Addressing the pharmacogenomic variance associated with infertility outcomes in south Asians and recommendations for customization of assisted reproductive technology techniques.

Addressing the pharmacogenomic variance associated with infertility outcomes in south Asians and recommendations for customization of assisted reproductive technology techniques.

Maham Zakir

A Thesis in the Field of Biotechnology Management

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

May 2024

Abstract

Female subfertility impacts approximately 15% of the world population. Those experiencing infertility rely on Assisted Reproductive Technologies (ART), which constitutes a variety of procedures including In Vitro Fertilization (IVF). The CDC estimates national ART success rates at 23.4% with numbers swiftly declining with age. IVF is financially and emotionally cumbersome on the patient, with prohibitive barriers for lower socio-economic groups. Despite scientific advances in the field, there are several unknowns that give rise to poor outcomes as standard fertility protocols are not customized based on genomic ethnic variation.

This thesis aims to address the pharmacogenomic variance in infertility outcomes in the South Asian population and attempts to provide a blueprint for a customized approach to infertility treatment. Additionally, the merits of multifactorial customization and its potential to improve fertility outcomes have been explored to provide a holistic depiction of the south Asian health landscape, socio economic factors and their impact to the presence and treatment of infertility in this unique population.

Literature searches were performed to identify 5 key genes involved in ovulatory dysfunction, recurrent pregnancy loss, unexplained factors and implantation (CDC infertility diagnosis categories). Analysis was conducted on genomic variation categorized by ancestry groups in the genomic aggregator database gnomAD, and cross referenced with other databases. Allele frequencies of pathogenic, predicted loss of function, and previously identified infertility associated gene variants were calculated and

compared against the backdrop of general allele frequencies. Clinical trial representation, disease burdens and future projections, lifestyle factors, reported genomic variations, cultural and socio-economic factors were analyzed for South Asians. IVF protocols were assessed from three leading fertility clinics and assessed for gaps and risks in the South Asian context and recommendations for customization have been postulated.

The 5 genes of interested assessed were LIF, IGF-1, HOXA10, HOXA11, and ESR1. Pathogenic, predicted loss of function, and previously identified infertility associated previously identified variants showed variation in allele frequencies in South Asian populations when compared to other ancestry groups.

Additionally, disease prevalence statistics and key genomic variations associated with PCOS, diabetes, drug metabolism and their potential links to infertility were identified. Potential recommendations for custom IVF protocol enhancements for South Asian infertility patients have been put forth with the goal to improve ART outcomes and additional commentary has been provided to address population differentiators and factors of influence.

This thesis sheds light on the existing “one size fits all” approach towards infertility and provides avenues for multifactorial customization of infertility treatment based on genomic ethnic variation. Additionally, it illuminates the need of ethnic inclusion in clinical trials and genomic databases to be able to remove health disparities globally. It also highlights the need for deep research in the identification of key genes, rare variants, proteins, and enzymes as therapeutic targets to improve fertility outcomes. Ultimately, a pharmacogenomic approach that couples emerging technologies such as artificial intelligence and machine learning with genomic research can result in predictive

algorithms to solving infertility with precision for inter and intra population groups. This coupled with a holistic understanding of socio-economic, cultural constraints, developing and global healthcare challenges will lead to reduced costs of infertility treatments, improved success outcomes, mitigation of ART associated risks and demonstrate benefits to policy makers, clinicians, insurance companies, and patients.

Dedication

This thesis is dedicated to the invisible warriors seeking answers to infertility and researchers who want to make the world of science a more equitable place of representation. I would also like to dedicate my work to my parents Zakir H. Nuruddin & Farhana Zakir, who have empowered my intellectual pursuits, to my spouse Farhan Bhaba for his support, my siblings Danish and Rida for their ever-motivating taunts and my friend Jeremy Cox for his constant encouragement. Most of all I would like to thank my daughter, the original proof of concept and inspiration, who has lit up my life with her toddler hood and made this experience challenging and yet enjoyable. And finally, I dedicate this to my exceptionally supportive thesis director Beth Zielinski Habershaw, who has provided me the intellectual and emotional space to express myself and have the confidence to execute my ideas.

Table of Contents

List of Tables

List of Figures

Chapter I.

Infertility: The Realm of the Invisible

In the heart's deep chamber, a silent plea,
Yearning whispers for a child to be,
Tumultuous emotions, hidden in invisible cloaks as I go about my day,
In the saga of life, I have no answers to this fray,
Science and hope entwined, a waltz of genes and fate,
Between these visceral base pairs, my eternal hope awaits.
--written by the author

Female (in)fertility and subfertility impact approximately 15% of the world population and the Centers for Disease Control and Prevention (CDC) estimates that about 12% of women aged 15 to 44 years in the United States have difficulty getting pregnant or carrying a pregnancy to term. According to the World Health Organization, millions of individuals are affected by infertility and one in every six people of reproductive age worldwide experience infertility in their lifetime.

Many factors contribute to impaired fertility in females which include endocrine disorders such as Polycystic Ovary Syndrome (PCOS), diminished ovarian reserve, functional hypothalamic amenorrhea (FHA), improper function of the hypothalamus and pituitary glands, uterine disorders, premature ovarian insufficiency, menopause, anatomical abnormalities etc. The risk of infertility also increases with age, smoking, alcohol use, weight gain/loss, as well as physical or emotional stress. Assisted Reproductive Technology (ART) commonly includes procedures such as Intrauterine

insemination (IUI), In vitro fertilization (IVF), Intracytoplasmic sperm injection (ICSI), Zygote intrafallopian transfer (ZIFT), use of donor eggs, embryos, and surrogacy.

National ART Statistics

The CDC provides ART surveillance at the national and state level and reports statistical data aggregated from clinics nationwide. It provides consumer statistics that offers visibility into diagnosis segmentation, success rates, and expected outcomes based on a plethora of variables. Although this information provides visibility into critical statistics pertaining to ART success rates for some variables it does not segment data attributes by population diversity. As IVF adoption increases nationally and internationally, data collection parameters and data modeling need to evolve to reflect the success rates/treatment response of various ancestry groups. This will provide much needed insight into why estimates for cumulative live birth rates are discordant for certain ethnicity groups and will provide a more realistic prediction for patients and clinicians weighing treatment options by ethnicity.

Diagnosis Segmentation

There are several diagnosis/reasons which may impact the success outcomes and treatment approach for ART candidates. The CDC reports these percentages aggregated from clinics nationwide based on a total of 413,776 reported ART cycles in 2021. These diagnosis categories include Tubal factor, Uterine factor, Ovulatory dysfunction, Diminished Ovarian Reserve, Male Factor, Endometriosis, Recurrent Pregnancy Loss, Egg or Embryo Banking, Preimplantation Genetic Testing, Gestational Carrier, Other factor and Unexplained factors . The “other factor” category (24%) refers to other known

reasons which may be related to infertility (Immunological, chromosomal, cancer chemotherapy, other illnesses) or known health issues not related to fertility.

National usage percentages based on diagnosis/reason are reported in table 1 below based on 413,776 ART cycles.

Table 1. Segmentation of ART Usage Based on Diagnosis Category.

Diagnosis/Reason	National %
Tubal factor	10.4%
Uterine factor	5.8%
Ovulatory dysfunction	14.3%
Diminished ovarian reserve	26.9%
Endometriosis	6.5%
Male factor	27.8%
Egg or embryo banking	40.9%
Recurrent pregnancy loss	6.3%
Preimplantation genetic testing	16.6%
Gestational carrier	1.4%
Other factor	30.0%
Unexplained factor	10.5%

Table 1. Segmentation of ART treatment usage based on diagnosis category of patient. Source: (https://nccd.cdc.gov/drh_art) based on 413,776 ART cycles.

Success Rates

Success rates are observed to drop dramatically with increasing age and as the general population chooses to have children later in life, this has a resounding impact on the success rates of ART with increasing age. The table below provides a high-level snapshot of total cycles, fertility preservation cycles, pregnancies, deliveries, and the number of live births in 2021 nationwide..

Table 2. National Summary Data on Outcomes of ART Cycles In 2021.

2021 National Summary	Number
Total cycles	413,776
Fertility preservation cycles	36,072
Pregnancies	112,088
Deliveries	91,906
Total infants born	97,128

Table above shows the National Summary Data from the Centers of Disease Control (CDC) on ART cycle outcomes in 2021.The table shows total numbers of cycles resulting in pregnancies, deliveries and infants born (live births). The success rate was calculated by dividing the total infants born by the total cycles. Source: https://nccd.cdc.gov/drh_art

The table above shows the national success rates of ART resulting in a live birth was 23.5% calculated by dividing total infants born 97,128 out of a total of 413,776 cycles that were conducted in 2021. This means that a large population of ART candidates will require repeat ART procedures or will need alternative options to achieve success. Additionally, challenges exist not only at the initial phase of achieving a positive

pregnancy diagnosis phase but also during the maintenance of a viable pregnancy during the gestation period to produce a live birth.

The table below shows the percentage of intended egg retrievals resulting in live-birth deliveries based on national information from 144,773 ART cycles by age group. This includes all diagnosis categories outlined in table 1.

Table 3. CDC 2021 National Summary Data on ART Associated Live Birth Deliveries

Patient Age	<35	35-37	38-40	>40
National %	50.7%	36.3%	23.3%	7.9%

For patients using their own eggs, the table shows cumulative percentages of intended egg retrieval cycles resulting in live-birth deliveries, categorized by age. The sample includes all diagnosis categories based on 144,773 ART cycles as reported by the CDC. The chances of live births reduces drastically with increasing age of patients. Source: https://nccd.cdc.gov/drh_art.

It is important to note that not all cycles are initiated with egg retrieval intent result in actual egg retrieval and some cycles may be halted prior to retrieval for a multitude of reasons. Therefore, the intended retrievals may be higher than actual retrievals. The ART lifecycle from start to finish presents many areas of optimization that can be leveraged to produce better live-birth delivery rates.

Categories of Interest

Drilling down into specific diagnosis categories, the success rates for live births for the diagnosis of "Uterine factor" (based on 8,245 ART cycles) were even lower than the national average by a few percentages in each age category. The success rates for "Recurrent Pregnancy Loss" (based on 8,606 ART cycles) were lower than the national

average in the <35 and 35-37 age groups, while the “Unexplained factor” (based on 14,216 ART cycles) and “Ovulatory dysfunction” (based on 18,767 ART cycles) performed marginally better in all categories compared to the national average of live birth deliveries.

Controlling for the variance in sample sizes for each of these categories, these statistics merit additional investigation. Several categories such as “uterine factor”, “recurrent pregnancy loss”, “unexplained infertility” remain ambiguous, lack explicit causation information and as a result there is a lack of clarity around effective treatment options. This underscores the possibility that there are several avenues of variability within the general population that may be important to consider.

Additionally, further research needs to be conducted in each diagnostic category to sub categorize and further identify the exact molecular mechanisms, genomic variances and associated pathways. Key insights should be extrapolated from treatment success rates/response coupled with additional data categories such as ethnicity. This can further elucidate specific ancestry groups which stand at greater risk of adverse events or those where certain diagnoses categories occur at greater frequencies.

Regulatory Bodies

In the United States the CDC, Food and Drug Administration (FDA), and Centers for Medicare and Medicaid Services (CMS) have regulatory responsibilities in ART regulation, monitoring, and reporting. Federal legislation on ART was passed by Congress in 1992. Public Law 102-493, the Fertility Clinic Success Rate and Certification Act (FCSRCA) mandates reporting of ART data and endorses standard definitions. The CDC publishes annual reports, aggregates data on infertility procedures,

diagnosis, and success rates nationally from clinics and provides consumer access to reporting. The FDA has established good tissue practices that are codified in 21 CFR 1271 to set requirements over the screening and testing of reproductive tissues (eggs and sperm), tissue donors, facility standards, medical histories, identification controls, and infectious disease management. CMS under the Clinical Laboratory Improvement Act (CLIA) ensures quality laboratory diagnostic testing by establishing standards for accuracy and reliability. It is important to note that most South Asian countries have ill-defined and poorly implemented standards for ART which stems from poor healthcare systems, lack of regulatory oversight, poor data collection, and cultural taboos associated with ART procedures. As a result, reporting data is sparse and unreliable. This adversely impacts the entire ART ecosystem. Patients and physicians have limited insights into ART data, success factors and public health and safety is debatable at best.

Mental Health Impacts

The financial and emotional impacts of infertility are associated with anger, depression, anxiety, marital problems, sexual dysfunction, and social isolation. The pooled prevalence values of depression among infertile women are 44.32% (95% CI: 35.65–52.99) in low- and middle-income countries and 28.03% (95% CI: 19.61–36.44) in high-income countries (Kiani, 2021).

Cultural opinions, stigmatization, financially prohibitive access to quality IVF clinics and repeat IVF failures can all negatively impact mental well-being. Rates of stress, anxiety and depression are higher in the IVF population compared to the general population which can further negatively impact IVF success rates (Aimagambetova, 2020)

Financial Burden of Infertility

The cost of IVF can vary on average between $19,830 and $65,850 and can steeply rise based on age, diagnostic testing requirements, number of cycles, medications, frozen embryo services etc. Leading fertility clinics were consulted for 2024 costs and are summarized below to provide a financial snapshot of a typical IVF cycle. Many patients require multiple cycles and additional diagnostic testing can add to costs. These may include pre-existing health conditions, age, BMI, type of procedure, IVF-associated adverse effects, oocyte quality, recurrent pregnancy loss, uterine factors etc. The existing one size fits all IVF approach does not account for genomic variances within populations that may respond better to customized treatments as discussed in detail in the following sections. Repeat cycle costs, the general lack of insurance coverage of infertility treatments in most states, and rising inflation are financial stressors that add to the overall emotional burden of the experience. 15 states have enacted legislation that requires private insurers to cover a portion of infertility costs with only 8 states mandating reimbursement for ART. This reduces access to lower socio-economic groups and compounds the burden of unsuccessful cycles. The increase in financial pressure to conceive can also impact choice of treatment options and increase IVF associated risks. Patients may opt to implant multiple embryos at a time which increases the risk of multiple births and adversely impact maternal and fetal health. The table below breaks down key cost buckets associated with an average IVF cycle which can drastically increase based on the individual's case.

Table 4. Breakdown Of Cost Components in an Average IVF Cycle

Clinic Name/ Location(s)	IVF+ FET	Medicine Costs	Diagnostic Testing	Genetic Testing	Cryo Services	FET Add on cycles

Mayo Clinic, Rochester (Minnesota)	$16,650 -$18500 (With embryo testing)	$8000-$16000	Female: ~$5000 (Services include: Consult, ovarian reserve testing AMH, Estradiol, TSH, FSH, Antral Follicle count US, HSG with mock sonogram transfer ($2047), baseline ultrasound, infectious disease panel, blood typing, antibody screen) Male: ~$2000 (Infectious disease panel and semen analysis)	Chromosomal analysis Karyotyping: $2536 Endometrial receptivity test (ERA): $460 Coupled with IVF sampling cycle: ~$4000	Embryo storage: $440/yr (Multiple years on average)	FET Cycle: $6500 Medicine: $1500-3500
Shady Grove Fertility (Various Locations)	$13,600	$4000-$8000	Similar	Similar	Similar	Similar
Colorado Center for Reproductive Medicine (Various Locations)	$17,745	$4000-$7000	Similar	Similar	Similar	FET Cycle: $5895 Medicine: $2000-$3000

Table 4 outlines the various cost components of the IVF cycle. Diagnostic and genomic services can vary by individual based on necessity. (FET: Frozen Embryo Transfer, IVF: In Vitro Fertilization, Cryo: Cryopreservation services for embryos). Source: Self-reported by conducting survey of one private clinic (single location) and three leading national level private clinics (with multiple locations).

South Asia primarily consists of low-and middle-income developing countries which are confronted with systemic health infrastructure challenges, prevalence of disease, lack of welfare services, and limited health care insurance. In the absence of financing and health care insurance models, the primary route to achieve IVF treatment is solely through prohibitive out-of-pocket expense and is often financially catastrophic to families. The absence of government ART policies and regulation results in fertility clinics operating without oversight further worsening the infertility treatment landscape.

The lack of representation in clinical trials and genome wide studies results in treatment options that may produce lower success rates and higher adverse events associated with ART treatment for South Asian populations, as these studies are primarily focused on European populations. By 2050, 155 out of 204 countries will not be able to sustain population sizes as global infertility rates rise (Fertility and Forecasting Collaborators, 2024). The figure below shows the trajectory of global fertility rates categorized by region.

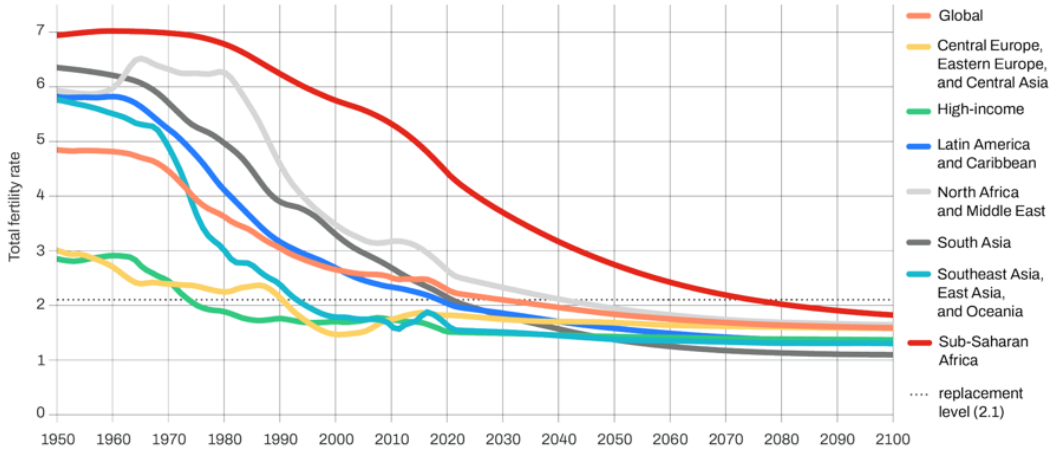


Figure 1 Global Total Fertility Rates 1950-2100

Figure shows a global decline in fertility rates which will not be adequate enough to sustain population growth over time in 97% of countries. Source: GBD 2021 Fertility and Forecasting Collaborators, 2024

Developed countries are better postured to deal with this changing fertility landscape and developing countries may face additional challenges in this health arena owing to immature health eco systems and lack of health equity in cutting edge research and clinical trials.

The Standard One Size Fits All IVF Protocol

Background

Since its introduction in 1978, in vitro fertilization (IVF) has revolutionized human procreation. Initially developed to assist infertile couples, the clinical applications of IVF have expanded rapidly to treat medical and genetic conditions, as well as fertility preservation. Although IVF accessibility and usage vary globally, it now contributes to over 1.9% of all newborns in the USA, 5% in certain European countries, 4.1% in Australia and New Zealand, and 1.7% in China. Changing demographics and societal norms are also fueling increased IVF use (Kushnir, 2022).

Improved access to education and career opportunities for women, coupled with effective contraception, has led to delayed childbearing and lower overall fertility rates globally. In many regions, fertility rates are now well below population replacement levels. The average age at first birth has surpassed 30 years in numerous metropolitan areas and highly developed countries, extending beyond peak fertility in the mid-20s. Consequently, more women are delaying childbearing, leading to age-related fertility decline and increased demand for fertility treatments like oocyte cryopreservation and IVF. Environmental changes in the form of sedentary lifestyles, processed diets, and population migration resulting from globalization may be influential factors in the gene-environment interaction which could influence the expression of phenotypes. It is imperative to take dynamic epigenetic factors of infertility etiology into consideration when treating certain ethnicities as this can impact gene expression. The role of genetic variability and the epigenome on infertility associated PCOS, implantation,

endometriosis, and unexplained infertility is an area of research that could lead to better fertility outcomes through a personalized medicine approach.

The Average IVF Cycle

IVF treatments are standard across clinics with some adjustments for age and ovarian reserve. Standardized protocols control costs for clinics and patients by reducing the need for customization, increasing clinical operational efficiency, standardizing staff trainings, ensuring patient safety, protocol efficacy and compliance with regulatory guidelines. An average IVF cycle involves several key steps which include the following:

1. Initial Consult and Diagnostic Bloodwork: This involves discussing plans, schedules, coordinating hormonal, blood count and infectious disease lab work, conducting sperm analysis, ultrasounds and scheduling the mock embryo transfer.
2. Ovarian Stimulation: The patient takes birth control to suppress uterine activity and begins fertility medications (exogenous hormones Follicle Stimulating Hormone and Luteinizing Hormones) to stimulate the ovaries to produce multiple eggs.
3. Monitoring: The progress of follicle growth is monitored using ultrasounds and blood tests.
4. Egg Retrieval: Once the follicles are mature, eggs are retrieved from the ovaries using a minor surgical procedure under sedation.
5. Uterine Lining: The uterine lining is bolstered to prepare the uterus for implantation by a prescribed dosage of Estrogen and Progesterone

6. Fertilization: The retrieved eggs are then fertilized with sperm in a laboratory, either through conventional insemination (mixing eggs and sperm) or intracytoplasmic sperm injection (ICSI) where a single sperm is directly injected into each egg.
7. Embryo Culture: The fertilized eggs (embryos) are cultured in the laboratory for several days.
8. Embryo Transfer: One or more embryos are selected and transferred into the woman's uterus using a thin catheter.
9. Luteal Phase Support: Medications are given to support the luteal phase, the time between embryo transfer and the pregnancy test.
10. Pregnancy Test: A blood test is preformed approximately two weeks after embryo transfer to determine if the cycle was successful.
11. Follow-Up: If the cycle is successful, pregnancy monitoring continues. If not, discussions about future treatment options may take place.

