

UNIVERSIDADE FEDERAL DE MINAS GERAIS

FACULDADE DE FARMÁCIA

PROGRAMA DE PÓS-GRADUAÇÃO EM ANÁLISES CLÍNICAS E
TOXICOLÓGICAS

NAYARA CRISTINA NUNES BARRETO

**PREDIÇÃO DO SUCESSO DE GESTAÇÃO UTILIZANDO ALGORITMOS DE
MACHINE LEARNING APÓS PROCEDIMENTOS DE FERTILIZAÇÃO *IN*
VITRO REALIZADOS POR UM SERVIÇO DE ATENDIMENTO PÚBLICO DE
SAÚDE**

BELO HORIZONTE

2021

NAYARA CRISTINA NUNES BARRETO

**PREDIÇÃO DO SUCESSO DE GESTAÇÃO UTILIZANDO ALGORITMOS DE
MACHINE LEARNING APÓS PROCEDIMENTOS DE FERTILIZAÇÃO *IN*
VITRO REALIZADOS POR UM SERVIÇO DE ATENDIMENTO PÚBLICO DE
SAÚDE**

Dissertação apresentada ao Programa de Pós-Graduação em Análises Clínicas e Toxicológicas, da Faculdade de Farmácia da Universidade Federal de Minas Gerais, como requisito para obtenção do título de mestre em Análises Clínicas e Toxicológicas

Orientadora: Prof^a. Dr^a. Karina Braga Gomes Borges

Co-orientadores: Dra. Inês Katerina Damasceno Cavallo Cruzeiro e Prof. Dr. Wagner Meira Júnior

BELO HORIZONTE – MG

2021

B812p Barreto, Nayara Cristina Nunes.
Predição do sucesso de gestão utilizando algoritmos de *Machine Learning* após procedimentos de fertilização *in vitro* realizados por um serviço de atendimento público de saúde / Nayara Cristina Nunes Barreto. – 2021.
82 f. : il.

Orientadora: Karina Braga Gomes Borges.
Coorientadores: Inês Katerina Damasceno Cavallo Cruzeiro.
Wagner Meira Júnior.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Farmácia, Programa de Pós-Graduação em Análises Clínicas e Toxicológicas.

1. Algoritmos – Teses. 2. Gravidez – Teses. 3. Fertilização *in vitro* – Teses. 4. Infertilidade feminina – Teses. 5. Infertilidade masculina – Teses. 6. Saúde pública – Teses. I. Borges, Karina Braga Gomes. II. Cruzeiro, Inês Katerina Damasceno Cavallo. III. Meira Júnior, Wagner. IV. Universidade Federal de Minas Gerais. Faculdade de Farmácia. V. Título.

CDD: 618.178



FOLHA DE APROVAÇÃO

PREDIÇÃO DO SUCESSO DE GESTAÇÃO UTILIZANDO ALGORITMOS DE MACHINE LEARNING APÓS PROCEDIMENTOS DE FERTILIZAÇÃO IN VITRO REALIZADOS POR UM SERVIÇO DE ATENDIMENTO PÚBLICO DE SAÚDE

NAYARA CRISTINA NUNES BARRETO

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ANÁLISES CLÍNICAS E TOXICOLÓGICAS, como requisito para obtenção do grau de Mestre em ANÁLISES CLÍNICAS E TOXICOLÓGICAS, área de concentração ANÁLISES CLÍNICAS E TOXICOLÓGICAS.

Aprovada em 05 de maio de 2021, pela banca constituída pelos membros:

Karina Braga Gomes Assinado de forma digital por
Karina Braga Gomes
Borges:03132898678
Dados: 2021.05.10 15:05:46 -03'00'

Profa. Karina Braga Gomes Borges - Orientadora
UFMG

Ines Katerina Digitally signed by Ines Katerina Damasceno
Cavallo Cruzeiro
DN: cn=Ines Katerina Damasceno Cavallo
Cruzeiro, o=UFMG, ou=Hospital das Clínicas,
email=ineskaterina@gmail.com, c=BR
Date: 2021.05.11 11:51:55 -03'00'

Profa. Inês Katerina Damasceno Cavallo Cruzeiro
UFMG

Alamanda Kfoury Assinado de forma digital por
Alamanda Kfoury Pereira:42077087668
Dados: 2021.05.05 16:41:19 -03'00'

Profa. Alamanda Kfoury Pereira

UFMG

Prof. Adnana de Paula Sabino
UFMG

Belo Horizonte, 5 de maio de 2021.



ATA DA DEFESA DA DISSERTAÇÃO DA ALUNA **NAYARA CRISTINA NUNES BARRETO**

Realizou-se, no dia 05 de maio de 2021, às 13:30 horas, Online, da Universidade Federal de Minas Gerais, a defesa de dissertação, intitulada *PREDIÇÃO DO SUCESSO DE GESTAÇÃO UTILIZANDO ALGORITMOS DE MACHINE LEARNING APÓS PROCEDIMENTOS DE FERTILIZAÇÃO IN VITRO REALIZADOS POR UM SERVIÇO DE ATENDIMENTO PÚBLICO DE SAÚDE*, apresentada por NAYARA CRISTINA NUNES BARRETO, número de registro 2019663974, graduada no curso de BIOMEDICINA, como requisito parcial para a obtenção do grau de Mestre em ANÁLISES CLÍNICAS E TOXICOLÓGICAS, à seguinte Comissão Examinadora: Profa. Karina Braga Gomes Borges - Orientadora (UFMG), Profa. Inês Katerina Damasceno Cavallo Cruzeiro (UFMG), Prof. Wagner Meira Junior (DCC/UFMG), Profa. Alamanda Kfoury Pereira (UFMG), Prof. Adriano de Paula Sabino (UFMG).

A Comissão considerou a dissertação:

(x) Aprovada

() Reprovada

Finalizados os trabalhos, lavrei a presente ata que, lida e aprovada, vai assinada por mim e pelos membros da Comissão. Belo Horizonte, 05 de maio de 2021.

Karina Braga Gomes
Borges:03132898678

Assinado de forma digital
por Karina Braga Gomes
Borges:03132898678
Dados: 2021.05.10
15:04:56 -03'00'

Profa. Karina Braga Gomes Borges (Doutora)

Ines Katerina Damasceno
Cavallo Cruzeiro

Digitally signed by Ines Katerina Damasceno Cavallo
Cruzeiro
DN: cn=Ines Katerina Damasceno Cavallo Cruzeiro,
ou=UFMG, ou=Hospital das Clínicas,
email=ineskaterina@gmail.com, c=BR
Date: 2021.05.11 11:51:03 -0300

Profa. Inês Katerina Damasceno Cavallo Cruzeiro (Doutora)

Alamanda Kfoury
Pereira:42077087668

Assinado de forma digital por
Alamanda Kfoury Pereira:42077087668
Dados: 2021.05.05 16:40:21 -03'00'

Profa. Alamanda Kfoury Pereira (Doutora)

Prof. Adriano de Paula Sabino (Doutor)

AGRADECIMENTOS

Agradeço à Deus pelo infinito amor, pelas graças derramadas e por todo cuidado comigo e com minha família. À Nossa Senhora, minha intercessora, mãe amada e protetora, sempre à minha frente.

Ao meu esposo, Albert, pelo companheirismo, cuidado e incentivo. Por me dar a calma necessária nos dias mais difíceis e por ficar ao meu lado em todos os momentos. Obrigada por tornar o meu sonho a sua prioridade e por ser tão especial em minha vida.

Às minhas irmãs e minha mãe pelas orações e pelo apoio incondicional.

À minha orientadora, Karina Braga, pelo aprendizado, paciência e acolhimento. Por todo incentivo, essencial na minha caminhada. Obrigada por ser tão atenciosa e exemplar.

À Giulia pela contribuição e por sempre estar disposta a ajudar.

Aos participantes deste estudo, pelo aprendizado e pelas importantes contribuições, essenciais para que chegássemos até aqui.

À Universidade Federal de Minas Gerais e ao Programa de Pós-Graduação em Análises Clínicas e Toxicológicas pela oportunidade e o conhecimento adquirido.

RESUMO

A infertilidade, definida como ausência de concepção após um período médio de 12 meses de relação sexual desprotegida, tornou-se um problema de saúde global. A infertilidade afeta cerca de 37 a 70 milhões de casais em todo o mundo, o que justifica o aumento do número de casais em busca de fertilização *in vitro* (FIV). Apesar dos avanços e aprimoramentos técnicos nos procedimentos de fertilização artificial, alguns casais não conseguem ter sucesso devido à alta complexidade da técnica e aos vários fatores, controláveis ou não, que podem comprometer o resultado final. *Machine Learning* (ML) é uma subárea da Inteligência Artificial que se dedica ao estudo de algoritmos e modelos estatísticos para executar uma determinada tarefa sem a necessidade do uso de instruções explícitas, com o intuito de gerar modelos preditivos do desfecho. O objetivo deste estudo foi aplicar modelos de ML para prever o sucesso de gestação em mulheres submetidas ao procedimento de FIV em um serviço público de saúde, incluindo variáveis pré-implantacionais. Este estudo incluiu a análise de um banco de dados compreendendo 771 casos de casais submetidos à FIV no Hospital das Clínicas da Universidade Federal de Minas Gerais, entre 2013 e 2019. Foram utilizados os seguintes algoritmos baseados em ML: Regressão Logística, *Random Forest*, *XG Boost* e *Support Vector Machines* (SVM). O algoritmo *Random Forest* apresentou o melhor desempenho, com melhor acurácia, sensibilidade e área sob a curva ROC, indicando as 20 variáveis, dentre as 90 totais do banco de dados, que são mais relevantes para prever o sucesso da gravidez após a FIV. Nosso algoritmo de ML pode ser útil para estimar a chance de gravidez em mulheres submetidas à FIV, bem como definir as variáveis nas quais a intervenção clínica pode melhorar o sucesso do tratamento.

Palavras-chave: Infertilidade, FIV / ICSI, modelo de predição, Machine Learning, Random Forest.

ABSTRACT

Infertility, defined as the absence of conception after an period of 12 months of unprotected sexual intercourse, has become a global health problem. Infertility affects about 37 to 70 million couples worldwide, which justifies the increase in the number of couples looking for *in vitro* fertilization (IVF). Despite advances and technical improvements in artificial fertilization procedures, some couples are unable to succeed due to the high complexity of the technique and several variables, controllable or not, that can compromise the final result. Machine Learning (ML) is a sub-area of Artificial Intelligence that is dedicated to the study of algorithms and statistical models to perform a procedure without the need for the use of explicit instructions, in order to generate predictive models of the outcome. The objective of this study was to apply ML models to predict the success of pregnancy after the IVF procedure in a public health service, including pre-implantation variables. This study included the analysis of a database comprising 771 cases of couples undergoing IVF at the Hospital das Clínicas of the Federal University of Minas Gerais, between 2013 and 2019. The ML-based algorithms were used: Logistic Regression, Random Forest, XG Boost and Support Vector Machines (SVM). The Random Forest algorithm showed the best performance, with better accuracy, sensitivity and area under the ROC curve, indicating the 20 variables, among the 90 totals in the database, that are most relevant to predict the success of pregnancy after IVF. Our ML algorithm can be useful for predicting pregnancy in women undergoing IVF, as well as defining the variables in which clinical intervention can improve treatment success.

Keywords: Infertility, FIV / ICSI, Prediction model, Machine Learning, Random Forest.

LISTA DE ABREVIÇÕES E SIGLAS

Adaboost: adaptive boosting
AMH: hormônio antimülleriano
BMI: body mass index
CART: classification and Regression Trees
CFA: contagem de Folículos Antrais
COEP: comitê de Ética em Pesquisa
DM: data mining
DIP: doença inflamatória pélvica
DST: doenças sexualmente transmissíveis
DHS: demographic and Health Surveys
DNN: deep neural networks
E2: estradiol
FIV: fertilização in vitro
FSH: hormônio folículo estimulante
gnRH: hormônio liberador de gonadotrofina
hCG: gonadotrofina coriônica humana
ICSI: injeção intracitoplasmática de espermatozoides
IA: inteligência Artificial
LH: hormônio luteinizante
LR: logistic Regression
ML: machine learning
MESA: microscopic epididymal sperm aspiration
micro – TESE: microdissecação testicular
MvLRM: modelo de regressão logística multivariada
MII: metaphase II
MLP: multilayer perceptron
NBC: naïve bayes classifier
OMS: organização Mundial da Saúde
1NN: one-nearest neighbor
PESA: percutaneous epididymal sperm aspiration
PGD: diagnóstico genético pré-implantação

RF: random forest

REDLARA: latin american network of assisted reproduction

RIMARC: ranking instances by maximizing the área under ROC curve

RNA: rede neural artificial

ROC curve: receiver operating characteristic curve

RPART: recursive Partitioning

SERA: success estimating using a ranking algorithm

SHAP: shapley additive explanations

SMOTE: super-sampling technique

SVM: support Vector Machines

SOP: síndrome dos ovários policísticos

TESA: punção do epidídimo ou no testículo

TESE: biópsia testicular

US: ultrassonografia

LISTA DE FIGURAS

Figura 1. Prevalência de infertilidade primária entre mulheres.....	23
Figura 2. Prevalência de infertilidade secundária entre mulheres.	23
Figura 3. Tipos de abordagem pelos algoritmos de Machine Learning	27
Figure 1. Subdivision of the training set and the model validation.	62
Figure 2. Area under the ROC curve for the pregnancy prediction model using Random Forest.....	63
Figure 3. Confusion matrix of the Random Forest model.	64
Figure 4. Individual SHAP result for the variables indication for (A) IVF treatment and (B) catheter model.	65
Figure 5: SHAP value for the Random Forest model.	66
Figure 6. Algorithm applied in the prediction of (A) a real positive pregnancy and (B) a real negative pregnancy.	67

LISTA DE TABELAS

Tabela 1. Resultados da AUC e precisão no conjunto de dados de fertilização in vitro.	33
Tabela 2. Resultados da AUC e precisão no conjunto de dados de fertilização in vitro.	34
Tabela 3. Resultados da sensibilidade e especificidade no conjunto de dados de fertilização in vitro.....	34
Table 1. Set of varied hyperparameters for each algorithm	57
Table 2. Variables included in the final algorithm models	58
Table 3. Characterization of the cases included in the study, according to the presence or absence of pregnancy in univariate analysis.	59
Table 4. Performance of machine learning models applied.	61
Supplementary material 1.....	67

SUMÁRIO

1 INTRODUÇÃO	13
2 REFERENCIAL TEÓRICO	16
2.1 Infertilidade	16
2.2 Causas de Infertilidade	17
2.3 Infertilidade Feminina	17
2.4 Infertilidade Masculina	20
2.5 Infertilidade inexplicável	21
2.6 Epidemiologia	22
2.7 Fertilização in vitro	24
2.8 Machine Learning	26
2.8.1 O que é Machine Learning	26
2.8.2 Quando utilizar o Machine Learning	27
2.8.3 Técnicas de Classificação de Machine Learning	29
2.8.4 Trabalhos que utilizaram Machine Learning aplicados na FIV	30
3 OBJETIVOS	34
3.1 Objetivo geral	34
3.2 Objetivos específicos	35
4 CAPÍTULO 1	36
5 DISCUSSÃO	70
6 CONCLUSÕES	72
7 PERSPECTIVAS	72
REFERÊNCIAS	73

1. INTRODUÇÃO

O termo infertilidade é aplicado nos casos em que há ausência de concepção após um ano de atividade sexual desprotegida [1]. Estima-se que a infertilidade atinge entre 8% e 12% dos casais em idade reprodutiva em todo o mundo [2]. Sendo assim, um em cada seis casais buscará auxílio de um especialista em algum momento de suas vidas devido às dificuldades de conceber [3]. A infertilidade é causada por diversos fatores como características ambientais, genéticas ou pessoais [4].

A fertilização *in vitro* (FIV) é comumente utilizada como tratamento para a infertilidade. O primeiro nascimento humano resultante dessa técnica ocorreu em 1978, na Inglaterra, sendo considerado naquela época um procedimento reprodutivo complexo [5] que envolve os processos de estimulação ovariana, recuperação de oócitos, fertilização, cultura de embriões e transferência de embriões para o útero [6]. Desde então, mais de 6 milhões de crianças nasceram por meio da FIV [7] e mesmo com os avanços e melhorias técnicas, a eficácia é considerada relativamente baixa [8], sendo em média, apenas um terço dos ciclos de FIV resultante em gravidez [9].

O sucesso da fertilização *in vitro* envolve vários aspectos, como a idade do casal, a causa da infertilidade, qualidade seminal, a qualidade embrionária, entre outras variáveis. A escolha do melhor embrião para a transferência ao útero, o qual é escolhido com base em características morfológicas por meio da avaliação visual, sendo esta uma tarefa complexa para o embriologista, que pode produzir resultados diferentes do esperado, o que pode justificar uma baixa taxa de sucesso da FIV [8]. Sendo assim, o emprego de recursos para auxiliar no prognóstico do tratamento de FIV é sempre bem visto tanto para os pacientes, quanto para os médicos [5]. Embora existam processos para melhorar a taxa de sucesso no tratamento, muitas vezes é necessário adotar métodos mais avançados de predição deste desfecho, como por exemplo, ferramentas automatizadas baseadas em Inteligência Artificial, de forma a tornar possível a análise de modo mais inteligente dos fatores que interferem na eficácia do tratamento [4].

