

ПРОГНОЗИРОВАНИЕ ИСХОДОВ ПРОГРАММ ЭКСТРАКОРПОРАЛЬНОГО ОПЛОДОТВОРЕНИЯ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ «СЛУЧАЙНЫЙ ЛЕС»

Г. М. Владимирский¹, М. А. Журавлева¹, А. Э. Дашиева²✉, И. Е. Корнеева², Т. А. Назаренко²

¹ Национальный исследовательский университет «Высшая школа экономики», Москва, Россия

² Национальный медицинский исследовательский центр акушерства, гинекологии и перинатологии имени В. И. Кулакова, Москва, Россия

Программы экстракорпорального оплодотворения (ЭКО) с переносом эмбрионов в настоящее время являются основными методами лечения всех форм бесплодия, однако успешные случаи, заканчивающиеся наступлением беременности, по-прежнему составляют лишь треть из числа всех проведенных циклов. Для оценки результатов лечения бесплодия методом ЭКО необходимо учитывать множество показателей и изучать взаимосвязь между ними. За последние десятилетия был разработан ряд моделей прогнозирования ЭКО для оценки индивидуальных результатов лечения, но лишь немногие из них оказались клинически значимыми, из-за их плохой прогностической способности. Целью исследования было создать нелинейные модели прогнозирования исходов ЭКО и выявить наиболее значимые факторы, влияющие на результат лечения. На основе базы данных, содержащей информацию по более чем 700 показателям 7004 женщин в возрасте от 18 лет до 45 лет, прошедших лечение в программе ЭКО на территории Российской Федерации с 2010 по 2020 г., была обучена модель «Случайного леса», прогнозирующая беременность в цикле ЭКО с ROC-AUC = 0,69. Описаны 20 самых успешных предикторов полученной модели с интерпретацией их вклада в прогнозирование. Из них индекс массы тела (ИМТ), число полученных и оплодотворившихся ооцитов описаны в научной литературе как предикторы исхода программ ЭКО, в то время как другие признаки, например анамнестические данные, количество и исходы предыдущих программ ЭКО, а также сывороточная концентрация АМГ, редко фигурируют в зарубежных прогностических моделях.

Ключевые слова: ЭКО, бесплодие, прогностическая модель, «Случайный лес»

Вклад авторов: Г. М. Владимирский — обучение прогностических моделей, анализ литературы, выбор методов исследования; М. А. Журавлева — предобработка и анализ данных, анализ литературы, написание рукописи; А. Э. Дашиева — обработка исходного материала, анализ результатов; И. Е. Корнеева, Т. А. Назаренко — разработка анкеты для базы данных, редактирование рукописи.

✉ **Для корреспонденции:** Аюна Эрдэмовна Дашиева
ул. Академика Опарина, д. 4Б, г. Москва, 117198, Россия, dr.dashieva@mail.ru

Статья получена: 24.11.2023 **Статья принята к печати:** 19.12.2023 **Опубликована онлайн:** 31.12.2023

DOI: 10.24075/vrgmu.2023.057

PREDICTING THE OUTCOMES OF *IN VITRO* FERTILIZATION PROGRAMS USING A RANDOM FOREST MACHINE LEARNING MODEL

Vladimirsky GM¹, Zhuravleva MA¹, Dashieva AE²✉, Korneeva IE², Nazarenko TA²

¹ Higher School of Economics National Research University, Moscow, Russia

² Kulakov National Medical Scientific Centre for Obstetrics, Gynecology and Perinatal Medicine, Moscow, Russia

Currently, *in vitro* fertilization (IVF) with embryo transfer is the main method of treatment of all forms of infertility, but successful cases ending in pregnancy still account for only a third of all cycles performed. It is necessary to take into account many parameters and investigate the connections between them in order to properly evaluate the results of IVF. Over the past decades, a number of IVF prediction models have been developed with the aim at assessing the outcomes in individual cases, but, given the generally poor prognostic capacity, only a few of them have proven to be clinically significant. This study aimed to create nonlinear IVF outcomes prediction models and identify the most significant factors affecting the said outcomes. Using a database containing information on more than 700 indicators of 7004 women aged 18 to 45 years who participated in the IVF program in Russia from 2010 to 2020, we trained a random forest model that predicted pregnancy in the IVF cycle with ROC-AUC = 0.69. This paper describes 20 most successful predictors of the resulting model and interprets their contribution to the prognosis. Of these, body mass index (BMI) and the number of received and fertilized oocytes have been covered in the scientific literature previously as predictors of IVF outcomes, but other parameters, such as anamnestic data, previous participation in an IVF program (number of cases and their results), as well as serum concentration of AMH, rarely appear in foreign prognostic models.

Keywords: IVF, infertility, prognostic model, random forest

Author contribution: Vladimirsky GM — predictive models training, literature analysis, choice of research methods; Zhuravleva MA — preprocessing and analysis of data, literature analysis, manuscript authoring; Dashieva AE — processing of source material, analysis of results; Korneeva IE, Nazarenko TA — development of the survey for the database, manuscript editing.