A typical IVF cycle from initial consult to pregnancy check takes 6-8 weeks and sometimes can take up to 4 months. As mentioned before, the success rate of live birth is on average is approximately 24% which means that many patients will have to undergo Frozen Embryo Transfer cycles if good quality embryos were achieved in the initial IVF cycle or may have to go through IVF all over again. These cycles are time intensive, expensive and take a significant emotional toll on the patient. The schematic below provides a visual representation of key milestones in a typical IVF lifecycle from initiation to pregnancy if successful. If the initial IVF cycle does not result in pregnancy

or results in pregnancy loss, then a frozen embryo cycle can be added on which is less intensive than a complete IVF cycle, as it removes the need for ovarian stimulation and egg retrieval. This, however, can only be added on if good quality viable embryos were collected in the initial IVF cycle.

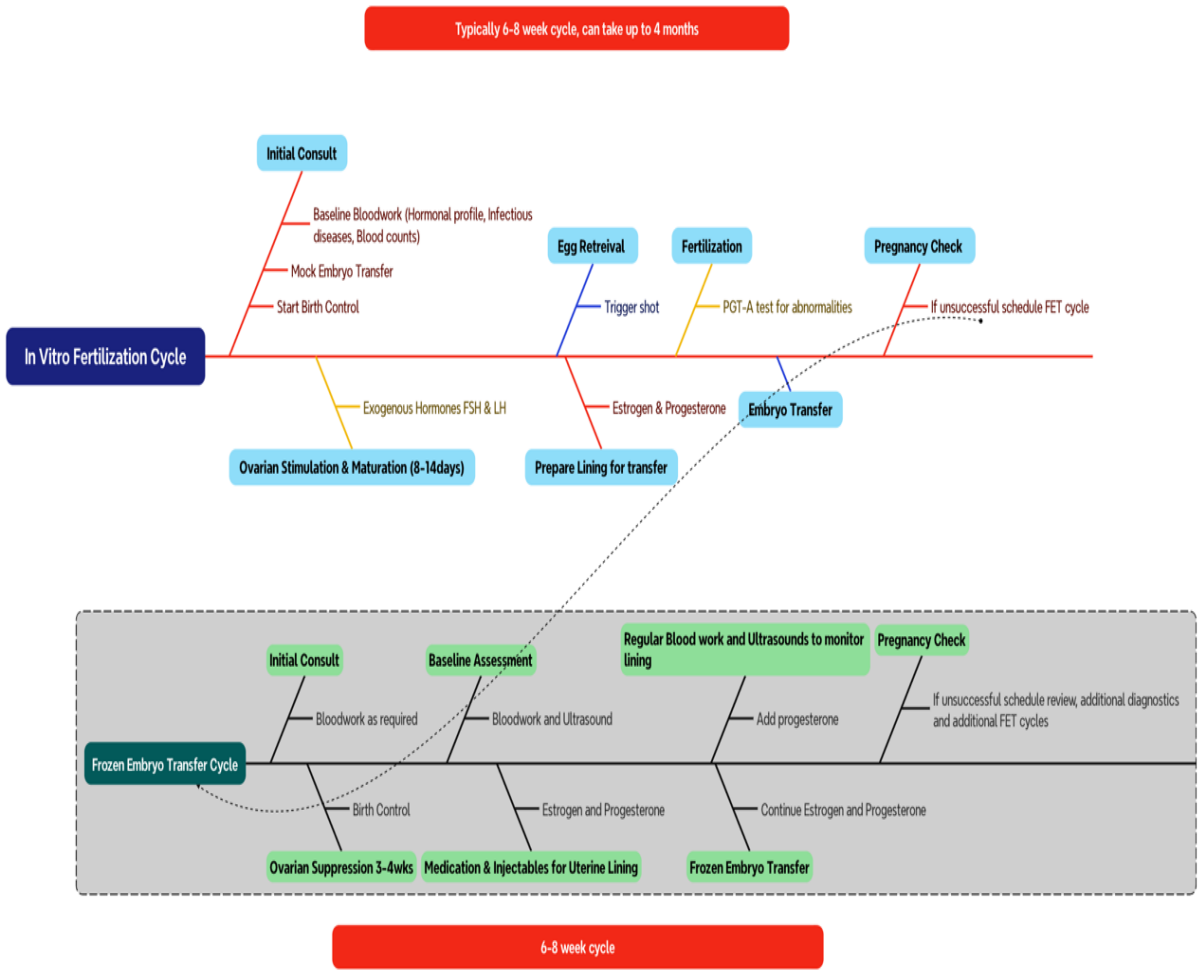


Figure 2. Schematic of a Standard IVF Cycle, Timelines, and Key Milestones

Figure based on analyzing standard IVF protocols and progressive Frozen Embryo Transfer protocols obtained from Shady Grove and Colorado Center for Medicine, see appendix. Source: Self-reported based on protocols from Shady Grove Fertility, CCRM and Mayo Clinic.

Adjunct Treatments

In the case of repeat failure, IVF patients may be presented with poorly regulated adjunct treatments and billed for various additional diagnostics/treatments that are suggested to enhance the likelihood of a successful live birth. Clinical Adjunct treatments may include growth hormone, aspirin, heparin, dehydroepiandrostenedione, testosterone, male and female antioxidants, and screening hysteroscopies (Kamath, 2019). However, these treatments may lack clinical evidence to substantiate their effectiveness and increase the overall cost. Despite this, most individuals experiencing infertility, particularly those whom have experienced repeat failures, tend to readily accept these additional treatments. This can add additional layers of complexity, financial burden, and emotional stress.

IVF Associated Risks

Standardized protocols have traditionally yielded success rates that border around 24% and can drastically change with age as reported by the CDC. While universal protocols ensure minimal deviation and contribute to patient safety, these protocols are applied indiscriminately on patients from all ethnicity backgrounds and pre-existing conditions (PCOS, unidentified infertility, uterine issues etc.). Additionally, dosage modulation is not a clinical practice geared towards patients with varying BMIs who end up receiving the same dosage of hormones for stimulation and uterine augmentation. Some infertility patients may have entirely different root causes of infertility and additionally may require dosage modulation owing to their age, body mass or predisposition to adverse effects.

The standard IVF procedure is not replete of risks. Some adverse events include serious complications with dosage related ovarian hyper stimulation syndrome (OHSS), multiple pregnancies which have increased risks for mothers and babies, preterm births, overuse of ineffective procedures, and ethical considerations pertaining to donor gametes and embryos (Sun, 2021). Additionally, the biological risks are also coupled with psychological ones in the form of trauma, anxiety, and depression. Not to mention financial repercussions in the form of increased costs due to multiple generic cycles that don't produce desired success outcomes.

Ovarian Hyper Stimulation Syndrome

Ovarian Hyper Stimulation Syndrome (OHSS) is a serious complication that is associated with ovarian stimulation during IVF. The syndrome is characterized by a spectrum of symptoms which include discomfort, dyspnea, abdominal distention, ovarian enlargement, ascites, hemoconcentration, hypercoagulability, and electrolyte imbalances. OHSS can be characterized by mild, moderate, severe, or critical severity and can also be classified by the timing of onset (early or late).

While severe OHSS affects 2%-6% of patients in a general population, certain groups, such as those with polycystic ovaries, face a five-fold higher risk. Additionally, patients with PCOS exhibit higher antral follicle counts (AFC), elevated anti-Mullerian Hormone (AMH) and serum Estradiol (E2) levels (Sun 2020).

Patients with at least 23 antral follicles have a four-fold increased risk of clinically significant OHSS, and studies show rates of severe OHSS in women with polycystic ovaries ranging from approximately 11% to 16% (Abbara, 2018).

Young age, a history of allergies, pre-existing PCOS, high doses of gonadotropins, high E2 levels, low body mass index, elevated AMH, high AFC are all risk factors for OHSS development. Currently Clinicians do not factor these variables when deciding dosage of IVF protocols. Additionally, OHSS is an underreported side effect in patients and the mandate for reporting this risk is not stringent or enforced adequately. Patients with mild and moderate OHSS are often left to resolve their symptoms naturally. OHSS has a negative effect on oocyte quality, embryo quality, and overall success during the IVF cycle. Patients are also recommended to wait an additional time period for their bodies to normalize after OHSS before scheduling additional rounds of ART which can extend treatment timelines.

Clinicians need to identify patients who stand at a high risk for developing OHSS and customize evidence-based strategies to educate patients of potential risks, customize protocols and prevent OHSS onset.

Multiple Pregnancies, Ectopic Pregnancy, Miscarriage, and Birth Defects

IVF increases the risk of multiple pregnancies, which are associated with higher rates of complications for both the mothers and babies, including premature birth, low birth weight, and developmental issues.

There is a slightly higher risk of ectopic pregnancy with IVF, where the fertilized egg implants outside the uterus, typically in the fallopian tubes. This can be a medical emergency and may require treatment to prevent complications. The risk of miscarriage is slightly higher with IVF compared to natural conception, especially in older women or those with certain medical conditions. (Hu, 2018)

While the overall risk of birth defects is low, some studies suggest a slightly increased risk of certain birth defects in babies conceived through IVF. However, it's important to note that most IVF babies are born healthy. Some studies suggest a possible link between certain types of ovarian stimulation used in IVF and a slightly increased risk of ovarian cancer due fertility drugs and dosage exposure (Rizutto, 2019). However, more research is needed to confirm this association as information and cohort sizes previously examined are insufficient and there is not enough conclusive evidence (Stewart, 2013).

Financial and Psychological Impacts

The emotional toll of undergoing IVF can be significant, especially if the treatment is not successful. IVF can be expensive, and the financial burden can be a significant consideration for many couples as described above. Additionally, insurance coverage for IVF varies widely, and not all treatments may be covered as previously elucidated. That coupled with the success rates associated with age can result in multiple cycles before a live birth is achieved. Existing Infertility treatment plans do not address the psychological consequences associated with emotional responses, stigmatization, and stress which can further impact success outcomes (Cousineau, 2007).

The investment of time, money, emotion, and physical health often leads to patients that feel resigned to their fates and abandon hope or seek homeopathic or alternative medicine regimens that are not regulated. Counseling and support services may be beneficial for individuals and couples undergoing IVF to improve success outcomes and better patient management.

Chapter II

Challenging the One Size Fits All Approach

Pharmacogenomics is the study of how an individual's genomic makeup affects their response to drug metabolism, therapy, and the potential to develop adverse side effects after exposure to drugs (MOC, 2020). The completion of the Human Genome Project has allowed scientists to characterize variations in the human genome and catalogue and identify common variants which include Single Nucleotide Polymorphisms (SNPs), insertions/deletions, copy number variants, large transposed sequences etc. Population based genome analysis using next-generation sequencing has made it possible to catalogue rare variants in any given population and comprehensive genome reference data sets are instrumental in creating comparisons (Southam, 2017).

Personalized or Precision medicine is an emerging practice and medical model that recognizes the individual's unique genetic makeup and uses this information to couple medical decisions, practices, and interventions, keeping in mind their predicted response or risk of disease. This practice takes a nuanced approach to incorporate the molecular, physiological, environmental exposure, behavioral factors, and other unique variations by employing DNA sequencing, proteomics, imaging protocols and health monitoring to define custom therapeutic interventions (Goetz, 2018).

The figure below shows a graphical representation of the multifactorial aspects of an individual's health and the technologies that can predict/monitor informed outcomes.

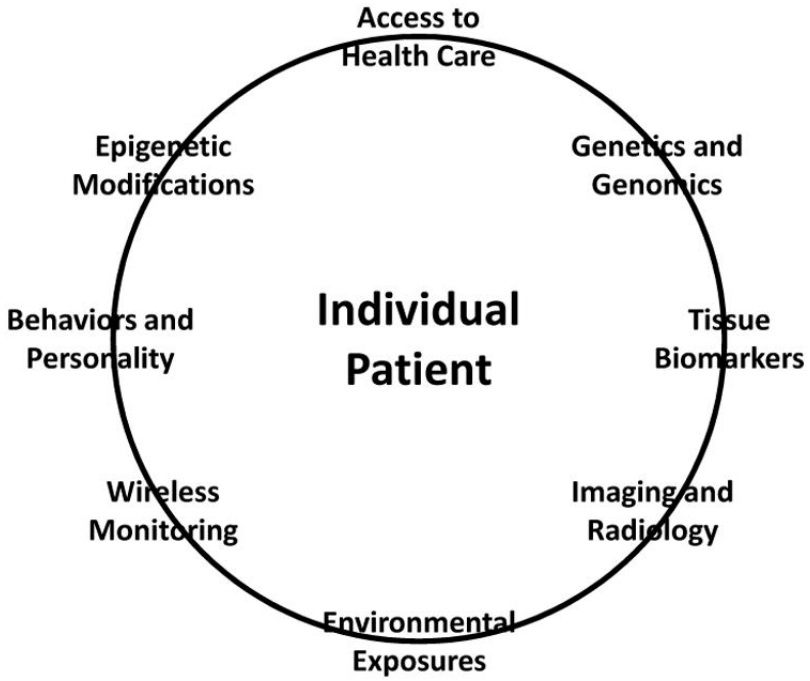


Figure 3. Factors Impacting the Individual's Health from a Personalized Medicine Lens

This graphical depiction summarizes the many inputs that define health outcomes and the technological arsenal of tools that can be utilized to impact successful health outcomes. Source (Goetz, 2018)

Ethnicity and inter ethnicity can affect differences in drug responses and toxicity. At a high level, to elucidate the presence of ethnicity variations and their impact on therapeutic outcomes, a study reported poor prognoses with in vitro fertilization in South Asian women compared to Caucasian women despite similar embryo quality (Shahine, 2009). The cause of this disparity remains unclear and requires additional investigation. Additionally, South Asians also have a markedly higher risk for diabetes and mortality

from cardiovascular disease compared to Caucasian populations (Bhopal, 2000). A higher rate of insulin resistance and increase in the incidence of polycystic ovarian syndrome (PCOS) has been shown in Indian women as compared to Caucasian women (Rodin, 1998).

Inter-ethnicity differences also exist and can impact drug mediated response and therapeutic outcomes. Identification of these variances are particularly important in regions with heterogenous ethnicity groups that are categorized together based on geographical regions such as Europe, South Asia, East Asia etc. South Asia has several distinct populations owing to migrations and invasions, which are lumped into one category. To expound the point of how inter-ethnicity affects predicted drug phenotypes, a study explored the pharmacogenomic variation in Latin American populations. The study demonstrated differences in drug metabolism in sub-groups based on allele frequency variation in CYP (Cytochrome P450) genes. Native American and Ibero-American populations revealed significant drug metabolism differences based on the variation of allele frequencies between the two populations. Additionally, differences between the North and South Native American population were also observed (Naranjo, 2018). Furthermore, in another study to support the need for personalized medicine, the inter-ethnic genomic differences (substantial genetic variations at 100 pharmacogenomic loci) of Southeast Asians populations were found to be significant enough to be taken into consideration to reliably predict drug safety and efficacy at the population level (Runcharoen, 2021).

Current IVF and IUI protocols are a one size fits all approach for patients and does not account for inter or intra ethnicity variation, and the standard approach may be

responsible for the low success rates observed year after year. Population pharmacogenomics is a promising frontier for global health and there is enough compounding evidence to support the need for protocol customization based on the individuals pharmacogenomic profile to improve ART outcomes. This area of focus will be explored in greater detail in the following sections.

Clinical Trial Representation

Of the 17 Infertility studies that currently exist on clinicaltrials.gov that were filtered for the following criteria 1) Female 2) Infertility and unexplained infertility treatment and 3) IVF, only 4 trials were conducted in the USA that had no specific criteria for ethnic inclusion (www.clinicaltrials.gov). This corroborates the premise that along with the standard IVF protocols offered by fertility centers, the existing IVF approach is not an inclusive one that incorporates population diversity. This may be contributing to the stagnant success rates and IVF risks observed in certain ethnic groups. It also raises concerns regarding the translatability of data across various ancestry groups nationally since the USA is a heterogenous mix of populations. There are subsequent global ramifications as countries rely on and adopt American innovation to treat infertility.

Commoditization of the IVF Industry

The absence of consistent insurance coverage for infertility, federal subsidization of infertility, and the privatization of IVF clinics has resulted in a rise in commercial entities driven by profits instead of research and diagnostic medicine. IVF practices such as Shady Grove Fertility and CCRM have metamorphosized into thriving commercial

national enterprises serving multiple locations and catering to staggering client volumes. Physician practices and hospitals have attracted private equity investments, and this has resulted in a surge of healthcare acquisitions. It is estimated that from 2010-2017 acquisitions increased 187% to $42.6 billion while health care deals increased by 48% (Gondi, 2019). According to Yahoo finance the global IVF market size will be worth $39.12 billion by 2032 with a compound annual growth rate of 5.9%.

As mentioned previously only seventeen states have passed laws that mandate insurance to offer some or all coverage for infertility diagnoses and treatments. In some states infertility treatment stops before IVF coverage and the burden remains largely out of pocket. This allows privately run clinics to operate purely on cash up front and necessitate payments for entire cycles to be made as a lump sum before treatment initiation, with no guarantee of success. This results in prohibitive cost barriers for most patients, limits access to lower socioeconomic groups and adds a layer of anxiety for couples hoping to achieve success on their first attempt with IVF.

Applicability and Translatability of Genome Wide Association Studies (GWAS) related discoveries to global populations

Genome wide association studies investigate associations between genetic variation that may be statistically associated with a disease risk or trait. They shed light on gene environment studies, Mendelian randomization studies, and polygenic risk scores (statistical calculation to predict medical conditions based on genomic variants).

GWAS can be utilized to gain insight into the biological underpinnings of a certain phenotype, estimation of heritability, and calculation of genetic correlations. This information can be leveraged to create clinical risk predictions, inform drug development

programs and infer potential causal relationships between risk factors and health outcomes (Uffelmann, 2021).

There are limitations to GWAS studies that restricts their statistical power and current approaches do not represent the complete spectrum of population diversity. Genetic associations may vary across ancestries and current studies are primarily European data as existing catalogued genomic databases do not contain equitable numbers of ancestry group data. The representation of European data versus South Asian data is in stark contrast. Recent commentaries and analysis suggest that biases (Baroso, 2021) related to population diversity. Prevailing attitudes towards inclusion criteria by scientists and institutions exist and may not only impact the equity of high impact genetics research but sometimes may exacerbate disparities (Martin, 2019). It is only in the recent past that genome databases are making efforts towards diversity inclusion to improve their data sets. These are particularly critical where noncommunicable diseases (NCDs) are analyzed for quantification as they are the leading causes of death and disease burdens for countries.

Existing genetic ancestry grouping has several limitations of bias, representation, and data contribution challenges. European participants are overrepresented in all data sets and there is poor representation of many populations and population sub-groups who may have rare variants which have not been analyzed and are of uncertain significance.

The gnomAD database is one of the most comprehensive, cross-linked aggregators of genomic information and provides variable views for investigation. However, parsing through several genes related to implantation failure, endometrial receptivity, PCOS and recurrent pregnancy loss, there were challenges with ancestry

group sample sizes, inconsistent ClinVar cross labeling, blank fields, and variant warnings (related to complex sequencing issues, variance presence in less than 50% of the population etc.). These inconsistencies contribute to limitations of the power of results generated. The figure below shows ancestry stratification represented by the commonly used reference databases and shows the proportion of population composition data sets available.

