Н.В. Корепанова

Машинное обучение для оптимизации лечения в подгруппах пациентов[1]

Аннотация. Клинические исследования показывают, что часто эффект от лечения оказывается зависимым от различных признаков пациента: клинических, антропологических, генетических, психологических, социальных и т.д. Выявление подобного рода зависимостей составляет задачу персонифицированной медицины и способствует созданию стратегий лечения, более адаптированных под конкретного пациента. В данной работе представлен обзор подходов к анализу данных клинических исследований для поиска признаков, влияющих на эффективность лечения, и выделения подгрупп пациентов, для которых есть существенные различия в эффективности экспериментального и контрольного лечения.

Ключевые слова: персонифицированная медицина, анализ подгрупп, клинические исследования, машинное обучение.

Введение

В последнее время персонифицированную медицину все чаще рассматривают как многообещающий инструмент повышения эффективности лечения ряда заболеваний. Первым шагом на пути к персонифицированному лечению можно считать выделение и анализ подгрупп пациентов на данных клинических исследований с целью оптимизации терапии в этих подгруппах. В результате длительных споров о полезности, опасности и корректности анализа подгрупп [1-9] большинство исследователей сошлись во мнении, что существует две задачи анализа подгрупп, отличающихся применимостью сформулированных принципов [10]:

1) Подтверждение заранее сформулированного небольшого числа гипотез о наличии эффекта в подгруппах, определенных заранее, в ходе клинического исследования (подтверждающий анализ).

2) Выявление и оценка подгрупп по данным проведенного клинического исследования (исследовательский анализ).

С точки зрения персонифицированной медицины обе задачи имеют право на существование, однако возможности первой сильно ограничены необходимостью заранее формулировать гипотезы. Вторая задача, фактически, направлена на исследование и поиск подгрупп, но, как правило, требует проведения новых клинических исследований для подтверждения наличия эффекта в выделенных подгруппах, поэтому крайне необходимо на имеющихся данных оценить устойчивость и воспроизводимость результатов.

Часто исследовательский анализ подгрупп рассматривается как особый случай выбора наилучшей в некотором смысле модели выделения подгрупп из пространства моделей, построенных, например, при помощи методов машинного обучения на имеющихся данных. При этом мета-параметры, позволяющие контролировать сложность пространства моделей и ориентироваться в нем, тоже оцениваются по данным, но критерии оптимальности значений мета-параметров оговариваются заранее. Как правило, моделирование выделения подгрупп, оценка и выбор наилучшей модели составляют

[1] Работа выполнена при финансовой поддержке РФФИ, грант №16-29-12982.

единую процедуру. Таким образом, заранее фиксируется пространство моделей и сама процедура моделирования и выбора, включая критерий оптимальности мета-параметров, а все остальное оценивается по данным.

В статье представлен обзор основных требований, предъявляемых к исследовательскому анализу подгрупп, а также обзор различных классов методов исследовательского анализа подгрупп для персонификации лечения. К сожалению, несмотря на актуальность задачи персонификации лечения в целом, в отечественной литературе не удалось обнаружить оригинальных методов анализа данных, предназначенных для решения этой задачи, чем объясняется отсутствие в Литературе на русскоязычные источники.

1. Исследовательский анализ подгрупп

В работе [10] по результатам анализа большого числа предшествующих работ были выделены необходимые составляющие исследовательского анализа подгрупп.

Ошибка первого рода или доля ложных отклонений нулевой гипотезы оценивается для всей процедуры выделения подгрупп в целом. Статистическая проверка различий в каждой отдельно взятой подгруппе без учета значений статистики в других подгруппах смысла не имеет. При этом наличие контроля ошибки первого рода или доли ложных отклонений нулевой гипотезы имеет большое значение при анализе клинических данных, так как цена ошибки в результате применения результатов анализа на практике очень высока. Это значит, что алгоритмы выявления подгрупп пациентов должны включать в себя статистическую проверку гипотез в подгруппах. Как правило, такое требование рассматривается как препятствие к использованию методов машинного обучения для персонификации лечения в подгруппах, так как тестируемые гипотезы формируются по тем же данным, на которых затем должны тестироваться. Однако для ряда методов машинного обучения были разработаны процедуры контроля ошибки при множественном тестировании. Например, в статье [11] представлен подход к проверке достоверности результатов регрессии при большом числе признаков, а авторы статьи [12] предложили аналитическую процедуру для проверки достоверности коэффициентов lasso регрессии. Более того, статистическая проверка достоверности и оценка воспроизводимости результатов может быть составной частью некоторых методов анализа данных [13-16]. Решением для аппроксимации ошибки могут служить подходы, основанные на ресемплировании [17] исходного набора данных. При этом при исследовательском анализе подгрупп рекомендуется не задаваться «магическим» значением уровня значимости в 5%, а оценивать «неслучайность» эффекта в подгруппах.

Контроль сложности модели для предотвращения переобучения. Как было отмечено выше, исследовательский анализ подгрупп можно рассматривать как задачу отбора моделей выделения подгрупп. Так как пространство моделей может быть очень большим, необходимо встраивать контроль сложности модели (регуляризация, прунинг и т.д.) в сам процесс его отбора, чтобы он не завершился выбором переобученной модели, которая даже при дополнительных модификациях не способна качественно описывать новые данные. Более того, встроенный контроль сложности также позволяет сократить суммарное число тестируемых гипотез, т.е. упрощает контроль ошибки первого рода или доли ложных отклонений нулевой гипотезы.

Снижение влияния смещения, вызванного различиями областей определения признаков. Часто для выделения подгрупп осуществляется перебор всевозможных разбиений множества пациентов, определяемых уникальными значениями признаков. Это приводит к тому, что вероятность выбора признака для определения подгруппы тем выше, чем больше у него уникальных значений. Поэтому процедура выбора подгрупп должна быть организована так, чтобы вероятность ошибочного выбора подгруппы с наличием различий не зависела от числа уникальных значений для каждого из признаков [10]. Эта проблема была исследована в работах [18, 19] в контексте алгоритмов рекурсивного разбиения пространства признаков.

Оценка воспроизводимости результатов. Чтобы выделенную подгруппу можно было считать проявлением скрытой закономерности в данных, а не случайности, необходимо, чтобы при аналогичном анализе новых данных эта подгруппа с наименьшими изменениями была в числе результатов выделения подгрупп. Чтобы полу-

чать оценку вероятности воспроизводимости чаще всего используют техники ресемплирования, вроде бутстрепа [20] и скользящего контроля [21]. Те подгруппы, которые с минимальными изменениями воспроизводятся на ресемплированных данных, называют стабильными.

Получение несмещенных оценок различия в эффективности экспериментального и контрольного лечения. Важно не только выделять подгруппы таким образом, чтобы они воспроизводились на новых данных, но и чтобы эффект в этих подгруппах на исходных данных и на новых данных не различался существенно. Для этого чаще всего применяют техники ресемплирования, например, как в методе Virtual Twins [22]. Также широкое распространение получили Байесовские методы оценки различий в эффективности в выбранных подгруппах [23-25] .

К сожалению, не все существующие на данный момент методы исследовательского анализа подгрупп содержат в себе все указанные выше составляющие, поэтому требуют доработки. Тем не менее, далее мы рассмотрим основные классы методов исследовательского анализа подгрупп пациентов.