Machine Learning (ML) é uma subárea da Inteligência Artificial, a qual possibilita computadores se submeterem a modos de autoaprendizagem, sem o uso de programação explícita [10], por meio da exposição de programas de computadores a um conjunto de dados [9]. Em termos gerais, a abordagem básica de todo ML é proporcionar às máquinas a habilidade de aprender, utilizando dados anteriores, e o reconhecimento de padrões para desenvolver algoritmos de classificação altamente precisos, e dessa forma, entregar respostas com elevado grau de confiança [11].

Sistemas baseados em ML, na maioria das vezes, objetivam realizar previsões, algo que na área médica pode fornecer valiosas informações durante a realização de diagnósticos de doenças ou no decorrer do tratamento. Dessa forma, a maioria dos estudos que abordam estes dois temas utiliza este tipo de algoritmo, pelo fato de serem capazes de empregar métodos estatísticos para encontrar padrões ao analisar um grande volume de informações provenientes de dados históricos ao longo da etapa de treinamento dos modelos de previsão [9,11, 12].

Nos últimos anos, o ML está rapidamente ganhando força nos setores de reprodução humana e embriologia. Somente no ano de 2018, a quantidade de estudos abordando o tema cresceu mais de sete vezes nos anais da Sociedade Americana de Medicina Reprodutiva (ASRM) e Sociedade Europeia de Reprodução Humana e Embriologia (ESHRE) [9].

Alguns estudos descreveram técnicas de ML para a predição de gravidez após fertilização *in vitro*. Um dos primeiros estudos foi realizado por Kaufmann et al. (1997) [3], utilizando Rede Neural Artificial (RNA), os quais identificaram quatro atributos que consideraram como sendo os mais relevantes (idade do paciente, número de óvulos recuperados, número de embriões transferidos e se houve congelamento de embriões). Estes atributos foram usados como dados de entrada para treinar a RNA e alcançou precisão de 59% na predição. Mais tarde, Durairaj e Thamilselvan (2013) [13] utilizaram a mesma técnica de RNA para prever o índice de sucesso do tratamento com FIV, mas dessa vez utilizando 14 atributos e uma precisão de 73% foi alcançada. Posteriormente, o estudo foi expandido ao integrar mineração de dados à RNA, que foram então utilizados para treinar um classificador que pôde prever com eficiência os resultados de gravidez em 90% dos casos. Já Uyar et al. (2015) [12] investigaram

vários tipos de classificadores na previsão de resultados de embriões da FIV usando 18 características no total, sendo que nove se relacionavam aos pacientes e os outros nove aos embriões. Por fim, Durairaj e Ramasamy (2017) [4] propuseram um algoritmo híbrido que combina a Colônia de Formigas e Redução Relativa existente para pré-processar os dados de FIV, e assim reduziram o número de atributos de 41 para 18 sem comprometer o treinamento do sistema, atingindo uma precisão de 90%. Apesar dos recentes trabalhos da literatura mostrarem que as técnicas de aprendizado de máquina são promissoras na orientação do plano de tratamento de FIV, seu desempenho carece de um maior refinamento para que seja utilizada na rotina médica, em especial à sua validação em condições diversas em que o procedimento é realizado [5].

Até o momento, nenhum estudo avaliou a aplicação das técnicas de ML na predição do sucesso de gestação após FIV em um serviço de atendimento público de saúde. Sabe-se que a origem do serviço pode impactar neste sucesso, seja por indisponibilidade de novas tecnologias, bem como pela condição de saúde do casal proponente e tempo de espera para realizar o tratamento. Desta forma, este trabalho se justifica pela possibilidade de avaliar, por meio de algoritmos de ML, variáveis que possam contribuir para o sucesso do procedimento em um serviço de saúde pública. Estes resultados, portanto, poderão fornecer ferramentas de otimização das variáveis de modo a aumentar a chance de gravidez após a realização da FIV, nesse grupo de pacientes.

2. REFERENCIAL TEÓRICO

2.1 Infertilidade

Nos últimos anos, o aprimoramento da qualidade na saúde materno-infantil tem mostrado sucesso em todo o mundo, em parte devido ao foco na saúde reprodutiva, tornando-se a cada dia prioridade nos sistemas de saúde público-privados. Entretanto, a infertilidade é um fator crítico nesse processo de evolução e vem sendo muitas vezes negligenciada. A incapacidade de ter filhos afeta homens e mulheres em todo o mundo, podendo causar distúrbios psicossociais como angústia e depressão, bem como discriminação e isolamento social [14].

O termo “infertilidade” é definido quando há ausência de concepção após um período de doze meses de relação sexual desprotegida e regular [1; 15].

A infertilidade afeta cerca de 37 a 70 milhões de casais em todo o mundo e este valor vem aumentando a cada ano [16]. Esse aumento deve-se a pelo menos quatro fatores:

- Atraso na gravidez, uma vez que as mulheres têm protelado a gestação;
- Alterações no sêmen devido a hábitos como tabagismo e alcoolismo;
- Mudanças no comportamento sexual da população;
- Eliminação de tabus.

Nas últimas décadas, ocorreram mudanças importantes no âmbito da saúde reprodutiva. Primeiramente, o estudo de técnicas reprodutivas abriu caminhos para o desenvolvimento de tecnologias de reprodução assistida como a fertilização *in vitro* (FIV) e a injeção intracitoplasmática de espermatozoides (do inglês *intra-cytoplasmic sperm injection* - ICSI. Segundo, ocorreram alterações em nossa sociedade, como o aumento da proporção de mulheres acima de 35 anos de idade em busca de gravidez. E terceiro, a evolução da biologia molecular

e genética tornou-se preponderante para o estudo, diagnóstico e avaliação de casais, cuja causa de infertilidade não era possível de ser determinada antes do surgimento destas novas análises [17].

2.2 Causas de Infertilidade

As principais causas de infertilidade podem variar entre homens e mulheres, sendo divididas de maneira geral em: fator masculino, diminuição da reserva ovariana, disfunção ovulatória, fator tubário, fator uterino e infertilidade inexplicável. Nos homens, inclui-se varicocele, azoospermia, oligospermia, astenospermia, teratospermia, disfunção ejaculatória [2, 18].

A avaliação da infertilidade pode ser direcionada às pacientes que possuem critérios ou alto risco de infertilidade. Mulheres com idade superior a 35 anos com mais de 6 meses de relações desprotegidas, sem gravidez, devem receber com celeridade avaliação e tratamento, se indicado. A obtenção do histórico médico abrangente, que incluam itens que destacam potenciais etiologias da infertilidade, deve ser obtido juntamente com o parceiro, caso seja possível. Na mulher, os testes são direcionados para a reserva ovariana, função ovulatória e anormalidades estruturais, uterinas e tubárias. Já nos homens, o histórico médico e a análise do sêmen faz-se necessário, e todos devem ser assistidos por um especialista em medicina reprodutiva [18].

2.3 Infertilidade Feminina

De modo geral, as mulheres costumam receber mais atenção no processo de avaliação da fertilidade devido à maior quantidade de diagnósticos e tratamentos disponíveis atualmente para elas, além do fato de precisarem manter a gravidez até o nascimento. Os seguintes fatores estão diretamente ligados à infertilidade feminina:

- **Idade**

Em todo o mundo, a idade média das mulheres na primeira gestação tem aumentado nas últimas décadas. Sendo assim, a postergação da gravidez contribui de forma significativa para a ocorrência de infertilidade em mulheres,

uma vez que durante o processo de envelhecimento, o organismo feminino sofre diminuição das reservas ovarianas, o que provoca uma maior incidência de problemas cromossômicos em seus oócitos. Além disso, tal condição é refletida pela mudança de hormônios e indicam que ela está se aproximando da perimenopausa [19, 20, 21].

Apesar das mulheres continuarem a ter ciclos menstruais regulares após os 35 anos, a porcentagem real de ciclos ovulatórios diminui consideravelmente. Dessa forma, a maioria delas possuem a falsa crença de que a presença contínua de um ciclo menstrual normal sinaliza que elas ainda possuem plena capacidade de engravidar [20, 22].

Embora a quantidade de oócitos disponíveis diminua com a idade, a principal razão do declínio da fertilidade relacionada à idade deve-se a piora da qualidade oocitária, refletida no aumento exponencial de aneuploidias oocitárias, que consistem na alteração no número normal de cromossomos no gameta [23].

- ***Diminuição de reserva ovariana***

A reserva ovariana prediz o potencial reprodutivo dos ovários, e é representada pelo número de oócitos disponíveis para fertilização. Essa avaliação pode ser feita por meio de testes séricos, como dosagem do hormônio folículo estimulante (FSH) em fase folicular, estradiol em fase folicular, hormônio antimülleriano, ou ultrassonografia com contagem de folículos antrais e a avaliação dessas análises devem ser em conjunto com o contexto da idade da paciente [24, 18].

Não existem critérios definitivos para a classificação de reserva ovariana diminuída, porém, os valores apresentados abaixo podem ser consistentes com o quadro clínico:

- valor do hormônio antimülleriano (do inglês *anti-müllerian hormone* - AMH) menor que 1 ng/mL
- contagem folicular antral menor que 5-7, e
- hormônio folículo-estimulante (FSH) maior que 10 UI/L, ou
- uma história de baixa resposta à estimulação para fertilização *in vitro* (menos de quatro oócitos na recuperação do óvulo).

Pode-se ainda avaliar a reserva ovariana com a medida conjunta do estradiol e o FSH entre os dias 2 e 5 do ciclo, sendo que o FSH superior a 10

UI/L está associado a uma menor responsividade à estimulação ovariana. O estradiol deve ser inferior a 60-80 pg/mL, sendo que níveis elevados podem ter efeito supressor nos níveis de FSH e são indicativos de diminuição da reserva ovariana [24, 18].

A contagem de Folículos Antrais (CFA) corresponde à avaliação do número de folículos antrais com medida de 2 e 10 mm presentes na avaliação ultrassonográfica dos ovários na fase folicular inicial. O CFA é considerado como bom preditor da resposta ovariana à estimulação com gonadotrofinas em tratamento de reprodução assistida e quanto maior o número de folículos antrais, melhor a reserva ovariana da mulher [24, 18].

- **Fator tubário**

Pelo fato de que a fertilização do embrião ocorre nas tubas uterinas, quando estas estão danificadas ou bloqueadas, tal barreira física impede que os espermatozoides cheguem até os oócitos. Dessa forma, o fator que mais contribui para o bloqueio tubário é a doença inflamatória pélvica (DIP), que se manifesta por meio de complicações graves de doenças sexualmente transmissíveis (DST), como a clamídia e a gonorreia. Tal condição afeta de 1% a 2% das mulheres com menos de 25 anos sexualmente ativas. O histórico de DSTs, a precoce iniciação sexual, múltiplos parceiros sexuais, a não utilização de preservativo, ducha vaginal, fumo, uso de álcool e droga são fatores de risco para a ocorrência da DIP [25, 26].

- **Disfunção ovulatória**

A disfunção ovulatória é geralmente representada por distúrbios menstruais, seja oligomenorréia ou amenorréia. As causas mais comuns da disfunção ovulatória são: síndrome dos ovários policísticos (SOP), obesidade, ganho ou perda de peso, exercício extenuante, disfunção tireoidiana e hiperprolactinemia. Entretanto, a causa específica da disfunção ovulatória muitas vezes permanece inexplicável [18, 61].

A história menstrual é uma importante ferramenta para análise da disfunção ovulatória. A maioria das mulheres apresentam ciclos menstruais regulares (a cada 25 a 35 dias). Entretanto, até um terço das mulheres com ciclos menstruais

normais são anovulatórias, e por isso a confirmação da ovulação deve ser considerada no estudo da infertilidade [18, 61].

A SOP é a desordem endócrina mais comum em mulheres em idade reprodutiva e afeta de 5% a 10% das mulheres com idades entre 15 a 44 anos, sendo importante na epidemiologia da disfunção ovulatória [27]. Devido à resistência à insulina que afeta aproximadamente 50% a 70% das mulheres com SOP, comorbidades podem ocorrer, como por exemplo, síndrome metabólica, hipertensão, dislipidemia, intolerância à glicose e diabetes mellitus. Além disso, as mulheres com SOP apresentam maior risco de distúrbios psiquiátricos, incluindo depressão, ansiedade, transtorno bipolar e transtorno da compulsão alimentar periódica [27, 28].

- ***Fator Uterino***

Aliado à permeabilidade tubária, o estabelecimento de uma cavidade uterina normal indica que a anatomia da paciente viabilizará uma gravidez normal, enquanto que a presença de anormalidades uterinas podem afetar negativamente o sucesso da gestação [1, 29].

As causas de infertilidade associadas aos fatores uterinos incluem pólipos endometriais, sinéquias, anomalias de Müllerianas e leiomiomas. A avaliação das alterações uterinas normalmente é feita por meio de testes de imagem e testes fisiológicos, que consistem em: histerossalpingografia, ultrassonografia transvaginal, histerossonografia e histeroscopia [1, 29].

- ***Endometriose e miomas***

Estima-se que aproximadamente 10% a 15% das mulheres em idade reprodutiva, e mais de 30% das mulheres com infertilidade, sofram de endometriose. Já os miomas uterinos, como por exemplo leiomiomas, afetam até 77% das mulheres em idade reprodutiva, nas quais 20% a 50% são sintomáticas. No entanto, devido ao sangramento uterino disfuncional e à dor pélvica, ambas as condições podem afetar diretamente a fertilidade e a qualidade de vida das mulheres afetadas [30, 31].

2.4 Infertilidade Masculina

Cerca de 7% dos homens são inférteis e as causas mais comuns são tabagismo, alcoolismo, exposição a substâncias tóxicas, caxumba após a puberdade, varicoceles, testículos atrofiados e DST's.

- **Condições adquiridas**

Condições adquiridas são comumente causadas por eventualidades que incluem infecção ou inflamação, imunoinfertilidade, trauma, intervenção cirúrgica nos órgãos reprodutivos, bem como a exposição a toxinas, como por exemplo agrotóxicos, ou a produtos químicos e anabolizantes [33].

As infecções podem deteriorar a qualidade do esperma, podendo ser o resultado de enfermidades que incluem tuberculose genital e prostatite. Já os fatores imunológicos são tipicamente secundários aos traumas - uma vez que o esperma não é formado até a puberdade, os espermatozoides são mantidos em ambientes imune-protetidos por barreiras sangue-testicular. Caso essa barreira seja rompida por trauma ou cirurgia, os espermatozoides passam a ser considerados elementos estranhos pelo organismo, levando ao desenvolvimento de anticorpos anti-espermatozoides [33].

- **Condições congênitas**

As condições congênitas são fatores presentes desde o nascimento do indivíduo. Dentre os diversos fatores, destaca-se a varicocele, que consiste no aumento das veias do cordão espermático produzindo uma massa tumoral compressível no escroto, mas que pode ser corrigida por meio de cirurgia [34]. Outro fator congênito é o testículo atrofiado. A idade no qual o reparo é realizado afeta diretamente na reversibilidade da infertilidade [35]. Existem também as condições congênitas ligadas a anomalias endócrinas do eixo hipotálamo-hipófise-gonadal e as disfunções genéticas como a síndrome de Klinefelter e outras anormalidades no cromossomo Y [33, 34, 36]. Por fim, estima-se que 1% a 2% de casos de infertilidade masculina sejam atribuídos à inexistência bilateral congênita do ducto deferente, observada em indivíduos portadores de mutações no gene relacionado à fibrose cística.

2.5 Infertilidade inexplicável

A infertilidade inexplicada ocorre quando os critérios de infertilidade são atendidos, porém todos os testes são normais. Como triagem, esses pacientes devem ter avaliação da ovulação, permeabilidade tubária e qualidade do sêmen [18].

2.6 Epidemiologia

Nos dias atuais, a infertilidade segue como sendo uma condição global prevalente. Em determinadas regiões do mundo, suas taxas são maiores que a média mundial, podendo chegar a 30% em alguns países, como por exemplo, na Ásia sul e central, no norte africano, na África Subsaariana, no Oriente Médio e na Europa Central e Oriental [37, 38, 39, 40].

Vale ressaltar que estimar o número total de pessoas inférteis no mundo é algo de difícil obtenção devido à heterogeneidade dos critérios atribuídos na definição do conceito de infertilidade, uma vez que algumas métricas utilizam um ano, dois anos ou até cinco anos de “tentativa” de engravidar. Outro fator complicador são as diferenças críticas entre estimativas de infertilidade baseadas em investigações populacionais em larga escala *versus* estudos epidemiológicos de infertilidade. Por último, tem-se o desafio de se padronizar se a infertilidade se refere somente às “mulheres”, ao “casal” ou à “pessoa” como indivíduo independentemente do sexo. Sendo assim, tais parâmetros analíticos muitas vezes são aplicados de maneira intercambiável ou sem precisão [41].

