall for one cycle of IVF with PGS versus three cycles of IVF.

In Chapter 4, I will present the data collection methods and the values used for both costs and clinical probabilities in the cost-effectiveness model. Further, I will explain the statistical methods used to analyze the data and test the robustness of the model. Lastly, the results of the model will be presented.

Data Collection

The key secondary datasets used in this study were costs and clinical probabilities. I obtained clinical probabilities from two sources. The first source was the published literature, and the second source was from publicly available datasets provided free of charge by the CDC and fertility clinic websites. Concerning costs, secondary data sets used were from multiple sources, including the literature, public sources, and from one private clinic. The next three tables detail the percentages used and the literature

sources used in the model. Table 4 details the clinical probabilities for cumulative live births.

Table 4

Cumulative Live Birth Rate Probabilities by Age, Prognostic-adjusted Estimate (%)

Cycle	Women aged < 40 years	Women aged 40-42	Women aged > 42
1st Cycle	32.3	12.3	3.7
2nd Cycle	48.7	19.8	6.3
3rd Cycle	58.0	24.7	8.3

Note. Adapted from "Live-Birth Rate Associated with Repeat In Vitro Fertilization Treatment Cycles," by Smith et al., 2015, *Journal of the American Medical Association, 98*, p. 2659.

Table 5 describes the clinical probabilities of pregnancy, live birth, and miscarriage used in the final model for the 3xIVF intervention arm, stratified by age group. The calculation of the averages were described in detail in Chapter 3.

Table 5

Pregnancy, Live Birth, and Calculated Miscarriage Rates by Age, (%)

Rates	Women aged < 40 years	Women aged 40-42	Women aged > 42
Pregnancy rate[a]	50.3	35.6	20
1st Cycle live birth rate[b]	32.3	12.3	3.7
2nd Cycle live birth rate[b]	27.1	10.1	3.3
3rd Cycle live birth rate[b]	24.3	8.6	3.3
1st Cycle miscarriage rate[c]	18.0	23.3	16.3
2nd Cycle miscarriage rate[c]	23.2	25.5	16.7
3rd Cycle miscarriage rate[c]	26.0	27.0	16.7

[a] Averaged for all ages and transfer type, from Centers for Disease Control and Prevention, ART Clinical Success Rates, 2016, p. 21. [b]From "Live-Birth Rate Associated with Repeat In Vitro Fertilization Treatment Cycles," by Smith et al., 2015, *Journal of the American Medical Association, 98,* p. 2659. [c]Calculated.

Table 6 describes the pregnancy rates, live birth rates, and miscarriage rates as used in the CEA models for the IVF-PGS intervention. Again, the description of the calculations are detailed in Chapter 3.

Table 6

Pregnancy, Miscarriage, and Live Birth Rate Probabilities by Age with PGS, (%)

	Women aged < 40 years	Women aged 40-42	Women aged > 42
Pregnancy rate[a]	82.0	47.8	28.1
Live Birth rate[b]	63.2	33.3	8.1
Miscarriage rate[c]	18.8	14.5	20

Notes: [a]Calculated. [b]Averaged for women aged less than 40, from "Preimplantation Genetic Screening is cost-effective in cost per delivery compared to routine in vitro fertilization." B. Hodes-Wertz, D. H. McCulloh, & J. Grifo. 2015. *Fertility and Sterility, 104, vol. 3,* p.e278. [c]From "IVF patients over age 39 experience decreased cost effectiveness and live birth rates with preimplantation genetic screening: a decision analytic model and cost effectiveness analysis," by Salem et al., 2016, *Fertility and Sterility, 106, vol. 3,* p.e61.

Table 7 details the costs used in the model by cost type. Costs were derived from secondary sources, including the literature, one private clinic, and public websites.

Table 7

Costs per procedure, in 2016 US Dollars ($)

Cost description	Mean	Standard Deviation	Minimum	Maximum	95% CI
Routine IVF cycle (includes oocyte retrieval, ICSI, and fresh embryo transfer, and cycle medications)	20,548	3,817	17,755	26,058	16,807-24288
Frozen embryo transfer	5,108	983	4,000	6,395	4,145-6,072
PGS costs (includes procedure, shipping, and reporting)	4,836	1,224	3,450	6,225	3,637-6,035
Clinical dilation and curettage (D&C) associated with miscarriage	2,517	1,679	1,304	5,000	872 - 4,163

Notes: Costs obtained from Murugappan et al., 2015, Private Clinic by permission, and public websites.

Results

Two research questions were asked:

RQ1-Quantitative: What is the cost-effectiveness of IVF paired with PGS in one cycle compared to three consecutive cycles of IVF in achieving a live birth?

The incremental cost-effectiveness for all women was ($2,657,388), with a confidence interval of ($3,087,000) – ($1,958,045), $p < .05$. A negative ICER indicates that IVF-PGS is the preferred strategy across all age groups. The CI does not include the null ($0), and the *p* value is less than .05 for the *z* test, and therefore the null hypothesis can be rejected and the alternate is accepted.

RQ2-Quantitative: What are the statistical differences in means between cost-effectiveness among women aged under 40, women aged 40-42, and women aged over 40 of women in each treatment arm?

For each age group, the ICER was examined and confidence intervals determined. The *z* score for each age group and *p* value was determined. The results of the initial CEA are that in the age groups of women aged under 40, and aged 40-42, the IVF-PGS strategy dominated the IVFx3 strategy for cost-effectiveness. However, neither strategy was dominant for women aged over 42. Concerning RQ2, there were significant ($p < .05$) statistical differences between age groups for ICER, and therefore we can reject the null hypothesis and accept the alternate hypotheses. The results show that IVF with PGS was a more cost-effective strategy for all age groups except for women over 42. Table 8 details the costs per live birth and ICERs per age group and strategy.

Table 8

Cost per Live Birth per Strategy, Age Group, and ICER ($)

	All ages	Women aged < 40 years	Women aged 40-42	Women aged > 42
IVFx3	366,630.41	90,770.30	250,799.64	758,321.30
IVF-PGS	174,770.57	48,995.85	92,662.17	382,653.70
ICER	-2,657,388	-4,620,123.31	-731,317.69	2,402,310.35
ICER 95% CI	-3,087,000 – -1,958,045	-9,405,548- -1,427,930	-813,727 - -611,649	2,252,176 – 2,683,557
Z score	-9.22	-2.27	-14.19	-21.83
One-tailed *p* value	.00001	.012	.00001	.00001

Cost-effectiveness graphs were plotted against the willingness-to-pay threshold for each age group. These graphs plot the costs per strategy against the effectiveness. If two strategies are undominated (or in other words, neither is preferred or more cost-effective), then those strategies are connected by line segments which form the cost-effective frontier. Only the strategies on the cost-effective frontier could be the optimal choice. The lowest cost option is always part of this frontier; if it dominates all comparators, the graph will have no lines. This is the case for the first two age groups. The willingness-to-pay slope intersects with the favored strategy. For each scenario, the graphs plot which strategy was undominated, or preferred strategy compared to the dominated strategy. Figures 3, 4, and 5 show the results of each.

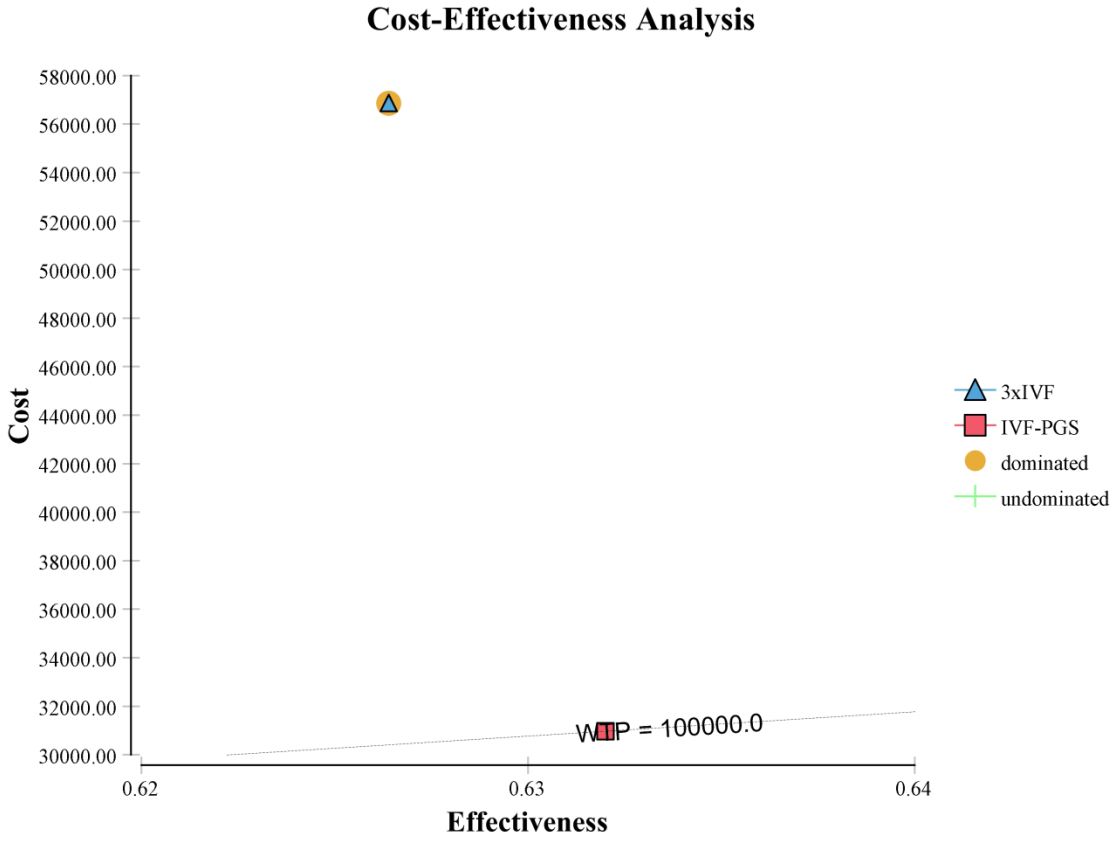

Figure 3. Cost v. effectiveness for women aged under 40 in achieving a live birth.