2. Общие обозначения

Будем предполагать, что анализируются данные клинического исследования, которое проводилось с целью сравнить экспериментальное лечение с контрольным (или плацебо). Обозначим: n – число пациентов в исследовании, p – количество признаков пациента, используемых для определения подгрупп. Пусть $\boldsymbol{X}=\{X_1, \ldots, X_p\}$ – признаки, T – стратегия лечения, а Y - исход, причем большие значения Y соответствуют лучшему исходу. При этом Y может быть как бинарной или количественной переменной, так и оценкой времени до наступления некоторого события. Тогда для i-го пациента $\boldsymbol{x}_i=\{x_{i1},\ldots,x_{ip}\}$ - вектор значений признаков, измеренных или оцененных до начала терапии, t_i - реализация выбора лечения в исследовании, а y_i - значение исхода этого пациента при условии, что он получал лечение t_i $(i = 1,\ldots,n)$. При этом $t_i = 0$, если i-й пациент получал контрольное лечение, и $t_i = 1$, если экспериментальное. Подгруппа $S(\boldsymbol{X})$ - подмножество пациентов, определяемое правилом, накладывающим ограничения на значения вектора $\boldsymbol{X}$. Например, $S(\boldsymbol{X}) = I\{X_1 > c\}$. Здесь и далее $I\{\cdot\}$ – индикаторная функция.

Пусть $f(\boldsymbol{x},t)=E(Y|\boldsymbol{X}=\boldsymbol{x}, T=t)$ – функция ответа пациента с признаками $\boldsymbol{x}$ на терапию t. Здесь и далее $E(\cdot)$ - математическое ожидание. Также обозначим за $z(\boldsymbol{X})$ функцию различия в ответе пациента на экспериментальное и контрольное лечение. Будем полагать, что $z(\boldsymbol{X})$ можно представить в виде $z(\boldsymbol{X}) = g(f(\boldsymbol{X},1), f(\boldsymbol{X},0))$, где $g(\cdot)$ - монотонная по всем аргументам функция. Тогда функция ожидаемого исхода принимает вид: $f(\boldsymbol{X},T) = g(h(\boldsymbol{X}) + l(z(\boldsymbol{X})T))$ [26], где $h(\cdot)$ - произвольная функция от вектора признаков, не зависящая от лечения, $l(\cdot)$ - монотонная функция. Функцию $h(\cdot)$ интерпретируют как базовый (не зависящий от лечения) исход пациента, а те признаки, от которых зависит значение этой функции, называют *прогностическими*. Для персонифицированной медицины больший интерес представляет функция $z(\cdot)$, так как она, по сути, представляет собой зависимость различий в ответе на экспериментальное и контрольное лечение. Признаки, которые имеют вклад в значение функции $z(\cdot)$, называют *предиктивными*.

Для ряда методов ключевым является понятие *потенциального исхода*. Так для i-го пациента есть два потенциальных исхода $\tilde{Y}_i(0)$ и $\tilde{Y}_i(1)$ - случайные величины, представляющие оценки значения исхода случайного пациента при получении лечения $T = 0$ и $T = 1$, соответственно. При этом предполагается что наблюдаемый исход совпадает с потенциальным исходом для того лечения, которое в действительности получал пациент (предположение состоятельности), т.е. $Y_i = \tilde{Y}_i(0)(1 - T_i) + \tilde{Y}_i(1)T_i$, $i=1,\ldots,n$. При этом функцию исхода $f(\boldsymbol{x},t)$ можно оценить $E(\tilde{Y}_i(t)|\boldsymbol{X}=\boldsymbol{x})$, если выбор лечения для каждого пациента независим от выбора лечения для других пациентов, и пациенты не подвергались никаким дополнительным вмешательствам, которые могли бы повлиять на исход лечения [27].

Далее речь пойдет о различных классах методов выделения подгрупп:

- Базовые методы на основе одномерной регрессии и регрессионных деревьев.
- Методы моделирования исхода, как функции от признаков пациента и реализованной стратегии лечения $f(\boldsymbol{X},T)$.
- Методы моделирования различия эффективности между экспериментальным и кон-

трольным лечением $z(\boldsymbol{X})$. К ним, в том числе, относятся методы определения оптимального индивидуального правила лечения и методы моделирования разницы между ожидаемыми ответами на экспериментальное и контрольное лечение (моделирование аплифта).

- Методы локального моделирования, направленные на выделение подгрупп с высоким значением функции $z(\boldsymbol{X})$ без моделирования функции $z(\boldsymbol{X})$ на всем признаковом пространстве.

Данная классификация с небольшими модификациями заимствована из [10].

3. Базовые методы

Базовые методы появились раньше остальных и, как правило, любой метод анализа и выделения подгрупп в первую очередь сравнивается с ними. В первую очередь, к базовым методам относятся одномерные регрессионные модели (чаще всего, регрессионные модели Кокса [28]). Для каждого признака (X_i, $i=1,...,n$) строится регрессионная модель вида $Y_i = \alpha X_i + \beta T_i + \gamma w(X_i, T)$, где α, β, и γ – параметры модели, а $w(X_i, T)$ – функция взаимодействия признака с лечением (например, X_iT или $I\{X_i<c\}T$). Далее, оставляют только те признаки, для которых значимость компоненты взаимодействия признака с лечением оказывается выше некоторого наперед заданного порогового значения. Каждый из отобранных признаков может в отдельности от остальных быть использован для определения подгруппы. Главный недостаток – отсутствие возможности учета одновременного влияния нескольких признаков на исход.

Также к базовым методам относят регрессионные деревья (например, на основе CART [29]), где лечение включается во множество признаков модели наряду с признаками пациента, а в качестве целевой переменной выступает исход. В отличие от одномерных регрессионных моделей регрессионные деревья позволяют улавливать более сложные зависимости, в том числе и между признаками, и дают возможность включать в определение подгруппы несколько признаков пациента одновременно. Также в случае количественных или порядковых признаков деревья регрессии позволяют автоматически выбирать пороги разбиения, в то время как при использовании одномерной регрессии эти пороги необходимо задавать заранее (например, при введении в регрессионную модель индикаторов вида $I\{X_j<c\}$).

4. Моделирование функции исхода

Примером параметрического метода служит применение регрессии на основе частичных полиномов [30] к моделированию зависимости исхода от значений количественных признаков при разных стратегиях лечения [31]. Применение этого метода оправдано лишь при небольшом количестве признаков, так как с ростом их числа резко растет число параметров модели.

Также к параметрическим методам относятся методы регрессионного моделирования функции исхода с регуляризацией, включающие как прогностические, так и предиктивные компоненты. Например, lasso-регуляризация [32] позволяет, фактически, осуществлять отбор признаков для формирования подгрупп пациентов. Примером ее применения при решении задачи выделения подгрупп служит метод FindIt [33], представляющий собой классификатор на основе опорных векторов с lasso-регуляризацией отдельно для прогностических и предиктивных компонентов модели. Необходимость отдельной регуляризации для прогностической и предиктивной составляющих модели авторы объясняют тем, что влияние вторых на исход экспериментального лечения, обычно, слабее, но в концепции персонифицированной медицины именно выявление предиктивных признаков играет ключевую роль.