Entretanto, nos últimos anos alguns estudos sobre o número de pessoas inférteis no mundo foram publicados. O primeiro, em consonância com a Organização Mundial da Saúde (OMS), utilizou dados de 47 pesquisas realizadas em países em desenvolvimento no âmbito do programa *Demographic and Health Surveys* (DHS). As pesquisas consistiram em entrevistas com mais de 495.000 mulheres entre 15 e 49 anos de idade que foram questionadas acerca do uso passado ou atual de métodos contraceptivos, infertilidade primária e secundária, infecundidade autorreferida e indicações de infecundidade secundária entre mulheres casadas em idade reprodutiva. Como resultado, estima-se que em meados de 2002, havia cerca de 186 milhões de mulheres casadas de 15 a 49 anos consideradas inférteis, devido à infertilidade primária ou secundária, nos países em desenvolvimento, excluindo a China. Tais

números na época correspondiam a mais de um quarto das mulheres casadas em idade reprodutiva nesses países [40, 42].

Sendo assim, define-se infertilidade primária quando há ausência de gestação, sem histórico de outra com sucesso e a infertilidade secundária quando há histórico anterior de gestação com sucesso. As Figuras 1 e 2 mostram a taxa de infertilidade primária e secundária no mundo, respectivamente [14]:

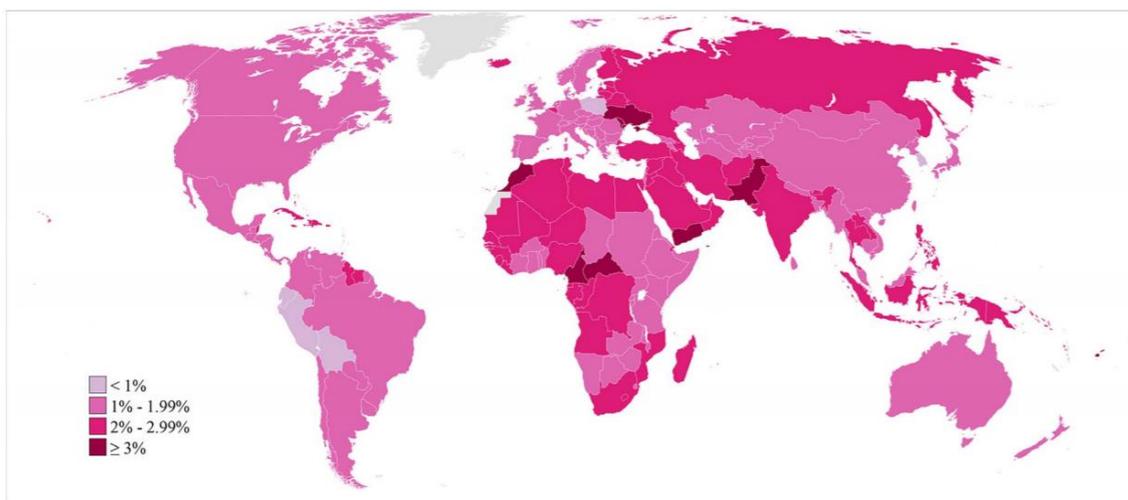


Figura 1. Prevalência de infertilidade primária entre mulheres (2010). Adaptado de Mascarenhas, et al (2012) [14].

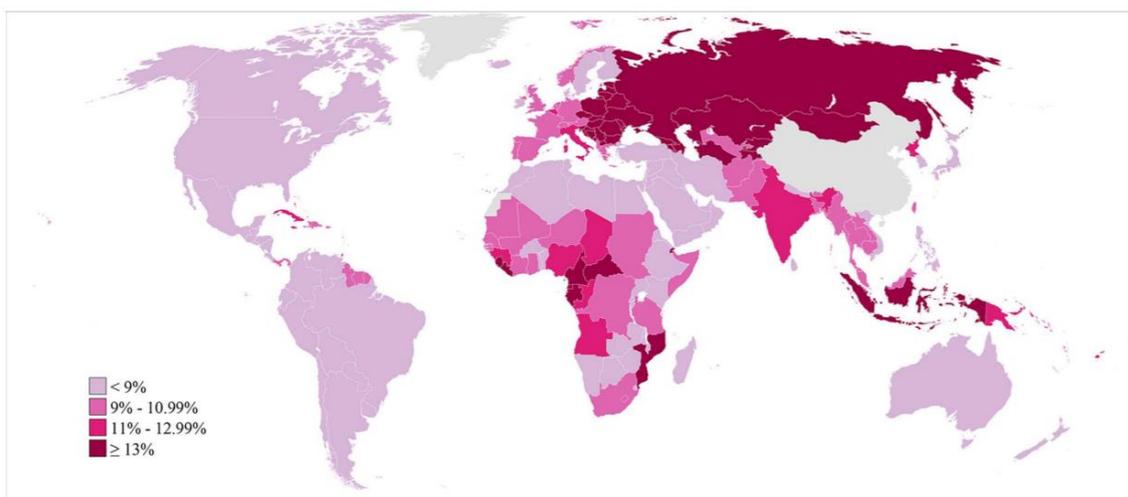


Figura 2. Prevalência de infertilidade secundária entre mulheres (2010). Adaptado de Mascarenhas et al (2012) [14].

Um segundo estudo, baseado em dados de 1990, utilizou 25 pesquisas populacionais de inúmeros países em desenvolvimento. Com uma base amostral de 172.413 mulheres analisadas e extrapolando as estatísticas para a população

mundial de 2007, estimou-se que 72,4 milhões de mulheres eram inférteis, sendo que destas, 40,5 milhões delas procuravam tratamento para engravidar [16].

No entanto, o estudo mais recente realizado por Mascarenhas et al. (2012) [14] e apoiado pela OMS e pela Fundação Bill & Melinda Gate, integrou o Estudo Global de Carga de Doenças de 2010 com o intuito de fornecer um panorama global das tendências de infertilidade mundial com base na análise de 277 pesquisas sobre saúde e reprodução disponíveis em 190 países, entre os anos de 1990 a 2010. Ao invés de usar as definições clínicas ou epidemiológicas de infertilidade da OMS (ou seja, ausência de concepção após 1 ou 2 anos de tentativas, respectivamente), a pesquisa foi permeada pelo conceito de infertilidade primária e secundária como medida de resultado ao longo de um período de exposição de 5 anos, com base no estado estável da união, falta de uso de contraceptivos e desejo de ter um filho. Por meio desta definição demográfica de infertilidade, estimou-se que 48,5 milhões de casais foram afetados pela infertilidade em 2010. Tal número é consideravelmente menor do que em relatórios anteriores. Todavia, segundo a OMS, ao se reduzir o período de 5 anos para 2 anos, o número total de casais inférteis aumentaria em 2,5 vezes, elevando para 121 milhões a estimativa de casais inférteis no mundo [43].

2.7 Fertilização in vitro

O primeiro nascimento por meio da fertilização *in vitro* (FIV) ocorreu em 1978 sendo um marco para a medicina reprodutiva. Desde então, estima-se que mais de 7 milhões de crianças tenham nascido por meio dessa técnica [44, 45].

A FIV é considerada um procedimento de alta complexidade e constitui-se na fertilização realizada em laboratório, no qual os gametas (óvulos e espermatozoides) são manipulados e unidos com o objetivo de formar embriões. Atualmente temos disponíveis dois tipos de FIV: a clássica e a ICSI (injeção intracitoplasmática de espermatozoides) [46].

A FIV clássica inicia-se com injeções subcutâneas de gonadotrofinas urinárias ou recombinantes, que consistem em hormônios folículo estimulante (FSH) e luteinizante (LH). O uso de agonistas ou antagonistas do hormônio liberador das gonadotrofinas (gnRH) têm o objetivo de impedir a ovulação

prematura durante o ciclo de indução. A monitorização dos folículos ovarianos ocorre por meio da ultrassonografia (US) transvaginal ou com dosagens de estradiol e progesterona, essa etapa possui duração média de 10 a 12 dias. Ao alcançarem tamanho entre 17 a 20 mm, o amadurecimento final dos oócitos é desencadeado com uma injeção de gonadotrofinas coriônica humana (hCG) que mimetiza o pico natural de LH pré-ovulatório. Durante o período de uso das medicações, é realizado então o acompanhamento com US vaginal para avaliação da resposta às injeções e definição do dia da coleta dos óvulos. Dado o dia da coleta dos óvulos, o procedimento é realizado por meio de uma agulha acoplada a uma sonda de US transvaginal. Os óvulos obtidos são levados ao laboratório, os quais são colocados em contato com os espermatozoides em placa de meio de cultura. Os embriões obtidos são cultivados entre 2 a 6 dias e são transferidos no estágio de duas a oito células, ou depois de 5 a 6 dias, no estágio de blastocistos. Os embriões são transferidos para o útero com o auxílio de um cateter de material inerte e atóxico, guiado por US, e após aproximadamente 14 dias, o teste de gravidez é realizado [47, 48, 49, 50].

Uma variação da técnica de FIV é a ICSI, indicada inicialmente para tratamento de casais com infertilidade masculina grave. Nesta técnica, um espermatozoide é injetado em cada óvulo no laboratório com o auxílio de um microscópio acoplado a um micromanipulador. Em situações em que não existem espermatozoides no ejaculado, pode-se obter por meio de punção do epidídimo ou no testículo (TESA). Em casos mais severos é necessária uma biópsia testicular (TESE) ou uma microdissecação testicular (micro – TESE). Através da ICSI é possível coletar um único espermatozoide e colocá-lo diretamente dentro do óvulo [51, 52].

A falha na fertilização ocorre em 3,52% a 20% dos ciclos de FIV e 1% a 3% dos ciclos ICSI. Os problemas gerados em torno da falha na fertilização, ocasionou maior utilização da ICSI em casos de infertilidade em fator não masculino, em especial em ciclos que um pequeno número de óocitos é recuperado, infertilidade inexplicável, idade materna avançada, falha prévia da fertilização clássica, diagnóstico genético pré-implantação (PGD), fertilização após maturação in vitro e fertilização de oócitos congelados e descongelados [53, 54].

Um estudo realizado pela ASRM e SART, demonstra que a utilização da ICSI é recomendada e eficaz para pacientes com infertilidade masculina e pode beneficiar pacientes submetidos à fertilização in vitro com PGD, em casos de falha de fertilização em tentativas anteriores de FIV clássica, óocitos amadurecidos in vitro e óocitos descongelados. No entanto a utilização da ICSI não pode melhorar os resultados clínicos dos pacientes com infertilidade inexplicada, baixo rendimento de ovócitos ou idade avançada [53, 54].

2.8 Machine Learning

2.8.1 O que é *Machine Learning*

Machine Learning (ML) é uma subárea da Inteligência Artificial (IA) que se dedica ao estudo científico de algoritmos e modelos estatísticos utilizados por computadores para executar uma determinada tarefa sem a necessidade do uso de instruções explícitas, baseando-se principalmente em padrões e inferências [55].

Os algoritmos de ML podem ser classificados, de maneira geral, em três tipos básicos:

- **Aprendizado supervisionado:** o algoritmo de aprendizado desenvolve um modelo matemático a partir de um conjunto de dados que contém os *inputs* e *outputs* desejados;
- **Aprendizado não-supervisionado:** o algoritmo desenvolve um modelo matemático a partir de um conjunto de dados que contém somente os *inputs*, mas nenhum *output*;
- **Aprendizado semi-supervisionado:** neste caso, o algoritmo desenvolve um modelo a partir de dados de treinamento que estão incompletos, nos quais parte dos dados possuem somente os *inputs* e outra parte possui tanto *inputs* quanto *outputs* [55];

A Figura 3 abaixo ilustra as três tarefas básicas utilizadas pelo algoritmo de ML:

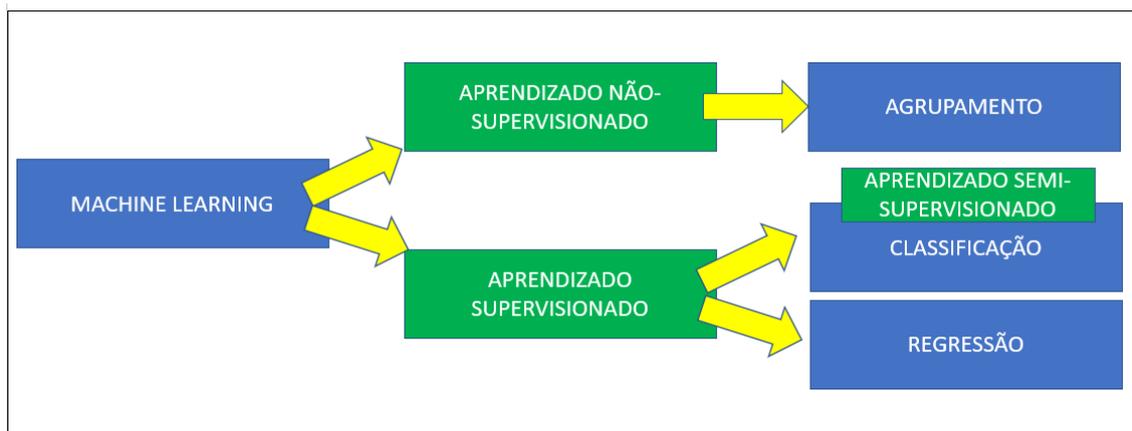


Figura 3. Tipos de abordagem pelos algoritmos de Machine Learning (Adaptado de BABITHA, 2019 [55])

2.8.2 Quando utilizar o *Machine Learning*

Os algoritmos de ML têm sido cada vez mais aplicados na área de saúde para auxiliar na previsão de desfechos clínicos. Segundo Brynjolfsson e Mitchell (2017) [56], as técnicas de *Machine Learning* (ML) apresentam maior probabilidade de êxito se utilizadas, de modo geral, em problemas que possuem as seguintes características:

- **Dados de inputs e outputs bem definidos** – os mais recentes trabalhos utilizam o ML como um mecanismo de classificação (por exemplo, identificar os registros médicos de acordo com a probabilidade de testarem positivo para câncer) e para efetuar previsões (por exemplo, a utilização de redes neurais para prever o risco de retorno de pacientes já tratados aos hospitais [57]);
- **conjunto de dados disponíveis relativamente grande** – quanto maior a disponibilidade de exemplos de treinamento acessíveis ao algoritmo, maior será a taxa de precisão do aprendizado;
- **feedbacks claros com objetivos concisos e métricas bem definidas** – O ML funciona melhor em situações em que é possível descrever claramente os objetivos, mesmo quando não se pode definir o melhor processo para alcançá-lo. Além disso, ter métricas claras como parâmetros de desempenho fornece um padrão-ouro para o sistema, uma

vez que o ML funciona melhor quando aplicado à este tipo de base de dados;

- **não envolve longas cadeias de lógica ou raciocínios que dependem de conhecimentos baseados em senso comum** - os sistemas de ML são muito efetivos para o aprendizado de associações empíricas dos dados, mas são menos eficazes quando a tarefa exige longas cadeias de raciocínio ou planejamento complexo que dependem do bom senso ou conhecimentos aprofundados que sejam desconhecidos pelo computador;
- **não há necessidade de explicação detalhada de como a decisão foi tomada** – de modo geral, os algoritmos de ML aprendem a tomar decisões ajustando sutilmente até centenas de milhões de dados numéricos que interconectam seus neurônios artificiais. Explicar o raciocínio de tais decisões para os seres humanos pode ser difícil porque as *deep neural networks* (DNN) geralmente não fazem uso das mesmas formas de raciocínios que os humanos;
- **existe uma tolerância a erros e não necessita da geração de soluções ideais** – a maioria dos algoritmos de ML geram suas soluções baseadas em técnicas estatísticas e probabilísticas. Como resultado, raramente é possível treiná-los com 100% de precisão. Até os melhores sistemas computacionais de fala, reconhecimento de objetos e diagnóstico clínico cometem erros (assim como os melhores seres humanos). Portanto, a tolerância a erros do sistema aprendido é um critério importante que restringe a sua aplicação;
- **a função de aprendizagem não muda rapidamente ao longo do tempo** - em geral, os algoritmos de ML funcionam bem apenas quando o conjunto de dados utilizado para efetuar as decisões futuras são semelhantes aos dados utilizados no treinamento. Se houver alterações ao longo do tempo, um novo treinamento será necessário e a acurácia do sistema dependerá se a velocidade de aquisição dos novos dados de treinamento será igual ou maior que a taxa de mudança dos dados; e
- **quando nenhuma destreza, habilidade física ou mobilidade especial é necessária** – as máquinas e robôs ainda são, de modo geral, bastante

limitados em comparação com os humanos ao lidar com deslocamentos ou tarefas que exijam destreza física. Tal fato pode comprometer a efetividade da solução do problema.

Esses critérios são diretrizes valiosas para a definição de quais métodos de ML são os mais adequados para serem aplicados para a realização de uma determinada tarefa [58].

2.8.3 Técnicas de Classificação de *Machine Learning*

A seguir, tem-se a apresentação de algumas técnicas de classificação que são baseadas nos conceitos do *Machine Learning*.