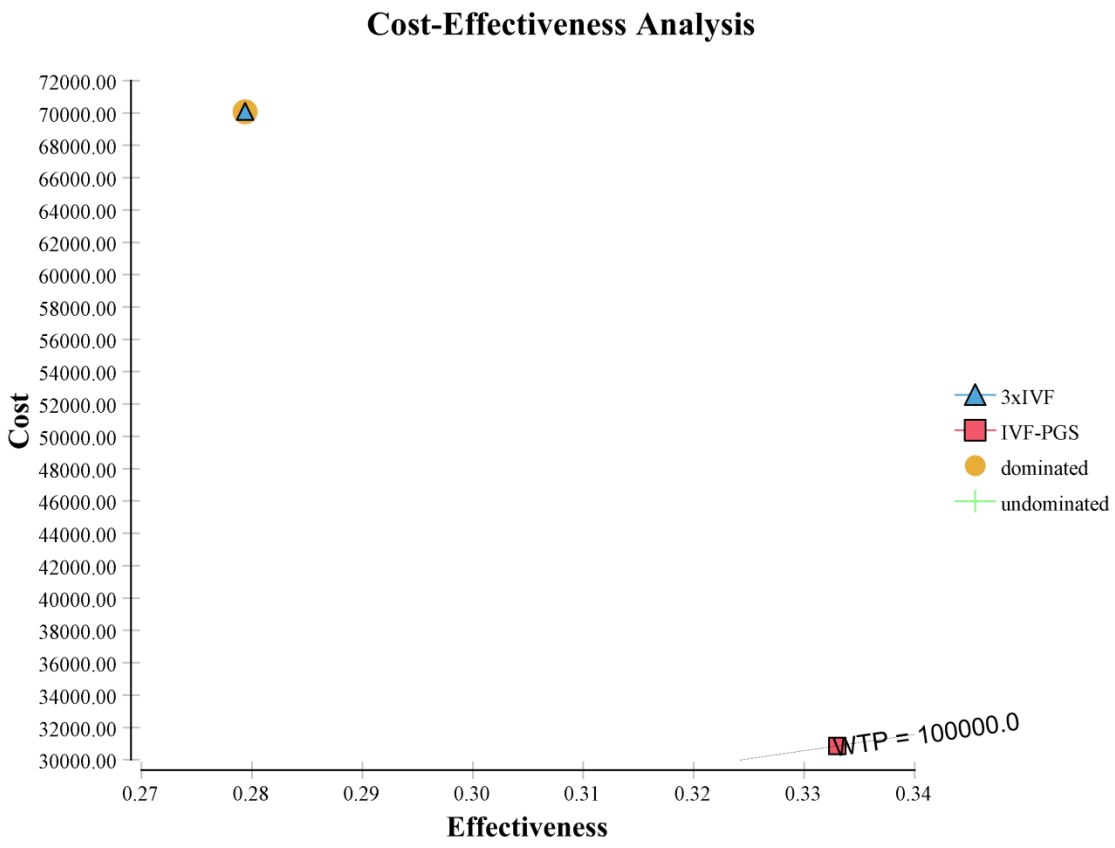

Figure 4. Cost v. effectiveness for women aged 40-42 in achieving a live birth.

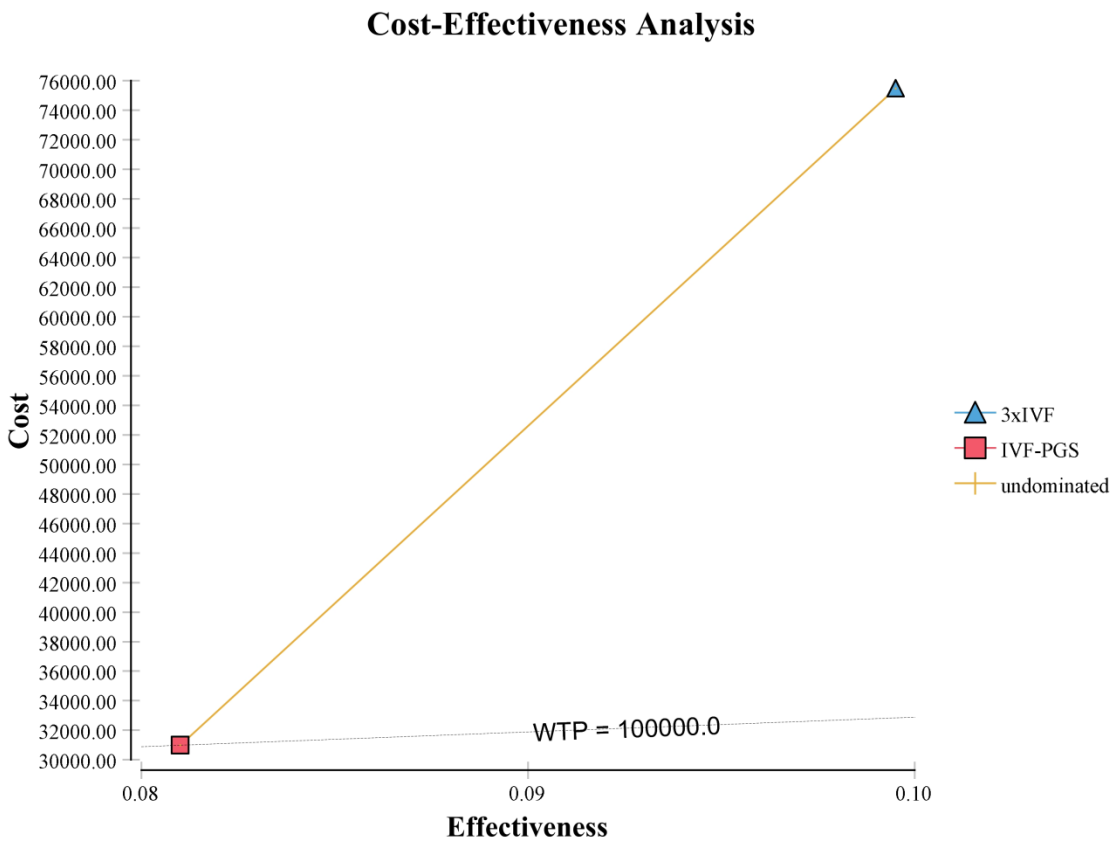

Figure 5. Cost v. effectiveness for women over 42 in achieving a live birth.

Univariate Analysis

Univariate analysis was performed from baseline to examine the effect of live birth rate and cost by strategy. The baseline model was women aged under 40. Live birth rate was varied by + or – 25% for each branch, and then ICERs were re-examined. Further, costs were varied by the same amount for each branch to determine effect of the variation. Results are shown in Figure 6 as a tornado diagram, where the expected value is the base case ICER of women aged under 40.

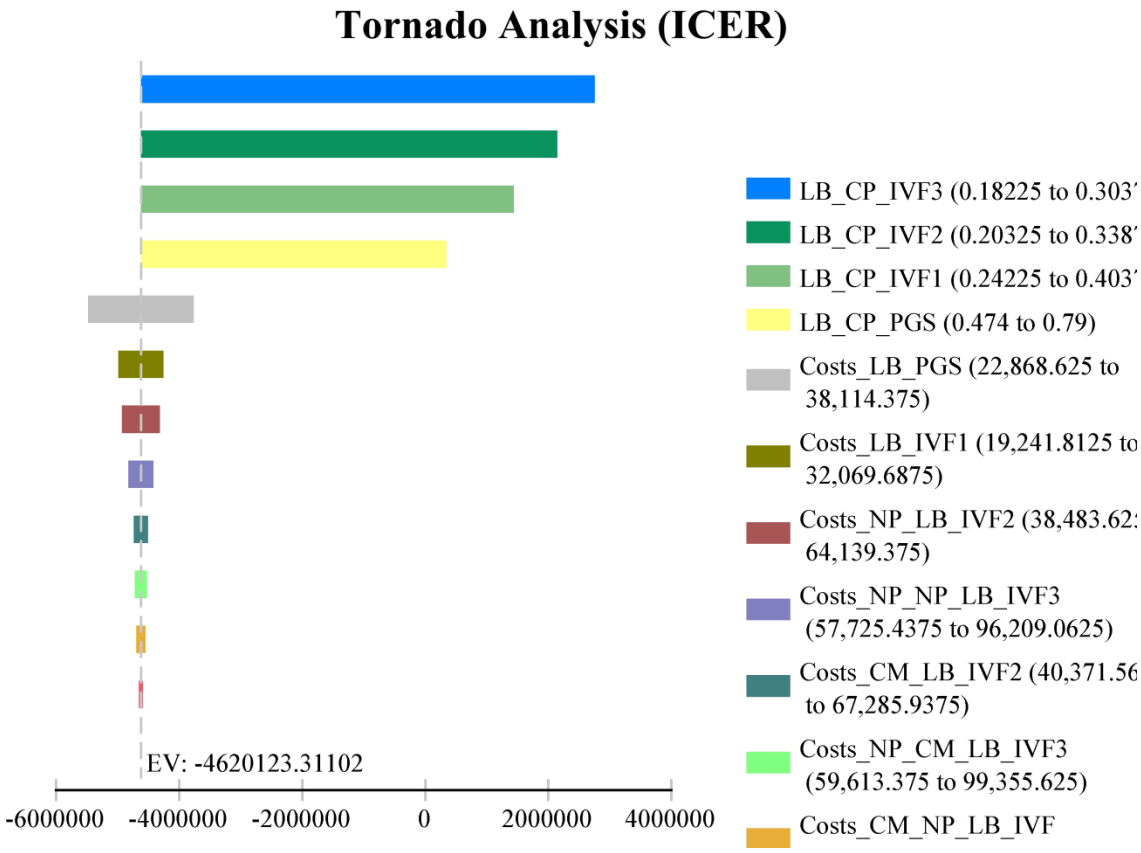

Figure 6. Tornado Diagram of univariate sensitivity analysis. *Notes:*

LB_CP_IVF3 – Live birth probability in the third cycle

LB_CP_IVF2 – Live birth probability in the second cycle

LB_CP_IVF1 – Live birth probability in the first cycle

LB_CP_PGS – Live birth probability in one cycle with PGS

Costs_LB_PGS – Costs to achieve a live birth in one cycle with PGS

Costs_LB_IVF1 – Costs to achieve a live birth in first cycle

Costs_NP_LB_IVF2 – Costs to achieve a live birth in the second cycle of IVF where the first cycle resulted in a no pregnancy

Costs_NP_NP_LB_IVF3 – Costs to achieve a live birth in the thirs IVF cycle where the first two cycles both resulted in no pregnancies

Costs_CM_LB_IVF2 – Costs to achieve a live birth in the second cycle where the first cycle resulted in a clinical miscarriage

Costs_NP_CM_LB_IVF3 – Costs to achieve a live birth in the third cycle where the first cycle resulted in a no pregnancy and the second cycle resulted in a clinical miscarriage

Costs_CM_NP_LB_IVF – Costs to achieve a live birth in the third cycle where the first cycle resulted in a clinical miscarriage and the second cycle resulted in a no pregnancy

The results of the univariate analysis show that the clinical probabilities have the largest effect on ICER, starting with the clinical probability of a live birth in cycle number 3 for the IVFx3 strategy. Costs, however, did not have a considerable impact on the expected value.