✉ **Correspondence should be addressed:** Ayuna E. Dashieva,
Akademika Oparina, 4B, Moscow, 117198, Russia, dr.dashieva@mail.ru

Received: 24.11.2023 **Accepted:** 19.12.2023 **Published online:** 31.12.2023

DOI: 10.24075/brsmu.2023.057

Бесплодие — проблема, затрагивающая десятки миллионов семей. Развитие вспомогательных репродуктивных технологий (ВРТ) принесло надежду таким парам стать родителями. Согласно отчету РАРЧ в 2020 г. в Российской Федерации (РФ) было выполнено 148 660 циклов ВРТ, родилось около 34 250 детей. Однако, несмотря на удовлетворенность потребности населения в этом методе лечения бесплодия, частота наступления клинической беременности в расчете на один перенос эмбриона составляет 34,8% [1]. Поскольку вероятность успеха

ЭКО зависит от многих факторов, усложняющих оценку результативности циклов, разработка инструмента принятия решений, основанного на анализе этих факторов, могла бы улучшить качество медицинской помощи и консультирования пациенток, проходящих лечение в программе ЭКО.

В научной литературе предложен ряд моделей машинного обучения, прогнозирующих исходы ЭКО и предлагающих идентификацию наиболее важных для предсказания характеристик женщин и протокола программы [2].

Наиболее распространенный подход к прогнозированию результатов ЭКО — использование линейных моделей. В обзоре 2020 г. выявлено 35 таких моделей, все из них основаны либо на логистической регрессии, либо на регрессии Кокса [3]. Часто в таких исследованиях качество моделей не оценивают, хотя для этого существуют предпочтительные способы, например ROC-AUC и с-статистика. Для качественной оценки прогностических моделей используют ROC-анализ с построением ROC-кривой (от англ. Receiver operating characteristic). ROC-анализ заключается в построении четырехпольных таблиц с оценкой чувствительности и специфичности модели. ROC-кривая — график, позволяющий дать оценку качеству модели по разделению двух классов. По оси ординат отложены значения частоты истинно положительных результатов (чувствительность), по оси абсцисс — частоты ложноположительных результатов (специфичность). Значения по осям соответствуют значениям от 0 до 1 (т. е. от 0 до 100%). Полученная кривая показывает зависимость верно классифицированных положительных примеров от количества неверно классифицированных отрицательных примеров. Для идеального классификатора график ROC-кривой проходит через верхний левый угол, где доля истинно положительных случаев составляет 1,0 или 100% (идеальная чувствительность), а доля ложноположительных примеров равна 0. Кроме графика ROC-кривой, для оценки качества модели используют характеристику площади под кривой ROC-AUC (Area under ROC). Чем больше показатель AUC, тем лучшей прогностической силой обладает модель. Чаще показатель AUC предназначен для сравнительного анализа нескольких моделей. Значения ROC-AUC при прогнозировании результатов ЭКО в литературных источниках варьируют от 0,58 до 0,73 [3–12].

Линейные прогностические модели успеха ЭКО обычно включают в себя около семи признаков. Наиболее распространенными оказываются возраст женщины, причины бесплодия, исход предыдущих беременностей и программ ЭКО, число полученных ооцитов и перенесенных эмбрионов и полученных ооцитов [4, 5, 9–11]. Некоторые исследователи предполагают, что ограниченное число признаков, переходящих от исследования к исследованию, приводит к скромной прогностической способности моделей, и выступают за выявление новых факторов, связанных с успехами и неуспехами процедуры [13].

Несмотря на распространенность моделей логистической регрессии, они имеют ряд недостатков. Так, например, в ряде исследований было обнаружено, что взаимосвязь между некоторыми ключевыми характеристиками, такими как возраст женщины, количество полученных яйцеклеток, год начала лечения, и успехом ЭКО нелинейна [10, 11]. В таких случаях возможно применение функции кубического сплайна с целью интерполяции данных (например, возраста) для придания линейным моделям нелинейности либо полиномиальное преобразование данных [8, 10, 11]. Тем не менее, подобные модификации линейных моделей все еще основаны на простой (полиномиальной) зависимости между целевой переменной и признаками.

Кроме того, модели логистической регрессии, будучи интерпретируемыми, не имеют высокой прогностической способности. Поэтому многие исследователи обратились к нелинейным неинтерпретируемым моделям машинного обучения, которые охватывают такие методы, как «Случайный лес», градиентный бустинг и нейронные сети. Среди этих методов «Случайный лес» и градиентный бустинг часто рассматривают как самые современные для

задач бинарной классификации с табличными данными, поскольку они, как правило, превосходят другие методы с точки зрения точности и обобщения [2]. Как правило, площадь под кривой ROC для подобных моделей, колеблется от 0,68 до 0,86, что выше, чем у линейных классификаторов [14–16].

Основное ограничение в использовании нелинейных неинтерпретируемых моделей — сложность оценки вклада отдельных признаков в предсказание моделей. Тем не менее, разработанные в последние годы методы предлагают возможность интерпретации признаков для любых моделей машинного обучения вне зависимости от их сложности. В настоящем исследовании применен метод SHAP, основанный на значении Шепли, концепции из кооперативной теории игр. В нем аппроксимированное значение Шепли (средний вклад признака во все коалиции признаков) использовано для вычисления вклада каждого признака в прогноз модели [17], что позволит более точно прогнозировать исходы программ ЭКО.