В рассматриваемом классе можно также отдельно выделить методы, использующие понятие потенциального исхода[22, 34-37]. На первом шаге для каждого пациента оценивается потенциальный исход при обеих стратегиях лечения, например при помощи регрессионной модели Кокса, и вычисляется их разность. Далее для разности тем или иным способом осуществляется выбор порогового значения вхождения в подгруппу с превосходством экспериментального лечения. Так, в непараметрическом методе Virtual Twins [22] при помощи случайного леса [38] получают оценки потенциального исхода. Далее разницу между потенциальными исходами используют в качестве целевой переменной регрессионного дерева, позволяющего выделить и описать подгруппы.

Метод STIMA [39] – пример комбинации параметрического и непараметрического подходов

к моделированию. На первом шаге строится линейная регрессионная модель для выявления прогностических эффектов. Далее отдельно для пациентов, получавших экспериментальное и контрольное лечение, строится дерево решения, где на каждом шаге максимизируется рост доли объясненной дисперсии после разбиения. После добавления разбиения в дерево к регрессионной модели добавляется соответствующий признак взаимодействия лечения и разбиения, выбранного на текущем шаге, и пересчитываются коэффициенты регрессии. Таким образом, получается последовательность регрессионных моделей, упорядоченных по возрастанию сложности модели, среди которых посредством прунинга на основе скользящего контроля выбирается оптимальная модель.

Байесовские методы выделения подгрупп посредством моделирования функции исхода также имеют место [23, 40]. Кроме того, есть примеры применения байесовских регрессионных моделей с регуляризацией [41] и непараметрических байесовских методов [25].

5. Моделирование функции различий

Отличительной особенностью этого класса методов является моделирование предиктивного эффекта без необходимости моделирования прогностического. В статьях [42-44] предлагается серия модификаций метода под названием Interaction Trees (IT), в основе которого– регрессионные деревья CART. В узлах IT обучаются модели вида $y_i = a_0 + a_1 s_i + a_2 t_i + a_3 t_i s_i + \varepsilon_i$, где s_i – индикаторная функция разбиения по одному из количественных признаков, не зависящая от лечения (например, $s_i = I\{X_1 < c\}$). Выбор разбиения происходит путем минимизации среднеквадратичной ошибки или p-значения теста на отличие от нуля коэффициента a_3. Поэтому полагается, что метод IT строит кусочно-постоянную оценку функции $z(\boldsymbol{X})$. В статье [43] также представлена схема прунинга для IT.

Другой метод выделения подгрупп на основе деревьев предлагается в статьях [45, 46], в основе которого рекурсивные деревья GUIDE [47]. Выбор деления в каждом узле происходит в два этапа: выбирается лучший признак для деления независимо от числа уникальных значений этого признака, а затем для выбранного признака выбирается лучшее значения для деления. Это позволяет снизить влияние вариативности признака на его важность для модели. Похожие идеи применялись в ряде других методов, в их числе – деревья условного вывода [19] и рекурсивное деление на основе моделей [48].

Метод QUINT [49] - еще один вариант на основе деревьев, с предусмотренной процедурой прунинга на основе ресемплирования. В его простейшей модификации при выборе разбиения в каждом узле дерева максимизируется

$$\sum_{l\in G_1} N_l(\bar{Y}_l(1) - \bar{Y}_l(0))/\sum_{l\in G_1} N_l - \sum_{l\in G_0} N_l(\bar{Y}_l(0) - \bar{Y}_l(1))/\sum_{l\in G_0} N_l,$$

где G_1 и G_0 – множества листов дерева с превосходством экспериментального и контрольного лечения соответственно; N_l – число пациентов в листе l; $\bar{Y}_l(t)$ – оценка среднего исхода в листе при лечении t.

Для этого метода также предусмотрена процедура прунинга на основе ресемплирования. Пример параметрического метода, моделирующего функцию различий, представлен в статье [50]. Утверждается, что в случае количественного исхода и равновероятной рандомизации на стратегии лечения $z(\boldsymbol{x}) = E(2Y(2T-1)|\boldsymbol{X}=\boldsymbol{x})$. Тогда достаточно сформировать целевой признак $Z = 2Y(2T-1)$ и построить регрессионную модель $z(X)$ без необходимости моделирования прогностического эффекта.

Байесовские методы моделирования функции различий также имеют место. Так в статье [51] представлен общий взгляд на байесовский подход к анализу в подгруппах.

6. Методы выбора индивидуального оптимального правила лечения

Методы, нацеленные на поиск «оптимального» лечения для каждого пациента, также можно отнести к классу методов, моделирующих $z(\boldsymbol{X})$. *Индивидуальное правило лечения* $d(\boldsymbol{X})$ - это функция, которая вектору признаков $\mathbf{X}$ ставит в соответствие одну из стратегий лечения. Потенциальный исход при лечении по правилу $d(\boldsymbol{X})$ определяется как $\tilde{Y}(d(\boldsymbol{X})) = \tilde{Y}(1)d(\boldsymbol{X}) + \tilde{Y}(0)(1 - d(\boldsymbol{X}))$. В статье [52] впервые была представлена *функция ценности* индивидуального правила: $V(d(\boldsymbol{X})) = E(Y(d(\boldsymbol{X})))$. *Оптимальное индивидуальное правило лечения* определяют как $d_{opt}(\boldsymbol{X}) = argmax_d V(d(\boldsymbol{X}))$. Самый простой способ оценки: $d_{opt}(\boldsymbol{X}) = I\{\hat{f}(\boldsymbol{X},1) > \hat{f}(\boldsymbol{X},0)\}$, где $\hat{f}(\boldsymbol{X},\cdot)$ – оценка функции исхода, которую, например, по-

лучают при помощи lasso-регрессии [52]. Также в [53] было показано, что если функция исхода имеет вид $f(\boldsymbol{X},T) = h(\boldsymbol{X}) + z(\boldsymbol{X})T$, то $d_{opt}(\boldsymbol{X}) = I\{z(\boldsymbol{X}) > 0\}$, поэтому оценка оптимального индивидуального правила равноценна моделированию функции различий [54, 55]. В работе [52] неявно было отмечено, что задача поиска оптимального правила эквивалентна задаче взвешенной бинарной классификации, т.е. можно достаточно просто адаптировать для решения этой задачи любой бинарный классификатор, что было подтверждено и использовано для разработки ряда методов [53, 56-60].

7. Моделирование аплифта

Методы моделирования *аплифта* под разными названиями были предложены в ряде работ, касающихся, прежде всего, решения маркетинговой задачи A/B-тестирования. По сути, моделирование аплифта представляет собой частный случай моделирования функции различий. Сам аплифт определяется как $m(\boldsymbol{x}) = E(Y|\ \boldsymbol{X}=\boldsymbol{x},\ T = 1) - E(Y|\ \boldsymbol{X}=\boldsymbol{x},\ T = 0)$.

Большинство регрессионных методов основываются на отдельном моделировании исхода при контрольном и экспериментальном лечении [61-64]. В работе [65] на примере логистической регрессии была продемонстрирована возможность параметрического моделирования аплифта при переходе к моделированию переменной $Z = I\{T = Y\}$.

В первом методе моделирования аплифта при помощи деревьев решений критерий ветвления основан на оценке статистической значимости разницы между вероятностями успешного исхода при двух стратегиях лечения [66, 67]. Другой пример адаптации деревьев решений был представлен в статье [68], в котором при разбиении максимизируется абсолютная разница значений аплифта в левой и правой ветвях разбиения. Другие методы моделирования аплифта на основе деревьев представлены в работах [69-71].