PSA / Monte Carlo Simulations

Secondary analysis of each model for each age group included Monte Carlo simulations (PSA) to determine the robustness of the model. A Dirichlet distribution was used for the clinical probabilities (to achieve a sum of .1000 for the three branches), and a lognormal distribution was used to model the costs using the mean and median. 1000 simulations were run yielding a mean of each model that was within one standard deviation of the original model.

The statistics for each age group are shown in Table 9. For the first two age groups, the IVF-PGS strategy dominated at a WTP of $100,000 and was dominant for women aged 40-42 in all scenarios. However, for women aged under 40, and as WTP increases, the model shows that IVFx3 could begin to become a preferred strategy, with preference defined as the most cost-effective strategy in the majority of iterations. Even at a WTP of $3M, IVF-PGS was the preferred strategy 76% of the time (see Figure 7).

For women over 42, IVF-PGS was the dominant strategy at a willingness-to-pay of $100,000. In other words, if $100,000 is the amount one was willing to pay to achieve a live birth and could not spend over that amount, then IVF-PGS would give you the best probability to achieve a live birth. However, IVFx3 is the dominant strategy around $2.4M, and if the WTP increases to $3,000,000, then IVFx3 is the dominant strategy 61.2% of the time. So, if the goal is to achieve a live birth irrespective of a willingness to pay threshold, then IVFx3 would be the preferred strategy. Figure 8 shows the preferred strategy statistics for women aged 40 to 42 with a WTP of $3M and figure 9 represents the same data for women over 42.

Table 9

Monte Carlo Simulation Results, Costs per Live birth in 2016 US Dollars ($)

CER	Mean	Standard Deviation	90% CI
Women aged under 40			
IVF-PGS	49,747.62	7,567.13 / .02	35,940.98 – 67,448.21
IVFx3	90,218.05	4,131.48 / .01	81,385.64 – 102,082.82
Women aged 40-42			
IVF-PGS	93,209.91	3,856.99 / .01	76,264.17 – 115,217.94
IVFx3	249,967.25	2,018.23 / .01	225,086.63 – 278,896.11
Women aged over 42			
IVF-PGS	387,430.75	89,660	333,234.11 - 460,074.43
IVFx3	754,703.10	87,211	676,497.64 - 850,820.67

Note. Standard deviation depicted as standard deviation of costs / standard deviation of effectiveness.

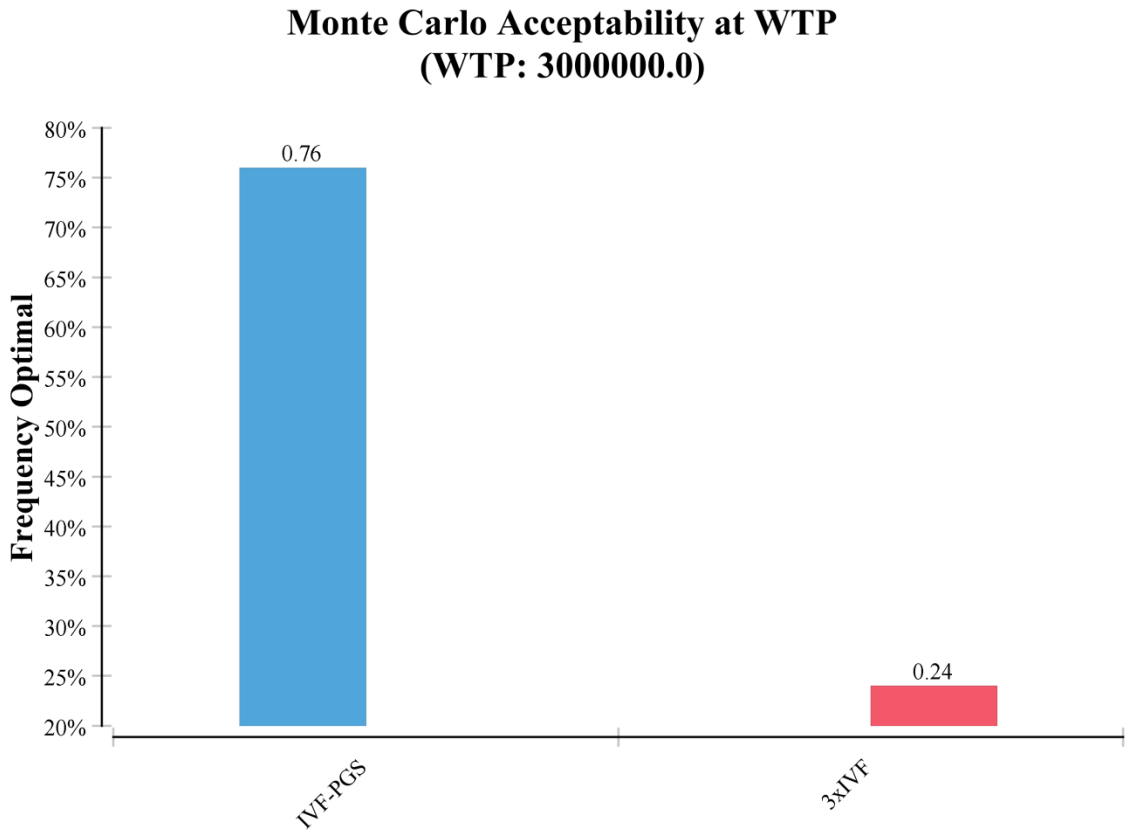

Figure 7. Preferred strategy at a WTP of $3M for women aged under 40.

