

Прогнозирование раннего исхода ЭКО с помощью методов машинного обучения Иващенко Д. В.¹, Чечехина Е.С.¹, Чечехин В.И.¹, Щербакова Л.Н.¹, Климов А.А.¹, Младова Е.С.² ¹Медицинский научно-образовательный институт, МГУ имени М.В. Ломоносова, Москва, Россия ²Институт репродуктивной медицины REMEDI, Москва, Россия

Научный руководитель – Щербакова Лия Ниязовна

Иващенко Дарья Владимировна

Аспирант

Московский государственный университет имени М.В.Ломоносова, Факультет фундаментальной медицины, Кафедра акушерства и гинекологии, Москва, Россия

E-mail: nanoty_1@icloud.com

Несмотря на развитие вспомогательных репродуктивных технологий, доля успешных протоколов переносов эмбрионов при ЭКО сохраняется на достаточно низком уровне - не более 40% [1.РАРЧ отчет 2022г]. Согласно статистическим данным, приблизительно 30% неудач имплантации происходят в виду эмбрионального фактора (анэуплоидии, моногенные мутации и тд.), остальные 70% приходятся на долю материнского фактора [Coughlan С., 2014, с.14–38]. Настоящее исследование посвящено изучению предикторов неудач имплантации при исключенном эмбриональном факторе с применением методов машинного обучения.

Цель исследования: разработка и валидация моделей машинного обучения для предсказания раннего исхода ЭКО (наступление клинической беременности) на когорте из 392 пациенток с исключенным эмбриональным фактором.

Были собраны данные медицинской документации 392 пациенток, проходивших процедуру ЭКО, на основе протоколов криоконсервированных эмбрионов высокого качества, прошедших ПГТ-а с 2022 по 2025гг на базе института репродуктивной медицины REMEDI. На этапе сбора датасета было отобрано 70 различных признаков, включая возраст пациенток, антропометрические и лабораторные показатели, репродуктивный анамнез и анамнез ВРТ. Затем была обучена модель Random Forest с двухэтапной процедурой отбора признаков. На первом этапе методом feature importance были ранжированы все признаки, после чего автоматически подбиралось оптимальное их число по максимуму F1-score при 5-кратной стратифицированной кросс-валидации. На втором этапе на оптимальном наборе признаков проводилась оптимизация гиперпараметров. Основной метрикой оценки был F1-score, дополнительно рассчитывались Accuracy, Precision и Recall. Значимость предикторов была подтверждена методом permutation importance.

Для раннего исхода ЭКО было определено 45 признаков (F1-score 0.750, Accuracy=0.704), что демонстрирует стабильную предсказательную способность модели. Наиболее значимыми предикторами оказались гормональные показатели – уровни ЛГ, ФСГ, пролактина в раннюю фолликулярную фазу естественного цикла, уровень ТТГ и Т4 на момент переноса эмбриона, а также качество эмбриона, что согласуется с клиническими данными.

Таким образом, разработанная модель машинного обучения обеспечивает стабильное предсказание ранних исходов ЭКО с F1-score 0.73-0.76 при использовании 30-45 наиболее значимых признаков из 70 доступных, демонстрируя при этом устойчивость к переобучению. Полученные результаты обеспечивают воспроизводимую основу для персонализированного подбора протоколов ЭКО, позволяя оптимизировать распределение клинических ресурсов и повышать эффективность лечения бесплодия.

Источники и литература

- 1) Регистр ВРТ Общероссийской общественной организации «Российская Ассоциация Репродукции Человека». Отчет за 2022 год
- 2) Coughlan C., Ledger W., Wang Q., D'Souza R., Quenby S., Li T.C. Recurrent implantation failure: definition and management. Reproductive BioMedicine Online. 2014. Vol. 28, No. 1. P. 14–38.