Также для моделирования аплифта разработано несколько методов на основе опорных векторов [72-74] и известны примеры применения ансамблевых подходов к аплифту моделирования: в работе [67] авторы упоминают об успешном применении бэггинга, а в [75] представлена адаптация случайного леса. Экспериментальное сравнение методов моделирования аплифта представлено в работе [76].

8. Локальное моделирование

Примерами методов локального моделирования являются PRIM [77] и SIDES [78] на основе рекурсивной процедуры bump hunting [79], которая позволяет выделять прямоугольные области в признаковом пространстве, в которых сосредоточено большое число пациентов с хорошим исходом при экспериментальном лечении или с плохим исходом при контрольном лечении.

Байесовские методы локального моделирования представлены в статьях [24, 80], рассматривающих эффект от экспериментального лечения в подгруппах пациентов как наличие подмоделей. Так, авторы [24] предлагают определять подмодель, используя 10 различных базовых структур. В рамках каждой из них эффект от экспериментального лечения моделируется локально путем выбора одного прогностического и одного предиктивного бинарного признака. После оценки априорной вероятности каждой подмодели из пространства моделей, образованного всевозможными комбинациями прогностических и предиктивных признаков, для каждого пациента на основе оценок апостериорных вероятностей рассчитывается усреднение по подмоделям, которое и является оценкой эффекта для этого пациента.

Также к классу локальных методов можно отнести метод на основе теории паросочетаний и деревьев решений [81]. Сначала на основании значений признаков вводятся понятия сходства и сравнимости пациентов. Например, пациенты считаются сравнимыми, если значения категориальных признаков совпадают, а степень сходства рассчитывается по значениям количественных признаков при помощи какой-либо выбранной заранее меры сходства. Таким образом, сравнимость и сходство позволяют задать частичный порядок на множестве пациентов. Далее для поиска пар пациентов со схожим набором признаков, получавших разное лечение, предлагается использовать алгоритм решения задачи об оптимальном паросочетании [82-84]. В качестве сочетаемых множеств берутся множества пациентов, соответствующие стратегиям лечения, а в качестве предпочтений – заданный на предыдущем шаге частичный порядок. Поиск оптимального, а не максимального, паросочетания позволяет сохранить наибольшее число пациентов для дальнейшего

анализа, что важно в случае небольшого объема исходных данных, а также избежать переобучения. Далее, при наличии критерия превосходства эффективности одного лечения над другим в зависимости от исхода лечения, пары пациентов разбиваются на несколько классов, например на 4, (лучше экспериментальное, лучше контрольное, оба неэффективны, оба эффективны), и оба пациента пары получают метку соответствующей пары. После этого для формирования подгрупп на множестве пациентов с метками классов применяются деревья классификации.

Еще одной попыткой локального моделирования эффекта можно отнести метод на основе узорных структур [85, 86], представленный в [87, 88]. Здесь подгруппа описывается интервальными ограничениями на количественные признаки и подмножествами значений категориальных признаков. Главная идея состоит в сокращении полного перебора всех возможных комбинаций значений признаков до перебора только максимальных комбинаций признаков. Иными словами, разные комбинации признаков могут описывать одну и ту же подгруппу пациентов, поэтому достаточно рассмотреть только ту комбинацию признаков, которая включает в себя все признаки, свойственные данной подгруппе. Это, с одной стороны, позволяет существенно сократить перебор, а с другой, в отличие от методов на основе деревьев решений, избавляет от необходимости жадного перебора, что, однако, негативно сказывается на времени работы такого подхода. Также в отличие от подгрупп, получаемых при использовании деревьев решений, получаемые подгруппы могут пересекаться, что несколько снижает интерпретируемость результатов. Решение этой проблемы в рамках данного метода не предложено.

Заключение

Развитие персонифицированной медицины спровоцировало бурное развитие методов выделения и анализа подгрупп пациентов для сравнения исхода при разных стратегиях лечения. Причем использование чисто статистических методов ограничивает пространство поиска, а применение методов машинного обучения и анализа данных без статистической поддержки оказывается недостаточным для того, чтобы обосновать надежность полученных результатов для применения в клинической практике. Так как часто к подгруппам предъявляется требование интерпретируемости, наибольшее распространение получили линейные регрессионные методы и древовидные методы регрессии и классификации. Дополнительную сложность представляют данные с цензурированным исходом, так как такой исход не вписывается ни в концепцию регрессии, ни в концепцию классификации. В данной работе представлены современный взгляд на анализ эффективности лечения в подгруппах и вариант классификации существующих на данный момент методов исследовательского анализа подгрупп.

Литература

1. Brookes S.T., Whitley E., Peters T.J., Mulheran P.A., Egger M., Smith G.D. Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives // Health Technology Assessment, Vol. 5, No. 33, 2001. pp. 1-56.
2. Cook D.I., Gebski V.J., Keech A.C. Subgroup analysis in clinical trials // Medical Journal of Australia, Vol. 180, No. 6, 2004. pp. 289-291.
3. Grouin J.M., Coste M., Lewis J. Subgroup Analyses in Randomized Clinical Trials: Statistical and Regulatory Issues // Journal of Biopharmaceutical Statistics, Vol. 15, No. 5, 2005. pp. 869-882.
4. Kent D.M., Rothwell P.M., Ioannidis J.P.A., Altman D.G., Hayward R.A. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal // Trials, Vol. 11, No. 1, 2010. P. 85.
5. Pocock S.J., Assmann S.E., Enos L.E., Kasten L.E. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems // Statistics in Medicine, Vol. 21, No. 19, 2002. pp. 2917-2930.
6. Rothwell P.M. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation // Lancet, Vol. 365, No. 9454, 2005. pp. 176-186.
7. Sleight P. Debate: Subgroup analyses in clinical trials — fun to look at, but don't believe them! // Current Controlled Trials in Cardiovascular Medicine, Vol. 1, No. 1, 2000. pp. 25-27.
8. Sun X., Briel M., Walter S.D., Guyatt G.H. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses // British Medical Journal, Vol. 340, 2010. pp. 850-854.
9. Wang R., Lagakos S.W., Ware J.H., Hunter D.J., Drazen J.M. Statistics in Medicine - Reporting of Subgroup Analyses in Clinical Trials // The New England Journal of Medicine, Vol. 357, 2007. pp. 2189-2194.
10. Lipkovich I., Dmitrienko A., DAgostino R.B. Tutorial in Biostatistics: Data-Driven Subgroup Identification and Analysis in Clinical Trials // Statistics in Medicine, Vol. 36, No. 1, 2016. pp. 136-196.
11. Meinshausen N., Meier L., Bühlmann P. p-Values for High-Dimensional Regression // Journal of the American Statistical Association, Vol. 104, No. 488, 2009. pp. 1671-1681.