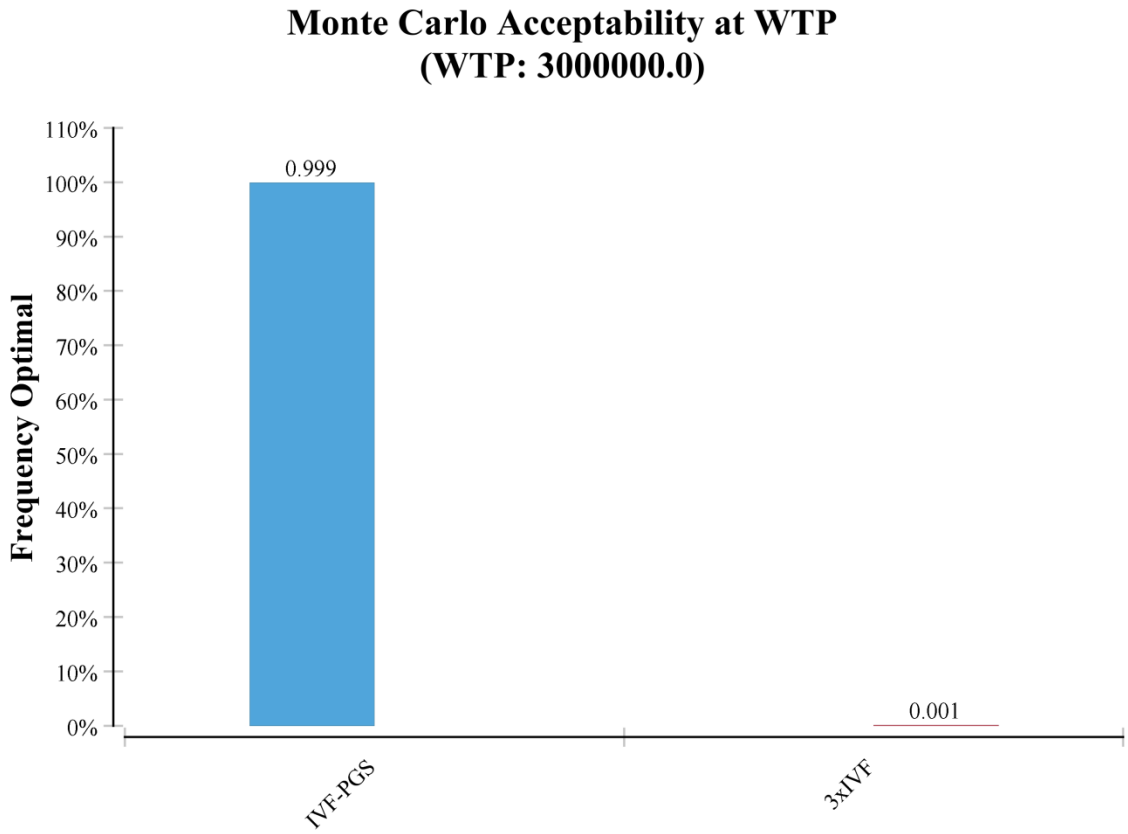

Figure 8. Preferred strategy at a WTP of $3M for women aged 40-42

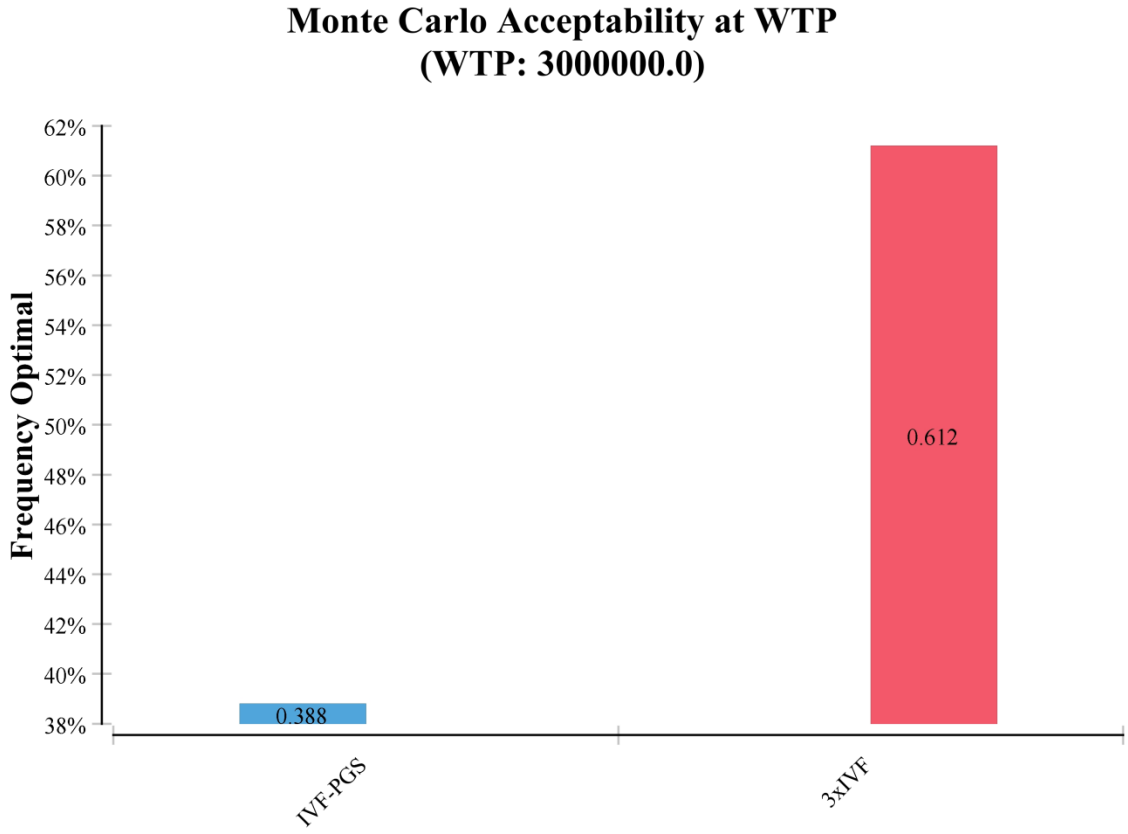

Figure 9. Preferred strategy at a WTP of $3M for women over 42.

Summary

Data was collected from the literature, public sources, and one private clinic. The data collected was costs and clinical probabilities for the cost-effectiveness model. The research questions were whether there were statistical differences in cost-effectiveness between strategies of one cycle of IVF with PGS versus three cumulative cycles of IVF without PGS, and if there were differences by age group. The results show negative ICERs for all age groups in aggregate, and the age groups of women aged under 40 and women aged 40-42. The results also show that neither strategy was dominant for women over 42, with a positive ICER of around $2.4 million. For both research questions, z-

scores provided p-values less than .05 based on confidence intervals around the ICERs for each age group. These results confirm that we can reject the null hypothesis and accept the alternate hypothesis for both research questions.

Univariate analysis was performed on a base case of women aged under 40. All costs and clinical probabilities were varied by plus or minus 25%. The results show that clinical probabilities of a live birth for the IVFx3 arm provide the most variability in the expected ICER value. Further, Monte Carlo simulations were performed to determine the robustness of the model with Dirichlet distributions used for the clinical probabilities and lognormal distributions used for the costs. For women aged 40-42, no change to the dominance of IVF-PGS as a cost-effective option was observed, regardless of willingness-to-pay. However, for women aged under 40 and women over 42 age group, IVF-PGS was not always dominant and was not the preferred strategy when there were no willingness-to-pay limits.

I will discuss these results in more detail in Chapter 5. The interpretation of the results will be discussed and the social change implications will be discussed. Further, the limitations of the study and recommendations for future research will be examined. Lastly, the study will be summarized and concluded.

Chapter 5: Discussion, Conclusions, and Recommendations

Introduction

In this study, I examined a cost-effectiveness model where two strategies were compared to determine the most cost-effective intervention to achieve a live birth. The two strategies were IVF in one cycle paired with PGS and IVF three consecutive times without PGS. I examined three different age groups: Women aged under 40, women aged 40-42, and women aged over 42.

The results show that IVF with PGS was the dominant strategy for women aged under 40 and women aged 40-42. However, the ICER for women aged over 42 was positive, although no dominant strategy could be determined. Further, univariate analysis of the base case of women aged under 40 showed that varying the clinical probability of live birth in the IVFx3 arm had the most impact to the expected value.

Lastly, PSA and Monte Carlo simulation showed that the findings were consistent across age groups. For women aged under 40 and aged 40-42, IVF paired with PGS was the dominant strategy at any willingness-to-pay (WTP) threshold. However, for women aged over 42, IVF with PGS was preferred at a lower WTP, but IVF performed three times consecutively was preferred at higher price points to achieve a live birth.

This chapter will include interpretation of these results along with the implications for clinical application and society in general. I will also discuss the limitations of the study and make recommendations for future research.

Interpretation of the Findings

The results of the study returned a statistical difference between strategies for cost-effectiveness and also showed differences between age groups. PGS paired with IVF alone is definitively more cost-effective than three consecutive cycles of IVF for women aged 40-42 and in the vast majority of cases for women aged under 40. For women aged over 42, no strategy was dominant and many factors could come into play to determine the best course of clinical action, including factors that go into the protection motivation theory.

Looking first at the ICERs, the results showed that for women aged under 40 and aged 40-42, cost per live birth was much less for IVF with PGS, and the effectiveness was better than with IVF alone performed three times consecutively. For women aged over 42, no strategy was dominant as the IVFx3 arm was more effective but more costly in achieving a live birth. This is consistent with some recent studies on CEA in IVF and PGS, where PGS is seen as cost-effective for younger women but not in higher age groups (Hodes-Wertz et al., 2015).

The results of this study point clearly to the clinical probability of the interventions as having the most impact on cost-effectiveness. Looking at the literature, in many studies where PGS was not cost-effective, researchers used a much lower clinical probability for PGS with IVF than the alternate intervention. For example, Mersereau et al. (2008) concluded that for women aged under 40, PGS was less cost effective than IVF with PGS in just one cycle, but that for women aged over 40, the cost-effectiveness was the same. Those findings were made prior to the advent of new PGS

technologies (PGS#2). Further, the clinical probability of PGS in the IVF-PGS arm of that study was 21.7% for both women aged under 40 and women aged over 40, where the IVF alone live birth rates were 37.8% and 20.7%, respectively. The IVF-PGS arm of our study for women aged under 40 was 63% based on the most recent studies using PGS#2 techniques, but for women aged over 40, it was an average of 20.7% also. Univariate analysis shows that a live birth rate as low as 48% for the IVF-PGS arm of our study would move the ICER to positive.

Other contrary findings that showed that PGS is not cost-effective in any age group used were based on very low clinical probabilities when evaluating PGS. Kushnir et al (2016) used a PGS live birth rate per cycle of 17.7% for women over 37. By comparison, I used a live birth rate of 33.3% for women aged 40-42 in this study. Our results for that age group showed that almost 100% of Monte Carlo simulations that IVF paired with PGS were more cost-effective than IVF alone in three cycles. This finding is consistent with specific studies for women aged 40-43. Lee et al. (2015) observed a live birth rate for that age group of 45.5%, much higher than the clinical probability used in this study.

For women aged under 40, IVF paired with PGS is the most cost-effective approach in the majority of cases. The findings most likely conflate multiple age groups under 40 to explain the fact that IVF performed three consecutive times was preferred in nearly a quarter of Monte Carlo simulations, as it is clear from almost any study involving IVF that the single-most important factor to achieving a live birth is maternal age (Sunderam et al., 2015).

The result that IVF performed three consecutive times is more costly but more effective in achieving a live birth for women aged over 42 complicates the discussion about how best to treat these patients. As discussed previously, IVFx3 was the dominant strategy around $2.4M, and if the WTP increases to $3,000,000, then IVFx3 is the dominant strategy 61.2% of the time. Where money is no object, clearly women should attempt IVF three times consecutively to achieve a live birth. Women of a higher socioeconomic class, or with higher incomes, may potentially choose to proceed with three cycles of IVF because their willingness to pay for a better outcome is higher than other women. Where the patient is the payer, which is usually the case in the United States, this result can be linked to higher incomes affecting the overall cost-effectiveness. Also, where insurance covers IVF up to that number, for example, but does not cover PGS, IVFx3 may be the right approach for this cohort. This finding is contrary to a recent study that showed that PGS#2 was cost-effective across all age groups (Resetkova et al., 2013). However, that study also averaged costs and probabilities across the age spectrum, and did not distinguish age groups.

Researchers in many of the CEA studies to date have compared the same number of cycles in each arm, and direct comparisons are difficult to make. Intuitively, one would imagine that performing the same procedure three times compared to that procedure paired with screening would obviously be more costly, which was true in all three scenarios. Costs per live birth were far higher for IVFx3 compared to IVF with PGS alone. However, the clinical probabilities of achieving a live birth is what drives the

overall analysis, and is what most certainly will drive the debate over the use of PGS in routine clinical practice.

For women over 42, where the cost-effectiveness is not as clear per strategy, other factors related to the protection motivation theory could be examined. These include the threat appraisals and coping appraisals of other parts of the IVF process, excluding costs. These could be the apprehension of not achieving a live birth, the fear of a prolonged timeline, and the emotional toll such a process could have. While costs and economic factors play into the PMT, they are not the only things to consider when determining if a measure like genetic screening should be employed (Helmes, 2002).

As mentioned in Chapter 2, researchers have reported a wide variation of the cost per live birth using IVF, anywhere from $19,416 to as high as $291,907, depending on age group and PGS status (Chambers et al., 2009; Hodes-Wertz et al., 2015; Resetkova et al., 2013). My study found a one-cycle cost per live birth with PGS to be $48,995.85 for women under the age of 40 and $382,653.70 for women over the age of 42. This is fairly consistent with the literature in showing a wide range between age groups, and is generally consistent with the costs per live birth based on what costs were included in the study.

Lastly, in this study I attempted to examine a real life dilemma for patients attempting to achieve a live birth using IVF. The choice to employ genetic screening or simply stay the course through multiple cycles is a common one for many patients. I employed the most recent findings for cumulative success rates in IVF and for IVF paired with PGS, discriminated by age group. The results highlight the debate around PGS but

also advance the body of evidence of PGS as a useful clinical tool in certain clinical circumstances. The findings have both limitations and implications for social change.