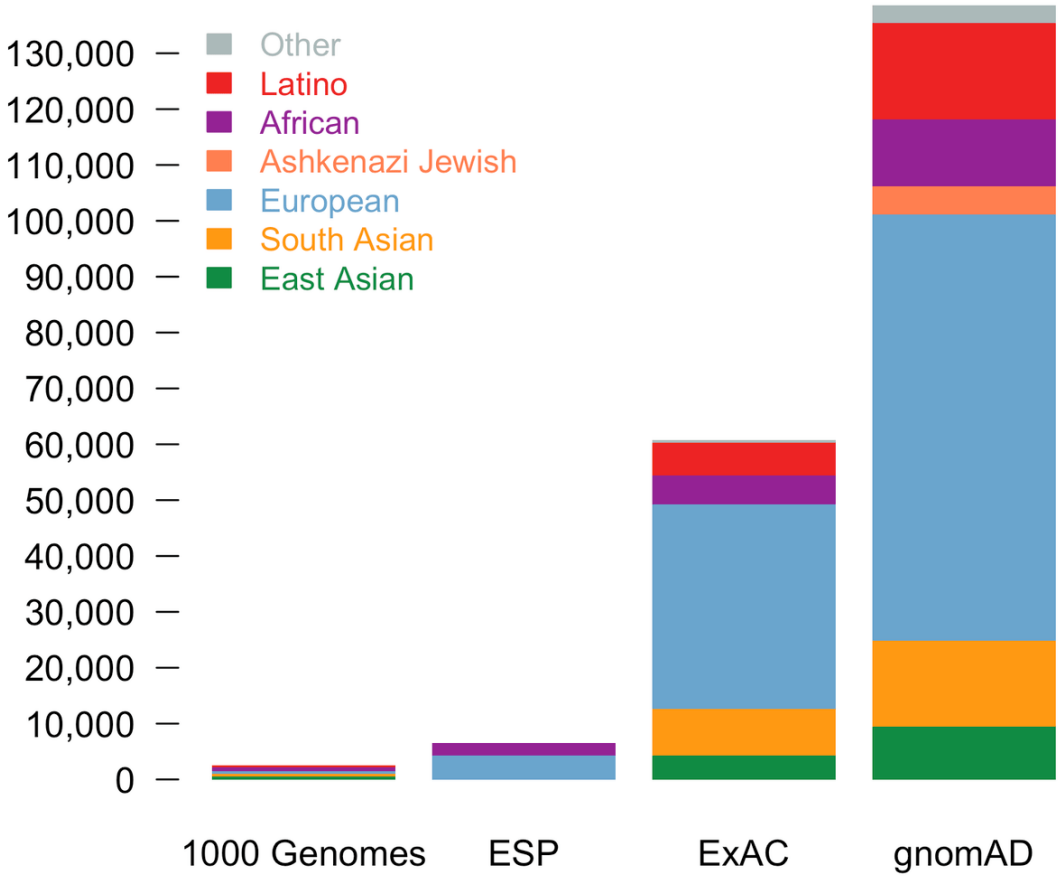


Figure 4. Ancestry Composition in Leading Databases.

Reference databases and their ancestry composition of key ancestry groups, Source: www.gnomad.com

European data sets are widely represented in stark contrast to South Asian data sets. Additionally, age and gender data is not discernable on the gnomAD interface as some of this data is available for a subset of individuals.

Despite the many glaring gaps of existing databases, genomic stratification of frequencies by group is important for variant classification and risk assessment. Certain variants occur at higher frequencies in specific genetic ancestries owing to historical events, geography, cultural practices, and this information can be leveraged to shed light on disease prevalence, disease risk and potential therapies (Gudmundsson, 2022).

The scientific community recognizes these disparities in representation and the effect this has on the data generated. The following figure below shows the fold increase of diverse populations in the reference population database gnomAD over time to build more robust data sets for analytical applicability and predictive accuracy.

The version 4 data set released in November 2023 also known as (GRCh38) currently spans 730,947 exome sequences and 76,215 whole-genome sequences from unrelated individuals of diverse ethnicities which have been sequenced, with a threefold increase in South Asian representation as shown below.

Population	ExAC	gnomAD v2	gnomAD v3	gnomAD v4*		
	#	#	#	#	%	Fold increase from v2
Admixed American	5,789	17,720	7,647	30,019	3.72%	1.7x
African	5,203	12,487	20,744	37,545	4.65%	3x
Ashkenazi Jewish	-	5,185	1,736	14,804	1.83%	2.9x
East Asian	4,327	9,977	2,604	22,448	2.78%	2.3x
European^	36,667	77,165	39,345	622,057	77.07%	8.1x
Middle Eastern	-	-	158	3,031	0.38%	19.2x
Remaining Individuals^	454	3,614	1,503	31,712	3.93%	8.8x
South Asian	8,256	15,308	2,419	45,546	5.64%	3x
Total	60,706	141,456	76,156	-	807,162	-

Figure 5. Ancestry Group Representation Improvements Over Time

Figure 5 represents the evolution of improvements in representation for ancestry groups in the gnomAD genomic database over time as a fold increase in each progressive version release. There is still a significant inequity of representation in data for various groups which results in an inadequate power of analysis. Source: www.gnomad.com

GWAS has the potential of revolutionizing the field of complex disease genetics and provide associations for complex traits, in particular the association between common single nucleotide polymorphisms (SNPs) and common diseases such as type 2 diseases, heart disease and psychiatric disorders. The hypothesis “common disease-common variant hypothesis” states that common diseases are likely influenced by genetic variation that may exist in a population. Common alleles have been shown to prove their role in susceptibility and diseases may be also governed my multiple alleles with small effect sizes. These may vary across different ethnicities and it is critical to improve the power of

GWASs in underrepresented populations and outline future strategies for GWAS meta - analyses.

Chapter III.

South Asia Focus: The Case for Custom ART Protocol Development

South Asia comprises of Afghanistan, Pakistan, India, Nepal, Bhutan, Bangladesh, Sri Lanka, and Maldives according to the World Bank. These countries are some the most densely populated geographical regions of the worlds and house one fourth of the world's population which is approximately 1.92 billion people.

Figure 6. Countries that make up South Asia.

A representation of populations comprising of the South Asian sub-continent. Source: The World Bank

According to the United States Census Bureau approximately 5,696,401 South Asians reside in the USA. South Asians constitute the largest ethnic minority in the United Kingdom. It is essential to better understand the impact of South Asians to ensure clinical treatments are well aligned to better health outcomes for this population.

South Asian Health Landscape: Unique Challenges

One fourth of the world's population lives in the South Asian countries combined and due to multiple factors, such as poor quality of health care, low access, lack of awareness, low literacy rates, nutritional deficits, cultural, linguistic, and socioeconomic challenges. South Asia suffers from significant disease burdens and trends in non-communicable diseases (NCDs) such as metabolic, cardiovascular, cancers, injuries, and mental health orders (Siegel, 2014). Coupled with the existing health disparities is the paucity of credible data on this population's unique pharmacogenomic make up, cancer risk profiles, and etiologic mechanisms. Additionally, inadequate representation in clinical trials and genomic databases results in greater risk of adverse reactions to drugs and therapies that have been tested in mainly European populations. Western models of diagnosis and therapy require investigations, and continuous modifications to achieve effective health outcomes. The South Asian population has unique biomarkers, nutritional epigenetics, lifestyle factors, and hazard exposure profiles that must be accounted for when identifying root cause analysis and therapeutic options for diseases.

The figure below provides a snapshot into the top ten risk factors for premature death and disability in South Asia juxtaposed to global levels. These emanate from poorly regulated working environments, hazard exposure, pollution, diet, and lifestyle challenges and are a testament to the sub optimal healthcare infrastructure and awareness.

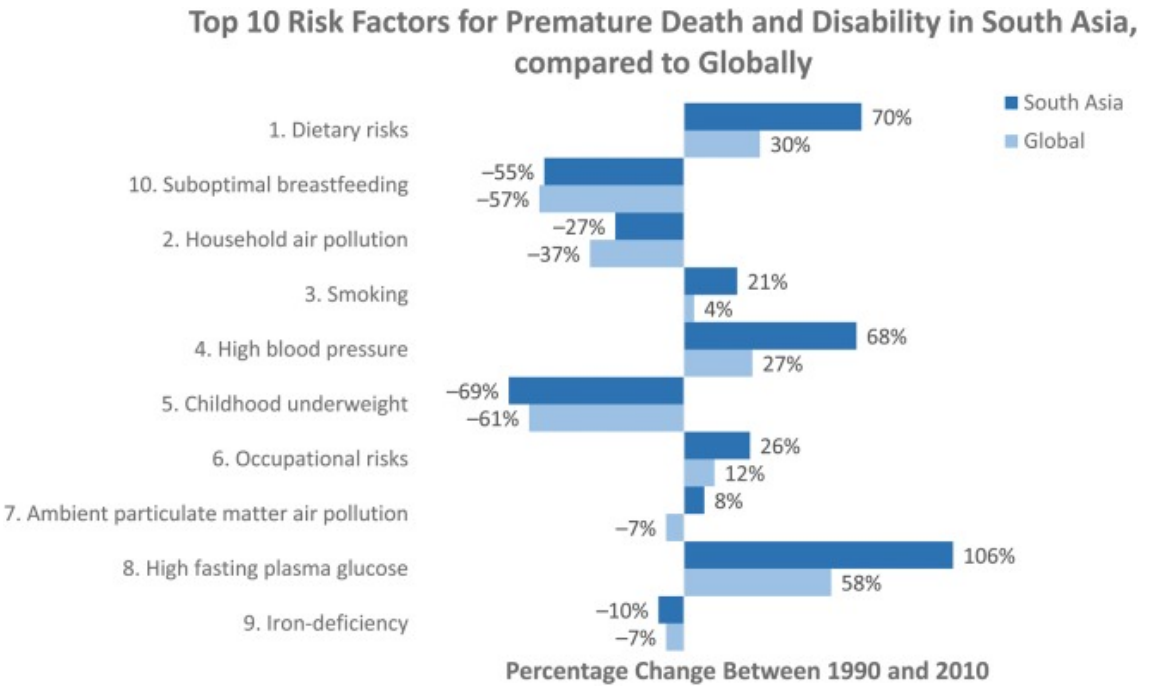


Figure 7. Relative change in Disability-adjusted life years (DALYs)

The figure shows relative change in DALYs for top 10 risk factors in South Asia, compared with global relative change between 1990 and 2010. Source*: Siegel KR, 2014*

Additionally, forward projections of diseases are expected to increase and three of the most populous countries in South Asia are on the top ten list of countries for diseases like diabetes (Wild, 2004). This statistic is of particular importance as women with PCOS are often insulin resistant and which increases their risk for type-2 diabetes and more than half of women with PCOS develop type 2 diabetes by the age of 40, according to the

CDC. This inextricable link between PCOS and diabetes onset has implications for infertility in the South Asian population.

Table 5. Countries with the Highest Estimated Cases for Diabetes 2000 and 2030.

Ranking	2000		2030	
	Country	People with diabetes (millions)	Country	People with diabetes (millions)
1	India	31.7	India	79.4
2	China	20.8	China	42.3
3	U.S.	17.7	U.S.	30.3
4	Indonesia	8.4	Indonesia	21.3
5	Japan	6.8	Pakistan	13.9
6	Pakistan	5.2	Brazil	11.3
7	Russian Federation	4.6	Bangladesh	11.1
8	Brazil	4.6	Japan	8.9
9	Italy	4.3	Philippines	7.8
10	Bangladesh	3.2	Egypt	6.7

The top ten countries include 3 of the major countries that are part of South Asia, India, Pakistan, and Bangladesh. Source: Wild, 2004

South Asians have a significantly higher risk of developing type-2 diabetes and cardiovascular disease compared to Europeans, with a 2- to 4-fold increased risk (Wild, 2004). However, they also exhibit a lower susceptibility to conditions like skin cancer. A major challenge in identifying genetic variants that influence disease susceptibility in the South Asian population has been the lack of understanding regarding their specific patterns of genetic variation.

The South Asian Genome and Population Diversity

At a genome level the gnomAD genomic database incorporates several sub populations within these countries which includes 45,546 individuals with the following sub populations which are meta data fields provided by contributing studies (https://gnomad.broadinstitute.org/stats).

Table 6. South Asian Sub Population Study Provided Labels in the gnomAD Database

Genetic Ancestry Group	Individuals	Study Provided labels	Study provided labels (<1%)
South Asian (SAS)	45,546	South Asian (61.38%) Indian (10.63%) Pakistani (5.79%) Other Asian (1.97%) No label provided (16.47%)	African, Asian, Asian or Asian British, Balochi, Bangladeshi, Bengali, Black, Brahui, British, Burusho, Caribbean, Chinese, Do not Know, Dutch, Finnish, French Canadian, Gujarati, Hazara, Indian Telugu, Kalash, Makrani, Mixed, Other, Other Black, Other Mixed, Other white, Pathan, Prefer not to answer, Punjabi, Qatari, Sindhi, Sri Lankan Tamil, Uygur, White, White and Asian, White and Black Caribbean

The table summarizes study provided labels that are attributed to South Asian (SAS) individuals in the gnomAD database. Column 3 shows the majority category labels and labels that make up less than 1% of labels in column 4. This gives a glimpse of label fidelity to the accuracy of data output from this database.

Whole genome sequencing of 168 South Asians and whole exome sequencing of 147 South Asians identified 12,962,155 autosomal sequence variants, including 2,946,861 new SNPs and 312,738 novel indels (Chambers, 2014).

In another study, whole genome sequence (WGS) data was gathered from 4806 individuals enrolled in healthcare systems in Pakistan, India, and Bangladesh. This dataset was supplemented by WGS data from 927 individuals belonging to isolated South Asian populations. The analysis revealed significant evidence of reproductive isolation, endogamy, and consanguinity (inbreeding) within the subcontinent, with varying levels across different regions. These factors contribute to exceptionally high rates of rare homozygotes, up to 100 times more than in outbred populations (Wall, 2023).

The potential for deep research is necessary for South Asia and its unique subgroups resulting from constant population migrations over time, catastrophic events, introduction of new populations through historical conquests. These micro populations have their own unique characteristics as these populations include castes, subgroups, and regions with diverse ancestry.

Pharmacogenomics of South Asia

IVF stimulation protocols often lead to ovarian hyperstimulation syndrome (OHSS) and result in cycle cancelation and poor ART outcomes. It is important to investigate the pharmacogenomic profile of South Asians to understand the mechanics of drug metabolism, and risk of adverse events in the context of IVF protocols. This will allow clinicians to customize dosage and select appropriate drugs geared towards the South Asian genome and build a case for pharmacogenomic approaches to infertility.

Studies conducted in India observed 91 drugs that show disruption transport, metabolism, targeting functions, and risk of adverse effects (Sahana, 2022). There are numerous highly differentiated SNPs, indels, and haplotypes which have not been assigned a high level of evidence for association with drug response or toxicity. The presence of 18 SNVs and 34 haplotype variants, including *HLA* alleles associated with 85 clinical annotations among Indians, may have relevance to the wider South Asian populations and has implications for forming informed guidelines for drug dosage or adverse drug reactions. Genotyping of *HLA* alleles prior to the phenytoin, abacavir, carbamazepine, and allopurinol therapies could prevent toxicity and improve patient outcomes. Polymorphisms in three variants were shown to affect warfarin pharmacodynamics.

Another study investigated the frequencies of pharmacogenetic variants and their clinical relevance amongst multiethnic groups (Balochi, Brahui, Burusho, Hazara, Kalash, Pashtun, Punjabi, and Sindhi) in Pakistan and uncovered 29 genes with 44 variants with high to moderate evidence of clinical association. These corresponded to drug metabolizing enzymes, transporters, and PGx gene regulators that are involved in drug response for acquired immune deficiency syndrome, transplantation, cancer, heart disease, and mental health therapy (Khan, 2022).

Implementing dosing guidelines based on variation in the PGx can inform the selection of alternate drugs with fewer population-level pathway disruptions and adverse events in South Asian populations. This is not only relevant to infectious and non-infectious disease treatment such as Diabetes, Hepatitis and Cardiovascular, but also has

applications for IVF. Research can inform the selection and optimal dosage regimens for drugs used in IVF, to reduce adverse events and improve ART outcomes.

Evidence of Variance Related to Infertility, Insulin Metabolism and PCOS in South Asians

Window of implantation

Patients with unidentified infertility, implantation failure and with repeated IVF failure have found some success in utilizing genomic diagnostics such as the Endometrial Receptivity Test to identify the precise Window of Implantation (WOI) for embryo transfer (Ruiz, 2013). Currently, in standard IVF protocols the WOI is treated as a generic time point (5 days post progesterone exposure) regardless of ethnicity. Studies suggest that the WOI may be unique and may be displaced for some women (Wilcox, 1999) (Ruiz, 2012). The window of implantation is a time period in which the uterus is receptive to the transferred embryo and involves complex genomic activation/inactivation profiles that prepares the uterus for implantation (Altmäe, 2017). Implantation consists of three stages a) Apposition, b) Invasion and c) Invasion. Various molecules, hormones, and cells are involved in embryo implantation. Complex dialogues occur between the blastocyst and the receptive endometrium during implantation. The figure below shows the complexity of signaling that occurs between the blastocyst and the endometrium.

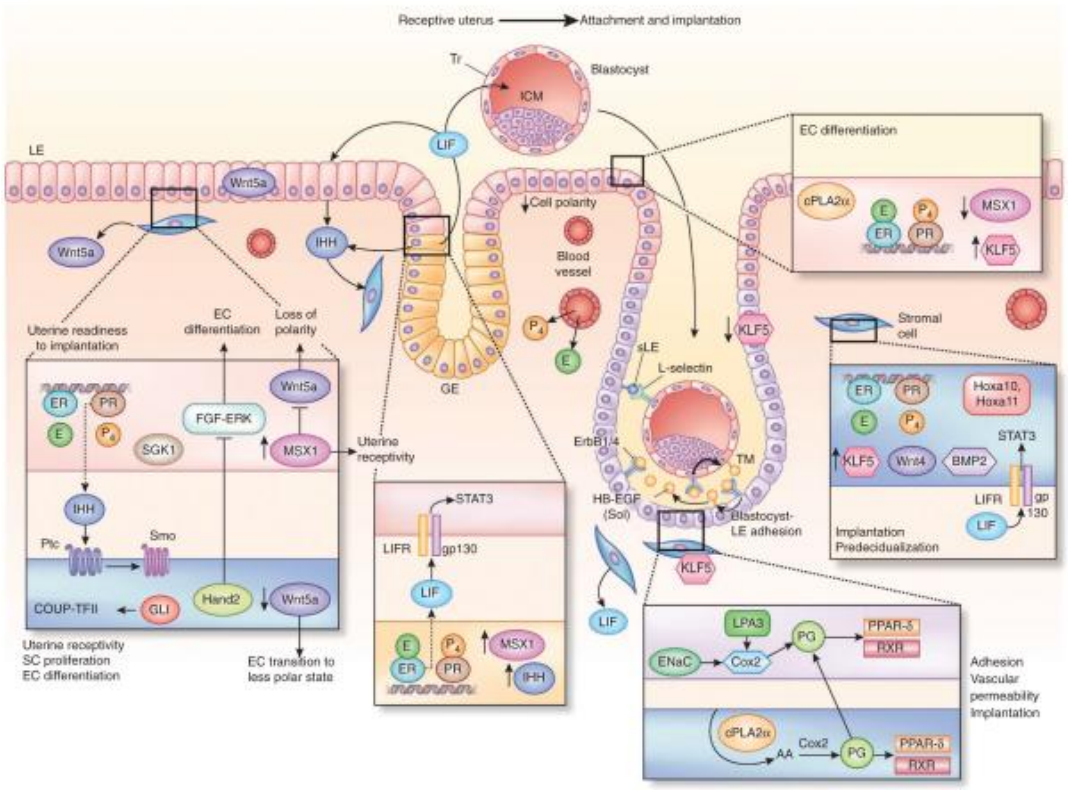


Figure 8. Signaling networks involved in uterine receptivity and implantation.

This schematic is based on mouse and human studies sheds light on the complexity of genes, molecules and pathways involved in uterine receptivity, implantation and decidualization. Source: (Cha, 2012)

The Endometrial Receptivity Test (ERA)

The ERA test is a bioinformatic predictor for endometrial dating, through the analysis of the transcriptomic profiles of approximately 248 selected genes. The ERA test identifies the receptivity of the uterus and the unique Window of Implantation (WOI) for the patient. Currently the Endometrial receptivity diagnostic test is not part of standard diagnostic testing in the IVF process and may be performed after repeat implantation failure at the discretion of the doctor or patient. This genomic analysis provides precise

and custom 12-hour windows of when the blastocyst should be transferred for implantation in the individual. This involves analysis of gene expression of the "suspected" implantation window compared against a reference to determine the exact timing of the transcriptomic signature that corresponds to the window of receptivity. The figure below shows the fold change in gene expression that is identified in an ERA test to designate Pre receptive, receptive and post receptive endometrial states.