12. Lockhart R., Taylor J., Tibshirani R.J., Tibshirani R. A Significance Test for the Lasso // Annals of Statistics, Vol. 42, No. 2, 2014. pp. 413-468.
13. Freidlin B., Simon R. Adaptive Signature Design: An Adaptive Clinical Trial Design for Generating and Prospectively Testing A Gene Expression Signature for Sensitive Patients // Clinical Cancer Research, Vol. 11, No. 21, 2005. pp. 7872-7878.
14. Meinshausen N., Bühlmann P. Stability selection // Journal of the Royal Statistical Society, Series B, Vol. 72, No. 4, 2010. pp. 417-423.
15. Gunter L., Zhu J., Murphy S. Variable Selection for Qualitative Interactions in Personalized Medicine while Controlling The Familywise Error Rate // Journal of Biopharmaceutical Statistics, Vol. 21, No. 6, 2011. pp. 1063-1078.
16. Simon R.M., Subramanian J., Li M.C., Menezes S. Using cross-validation to evaluate predictive accuracy of survival risk classifiers based on high-dimensional data // Briefings in Bioinformatics, Vol. 12, No. 3, 2011. pp. 203-214.
17. Good P. Resampling Methods: a Practical Guide to Data Analysis. 3rd ed. Boston: Birkhauser, 2005.
18. Loh W.Y., Shih Y.S. Split Selection Methods for Classification Trees // Statistica Sinica, Vol. 7, 1997. pp. 815-840.
19. Hothorn T., Hornik K., Zeileis A. Unbiased Recursive Partitioning: A Conditional Inference Framework // Journal of Computational and Graphical Statistics, Vol. 15, No. 3, 2006. pp. 651-674.
20. Hesterberg T., Moore D.S., Monaghan S., Clipson A., R. E. Bootstrap methods and permutation tests. Vol 5. // In: The Practice of Business Statistics / Ed. by Moore D.S. W. H. Freeman, 2005. pp. 1-70.
21. Varma S., Simon R. Bias in error estimation when using cross-validation for model selection // BMC Bioinformatics, Vol. 7, 2006. P. 91.
22. Foster J.C., Taylor J., Ruberg S.J. Subgroup identification from randomized clinical trial data. // Statistics in Medicine, Vol. 30, No. 24, 2011. pp. 2867-2880.
23. Dixon D.O., Simon R. Bayesian Subset Analysis // Biometrics, Vol. 47, No. 3, 1991. pp. 871-881.
24. Berger J.O., Wang X., Shen L. A Bayesian Approach to Subgroup Identification // Journal of Biopharmaceutical Statistics, Vol. 24, No. 1, 2014. pp. 110-129.
25. Xu Y., Trippa L., Müller P., Ji Y. Subgroup-Based Adaptive (SUBA) Designs for Multi-Arm Biomarker Trials // Statistics in Biosciences, Vol. 8, No. 1, 2016. pp. 159-180.
26. Xu Y., Yu M., Zhao Y.Q., Li Q., Wang S., Shao J. Regularized Outcome Weighted Subgroup Identification for Differential Treatment Effects // Biometrics, Vol. 71, No. 3, Sep 2015. pp. 645-653.
27. Little R.J., Rubin D.R. Causal effects in clinical and epidemiological studies via potential outcomes. // Annual Review of Public Health, Vol. 21, 2000. pp. 121-145.
28. Cox D.R. Regression Models and Life-Tables // Journal of the Royal Statistical Society. Series B, Vol. 34, No. 2, 1972. pp. 187-220.
29. Breiman L., Friedman J.H., Olshen R.A., Stone C.J. Classification and Regression Trees. Wadsworth: Belmont, CA, 1984.
30. Royston P., Altman D.G. Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling // Journal of the Royal Statistical Society. Series C, Vol. 43, No. 3, 1994. pp. 429-467.
31. Royston P., Sauerbrei W. A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials // Statistics in Medicine, Vol. 23, No. 16, 2004. pp. 2509-2525.
32. Tibshirani R. Regression shrinkage and selection via the lasso. // Journal of the Royal Statistical Society. Series B, Vol. 58, No. 1, 1996. pp. 267-288.
33. Imai K., Ratkovic M. Estimating Treatment Effect Heterogeneity in Randomized Program Evaluation // Tha Annals of Applied Statistics, Vol. 7, No. 1, 2013. pp. 443-470.
34. Cai T., Tian L., Peggy H W., Wei L.J. Analysis of randomized comparative clinical trial data for personalized treatment selections // Biometrics, Vol. 12, No. 2, 2011. pp. 270-282.
35. Song X., Pepe M.S. Evauating Markers for Selecting a Patient's Treatment // Biometrics, Vol. 60, No. 4, 2004. pp. 874-883.
36. Huang Y., Gilbert P.B., Janes H. Assessing Treatment-Selection Markers using a Potential Outcomes Framework // Biometrics, Vol. 68, No. 3, 2012. pp. 687-696.
37. Zhao L., Tian L., Cai T., Claggett B., Wei L.J. Effectively Selecting a Target Population for a Future Comparative Study // Journal of the American Statistical Association, Vol. 108, No. 502, 2013. pp. 527-539.
38. Breiman L. Random forests // Machine Learning, Vol. 45, No. 1, 2001. pp. 5-32.
39. Dusseldorp E., Conversano C., Van Os B.J. Combining an Additive and Tree-Based Regression Model Simultaneously: STIMA // Journal of Computational and Graphical Statistics, Vol. 19, No. 3, 2010. pp. 514-530.
40. Hodges J.S., Cui Y., Sargent D.J., Carlin B.P. Smoothing Balanced Single-Error-Term Analysis of Variance // Technometrics, Vol. 49, No. 1, 2007. pp. 12-25.
41. Gu X., Yin C., Lee J.J. Bayesian Two-step Lasso Strategy for Biomarker Selection in Personalized Medicine Development for Time-to-Event Endpoints // Contemporary Clinical Trials, Vol. 36, No. 2, 2013. pp. 642-650.
42. Negassa A., Ciampi A., Abrahamowicz M., Shapiro S., Boivin J.F. Tree-structured subgroup analysis for censored survival data: Validation of computationally inexpensive model selection criteria // Statistics and Computing, Vol. 15, No. 3, 2005. pp. 231-239.
43. Su X., Tsai C.L., Wang H., Nickerson D.M., Li B. Subgroup Analysis via Recursive Partitioning // Journal of Machine Learning Research, Vol. 10, 2009. pp. 141-158.
44. Su X., Zhou T., Yan X. Interaction Trees with Censored Survival Data // The International Journal of Biostatistics, Vol. 4, No. 1, 2008. P. 2.
45. Loh W.W., He X., Man M. A regression tree approach to identifying subgroups with differential treatment effects // Statistics in Medicine, Vol. 34, No. 11, 2015. pp. 1818-1833.
46. Loh W.Y., Fu H., Man M., Champion V., Yu M. Identification of subgroups with differential treatment effects for longitudinal and multiresponse variables // Statistics in Medicine, Vol. 35, No. 26, 2016. pp. 4837-4855.
47. Loh W.Y. Regression Trees with Unbiased Variable Selection and Interaction Detection // Statistica Sinica, Vol. 12, 2002. pp. 361-386.
48. Zeileis A., Hothorn T., Hornik K. Model-based Recurcive Partitioning // Journal of Computational and Graphical Statistics, Vol. 17, No. 2, 2008. pp. 492-514.
49. Dusseldorp E., Mechelen I.V. Qualitative Interaction Trees: a tool to identify qualitative treatment-subgroup interactions // Statistics in Medicine, Vol. 33, No. 2, 2014. pp. 219-237.
50. Tian L., Alizadeh A.A., Gentles A.J., Tibshirani R. A Simple Method for Estimating Interactions between a