Limitations of the Study

The study had multiple limitations that could have prevented a more accurate picture of the cost-effectiveness of the two interventions studied. First, the model is simply theoretical, and used secondary data sources from the literature and other public sources to build a model that explores these interventions. I believe that the model is very representative of the most current costs and probabilities for the United States. The literature I used to populate the probabilities, particularly the Smith et al. (2015) that had hundreds of thousands of data points, was well designed and used the most current technologies. Any limitation found in the collection of the data from other sources is inherent in the model. Particularly, the fact that collection of the data from Smith et al. was done in the United Kingdom, not the United States. All other literature sources were from the United States and the CDC data is from the US. All costs were in 2016 dollars and so should not have created any bias.

Further, certain assumptions were made to make the model as usable as possible. For example, it did not explore the costs associated with multiple births, which is a major issue with IVF (Kulkarni et al., 2013). Comparing the rate at which each arm produced multiple gestations and the costs associated with multiples could have changed the analysis. We also assumed that every transfer was a single-embryo transfer, which is not the usual routine practice of most fertility physicians today (De Sutter, Gerris, & Dhont, 2002).

Further, the study made certain assumptions about fresh and frozen transfers and the use of ICSI, and also included the costs of medications which is typically not seen in many of the CEA studies for IVF and PGS (Dahdouh et al., 2015). We also made assumptions for the clinical probabilities for the cumulative IVF group based on certain drop-out rates between cycles.

The study also did not examine any other outcomes of an IVF cycle such ectopic pregnancies or elective abortions. Other assumptions of the model were that every woman in every cohort had a euploid embryo after getting PGS, which is not always the case, especially in older women (Hodes-Wertz et al., 2012).

Recommendations

Recommendations based on the findings of this study are threefold. First, a prospective, multi-center trial comparing IVF in one cycle with PGS and IVF performed three consecutive times should be attempted. Such a trial should account for all of the real world factors and costs associated with a woman's journey to achieve a live birth, and should utilize the most recent technology for PGS (aCGH or NGS). The study could provide insight into PGS live birth rates based on varying age groups and disease states and could use real-world costs to determine cost-effectiveness (Orvieto, 2016).

Second, clinical societies such as the ASRM and ACOG that govern IVF should examine this study and others on the topic of PGS and make firm guideline recommendations based on multiple factors, including cost-effectiveness. The most recent guidelines from the ACOG are almost 9 years old now and do not reflect the

current technologies for PGS or the studies that examine the clinical efficacy and cost-effectiveness ("ACOG Committee Opinion No. 430," 2009).

Lastly, I recommend further studies into the economic impact of PGS and new genetic screening techniques on the health care system at large. One such study could be a cost-benefit analysis of how PGS impacts society at large and the overall economic impact of delivering healthy, single infants as a result of IVF. One impact of PGS not examined in this study is the prevalence of healthy infants without genetic defects. Children born with a genetic birth defect in the US cost the healthcare system over $23 Billion every year (Arth, 2017). New genetic screening tests could impact the economics of health care in general, but require more research to fully understand the effect.

Implications

The study has significant social implications based on the results, especially in the United States where the vast majority of patients pay their own way to get assistance in achieving a live birth. Many clinics in the United States have moved to novel payment techniques to keep patients in IVF longer, including guarantees on a live birth and one price for three consecutive cycles of IVF (Cha, 2017). Cost-effectiveness tools can help these parents and other stakeholders to make the best clinical decisions based on limited resources. The implications for social change mainly affect four different groups, discussed in order: patients, physicians (RE), insurers, and policy-makers.

For patients, having as much information as possible prior to pursuing IVF or ART is critical, especially given the competitive nature of many IVF clinics. Organizations such as Resolve (resolve.org) seek to help infertile women understand

pregnancy rates and live birth rates between different clinics. The CDC publishes data every year and seeks to educate patients about what clinics have the highest success rates and what those rates mean (Centers for Disease Control and Prevention, 2016). Genetic screening is a particularly difficult topic to understand for most patients and the onus of explaining such a difficult concept falls to the physician, or in some cases, genetic counselor (Hamilton et al., 2014). However, often times costs associated with genetic screening and the outcomes are not clearly articulated. The results of this study can help patients understand their options, particularly if costs are not an issue in situations where IVF is covered but PGS is not (Schwartz et al., 2017).

As discussed, reproductive endocrinologists are often the gatekeeper to genetic testing in the IVF clinic. They typically have the most information about the options available to a patient undergoing IVF. Understanding not only what the most recent literature says about the clinical efficacy of PGS but the cost-effectiveness by age group will enable physicians to educate their patients about the best options for them. Other factors associated with each intervention should also be factored into a decision that is ultimately made by the patient about their own treatment. This study helps to add to the available literature on PGS and can help inform practice guidelines or future research.

Insurance companies to date still do not cover IVF or PGS in the vast majority of states (Boulet et al., 2015). Many employers may seek to provide such coverage for their employees as an added benefit, and some insurance companies may provide it at a higher premium (Collura & Adamson, 2017). CEAs such as this one could inform financial

decision-making in order to best cover these patients and possibly add new patients under coverage.

A final implication of the results of the study are around the univariate analysis that indicates that the clinical probabilities are the key to cost-effectiveness of these interventions, particularly in later cycles of IVF. If IVF progressed to show improved clinical probabilities in the second and third cycles without PGS, the cost-effectiveness models would change greatly. Other techniques involved with IVF should be considered to improve those probabilities to impact the cost-effectiveness relationship between interventions. This also implies that further research needs to constantly take into account the most recent techniques involved with IVF in general, including vitrification, transferals, and IVF techniques.

Conclusion

Society has a vested interest in understanding the cost-effectiveness of different clinical interventions to achieve certain health benefits or outcomes. For assisted reproductive techniques, the technology involved is rapidly evolving and determining the best clinical interventions to achieve a live birth is difficult from the perspective of the patient. New technology can be useful in achieving the desired outcome but can be controversial as to whether it should be implemented in routine clinical practice. The cost-effectiveness of these interventions is one key factor that clinicians and patients should take into consideration as part of their decision-making.

This study sought to understand the cost-effectiveness of two different interventions used in the field of IVF. The first was using the controversial genetic

screening technique PGS in one cycle of IVF. The second intervention was in pursuing IVF without PGS in three consecutive cycles. These two interventions had never been compared for cost-effectiveness before and therefore the study was novel. The research questions asked were whether there was a cost-effectiveness difference between the interventions and in three different age groups.

Looking at the incremental cost-effectiveness ratios yielded results that allowed us to reject the null hypothesis for both research questions. For women aged under 40 years old, IVF paired with PGS was less expensive and more effective than the IVFx3 intervention. However in Monte Carlo simulations, the preferred method depended on the willingness-to-pay ratio. For women aged 40-42, IVF with PGS was the dominant strategy in all scenarios. Lastly, for women aged over 42, no strategy was dominant and if the willingness-to-pay was sufficiently high enough, IVF three consecutive times would be the preferred choice.

Univariate analysis showed that the key factor was in the clinical probabilities, particularly in the rounds 2 and 3 of IVF for the IVFx3 arm. This is consistent with a review of the literature, where past studies used a much lower clinical probability for a live birth when using PGS. Those studies mainly were based on PGS#1 technologies and did not reflect current clinical results.

The implications of the study are important for social change. IVF is a $16B industry in the United States, mainly funded out-of-pocket by patients. Understanding the cost-effectiveness of these interventions will make patients, clinicians, payers, and

policymakers more informed about the right clinical decisions to make to save the most money.

Further research is required to understand all of the real world costs involved with these interventions, some of which were ignored or assumed to be nominal. Other studies examining the overall economic impact of PGS should also be considered. Lastly, this study can help governing bodies to provide guidelines to clinicians when new technology is introduced into clinical practice.

References

ACOG Committee Opinion No. 430: Preimplantation genetic screening for aneuploidy. (2009). *Obstetrics and Gynecology, 113*(3), 766–767. https://doi.org/10.1097/AOG.0b013e31819e9f05

Accreditation Council for Graduate Medical Education. (2016, June 12). ACGME Program Requirements for Graduate Medical Education in Reproductive Endocrinology and Infertility. Retrieved from ACGME.org: https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/235_reproductive_endocrinology_infertility_06122016.pdf

Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology. (2011). The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Human Reproduction (Oxford, England), 26*(6), 1270–1283. https://doi.org/10.1093/humrep/der037

American Society of Reproductive Medicine. (2015). *Assisted Reproductive Technology: A guide for Patients.* Retrieved November 19, 2016, from ASRM.org: https://www.asrm.org/uploadedFiles/ASRM_Content/Resources/Patient_Resources/Fact_Sheets_and_Info_Booklets/ART.pdf

American Society for Reproductive Medicine. (2014). *Is In Vitro Fertilization Expensive?* Retrieved from ASRM.org: http://www.asrm.org/detail.aspx?id=3023

American Society for Reproductive Medicine. (2013). Blastocyst culture and transfer in clinical-assisted reproduction: A committee opinion. *Fertility and Sterility, 99*(1), 667-672.

American Society for Reproductive Medicine. (2008, November). Preimplantation genetic testing: a Practice Committee opinion. *Fertility and Sterility, 90*(3), S136-S143.