Цель настоящего исследования — построить нелинейные модели прогнозирования исходов ЭКО и выявить наиболее значимые факторы, влияющие на результат лечения.

ПАЦИЕНТЫ И МЕТОДЫ

Клинический материал исследования

Для построения модели использовали данные, включающие характеристики и результаты программ ЭКО 7004 женщин, прошедших лечение в 17 клиниках ВРТ на территории РФ в период с 2011 по 2020 г. Критерии включения: возраст женщин от 18 лет до 45 лет с любыми факторами бесплодия (N97). Критерии исключения: противопоказания для проведения ВРТ и вынашивания беременности (согласно Приказу МЗ РФ от 31 июля 2020 г. № 803н «О порядке использования вспомогательных репродуктивных технологий, противопоказаниях и ограничениях к их применению»).

Распределение клиник, участвующих в исследовании в субъектах РФ, представлено на рис. 1. Сбор материала осуществляли путем заполнения специалистами клиник специально разработанных анкет, содержащих 770 признаков. Получаемые данные содержали несколько блоков информации: социальные характеристики пациентов (124 вопроса), данные анамнеза (171 вопроса), куда вошли данные о состоянии соматического здоровья (58 вопросов), состоянии гинекологического здоровья (108 вопросов), истории бесплодия и методов лечения (73 вопроса), данные лабораторного обследования (6 пунктов), данные о партнере пациентки (210 вопросов), данные протокола овариальной стимуляции (7 вопросов) и эмбриологического этапа (30 вопросов), поддержки лютеиновой фазы, исхода программы ЭКО.

Обработка и анализ данных

Предобработка данных для модели включала в себя отбор минимального значения среди нескольких анализов уровня гормонов в сыворотке крови (антимюллеровый гормон, или АМГ; фолликулостимулирующий гормон, или ФСГ; лютеинизирующий гормон, или ЛГ; тиреотропный гормон, ТТГ; пролактин). Разреженные и дублирующиеся данные были удалены, что сократило число признаков до 408. Пропуски в данных были заполнены средними



Рис. 1. Географический охват исследования

значениями. Для статистического анализа использовали отношение шансов (ОШ) [18]. Расчет p -value по полученному отношению шансов производили в соответствии с источником [19].

Отбор и интерпретация важности признаков

В данном исследовании применяли метод «Случайного леса» (Random forest) — метод машинного обучения, использующий ансамбль деревьев решений (decision tree) для задач классификации. Каждое отдельное дерево в таком лесу дает предсказание класса, и набравший наибольшее количество голосов класс становится предсказанием. Данная работа направлена на предсказание наступления беременности после процедуры программы ЭКО.

Для интерпретации важности признаков после использования модели «Случайного леса» использовали

показатель важности Джини. Он позволяет сравнить распределение признака в выборке с различным числом единиц [20]. Использованная для выделения признаков модель имела гиперпараметры (параметры, настраиваемые вручную перед обучением модели), максимизирующие значение метрики ROC-AUC на пятикратной кросс-валидации, и была впоследствии обучена на полном датасете (наборе данных в машинном обучении). С целью отбора оптимального числа признаков для предсказания использовали метод рекурсивного отбора признаков с пятикратной кросс-валидацией, в рамках которого на каждом этапе из признаков модели удалял признак, имеющий минимальную важность Джини в модели. Все вышеперечисленные методы были использованы в реализации библиотеки scikit-learn [21]. В рамках расширенной интерпретации признаков задействовали метод SHAP [17], подходящий для интерпретации значимости признаков для нелинейных моделей.

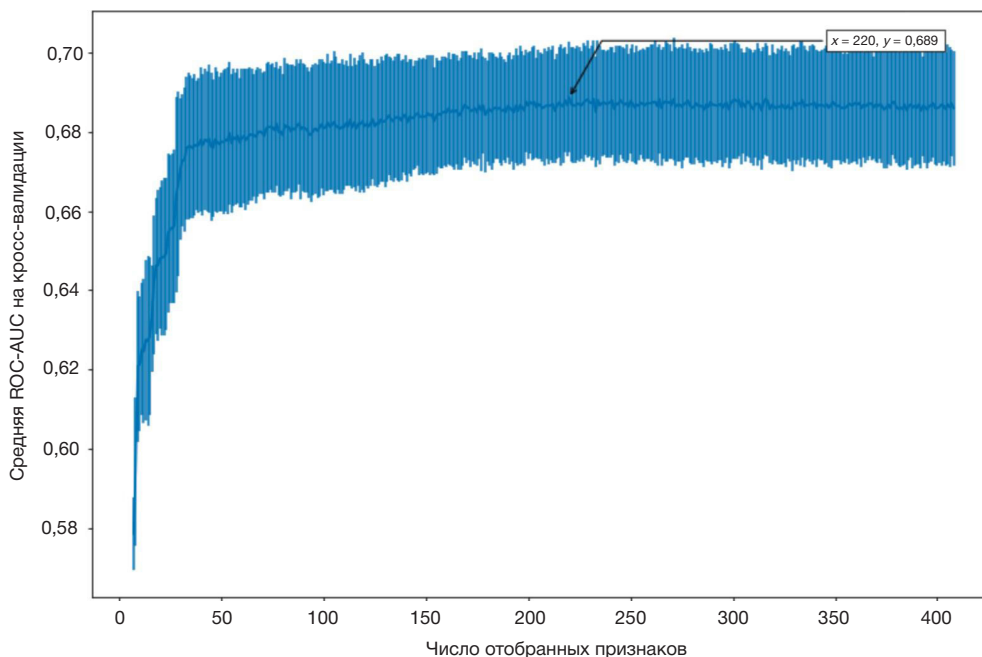


Рис. 2. График среднего ROC-AUC на кросс-валидации при элиминации признаков. На графике отмечена точка с максимальным ROC-AUC, а также среднее квадратичное отклонение ROC-AUC в кросс-валидации