- Os *Naive Bayes Classifiers* fazem parte de uma família de classificadores probabilísticos simples, baseados na aplicação do teorema de Bayes a partir do princípio da independência entre os recursos. O Teorema de Bayes fundamenta-se na probabilidade de um evento ocorrer levando-se em conta a probabilidade de um outro evento que já ocorreu, sendo representado matematicamente da seguinte maneira:

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)} \quad (1)$$

Equação 1. Teorema de Bayes. Probabilidade de um evento ocorrer

Onde A e B são eventos e $P(B) \neq 0$, $P(A)$ é a probabilidade anterior, $P(A|B)$ é uma probabilidade posterior [55].

- A Árvore de Decisão é uma técnica de classificação simples e muito utilizada no meio científico. Ela consiste em um classificador que apresenta uma série de perguntas elaboradas a partir dos atributos provenientes dos dados coletados. Toda vez que o sistema recebe uma resposta, uma pergunta de acompanhamento é realizada para que então seja entregue uma conclusão [55].

- O *Random Forest* constitui-se em um método de aprendizado, baseado em classificação, cujo objetivo é construir diversas árvores de decisão durante o processo de treinamento. A partir disso, utiliza-se as saídas médias das diversas árvores de decisão para decidir qual a classe melhor se adequa ao problema proposto [55].
- Na técnica denominada *Support Vector Machines* (SVM), os algoritmos de ML utilizados para a classificação do sistema são do tipo aprendizado supervisionado. Esta é uma metodologia de classificação binária que elenca as instâncias com base na probabilidade de elas terem um rótulo positivo, sendo baseada na curva de ROC e objetiva maximizar a área sob essa curva. Dessa forma, o algoritmo aprende uma função de classificação, sendo esta uma combinação linear de funções de pontuação não-lineares construídas para cada recurso separadamente [55].
- Já as redes neurais artificiais (RNA) baseiam-se na formação de nós interconectados denominados neurônios artificiais, criados a partir de um modelo simplificado dos neurônios cerebrais. Cada um desses neurônios contribui individualmente para o aprendizado do sistema e possuem um peso que se ajusta à medida que o aprendizado avança. Vale destacar a crescente utilização desta técnica nos últimos anos em diversos campos que vão muito além das áreas biológicas ou médicas, enveredando-se para engenharia, ciências sociais, robótica, *marketing* e negócios. A RNA possui o potencial de substituir os métodos tradicionais de previsão estatística, por deter uma taxa de precisão bastante superior [55].

2.8.4 Trabalhos que utilizaram *Machine Learning* aplicados na FIV

O *Machine Learning*, pelo seu poder de predição, pode ser uma ferramenta importante para auxiliar profissionais da área de reprodução assistida

na tomada de decisão. Alguns destes classificadores foram expandidos para a análise do sucesso do tratamento da FIV, a partir de modelos que incluem variáveis relacionadas ao desfecho da gestação, conforme pode-se observar em trabalhos publicados anteriormente.

Guh et al. (2011) [59], desenvolveram um método de inteligência híbrida que integra técnicas de árvore de decisão e *Data Mining* (DM), na mineração de dados médicos de FIV com o intuito de identificar os atributos mais significativos para se determinar a taxa de gravidez. Foi utilizado um banco de dados com 5275 registros de tratamento, coletados no centro de reprodução artificial de um hospital universitário de Los Angeles, EUA, de dezembro de 1999 a julho de 2007. A faixa etária dos pacientes maternos é entre 17 e 57 anos, com idade média de 39 anos. As pacientes com idade igual ou superior a 40 constituíram 42,6% (2250 pacientes) desse conjunto de dados.

Dessa forma, foram selecionados 28 atributos classificados como principais, por exemplo, idade da paciente, número de embriões transferidos, número de embriões congelados e dias de cultura do embrião. Como resultado, o modelo teve uma precisão de 73% e sensibilidade e especificidade de 71% e 73%, respectivamente. Os autores concluíram que tal algoritmo é capaz de auxiliar os médicos a identificarem a probabilidade de sucesso do tratamento, porém não os ajuda na escolha dos embriões mais aptos para o sucesso da gestação. Além disso, os resultados empíricos mostram que o método de DM proposto também pode ser aplicado na geração de parâmetros de auxílio aos profissionais, na adaptação do tratamento de FIV de acordo com as características de cada paciente, com o objetivo de melhorar a taxa de gravidez [59].

Blank et al (2019) [7] propôs um modelo de *random forest* para prever o potencial de implantação de um embrião transferido e compará-lo com um modelo de regressão logística multivariada (MvLRM), em um estudo retrospectivo de dois anos, baseado em 1052 mulheres, submetidas à transferência de embrião único em estágio de blastocisto no tratamento de ICSI do Hospital Universitário de Ghent, situado na Bélgica. Nesta análise as variáveis de maior importância foram: idade da mulher, nível de hormônio antimülleriano, paridade e gravidade. A análise da curva ROC resultou em uma acurácia de 0,74 +/- 0,03 para o *Random forest* e 0,66 +/- 0,05 para o MvLRM na previsão de

gestações em andamento de 11 semanas. Os resultados de sensibilidade para MvLRM e Random forest foram respectivamente de 0,66 +/- 0,08 e 0,84 +/- 0,07 e a especificidade de 0,58 +/- 0,08 e 0,48 +/- 0,07. Este estudo demonstra que a utilização do *Random forest* pode melhorar o desempenho para prever a implantação em andamento em relação ao MvLRM.

Durairaj e Thamilselvan (2013) [13], utilizaram a Rede Neural Artificial (RNA) para apresentar a taxa de sucesso da FIV aplicando 27 atributos, incluindo idade da mulher, duração da infertilidade, índice de massa corporal (IMC), cirurgia anterior, gravidez anterior, endometriose, fator tubário, ovulatório, concentração de espermatozoides, vitalidade espermática, número de oócitos recuperados, número de embriões transferidos e fatores psicológicos. Foi estudada uma população de 250 pacientes de origem indiana de centros de pesquisas de FIV, clínicas de fertilidade e maternidades, no qual, os dados foram treinados, validados cruzadamente, testados e então apresentados como taxa de sucesso da FIV obtendo 79% de precisão.

Em Güvenir et. al (2015) [60], foi utilizado o algoritmo RIMARC (algoritmo de ML, maximizando a área sob a curva ROC) para classificar os atributos e os seus valores que afetam o resultado do tratamento de FIV. Neste trabalho, os autores apresentaram uma nova técnica, chamada estimativa de sucesso, usando um algoritmo de classificação (SERA), baseado em SVM e compararam os resultados obtidos com os algoritmos *Naïve Bayes Classifier* e *Random Forest*. Os parâmetros de comparação foram área sob a curva ROC (AUC) e precisão.

Um conjunto de dados de 1456 pacientes obtidos através da unidade de fertilização in vitro no Hospital de Ensino, Pesquisa e Saúde da Mulher Etlik Zubeyde Hanim, localizado em Ancara, na Turquia. Ao todo foram consideradas 64 características, sendo 52 delas relacionadas à mulher e 12 relacionadas ao homem. Os fatores mais relevantes foram a idade da mulher, a causa da infertilidade, hidrossalpinge, miomas uterinos, tabagismo e obesidade. O algoritmo RIMARC calculou os pesos dos atributos e criou regras de simples interpretação pelos médicos [60].

Dessa forma, tal algoritmo superou outros classificadores usados no estudo em termos de área sob curva de ROC e em precisão, conforme a Tabela 1 a seguir:

Tabela 1. Resultados da AUC e precisão no conjunto de dados de fertilização in vitro.

Algoritmo	Área sobre a Curva ROC (AUC) (%)	Precisão (%)
SERA	0.833 (± 0.003)	0.844 (± 0.004)
NBC	0.794 (± 0.002)	0.783 (± 0.002)
Random forest	0.769 (± 0.009)	0.792 (± 0.008)

Adaptado de Güvenir et. al (2015) [60]

Dessa forma, os autores concluíram que os resultados indicam que o algoritmo SERA proposto tem potencial para ser utilizado com sucesso para estimar a probabilidade de sucesso no tratamento médico. O trabalho mostrou que é possível fazer com que um modelo de ML aprenda, a partir de um conjunto de casos passados, a estimar o resultado do tratamento de FIV. Além disso, o algoritmo RIMARC, usado pela SERA, possui três características importantes para aplicações médicas. Em primeiro lugar, aprende regras sobre os dados, que podem ser analisadas posteriormente por médicos. Em segundo lugar, ele não possui parâmetros que precisam ser otimizados após a adição de novos registros de pacientes. E por fim, demonstrou que a falta de dados, o que é comum em conjuntos de dados médicos é relevante para o resultado do modelo. Além disso, o algoritmo RIMARC calcula os pesos dos atributos e cria regras que são legíveis por humanos e fáceis de interpretar pelos médicos. Essa característica do RIMARC permite que os médicos validem o modelo construído.

Hafiz et al (2017), [6], realizou um estudo comparativo entre os classificadores Support Vector Machines (SVM), Recursive Partitioning (RPART), Random Forest (RF), Adaptive Boosting e One-nearest neighbor (1NN) e também teve como objetivo encontrar os fatores mais efetivos para predição do sucesso da FIV em casais inférteis. Foram utilizados os dados de 486 pacientes, coletados do Centro de Pesquisa em Infertilidade do Hospital Mãe e Filho em Shiraz, Irã, de 2009 a 2015, sendo 131 casos positivos para gravidez e 355 negativos. A média de idade das mulheres foi entre 18 a 47 anos e dos homens de 23 a 70 anos.

Os valores comparativos entre os resultados obtidos em relação à Área sob a curva ROC (AUC) e precisão estão descritos na Tabela 2 abaixo:

Tabela 2. Resultados da AUC e precisão no conjunto de dados de fertilização in vitro.

Algoritmo	Área sobre a Curva ROC (AUC) (%)	Precisão (%)
SVM	57.57 ± 1.51	68.3 ± 1.05
AB	47.52 ± 4.5	66.99 ± 2.85
RPART	82.05 ± 2.34	83.56 ± 0.99
RF	84.23 ± 0.91	83.96 ± 0.62
1NN	50 ± 0	64.84 ± 1.46

Adaptado de Hafiz et. al (2017) [6]

Já os valores comparativos entre os resultados de sensibilidade e especificidade estão descritos na Tabela 3 abaixo:

Tabela 3. Resultados da sensibilidade e especificidade no conjunto de dados de fertilização in vitro

Algoritmo	Sensibilidade (%)	Especificidade (%)
SVM	48.85	98.03
Adaboost	59.54	91.83
RPART	54.96	70.42
RF	14.5	88.73
1NN	35.88	73.52

. Adaptado de Hafiz et. al (2017) [6]

Como conclusão, tem-se que, de acordo com os resultados obtidos no presente estudo comparativo, os classificadores de RF e RPART superaram os outros métodos de previsão de gravidez com AUC de 84,23% e 82,05%, respectivamente. Além da questão dos classificadores, concluiu-se que a idade da mulher, o número de embriões desenvolvidos e o nível sérico de estradiol (E2) no dia da administração de gonadotrofina coriônica humana (hCG) foram as variáveis mais importantes para a predição do sucesso da FIV.

3. OBJETIVOS

3.1 Objetivo geral

Elaborar um modelo de predição do sucesso de gestação após procedimento de fertilização *in vitro* (FIV), por meio de algoritmos de *Machine Learning* (ML),

utilizando variáveis de um banco de dados de um serviço de atendimento público de saúde.

3.2 Objetivos específicos

- Selecionar variáveis que possam estar envolvidas no sucesso da gestação após fertilização *in vitro* a partir de um banco de dados de um serviço de atendimento público de saúde;
- Avaliar diferenças pré-implantacionais entre o grupo de mulheres que tiveram sucesso e aquelas que apresentaram insucesso na fertilização.
- Aplicar modelos algorítmicos de ML para avaliar a capacidade de predição dessas variáveis para a gestação.
- Selecionar o modelo de melhor acurácia e validá-lo na população teste, formada como subgrupo da população total.

4. CAPÍTULO 1

***PREDICTION OF SUCCESS OF PREGNANCY AFTER IN VITRO
FERTILIZATION PERFORMED BY A PUBLIC HEALTHY SERVICE USING
MACHINE LEARNING ALGORITMS***

ABSTRACT

Introduction: Infertility has become a global health problem, increasing the number of couples looking for *in vitro* fertilization (IVF). Despite advances and technical improvements, some couples remain childless due to the high

complexity of the technique. The use of Machine Learning (ML) in the prediction of pregnancy, computing factors that could interfere in the effectiveness of the treatment, is an important tool to optimize these factors and reach the success of pregnancy. **Objective:** The aim of this study was to apply ML models to predict the success of IVF in a public health service, including pre-implantation variables. **Material and methods:** This study included 771 women who underwent IVF treatment at Hospital das Clínicas, Federal University of Minas Gerais, between 2013 and 2019. We used the following algorithms based on Machine Learning: Logistic Regression, Random Forest, XG Boost and Support Vector Machines (SVM). **Results:** The Random Forest algorithm obtained the best performance, with better accuracy, sensitivity and area under the ROC curve, showing 20 variables that are most relevant to predict the success of pregnancy after IVF. **Conclusions:** Our ML algorithm can be useful to predict pregnancy in women submitted to IVF, as well as to define the variables on which the clinical intervention may improve the success of the treatment.

Keywords: Infertility, FIV / ICSI, Prediction model, Machine Learning, Random Forest.

1. INTRODUCTION

Infertility, defined as failure to conceive after 12 months of unprotected intercourse, is a condition that affects between 8% and 12% of couples in reproductive age worldwide [1,2]. Thus, one in six couples will seek help from a specialist at some point in their lives due to difficulties in conceiving [3]. Infertility is caused by several factors such as environmental, genetic or personal characteristics [4].

In vitro fertilization (IVF) is commonly used as a treatment for infertility. The first human birth resulting from this technique occurred in 1978, being considered at that time a complex reproductive procedure [5] that involves the processes of ovarian stimulation, oocyte retrieval, fertilization, embryo culture and embryo transfer to the uterus [6]. Since then, more than 6 million children have been born through IVF [7]. However, even with technical advances and improvements, effectiveness is considered relatively low [8], with an average of only one third of the cycles of IVF resulting in pregnancy [9]. In addition, IVF is inaccessible for the most infertile couples due to the scarcity of specialized clinics and the high cost around the world, with few public services that perform this type of procedure [10].

The success of IVF involves the choice of the best embryo for transfer to the uterus, which is selected based on morphological characteristics through visual assessment, however this method is not accurate enough to predict pregnancy or live birth [8]. Therefore, the use of resources that permit the prediction of pregnancy results from an IVF treatment is very important for both patients and physicians [5]. Although there are processes to improve the success rate in the treatment, it is often necessary to adopt more advanced methods of studies, such as automated tools based on Artificial Intelligence, in order to make possible the global analysis of the factors that interfere in the effectiveness of treatments [4].

Machine Learning (ML) is a sub-area of Artificial Intelligence, which allows computers and machines to undergo self-learning models, without the use of explicit programming [11] by presenting computer programs to a set of data [9].

In general terms, the basic approach of all ML is to provide machines with the ability to learn, using previous data and pattern of recognition, and to develop highly accurate classification algorithms, which deliver responses with a high degree of confidence. ML-based systems aim to make predictions in most cases, which can provide valuable information during the diagnosis of diseases or during treatment [9, 12, 13].

In recent years, ML is rapidly gaining adhesion in the human reproduction and embryology sectors. In the year 2018, the number of studies addressing the topic grew more than seven times in the annals of the American Society for Reproductive Medicine and European Society for Human Reproduction and Embryology [9]. Despite recent studies showing that ML techniques are promising in guiding the IVF treatment plan, its performance needs to be further refined in order to be used in the medical routine [5].

The variables involved in the success of pregnancy may vary according to the type of service, public or private, since different demographic and clinical characteristics of the patients may directly interfere with the results. Thus, the objective of this study was to apply ML models to predict the success of IVF in a public health service, including pre-implantation variables.

2. MATERIALS AND METHODS

2.1. Population characteristics

This study was conducted including the clinical data of women treated at a public teaching hospital - Hospital das Clínicas, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, between 2013 and 2019. Patients had their medical records analyzed and cases without outcome information (pregnancy or non-pregnancy) were excluded. Finally, 771 cases and variables related to the couple or the IVF procedure were included. All personal identification data were removed and replaced with unique identification codes. The outcome used to classify the groups was positive or negative serum β HCG test after IVF process [14].

This study was approved by the Ethics Committee of the Federal University of Minas Gerais (COEP) under number 3.613.334, following the recommendations of the Declaration of Helsinki.

2.2. Data analysis

Four supervised learning algorithms were used, based on Machine Learning, accomplished of handling problem classification: Logistic Regression, Random Forest, XGBoost and Support Vector Machines (SVM).

Logistic Regression and SVM use binary classification methods that list the variables based on their probability to present a high degree of significance for the model. While logistic regression uses mainly statistical approaches, SVM is based on the geometric properties of the data. In the latter, through a kernel function, the data is mapped to a multidimensional space. From this new representation of the data, the objective is to seek a hyperplane that better divides the classes [5, 15].