Altman, D. G., & Bland, J. M. (2011). How to obtain the P value from a confidence interval. *BMJ, 343*, d2304. https://doi.org/10.1136/bmj.d2304

Andronis, L., Barton, P., & Bryan, S. (2009). Sensitivity analysis in economic evaluation: An audit of NICE current practice and a review of its use and value in decision-making. *Health Technology Assessment (Winchester, England), 13*(29), iii, ix–xi, 1-61. https://doi.org/10.3310/hta13290

ART Success Rates (ART) Data | Reproductive Health | CDC. (n.d.). Retrieved October 11, 2016, from http://www.cdc.gov/art/artdata/index.html

Arth AC, Tinker SC, Simeone RM, Ailes EC, Cragan JD, Grosse SD. Inpatient Hospitalization Costs Associated with Birth Defects Among Persons of All Ages — United States, 2013. MMWR Morb Mortal Wkly Rep 2017;66:41–46. DOI: http://dx.doi.org/10.15585/mmwr.mm6602a1

Boulet, S. L., Crawford, S., Zhang, Y., Sunderam, S., Cohen, B., Bernson, D., . . . States Monitoring ART Collaborative. (2015). Embryo transfer practices and perinatal outcomes by insurance mandate status. *Fertility and Sterility*, *104*(2), 403–409. https://doi.org/10.1016/j.fertnstert.2015.05.015

Brezina, P. R., & Kutteh, W. H. (2015). Clinical applications of preimplantation genetic testing. *BMJ, 350*(feb19 3), g7611–g7611. https://doi.org/10.1136/bmj.g7611

Briggs, A. H., Weinstein, M. C., Fenwick, E. A. L., Karnon, J., Sculpher, M. J., & Paltiel,

A. D. (2012). Model parameter estimation and uncertainty: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value in Health*, *15*(6), 835–842. https://doi.org/10.1016/j.jval.2012.04.014

Brigham, S. A., Conlon, C., & Farquharson, R. G. (1999). A longitudinal study of pregnancy outcome following idiopathic recurrent miscarriage. *Human Reproduction*, *14*(11), 2868–2871. https://doi.org/10.1093/humrep/14.11.2868

Buchanan, J., Wordsworth, S., & Schuh, A. (2013). Issues surrounding the health economic evaluation of genomic technologies. *Pharmacogenomics*, *14*(15), 1833–1847. https://doi.org/10.2217/pgs.13.183

Capalbo, A., Treff, N. R., Cimadomo, D., Tao, X., Upham, K., Ubaldi, F. M., … Scott, R. T., Jr. (2015). Comparison of array comparative genomic hybridization and quantitative real-time PCR-based aneuploidy screening of blastocyst biopsies. *European Journal Of Human Genetics: EJHG*, *23*(7), 901–906. https://doi.org/10.1038/ejhg.2014.222

Centers for Disease Control and Prevention. (2016). *2014 Assisted Reproductive Technology Fertility Clinic Success Rates Report.* Atlanta, GA.

Centers for Disease Control and Prevention. (2016, February 23). Policy Document | Assisted Reproductive Technology (ART) | Reproductive Health | CDC. Retrieved from http://www.cdc.gov/art/nass/policy.html

Cha, A. E. (2017, October 21). Discounts, guarantees and the search for “good” genes: The booming fertility business. *Washington Post*. Retrieved from https://www.washingtonpost.com

Chambers, G. M., Sullivan, E. A., Ishihara, O., Chapman, M. G., & Adamson, G. D. (2009). The economic impact of assisted reproductive technology: A review of selected developed countries. *Fertility and Sterility*, *91*(6), 2281–2294. https://doi.org/10.1016/j.fertnstert.2009.04.029

Chang, J., Boulet, S. L., Jeng, G., Flowers, L., & Kissin, D. M. (2016). Outcomes of in vitro fertilization with preimplantation genetic diagnosis: An analysis of the United States Assisted Reproductive Technology Surveillance Data, 2011-2012. *Fertility and Sterility*, *105*(2), 394–400. https://doi.org/10.1016/j.fertnstert .2015.10.018

Chen, M., Wei, S., Hu, J., & Quan, S. (2015). Can comprehensive chromosome screening technology improve IVF/ICSI outcomes? A meta-analysis. *Plos One*, *10*(10), e0140779–e0140779. https://doi.org/10.1371/journal.pone.0140779

Collura, B. L., & Adamson, D. (2017). Do health insurance infertility benefits impact the employee/employer relationship. *Fertility and Sterility*, *108*(3, Supplement), e9. https://doi.org/10.1016/j.fertnstert.2017.07.041

Crawford, S., Boulet, S. L., Mneimneh, A. S., Perkins, K. M., Jamieson, D. J., Zhang, Y., & Kissin, D. M. (2016). Costs of achieving live birth from assisted reproductive technology: a comparison of sequential single and double embryo transfer approaches. *Fertility and Sterility*, *105*(2), 444–450. https://doi.org/10.1016/j .fertnstert.2015.10.032

Dahdouh, E. M., Balayla, J., & García-Velasco, J. A. (2015). Comprehensive chromosome screening improves embryo selection: a meta-analysis. *Fertility and*

Sterility, *104*(6), 1503–1512. https://doi.org/10.1016/j.fertnstert.2015.08.038

De Sutter, P., Gerris, J., & Dhont, M. (2002). A health-economic decision-analytic model comparing double with single embryo transfer in IVF/ICSI. *Human Reproduction*, *17*(11), 2891–2896.

Definitions of infertility and recurrent pregnancy loss: a committee opinion. (2013). *Fertility and Sterility*, *99*(1), 63. https://doi.org/10.1016/j.fertnstert.2012.09.023

Eftekhar, M., Mohammadian, F., Yousefnejad, F., Molaei, B., & Aflatoonian, A. (2012). Comparison of conventional IVF versus ICSI in non-male factor, normoresponder patients. *Iranian Journal of Reproductive Medicine*, *10*(2), 131–136.

ESHRE Capri Workshop Group. (2015). Economic aspects of infertility care: a challenge for researchers and clinicians. *Human Reproduction (Oxford, England)*, *30*(10), 2243–2248. https://doi.org/10.1093/humrep/dev163

Feichtinger, M., Stopp, T., Göbl, C., Feichtinger, E., Vaccari, E., Mädel, U., … Neesen, J. (2015). Increasing live birth rate by preimplantation genetic screening of pooled polar bodies using array comparative genomic hybridization. *PLoS ONE*, *10*(5). https://doi.org/10.1371/journal.pone.0128317

Fiddelers, A. A. A., Dirksen, C. D., Dumoulin, J. C. M., van Montfoort, A. P. A., Land, J. A., Janssen, J. M., . . . Severens, J. L. (2009). Cost-effectiveness of seven IVF strategies: results of a Markov decision-analytic model. *Human Reproduction*, *24*(7), 1648–1655. https://doi.org/10.1093/humrep/dep041

Ford, H., & Schust, D. J. (2009). Recurrent pregnancy loss - etiology, diagnosis and therapy. *Review of Obstetrics and Gynecology*, *Spring 2*(2), 76-83.

Forman, E. J., Tao, X., Ferry, K. M., Taylor, D., Treff, N. R., & Scott, R. T. (2012). Single embryo transfer with comprehensive chromosome screening results in improved ongoing pregnancy rates and decreased miscarriage rates. *Human Reproduction*, *27*(4), 1217–1222. https://doi.org/10.1093/humrep/des020

Foyouzi, N., Cedars , M., & Huddlest, H. G. (2012). Cost-effectiveness of cytogenetic evaluation of products of conception in the patient with a second pregnancy loss. *Fertility and Sterility, 98*(15), 1-5.

Garber, A., & Phelps, C. (1997). Economic Foundations of Cost-effectiveness Analysis. *Journal of Health Economics, 16,* 1-31.

Gebhart, M. B., Hines, R. S., Penman, A., & Holland, A. C. (2016). How do patient perceived determinants influence the decision-making process to accept or decline preimplantation genetic screening? *Fertility and Sterility*, *105*(1), 188–193. https://doi.org/10.1016/j.fertnstert.2015.09.022

Gleicher, N., Kushnir, V. A., & Barad, D. H. (2014). Preimplantation genetic screening (PGS) still in search of a clinical application: a systematic review. *Reproductive Biology and Endocrinolology*, *12*(1), 22.

Glujovsky, D., & Farquhar, C. (2016). Cleavage-stage or blastocyst transfer: What are the benefits and harms? *Fertility and Sterility*, *106*(2), 244–250. https://doi.org/10.1016/j.fertnstert.2016.06.029

Goldhaber-Fiebert, J. D., & Brandeau, M. L. (2015). Evaluating cost-effectiveness of interventions that affect fertility and childbearing: How health effects are measured matters. *Medical Decision Making*, *35*(7), 818–846.

https://doi.org/10.1177/0272989X15583845

Greco, E., Bono, S., Ruberti, A., Lobascio, A. M., Greco, P., Biricik, A., . . . Fiorentino, F. (2014). Comparative genomic hybridization selection of blastocysts for repeated implantation failure treatment: A pilot study. *BioMed Research International*, *2014*, e457913. https://doi.org/10.1155/2014/457913

Grosse, S. D., Wordsworth, S., & Payne, K. (2008). Economic methods for valuing the outcomes of genetic testing: beyond cost-effectiveness analysis. *Genetics in Medicine*, *10*(9), 648–654. https://doi.org/10.1097/GIM.0b013e3181837217

Halpern, E. F., Weinstein, M. C., Hunink, M. G. M., & Gazelle, G. S. (2000). Representing both first- and second-order uncertainties by Monte Carlo simulation for groups of patients. *Medical Decision Making*, *20*(3), 314–322. https://doi.org/10.1177/0272989X0002000308

Hamilton, A. B., Oishi, S., Yano, E. M., Gammage, C. E., Marshall, N. J., & Scheuner, M. T. (2014). Factors influencing organizational adoption and implementation of clinical genetic services. *Genetics in Medicine*, *16*(3), 238–245. https://doi.org/10.1038/gim.2013.101

Handyside, A. H. (2013). 24-chromosome copy number analysis: a comparison of available technologies. *Fertility and Sterility*, *100*(3), 595–602. https://doi.org/10.1016/j.fertnstert.2013.07.1965

Harper, J. C., Wilton, L., Traeger-Synodinos, J., Goossens, V., Moutou, C., SenGupta, S. B., … Harton, G. (2012). The ESHRE PGD Consortium: 10 years of data collection. *Human Reproduction Update*, *18*(3), 234–247.

https://doi.org/10.1093/humupd/dmr052

Harper, Joyce C., & Sengupta, S. B. (2012). Preimplantation genetic diagnosis: state of the art 2011. *Human Genetics*, *131*(2), 175–186. https://doi.org/10.1007/s00439-011-1056-z

Helmes, A. W. (2002). Application of the Protection Motivation Theory to Genetic Testing for Breast Cancer Risk. *Preventive Medicine*, *35*(5), 453–462. https://doi.org/10.1006/pmed.2002.1110

Hodes-Wertz,, Grifo, J., Ghadir, S., Kaplan, B., Laskin, C. A., Glassner, M., & Munné, S. (2012). Idiopathic recurrent miscarriage is caused mostly by aneuploid embryos. *Fertility and Sterility*, *98*(3), 675–680. https://doi.org/10.1016/j.fertnstert.2012.05.025