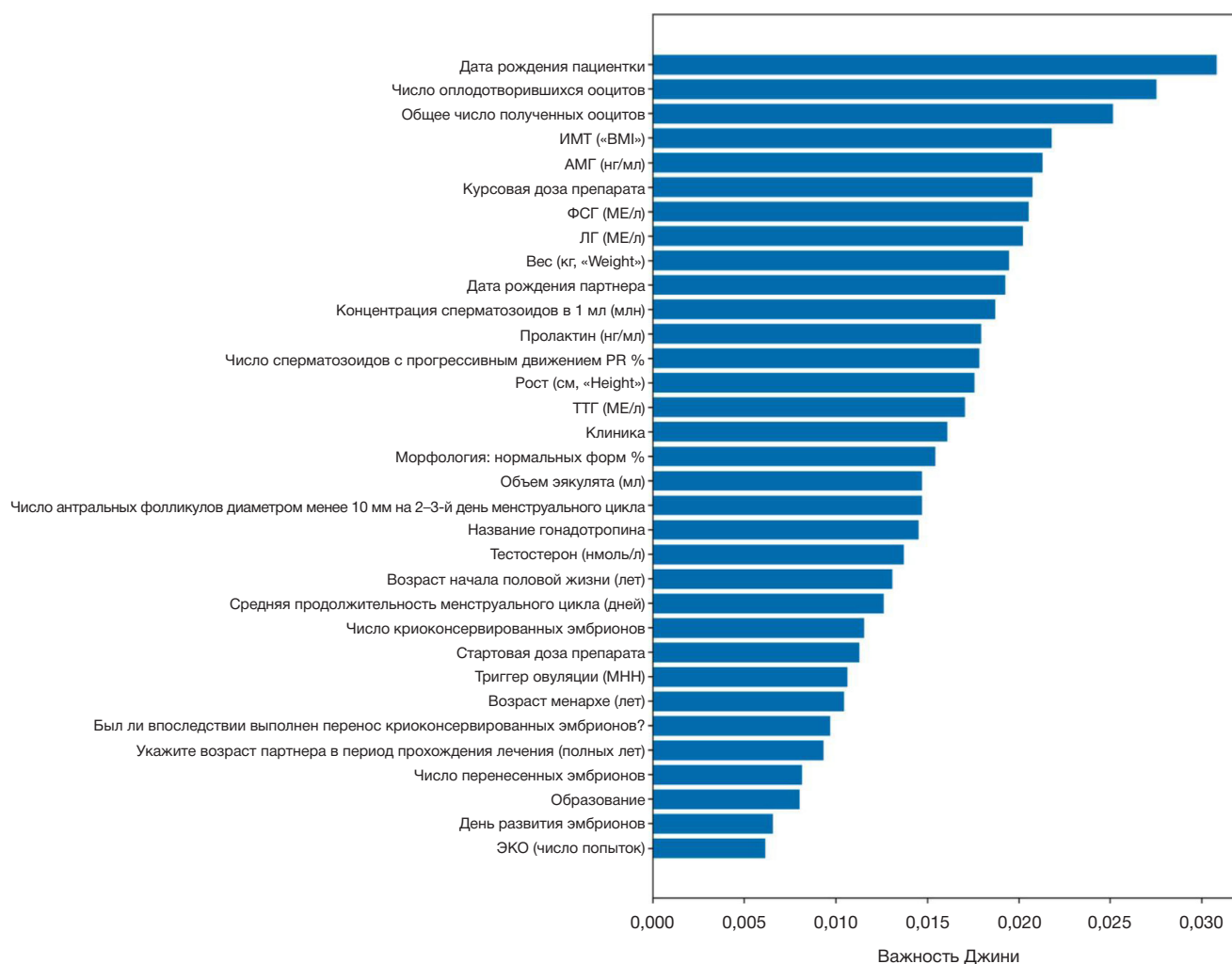


Рис. 3. Важность Джини для 33 лучших признаков, полученная при обучении «Случайного леса» на 220 признаках

Используемые модели

Ввиду наличия большого количества бинарных и категориальных признаков, а также существования нелинейных зависимостей между признаками и целевой переменной, в качестве основного классификатора была выбрана модель «Случайного леса», реализованная в библиотеке scikit-learn [21]. Параметры модели были выбраны с использованием метода GridSearch с пятикратной кросс-валидацией [21], оценка качества модели происходила с использованием показателя ROC-AUC, который менее чувствителен к дисбалансу классов, наблюдаемому в данных. Наилучшими параметрами для модели «Случайного леса» оказались максимальная глубина, равная 50, минимальное число объектов в листе, равное 2, и общее число деревьев, равное 2000. Кроме того, в этом исследовании была протестирована модель классификатора Catboost [22]. Catboost был выбран из-за встроенной поддержки категориальных признаков, отличающей его от других реализаций алгоритма градиентного бустинга. Роль целевой переменной для всех обученных моделей играло наличие беременности у женщины.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Использование рекурсивного отбора признаков показало, что максимального значения метрики ROC-AUC можно

достигнуть при обучении «Случайного леса» на 220 признаках, и оно составляет 0,69. Максимальный ROC-AUC Catboost на кросс-валидации составил 0,68, поэтому для дальнейшего анализа использовали более удобную в интерпретации модель «Случайного леса». Графическая иллюстрация изменения метрики ROC-AUC при постепенном удалении признаков для модели «Случайного леса» представлена на рис. 2.