The XGBoost and Random Forest algorithms consist of ensemble models (combination of models) that aim to generate prediction models through a set of data interconnected by several links, known as decision trees. The combination of the average results of these links determines which class of results better matches the proposed problem. The main difference between the two techniques consists of the way that the trees are built. Random Forest trains each tree independently, based on random sampling of the data attributes. XGBoost builds the trees sequentially, in which each new tree goals to correct the errors of the previous one and improve the accuracy of the results [14].

Due to the high complexity related to the pregnancy process, IVF treatment is influenced by several characteristics. In this study, the original database consisted of a total of 90 variables (**Supplementary material 1**). However, throughout the model validation process, it was observed that not all of them affected the prediction model. Therefore, only factors whose effects were shown to be relevant to the system were considered. Thus, the model was reduced to only 32 variables, which are clearly defined, standardized and capable of being replicated (**Table 2**).

The normality of the variables was verified by the Shapiro-Wilk test. Categorical variables were described as absolute value and percentage, while continuous variables were expressed as median and interquartile range. Differences in the frequency of variables between women who became pregnant and those who did not were assessed by the chi-square test, or the median difference was assessed by the Mann-Whitney test. Statistical analyses were conducted using SPSS version 21 software.

The machine learning and plotting algorithms were performed using Python software version 3.6.9. Python Sklearn and XGBoost packages were used. Categorical features were coded using an one-hot coding scheme. After processing the variables, 75% of the data set was selected randomly as a training set for establishing the forecasting model, and the remaining 25% was used for testing. A stratified random sampling method was employed to ensure that the proportions of women who became pregnant and those who did not were the same in the training, validation and test sets, as in the original data set. To balance the classes, the SMOTE super-sampling technique was used for training data.

The grid search with cross-validation k-fold ($k = 5$) was used to find the ideal hyperparameters of the four classifiers mentioned above. The training set was divided into k folds. Each time, $k - 1$ folds were used for training and the remaining for validation, as shown in **Figure 1**. For each algorithm, the models with different hyperparameters were scored for their average precision and accuracy. We chose the set of hyperparameters that maximized the area under the ROC curve (AUC) and adjusted the model to the entire set of training data. To estimate the performance of each machine learning classifier, we evaluated the accuracy, sensitivity, specificity and AUC. The set of varied hyperparameters for each algorithm is shown in **Table 1**.

3. RESULTS

3.1. Characteristics of the population

The data of 1049 women were collected and reviewed according to the information of the service's database. From these, 144 patients without information about pregnancy, 55 patients in which no embryo transfer occurred,

09 recipients of oocytes or embryos from third-party donors, and 70 patients with missing data were excluded. Therefore, a total of 771 patients were included in this study. There were a total of 276 (36%) patients with positive β HCG and 495 (64%) with negative β HCG tests.

The age of men (the partner) ranged from 21 to 60 years, the age of women from 22 to 43 years and the body mass index (BMI) of women ranged from 4.81 to 33.5 kg/m². The antral follicle count ranged from 0 to 47 follicles. The total dose of gonadotropins was defined as the amount of ampoules administered to the patient, namely, Gonal F[®], Menopur[®] or Pergoveris[®], and the total dose ranged from 900 to 5400 IU. Preovulatory follicles (over 17 mm in diameter) ranged from 0 to 31. The total number of oocytes retrieved ranged from 1 to 42, the number of metaphase II oocytes ranged from 0 to 37. The number of oocytes inseminated by conventional IVF or ICSI ranged from 0 to 28, and the total number of zygotes ranged from 0 to 31. The total number of cleaved embryos varied from 0 to 28 and blastocysts from 0 to 15. The total number of embryos transferred ranged from 0 to 5, with morphological classification spanning all categories [6; 16; 17]. The number of blastocysts transferred ranged from 0 to 3.

The indications for IVF treatment included endometriosis, male factor, ovarian factor, tuboperitoneal factor, unexplained infertility or other factors (seropositive couple, infectious diseases, fibroids, recurrent pregnancy loss, independent production, unicorn uterus - indications that were grouped due to the low frequency). The sperm source was predominantly the ejaculate from the male partner, but also included donor sperm, testicular sperm extraction (TESE), microscopic epididymal sperm aspiration (MESA), microsurgical TESE, and percutaneous epididymal sperm aspiration (PESA). The semen was also classified into fresh or cryopreserved, and the semen preparation techniques included density gradient, seminal wash, direct swim-up or traditional swim-up. Gonadotropin types used for ovarian stimulation were recombinant, urinary or a combination of both. The stages of embryo transfer were cleavage (day 3) or blastocyst. The catheter models used for embryo transfer were CCD, Sydney, Transfer slow and Wallace.

The group of non-pregnant women was older ($p = 0.015$), while higher man age was observed for the group of pregnant women ($p = 0.049$, **Table 3**). Regarding the cause of infertility, a significant difference was observed between the two groups ($p = 0.016$). Endometriosis was the most frequent factor in the non-pregnant women group, while unexplained infertility and other factors were the more frequent in the pregnant group. The variables sperm source, semen status, as well as IVF or ICSI did not differ between groups in this univariate analysis ($p = 0.063$, 0.637 and 0.075 , respectively).

3.2. Machine learning models

The results of accuracy, sensitivity, specificity, and area under the ROC curve (AUC ROC) for each model are listed in **Table 4**. The four models were similar, but the logistic regression and the random forest had slightly better values for the parameters sensitivity, accuracy, specificity, and AUC ROC curve than the other models. The XGBoost and logistic regression models had better results only for specificity. Thus, the random forest model was chosen because it allows the discrimination of the most important variables for the model.

3.3 Random Forest model application

The Random Forest model presented an AUC ROC curve for the prediction of pregnancy of 0.64 (**Figure 2**). Taking a sample of 25% of cases ($n = 193$) to validate the model, it was able to correctly classify 76 negative cases and 41 positive cases, but 28 positive cases were classified as negative, and 48 negative cases were classified as positive by the model (**Figure 3**).

The importance of each variable within the model was estimated using the Shapley additive explanations (SHAP) algorithm [18]. For each patient, the effect of the absence of the variable was calculated using the distance from the original forecast. A negative distance is obtained if the patient has a value that reduces the risk of pregnancy compared to the average risk for the population studied, and a positive distance is obtained when this risk in favor of pregnancy is greater than the average for the population evaluated. An average of absolute

distances between the variables is used to generate a ranking list according to the effect on the predictions of all the results together.

Following a decreasing order of importance, the SHAP algorithm showed that the variables age of the woman, embryo classification A, total number of viable cleavage stage embryos, number of blastocysts, total fertilized pro nuclei, antral follicle count, sperm source (ejaculated sperm more relevant in non-pregnancy group), number of follicles greater than 17 mm, number of MII oocytes, indication for IVF treatment (endometriosis factor was the most relevant for non-pregnancy, but tuboperitoneal and others factors were the most important for pregnancy – **Figure 4A**), number of type A embryos transferred, number of oocytes retrieved, total dose of gonadotropins, number of type B embryos transferred, age of man, embryos with the classification B and C, number of embryos transferred, number of IVF oocytes fertilized and catheter model user for embryo transfer (the CCD model was the most relevant for non-pregnancy – **Figure 4B**) were the variables considered most important for the model among the 32 included. The result of the main variables of the model is shown in the SHAP, regarding the positive or negative β HCG (**Figure 5**).

As example of an individual application of the model, **Figure 6** shows the result of a positive case predicted positive by the model (true positive) and a negative case predicted negative by the model (true negative). Considering the probability of the prediction and the decision threshold = 0.5, an individual probability above 0.5 value is predicted as positive for pregnancy, and below 0.5 it is predicted as negative for pregnancy.

According to **Figure 6A**, the output value predicts a positive β hCG result. The variables that had the higher influence to increase the probability in this individual case: embryos with classification A =7, number of blastocysts = 6, age of the woman <35 years, antral follicle count = 18, number of viable cleavage stage embryos = 14, age of man = 33 years, number of oocytes retrieved = 19, and embryos with the classification B = 5. The variables that contributed to decrease the probability in this analysis were: number of type A embryos transferred = 0, ejaculated semen, preparation of the direct swim-up semen. The sum of the weight of each variable determined a final value of 0.64 (> 0.5,

therefore predicted positive for β HCG). In fact, this result applied to a real positive case.

Figure 6B shows the prediction of a negative β hCG result. The only variable that contributed to increase the probability of exit was the follicle count greater than 17mm = 3. The variables that determined the lowest probability value were fertilized cells = 1, male age 41 or 42 years, CCD catheter, number of embryos with classification A = 0, unexplained infertility as an indication for IVF treatment, number of MII oocyte cells = 2 and blastocysts = 0. The final probability was 0.21 (<0.5 , therefore predicted negative for β hCG). This result was in agreement with the case, which was therefore a true negative.

4. DISCUSSION

In this study, four Machine Learning models were developed in order to assess the ability to predict pregnancy success after performing IVF, using pre-implantation variables. The Random Forest model obtained the best performance, regarding the accuracy, sensitivity and area under the ROC curve. This model was able to select 20 variables that showed to be directly related to the success of IVF. The variables that were most relevant to the model were the following: age of the woman, number of embryos with classification A, total number of viable cleavage stage embryos, number of blastocysts, total of fertilized pro-nuclei, antral follicle count, sperm source, number of follicles greater than 17 mm, number of MII oocyte, indication for IVF, number of type A embryos transferred, number of oocytes, total dose of gonadotropins, number of type B embryos transferred, age of man, number of embryos with the classification B and C, number of embryos transferred, number of IVF oocytes and catheter model used for embryo transfer.

Random Forest contains a combination of characteristics that make them suitable for application in complex result prediction models. One of its main advantages is the ability to detect the nonlinear association of patterns between the predictors and the output, which becomes important when dealing with complex nonlinear events. In addition, it does not require specifying an underlying data model and it is capable of dealing with highly correlated predictor variables,

determining which variables are important in the regression model. Thus, the Random Forest models work efficiently on large data sets and can be more easily parallelized, showing robustness even in presence of outliers and noise. Finally, it offers the flexibility to perform several types of statistical data analysis and it is computationally lighter than other tree based methods [19].

We observed that women who did not progress to pregnancy had a higher age, represented by a wider interquartile range, when compared to those who became pregnant. In the model studied, this factor proved to be the most important for predicting the chances of pregnancy. The age-related decline in female fertility is due in part meiotic errors in the gametogenesis and in part to the exhaustion of the ovarian reserve. This occurs due to increased follicular apoptosis or as consequence of external factors that can reduce the reserve or primordial follicles, such as smoking, exposure to radiation and oxidizing agents, ovarian endometrioma, other benign ovarian tumors and oncological diseases [20]. Such conditions corroborate the results of the model, showing an inverse relationship between the increase in the woman's age and the number of MII oocytes, as well as the chances of successful pregnancy.

Man's aging is also an important factor, often responsible for the decline in semen quality, which involves sperm concentration, motility and morphology [2; 20]. Our unexpected results showed that the median age of men was higher in the group who achieved pregnancy, a data corroborated by the SHAP algorithm. In the United States, the birth rate in the past three decades has decreased for men under the age of 30 and increased for men aged 30 or over [21]. The birth rate for fathers aged 25 to 29 years declined by 27%, and in contrast, the birth rate for fathers aged 30 to 34 years, 35 to 39, 40 to 44 and 45 to 49 increased respectively 15%, 61%, 63 % and 52%. In addition, Rochebrochard et al. (2003) [22] showed that at a cutoff point for the paternal age of 40, couples with one of them up to 34 years old showed no significant effect of the paternal age on not conceiving in 12 months. However, among couples with women aged 35 to 39, the risk of late conception was significantly higher when the father's age was over 40 than.

In the present study, no difference was observed regarding BMI between the two outcome groups. This result corroborates the study by Matalliotakis et al.

(2008) [23], indicating that no significant difference was found in the rates of clinical pregnancy and abortion among women with a BMI above 24 kg/m². Their results indicate a lower stimulation response in women with a high BMI, but no adverse effect on IVF outcome.

We observed that women with endometriosis were more frequent among those who did not become pregnant, while those with other infertility factors were more frequent in the group with success in IVF. The SHAP model also indicates that the endometriosis factor contributed to the group that did not become pregnant, and the tuboperitoneal factor showed higher contribution to successful of pregnancy, followed by other factors. Endometriosis is a major cause of infertility, but it is unknown to what extent this disease affects IVF results. Cohort studies in Italy [24] Latin America [25] and India [26] did not find any significant decrease in IVF pregnancy rates of women with endometriosis compared to women without endometriosis. One study suggested that IVF results are worse in the presence of deep endometriosis associated with ovarian endometrioma compared to endometrioma alone [27]. Importantly, a history of surgery for endometriosis may not only harm ovarian reserve [28] but also lower IVF pregnancy rates [29]. Our database does not allow to stratify the participants by the type of endometriosis treatment, but all cases of endometriosis had been diagnosed by surgery, suggesting that a history surgery for endometriosis might be the true factor behind their lower likelihood of achieving pregnancy [29].

Among the different sperm source, the sperm obtained by ejaculation was more frequent in the group of non-pregnant women, and associated to non-pregnancy in SHAP model. In other words, the cases in which the sperm was obtained by surgical methods or from third-party donors, due to severe male factor and absence of spermatozoa in the ejaculate (azoospermia) had a pregnancy rate higher than the general study population, formed by couples with several causes of infertility. Compared with tubal factor or unexplained infertility, male factor infertility has been associated with higher live birth rates after IVF [30].

Hafiz et al. (2017) [6], Zhao et al. (2018) [16], Lai et al. (2020) [17] showed the importance of using embryos classified according to their morphological characteristics. According to Zhao et al. (2018) [16] and Lai et al. (2020) [17], embryos initially received a numerical classification based on their degree of

expansion and incubation status. In such studies, the use of embryos with A and B classification obtained higher rates of pregnancy success when compared to cases in which type C embryos were used during the IVF process. This aspect corroborates with the model obtained, since a greater number of embryos classified as A and B, reported by SHAP, converged to the positive β HCG, in contrast to the higher number of embryos of classification C. This suggests that the grade of expansion and incubation status of the embryos is associated with the success of IVF.

The total number of oocytes collected from both ovaries correlates with the success of IVF [31]. In other words, the increase in the number of oocytes to certain limits also leads to an increase in the rate of pregnancy. However, there is no universal agreement on the ideal number of oocytes for the treatment of IVF today. According to Sunkara et al. (2011) [32], the chance of successful pregnancy rises by 37% when 15 oocytes are collected, but other authors indicate that the pregnancy rate by embryo transfer reached the absolute maximum (31%) when between 11 and 15 mature oocytes were available for injection of ICSI [33]. Among the studies, there is a suggestion for the so-called “gold standard” ranging from 5 to 15 oocytes [34]. Bals-Pratsch et al. (2010) [35] indicate that the number of oocytes is not as important as the number of mature oocytes in metaphase 2, with an ideal average number of nine. The model of the present study is in agreement with these findings, demonstrating that a higher number of oocytes tends to converge to a positive β HCG value.

Over the years, some studies have been carried out to determine the influence of the catheter models during embryo transfer. Gambadauro e Navaratnarajah (2015) [36] and Brown et al. (2016) [37] indicated that the choice of the catheter model might be one of the most important variables associated with the results of embryo transfer. In general, catheters are classified as soft or rigid. Such authors suggest that, as it prevents trauma to the cervix and endometrium, the soft model is ideal. Rigid catheters can cause bleeding or induce uterine contractions. Ren et al. (2016) [38] carried out a comparative study using catheters model Cook Sydney and Frydman-CCD to perform the embryo transfer and found higher rates of clinical pregnancy and live births in the group that used the first model. In agreement, our study showed that CCD catheter model contributed negatively to the treatment outcome.

When performing an analysis of the dose of gonadotropins, higher doses were associated to non-pregnancy. Interestingly, dose of medication was positively correlated to age of woman. In addition, an inverse correlation was observed between dose of gonadotropins and antral follicle count, since lower follicle count requires higher doses of medication, which explains the present findings.

Our algorithm showed an accuracy of 61%, AUC of 64%, a sensitivity of 61% and a specificity of 59% for Random Forest model. Hafiz et al. (2017) [6] used the Random Forest, SVM and Recursive Partitioning (RPART) algorithms to predict the success of pregnancy with IVF treatments. When comparing the results obtained, they observed that the Random Forest model reached an AUC superior to the other ML algorithms. According to their study, the variables considered as the most relevant for the model were age of woman, number of embryos developed, classification of embryos and type of treatment (IVF or ICSI). An accuracy of 83.96%, AUC of 84.23%, sensitivity of 48.85% and specificity of 98.03% were found. The main differences between our study and that of Hafiz et al. [6] were their sample size (486), the positive: negative pregnancy ratio (131:355) and the age of the parents (18 to 47 years for the women, and 23 to 70 years of age for the men), which could explain the different performance of the model.