Gene symbol	Gene name	Fold change	No. of probes
GPX3[a,b,c]	Glutathione peroxidase 3 (plasma)	35.49	2
PAEP[b,c]	Progestagen-associated endometrial protein (placental protein 14, pregnancy-associated endometrial alpha-2-globulin, alpha uterine protein) (PAEP), transcript variant 2	31.43	1
COMP[c]	Cartilage oligomeric matrix protein	30.95	2
SLC1A1[c]	Solute carrier family 1 (neuronal/epithelial high affinity glutamate transporter, system Xag), member 1	17.57	3
LIF[b]	Leukemia inhibitory factor (cholinergic differentiation factor)	15.03	3
TCN1[c]	Transcobalamin I (vitamin B12-binding protein, R binder family)	14.76	1
CXCL14[c]	Chemokine (C-X-C motif) ligand 14	14.02	2
C4BPA[c]	Complement component 4 binding protein, alpha	13.14	2
TSPAN8[c]	Tetraspanin 8	12.90	2
LAMB3[c]	Laminin, beta 3, transcript variant 2	11.32	1
MAOA[c]	Monoamine oxidase A, nuclear gene encoding mitochondrial protein	9.39	2
SOD2[c]	Superoxide dismutase 2, mitochondrial, nuclear gene encoding mitochondrial protein, transcript variant 2	9.06	2
GADD45A[c]	Growth arrest and DNA damage inducible, alpha	8.25	1
MUC16	Mucin 16, cell surface associated	8.01	8
THBD[c]	Thrombomodulin	7.84	3
NNMT[c]	Nicotinamide *N*-methyltransferase	7.74	2
DPP4[b,c]	Dipeptidylpeptidase 4 (CD26, adenosine deaminase complexing protein 2)	7.72	3
SCGB2A2[c]	Secretoglobin, family 2A, member 2	7.43	2
S100P[c]	S100 calcium-binding protein P	6.95	1
SNX10[c]	Sorting nexin 10	6.56	2
CP[c]	Ceruloplasmin (ferroxidase)	6.34	2
G0S2	Putative lymphocyte G0/G1 switch gene	6.20	2
C4.4A[c]	GPI-anchored metastasis-associated protein homologue	6.03	1
ANG[c]	Angiogenin, ribonuclease, RNase A family, 5	5.98	2
ABCC3[c]	ATP-binding cassette, subfamily C (CFTR/MRP), member 3	5.98	1
XCL1	Chemokine (C motif) ligand 1	5.80	3
ADRA2A	Adrenergic, alpha-2A, receptor	5.78	2
EFNA1	Ephrin-A1, transcript variant 1	5.77	3
KLRC1	Killer cell lectin-like receptor subfamily C, member 1, transcript variant 2	5.75	2
TAGLN[c]	Transgelin	5.71	3
SLC15A1	Solute carrier family 15 (oligopeptide transporter), member 1	5.59	2
IGFBP1	Insulin-like growth factor-binding protein 1	5.35	3
PTGER2	Prostaglandin E receptor 2 (subtype EP2), 53 kDa	5.18	2
THBS2	Thrombospondin 2	5.17	3
HPSE[c]	Heparanase	5.17	1
SERPING1[c]	Serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary), transcript variant 1	5.16	2
CFD[c]	D component of complement (adipsin) (DF)	5.13	2
CRISP3	Cysteine-rich secretory protein 3	5.09	1
C14orf161[c]	Chromosome 14 open reading frame 161	5.07	2
RPRM[c]	Reprimo, TP53 dependent G2 arrest mediator candidate	5.03	3
GAST[c]	Gastrin	5.00	2
EDNRB[b,c]	Endothelin receptor type B, transcript variant 2	4.89	4
RARRES1	Retinoic acid receptor responder (tazarotene induced) 1, transcript variant 2	4.87	1
ABLIM3	Actin-binding LIM protein family, member 3	4.87	4
GLIPR1	GLI pathogenesis-related 1 (glioma)	4.83	2
AOX1[c]	Aldehyde oxidase 1	4.82	2
CYP3A5	Cytochrome P450, family 3, subfamily A, polypeptide 5	4.82	3
CTSW	Cathepsin W	4.78	3
DHRS3	Dehydrogenase/reductase (SDR family) member 3	4.72	1
MYL9	Myosin, light polypeptide 9, regulatory, transcript variant 2	4.72	3
CLU[c]	Clusterin (CLU), transcript variant 2	4.70	2
IER3	Immediate early response 3, transcript variant long	4.69	4
GPRC5C	G protein-coupled receptor, family C, group 5, member C, transcript variant 1	4.69	6
C3	Complement component 3	4.67	3

Diaz-Gimeno. Customized array for endometrial receptivity. Fertil Steril 2011.

Figure 9. List of Genes involved in Endometrial Receptivity

Complete understanding of molecular pathways and genes involved in implantation is required to improve ART outcomes. Identification of genomic variances between ancestry groups is essential to be able to pinpoint the most opportune time for blastocyst transfer to increase success of ART outcomes.

Fresh Versus Frozen Embryo Transfer

In a retrospective cohort study, significantly better IVF outcomes (live births) have been reported using frozen versus fresh embryo transfers in South Asians compared to Caucasian women. The live birth rate was significantly lower in South Asians at 21% as compared to 37% in Caucasians upon fresh blastocyst transfer (Shah, 2016). These differences may be attributed to South Asian specific variances in the endometrial environment in response to hormonal stimulation in a fresh cycle. Another study demonstrated higher cancelation rates and lower live birth rates (9.1% vs. 22.7%) in South Asian versus Caucasian women (Mahmud, 1995).

Vitamin D Deficiency

South Asians along with Middle Eastern populations have high rates of Vitamin D deficiency, reported to be in the range of 67-82% in South Asia (Harinarayan, 2011). Additionally, the prevalence of Vitamin D deficiency is high in infertile women with PCOS and stimulates the anti-Mullerian hormone (AMH) production (Dabrowski, 2015). Women with high serum levels of AMH are at a higher risk for OHSS which can result in cycle cancellation, increased risk of miscarriage and poor pregnancy outcomes. Vitamin D receptors are present in the endometrium, placenta, ovaries, and pituitary glands. As a result this has implications in implantation and endometrial receptivity. Additionally, GC gene polymorphisms (rs17467825, rs3755967, rs2282679, rs7041 and rs2298850) have been found to be associated with 25(OH)D serum levels in Arabs and South Asians (Elkum, 2014).

PCOS

Stein-Leventhal syndrome, also known as PCOS, is a common endocrine disorder impacting 6-10% of reproductive age women globally and is one of the most common causes of infertility. PCOS is a complex disorder with a multifactorial etiology that includes ovarian function, hormonal dysregulation, metabolism, insulin signaling and multiple genetic risk alleles and has high heritability.

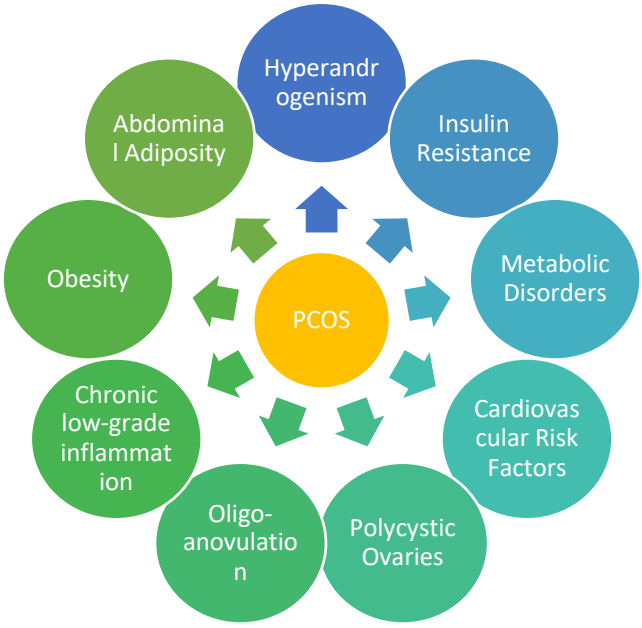


Figure 10. PCOS Associated Risk Factors and Disease Manifestations

PCOS is also linked to conditions including stroke, heart disease, high blood pressure, gestational diabetes, sleep apnea, cholesterol, depression, and anxiety etc. Additionally, according to the CDC more than half of women with PCOS develop type 2 diabetes by age 40. PCOS presents a broad spectrum of indications, and it is estimated that approximately 70% of people with PCOS remain undiagnosed attributing to poor quality of life including financial and societal burdens (March, 2010). This also means

that overall criticality and burden of this disease is underestimated and poorly understood. The figure below highlights the genetic, hormonal, environmental, and lifestyle factors that are inextricably connected in PCOS.

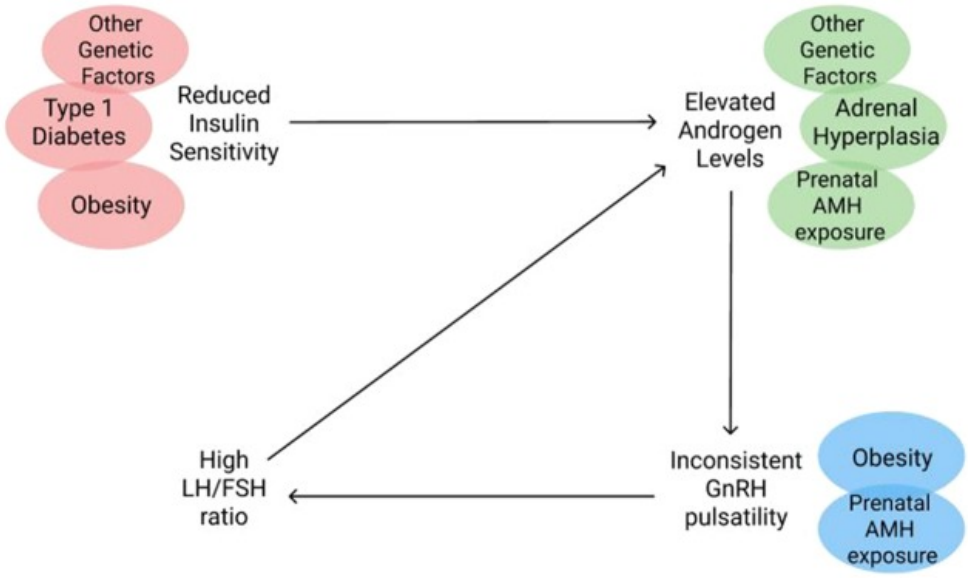


Figure 11. Feedback Mechanisms Involved in Hormonal Disruption Associated with PCOS

The figure highlights the cross linkage of genetic factors, obesity, prenatal exposures and associated hormone in the augmentation of the PCOS phenotype continuum. Source: Charifson, M.A., 2019

The predisposition of South Asians to diabetes, its prevalence, and the link between diabetes, PCOS and infertility, indicate that these factors should be weighed when treating South Asians for infertility. To better understand the unique nature of fertility in South Asians, literature searches were performed to capture data that establishes the population's markers in disease heritability, prognosis, and manifestation.

- Based on a study which sampled 212 South Asian Indian women (Aged 18-40), the prevalence of PCOS was at 52% and had significant clinical associations (Rodin 1998).
- A PCOS prevalence comparison in a community-based study in the United Kingdom between women of South Asian (WSA) and Caucasian women (WC) was significantly different between (WSA= 52% and WC=22%) (Clayton, 1992).
- South Asian women with PCOS presenting anovular PCOS were significantly younger, had more severe hirsuitism, higher prevalence of acanthosis nigricans, higher insulin concentrations and lower insulin sensitivity compared to Caucasian women with PCOS (Wijeyaratne 2002)
- South Asian women exhibit a higher prevalence of insulin resistance and type 2 diabetes which may increase long term morbidity among those with PCOS (Wijeyaratne 2002).
- The mitochondrial genome is a hot spot for acquired mutations and the D loop is a major regulatory site. South Indian patients with PCOS showed significantly lower mtDNA copy number carrying D310 and 189G alleles when compared to non-carriers and had significantly elevated LH/FSH ratios which constitutes an inheritable risk factor for PCOS in this subgroup (Reddy, 2017).
- A clinical case-controlled study shows the heritability of PCOS in mice by establishing PCOS has transgenerational transmission of reproductive and metabolic dysfunction in the male progeny. Sequencing the F1-F3 sperm revealed differentially expressed small noncoding RNAs which highlights the effects of maternal hyperandrogenism and points to the underrepresented risk of transmission or

reproductive and metabolic dysfunction (Risal 2023). Extrapolating this to the South Asian genome highlights the exponential increase of PCOS related infertility and metabolic disorders in subsequent generations as population rates increase.

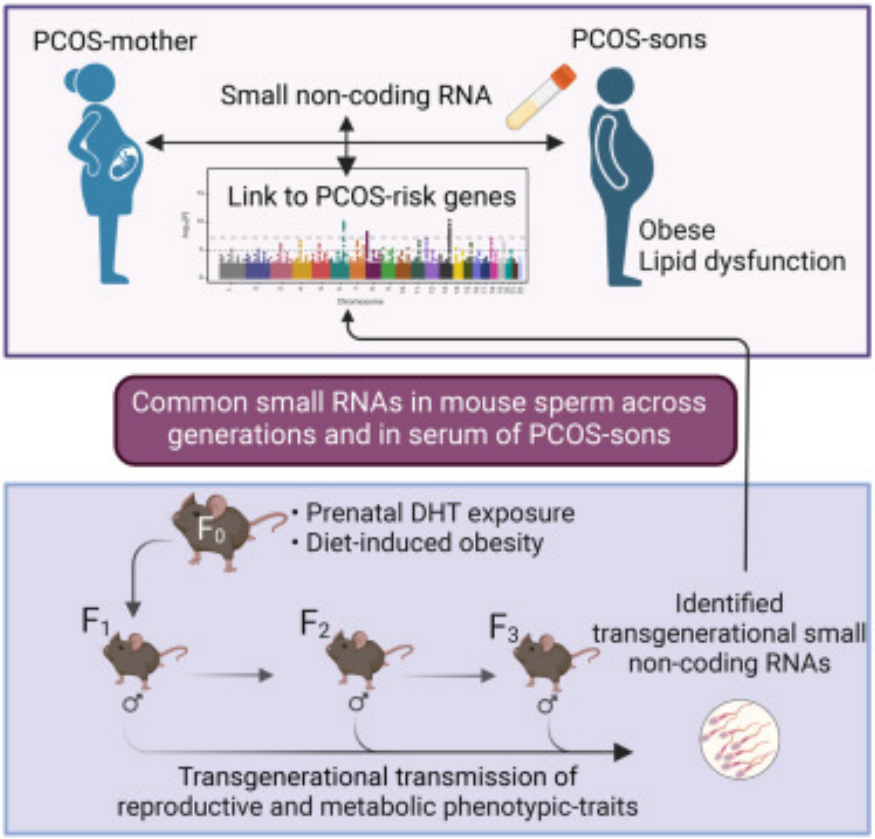


Figure 12. Transgenerational Transmission of PCOS.

Figure 9 shows the role of small non-coding RNA's in the transgenerational transmission of PCOS associated reproductive and metabolic phenotypic traits. Source: Flood D, 2022

Evidence from literature presented above highlight the theory that the South Asian population is indeed impacted differently by PCOS than other ethnicities. The

association between high prevalence of Diabetes, insulin resistance, PCOS and infertility is well documented. Additionally, the presence of PCOS severity, gene polymorphisms and the heritable nature of PCOS underscores the importance of exploring and characterizing the coding and non-coding regions of the human genome that control gene expression, and its variability in the South Asian genome. There is a critical need for diversified studies and custom treatments for South Asians when treating infertility patients with ART procedures.

This is even more critical considering that South Asian countries fall under the low- and middle-income countries (LMIC) category. A large majority of the population doesn't have access to sophisticated fertility clinics and cannot afford multiple rounds of expensive IVF treatments. Together, these findings support the development of a custom approach to alleviate infertility associated challenges and discovery of underlying associated pathologies such as PCOS, diabetes and other NCDs which impact the quality of life in these countries..

Urbanization Impacts on Infertility

Rural populations face a plethora of challenges owing to systemic multifactorial health inequities with disease prevalence and management and achieve significantly lower health system performance measures. This is especially true in the case of diabetes prevalence and management (Hales, 1992). The incidence of diabetes has increased rapidly in all rural areas globally with a higher relative growth in low-middle income countries (LMIC). The prevalence of diabetes in rural areas is estimated to be one-quarter that of urban areas for Bangladesh, Bhutan, India, the Maldives, Nepal, and Sri Lanka (Geldsetzer, 2022).

There is an increased trend of urban migrations of rural populations in search for superior economic opportunities which are then coupled with sedentary lifestyles associated with urbanization. Some studies suggest that the inherent genetic variation amongst South Asians involving gene clusters linked to core metabolic traits (lipid metabolism, adipogenesis, insulin signaling, type 2 diabetes etc.) that may have once been advantageous to survival during food scarcity, famines and rural lifestyles are now contributing to increased risks of obesity in urbanized environments and westernized societies (Hales, 1992). Considering the potential linkage between diabetes, insulin signaling, PCOS and infertility, this massive lifestyle shift is exerting its effects on infertility as well. There are intrinsic links between diet modifications, exercise, behavioral modification, and obesity in the improvement of reproductive disorders such as PCOS (Gu, 2022).

The mismatch between ancestral environments and current obesogenic environments influences the gene-environment interactions that emanate from higher

caloric intakes, low levels of physical activity, exacerbated by socioeconomic backgrounds, industrialized lifestyles, weakened healthcare infrastructure and the existing predispositions for PCOS. South Asians stand at an even greater risk for developing PCOS as parallels can be drawn to observations in the Northern Australian Aborigines population. This indigenous female population has undergone the transition from rural to urban where a significant relationship (Boyle, 2022) has been established between obesity, urbanization and a high prevalence of PCOS. In this population 21% of reproductive age Aborigines women have PCOS compared to the 6% global average.

Keeping the above multifactorial nature of the severity of the PCOS continuum, it is even more important to make informed health policy decisions for unique populations predisposed to PCOS and as a result infertility, match effective therapies that produce high success rates and low adverse effects to tackle the future of health issues on the South Asian infertility horizon.

Chapter IV

Genomic Analysis of Key Infertility Associated Genes

Implantation failure is one of the most frequent reasons for pregnancy loss and recurrent pregnancy loss after IVF. The cross talk between the embryo and endometrium is governed by genetic factors that have not been fully characterized. Genetic factors mediate invasion, regulation, and angiogenesis that can lead to implantation failure (Goodman 2008). Genetic polymorphisms are implicated in implantation failure and this area of research needs additional attention, especially where ancestry groups are concerned. Endometrial receptivity is a complex synchrony of timed gene expression as the body gears up for implantation and goes from pre receptive, receptive to post receptive state. It is this window of implantation in the receptive state that allows the embryo to successfully implant in the uterus.

To demonstrate genomic variability in South Asians, key genes were identified based on comprehensive literature searches for genes implicated in PCOS, Implantation, Endometrial Receptivity, and Recurrent Pregnancy Loss. To shed light on the allelic frequencies of pathogenic variations, these genes were investigated via the gnomAD genomic database to assess genomic variation in the "Pathogenic Allele Frequency" category in the South Asian population This was calculated against the backdrop of general allele frequencies encompassing all populations. As a caveat, rare variations also reside in areas that haven't been probed comprehensively such as the "Uncertain significance", "Benign/Likely Benign" and "Other Categories". Also, for the purposes of this study the focus was on "Predicted Loss of function" (pLof) variants, even though "Missense/Inframe indel", "Synonymous" and "Other categories" also merit

comprehensive investigation. Due to lower South Asian genome representation in the database in comparison to other populations (E.g., European), comprehensive data exists in some genes whereas in others there may be a general lack of availability of information for analysis. Additionally, this may be a general caveat for all populations as well. The presence of data errors warnings are highlighted in gnomAD which may present skewed allelic balances due to homozygous/heterozygous nature of the samples, number of participants, or other data issues. Such challenges are a testament to the need for additional research, precise cataloguing/labeling of genomic data, strategic and comprehensive study designs that can allow for better ancestry representation and robust modeling.

In this thesis, genes were investigated based on the level of information available to present differences in variant frequencies in populations. In some cases, there were several South Asian dominant pathogenic variants while in others they were not as relevant. The intent is to highlight these differences in a subset of genes to build a case for additional targeted research into not only pathogenic and curated variants but to build merit around investigating the other categories of "Uncertain significance", "Benign/Likely Benign" and "Other Categories". This will allow researchers to unearth rare variations, to shed light on potential diagnostic and therapeutic targets for South Asians, and to build a case for adequate inclusion in clinical trials, and genomic databases.

Chapter IV

Genomic Analysis of Genes of Interest

The World Health Organization classifies infertility as a disease of the reproductive system and is defined as the inability of a couple to achieve pregnancy after 12 months of sexual intercourse. Several factors contribute to infertility which includes endocrine conditions such as PCOS, age, obesity, male factor, genotypes, and karyotypes etc.

The molecular and genomic aspects of human infertility remain poorly understood and require additional research at the genome and exome level. In addition, polymorphisms must also be investigated and catalogued at the ancestry level to provide additional insights into population markers. Complex diseases such as PCOS involve multiple biochemical pathways in pathogenesis and the linkage to biosynthesis, metabolism, obesity, energy regulation, and insulin section. There are various interlinking factors that affect the expression of associated genes, that merit additional investigation.

A subset of genes found to be implicated in infertility associated issues consistent with implantation failure, recurrent pregnancy loss (RPL), PCOS, and endometrial receptivity. These genes include Leukemia Inhibitory Factor (LIF), Insulin-like Growth Factor 1(IGF-1), Homeobox A10/11 (HOXA10/11), and Estrogen Receptor 1 (ESR 1).

The gene associated variants were analyzed in the genome aggregator database gnomAD to highlight variation in the South Asian genome in comparison to the general population by calculating allele frequencies of tagged “predicted Loss of Function” (pLof) variants, “Pathogenic” variants tagged within the database and other variants that have been cited in literature with may be associated with infertility.

The table below describes 5 genes of interest and their role in infertility which were assessed in the genomic database gnomAD for allele frequencies in the South Asian ancestry group (SAS) compared to the general population frequency (other ancestry groups).

