

**Прогнозирование результатов хирургического лечения
атеросклероза на основе анализа клинических
и иммунологических данных**

*Кузнецов М. Р., Туркин П. Ю., Воронцов К. В.,
Дьяконов А. Г., Ивахненко А. А., Сиваченко Е. А.*
pavelturkin@km.ru, voron@ccas.ru

Москва, Российский государственный медицинский университет;
Москва, Вычислительный центр РАН

Проблема облитерирующего атеросклероза артерий считается одной из наиболее важных в современной медицинской науке. Проявлениями данной патологии являются такие социально значимые заболевания, как ишемическая болезнь сердца, острый инфаркт миокарда, острое и хроническое нарушение мозгового кровообращения и др. В возрасте 50 лет ими в той или иной форме страдают до 90%, а в возрасте свыше 60 лет — 100% населения. В структуре смертности атеросклероз и его проявления занимают основное место, являясь причиной 80% летальных исходов. Тем не менее, до сих пор не выявлены ключевые механизмы возникновения атеросклероза. Вызвано это отчасти тем, что фундаментальные и клинические изыскания в данной области практически не имеют точек соприкосновения. Клинические исследования в основном рассматривают вопросы терапии частных проявлений атеросклероза и практически не изучают основы патогенеза атеросклероза в целом.

Задача прогнозирования отдалённых последствий анастомоза

В данной работе решается задача прогнозирования результатов хирургического лечения атеросклероза, в ходе которого производится имплантации шунта с наложением соустьев (анастомозов) между артерией и протезом [1]. Рассматриваются только двузначные прогнозы — либо шунт приживётся, либо нет, т. е. прогнозирование сводится к классификации на два класса.

Данная задача классификации характеризуется большой размерностью (число признаков превышает число прецедентов), разнотипностью признаков, наличием пропусков в данных, низкой точностью измерения признаков. Перечисленные особенности свойственны многим задачам медицинской диагностики. Отличительной особенностью данной задачи является то, что данные собираются из двух независимых источников:

- *Клинические данные* [2]: скорости кровотока в зоне анастомоза, степень сужения линии анастомоза, свёртываемость, вязкость, агрегация тромбоцитов и эритроцитов, и др. Всего 18 признаков, измеряемых сразу после операции, спустя неделю, 1, 2, 3 и 6 месяцев.

- *Данные иммунологического обследования:* показатели местного иммунного статуса в поражённой зоне, показатели гуморального иммунитета, цитокинового статуса и т. д. Всего 17 признаков, измеряемых до и после операции.

Совместный анализ этих данных позволяет впервые в практике изучения атеросклероза оценить, какие иммунные факторы и особенности клеточных реакций в действительности оказывают значимое влияние на течение атеросклероза.

Задачи анализа данных

С точки зрения анализа данных цели исследования заключаются в следующем:

- Построить алгоритм прогнозирования (классификации), обладающий достаточно высокой точностью предсказания.
- Выявить значимые, нетривиальные, интерпретируемые закономерности и взаимосвязи между данными клинических и иммунологических обследований.

Для достижения указанных целей не достаточно только лишь выбрать наиболее точный (для данной конкретной задачи) метод классификации. Выбранный метод должен суметь «извлечь максимум выгоды» из совмещения клинических и иммунологических данных. Таким образом, появляется дополнительный критерий выбора модели классификации: алгоритм, построенный по объединённому набору признаков должен заметно превосходить по качеству прогнозов алгоритмы, построенные только по клиническим или только по иммунологическим данным.

Вычислительный эксперимент

Исходная выборка данных, собранная в клинике факультетской хирургии РГМУ, содержала описания 72 случаев. Долгосрочный результат операции был благоприятным в 54 и неблагоприятным в 18 случаях.

Тестировались следующие методы классификации:

- MLP — классический двухслойный персептрон;
- RBF — нейронная сеть с радиальными базисными функциями;
- SVM — метод опорных векторов;
- SVM-f — SVM с предварительной фильтрацией объектов-выбросов;
- DT — решающие деревья.

Для всех методов, кроме DT, выполнялся предварительный отбор признаков. Сначала все признаки были проранжированы по индивидуальной разделяющей способности. Затем были найдены парные корреляции признаков и все признаки были проранжированы по числу значимых корреляций с другими признаками. Далее производилась генерация

Алгоритм	Иммунологические	Клинические	Совмещённые
MLP	81 ± 5	76 ± 4	82 ± 4
RBF	78 ± 4	77 ± 4	71 ± 4
SVM	82 ± 3	78 ± 4	81 ± 3
SVM-f	87 ± 3	81 ± 3	84 ± 2
DT	87 ± 5	85 ± 6	88 ± 3

Таблица 1. Оценки частоты правильных прогнозов (%) по скользящему контролю для 5 алгоритмов классификации и 3 вариантов состава данных: только иммунологические, только клинические, и полные данные. Указаны 90%-е доверительные интервалы.

наборов признаков, причём признакам с высокой разделяющей способностью и ббльшим числом корреляций назначалась более высокая вероятность включения в набор. Для каждого из полученных наборов производилась оценка качества классификации по 10-кратному скользящему контролю (10-fold cross-validation). Описанная эвристическая стратегия направлена на построение информативных наборов признаков и одновременно на получение не сильно смещённых оценок качества классификации в условиях выборки малой длины.

Анализ результатов, представленных в Таблице 1, позволяет сделать следующие выводы. Предсказательная способность иммунологических данных несколько выше, чем клинических. Фильтрация выбросов (нетипичных объектов) из обучающей выборки позволяет построить алгоритм, точность которого на несколько процентов выше. Наилучшим качеством предсказаний обладают решающие деревья, однако они наименее устойчивы к изменению состава данных; для их адекватной настройки требуется выборка большей длины. Другие методы не позволяют увеличить точность прогнозов при совмещении клинических и иммунологических данных.

Работа выполнена при поддержке РФФИ, проекты № 06-07-08102-офи, № 07-01-00734, № 07-07-00380.

Литература

- [1] Кузнецов М. Р., Хаитов М. Р., Туркин П. Ю., Москаленко Е. П., Пинегин Б. В. Роль нарушений гуморального и клеточного иммунитета в генезе стеноза сосудистых анастомозов после реконструктивных вмешательств на артериях таза и нижних конечностей // Грудная и сердечно-сосудистая хирургия. — 2005. — № 2. — С. 29–33.
- [2] Кузнецов М. Р., Вирганский А. О., Капранов С. А., Москаленко Е. П., Туркин П. Ю. и др. Способ диагностики функциональной полноценности сосудистого анастомоза после реконструктивных хирургических вмешательств. Патент РФ № 2266711 от 27.12.2005.