При постепенном удалении признаков значение метрики в начале изменяется незначительно, начиная резко падать только при количестве признаков меньше 33. Поэтому в качестве оптимального количества признаков в модели было выбрано число 33, а модель, обученная на этих признаках, позволяет добиться значения ROC-AUC, равного 0,69.

Важность 20 признаков, дающих наибольший вклад в предсказание полученной модели, была проанализирована при помощи критерия важности Джини (рис. 3).

Наиболее значимыми для предсказания признаками оказались дата рождения пациентки (возраст), количество оплодотворившихся ооцитов и общее число полученных ооцитов — показатели, сопоставимые с международными исследованиями [16].

Уровни АМГ в сыворотке крови имели наибольший вес в предсказании по сравнению с другими показателями содержания гормонов, однако в международных исследованиях этот признак используют значительно реже, чем параметры гонадотропных гормонов (ФСГ

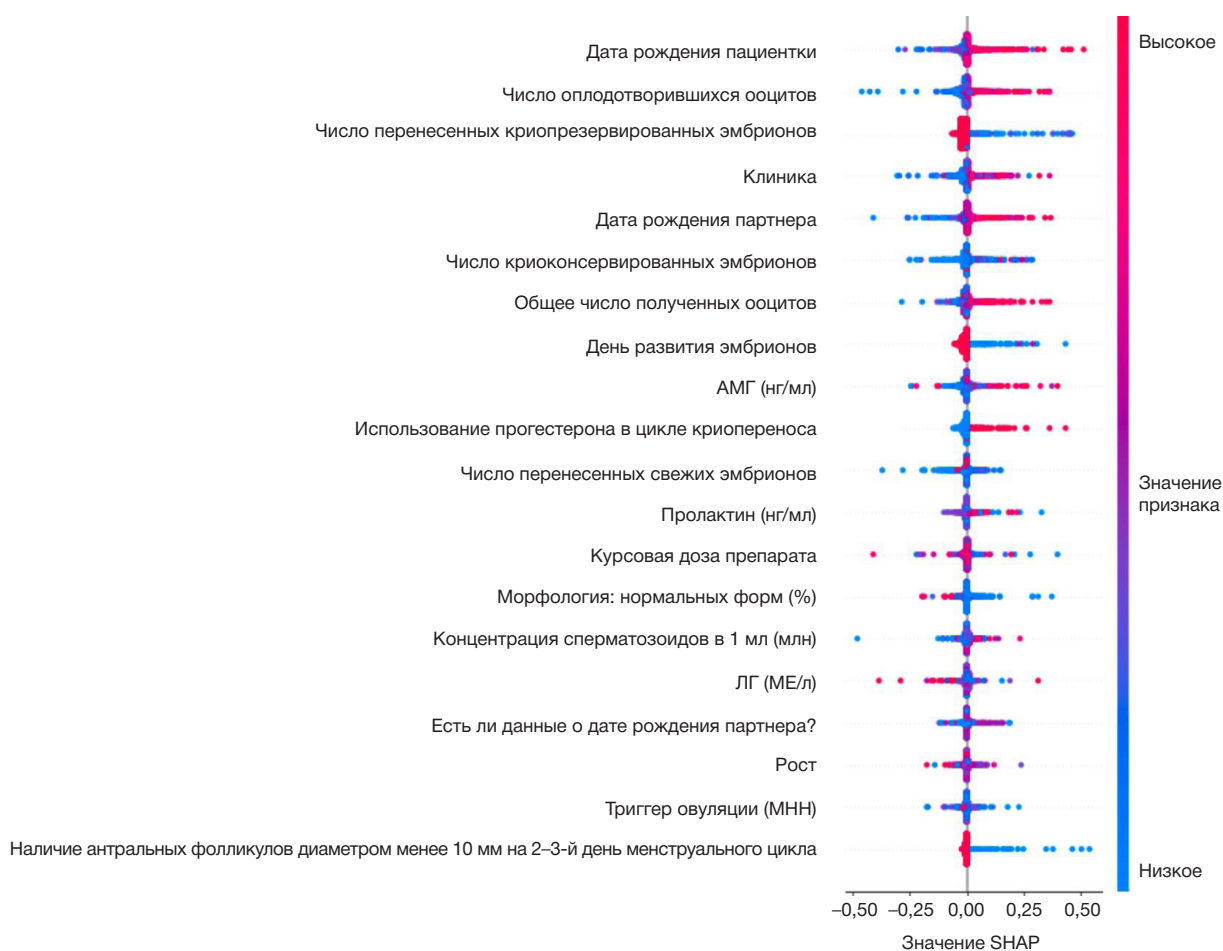


Рис. 4. Диаграмма, отражающая значения SHAP для 20 самых важных, в соответствии с методом SHAP, признаков. Положительный класс означает успех ЭКО, а отрицательный — неудачу процедуры

и ЛГ [8]. Заметный вклад в предсказание вносила также клиника, в которой проводили ЭКО.