Table 7 Genes of Interest: Role in Associated Reproductive Issues

Genes of Interest	Brief Description Relevant to Fertility	Associated Reproductive Issues
LIF (Leukemia Inhibitory Factor)	Member of the interleukin 6 family of cytokines known for its pleiotropic effects, possesses the ability to trigger terminal differentiation in leukemic cells. LIF plays a crucial role in blastocyst implantation by preparing the endometrium for embryo attachment, thereby contributing to endometrial receptivity.	1)Implantation Failure 2)Recurrent Pregnancy Loss 3)Endometriosis 4)Preterm Birth 5)PCOS
IGF1 (Insulin-like Growth Factor 1):	This gene encodes a protein that shares functional and structural similarities with insulin, belonging to a protein family crucial for regulating growth and development. IGF-1 is involved in the development and maintenance of endometrial receptivity, vascularization, embryo development, which is essential for successful embryo implantation.	1)Implantation Failure 2)Endometrial 3) Receptivity 4)PCOS and Insulin Resistance
HOXA10 and HOXA11 (Homeobox A10/A11)	These genes belong to the homeobox family of genes, which encode transcription factors involved in the regulation of embryonic development and cell differentiation. In the context of fertility, HOXA10 plays a critical role in female reproductive health, particularly in the establishment and maintenance of pregnancy	1)Implantation Failure 2)Endometrial development & Receptivity 3)Menstrual disorders
ESR1 (Estrogen Receptor 1)	ESR1 is a key regulator of reproductive function in females, influencing ovarian function, uterine receptivity, cervical function, and the regulation of the menstrual cycle. Its role in fertility underscores the importance of estrogen signaling in female reproductive health	1)PCOS 2)Cervical function 3)Pituitary function 4)Uterine function, proliferation, differentiation, and implantation 5)Ovarian function 3)Menstrual dysfunction

Table 7 outlines a subset of genes of interest implicated in implantation, endometrial receptivity, recurrent pregnancy loss and PCOS for analysis. These genes will be investigated for their South Asian (SAS) population frequencies in the following section to shed light on variances between SAS and the general population frequencies via the gnomAD genomic database.

LIF Interleukin 6 Family Cytokine

Ensembl gene ID: ENSG00000128342.5

Region: Chromosome 22:30240453-30246759

Brief Description

Member of the interleukin 6 family of cytokines known for its pleiotropic effects, possesses the ability to trigger terminal differentiation in leukemic cells. LIF plays a crucial role in blastocyst implantation by preparing the endometrium for embryo attachment, thereby contributing to endometrial receptivity.

Fertility Implications

Implantation failure, recurrent pregnancy loss (RPL), Endometriosis, Preterm Birth, PCOS

Relevant Literature

Women with PCOS showed lower levels of LIF in their serum and follicular fluid compared to controls. However, those who conceived through IVF had higher LIF levels in the embryo culture medium compared to non-conceivers, across both PCOS and control groups. These results suggest that decreased LIF levels in serum and follicular fluid might play a role in the disrupted folliculogenesis observed in PCOS. Additionally, LIF levels in the embryo culture medium could potentially serve as a predictive marker for IVF success (Li, 2018). Other studies also incriminate the negative role of LIF gene mutations on IVF in women diagnosed with unexplained infertility and endometriosis, although additional research needs to be done (Novotný, 2009).

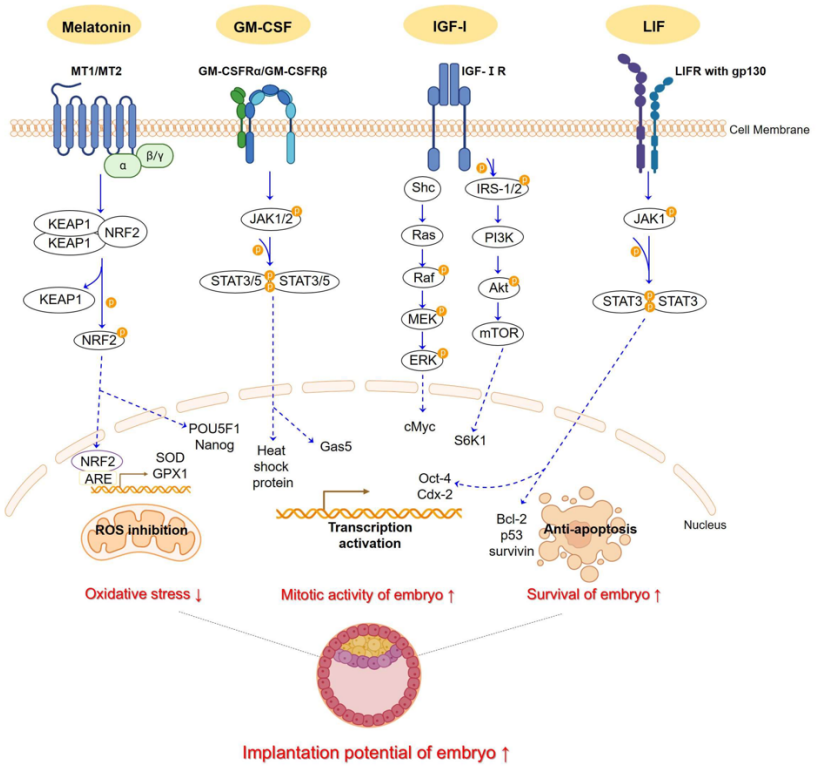


Figure 13. Role of LIF and IGF-1 in Implantation and Embryo Survival.

Schematic shows the pathways associated with LIF and IGF-1 amongst others in embryo culture media, and the signaling that affects implantation of the embryo. LIF binds to its receptor and results in receptor heterodimerization, triggerting the JAK1, 2/STAT3 pathway resulting in the regulation of apoptosis-related gene. It also activated critical transcription genes (Oct-4 AND cDX-2) which reduce oxidative stress and increase mitotic activity and survival of the embryo, improving the chances of a successful implantation. Source: Choi, 2024

Pharmaceutical Target

The drug targeting embryonic implantation in women is currently in phase II clinical trials. This is applicable to women who have failed to become pregnant despite ART intervention (www.uniprot.com).

Analysis of Variation

The LIF gene was filtered for Predicted Loss of Function Variants (pLoF) in all populations and 18 variants uncovered, allele frequencies were calculated and the data for all groups was juxtaposed against each other to shed light on SAS variants with higher frequencies compared to the rest of the population.

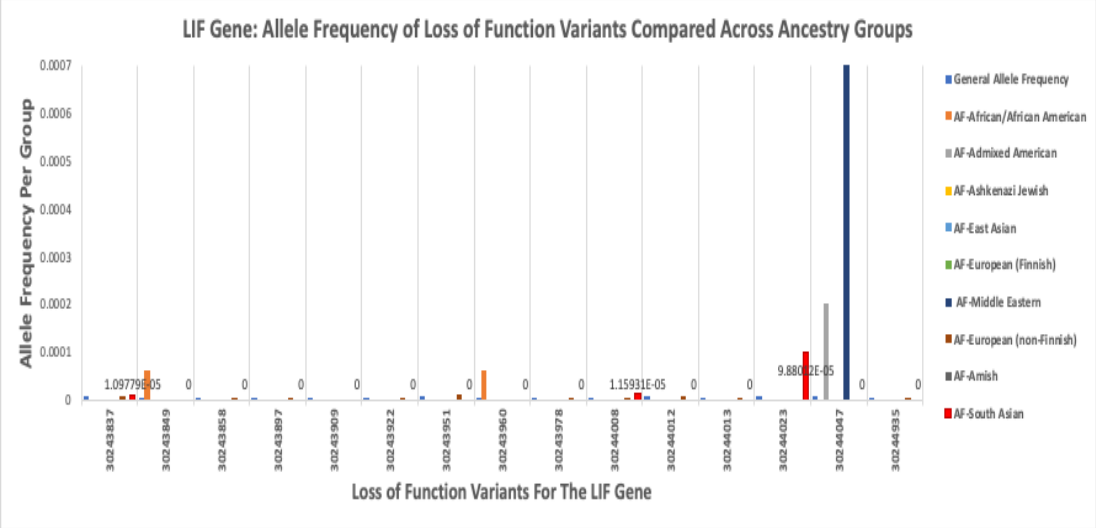


Figure 14. Allele Frequencies of Loss of Function Variants for the LIF Gene Across All Ancestry Groups

An analysis was performed in the gnomAD genomic database to capture allele frequencies for all ancestry groups for the LIF gene and its loss of function variants. Allele frequencies were calculated for all ancestry groups and presented as a graph. Source: gnomAD genomic database

Variants ID: 30243837, 30244008 and 30244023

Three of the Loss of function variants from the analysis namely variants ID: 30243837, 30244008 and 30244023 were observed at higher frequencies in the SAS population. These are presented in the figure below juxtaposed with the average general population frequency for comparison.

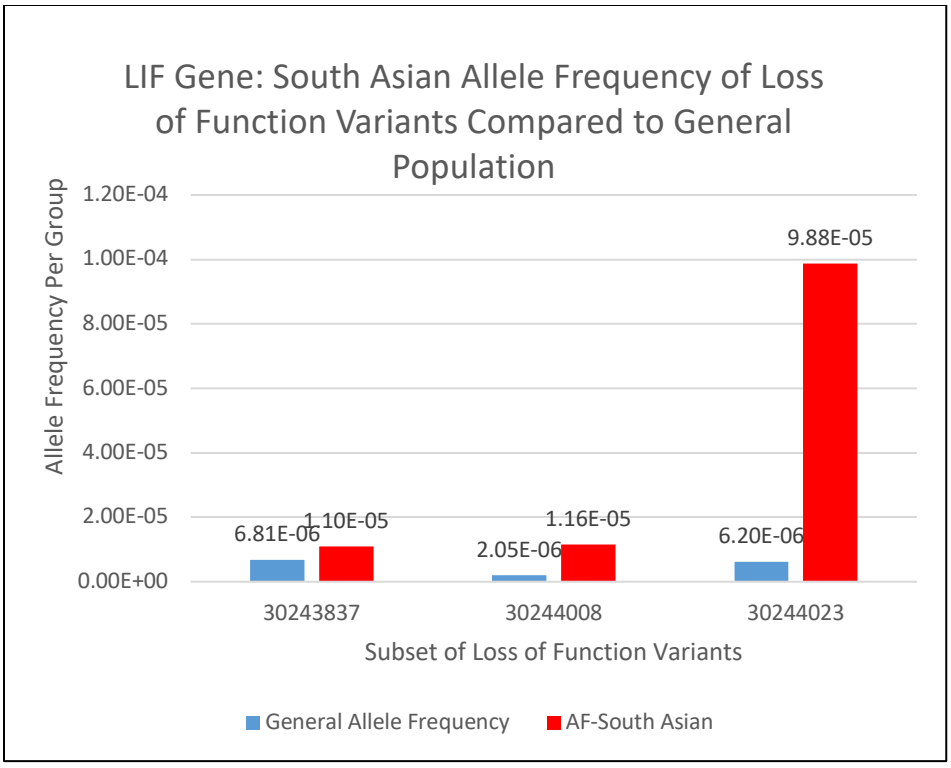


Figure 15. Comparison of Allele Frequencies in SAS LIF Gene Loss of Function Variants

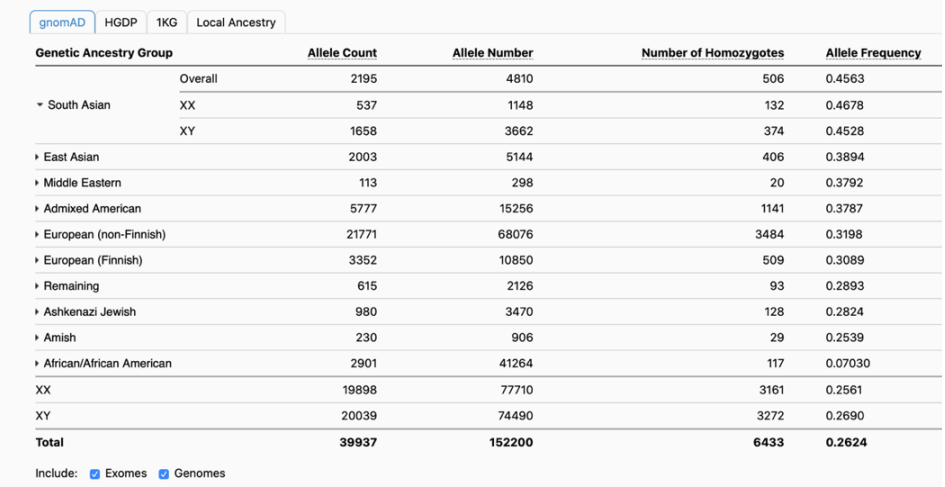
The SAS Loss of Function subset is captured in table 8 below to more clearly depict the allele frequencies compared to the general population frequency.

Table 8. SAS LIF Gene Loss of Function Variant's and their Allele Frequencies

Variant ID	General Allele Frequency	South Asian Allele Frequency
30243837	6.8149E-06	1.09779E-05
30244008	2.0532E-06	1.15931E-05
30244023	6.2027E-06	9.88012E-05

LIF SNP Variant ID: 22-30242237-T-G (rs929271)

A specific gene polymerism found in literature established an association between LIF gene polymorphism and pregnancy outcomes after ART. The LIF single nucleotide polymorphism (SNP) thymine (T)/guanine (G) (rs929271) (Oliviera, 2016) was associated with lower implantation rate (T/G:16.2% Vs G/G: 27%; $p<0.05$), a lower ongoing pregnancy rate/patient (T/G:36.1% Vs G/G:53.7%; $p<0.05$) and a lower ongoing pregnancy rate/transfer (T/G:20.2% Vs G/G: 36.7%; $P<0.05$).



gnomAD | HGDP | 1KG | Local Ancestry

Genetic Ancestry Group		Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
	Overall	2195	4810	506	0.4563
▾ South Asian	XX	537	1148	132	0.4678
	XY	1658	3662	374	0.4528
▸ East Asian		2003	5144	406	0.3894
▸ Middle Eastern		113	298	20	0.3792
▸ Admixed American		5777	15256	1141	0.3787
▸ European (non-Finnish)		21771	68076	3484	0.3198
▸ European (Finnish)		3352	10850	509	0.3089
▸ Remaining		615	2126	93	0.2893
▸ Ashkenazi Jewish		980	3470	128	0.2824
▸ Amish		230	906	29	0.2539
▸ African/African American		2901	41264	117	0.07030
XX		19898	77710	3161	0.2561
XY		20039	74490	3272	0.2690
Total		**39937**	**152200**	**6433**	**0.2624**

Include: ☑ Exomes ☑ Genomes

Figure16. LIF Gene Polymerism Variant ID: 22-30242237-T-G SAS Allele Frequency.

High Allele Frequency in the South Asian (SAS) population for this particular variant as compared to other ancestry groups.This variant is associated with a lower implantation rate, lower ongoing pregnancy rate/pregnancy and lower ongoing pregnancy rate/transfer. Source: gnomAD database variant table

Upon investigation in gnomAD for allele frequencies in various populations it was found that South Asians had the highest allele frequency for this SNP compared to any other ancestry group. This specific genotype was found to be more frequently present in the Recurrent Implantation Failure group of women investigated for polymorphisms related to implantation challenges (Vagnini, 2019).

IGF-1 Insulin Like Growth Factor 1

Ensembl gene ID: ENSG00000017427.17

Region: Chromosome 12:102395874-102481744

Brief Description

This gene encodes a protein that shares functional and structural similarities with insulin, belonging to a protein family crucial for regulating growth and development. IGF-1 is involved in the development and maintenance of endometrial receptivity, vascularization, embryo development, which is essential for successful embryo implantation.

Fertility Implications

Implantation Failure, Endometrial, Receptivity, PCOS and Insulin Resistance

Pharmaceutical Target

Growth Hormones are commonly used as a supplementary agent in assisted reproductive technology (ART) to enhance several clinical outcomes. These include increasing the number of collected oocytes, enhancing oocyte and embryo quality, and improving pregnancy and live birth rates (Regan, 2018).

Relevant Literature

A study conducted in mice demonstrated that the window of uterine receptivity was closed by the estrogen hormone Estradiol (E2), which was mediated by the IGF1 pathway (Kobayashi, 2017). In another mouse study the presence of 10 ng/mL of IGF-I

notably increased the blastocyst development rate (Green, 2013). These studies may have implications for humans as well and the role of IGF-1 in IVF culture media. The role of IGF-1 Follicular fluid insulin-like growth factor-1 (FF IGF-1) serves as a biochemical indicator of embryo quality and implantation rates in in vitro fertilization cycles (Chimote, 2013). The Insulin Growth Factors are closely related to growth hormones and are found in various reproductive tissues. They are involved in a variety of roles (endocrine, paracrine, and autocrine factors) that include promotion of cell growth, proliferation, differentiation, and anti-apoptosis. They are also involved in pathways which may have a role to play the molecular explanation of pregnancy loss and are involved in cell division and hormone production. Some examples of pathways include such as the mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinase (ERK), Jak/STAT, and phosphoinositide 3-kinase (PI3K)/protein kinase B (Akt) (Bertrand, 2006).

Variant Analysis

The IGF-1 gene was filtered for Predicted Loss of Function Variants (pLoF) across all ancestry group populations and 20 variants uncovered. Allele frequencies were calculated and the data for all groups was juxtaposed against each other to shed light on SAS variants with higher frequencies compared to the rest of the population.

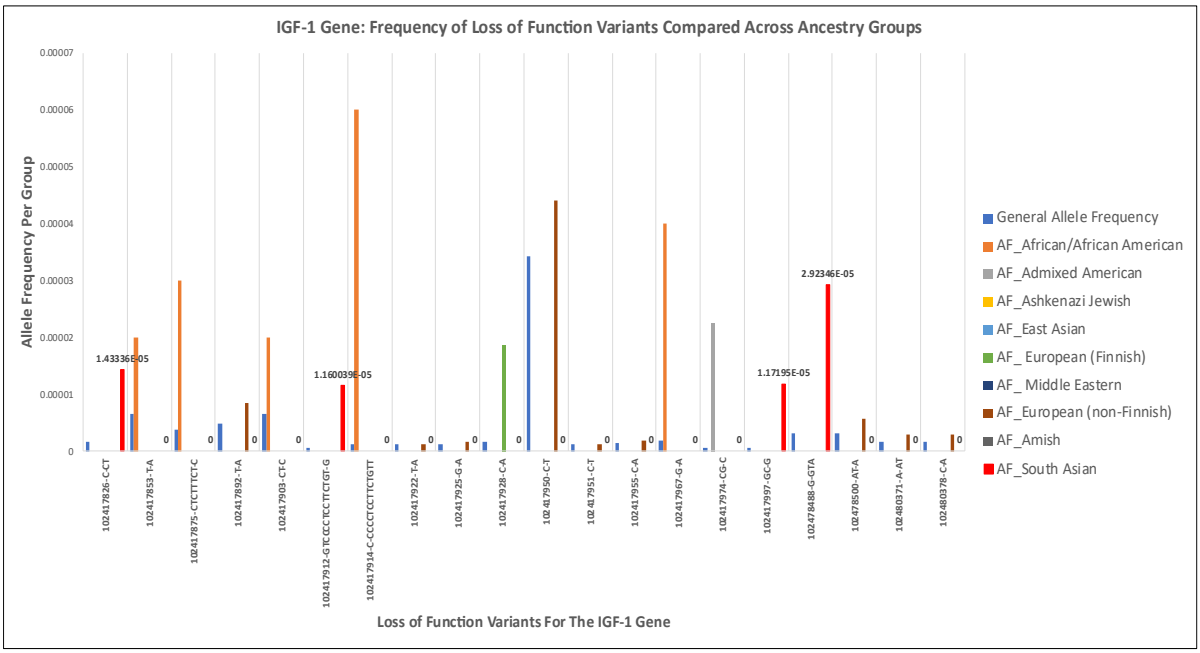


Figure 17. IGF-1 Gene Loss of Function Variants and their Allele Frequencies across all Ancestry Groups.

An analysis was performed in the gnomAD genomic database to capture allele frequencies for all ancestry groups for the IGF-1gene and its loss of function variants. Allele frequencies were calculated for all ancestry groups and presented as a graph. Source: gnomAD genomic database.

Variants IDS: 102417826, 102417912, 102417997, 102478488

4 key pLoF variants were found at a comparatively higher allelic frequency for the South Asian group which resulted in loss of function of the gene as compared to other population groups. pLoF variants ID: 102417826-C-CT, 102417912-GTCCCCTCCTTCTGT-G, 102417997-GC-G, 102478488-G-GTA were observed at higher frequencies in the South Asian population are presented in the table below.

Table 9: Allele Frequency Table for IGF-1 Gene Loss of Function South Asian Variants Compared to the General Population

Variant ID	General Allele Frequency	South Asian Allele Frequency
102417826	1.59153E-06	1.43336E-05
102417912	6.84E-07	1.160039E-05
102417997	6.86E-07	1.17195E-05
102478488	3.20287E-06	2.92346E-05

IGF-1 gene loss of function South Asian (SAS) variants were calculated and their allele frequencies were compared to the general population's allele frequencies. These 4 variants were found to be higher in SAS group. Source: calculated from the gnomAD database.