Treatment and a Large Number of Covariates // Journal of the Americal Statistical Association, Vol. 109, No. 508, 2014. pp. 1517-1532.
51. Jones H.E., Ohlssen D.I., Neuenschwander B., Racine A., Branson M. Bayesian Models for Subgorup Analysis in Clinical Trials // Clinical Trials, Vol. 8, No. 2, 2011. pp. 129-143.
52. Qian M., Murphy S.A. Performance guarantees for individualized treatment rules // The Annals of Statistics, Vol. 39, No. 2, 2011. pp. 1180-1210.
53. Zhao Y., Zheng D., Rush A.J., Kosorok M.R. Estimating individualized treatment rules using outcome weighted learning. // Journal of the American Statistical Association, Vol. 107, No. 449, 2012. pp. 1106-1118.
54. Lu W., Zhang H.H., Zeng D. Variable Selection for Optimal Treatment Decision // Statistical Methods in Medical Research, Vol. 22, No. 5, 2013. pp. 493-504.
55. Foster J., Taylor J.M.G., Kaciroti N., Nan B. Simple subgroup approximations to optimal treatment regimes from randomized clinical trial data // Biostatistics, Vol. 16, No. 2, 2015. pp. 368-382.
56. Zhang B., Tsiatis A.A., Davidian M., Zhang M., Laber E. Estimating Optimal Treatment Regimes from a Classification Perspective // Statistics, Vol. 1, No. 1, 2012. pp. 103-114.
57. Zhang B., Tsiatis A.A., Laber E.B., Davidian M. Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions // Biometrika, Vol. 100, No. 3, 2013. pp. 681-694.
58. Laber E.B., Zhao Y.Q. Tree-base methods for individualized treatment regimes // Biometrika, Vol. 102, No. 3, 2015. pp. 501-514.
59. Zhang Y., Laber E.B., Tsiatis A., Davidian M. Using decision lists to construct interpretable and parsimonious treatment regimes // Biometrics, Vol. 71, No. 4, 2015. pp. 895-904.
60. Fu H., Zhou J., Faries D.E. Estimating optimal treatment regimes via subgroup identification in randomized control trials and observational studies // Statistics in Mdeicine, Vol. 35, No. 19, 2016. pp. 3285-3302.
61. Lo V.S.Y. The true lift model: a novel data mining approach to response modeling in database marketing. // SIGKDD Explorations, Vol. 4, No. 2, 2002. pp. 78-86.
62. Larsen K. Net lift models: optimizing the impact of your marketing. // Predictive analytics world. 2011. Vol. Workshop presentation.
63. Robins J.M. Correcting for non-compliance in randomized trials using structural nested mean models // Communications in Statistics - Theory and Methods, Vol. 23, No. 8, 1994. pp. 2379-2412.
64. Robins J., Rotnitzky N. Estimation of treatment effects in randomised trials with non-comliance and a dichotomous outcome using structural mean models // Biometrika, Vol. 91, No. 4, 2004. pp. 763-783.
65. Jaskowski M., Jaroszewicz S. Uplift modeling for clinical trial data // ICML, 2012 workshop on machine learning for clinical data analysis. Edinburgh. Scotland. 2012.
66. Radcliffe N.J., Surry P.D. Differential analysis: modeling true response by isolating the effect of a single action. // Proceedings of credit scoring and credit control VI. 1999.
67. Radcliffe N.J., Surry P.D. Real-world uplift modeling with significance-based uplift trees., Portrait Technical Report TR-2011-1, stochastic solutions, Tech. rep. 2011.
68. Hansotia B., Rukstales B. Incremental Value Modeling // Journal of Interactive Marketing, Vol. 16, No. 3, 2002. pp. 35-46.
69. Chickering D.M., Heckerman D. A decision theoretic approach to targeted advertising. // Proceedings of the 16th conference in uncertainty in artificail intelligence (UAI'00). 2000. pp. 82-88.
70. Rzepakowski P., Jaroszewicz S. Decision trees for uplift modeling // Proceedings of the 10th IEEE International conference on data mining (ICDM). Sydney. Australia. 2010. pp. 441-450.
71. Rzepakowski P., Jaroszewicz S. Decision trees for uplift modeling wth single and multiple treatments // Knowledge and Information Systems, Vol. 32, No. 2, 2012. pp. 303-327.
72. Kuusisto F., Costa V.S., Nassif H., Burnside E., Page D., Shavlik J. Support vector machines for differenctial prediction // Proceedings of the ECML-PKDD. 2014.
73. Jaroszewicz S., L. Zaniewicz Ł. Székely regularization for uplift modeling. // In: Challenges in computational statistics and data mining. Springer International Publishing, 2016. pp. 135-154.
74. Zaniewicz L., Jaroszewicz S. Lp - Support vector machines for uplift modeling // Knowledge and Information Systems, Vol. 53, No. 1, 2017. pp. 269-296.
75. Guelman L., Guillen M., Perez-Marin A.M. Random forests for uplift modeling: an insurance customer retention case. Vol 115. // In: Modeling and simulation in engingeering, economics and management. Springer, Berlin, 2012. pp. 123-133.
76. Soltys M., Jaroszewicz S., Rzepakowski P. Ensemble methods for uplift modeling // Data Mining and Knowledge Discovery, Vol. 29, No. 6, 2015. pp. 1531-1559.
77. Chen G., Zhong H., Belousov A., Devanarayan V. A PRIM approach to predictive-signature development for patient stratification // Statistics in Medicine, Vol. 34, No. 2, 2015. pp. 317-342.
78. Lipkovich I., Dmitrienko A., Denne J., Enas G. Subgroup identification based on differential effect search—A recursive partitioning method for establishing response to treatment in patient subpopulations // Statistics in Medicine, Vol. 30, No. 21, 2011. pp. 2601-2621.
79. Friedman J.H., Fisher N.I. Bump Hunting in High-Dimensional Data // Statistics and Computing, Vol. 9, No. 2, 1999. pp. 123-243.
80. Sivaganesan S., Laudb P.W., Müller P. A Bayesian subgroup analysis with a zero-enriched Polya Urn scheme // Statistics in Medicine, Vol. 30, No. 4, 2010. pp. 312-323.
81. Korepanova N., Kuznetsov S.O., Karachunskiy A.I. Matchings and Decision Trees for Determining Optimal Therapy // In: Analysis of Images, Social Networks and Texts Third International Conference, AIST 2014, Yekaterinburg, Russia, April 10-12, 2014, Revised Selected Papers. Springer International Publishing, 2014. pp. 101-110.
82. Gale D., Shapley L.S. College Admissions and the Stability of Marriage // The American Mathematical Monthly, Vol. 69, No. 1, 1962. pp. 9-15.
83. Roth A.E. Differed acceptance algorithm: history, theory, practice, and open questions, Harvard University, 2007.
84. Alkan A., Gale D. Stable schedule matching under revealed preference // Journal of Economic Theory, Vol. 112, 2003. pp. 289-306.
85. Ganter B., Kuznetsov S.O. Pattern Structures and Their Projections // 9th International Conference on Conceptual Structures (ICCS 2001). 2001. Vol. 2120. pp. 129-142.
86. Kuznetsov S.O. Pattern Structures for Analyzing Complex Data // 12th International Conference on Rough Sets,

Fuzzy Sets, Data Mining and Granular Computing (RSFDGrC 2009). 2009. Vol. 5908. pp. 33-44.
87. Korepanova N., Kuznetsov S.O. Pattern Structures for Treatment Optimization // In: CLA 2016: Proceedings of the Thirteenth International Conference on Concept Lattices and Their Applications. CEUR Workshop Proceedings. Moscow: Higher School of Economics, National Research University, 2016. pp. 217-228.
88. Корепанова Н.В., Кузнецов С.О. Выбор терапии онкологического заболевания в подгруппах пациентов на основе анализа замкнутых описаний // Пятнадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2016 (3-7 октября 2016 г., г. Смоленск, Россия). Труды конференции. В 3-х томах. Смоленск. Россия. 2016. Т. 1. С. 352-359.