Hassan et al. (2018) [5] used the SVM, Random Forest, Multiplayer Perception (MLP) and Classification and Regression Trees (CART) algorithms to predict IVF results. For the Random Forest model, the average accuracy obtained was 98.25% and the average AUC was 98.9%. The most relevant variables were woman's age, FSH levels, state and sperm source, and ovarian stimulation. Contrary to our study, Hassan et al. (2018) [5] investigated the variables in a private infertility clinic, and their study population had a higher pregnancy rate.

Our study presents some limitations, such as sample size, the number of models tested and several missing data that excluded the cases. Therefore, further studies with the same population characteristics should be conducted in order to validate our findings.

5. CONCLUSION

The present study evaluated the main variables related to the success of pregnancy using IVF in a public health service. Our results indicate that the optimization of controlled variables could ameliorate the frequency of women achieving success in pregnancy after IVF in services with the same characteristics.

Disclosure of interest

The authors declare that they have no conflict of interest.

Acknowledgments

Research in the authors' laboratories is funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grants to FMR and KBG. WMJ is funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo a Pesquisa de Minas Gerais (Fapemig), as well as projects Atmosphere and INCT-Cyber.

REFERENCES

- [1] Zegers-Hochschild F., Adamson G.D., de Mouzon J., Ishihara O., Mansour R., Nygren K., Sullivan E., Vanderpoel S.; International Committee for Monitoring Assisted Reproductive Technology; World Health Organization. International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART terminology, 2009. *Fertil Steril*. 2009 Nov, Volume 92, Issue 5, p. 1520-4.
- [2] Kumar N., Singh A.K. Trends of male factor infertility, an important cause of infertility: A review of literature. *J Hum Reprod Sci*. 2015 Oct-Dec, Volume 8, Issue 4, p.191-6.

- [3] Kaufmann S.J., Eastaugh J.L., Snowden S., Smye S.W., Sharma V. The application of neural networks in predicting the outcome of in-vitro fertilization. *Hum. Reprod.* Jul 1997, Volume 12, Issue 7, p. 1454-7.
- [4] Ramasamy N., Durairaj M. Feature Reduction by Improvised Hybrid Algorithm for Predicting the IVF Success Rate, 2017, Volume 8, n. 1, p. 37–40.
- [5] Hassan, R.; Al-Insaif, S.; Hossain, M. I.; Kamruzzaman, J. A machine learning approach for prediction of pregnancy outcome following IVF treatment. *Neural Computing and Applications*, Ago. 2018.
- [6] Hafiz, P.; Nematollahi, M.; Boostani, R.; Jahromi, B.N. Predicting Implantation Outcome of In Vitro Fertilization and Intracytoplasmic Sperm Injection Using Data Mining Techniques. *International Journal of Fertility and Sterility*, Oct. 2017, Volume 11, No. 3, p. 184-190.
- [7] Blank C., Wildeboer R.R., DeCruo I., Tilleman K., Weyers B., de Sutter P., Mischi M., Schoot B.C. Prediction of implantation after blastocyst transfer in in vitro fertilization: a machine-learning perspective. *Fertil Steril.* Feb. 2019, Volume 111, Issue 2, p.318-326.
- [8] Khosravi P, Kazemi E, Zhan Q, Malmsten JE, Toschi M, Zisimopoulos P, Sigaras A, Lavery S, Cooper LAD, Hickman C, Meseguer M, Rosenwaks Z, Elemento O, Zaninovic N, Hajirasouliha I. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. *NPJ Digit Med.* Apr. 2019, p.2-21.

- [9] Curchoe, C. L.; Bormann, C. L. Artificial intelligence and machine learning for human reproduction and embryology presented at ASRM and ESHRE 2018. *J Assist. Reprod. Genet.* Apr. 2019, Volume 36, Issue 4, p.591-600.
- [10] Inhorn M., Patrizio P. Infertility around the globe: New thinking on gender, reproductive technologies and global movements in the 21st century. *Human reproduction update*, Aug 201, Volume 21, Issue 4, p. 411-26.
- [11] Beam, A. L.; Kohane, I. S. Big Data and Machine Learning in Health Care. *JAMA*, Apr. 2018, Volume 319, Issue 13, p.1317-1318.
- [12] Uyar A., Bener A., Ciray H. N. Predictive Modeling of Implantation Outcome in an In Vitro Fertilization Setting: An Application of Machine Learning Methods, Aug. 2015, Volume 35, Issue 6, p.714-725.
- [13] Crown, W. H. Potential Application of Machine Learning in Health Outcomes Research and Some Statistical Cautions. *Value in Health*, 2015, Volume 18, n. 2, p. 137–140.
- [14] Qiu, J., Li, P., Dong, M. et al. Personalized prediction of live birth prior to the first in vitro fertilization treatment: a machine learning method. *J. Transl Med*, 2019, Volume 17, p.317.
- [15] Babitha, M.; A Survey on the Machine Learning Techniques used in IVF Treatment to Improve the Success Rate. *International Journal Of Engineering Research and Technology (IJERT)*, 2019, Volume 7, Issue 08.
- [16] Zhao Y.Y., Yu Y., Zhang X.W. Overall Blastocyst Quality, Trophectoderm Grade, and Inner Cell Mass Grade Predict Pregnancy Outcome in Euploid

Blastocyst Transfer Cycles. Chinese Medical Journal, 2018, Volume 131, p. 1261-67.

[17] Lai I., Neal M., Gervais N., Amin S., Taerk E., Faghieh M. Transfers of lower quality embryos based on morphological appearance result in appreciable live birth rates: a Canadian center's experience, F&S Reports, 2020, , Volume 1, Issue 3, p. 264-269.

[18] Rodríguez-Pérez R., Bajorath J. Interpretation of Compound Activity Predictions from Complex Machine Learning Models Using Local Approximations and Shapley Values. J Med Chem. Aug. 2020, Volume 63, Issue 16, p.8761-8777.

[19] Kühnlein M., Appelhans T., Thies B., Nauss T. Improving the accuracy of rainfall rates from optical satellite sensors with machine learning — A random forests-based approach applied to MSG SEVIRI. Remote Sensing of Environment, 2014, Volume 141, p. 129-143.

[20] Brugo-Olmedo S., Chillik C., Kopelman S. Definition and causes of infertility. Reprod Biomed Online, 2001, Volume 2, Issue 1, p. 41–53.

[21] Hamilton, B.E., Martin J.A., Osterman M.J., Curtin S.C., Matthews T.J. Births: Final Data for 2014. Natl Vital Stat Rep. 2015, Volume 64, Issue 12, p.1-64.

[22] Rochebrochard E. L.; Thonneau P. Paternal age \geq 40 years: an important risk factor for infertility. American Journal of Obstetrics and Gynecology, Elsevier, 2003, Volume 189, Issue 4, p.901-905.

[23] Matalliotakis I., Cakmak H., Sakkas D., Mahutte N., Koumantakis G., Arici A. Impact of body mass index on IVF and ICSI outcome: a retrospective study. *Reprod Biomed Online*, Jun 2008, Volume 16, Issue 6, p. 778-83.

[24] Rossi A.C., Prefumo F. The effects of surgery for endometriosis on pregnancy outcomes following in vitro fertilization and embryo transfer: a systematic review and meta-analysis. *Arch Gynecol Obstet*. Sep. 2016, Volume 294, Issue 3, p.647-655.

[25] Murta M., Machado R.C., Zegers-Hochschild F., Checa M.A., Sampaio M., Geber S. Endometriosis does not affect live birth rates of patients submitted to assisted reproduction techniques: analysis of the Latin American Network Registry database from 1995 to 2011. *J Assist Reprod Genet*. 2018 Aug, Volume 35, Issue 8, p. 1395-1399.

[26] Sharma S., Bathwal S., Agarwal N., Chattopadhyay R., Saha I., Chakravarty B. Does presence of adenomyosis affect reproductive outcome in IVF cycles? A retrospective analysis of 973 patients. *Reprod. Biomed. Online*, Jan 2019, Volume 38, Issue 1, p.13-21.

[27] Ballester M., d'Argent E. M., Morcel K., Belaisch-Allart J., Nisolle M., Daraï E. Cumulative pregnancy rate after ICSI-IVF in patients with colorectal endometriosis: results of a multicentre study. *Hum. Reprod*. Apr 2012, Volume 27, Issue 4, p.1043-9.

[28] Ferrero S., Scala C., Racca A., Calanni L., Remorgida V., Venturini P.L., Maggiore U. L. R. Second surgery for recurrent unilateral endometriomas and impact on ovarian reserve: a case-control study. *Fertil Steril*. May 2015, Volume 103, Issue 5, p. 1236-43.

[29] Maignien C., Santulli P., Gayet V., Lafay-Pillet M.C., Korb D., Bourdon M., Marcellin L., de Ziegler D., Chapron C. Prognostic factors for assisted reproductive technology in women with endometriosis-related infertility. *Am J Obstet Gynecol*, Mar 2017, Volume 216, Issue 3, p. 280-289.

[30] Vaegeter K.K., Lakic T.G., Olovsson M., Berglund L., Brodin T., Holte J. Which factors are most predictive for live birth after in vitro fertilization and intracytoplasmic sperm injection (IVF/ICSI) treatments? Analysis of 100 prospectively recorded variables in 8,400 IVF/ICSI single-embryo transfers. *Fertil Steril*, Mar 2017.

[31] Choe S.A., Ku S.Y., Jee B.C., Suh C.S., Kim S.H., Choi Y.M., Gu Kim J., Yong Moon S. Symmetry in number of retrieved oocytes between two ovaries: a possible predictor of in vitro fertilization outcome. *Gynecol. Endocrinol.* Dec 2011, Volume 27, Issue 12, p.997-1000.

[32] Sunkara S.K., Rittenberg V., Raine-Fenning N., Bhattacharya S., Zamora J., Coomarasamy A. Association between the number of eggs and live birth in IVF treatment: an analysis of 400 135 treatment cycles. *Hum Reprod.* Jul 2011, Volume 26, Issue 7, p.1768-1774.

[33] Steward R.G., Lan L., Shah A.A., Yeh J.S., Price T.M., Goldfarb J.M., et al. Oocyte number as a predictor for ovarian hyperstimulation syndrome and live birth: an analysis of 256,381 in vitro fertilization cycles. *Fertil Steril* 2014, Volume 101, p.967-973.

[34] Timeva T., Milachich T., Antonova I., Arabaji T., Shterev A., Omar H. Correlation Between Number of Retrieved Oocytes and Pregnancy Rate After In

Vitro Fertilization/IntraCytoplasmic Sperm Infection. The Scientific World Journal, Apr 2006, Volume 6, p.686-90.

[35] Bals-Pratsch M., Bühler K., Krüssel J., Wendelken M., Dahncke W., Kupka MS. Extended Analyses of the German IVF Registry (DIR): Andrological Aspects, Medical-Economical Assumptions Related to the Shift from IVF to ICSI and Stimulation with Gonadotropins. J. Reproduktionsmed. Endokrinol, 2010, Volume 7, Issue 1, p.40-44.

[36] Gambadauro P., Navaratnarajah R. Reporting of embryo transfers methods in IVF research: a cross-sectional study. Reprod Biomed Online. 2015, Volume 30, Issue 2, p. 137-143.

[37] Brown J., Buckingham K., Buckett W., Abou-Setta A.M. Ultrasound versus 'clinical touch' for catheter guidance during embryo transfer in women. Cochrane Database of Systematic Reviews, 2016, Issue 3. Art. No.: CD006107.

[38] Ren X.L., Liu P., Lian Y., Huang J., Zheng X.Y., Zhu Y.J., Qiao J. Effect of catheter choice during embryo transfer on the clinical outcome of in vitro fertilization-embryo transfer. Beijing Da Xue Xue Bao Yi Xue Ban. Oct 2016, Volume 48, Issue 5, p.905-909.

Table 1: Set of varied hyperparameters for each algorithm

Algorithm	Hyperparameters
LR	C = {0.1, 1.0}; penalty = {l1, l2}
SVM	kernel = {rbf, linear}; C = {0.001, 0.01, 0.05, 0.1, 0.5, 1, 10}
RF	number of trees = {1, 2, 3, ..., 50}; depth of estimators = {1, 2, 3, ..., 20}
XGBoost	number of trees = {1, 2, 3, ..., 50}; learning rate = {0.01, 0.05, 0.1}; max depth of estimators = {1, 2, 3, ..., 19}

Table 2: Variables included in the final algorithm models

Variables	
Age of the man	Number of blastocysts
Age of the Woman	Number of embryos transferred
BMI	Number of type embryos transferred A
Antral follicle count	Number of type embryos transferred B
Medication Total dose	Number of type embryos transferred C
Number of follicles greater than 17 mm	Number of type embryos transferred D
Number of oocytes	Number of blastocysts transferred
Number of MII oocyte	Indication of IVF
Number of IVF oocytes	Sperm source
Number of oocytes injected in the insemination	Semen status
Total of fertilized pro-nuclei	Semen preparation
Number of embryos with the classification A	Type of gonadotropin
Number of embryos with the classification B	IVF or ICSI
Number of embryos with the classification C	Transferred embryonic stage
Number of embryos with the classification D	Catheter model used for embryo transfer
Total number of viable cleavage stage embryos	β HCG positive/negative*

*BMI: body mass index; IVF: In vitro fertilization; ICSI: intracytoplasmic sperm injection; MII: Metaphase 2. * Outcome.*

Table 3: Characterization of the cases included in the study, according to the presence or absence of pregnancy in univariate analysis.

Variables	Pregnancy (n=276)	Non-pregnancy (n=475)	p
Woman Age	37 (5)	37 (6)	0.015
BMI	25.58 (7)	26.63 (8)	0.113
Man Age	40 (9)	38 (8)	0.049
Indication			
Male Factor	149 (54.4%)	239 (48.5%)	0.016
Endometriosis	22(8.0%) a	68 (13.8%) b	
Tuboperitoneal Factor	64 (23.3%)	129 (26.2%)	
Ovarian Factor	3 (1.1%)	7 (1.4%)	
Unexplained	29 (10.6%) b	48 (9.7%) a	
Others	7 (2.6%) b	2 (0.4%) a	
Origin of Semen			
Donated Sperm	17 (6.2%)	20 (4.0%)	0.063
Ejaculate + thesis	0 (0.0%)	2 (0.4%)	
Ejaculated Sperm	209 (76.0%)	419 (84.9%)	
Heterologous	1 (0.4%)	0 (0.0%)	
MESA	2 (0.7%)	4 (0.8%)	
PESA	37 (13.4%)	36 (7.3%)	
PESA + TESE	0 (0.0%)	1 (0.2%)	
TESE	9 (3.3%)	11 (2.2%)	
TESE + Heterologous	0 (0.0%)	1 (0.2%)	
Semen State			
Fresh	257 (93.5%)	465 (94.3%)	0.637
Cryopreserved	18 (6.5%)	28 (5.7%)	
FIV or ICSI			

FIV	98 (36.4%)	186 (38.4%)	0.075
ICSI	163 (60.6%)	294 (60.8%)	
FIV/ICSI	8 (3.0%)	4 (0.8%)	

Continuous variables are shown as medians and interquartile ranges. BMI: body mass index; a= less frequent (residue <-1.96), b= more frequent (residue >+1.96)

Table 4: Performance of machine learning models applied.

Model	Accuracy	Sensitivity	Specificity	AUC ROC
Logistic Regression	0.62	0.58	0.70	0.64
Randon Forest	0.61	0.61	0.59	0.64
XGBoost	0.58	0.56	0.62	0.61
SVM	0.58	0.56	0.59	0.61

SVM: Support Vector Machine

Figure Legends

Figure 1. Subdivision of the training set and the model validation.

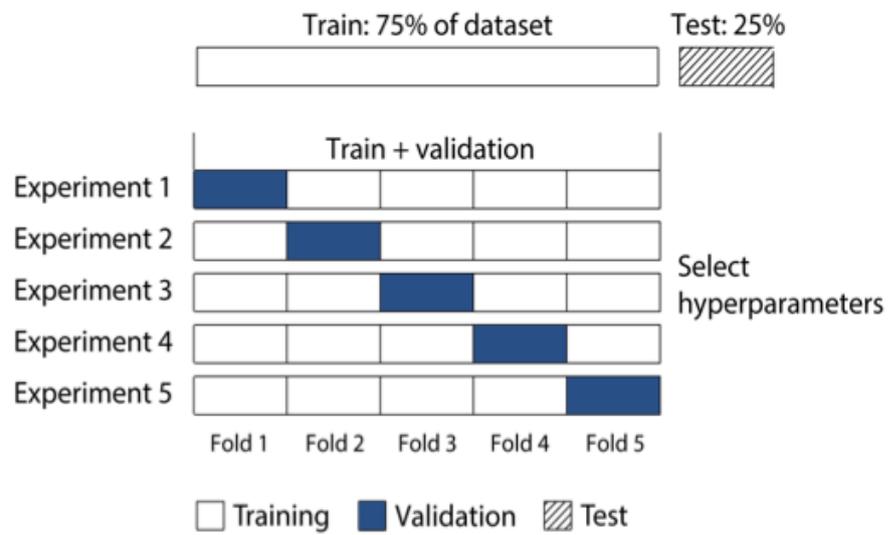


Figure 2. Area under the ROC curve for the pregnancy prediction model using Random Forest.