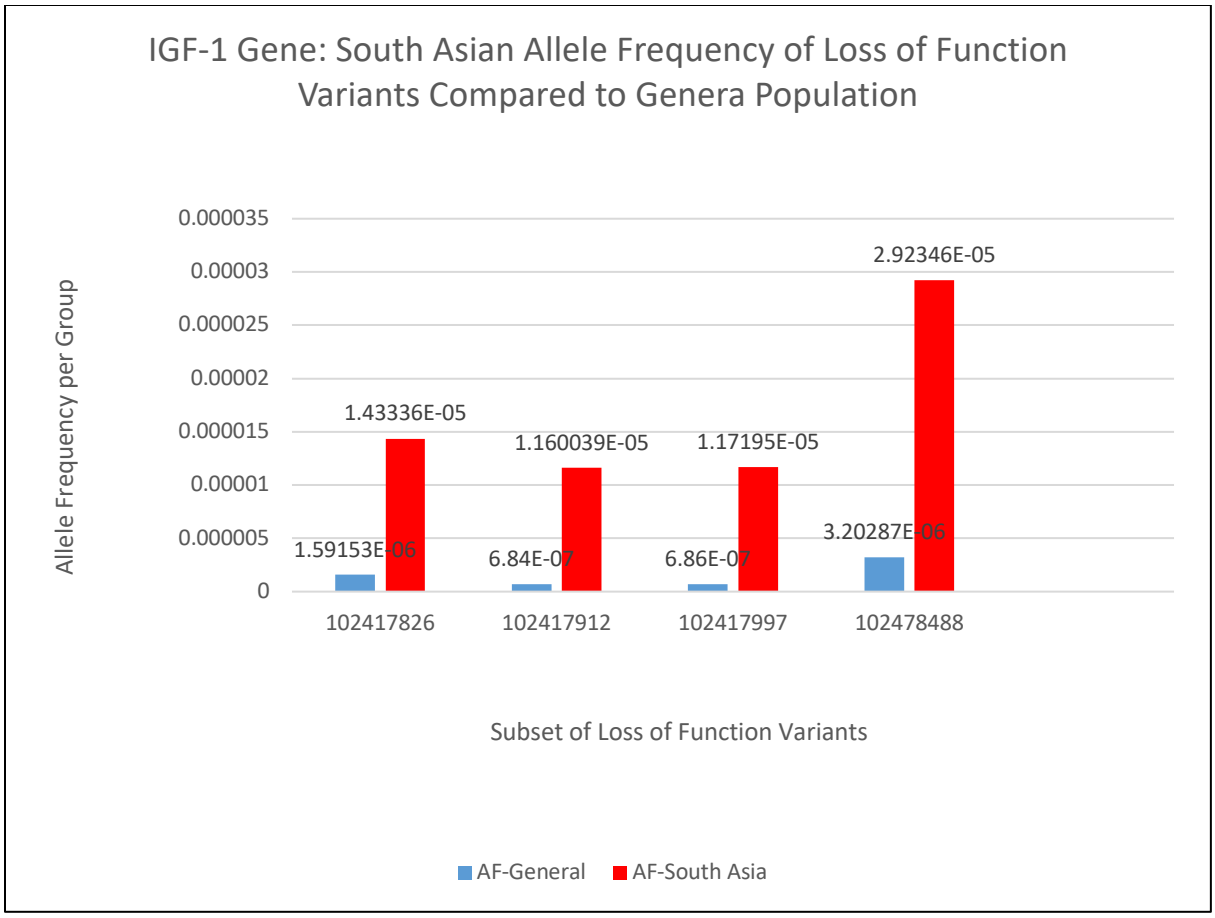


Figure 17. IGF-1 Gene Allele Frequencies of South Asian Loss of Function Variants Compared to the General Population Allele Frequency.

The figure shows 4 variants 102417826, 102417912, 102417997, 102478488 with elevated levels found in the South Asian ancestry group as compared to the general population

HOXA10 Homeobox A10

Ensembl gene ID: ENSG00000253293.6

Region: Chromosome 7:27170605-27180261

Brief Description

These genes belong to the homeobox family of genes, which encode transcription factors involved in the regulation of embryonic development and cell differentiation. In the context of fertility, HOXA10 plays a critical role in female reproductive health, particularly in the establishment and maintenance of pregnancy.

Fertility Implications

Implantation Failure, Endometrial development & Receptivity, Menstrual disorders

Pharmaceutical Target

None identified in the context of infertility.

Relevant Literature

HOXA10 expression in the endometrium must rise to facilitate embryo implantation. Yet, endometrial biopsies from women with PCOS in ovulatory cycles indicate reduced HOXA10 expression during the secretory phase compared to normal fertile women. In vitro studies suggest that testosterone suppresses HOXA10 expression (Cermik, 2003)

β3-integrin, a marker of endometrial receptivity to embryo implantation, is a target gene of HOXA10 directly controlled by HOXA10 in endometrial cells. The expression of this marker is lower in the endometrium of women with PCOS compared to fertile controls (Apparo 2002). Decreased expression of this gene results in impaired implantation and uterine developmental abnormalities.

Variant Analysis

Variant ID: SNV:7-27173892-G-A

Pathogenic variants were screened on gnomAD and one variant SNV:7-27173892-G-A had a relatively higher allele frequency compared to the general population allele frequency (South Asian: 0.00001271 Vs General Population 0.000002873). This variant falls on 4 transcripts in 3 genes with a VEP annotation of Stop Gained. No publications were reported.

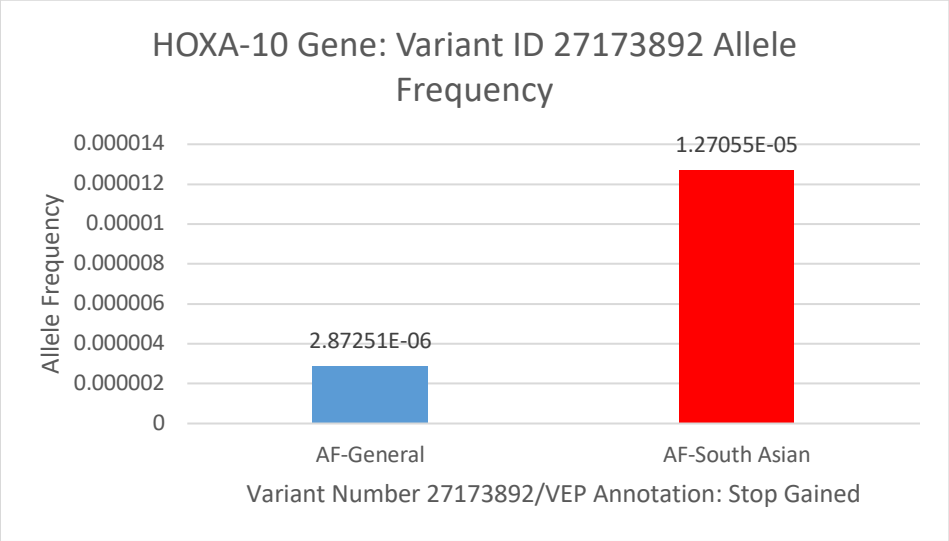


Figure 18. HOXA-10 Gene South Asian Pathogenic Variant of Interest Allele Frequency Compared to the General Population Allele Frequency

HOXA11 Homeobox A11

Ensembl gene ID: ENSG00000005073.6

Region: Chromosome 7:27181157-27185232

The HOXA11 has similar functions to the HOXA10 gene and belong to the same family of genes that play a part in endometrial receptivity and implantation. HOXA11 was also analyzed for pathogenic variants and two variants were discovered on gnomAD that had elevated allele frequencies as shown in the figure below namely SNV ID: 7-271828842-AT-A (Frameshift) and 7-27185155-C-T (Splice donor). No publications were reported.

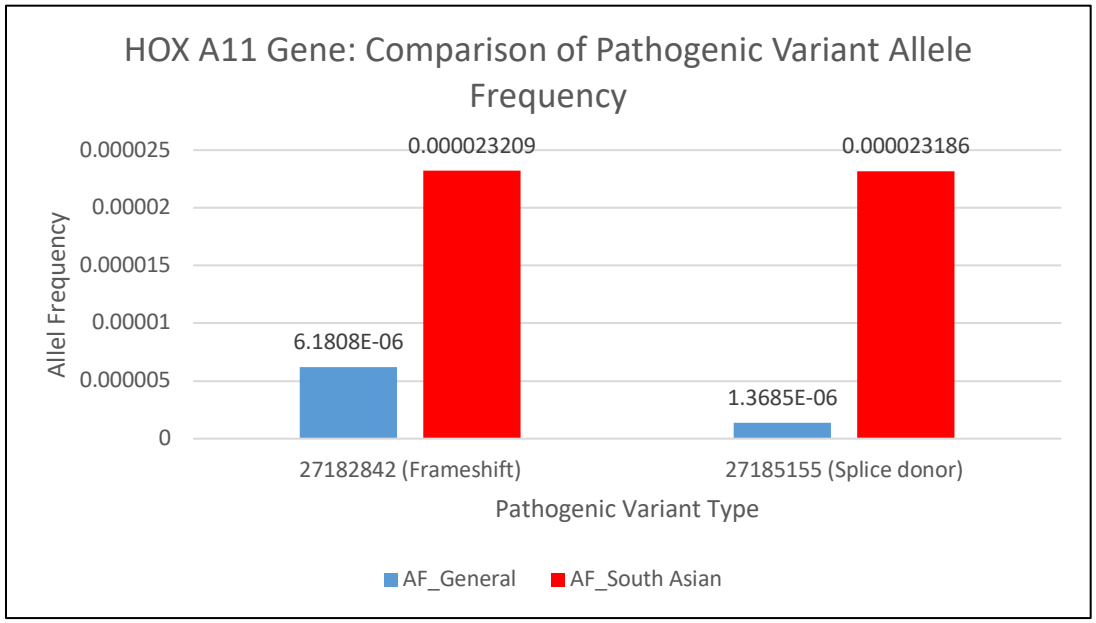


Figure 19. HOXA-11 Gene South Asian Pathogenic Variants of Interest Allele Frequency Compared to the General Population Allele Frequency

ESR1 Estrogen Receptor 1

Ensembl gene ID: ENSG00000091831.25

Region: Chromosome: 6:151656691-152129619

Brief Description

ESR1 is a key regulator of reproductive function in females, influencing ovarian function, uterine receptivity, cervical function, and the regulation of the menstrual cycle. Its role in fertility underscores the importance of estrogen signaling in female reproductive health.

Fertility Implications

PCOS, Cervical function, Pituitary function, Uterine function, proliferation, differentiation

Pharmaceutical Target

Anti-estrogens, selective estrogen receptor modulators (SERMs) and selective estrogen receptor degraders (SERDs) in cancer treatment

Relevant Literature

Associations between ESR1 and LIF polymorphisms are predicting factors for recurrent implantation failure (RIF) and the presence of ESR1/AA (rs12199722) and LIF/GT (rs929271) genotypes was more frequent in the RIF group were found more

frequently in the RIF group which increases presentation of RIF by 7.9% after IVF (Vangini, 2018).

Other variants such as the ESR1 rs9340799 (Paskulin, 2013) have been associated with endometriosis-related infertility and IVF failure.

Mutations in the estrogen receptor genes *ESR1 and ESR2*, involved in normal follicular development and ovulation, can contribute to development of the PCOS. Strong associations between ten SNPs present in *ESR1* and *ESR2* genes with PCOS has been shown in a study affecting South Asian Pakistanis (Muccee, 2024).

Analysis

The presence of ESR1/AA (rs12199722) and LIF/GT (rs929271 Variant ID: 22-30242237-T-G) genotypes was more frequent in the RIF group, leading to a 7.9-fold increase in the chance of women presenting with RIF when compared with women who became pregnant on their first cycle of IVF/intracytoplasmic sperm injection and a 2.8-fold increase when compared with women who became pregnant without treatment. The LIF variant has been demonstrated earlier (LIF Section) to be present at the highest Allele frequency in South Asians compared to the general population frequency and the ESR1 variant is demonstrated below.

Variant ID: SNV 6-151880764-GA-G

This multiallelic variant has alternate alleles. This variant is caused by a deletion (1 base) that results in a frameshift. The HGVS Consequence is p.Gly253ValfsTer13

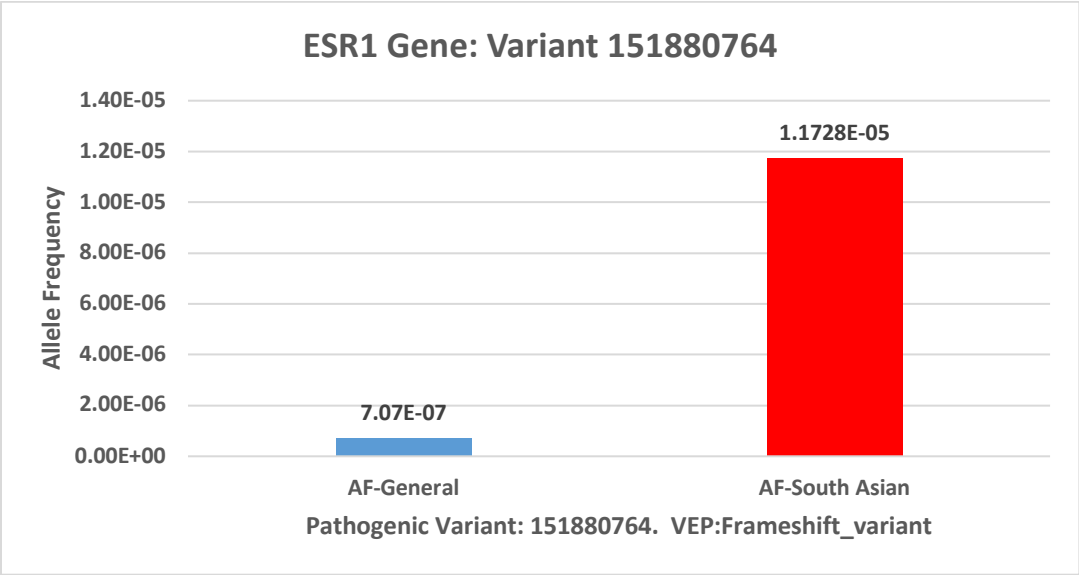
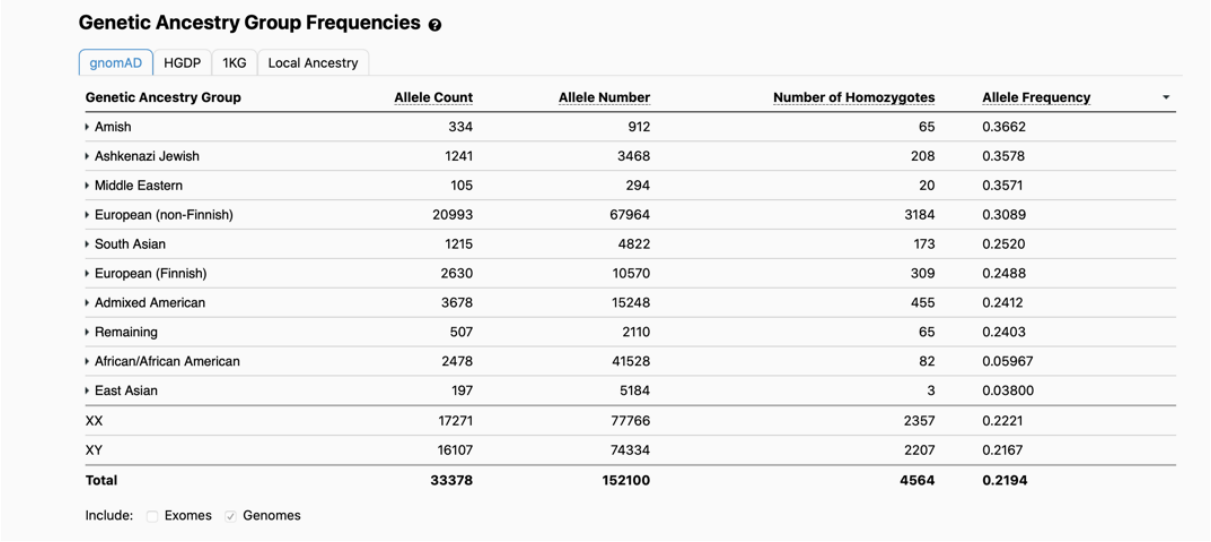


Figure 20. ESR1 Pathogenic Variant 151880764, South Asian Allele Frequency Compared to the General Population Frequency

This variant is present at a much higher allele frequency in South Asians as compared to the general Allele Frequency and is implicated in recurrent implantation failure.

Variant ID: SNV 6-151913765-A-G, rs12199722.

This variant is implicated in Recurrent Implantation Failure in concert with the LIF variant ID: SNV 22-30242237-T-G rs929271 which has a high allele frequency in South Asians (SAS: 0.4563 GF:0.2624), the literature reference is demonstrated in the LIF section. The ESR1 Variant ID: SNV 6-151913765-A-G is slightly elevated in South Asians when compared to the general population frequency.



Genetic Ancestry Group Frequencies

gnomAD | HGDP | 1KG | Local Ancestry

Genetic Ancestry Group	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
Amish	334	912	65	0.3662
Ashkenazi Jewish	1241	3468	208	0.3578
Middle Eastern	105	294	20	0.3571
European (non-Finnish)	20993	67964	3184	0.3089
South Asian	1215	4822	173	0.2520
European (Finnish)	2630	10570	309	0.2488
Admixed American	3678	15248	455	0.2412
Remaining	507	2110	65	0.2403
African/African American	2478	41528	82	0.05967
East Asian	197	5184	3	0.03800
XX	17271	77766	2357	0.2221
XY	16107	74334	2207	0.2167
Total	**33378**	**152100**	**4564**	**0.2194**

Include: Exomes Genomes

Figure 21. ESR1 Variant rs12199722. SNV: 6-151913765-A-G. South Asian Allele Frequency compared to other Populations.

This Variant is implicated in Recurrent Implantation Failure and is present at slightly elevated levels in the South Asian population as compared to general population frequency

Variant ID: SNV:6-151842246-A-G(GRCh38) rs9340799

The ESR1 variant SNV:6-151842246-A-G shown in the figure below has been implicated in literature to be implicated in endometriosis related infertility in the South Asian Pakistani population. The figure below shows the allele frequency breakdown by ancestry group (gnomAD database).

gnomAD | HGDP | 1KG | Local Ancestry

Genetic Ancestry Group	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
Middle Eastern	120	294	24	0.4082
Ashkenazi Jewish	1339	3466	259	0.3863
South Asian	1773	4822	343	0.3677
European (non-Finnish)	23845	67946	4185	0.3509
Remaining	683	2112	107	0.3234
Amish	290	908	45	0.3194
Admixed American	4388	15278	620	0.2872
African/African American	11864	41440	1716	0.2863
European (Finnish)	2462	10554	275	0.2333
East Asian	1037	5156	124	0.2011
XX	24905	77688	4086	0.3206
XY	22896	74288	3612	0.3082
Total	**47801**	**151976**	**7698**	**0.3145**

Figure 22. Allele Frequencies for ESR1 Variant ID: SNV:6-151842246-A-G(GRCh38) rs9340799

This variant has been shown to be elevated in one of the major South Asian sub-groups from the country Pakistan. Patients with this variant present endometriosis related infertility.

Chapter Conclusion

5 genes of interest (LIF, IGF1, HOXA10/11, ESR1) implicated in implantation failure, PCOS, Recurrent pregnancy loss and endometrial receptivity were analyzed through the genome aggregator database gnomAD to determine allele frequencies in South Asian ancestry (SAS) versus other groups.

LIF gene: 3 predicted Loss of function variants (pLOF)were observed at higher allele frequencies (AF) in the (SAS)group compared to general population (GP) allele frequencies (AF) (Variant IDs: 30243837, 30244008 and 30244023).

LIF SNP Variant ID: 22-30242237-T-G associated with lower implantation rate, pregnancy rate, and ongoing pregnancy rate was found to be the highest in South Asians compared to all other ancestry groups.

IGF-1 gene: 4 key pLoF variants (102417826, 102417912, 102417997, 102478488) were found at a comparatively higher AF for the SAS group which resulted in loss of function of the gene as compared to GP.

HOXA10 gene: Pathogenic Variant ID:27173892 had a relatively higher SAS-AF compared to the GP-AF.

HOXA11gene: Two Pathogenic Variant ID:7-271828842 and 7-27185155 had elevated allele frequencies in SAS compared to GP.

ESR1 gene: Pathogenic Variant ID:151880764 implicated in Recurrent Implantation Failure (RIF) was present at a significantly higher frequency in SAS groups compared to GP.

Variant ID:151913765 implicated in RIF was slightly elevated in SAS compared to GP. Variant ID:151842246 implicated in endometriosis related to infertility in South Asian Pakistani populations was at a higher frequency in the SAS compared to the GP.