Hodes-Wertz, B., McCulloh, D. H., & Grifo, J. (2015). Preimplantation genetic screening is cost effective in cost per delivery compared to routine in vitro fertilization. *Fertility and Sterility*, *104*(3), e278. https://doi.org/10.1016/j.fertnstert.2015.07.873

Holt, S., Bertelli, G., Humphreys, I., Valentine, W., Durrani, S., Pudney, D., … Phillips, C. J. (2013). A decision impact, decision conflict and economic assessment of routine Oncotype DX testing of 146 women with node-negative or pNImi, ER-positive breast cancer in the UK. *British Journal of Cancer*, *108*(11), 2250–2258. https://doi.org/10.1038/bjc.2013.207

Hornberger, J., Hirsch, F. R., Li, Q., & Page, R. D. (2015). Outcome and economic implications of proteomic test-guided second- or third-line treatment for advanced

non-small cell lung cancer: Extended analysis of the PROSE trial. *Lung Cancer*, *88*(2), 223–230. https://doi.org/10.1016/j.lungcan.2015.03.006

IVF Success Rates with PGS | Arizona IVF Clinic. (n.d.). Retrieved February 2, 2017, from http://www.acfs2000.com/pgs-ivf-success-rates.html

Jungheim, E. S., Meyer, M., & Broughton, D. E. (2015). Best practices for controlled ovarian stimulation in IVF. *Seminars in Reproductive Medicine*, *33*(2), 77–82. https://doi.org/10.1055/s-0035-1546424

Keltz, M. D., Vega, M., Sirota, I., Lederman, M., Moshier, E. L., Gonzales, E., & Stein, D. (2013). Preimplantation Genetic Screening (PGS) with Comparative Genomic Hybridization (CGH) following day 3 single cell blastomere biopsy markedly improves IVF outcomes while lowering multiple pregnancies and miscarriages. *Journal of Assisted Reproduction and Genetics*, *30*(10), 1333–1339. https://doi.org/10.1007/s10815-013-0070-6

Kulkarni, A. D., Jamieson, D. J., Jones, H. W. J., Kissin, D. M., Gallo, M. F., Macaluso, M., & Adashi, E. Y. (2013). Fertility Treatments and Multiple Births in the United States. *New England Journal of Medicine*, *369*(23), 2218–2225. https://doi.org/10.1056/NEJMoa1301467

Kushnir, V. A., Darmon, S. K., Albertini, D. F., Barad, D. H., & Gleicher, N. (2016). Effectiveness of in vitro fertilization with preimplantation genetic screening: a reanalysis of United States assisted reproductive technology data 2011–2012. *Fertility and Sterility*, *106*(1), 75–79. https://doi.org/10.1016/j.fertnstert.2016.02.026

Lee, E., Illingworth, P., Wilton, L., & Chambers, G. M. (2015). The clinical effectiveness of preimplantation genetic diagnosis for aneuploidy in all 24 chromosomes (PGD-A): systematic review. *Human Reproduction (Oxford, England)*, *30*(2), 473–483. https://doi.org/10.1093/humrep/deu303

Lee, H.-L., McCulloh, D. H., Hodes-Wertz, B., Adler, A., McCaffrey, C., & Grifo, J. A. (2015). In vitro fertilization with preimplantation genetic screening improves implantation and live birth in women age 40 through 43. *Journal of Assisted Reproduction and Genetics*, *32*(3), 435–444. https://doi.org/10.1007/s10815-014-0417-7

Liu, J., Wang, W., Sun, X., Liu, L., Jin, H., Li, M., … Gill, J. (2012). DNA Microarray Reveals That High Proportions of Human Blastocysts from Women of Advanced Maternal Age Are Aneuploid and Mosaic. *Biology of Reproduction*, *87*(6), 148–148. https://doi.org/10.1095/biolreprod.112.103192

Luke, B., Brown, M. B., Wantman, E., Lederman, A., Gibbons, W., Schattman, G. L., … Stern, J. E. (2012). Cumulative Birth Rates with Linked Assisted Reproductive Technology Cycles. *New England Journal of Medicine*, *366*(26), 2483–2491. https://doi.org/10.1056/NEJMoa1110238

Mastenbroek, S, Twisk, M., van der Veen, F., & Repping, S. (2011). Preimplantation genetic screening: a systematic review and meta-analysis of RCTs. *Human Reproduction Update*, *17*(4), 454–466. https://doi.org/10.1093/humupd/dmr003

Mastenbroek, S., & Repping, S. (2014). Preimplantation genetic screening: back to the future. *Human Reproduction*, *29*(9), 1846–1850.

https://doi.org/10.1093/humrep/deu163

Maxwell, S. M., Melzer-Ross, K., McCulloh, D. H., & Grifo, J. A. (2015). A comparison of pregnancy outcomes between day 3 and day 5/6 embryo transfers: does day of embryo transfer really make a difference? *Journal of Assisted Reproduction and Genetics*, *32*(2), 249–254. https://doi.org/10.1007/s10815-014-0404-z

Meltzer, David. (2001). Theoretical Foundations of Medical Cost Effectiveness Analysis: Implications for the Measurement of Benefits and Costs of Medical Interventions. In D. M. Cutler (Ed.), *Medical care output and productivity*. Chicago, IL: University of Chicago Press.

Mersereau, J., Plunkett, B., & Cedars, M. (2008). Preimplantation genetic screening in older women: a cost-effectiveness analysis. *Fertility and Sterility*, *90*(3), 592–598. https://doi.org/10.1016/j.fertnstert.2007.07.1307

Ming-Yu, F., & Xiao-Hua, Z. (2007). A simulation study to compare methods for constructing confidence intervals for the incremental cost-effectiveness ratio. *Health Services & Outcomes Research Methodology; Dordrecht*, *7*(1–2), 57–77. https://doi.org/http://dx.doi.org.ezp.waldenulibrary.org/10.1007/s10742-006-0017-9

Munné, S. (2012). Preimplantation genetic diagnosis for aneuploidy and translocations using array comparative genomic hybridization. *Current Genomics*, *13*(6), 463.

Murugappan, G., Ohno, M. S., & Lathi, R. B. (2015). Cost-effectiveness analysis of preimplantation genetic screening and in vitro fertilization versus expectant management in patients with unexplained recurrent pregnancy loss. *Fertility and*

Sterility, *103*(5), 1215–1220. https://doi.org/10.1016/j.fertnstert.2015.02.012

National Center for Chronic Disease Prevention and Health Promotion. (2014, November). *2012 Assisted Reproductive Technology National Summary Report.* Retrieved from CDC.gov: http://www.cdc.gov/art/pdf/2012-report/national-summary/art_2012_national_summary_report.pdf

National Conference of State Legislatures. (2014, June 1). *STATE LAWS RELATED TO INSURANCE COVERAGE FOR INFERTILITY TREATMENT.* Retrieved January 19, 2017, from NCSL.org: http://www.ncsl.org/research/health/insurance-coverage-for-infertility-laws.aspx

Neumann, P. J., Cohen, J. T., & Weinstein, M. C. (2014). Updating Cost-Effectiveness — The Curious Resilience of the $50,000-per-QALY Threshold. *New England Journal of Medicine*, *371*(9), 796–797. https://doi.org/10.1056/NEJMp1405158

Niakan, K. K., Han, J., Pedersen, R. A., Simon, C., & Pera, R. A. R. (2012). Human pre-implantation embryo development. *Development*, *139*(5), 829–841. https://doi.org/10.1242/dev.060426

Noble, G., Harvey, C., Napier, C., Curtis-Barton, M., McGarraghy, M., & Maheshwari, A. (2010). Preimplantation genetic screening in advanced maternal age: a systematic review. *Expert Review of Obstetrics & Gynecology*, *5*(6), 707–716. https://doi.org/10.1586/eog.10.55

Noyes, K., & Holloway, R. G. (2004). Evidence from Cost-Effectiveness Research. *NeuroRx*, *1*(3), 348-355.

O'Brien, B. (1996). Economic evaluation of pharmaceuticals. Frankenstein's monster or

vampire of trials? *Medical Care*, *34*(12 Suppl), DS99-108.

Orvieto, R. (2016). Preimplantation genetic screening- the required RCT that has not yet been carried out. *Reproductive Biology And Endocrinology: RB&E*, *14*(1), 35–35. https://doi.org/10.1186/s12958-016-0171-z

Phillips, K. A., Sakowski, J. A., Liang, S.-Y., & Ponce, N. A. (2013). Economic Perspectives on Personalized Health Care and Prevention. *Forum for Health Economics and Policy*, *16*(2). https://doi.org/10.1515/fhep-2013-0010

Ralph, A. F., Ager, B., Bell, M. L., Collins, I. M., Andrews, L., Tucker, K., … Butow, P. (2014). Women's preferences for selective estrogen reuptake modulators: an investigation using protection motivation theory. *Patient Education And Counseling*, *96*(1), 106–112. https://doi.org/10.1016/j.pec.2014.04.011

Ramsey, S. D., Willke, R. J., Glick, H., Reed, S. D., Augustovski, F., Jonsson, B., … Sullivan, S. D. (2015). Cost-Effectiveness Analysis Alongside Clinical Trials II—An ISPOR Good Research Practices Task Force Report. *Value in Health*, *18*(2), 161–172. https://doi.org/10.1016/j.jval.2015.02.001

Ramsey, S., Willke, R., Briggs, A., Brown, R., Buxton, M., Chawla, A., … Reed, S. (2005). Good Research Practices for Cost-Effectiveness Analysis Alongside Clinical Trials: The ISPOR RCT-CEA Task Force Report. *Value in Health*, *8*(5), 521–533. https://doi.org/10.1111/j.1524-4733.2005.00045.x

Resetkova, N., Tobler, K. J., Kearns, W. G., & Werner, E. F. (2013). In vitro fertilization (IVF) with 23-chromosome pair preimplantation genetic screening (PGS) is cost effective to achieve a live birth compared to IVF alone for recurrent pregnancy

loss (RPL). *Fertility and Sterility*, *100*(3), S99–S100. https://doi.org/10.1016/j.fertnstert.2013.07.1704

Rogers, R. (1975). A protection motivation theory of fear appeals and attitude change. *Journal of Psychology*, *91*, 93-114.