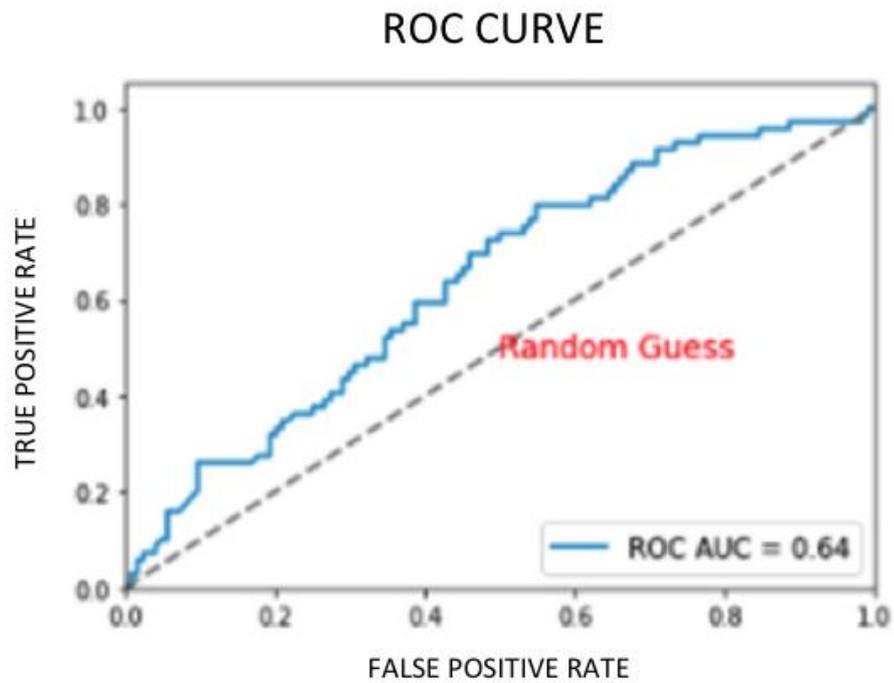


Figure 3. Confusion matrix of the Random Forest model.

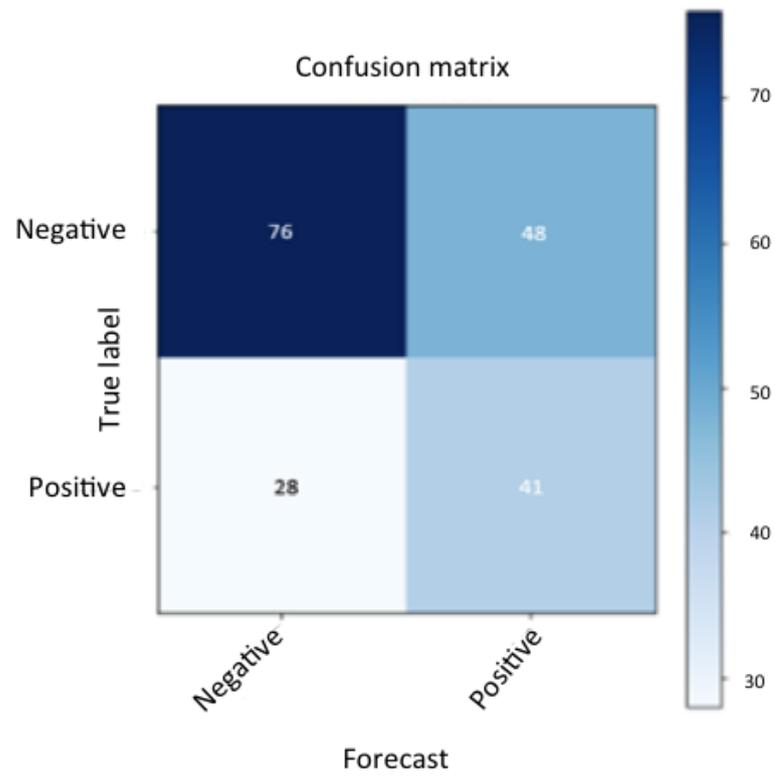


Figure 4. Individual SHAP result for the variables indication for (A) IVF treatment and (B) catheter model.

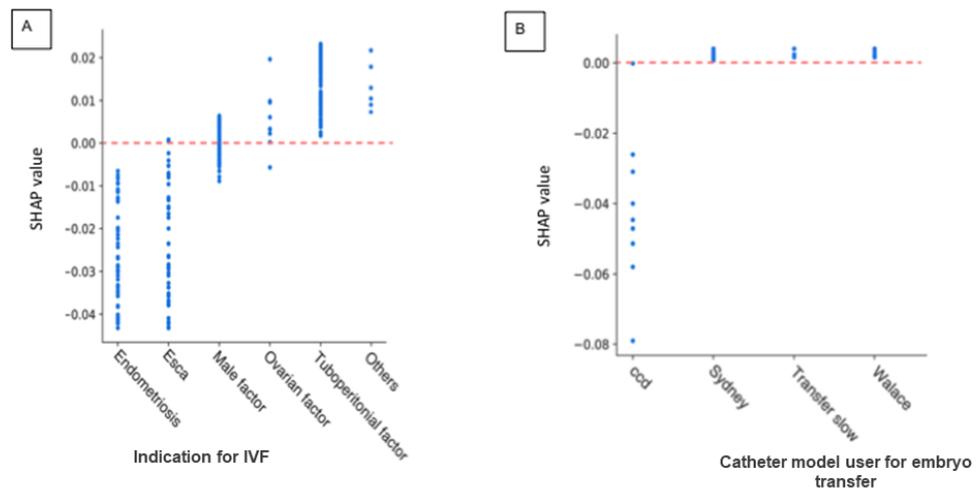


Figure 5: SHAP value for the Random Forest model.

Values below 0 on the X axis tend to the negative β HCG result and values above 0 tend to the positive result. The red color indicates the greater the value of the variable and the blue color indicates the lower value. For example, the variable “Age of Women” showed that the higher value is associated to a tendency for the negative result of β HCG.

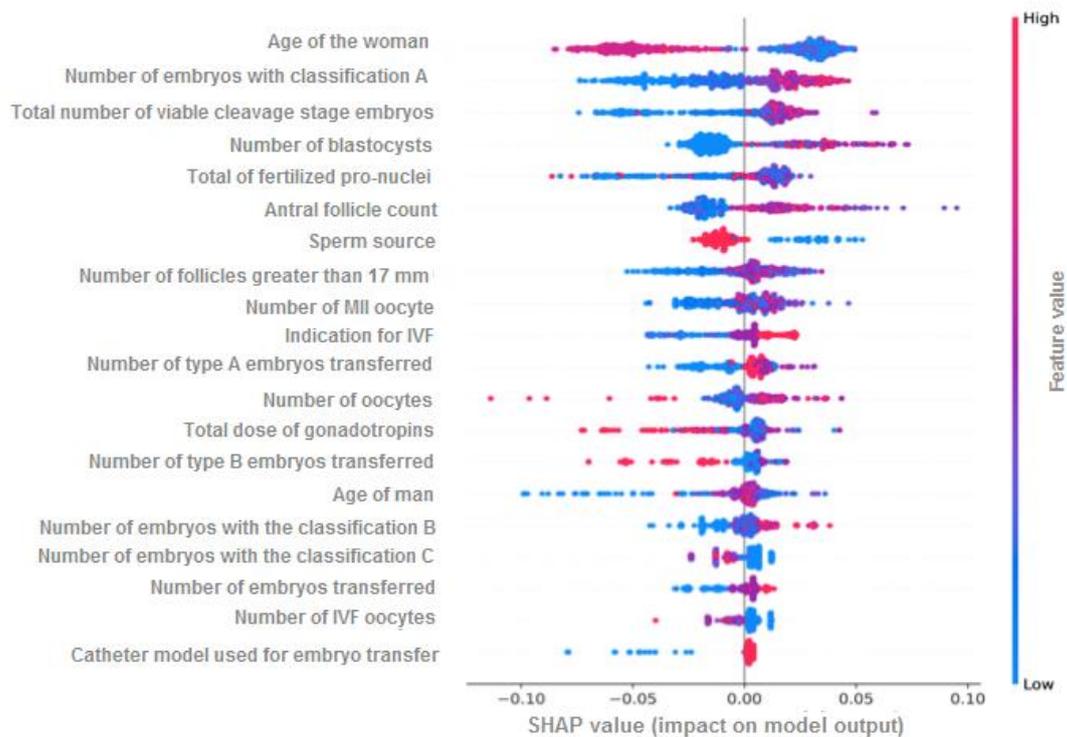
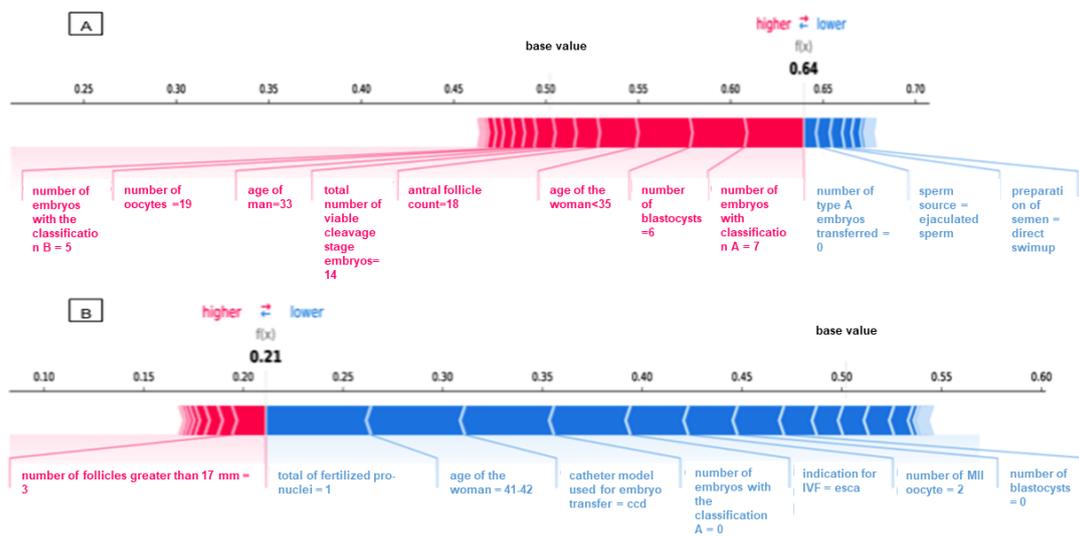


Figure 6. Algorithm applied in the prediction of (A) a real positive pregnancy and (B) a real negative pregnancy.

The variables that contributed to decrease the probability value of the output are represented in blue color, and the variables that contributed to increase the output value are represented in red color. The strength of the importance of the output variable is indicated by the size of the arrow.



Supplementary material 1

Variables	
Age of the man	Number of blastocysts
Age of the woman	Number of embryos transferred
BMI	Number of type embryos transferred A
Antral follicle count	Number of type embryos transferred B
Total dose of gonadotropins	Number of type embryos transferred C
Number of follicles greater than 17 mm	Number of type embryos transferred D
Number of oocytes	Number of blastocysts transferred
Number of MII oocyte	Indication of IVF
Number of IVF oocytes	Sperm source
Number of oocytes injected in the insemination	Semen status
Total of fertilized pro-nuclei	Semen preparation
Number of embryos with the classification A	Type of gonadotropin
Number of embryos with the classification B	IVF or ICSI
Number of embryos with the classification C	Transferred embryonic stage
Number of embryos with the classification D	Catheter model used for embryo transfer
Total number of viable cleavage stage embryos	β HCG positive/negative*
Woman date of birth	Cycle Type
GnRH Drugs	GnRHa Type
GnRH Responsible	Analog date
Total Cycle Number	Date of Medication
Infertility time	FSH Type
Primary or secondary infertility	LH Type
Women's Weight	Other medication
Women's Height	Number of ampoules Gonal®
FSH Third day	Number of ampoules Menopur®
Indication 2 REDLARA	Pergoveris
Indication 2 HRL	HCG Responsable
Semen Observation	Collection Date
Lumen Needle	Follicules Number

Follicle Collection Responsible	Follicle count greater than 12mm and smaller than 17mm
Donor Oocytes Number	Follicle count smaller than 12 millimeters
Number Germinal Vesicle	Follicle Collection Responsible Lab
Frozen Oocytes	Receiver Oocytes Number
Fertilized (2pn)	Oocyte Observation
Fertilized (3pn)	Lab insemination Responsible
Discarded Embryos	Fertilized (1pn)
Preimplantation genetic diagnosis	Number of embryos kept in culture Day 3 to Day 5/Day 6
Medical responsible for the transfer	Assisted Hatching
Post-collection Day	Number of biopsed
Transfer with or without trauma	Embryo Transferred
Lutea phase support	Lab responsible for the transfer
Frozen Embryo Stage	New or Reprocessed Catheter
Freezing Technique	Lutea phase support - Redlara
Number of unfeasible embryos thawed	Number of frozen embryos
Freezing Day	Embryo Class

*BMI: body mass index; IVF: In vitro fertilization; ICSI: intracytoplasmic sperm injection; MII: Metaphase 2; REDLARA - Latin American Network of Assisted Reproduction. HRL: human reproduction laboratory. * outcome*

5. DISCUSSÃO

Neste estudo, foram desenvolvidos quatro modelos de *Machine Learning* como forma de avaliar a capacidade de predição do sucesso de gravidez após realização de FIV usando variáveis pré-implantacionais. O modelo *Random Forest* mostrou o melhor desempenho e foi capaz de selecionar 20 variáveis diretamente relacionadas à predição da gestação.

A variável “idade da mulher” mostrou que aquelas que não evoluíram para a gestação, apresentaram maior intervalo interquartilico de idade quando comparadas àquelas que engravidaram. O modelo foi capaz de mostrar o impacto da idade da mulher em outras variáveis, como a correlação inversa entre o a idade da mulher e os valores absolutos das variáveis total de pro-núcleos fertilizados, contagem de folículos antrais, contagem de folículos maiores que 17mm, número de MII e número de oócitos. O envelhecimento do homem também foi um fator importante. Nossos resultados mostraram que a idade mediana dos homens foi maior no grupo que alcançou a gravidez, dado corroborado pelo algoritmo SHAP.

Em relação ao IMC, não foi observada diferença em relação entre os dois grupos de desfecho. Esse resultado corrobora o estudo de Matalliotakis et al. (2008) [62].

A análise dos fatores indicativos de FIV mostrou que mulheres com endometriose foram mais frequentes entre aquelas que não engravidaram, enquanto aquelas com “outros fatores” (casal soropositivo, doenças infecciosas, fibroides uterinos, perda gestacional recorrente, produção independente, útero unicorno) tiveram maior sucesso na FIV. O modelo SHAP também indicou que o fator endometriose contribuiu no grupo daquelas mulheres que não engravidaram e o grupo do fator tuboperitoneal tiveram maior contribuição para o sucesso na gestação, seguido de “outros fatores”.

Dentre as diferentes origens do sêmen, o esperma obtido por ejaculação foi mais frequente no grupo das mulheres não grávidas, resultado evidenciado pelo modelo do SHAP. Em relação aos métodos de criopreservação do sêmen, em nosso estudo, somente 5% da população fez uso do sêmen criopreservado,

frente a 95% que utilizou o sêmen fresco e indicou não haver diferença significativa de resultados entre os grupos.

Zhao et al. (2018) [63] e Lai et al. (2020) [64] demonstram a importância da classificação de embriões e que a utilização dos tipos A e B obtiveram maiores taxas de sucesso da gestação. Tal aspecto corrobora com o modelo obtido, uma vez que maior número de embriões com as classificações A e B ou transferidos do tipo A, reportados pelo SHAP, convergiram para o β HCG positivo.

A variável modelo do cateter mostrou-se como fator importante. Ren et al. (2016) [65] realizaram um estudo comparativo utilizando cateteres modelo Cook Sydney e Frydman-CCD para a transferência de embriões e encontraram maiores taxas de gravidez clínica e nascidos vivos no grupo que utilizou o primeiro modelo. Em concordância, nosso estudo mostrou que o modelo de cateter CCD contribuiu negativamente para o resultado do tratamento.

Em relação ao número de oócitos coletados, segundo Sunkara et al. (2011) [66], a chance de sucesso de gravidez sobe em 37% quando 15 oócitos são coletados. O presente estudo está de acordo com esses achados, demonstrando que um maior número de oócitos tende a convergir para um valor positivo de β HCG.

Ao realizar uma análise da dose de gonadotropinas, doses maiores foram associadas à não gravidez. A dose do medicamento foi positivamente correlacionada com a idade da mulher e foi observada correlação inversa entre a dose de gonadotropinas e a contagem de folículos antrais, uma vez que a contagem de folículos menores requer doses maiores de medicamentos, o que explica os presentes achados.