Методом SHAP был проведен дополнительный анализ 20 лучших признаков. Увеличение значений таких признаков, как количество оплодотворенных ооцитов, дата рождения пациентки, уровень АМГ, использование прогестерона в лютеиновой фазе цикла, увеличивают вероятность предсказания успеха ЭКО моделью (рис. 4).

Нами были построены модели, предсказывающие исход программы ЭКО для отдельных диагнозов бесплодия. Полученные значения ROC-AUC не превосходили значения метрики, полученной для всей выборки, из чего можно сделать вывод, что использование моделей для отдельных типов бесплодия нецелесообразно (таблица).

У женщин с безуспешными программами в прошлом вероятность успешного исхода ЭКО ниже, чем у женщин с первой попыткой ЭКО или с прошлыми успешными попытками (ОШ = 0,7675; $p < 0,0001$), поэтому неудивительно, что переменная, показывающая число прошлых попыток, фигурирует в значимых признаках, отобранных «Случайным лесом» — классификатором, показавшим лучшие результаты при кросс-валидации. Тем не менее, мы считаем модель, описанную нами, более релевантной для российской популяции, чем зарубежные модели, описанные в литературных источниках.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Согласно полученным нами данным, качество текущей прогностической модели «Случайного леса» (ROC-

AUC = 0,69) было сопоставимо с качеством аналогичных зарубежных моделей. Так лучшая модель в недавнем исследовании показала качество ROC-AUC 0,68 [14].

Несмотря на сопоставимость значений ROC-AUC нашей модели и моделей, описанных в литературе, даже если эти модели имеют одинаковую целевую переменную они, чаще всего, включают в себя различные критерии отбора пар. Так, часть моделей, представленных в литературе, не содержит информации о прошлых попытках ЭКО женщины, и предполагается только вероятность успеха в первом цикле лечения [14]. Полученная нами модель ЭКО опирается на результаты предыдущих программ. В обучающих данных, использованных нами, 40,9% женщин имели в анамнезе безуспешные попытки ЭКО. Преимуществом нашей модели является возможность предсказывать вероятность успеха ЭКО для одного цикла. В то время как некоторые модели, опубликованные в литературе, предсказывают кумулятивный успех для нескольких циклов ЭКО [10, 11].

Все эти факторы делают сравнение моделей по численным показателям не до конца объективными. Например, можно предположить, что текущая модель имеет ROC-AUC выше, чем модель, описанная ранее [14], поскольку включает в себя информацию о предыдущих успехах ЭКО: так среди 40,9% женщин с прошлыми безуспешными попытками ЭКО, в 92,95% случаев следующие попытки ЭКО закончатся неудачами.

Оценивая результативность разработанной модели с клинических позиций, следует отметить, что нам удалось выделить наиболее важные предикторы успеха программы

Таблица. Результаты пятикратной кросс-валидации модели «Случайного леса» для отдельных диагнозов бесплодия при подборе гиперпараметров методом GridSearch

Диагноз бесплодия (МКБ-10)	Число пациентов	ROC-AUC
Женское бесплодие, связанное с отсутствием овуляции (N 97.0)	1279	0,69 ± 0,02
Женское бесплодие трубного происхождения (N 97.1)	3191	0,66 ± 0,02
Женское бесплодие маточного происхождения (N 97.2)	1248	0,66 ± 0,05
Женское бесплодие, связанное с мужским фактором (N 97.4)	3226	0,68 ± 0,01

ЭКО в популяции российских бесплодных пар. В числе этих признаков как хорошо известные (возраст женщины, число оплодотворенных ооцитов, число полученных ооцитов, ИМТ, уровень АМГ), так и не рассматриваемые ранее параметры. Например, было показано, что поддержка лютеиновой фазы препаратами прогестерона ассоциирована с успехом лечения. Хотя назначение прогестерона — рутинная клиническая тактика, до настоящего времени математически доказанного обоснования необходимости поддержки лютеиновой фазы индуцированного цикла не существовало. Кроме того, объективно подтверждено негативное влияние предыдущих неуспешных попыток ЭКО на последующую. Этот факт обуславливает, по всей видимости, необходимость пересмотра тактики лечения при нескольких (в настоящем исследовании — четырех) неудачных попытках ЭКО. Несмотря на то что на сегодняшний день ЭКО — рутинный метод преодоления бесплодия, и казалось бы, все клиники работают по стандартным протоколам и технологиям, модель выявила различия в результатах лечения в зависимости от клиники, предоставившей данные, что может стать основанием для клиник провести анализ своей работы. Интересным оказался факт отсутствия зависимости исходов программ

ЭКО от фактора бесплодия при наличии идентичности остальных, значимых параметров, что противоречит многочисленным исследованиям, пытающимся связать исходы ЭКО с нозологической формой бесплодия.