Roque, M., Lattes, K., Serra, S., Solà, I., Geber, S., Carreras, R., & Checa, M. A. (2013). Fresh embryo transfer versus frozen embryo transfer in in vitro fertilization cycles: a systematic review and meta-analysis. *Fertility and Sterility*, *99*(1), 156–162. https://doi.org/10.1016/j.fertnstert.2012.09.003

Salem, W., Ho, J. R., Bendikson, K. A., Chung, K., & Paulson, R. (2016). IVF patients over age 39 experience decreased cost effectiveness and live birth rates with preimplantation genetic screening: a decision analytic model and cost effectiveness analysis. *Fertility and Sterility*, *106*(3), e60–e61. https://doi.org/10.1016/j.fertnstert.2016.07.184

Schoolcraft, W. B., Fragouli, E., Stevens, J., Munne, S., Katz-Jaffe, M. G., & Wells, D. (2010). Clinical application of comprehensive chromosomal screening at the blastocyst stage. *Fertility and Sterility*, *94*(5), 1700–1706. https://doi.org/10.1016/j.fertnstert.2009.10.015

Schwartz, A., Juneau, C. R., Patounakis, G., Kaser, D. J., Molinaro, T., Maguire, M., & Scott, R. T. (2017). Preimplantation genetic screening (PGS) utilization is increased in patients without in vitro fertilization insurance coverage. *Fertility and Sterility*, *108*(3, Supplement), e113–e114. https://doi.org/10.1016/j.fertnstert.2017.07.347

Scott, R. T., Upham, K. M., Forman, E. J., Hong, K. H., Scott, K. L., Taylor, D., … Treff, N. R. (2013). Blastocyst biopsy with comprehensive chromosome screening and fresh embryo transfer significantly increases in vitro fertilization implantation and delivery rates: a randomized controlled trial. *Fertility and Sterility*, *100*(3), 697–703. https://doi.org/10.1016/j.fertnstert.2013.04.035

Smith, A. D. A. C., Tilling, K., Nelson, S. M., & Lawlor, D. A. (2015). Live-Birth Rate Associated With Repeat In Vitro Fertilization Treatment Cycles. *JAMA*, *314*(24), 2654. https://doi.org/10.1001/jama.2015.17296

Society for Assisted Reproductive Technologies. (2014). *Assisted Reproductive Technologies.* Retrieved from SART.org: http://www.sart.org/SART_Assisted_Reproductive_Technologies/

Stason, W. B., & Weinstein, M. C. (1977). Allocation of Resources to Manage Hypertension [review-article]. Retrieved January 6, 2017, from http://www.nejm.org/doi/full/10.1056/NEJM197703312961307

Sunderam, S., Kissin, D. M., Crawford, S. B., Folger, S. G., Jamieson, D. J., Warner, L., … Centers for Disease Control and Prevention. (2015). Assisted Reproductive Technology Surveillance - United States, 2013. *Morbidity and Mortality Weekly Report. Surveillance Summaries (Washington, D.C.: 2002)*, *64*(11), 1–25. https://doi.org/10.15585/mmwr.ss6411a1

Trochim, W. (2006). *Statistical power*. Retrieved from http://www.socialresearchmethods.net/kb/power.php

U.S. Department of Health and Human Services, Office on Women's Health. (2009, July

1). *Infertility.* Retrieved from Womenshealth.gov: https://www.womenshealth.gov/files/assets/docs/fact-sheets/infertility.pdf

Van Voorhis, B. J. (2007). In vitro fertilization. *New England Journal of Medicine*, *356*(4), 379–386.

Van Voorhis, B. J., Thomas, M., Surrey, E. S., & Sparks, A. (2010). What do consistently high-performing in vitro fertilization programs in the U.S. do? *Fertility and Sterility*, *94*(4), 1346–1349. https://doi.org/10.1016/j.fertnstert.2010.06.048

Verberg, M. F. G., Eijkemans, M. J. C., Heijnen, E. M. E. W., Broekmans, F. J., Klerk, C. de, Fauser, B. C. J. M., & Macklon, N. S. (2008). Why do couples drop-out from IVF treatment? A prospective cohort study. *Human Reproduction*, *23*(9), 2050–2055. https://doi.org/10.1093/humrep/den219

Weinstein, M. C., Siegel, J. E., Gold, M. R., Kamlet, M. S., & Russell, L. B. (1997). Recommendations of the Panel on Cost-Effectiveness in Health and Medicine: *JAMA*, *276*(15), 1253–1258. https://doi.org/10.1097/00132586-199712000-00019

Yang, Z., Lin, J., Zhang, J., Fong, W. I., Li, P., Zhao, R., … Liu, J. (2015). Randomized comparison of next-generation sequencing and array comparative genomic hybridization for preimplantation genetic screening: a pilot study. *BMC Medical Genomics*, *8*(1). https://doi.org/10.1186/s12920-015-0110-4

Yang, Z., Liu, J., Collins, G. S., Salem, S. A., Liu, X., Lyle, S. S., … Salem, R. D. (2012). Selection of single blastocysts for fresh transfer via standard morphology assessment alone and with array CGH for good prognosis IVF patients: results from a randomized pilot study. Molecular Cytogenetics, 5(1), 1.

Appendix A: List of Acronyms

aCGH / CGH – Array Comparative Genomic Hybridization

ART - Assisted Reproductive Technology

CCS - Comprehensive Chromosomal Screening

CDC - Centers for Disease Control

CEA – Cost-effectiveness Analysis

CER – Cost-Effectiveness Ratio

CMS – Center For Medicaid and Medicare Services

eSET – Elective Single embryo Transfer

ET – Embryo Transfer

FET – Frozen Embryo Transfer

FISH – Fluorescence In-Situ Hybridization

FT – Fresh Embryo Transfer

HIPAA – Health Information Portability and Accountability Act

ICER – Incremental Cost-effectiveness Ratio

ICSI – Intracytoplasmic sperm injection

IVF – In Vitro Fertilization

NGS – Next Generation Sequencing

OB/GYN – Obstetrician / Gynecologist

PCR – Polymerase Chain Reaction

PGD – Preimplantation Genetic Diagnosis

PHI – Personal Health Information

PGS – Preimplantation Genetic Screening

PSA – Probabilistic Sensitivity Analysis

RE / REI – Reproductive Endocrinology / Infertility

RCT – Randomized Controlled Trials

RPL – Recurrent Pregnancy Loss

SART – Society for Assisted Reproductive Technology

Appendix B: Data Use Agreement

DATA USE AGREEMENT

This Data Use Agreement ("Agreement"), effective as of (Enter date.) ("Effective Date"), is entered into by and between William Moye ("Data Recipient") and Brooke Hodes-Wertz ("Data Provider"). The purpose of this Agreement is to provide Data Recipient with access to a Limited Data Set ("LDS") for use in research in accord with the HIPAA and FERPA Regulations.

1. Definitions. Unless otherwise specified in this Agreement, all capitalized terms used in this Agreement not otherwise defined have the meaning established for purposes of the "HIPAA Regulations" codified at Title 45 parts 160 through 164 of the United States Code of Federal Regulations, as amended from time to time.

2. Preparation of the LDS. Data Provider shall prepare and furnish to Data Recipient a LDS in accord with any applicable HIPAA or FERPA Regulations

Data Fields in the LDS. **No direct identifiers such as names may be included in the Limited Data Set (LDS).** The researcher will also not name the organization in the doctoral project report that is published in Proquest. In preparing the LDS, Data Provider or shall include the **data fields specified as follows**, which are the minimum necessary to accomplish the research: Costs associated with PGS and IVF.

3. Responsibilities of Data Recipient. Data Recipient agrees to:

   a. Use or disclose the LDS only as permitted by this Agreement or as required by law;

   b. Use appropriate safeguards to prevent use or disclosure of the LDS other than as permitted by this Agreement or required by law;

   c. Report to Data Provider any use or disclosure of the LDS of which it becomes aware that is not permitted by this Agreement or required by law;

   d. Require any of its subcontractors or agents that receive or have access to the LDS to agree to the same restrictions and conditions on the use and/or disclosure of the LDS that apply to Data Recipient under this Agreement; and

   e. Not use the information in the LDS to identify or contact the individuals who are data subjects.

4. Permitted Uses and Disclosures of the LDS. Data Recipient may use and/or disclose the LDS for its research activities only.

5. Term and Termination.

   a. Term. The term of this Agreement shall commence as of the Effective Date and shall continue for so long as Data Recipient retains the LDS, unless sooner terminated as set forth in this Agreement.

   b. Termination by Data Recipient. Data Recipient may terminate this agreement at any time by notifying the Data Provider and returning or destroying the LDS.

   c. Termination by Data Provider. Data Provider may terminate this agreement at any time by providing thirty (30) days prior written notice to Data Recipient.

   d. For Breach. Data Provider shall provide written notice to Data Recipient within ten (10) days of any determination that Data Recipient has breached a material term of this Agreement. Data Provider shall afford Data Recipient an opportunity to cure said alleged material breach upon mutually agreeable terms. Failure to agree on mutually agreeable terms for cure within thirty (30) days shall be grounds for the immediate termination of this Agreement by Data Provider.

   e. Effect of Termination. Sections 1, 4, 5, 6(e) and 7 of this Agreement shall survive any termination of this Agreement under subsections c or d.

6. Miscellaneous.

   a. Change in Law. The parties agree to negotiate in good faith to amend this Agreement to comport with changes in federal law that materially alter either or both parties' obligations under this Agreement. Provided however, that if the parties are unable to agree to mutually acceptable amendment(s) by the compliance date of the change in applicable law or regulations, either Party may terminate this Agreement as provided in section 6.

   b. Construction of Terms. The terms of this Agreement shall be construed to give effect to applicable federal interpretative guidance regarding the HIPAA Regulations.

   c. No Third Party Beneficiaries. Nothing in this Agreement shall confer upon any person other than the parties and their respective successors or assigns, any rights, remedies, obligations, or liabilities whatsoever.

   d. Counterparts. This Agreement may be executed in one or more counterparts, each of which shall be deemed an original, but all of which together shall constitute one and the same instrument.

e. Headings. The headings and other captions in this Agreement are for convenience and reference only and shall not be used in interpreting, construing or enforcing any of the provisions of this Agreement.

IN WITNESS WHEREOF, each of the undersigned has caused this Agreement to be duly executed in its name and on its behalf.

DATA PROVIDER

Signed: _____

Print Name: Brooke Hobbs-Wertz

Print Title: MD

8/29/17

DATA RECIPIENT

Signed: _____

Print Name: William Moye

Print Title: Doctoral Student