Hafiz et al. (2017) [6] também utilizaram os algoritmos Random Forest, SVM, dentre outros, para prever o sucesso da gravidez de tratamentos de FIV. Ao comparar os resultados obtidos, tal estudo indicou que o modelo Random Forest atingiu uma área sob a curva ROC superior aos demais algoritmos de ML. Hassan et al. (2018) [5] fizeram o uso dos algoritmos SVM, Random Forest, Multiplayer Perception (MLP) e Classification and Regression Trees (CART) na predição dos resultados de FIV. Para o modelo de Random Forest, a acurácia média obtida foi de 98,25% e o AUC médio de 98,9%. Já o presente estudo

obteve uma acuraria de 61% e uma AUC de 64%, uma sensibilidade de 61% e uma especificidade de 59% para o Random Forest, o que pode ser justificado pela análise de diferentes grupos entre os dois estudos.

Nosso estudo apresentou algumas limitações, como o tamanho amostral limitado, ausência de população externa para validação do modelo, o número de limitado de modelos testados e a ocorrência de dados ausentes que obrigaram a exclusão de vários casos. No entanto, podemos afirmar a sua importância, na medida em que possibilitou o conhecimento de variáveis, muitas delas controláveis, que podem contribuir para ao sucesso da gestação em mulheres submetidas à FIV em um serviço público. Outros estudos, incluindo diferentes populações, deverão aplicar o algoritmo de modo a validar nossos achados.

6. CONCLUSÕES

O presente estudo foi capaz de desenvolver um algoritmo baseado em Random Forest, agregando 20 variáveis, que são capazes de prever, com 61% de acurácia, a chance de sucesso na gestação após procedimento de FIV.

7. PERSPECTIVAS

- 1) Aplicação do algoritmo desenvolvido em casos novos no Ambulatório de Reprodução Humana do Hospital das Clínicas da UFMG para sua validação.
- 2) Aplicação do mesmo algoritmo em outros serviços de fertilização in vitro para sua validação em outras populações, como serviços privados de atendimento.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Cortina M., Ozan J. Basic Infertility Evaluation. *Ambulatory Gynecology*, May 2018, p. 247-259.
- [2] Kumar N., Singh A.K. Trends of male factor infertility, an important cause of infertility: A review of literature. *J Hum Reprod Sci*. 2015 Oct-Dec, Volume 8, Issue 4, p.191-6.
- [3] Kaufmann S.J., Eastaugh J.L., Snowden S., Smye S.W., Sharma V. The application of neural networks in predicting the outcome of in-vitro fertilization. *Hum. Reprod*. Jul 1997, Volume 12, Issue 7, p. 1454-7.
- [4] Ramasamy N., Durairaj M. Feature Reduction by Improvised Hybrid Algorithm for Predicting the IVF Success Rate, 2017, Volume 8, n. 1, p. 37–40.
- [5] Hassan, R.; Al-Insaif, S.; Hossain, M. I.; Kamruzzaman, J. A machine learning approach for prediction of pregnancy outcome following IVF treatment. *Neural Computing and Applications*, Ago. 2018.
- [6] Hafiz, P.; Nematollahi, M.; Boostani, R.; Jahromi, B.N. Predicting Implantation Outcome of In Vitro Fertilization and Intracytoplasmic Sperm Injection Using Data Mining Techniques. *International Journal of Fertility and Sterility*, Out. 2017, Volume 11, No. 3, p. 184-190.
- [7] Blank C., Wildeboer R.R., DeCruo I., Tilleman K., Weyers B., de Sutter P., Mischi M., Schoot B.C. Prediction of implantation after blastocyst transfer in in vitro fertilization: a machine-learning perspective. *Fertil Steril*. Feb. 2019, Volume 111, Issue 2, p.318-326.

[8] Khosravi P, Kazemi E, Zhan Q, Malmsten JE, Toschi M, Zisimopoulos P, Sigaras A, Lavery S, Cooper LAD, Hickman C, Meseguer M, Rosenwaks Z, Elemento O, Zaninovic N, Hajirasouliha I. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. *NPJ Digit Med.* Apr. 2019, p.2-21.

[9] Curchoe, C. L.; Bormann, C. L. Artificial intelligence and machine learning for human reproduction and embryology presented at ASRM and ESHRE 2018. *J Assist. Reprod. Genet.* Apr. 2019, Volume 36, Issue 4, p.591-600.

[10] Beam, A. L.; Kohane, I. S. Big Data and Machine Learning in Health Care. *JAMA,* Apr. 2018, Volume 319, Issue 13, p.1317-1318.

[11] Crown, W. H. Potential Application of Machine Learning in Health Outcomes Research and Some Statistical Cautions. *Value in Health,* 2015, Volume 18, n. 2, p. 137–140.

[12] Uyar A., Bener A., Ciray H. N. Predictive Modeling of Implantation Outcome in an In Vitro Fertilization Setting: An Application of Machine Learning Methods, Aug. 2015, Volume 35, Issue 6, p.714-725.

[13] Durairaj, M.; Thamilselvan, P.; Applications of artificial neural network for IVF data analysis and prediction. *Journal of Engineering, Computers & Applied Sciences,* Sep 2013, Volume 2, No. 9, p.11–15.

[14] Mascarenhas M.N., Flaxman S.R., Boerma T., Vanderpoel S., Stevens G.A. National, Regional, and Global Trends in Infertility Prevalence Since 1990: A Systematic Analysis of 277 Health Surveys. *PLOS Med.* 2012, Volume 9, Issue 12.

- [15] Jahanbani Y., Davaran S., Ghahremani-Nasab M., Aghebati-Maleki L., Yousefi M. Scaffold-based tissue engineering approaches in treating infertility. *Life Sci*, Jan 2020, Volume 240.
- [16] Boivin, J.; Bunting, L.; Collins, J. A.; Nygren, K. G.; International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care., *Human Reproduction*, Jun 2007, Volume 22, Issue 6, p. 1506–1512.
- [17] Brugo-Olmedo S., Chillik C., Kopelman S. Definition and causes of infertility. *Reprod Biomed Online*, 2001, Volume 2, Issue 1, p. 41–53.
- [18] ACOG Committee Opinion No. 762: Prepregnancy Counseling, *Obstetrics & Gynecology*, Jan 2019, Volume 133, Issue 1, p78-89.
- [19] Kohler H. P., Ortega, J. A. Tempo-adjusted period parity progression measures, fertility postponement and completed cohort fertility. *Demographic Research*, 2002, Volume 6, Issue 6, p.91–144.
- [20] Small C.M., Manatunga A. K., Klein M, Feigelson H. S., Dominguez, C.E., McChesney R., Marcus M. Menstrual Cycle Characteristics: Associations with Fertility and Spontaneous Abortion. *Epidemiology*: January 2006, Volume 17, Issue 1, p. 52-60.
- [21] Finer, L. B., Philbin, J. M. Trends in ages at key reproductive transitions in the United States, 1951–2010. *Women’s Health Issues*, 2014, Volume 24, Issue 3, p. 271–279.
- [22] Seifer D., Naftolin, F. Moving toward an earlier and better understanding of the perimenopause. *Fertility and Sterility*, 1998, Volume 69, Issue 3, p. 387–388.

[23] Hunt, P., Hassold, T. Female meiosis: Coming unglued with age. *Current Biology*, 2010, Volume 20, Issue 17, p.699-702.

[24] Practice Committee of the American Society for Reproductive Medicine. Testing and interpreting measures of ovarian reserve: a committee opinion. *Fertil Steril*. Mar 2015, Volume 103, Issue 3, p.9-17.

[25] Crossman, S. H. The challenge of pelvic inflammatory disease. *Am Fam Physician*, 2006 Dec 15, Volume 74, Issue 12, 2024.

[26] Stenkeler, J. A., Woodfield, C. A., Lazarus, E., Hillstrom, M. M. Female infertility: A systematic approach to radiologic imaging and diagnosis. *Radiographics*, 2009, Volume 29, Issue 5, p. 1353–1370.

[27] Sirmans S. M., Pate, K. A. Epidemiology, diagnosis, and management of polycystic ovary syndrome. *Clinical Epidemiology*, 2014, Volume 6, p.1–13.

[28] Bergh, C. M., Moore, M., Gundell, C. Evidence-based management of infertility in women with polycystic ovary syndrome. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, 2016, Volume 45, Issue 1, p.111–122.

[29] Practice Committee of the American Society for Reproductive Medicine. Removal of myomas in asymptomatic patients to improve fertility and/or reduce miscarriage rate: a guideline. *Fertil Steril*. 2017 Sep, Volume 108, Issue 3, p. 416-425.

[30] Nnoaham, K. E., Hummelshoj, L., Webster, P., d'Hooghe, T., de Cicco Nardone, F., de Cicco Nardone, C., Zondervan, K. T. et al. Impact of endometriosis on quality of life and work productivity: A multicenter study across ten countries. *Fertility and Sterility*, 2011, Volume 96, Issue 2, p. 366–373.

- [31] Ciarmela, P., Critchley, H., Christman, G. M., Reis, F. M. Pathogenesis of endometriosis and uterine fibroids. *Obstetrics and Gynecology International*, 2013.
- [32] Krausz, C. Male infertility: Pathogenesis and clinical diagnosis. *Best Practice & Research Clinical Endocrinology & Metabolism*, 2011, Volume 25, Issue 2, p. 71–85.
- [33] Stevenson E., Hershberger P., Bergh P. Evidence-Based Care for Couples with Infertility. *Journal of Obstetric, Gynecologic & Neonatal Nursing*, 2016, Volume 45, p. 100-110.
- [34] Wald, M. Male infertility: Causes and cures. *Sexuality, Reproduction & Menopause*, 2005, Volume 3, Issue 2, p.83–87.
- [35] Kolon, T. F., Herndon, C. D., Baker, L. A., Baskin, L. S., Baxter, C. G., Cheng, E. Y., Barthold, J. S. et al. Evaluation and treatment of cryptorchidism: AUA guideline. *Journal of Urology*, Aug 2014, Volume 192, Issue 2, p. 337–345.
- [36] HOTALING, J. M. Genetics of male fertility. *Urologic Clinics of North America*, 2014, Volume 41, Issue 1, p. 1–17.
- [37] Nachtigall, RD. International disparities in access to infertility services. *Fertil Steril* Apr 2006, Volume 85, Issue 4, p. 871–875.
- [38] Ombelet W., Cooke I., Dyer S., Serour G., Devroey P. Infertility and the provision of infertility medical services in developing countries. *Human Reprod Update* Nov-Dec 2008, Volume 14, Issue 6, p. 605–621.

[39] Ombelet W., Campo R. Affordable IVF for Developing Countries. Reproductive Biomedicine Online, Jul 2007, Volume 15, Issue 3 p.257-265.

[40] Inhorn M., Patrizio P. Infertility around the globe: New thinking on gender, reproductive technologies and global movements in the 21st century. Human reproduction update, Aug 201, Volume 21, Issue 4, p. 411-26.

[41] Gurunath S., Pandian Z., Anderson R.A., Bhattacharya S. Defining infertility—a systematic review of prevalence studies. Human Reprod Update 2011, Volume 17, p. 575–588.

[42] Rutstein S.O., Shah I.H. Infecundity, infertility, and childlessness in developing countries. DHS Comparative Reports No. 9. World Health Organization, Geneva, Switzerland, 2004.

[43] Sexual and reproductive health: infertility is a global public health issue. World Health Organization. Disponível em: <http://www.who.int/reproductivehealth/topics/infertility/perspective/en/> Acesso em: 06 Junho de 2019.

[44] Wang J.; Sauer M.V.; In vitro fertilization (IVF): a review of 3 decades of clinical innovation and technological advancement. Therapeutics and Clinical Risk Management, 2006, Volume 2, Issue 4, p. 355-364.

[45] Adamson, G.; Tabangin, M.; Macaluso, M.; The number of babies born globally after treatment with the assisted reproductive technologies (ART). Fertility and Sterility, Sep. 2013, Volume 100, No. 3, S42.

[46] Merchant R., Gandhi G.; Allahbadia G.N.; In Vitro Fertilization /Intracytoplasmic Sperm Injection For Male Infertility. Indian Journal Of Urology, Mar 2011, Volume 27, Issue 1, p.121-132.

[47] Kumar P.; Sait S.F.; Luteinizing hormone and its dilemma in ovulation induction. Journal of Human Reproductive Sciences, Jan 2011, Volume 4, Issue 1, p.2-7.

[48] Sepúlveda S. J., Portella J. R., Noriega L. P., Escudero E. L., Noriega L. H. Extended culture up to the blastocyst stage: A strategy to avoid multiple pregnancies in assisted reproductive technologies. Biological Research, 2011, Volume 44, p.195-199.

[49] Copperman, A.; Benadiva, C. Optimal usage of the GnRH antagonists: A review of the literature. Reproductive Biology and Endocrinology, 2013, Volume 11. No. 20.

[50] Rani K., Paliwal S. A Brief Review on In-Vitro fertilization (IVF): An advanced and miraculous gateway for infertility Treatments. World Journal of Pharmacy and Pharmaceutical Sciences, 2014, Volume 3, Issue 4, p.647-658.

[51] Karpman E., Williams D., Lipshultz L. IVF and ICSI in Male Infertility: Update on Outcomes, Risks, and Costs. The Scientific World Journal, 2005, Volume 5, p.922-32.

[52] Proctor M.; Johnson N.; Van Peperstraten A.M.; Phillipson G.; Techniques for surgical retrieval of sperm prior to intracytoplasmic sperm injection (ICSI) for azoospermia. Cochrane Database of Systematic Reviews, 2008, Issue 2. Art. No.: CD002807.

[53] Combelles C. M., Morozumi K., Yanagimachi R., Zhu L., Fox J. H., Racowsky C. Diagnosing cellular defects in an unexplained case of total fertilization failure. *Hum Reprod*, 2010, Volume 25, p.1666-71.

[54] Ming L., Liu P., Qiao J., et al. Synchronization between embryo development and endometrium is a contributing factor for rescue ICSI outcome. *Reprod Biomed Online*, 2012, Volume 24, Issue 5, p.527-531.

[55] Babitha, M. A Survey on the Machine Learning Techniques used in IVF Treatment to Improve the Success Rate. *International Journal Of Engineering Research and Technology (IJERT)*, 2019, Volume 7, Issue 08.

[56] Brynjolfsson, E.; Mitchell, T. What can machine learning do? Workforce implications. *Science*, Dez 2017, Volume 358, Issue 6370, p.1530–1534.

[57] Jamei M., Nisnevich A., Wetchler E., Sudat S., Liu E. Predicting all-cause risk of 30-day hospital readmission using artificial neural networks. *PLOS ONE*, 2017, Volume 12.

[58] Simeone O. A Brief Introduction to Machine Learning for Engineers. *Foundations and Trends R in Signal Processing*, 2018, Volume 12, No. 3-4, p. 200–431.

[59] Guh R. S., Tsung-Chieh J. W., Weng S. P. Integrating genetic algorithm and decision tree learning for assistance in predicting in vitro fertilization outcomes. *Expert Systems with Applications*, Abr 2011, Volume 38, Issue 4, p. 4437–4449.

[60] Güvenir, H. A.; Misirli, G.; Dilbaz, S.; Ozdegirmenci, O.; Demir, B.; Dilbaz, B. Estimating the chance of success in IVF treatment using a ranking algorithm. *Med. Biol. Eng. Comput.*, Sep 2015, Volume 53, Issue 9, p. 911-920.

[61] Practice Committee of the American Society for Reproductive Medicine. Diagnostic evaluation of the infertile female: a committee opinion. *Fertil Steril*. Jun 2015, Volume 103, Issue 6, p.44-50.

[62] Matalliotakis I., Cakmak H., Sakkas D., Mahutte N., Koumantakis G., Arici A. Impact of body mass index on IVF and ICSI outcome: a retrospective study. *Reprod Biomed Online*, Jun 2008, Volume 16, Issue 6, p. 778-83.

[63] Zhao Y.Y., Yu Y., Zhang X.W. Overall Blastocyst Quality, Trophectoderm Grade, and Inner Cell Mass Grade Predict Pregnancy Outcome in Euploid Blastocyst Transfer Cycles. *Chinese Medical Journal*, 2018, Volume 131, p. 1261-67.

[64] Lai I., Neal M., Gervais N., Amin S., Taerk E., Faghih M. Transfers of lower quality embryos based on morphological appearance result in appreciable live birth rates: a Canadian center's experience, *F&S Reports*, 2020, , Volume 1, Issue 3, p. 264-269.

[65] Ren X.L., Liu P., Lian Y., Huang J., Zheng X.Y., Zhu Y.J., Qiao J. Effect of catheter choice during embryo transfer on the clinical outcome of in vitro fertilization-embryo transfer. *Beijing Da Xue Xue Bao Yi Xue Ban*. Oct 2016, Volume 48, Issue 5, p.905-909.

[66] Sunkara S.K., Rittenberg V., Raine-Fenning N., Bhattacharya S., Zamora J., Coomarasamy A. Association between the number of eggs and live birth in IVF treatment: an analysis of 400 135 treatment cycles. *Hum Reprod*. Jul 2011, Volume 26, Issue 7, p.1768-1774.