Table 10 Genes of Interest: Role in Associated Reproductive Issues and identified SAS Allele Frequencies

Genes of Interest	Brief Description Relevant to Fertility	Associated Reproductive Issues	Variability Results for the SAS Ancestry Group
LIF (Leukemia Inhibitory Factor)	Member of the interleukin 6 family of cytokines known for its pleiotropic effects, possesses the ability to trigger terminal differentiation in leukemic cells. LIF plays a crucial role in blastocyst implantation by preparing the endometrium for embryo attachment, thereby contributing to endometrial receptivity.	1)Implantation Failure 2)Recurrent Pregnancy Loss 3)Endometriosis 4)Preterm Birth 5)PCOS	3 pLof Variants with Higher AF compared to the GP 1 variant implicated in RPL found to be highest in South Asians
IGF1 (Insulin-like Growth Factor 1):	This gene encodes a protein that shares functional and structural similarities with insulin, crucial for regulating growth and development. IGF-1 is involved in the development and maintenance of endometrial receptivity, vascularization, embryo development, essential for successful embryo implantation.	1)Implantation Failure 2)Endometrial 3) Receptivity 4)PCOS and Insulin Resistance	4 pLof Variants with Higher AF compared to the GP
HOXA10 and HOXA11	The homeobox family of genes, encode transcription factors involved in the regulation of embryonic development and cell differentiation. HOXA10 plays a critical role in female reproductive health, particularly in the establishment and maintenance of pregnancy	1)Implantation Failure 2)Endometrial development & Receptivity 3)Menstrual disorders	HOXA10: 1 Pathogenic Variant at a higher AF HOXA11: 2 pathogenic variants found at higher AF
ESR1 (Estrogen Receptor 1)	ESR1 is a key regulator of reproductive function in females, influencing ovarian function, uterine receptivity, cervical function, and the regulation of the menstrual cycle. Its role in fertility underscores the importance of estrogen signaling in female reproductive health	1)PCOS 2)Cervical function 3)Pituitary function 4)Uterine function, proliferation, differentiation, and implantation 5)Ovarian/Menstrual dysfunction	1 pathogenic variant implicated in RIF found with High AF 2 variants implicated in RIF and endometriosis related infertility at high AF

Table 9 outlines a subset of genes of interest implicated in implantation, endometrial receptivity, recurrent pregnancy loss and PCOS for analysis. These genes were investigated for their South Asian (SAS) population frequencies and compared to the general population allele frequencies to highlight variances in the SAS group. Results are documented in the last column

Based on the genomic analysis of variant allele frequencies in the South Asian ancestry group, certain variants were found to exist at higher allele frequencies in the South Asian population which are attributed to dysfunctions associated to infertility. Only variants that have been designated as predicted Loss of function, pathogenic variants or identified in previous literature were analyzed in gnomAD. This sheds light on the massive effort required to analyze rare variants and those that have not been categorized and bucketized yet. Investigating other categories of variants will further highlight variation in the exome and genome of SAS and shed light on therapeutic strategies, predictors of disease onset and provide insights into custom pharmacogenomic end points which will allow physicians to make informed decisions in the treatment of infertility.

Chapter V

Discussion & Recommendations for a Customized Approach to ART to bolster South Asian Success Outcomes

The literature review and genomic analysis conducted in this thesis points to the possibility that genomic diagnosis/customization is critical not after repeated failure but strategically at the onset of fertility treatments. The case for the South Asian unique ethnic identity has now been well established and the jigsaw pieces of South Asian infertility landscape are slowly coming together. A holistic approach to infertility is mandated when global populations are in consideration and additional targeted research is required to paint an accurate view into the unique markers of disease onset, inheritance, etiology, manifestation, and prognosis. Prevalence of disease burdens, lifestyle, nutrition, health care access, migration patterns must be factored into the formula of treatment. Global trends such as our newly adopted non-sedentary lifestyles are discordant with centuries of human evolution. It is important to recognize that these environmental factors have impacts on gene mutations, metabolic processing, and biological compensatory mechanisms that can only be superficially addressed by standard pharmacological intervention. A customized approach is mandated to consider all factors, genetic and environmental, that result in disease manifestation.

Recommendations for Customization

Empowering Families, Delivering Life: Medicine's Gift of Parenthood

The elusive applicability of standard IVF protocols as a paradigm across ethnicities needs to be analyzed and customized. The commoditization and commercialization of private fertility centers and private equity investments casts a transactional shadow to the business of fertility, detached from its fundamental goal of providing high-quality, patient-centered, compassionate care. Recognition of patient individuality maximizes the chances of a successful pregnancy, streamlining costs and minimizing adverse effects.

In the preceding sections, this thesis has aimed to unveil a compelling narrative woven from the threads of existing literature, genomic database analysis, meticulous scrutiny of IVF protocols, the glaring absence of representation in clinical trials, the paucity of comprehensive genomic data, and the sobering prevalence of disease markers among South Asians. Yet, amidst this tapestry, looms my own visceral odyssey spanning fifteen years as a patient. Through the crucible of personal experience, one truth emerges with undeniable clarity: the customization of IVF protocols isn't just advisable—it's imperative. For within this customization lies the beacon of hope, guiding families through the labyrinth of unfruitful cycles that threaten to shatter not only their finances but also their very experience of life.

Evolving the Holistic Approach

In this epoch of pharmacogenomics, the contemporary medical practitioner finds themselves armed with the indispensable knowledge necessary to refine the prevailing paradigm of diagnosing and treating infertility. In the pursuit of elevating the current success rates, which currently stand at 24%, it behooves us to engage in a scholarly discourse concerning several overarching considerations. These high-level factors serve as the cornerstone for crafting comprehensive strategies aimed at addressing infertility in its entirety.

1. Evidence-Based Treatment: Effective and evidence-based treatments can be incorporated that are tailored to individual patient needs. This may include personalized treatment plans based on factors such as age, PCOS, ovarian reserve, and previous treatment history.
2. Patient Education: Thorough and clear education needs to be presented to patients regarding the IVF process, including potential risks, success rates, and alternative treatment options. Patients should be empowered to make informed decisions about their care based on their unique situation.
3. Emotional Support: IVF treatment can be emotionally challenging, options for compassionate and supportive care to patients throughout the process may impact treatment positively and improve the patient experience. This may include counseling services and support groups.

4. Ethical and Transparent Practices: Adherence to ethical guidelines and ensuring transparency in the IVF lifecycle, clinic success rates, adjunct treatments as well as cost fluctuations.
5. Continued Research and Innovation: Engagement in research and innovation to improve treatment outcomes and advance the field of reproductive medicine.
6. Collaboration and Referral: Collaboration with other healthcare providers, such as reproductive endocrinologists, urologists, and mental health professionals, to provide comprehensive care. They should also be willing to refer patients to specialists when needed to alleviate the root cause of infertility.
7. Regulations and Compliance: Compliance and enhanced reporting of treatment plans, success rates, adverse events and adjunct treatments being offered to desperate patients are currently underreported.

Overall, the goal of IVF centers should be to provide comprehensive, customized, compassionate, and evidence-based care to help patients achieve their goal of building a family.

Focus Areas of Cycle Customization-Positively Racializing Infertility

The recommendations are tailored specifically for patients of South Asian descent seeking infertility services and undergoing IVF protocols. This focus is of paramount importance, particularly in light of previous evidence highlighting comparatively lower levels of IVF success within this ancestral group.

The accompanying figure illustrates crucial areas for potential customization (flagged) that can be tailored to the distinctive ancestry, prevailing disease profiles, lifestyle considerations, and various socio-economic factors that may significantly impact the patient journey.

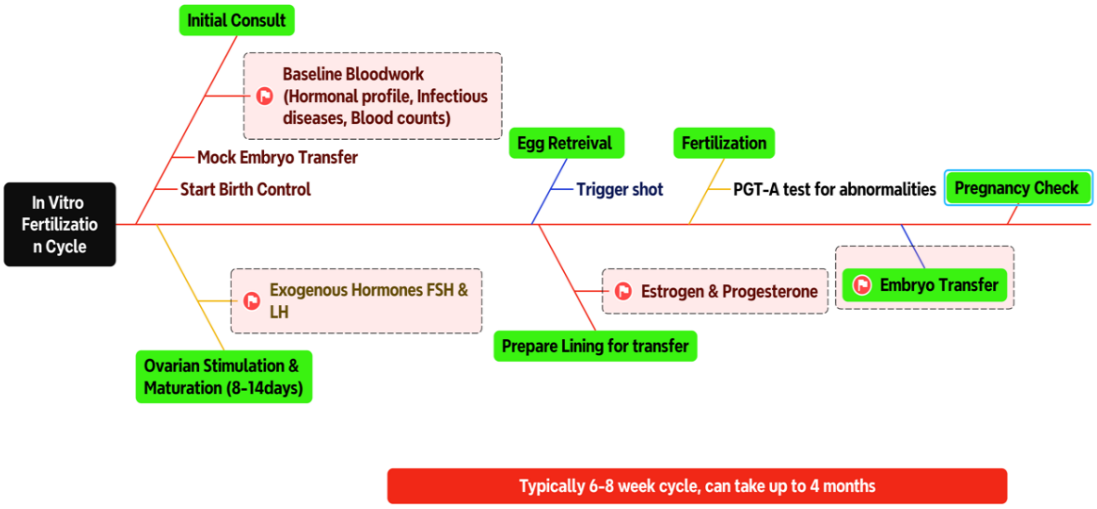


Figure 23. Potential Areas of Assessment and Modulation for Customization to South Asian IVF Treatment Offerings.

The flags and boundary areas highlighted in pink are key areas of personalization based on the unique health profile of the South Asian Patient. These customizations are two-fold: At the South Asian level and at the individual patient level, factoring in ancestry and current health profile

Front line diagnostics and assessment of predisposed disease risks

At the time of initial consult, age, BMI, smoking status and PCOS status of patients is documented. These factors are not mapped to ethnicity or factored into its impact on the stimulation protocol Gonadotropin dosage. Additionally, the implantation failure rate and repeated pregnancy loss associated with ethnic backgrounds is not factored into the IVF protocol strategy. Patients are also not advised on lifestyle factors such as nutrition or exercise, which as previously discussed have a positive impact on fertility outcomes. The reason for this lack of holistic approach can be attributed to the fact that fertility clinics view the IVF protocol as infertility treatment which supplants existing issues. Supplementation recommendations for vitamin D and thyroid levels are routinely assessed and adjusted, however, correcting the underlying root cause is not an area of focus for IVF centers. Multiple centers have a "We don't care about that issue", "IVF takes care of it" approach when accosted with questions on how to improve overall fertility health.

Considering the static nature of ART success rates at 24%, and the statistics pertaining to diagnostic categories such as "Unidentified Infertility" (11%) and "Recurrent pregnancy loss" (6%), "Uterine factor" (6%) and "Other factor" (30%), conclusive diagnosis upon repeat IVF cycle failure remains elusive. Fertility experts recommend repeating IVF protocols to improve chances with little informed modification to standard protocols. As discussed earlier, infertility is exacerbated by PCOS severity and genomic markers of infertility and other lifestyle factors all contribute to an individual's fertility potential. Patients will greatly benefit from a diagnostic and research-based approach that factors in the various health factors implicated in infertility.

Individualized Gonadotropin dosing algorithm:

Currently, BMI adjustments are not always factored into the dosage and cycle length of hormone stimulation drugs to prevent the possibility of Ovarian Hyperstimulation Syndrome. South Asian patients stand at an increased risk as shown previously. Controlled ovarian stimulation using gonadotropin releasing hormones directly impacts the number of harvested oocytes. Age, BMI, Anti-Mullerian Hormone (AMH), initial serum follicle stimulating hormone, ovarian stimulation days are all variables that can impact the oocyte sensitivity index as shown by a stimulation study conducted on Japanese patients (Kobanawa, 2023). No such calculation is currently in place for IVF patients. With the advent of artificial intelligence and machine learning research facilities and IVF centers should collaborate on creating a formula that can predict the ideal dosage and stimulation days with the associated variables to simulate the most precise drug exposure to patients and factor in genomic factors for an algorithm that can be applicable to patients of all background.

National standardization will allow physicians and to determine the exact protocol and prevent inefficiency, poor oocyte numbers and take away the risk of OHSS altogether. As a bare minimum Retrospective analysis of prior medical records should be taken seriously for patients who have had multiple unsuccessful IVFs and OHSS and the dosage should be metered which is not a usual practice for most infertility clinics. (Shah, 2016) demonstrated higher success outcomes (live births) In South Asians with FET cycles versus Fresh cycles. The live birth rate was significantly lower in South Asians at 21% as compared to 37% in Caucasians in fresh cycles. These could potentially point to the interaction of stimulation protocols and South Asian endometrial variances that

respond to hormonal stimulation, which further fortifies the need for an individualized Gonadotropin dosing algorithm for patients of South Asian descent.

Lifestyle, Nutrition, Vitamin D deficiency, thyroid & PCOS Severity:

PCOS severity in patients should be assessed and ethnicity should be factored into the patient profile. Patients should be made aware of potential impacts on the success of the cycle to better prepare them on potential outcomes. Lifestyle modifications and dietary changes should be discussed to prevent a worsening of hormonal profiles and health outcomes during treatment, which may impact IVF success. Vitamin D has been shown to affect oocyte quality and is implicated in PCOS and fertility outcomes, Vitamin D deficiency is rampant in the South Asian population as shown previously. Greater monitoring is required for South Asian patients and nutrient deficiencies should be resolved prior to IVF start and taken as a priority for monitoring throughout the pregnancy to improve outcomes.

Psychological Assessments & Continued Monitoring:

Assessments on the infertility history, cultural challenges, and patient awareness should be assessed at the beginning of treatment. An emotional support plan should be created for patients demonstrating need based on their abilities to comprehend and tolerate stressors during the treatment. South Asian patients face significant cultural impacts of infertility which include isolation, stigmatization, and discrimination which compounds emotional pain. These patients may require additional attention to ensure a positive mindset which can impact the outcome of these treatments. Socioeconomic factors also impact fertility that may not be a point of consideration western environments. South Asian women may have socio economic challenges, reduced access

to resources and services, and may require additional education, resources, and services to establish patient health equity.

Personalization of Window of Implantation:

The precise determination of the window of implantation has been shown to significantly improve ART outcomes by improving pregnancy rates, progression of pregnancy, and reducing pregnancy loss. 34.18% of patients assessed had a displaced WOI (Enciso, 2021). Endometrial Receptivity defines the narrow 48hr window of implantation (WOI) where the endometrium undergoes gene expression changes and is receptive to competent embryos. Deviation of just 12hours showed considerable decline in pregnancy rates. The figure below is a clinical sample of a South Asian patient with a displaced WOI (12 hours).

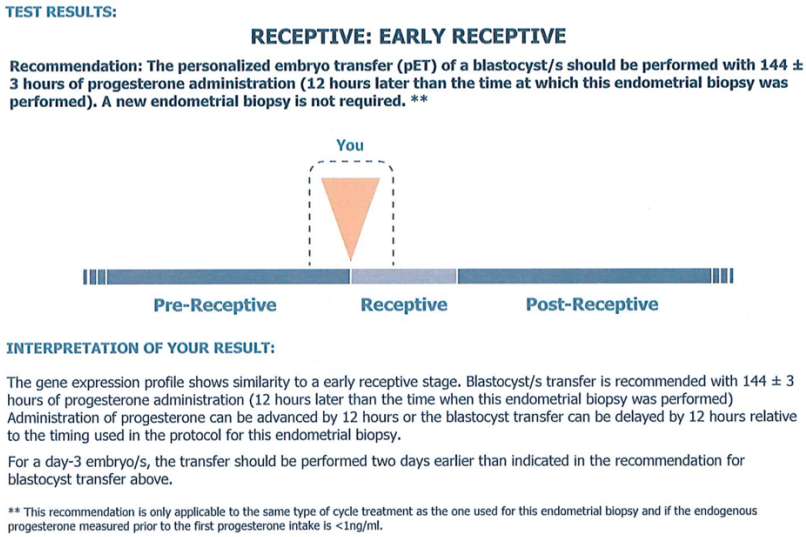


Figure 24 Window of Implantation Test Results from a South Asian Infertility Patient.

The ERA transcriptomic test identifies the precise window of implantation when the patient's endometrium is receptive to the embryo. Deviation of 12 hours negatively impacts pregnancy rates and pregnancy progression in ART outcomes based on

literature. This patient's gene expression profile aligns with an early receptive stage and would need to be adjusted by 12 hours (Either progesterone administration or blastocyst transfer timing). Distinction would also have to be made for Day 3 embryos versus blastocyst transfer timings.

Had the above patient followed a standard blastocyst transfer, this precise WOI would have been missed. This is especially relevant when treating South Asian patients. As an example the LIF gene polymerism (Variant ID: 22-30242237-T-G) is associated with lower implantation rates and has been found to exist at higher allele frequencies in South Asians as demonstrated in the LIF analysis section and as corroborated in literature (Oliveira, 2016) (Vagnini, 2019)

Existing embryo transfer strategies at most infertility centers, does not take into account the precise WOI and most embryo transfers are performed 5 days after progesterone administration. This is an opportunity for personalization of endometrial preparation with demonstrated results, that has the potential to significantly improve ART outcomes at the individual level.

The identification of possible displacements of the WOI prior to the start of an IVF cycle can potentially save patients thousands of dollars, several months and several unsuccessful IVF marathons, maximizing their chances of a successful pregnancy, preventing wastage of invaluable embryos.

Fresh Vs Frozen Embryo Transfer:

As demonstrated earlier there are racial disparities in fresh versus frozen embryo transfer success rates in South Asians compared to Caucasian women. The live birth rate was significantly lower in South Asians at 21% as compared to 37% in Caucasians (Shah, 2016) This should be discussed and built into the protocol and timing of Embryonic

Transfers upfront to improve live pregnancy outcomes. Currently, IVF clinics do not present this as a recommendation to South Asians and this results in repeated cycles that result in financial burden, loss of time, and emotional challenges for the patient. These factors present a major gap in understanding the success factors for South Asian patients.

Concluding Remarks & Future Perspectives

Infertility is a complex and multifactorial phenomenon that impacts every aspect of life for those who experience it. It is mediated by several genes and complex pathways and impacted by underlying disease, lifestyle factors, urbanization, and socio-economic profiles.

In this landscape of adversity, pharmacogenomics and precision medicine emerge as a beacon of hope. It promises tailor-made solutions, sculpted meticulously to fit the unique contours of individual genetic blueprints. While its triumphs echo in the annals of life-threatening illnesses such as cancer, HIV, cardiovascular disorders and diabetes—its strides in the realm of infertility treatment lag behind. Polymorphisms in several genes are associated with infertility related disorders and there is a massive scope of pharmacogenomics and personalized medicine for treatment of infertility owing to the increasing body of research on genes and associated SNPs.

It is imperative to view the challenge of infertility from a holistic lens as well as from a population focused one, to address its complexity and specificity. The one size fits all standard approach is riddled with ethnicity associated disparities that results in poor outcomes for certain groups. As genome wide association studies become more inclusive, key insights can be derived for specific populations that are clustered by variations to genes that govern infertility. This coupled with regional health profiles, identification of

rare variants, drug metabolism profiles, and unique socio-economic profiles can provide a complete picture of targeted therapy that can produce positive outcomes in ART for specific ethnicity groups.

The schematic below is built on genomic analysis, South Asian health factors, disease burdens, and unique IVF challenges that were investigated in this thesis. These inputs were assessed to recommend potential improvements for the creation of custom IVF protocols that address the unique landscape of South Asian infertility.

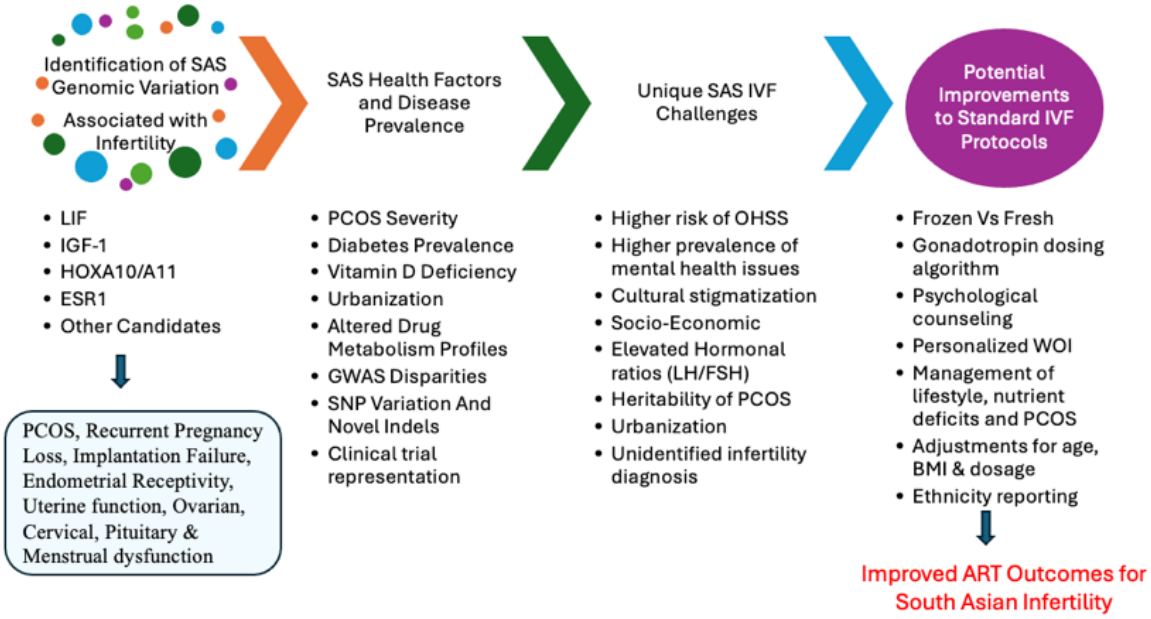


Figure 25. Schematic of Analysis and Recommendations for potential improvements in ART Outcomes for South Asian Infertility

Novel investigative pathways are imperative, including exploring the immunogenetic aspect of pharmacogenomics and delving into overlooked rare genetic

variations, which are purported to significantly influence differences in drug metabolism. Emerging technologies such as the fusion of machine learning and transcriptomics has given rise to computational frameworks aimed at systematically forecasting patient reactions to medications.