Корепанова Наталья Владимировна. Стажер-исследователь Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ). Окончила НИУ ВШЭ в 2016 году. Количество печатных работ: 7. Область научных интересов: анализ данных, машинное обучение, медицинская информатика. E-mail: nkorepanova@hse.ru

Machine learning for treatment optimization in subgroups of patients

N.V. Korepanova

In clinical trials comparing experimental and control treatment the effect of treatment often depends on the range of patient's characteristics (biomarkers) such as clinical, anthropological, genetic, psychological, social characteristics and others. Personalized medicine aims at finding such dependencies to tailor treatment strategies to a patient. This paper presents an overview of the approaches to data analysis of clinical trials intended for identification of influential biomarkers and subgroups of patients, where experimental and control treatment differ significantly in efficiency.

Keywords: personalized medicine, subgroup analysis, clinical trials, machine learning.

References

1. Brookes S.T., Whitley E., Peters T.J., Mulheran P.A., Egger M., Smith G.D. Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives // Health Technology Assessment, Vol. 5, No. 33, 2001. pp. 1-56.
2. Cook D.I., Gebski V.J., Keech A.C. Subgroup analysis in clinical trials // Medical Journal of Australia, Vol. 180, No. 6, 2004. pp. 289-291.
3. Grouin J.M., Coste M., Lewis J. Subgroup Analyses in Randomized Clinical Trials: Statistical and Regulatory Issues // Journal of Biopharmaceutical Statistics, Vol. 15, No. 5, 2005. pp. 869-882.
4. Kent D.M., Rothwell P.M., Ioannidis J.P.A., Altman D.G., Hayward R.A. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal // Trials, Vol. 11, No. 1, 2010. P. 85.
5. Pocock S.J., Assmann S.E., Enos L.E., Kasten L.E. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems // Statistics in Medicine, Vol. 21, No. 19, 2002. pp. 2917-2930.
6. Rothwell P.M. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation // Lancet, Vol. 365, No. 9454, 2005. pp. 176-186.
7. Sleight P. Debate: Subgroup analyses in clinical trials — fun to look at, but don't believe them! // Current Controlled Trials in Cardiovascular Medicine, Vol. 1, No. 1, 2000. pp. 25-27.
8. Sun X., Briel M., Walter S.D., Guyatt G.H. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses // British Medical Journal, Vol. 340, 2010. pp. 850-854.
9. Wang R., Lagakos S.W., Ware J.H., Hunter D.J., Drazen J.M. Statistics in Medicine - Reporting of Subgroup Analyses in Clinical Trials // The New England Journal of Medicine, Vol. 357, 2007. pp. 2189-2194.
10. Lipkovich I., Dmitrienko A., DAgostino R.B. Tutorial in Biostatistics: Data-Driven Subgroup Identification and Analysis in Clinical Trials // Statistics in Medicine, Vol. 36, No. 1, 2016. pp. 136-196.
11. Meinshausen N., Meier L., Bühlmann P. p-Values for High-Dimensional Regression // Journal of the American Statistical Association, Vol. 104, No. 488, 2009. pp. 1671-1681.
12. Lockhart R., Taylor J., Tibshirani R.J., Tibshirani R. A Significance Test for the Lasso // Annals of Statistics, Vol. 42, No. 2, 2014. pp. 413-468.
13. Freidlin B., Simon R. Adaptive Signature Design: An Adaptive Clinical Trial Design for Generating and Prospectively Testing A Gene Expression Signature for Sensitive Patients // Clinical Cancer Research, Vol. 11, No. 21, 2005. pp. 7872-7878.
14. Meinshausen N., Bühlmann P. Stability selection // Journal of the Royal Statistical Society, Series B, Vol. 72, No. 4, 2010. pp. 417-423.
15. Gunter L., Zhu J., Murphy S. Variable Selection for Qualitative Interactions in Personalized Medicine while Controlling The Familywise Error Rate // Journal of Biopharmaceutical Statistics, Vol. 21, No. 6, 2011. pp. 1063-1078.
16. Simon R.M., Subramanian J., Li M.C., Menezes S. Using cross-validation to evaluate predictive accuracy of survival risk classifiers based on high-dimensional data // Briefings in Bioinformatics, Vol. 12, No. 3, 2011. pp. 203-214.
17. Good P. Resampling Methods: a Practical Guide to Data Analysis. 3rd ed. Boston: Birkhauser, 2005.