Выводы

За последние десятилетия был разработан ряд моделей прогнозирования ЭКО для оценки индивидуальных результатов лечения, но лишь немногие из них оказались клинически значимыми из-за их недостаточной прогностической способности и используемых статистических методов. Машинное обучение, позволяющее интерпретировать данные и строить прогностические модели, все чаще применяют в клинической практике, особенно в сложных системах с несколькими переменными. В настоящем исследовании построена модель, предсказывающая исход циклов ЭКО с удовлетворительной эффективностью прогнозирования. Выделены важные факторы результативности ЭКО, обнаружены взаимодействия между ними. Мы продолжим работу по практическому применению модели для оценки влияния переменных на результативность лечения.

Литература

1. Корсак В. С., Смирнова А. А., Шурыгина О. В. Регистр ВРТ Общероссийской общественной организации «Российская ассоциация репродукции человека». Отчет за 2020 год. Проблемы репродукции. 2022; 28 (6): 12–27. DOI: 10.17116/repro20222806112.
2. Hassan MR, Al-Insaf S, Hossain MI, Kamruzzaman J. A machine learning approach for prediction of pregnancy outcome following IVF treatment. *Neural Computing and Applications*. 2018; 32: 2283–97.
3. Ratna MB, Bhattacharya S, Abdulrahim B, McLernon DJ. A systematic review of the quality of clinical prediction models in in vitro fertilisation. *Human Reproduction*. 2020; 35: 100–116.
4. Lintsen A, Eijkemans M, Hunault C, Bouwmans C, Hakkaart L, Habbema J, et al. Predicting ongoing pregnancy chances after IVF and ICSI: A national prospective study. *Human Reproduction*. 2007; 22: 2455–62.
5. van Weert JM, Repping S, van der Steeg JW, Steures P, van der Veen F, Mol BW. A prediction model for ongoing pregnancy after in vitro fertilization in couples with male subfertility. *J Reprod Med*. 2008; 53 (4): 250–6. PMID: 18472647.
6. Jones CA, Christensen AL, Salihi H, Carpenter W, Petrozzino J, Abrams E, et al. Prediction of individual probabilities of livebirth and multiple birth events following in vitro fertilization (IVF): A new outcomes counselling tool for IVF providers and patients using HFEA metrics. *Journal of experimental clinical assisted reproduction*. 2011; 8: 3.
7. Nelson SM, Lawlor DA. Predicting live birth, preterm delivery, and low birth weight in infants born from in vitro fertilisation: A prospective study of 144,018 treatment cycles. *PLoS Medicine*. 2011; 8: e1000386.
8. van Loendersloot LL, van Wely M, Repping S, Bossuyt PMM, van der Veen F. Individualized decision-making in IVF: Calculating the chances of pregnancy. *Human Reproduction*. 2013; 28: 2972–80.
9. Dhillon R, McLernon D, Smith P, Fishel S, Dowell K, Deeks J, et al. Predicting the chance of live birth for women undergoing IVF: A novel pretreatment counselling tool. *Human Reproduction*. 2015; 31: 84–92.
10. McLernon DJ, Steyerberg EW, te Velde ER, Lee AJ, Bhattacharya S. Predicting the chances of a live birth after one or more complete cycles of in vitro fertilisation: Population based study of linked cycle data from 113 873 women. *BMJ*. 2016; p. i5735.
11. McLernon DJ, Raja EA, Toner JP, Baker VL, Doody KJ, Seifer DB. Predicting personalized cumulative live birth following in vitro fertilization. *Fertility and Sterility*. 2022; 117 (2): 326–38.
12. Vaegter KK, Lakic TG, 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. *Fertility and Sterility*. 2017; 107 (3): 641–48.
13. Ottosen LDM, Kesmodel U, Hindkjaer J, Ingerslev HJ. Pregnancy prediction models and eSET criteria for IVF patients—do we need more information? *Journal of assisted reproduction and genetics*.

- 2007; 24 (1): 29–36.
14. Xu T, d. F. Veiga A, Hammer KC, Paschalidis ICh, Mahalingaiah Sh. Informative predictors of pregnancy after first IVF cycle using eIVF practice highway electronic health records. *Scientific Reports*. 2022; 12: 1–9.
 15. C N Barreto N, Castro GZ, Pereira RG, Pereira FAN, Reis FM, Junior WM, et al. Predicting in vitro fertilization success in the Brazilian public health system: a machine learning approach. *Med Biol Eng Comput*. 2022; 60 (7): 1851–61. DOI: 10.1007/s11517-022-02569-1. Epub 2022 May 4. PMID: 35508786.
 16. Barnett-Itzhaki Z, Elbaz M, Buttermann R, Amar D, Amitay M, Racowsky C, et al. Machine learning vs. classic statistics for the prediction of IVF outcomes. *Journal of Assisted Reproduction and Genetics*. 2020; 37: 2405–12.
 17. Lundberg S, Lee S-I. A unified approach to interpreting model predictions. Available from: <https://arxiv.org/abs/1705.07874>, may 22 2017.
 18. Morris JA, Gardner MJ. *Statistics in Medicine: Calculating confidence intervals for relative risks (odds ratios) and standardised ratios and rates*. BMJ. 1988; 296: 1313–6.
 19. Sheskin DJ. *Handbook of parametric and nonparametric statistical procedures*. Chapman and Hall/CRC, 2003.
 20. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and regression trees*. Routledge, 2017.
 21. Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*. 2011; 12: 2825–30.
 22. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulina A. Catboost: Unbiased boosting with categorical features. 2017. Available from: <https://arxiv.org/abs/1706.09516>.