Leveraging artificial intelligence alongside transcriptomics has the potential to tailor match patients with infertility medications. The accuracy of these predictions escalates in tandem with the caliber and volume of the underlying data. Therefore, it is critical to remove ethnic disparities that currently exist in genome wide association studies. Lastly, clinical integration of novel ethnicity focused information, adaptation of federal regulations through the development of supportive practices, policies, tools, and infrastructures, will impact the entire ART ecosystem positively.

This leap forward heralds a transformative era, not merely for infertility itself, but for the empowerment of marginalized communities. Armed with the ability to customize treatments, physicians stand poised to shatter barriers and ignite a revolution in ART success rates on a global scale. This isn't just progress—it's a symphony of hope, resonating with the promise of brighter tomorrows for all those yearning to build their families.

So, here's to those who've faced the night,
Who've battled on with all their might,
In twists and turns, the reverberating symphony of our base pairs,
Each gene a story, with variations, many rare,
The challenge of science is but an interlude,

In humanity’s garden, no two flowers are quite the same in their plume,

For in the end, the light shines through,

And dreams come true, for me and you.

List of Definitions

Term	Definition
AMH	Anti-Mullerian hormone
ART	Assisted Reproductive Technology
BMI	Body Mass Index
CDC	Centers of Disease Control and Prevention
CMS	Centers for Medicare and Medicaid Services
Cryo	Cryopreservation
DALYs	Disability-adjusted life years
Endometrial Receptivity	Endometrial receptivity is the ability of the endometrium to successfully attach the blastocyst and nourish it.
FDA	Food and Drug Administration
FET	Frozen Embryo Transfer
GWAS	Genome Wide Association Studies
HLA	Human Leukocyte Antigen
Intracytoplasmic Sperm Injection (ICSI)	A micropipette is used to inject a single sperm into the center of the egg.
Implantation	the attachment of the fertilized egg or blastocyst to the wall of the uterus at the start of pregnancy
IUI	Intrauterine insemination is a type of artificial insemination is a procedure for treating infertility
IVF	In Vitro Fertilization is a type of artificial insemination is a procedure for treating infertility
LMICs	Low and Middle Income Countries
NCD	Non communicable disease
Oocyte	A cell in an ovary which may undergo meiotic division to form an ovum
Ovarian Hyperstimulation (OHSS)	Ovarian hyperstimulation syndrome is an exaggerated response to excess hormones. It usually occurs in women taking injectable hormone medications to stimulate the development of eggs in the ovaries.
PCOS	Polycystic ovary syndrome (PCOS) is a condition in which the ovaries produce an abnormal number of androgens, male sex hormones
PGx	Pharamcogenomics
Pharmacogenomics	Pharmacogenomics is the study of how a person's unique genetic makeup (genome) influences his or her response to medications.
pLOF	Predicted Loss of Function
RPL	Recurrent Pregnancy Loss
SAS	South Asian
SNP	Single Nucleotide Polymorphisms

Term	Definition
Transcriptomic	The study of all RNA molecules in a cell. RNA is copied from pieces of DNA and contains information to make proteins and perform other important functions in the cell.
Window of implantation (WOI)	Window of implantation is the span of a few days in which a female's endometrium lining is at its most optimal state to receive an embryo for implantation.
WGS	Whole genome sequence
Zygote intrafallopian transfer (ZIFT)	Ovum from donor female and sperm from donor male are fertilized in laboratory. The zygote or early embryo (up to 8 blastomeres stage) then transferred to fallopian tube

References

Abbara A. (2018), Clinical Parameters of Ovarian Hyperstimulation Syndrome Following Different Hormonal Triggers Of Oocyte Maturation in IVF Treatment. Clin Endocrinol (Oxf). doi: 10.1111/cen.13569. Epub 2018 Mar 6. PMID: 29446481; PMCID: PMC6001461.

Altmäe, S. et al. (2017) Meta-signature of human endometrial receptivity: a meta-analysis and validation study of transcriptomic biomarkers. Sci Rep 7, 10077. https://doi.org/10.1038/s41598-017-10098-3

Apparao KB. (2002) Elevated endometrial androgen receptor expression in women with polycystic ovarian syndrome. Biol Reprod 66: 297–304.

Bertrand, F.E.; Steelman, L.S.; Chappell, W.H.; Abrams, S.L.; Shelton, J.G.; White, E.R.; Ludwig, D.L.; McCubrey, J.A. Synergy between an IGF-1R antibody and Raf/MEK/ERK and PI3K/Akt/mTOR pathway inhibitors in suppressing IGF-1R-mediated growth in hematopoietic cells. Leukemia 2006, 20, 1254–1260.

Bhopal, R. (2000). What is the risk of coronary heart disease in South Asians? A review of UK research. Journal of Public Health, 22(3), 375-385.

Boyle, J. A. (2012). Prevalence of polycystic ovary syndrome in a sample of Indigenous women in Darwin, Australia. The Medical journal of Australia, 196(1), 62–66. https://doi.org/10.5694/mja11.10553

Center for Disease Control and Prevention. Assisted Reproductive Technology. Retrieved March 01, 2024, from https://nccd.cdc.gov/drh_art

Center for Disease Control and Prevention. Reproductive Health. Infertility. Retrieved March 01, 2024, from https://www.cdc.gov/reproductivehealth/infertility/index.htm

Cermik D. (2003). Regulation of HOXA10 expression by testosterone in vitro and in the endometrium of patients with polycystic ovary syndrome. J Clin Endocrinol Metab 88: 238–243

Cha, J., Sun, X., & Dey, S. K. (2012). Mechanisms of implantation: strategies for successful pregnancy. Nature medicine, 18(12), 1754–1767. https://doi.org/10.1038/nm.3012

Chambers JC. (2014) The South Asian genome. PLoS One.9(8):e102645. doi: 10.1371/journal.pone.0102645. PMID: 25115870; PMCID: PMC4130493.

Chimote B. (2013). Follicular fluid insulin like growth factor-1 (FF IGF-1) is a biochemical marker of embryo quality and implantation rates in in vitro

fertilization cycles. Journal of human reproductive sciences. 6. 140-146. 10.4103/0974-1208.117171.

Choi, J.-W. (2024) Effects of Melatonin, GM-CSF, IGF-1, and LIF in Culture Media on Embryonic Development: Potential Benefits of Individualization. *Int. J. Mol. Sci.* 2024, *25*, 751. https://doi.org/10.3390/ijms25020751

Clayton RN. (1992) How common are polycystic ovaries in normal women and what is their significance for the fertility of the population? Clin Endocrinol (Oxf). 37(2):127-34. doi: 10.1111/j.1365-2265.1992.tb02296.x. PMID: 1395063.

Cousineau, T. M. (2007). Psychological impact of infertility. Best practice & research. Clinical obstetrics & gynaecology, 21(2), 293–308. https://doi.org/10.1016/j.bpobgyn.2006.12.003

Dabrowski FA. (2015) The role of vitamin D in reproductive health--a Trojan Horse or the Golden Fleece? Nutrients. 7(6):4139-53. doi: 10.3390/nu7064139. PMID: 26035242; PMCID: PMC4488777.

Elkum N.(2014) Vitamin D insufficiency in Arabs and South Asians positively associates with polymorphisms in GC and CYP2R1 genes. PLoS One. 2014;9(11): e113102. doi:10.1371/journal.pone.0113102

Enciso, M. (2021). The precise determination of the window of implantation significantly improves ART outcomes. Scientific reports, 11(1), 13420. https://doi.org/10.1038/s41598-021-92955-w

GBD 2021 Fertility and Forecasting Collaborators (2024). Global fertility in 204 countries and territories, 1950-2021, with forecasts to 2100: a comprehensive demographic analysis for the Global Burden of Disease Study 2021. Lancet (London, England), S0140-6736(24)00550-6. Advance online publication.

Geldsetzer FD. (2022) Rural-Urban Differences in Diabetes Care and Control in 42 Low- and Middle-Income Countries: A Cross-sectional Study of Nationally Representative Individual-Level Data. Diabetes Care. 45(9):1961-1970. doi: 10.2337/dc21-2342. PMID: 35771765; PMCID: PMC9472489.

Global In Vitro Fertilization Market Size to Worth USD 39.12 Billion by 2032 | CAGR of 5.5%. Yahoo Finance, Retrieved Feb 01, 2024 from https://finance.yahoo.com/news/global-vitro-fertilization-market-size-080000927.html

Goetz (2018). Personalized medicine: motivation, challenges, and progress. Fertility and sterility, 109(6), 952–963. https://doi.org/10.1016/j.fertnstert.2018.05.006

Gondi S. (2019) Potential Implications of Private Equity Investments in Health Care Delivery. JAMA.321(11):1047-1048. doi: 10.1001/jama.2019.1077. PMID: 30816912; PMCID: PMC6682417.

Goodman, C. (2008). Vascular endothelial growth factor gene polymorphism and implantation failure. Reproductive biomedicine online, 16(5), pp.720-723.

Green, C.J.; Day, M.L. Insulin-like growth factor 1 acts as an autocrine factor to improve early embryogenesis in vitro. Int. J. Dev. Biol. 2013, 57, 837–844.

Gu Y. (2022) Life Modifications and PCOS: Old Story But New Tales. Front Endocrinol (Lausanne). 13:808898. doi: 10.3389/fendo.2022.808898. PMID: 35498415; PMCID: PMC9045543.

Gudmundsson S. (2022). Genome Aggregation Database Consortium; Rehm HL, MacArthur DG, O'Donnell-Luria A. Variant interpretation using population databases: Lessons from gnomAD. Hum Mutat.;43(8):1012-1030. doi: 10.1002/humu.24309. Epub 2021 Dec 16. PMID: 34859531; PMCID: PMC9160216.

Hales CN. (1992) Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. Diabetologia. 35(7):595-601. doi: 10.1007/BF00400248. PMID: 1644236.

Harinarayan, C.V. . (2011) Vitamin D status and bone mineral density in women of reproductive and postmenopausal age groups: a cross-sectional study from south India. J Assoc Physicians India, 59, pp. 698-704

Hu L. (2018) Influencing Factors of Pregnancy Loss and Survival Probability of Clinical Pregnancies Conceived Through Assisted Reproductive Technology. Reprod Biol Endocrinol;16(1):74. doi: 10.1186/s12958-018-0390-6. PMID: 30086781; PMCID: PMC6081896.

Kamath, M. S., Mascarenhas, M., Franik, S., Liu, E., & Sunkara, S. K. (2019). Clinical adjuncts in in vitro fertilization: a growing list. Fertility and sterility, 112(6), 978–986. https://doi.org/10.1016/j.fertnstert.2019.09.019

Khan, A. R. (2022). The Prevalence of Pharmacogenomics Variants and Their Clinical Relevance Among the Pakistani Population. Evolutionary bioinformatics online, 18, 11769343221095834. https://doi.org/10.1177/11769343221095834

Kiani, Z. (2021) The prevalence of depression symptoms among infertile women: a systematic review and meta-analysis. Fertil Res and Pract

Kobanawa M. (2023). The gonadotropins starting dose calculator, which can be adjusted the target number of oocytes and stimulation duration days to achieve individualized controlled ovarian stimulation in Japanese patients. Reprod Med

Biol. 22(1):e12499. doi: 10.1002/rmb2.12499. PMID: 36699956; PMCID: PMC9853467.

Kobayashi. (2017). The Window of Implantation Is Closed by Estrogen via Insulin-Like Growth Factor 1 Pathway. Journal of Reproduction and Infertility. 18. 231-241.

Kushnir VA. (2022) The Future of IVF: The New Normal in Human Reproduction. Reprod Sci. 2022 Mar;29(3):849-856. doi: 10.1007/s43032-021-00829-3. Epub 2022 Jan 3. PMID: 34981459; PMCID: PMC8722744.

Li Z, (2018) Leukaemia inhibitory factor in serum and follicular fluid of women with polycystic ovary syndrome and its correlation with IVF outcome. Reproductive biomedicine online, 36(4), 483–489. https://doi.org/10.1016/j.rbmo.2017.12.020

Mahmud G. (1995) A controlled assessment of the in vitro fertilization performance of British women of Indian origin compared with white women.Fertil Steril. 1995; 64: 103-106

March WA. (2010). The prevalence of polycystic ovary syndrome in a community sample assessed under contrasting diagnostic criteria. Hum Reprod.25(2):544-51. doi: 10.1093/humrep/dep399. Epub 2009 Nov 12. PMID: 19910321.

Martin AR. (2019) Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. 2019 Apr;51(4):584-591. doi: 10.1038/s41588-019-0379-x. Erratum in: Nat Genet. 2021 May;53(5):763. PMID: 30926966; PMCID: PMC6563838.

MOC, C. (2020). Pharmacogenomics: an evolving clinical tool for precision medicine. Cleveland Clinic journal of medicine, 87(2), 91.

Muccee, F., S. et al. (2024) Exploring the association of ESR1 and ESR2 gene SNPs with polycystic ovary syndrome in human females: a comprehensive association study. J Ovarian Res 17, 27. https://doi.org/10.1186/s13048-023-01335-7

Naranjo, M. G. et al. (2018). Interethnic variability in CYP2D6, CYP2C9, and CYP2C19 genes and predicted drug metabolism phenotypes among 6060 Ibero- and Native Americans: RIBEF-CEIBA Consortium Report on Population Pharmacogenomics. OMICS 22, 575–588

National Institutes of Health. National Library of Medicine. Clinical Trials. Last retrieved from www.clinicaltrials.gov

Novotný Z. (2009) Leukaemia inhibitory factor (LIF) gene mutations in women diagnosed with unexplained infertility and endometriosis have a negative

impact on the IVF outcome. A pilot study. Folia Biol (Praha). 55(3):92-7. PMID: 19545488.

Ochoa-Bernal MA. (2020) Physiologic Events of Embryo Implantation and Decidualization in Human and Non-Human Primates. International Journal of Molecular Sciences. 21(6):1973. https://doi.org/10.3390/ijms21061973

Oliveira JB. (2016). Association between leukaemia inhibitory factor gene polymorphism and pregnancy outcomes after assisted reproduction techniques. Reproductive biomedicine online, 32(1), 66–78. https://doi.org/10.1016/j.rbmo.2015.09.018

Paskulin DD. (2013) ESR1 rs9340799 is associated with endometriosis-related infertility and in vitro fertilization failure. Dis Markers. 35(6):907-913. doi:10.1155/2013/796290

Reddy TV. (2017) Impact of mitochondrial DNA copy number and displacement loop alterations on polycystic ovary syndrome risk in south Indian women. Mitochondrion. 44:35-40. doi: 10.1016/j.mito. PMID: 29278759.

Regan SLP, A. (2018) Growth hormone during in vitro fertilization in older women modulates the density of receptors in granulosa cells, with improved pregnancy outcomes. Fertil Steril. 110(7):1298-1310. doi:10.1016/j.fertnstert.2018.08.018

Risal, S. (2023). Transgenerational transmission of reproductive and metabolic dysfunction in the male progeny of polycystic ovary syndrome. Cell Reports Medicine, 4(5), 101035. https://doi.org/10.1016/j.xcrm.2023.101035.

Rizzuto, I., Behrens, R. F., & Smith, L. A. (2019). Risk of ovarian cancer in women treated with ovarian stimulating drugs for infertility. The Cochrane database of systematic reviews, 6(6), CD008215. https://doi.org/10.1002/14651858.CD008215.pub3

Rodin DA, SS. (1998) Polycystic ovaries and associated metabolic abnormalities in Indian subcontinent Asian women. Clin Endocrinol (Oxf). 49(1):91-9. doi: 10.1046/j.1365-2265.1998.00492.x. PMID: 9797852.

Rodin, D. A. (1998). Polycystic ovaries and associated metabolic abnormalities in Indian subcontinent Asian women. Clinical endocrinology, 49(1), 91-99.

Ruiz-Alonso M. (2012) The genomics of the human endometrium. Biochim Biophys Acta;1822:1931–42.

Ruiz-Alonso M. (2013) The endometrial receptivity array for diagnosis and personalized embryo transfer as a treatment for patients with repeated implantation failure. Fertil Steril;100:818–24.

Runcharoen, C. et al. (2021). Prevalence of pharmacogenomic variants in 100 pharmacogenes among Southeast Asian populations under the collaboration of the Southeast Asian Pharmacogenomics Research Network (SEAPharm). Hum Genome Var 8, 7

Sahana S. (2022) Pharmacogenomic landscape of Indian population using whole genomes. Clin Transl Sci 15(4):866-877. doi: 10.1111/cts.13153. Epub 2022 Mar 26. PMID: 35338580; PMCID: PMC9010271.

Shah, M. S. (2016). In vitro fertilization outcomes after fresh and frozen blastocyst transfer in South Asian compared with Caucasian women. Fertility and sterility, 105(6), 1484–1487. https://doi.org/10.1016/j.fertnstert.2016.02.027

Shahine LK. (2009) Poor Prognosis with In Vitro Fertilization in Indian Women Compared to Caucasian Women Despite Similar Embryo Quality. PLoS ONE 4(10): e7599. https://doi.org/10.1371/journal.pone.0007599

Siegel KR. (2014) Non-communicable diseases in South Asia: contemporary perspectives. Br Med Bull. 111(1):31-44. doi: 10.1093/bmb/ldu018. PMID: 25190759; PMCID: PMC4416117.

Southam, L. (2017). Whole genome sequencing and imputation in isolated populations identify genetic associations with medically-relevant complex traits. Nature communications, 8(1), 1-11.

Stewart, L. M. (2013). In vitro fertilization, endometriosis, nulliparity and ovarian cancer risk. Gynecologic oncology, 128(2), 260–264. https://doi.org/10.1016/j.ygyno.2012.10.023

Sun, B (2021). Factors Associated with Ovarian Hyperstimulation Syndrome (OHSS) Severity in Women with Polycystic Ovary Syndrome Undergoing IVF/ICSI. Frontiers in endocrinology, 11, 615957. https://doi.org/10.3389/fendo.2020.615957

The Genome Aggregation Database (gnomAD). Statistics. Retrieved March 01, 2024, from https://gnomad.broadinstitute.org/stats

The World Bank. South Asia. Retrieved March 01, 2024, from https://data.worldbank.org/indicator/SP.POP.TOTL?end=2022&locations=8S&most_recent_value_desc=false&start=1960&view=map

Uffelmann, E. (2021). Genome-wide association studies. Nat Rev Methods Primers 1, 59 https://doi.org/10.1038/s43586-021-00056-9

Vagnini LD, et al. (2019) Association between estrogen receptor 1 (ESR1) and leukemia inhibitory factor (LIF) polymorphisms can help in the prediction of

recurrent implantation failure. Fertil Steril. 111(3):527-534. doi:10.1016/j.fertnstert.2018.11.016.

Waldby. (2019, Oct 13). The Business of IVF: How Human Egg Cells went from Simple Cells to Valuable Commodity. The Conversation. Retrieved March 01, 2024 from https://theconversation.com/the-business-of-ivf-how-human-eggs-went-from-simple-cells-to-a-valuable-commodity-119168

Wall JD. (2023) South Asian medical cohorts reveal strong founder effects and high rates of homozygosity. Nat Commun;14(1):3377. doi: 10.1038/s41467-023-38766-1. PMID: 37291107; PMCID: PMC10250394.

WHO. WHO fact sheet. Retrieved March 01, 2024, from http://www.who.int/news-room/fact-sheets/detail/infertility

Wijeyaratne CN. (2002) Clinical manifestations and insulin resistance (IR) in polycystic ovary syndrome (PCOS) among South Asians and Caucasians: is there a difference? Clin Endocrinol (Oxf). 57(3):343-50. doi: 10.1046/j.1365-2265.2002.01603.x. PMID: 12201826.

Wilcox AJ.(1999) Time of implantation of the conceptus and loss of pregnancy. N Engl J Med;340:1796–9

Wild S. (2004) Global Prevalence of Diabetes: Estimates for the year 2000 and projections for 2030 . Diabetes Care 27 (5): 1047–1053. https://doi.org/10.2337/diacare.27.5.1047

www.Uniprot.org