References

1. Korskak VS, Smirnova AA, Shurygina OV. Registr VRT Obshherossijskoj obshhestvennoj organizacii «Rossijskaja asociacija reprodukcii cheloveka». Otchet za 2020 god. *Problemy reprodukcii*. 2022; 28 (6): 12–27. DOI: 10.17116/repro20222806112. Russian.
2. Hassan MR, Al-Insaf S, Hossain MI, Kamruzzaman J. A machine learning approach for prediction of pregnancy outcome following IVF treatment. *Neural Computing and Applications*. 2018; 32: 2283–97.
3. Ratna MB, Bhattacharya S, Abdulrahim B, McLernon DJ. A systematic review of the quality of clinical prediction models in in vitro fertilisation. *Human Reproduction*. 2020; 35: 100–116.
4. Lintsen A, Eijkemans M, Hunault C, Bouwmans C, Hakkaert L, Habbema J, et al. Predicting ongoing pregnancy chances after IVF and ICSI: A national prospective study. *Human Reproduction*. 2007; 22: 2455–62.
5. van Weert JM, Repping S, van der Steeg JW, Steures P, van der Veen F, Mol BW. A prediction model for ongoing pregnancy after in vitro fertilization in couples with male subfertility. *J Reprod Med*. 2008; 53 (4): 250–6. PMID: 18472647.
6. Jones CA, Christensen AL, Salihi H, Carpenter W, Petrozzino J, Abrams E, et al. Prediction of individual probabilities of livebirth and multiple birth events following in vitro fertilization (IVF): A new outcomes counselling tool for IVF providers and patients using HFEA metrics. *Journal of experimental clinical assisted reproduction*. 2011; 8: 3.
7. Nelson SM, Lawlor DA. Predicting live birth, preterm delivery, and low birth weight in infants born from in vitro fertilisation: A prospective study of 144,018 treatment cycles. *PLoS Medicine*. 2011; 8: e1000386.
8. van Loendersloot LL, van Wely M, Repping S, Bossuyt PMM, van der Veen F. Individualized decision-making in IVF: Calculating the chances of pregnancy. *Human Reproduction*. 2013; 28: 2972–80.
9. Dhillon R, McLernon D, Smith P, Fishel S, Dowell K, Deeks J, et al. Predicting the chance of live birth for women undergoing IVF: A novel pretreatment counselling tool. *Human Reproduction*. 2015; 31: 84–92.
10. McLernon DJ, Steyerberg EW, te Velde ER, Lee AJ, Bhattacharya S. Predicting the chances of a live birth after one or more complete cycles of in vitro fertilisation: Population based study of linked cycle data from 113 873 women. *BMJ*. 2016; p. i5735.
11. McLernon DJ, Raja EA, Toner JP, Baker VL, Doody KJ, Seifer DB. Predicting personalized cumulative live birth following in vitro fertilization. *Fertility and Sterility*. 2022; 117 (2): 326–38.
12. Vaegter KK, Latic TG, 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. *Fertility and Sterility*. 2017; 107 (3): 641–48.
13. Ottosen LDM, Kesmodel U, Hindkjaer J, Ingerslev HJ. Pregnancy prediction models and eSET criteria for IVF patients—do we need more information? *Journal of assisted reproduction and genetics*. 2007; 24 (1): 29–36.
14. Xu T, d. F. Veiga A, Hammer KC, Paschalidis ICh, Mahalingaiah Sh. Informative predictors of pregnancy after first IVF cycle using eIVF practice highway electronic health records. *Scientific Reports*. 2022; 12: 1–9.
15. C N Barreto N, Castro GZ, Pereira RG, Pereira FAN, Reis FM, Junior WM, et al. Predicting in vitro fertilization success in the Brazilian public health system: a machine learning approach. *Med Biol Eng Comput*. 2022; 60 (7): 1851–61. DOI: 10.1007/s11517-022-02569-1. Epub 2022 May 4. PMID: 35508786.
16. Barnett-Itzhaki Z, Elbaz M, Buttermann R, Amar D, Amitay M, Racowsky C, et al. Machine learning vs. classic statistics for the prediction of IVF outcomes. *Journal of Assisted Reproduction and Genetics*. 2020; 37: 2405–12.
17. Lundberg S, Lee S-I. A unified approach to interpreting model predictions. Available from: <https://arxiv.org/abs/1705.07874>, may 22 2017.
18. Morris JA, Gardner MJ. *Statistics in Medicine: Calculating confidence intervals for relative risks (odds ratios) and standardised ratios and rates*. BMJ. 1988; 296: 1313–6.
19. Sheskin DJ. *Handbook of parametric and nonparametric statistical procedures*. Chapman and Hall/CRC, 2003.
20. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and regression trees*. Routledge, 2017.
21. Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*. 2011; 12: 2825–30.
22. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulina A. Catboost: Unbiased boosting with categorical features. 2017. Available from: <https://arxiv.org/abs/1706.09516>.