18. Loh W.Y., Shih Y.S. Split Selection Methods for Classification Trees // Statistica Sinica, Vol. 7, 1997. pp. 815-840.
19. Hothorn T., Hornik K., Zeileis A. Unbiased Recursive Partitioning: A Conditional Inference Framework // Journal of Computational and Graphical Statistics, Vol. 15, No. 3, 2006. pp. 651-674.
20. Hesterberg T., Moore D.S., Monaghan S., Clipson A., R. E. Bootstrap methods and permutation tests. Vol 5. // In: The Practice of Business Statistics / Ed. by Moore D.S. W. H. Freeman, 2005. pp. 1-70.
21. Varma S., Simon R. Bias in error estimation when using cross-validation for model selection // BMC Bioinformatics, Vol. 7, 2006. P. 91.
22. Foster J.C., Taylor J., Ruberg S.J. Subgroup identification from randomized clinical trial data. // Statistics in Medicine, Vol. 30, No. 24, 2011. pp. 2867-2880.
23. Dixon D.O., Simon R. Bayesian Subset Analysis // Biometrics, Vol. 47, No. 3, 1991. pp. 871-881.
24. Berger J.O., Wang X., Shen L. A Bayesian Approach to Subgroup Identification // Journal of Biopharmaceutical Statistics, Vol. 24, No. 1, 2014. pp. 110-129.
25. Xu Y., Trippa L., Müller P., Ji Y. Subgroup-Based Adaptive (SUBA) Designs for Multi-Arm Biomarker Trials // Statistics in Biosciences, Vol. 8, No. 1, 2016. pp. 159-180.
26. Xu Y., Yu M., Zhao Y.Q., Li Q., Wang S., Shao J. Regularized Outcome Weighted Subgroup Identification for Differential Treatment Effects // Biometrics, Vol. 71, No. 3, Sep 2015. pp. 645-653.
27. Little R.J., Rubin D.R. Causal effects in clinical and epidemiological studies via potential outcomes. // Annual Review of Public Health, Vol. 21, 2000. pp. 121-145.
28. Cox D.R. Regression Models and Life-Tables // Journal of the Royal Statistical Society. Series B, Vol. 34, No. 2, 1972. pp. 187-220.
29. Breiman L., Friedman J.H., Olshen R.A., Stone C.J. Classification and Regression Trees. Wadsworth: Belmont, CA, 1984.
30. Royston P., Altman D.G. Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling // Journal of the Royal Statistical Society. Series C, Vol. 43, No. 3, 1994. pp. 429-467.
31. Royston P., Sauerbrei W. A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials // Statistics in Medicine, Vol. 23, No. 16, 2004. pp. 2509-2525.
32. Tibshirani R. Regression shrinkage and selection via the lasso. // Journal of the Royal Statistical Society. Series B, Vol. 58, No. 1, 1996. pp. 267-288.
33. Imai K., Ratkovic M. Estimating Treatment Effect Heterogeneity in Randomized Program Evaluation // Tha Annals of Applied Statistics, Vol. 7, No. 1, 2013. pp. 443-470.
34. Cai T., Tian L., Peggy H W., Wei L.J. Analysis of randomized comparative clinical trial data for personalized treatment selections // Biometrics, Vol. 12, No. 2, 2011. pp. 270-282.
35. Song X., Pepe M.S. Evauating Markers for Selecting a Patient's Treatment // Biometrics, Vol. 60, No. 4, 2004. pp. 874-883.
36. Huang Y., Gilbert P.B., Janes H. Assessing Treatment-Selection Markers using a Potential Outcomes Framework // Biometrics, Vol. 68, No. 3, 2012. pp. 687-696.
37. Zhao L., Tian L., Cai T., Claggett B., Wei L.J. Effectively Selecting a Target Population for a Future Comparative Study // Journal of the American Statistical Association, Vol. 108, No. 502, 2013. pp. 527-539.
38. Breiman L. Random forests // Machine Learning, Vol. 45, No. 1, 2001. pp. 5-32.
39. Dusseldorp E., Conversano C., Van Os B.J. Combining an Additive and Tree-Based Regression Model Simultaneously: STIMA // Journal of Computational and Graphical Statistics, Vol. 19, No. 3, 2010. pp. 514-530.
40. Hodges J.S., Cui Y., Sargent D.J., Carlin B.P. Smoothing Balanced Single-Error-Term Analysis of Variance // Technometrics, Vol. 49, No. 1, 2007. pp. 12-25.
41. Gu X., Yin C., Lee J.J. Bayesian Two-step Lasso Strategy for Biomarker Selection in Personalized Medicine Development for Time-to-Event Endpoints // Contemporary Clinical Trials, Vol. 36, No. 2, 2013. pp. 642-650.
42. Negassa A., Ciampi A., Abrahamowicz M., Shapiro S., Boivin J.F. Tree-structured subgroup analysis for censored survival data: Validation of computationally inexpensive model selection criteria // Statistics and Computing, Vol. 15, No. 3, 2005. pp. 231-239.
43. Su X., Tsai C.L., Wang H., Nickerson D.M., Li B. Subgroup Analysis via Recursive Partitioning // Journal of Machine Learning Research, Vol. 10, 2009. pp. 141-158.
44. Su X., Zhou T., Yan X. Interaction Trees with Censored Survival Data // The International Journal of Biostatistics, Vol. 4, No. 1, 2008. P. 2.
45. Loh W.W., He X., Man M. A regression tree approach to identifying subgroups with differential treatment effects // Statistics in Medicine, Vol. 34, No. 11, 2015. pp. 1818-1833.
46. Loh W.Y., Fu H., Man M., Champion V., Yu M. Identification of subgroups with differential treatment effects for longitudinal and multiresponse variables // Statistics in Medicine, Vol. 35, No. 26, 2016. pp. 4837-4855.
47. Loh W.Y. Regression Trees with Unbiased Variable Selection and Interaction Detection // Statistica Sinica, Vol. 12, 2002. pp. 361-386.
48. Zeileis A., Hothorn T., Hornik K. Model-based Recurcive Partitioning // Journal of Computational and Graphical Statistics, Vol. 17, No. 2, 2008. pp. 492-514.
49. Dusseldorp E., Mechelen I.V. Qualitative Interaction Trees: a tool to identify qualitative treatment-subgroup interactions // Statistics in Medicine, Vol. 33, No. 2, 2014. pp. 219-237.
50. Tian L., Alizadeh A.A., Gentles A.J., Tibshirani R. A Simple Method for Estimating Interactions between a Treatment and a Large Number of Covariates // Journal of the Americal Statistical Association, Vol. 109, No. 508, 2014. pp. 1517-1532.
51. Jones H.E., Ohlssen D.I., Neuenschwander B., Racine A., Branson M. Bayesian Models for Subgorup Analysis in Clinical Trials // Clinical Trials, Vol. 8, No. 2, 2011. pp. 129-143.

52. Qian M., Murphy S.A. Performance guarantees for individualized treatment rules // The Annals of Statistics, Vol. 39, No. 2, 2011. pp. 1180-1210.
53. Zhao Y., Zheng D., Rush A.J., Kosorok M.R. Estimating individualized treatment rules using outcome weighted learning. // Journal of the American Statistical Association, Vol. 107, No. 449, 2012. pp. 1106-1118.
54. Lu W., Zhang H.H., Zeng D. Variable Selection for Optimal Treatment Decision // Statistical Methods in Medical Research, Vol. 22, No. 5, 2013. pp. 493-504.
55. Foster J., Taylor J.M.G., Kaciroti N., Nan B. Simple subgroup approximations to optimal treatment regimes from randomized clinical trial data // Biostatistics, Vol. 16, No. 2, 2015. pp. 368-382.
56. Zhang B., Tsiatis A.A., Davidian M., Zhang M., Laber E. Estimating Optimal Treatment Regimes from a Classification Perspective // Statistics, Vol. 1, No. 1, 2012. pp. 103-114.
57. Zhang B., Tsiatis A.A., Laber E.B., Davidian M. Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions // Biometrika, Vol. 100, No. 3, 2013. pp. 681-694.
58. Laber E.B., Zhao Y.Q. Tree-base methods for individualized treatment regimes // Biometrika, Vol. 102, No. 3, 2015. pp. 501-514.
59. Zhang Y., Laber E.B., Tsiatis A., Davidian M. Using decision lists to construct interpretable and parsimonious treatment regimes // Biometrics, Vol. 71, No. 4, 2015. pp. 895-904.
60. Fu H., Zhou J., Faries D.E. Estimating optimal treatment regimes via subgroup identification in randomized control trials and observational studies // Statistics in Mdeicine, Vol. 35, No. 19, 2016. pp. 3285-3302.
61. Lo V.S.Y. The true lift model: a novel data mining approach to response modeling in database marketing. // SIGKDD Explorations, Vol. 4, No. 2, 2002. pp. 78-86.
62. Larsen K. Net lift models: optimizing the impact of your marketing. // Predictive analytics world. 2011. Vol. Workshop presentation.
63. Robins J.M. Correcting for non-compliance in randomized trials using structural nested mean models // Communications in Statistics - Theory and Methods, Vol. 23, No. 8, 1994. pp. 2379-2412.
64. Robins J., Rotnitzky N. Estimation of treatment effects in randomised trials with non-comliance and a dichotomous outcome using structural mean models // Biometrika, Vol. 91, No. 4, 2004. pp. 763-783.
65. Jaskowski M., Jaroszewicz S. Uplift modeling for clinical trial data // ICML, 2012 workshop on machine learning for clinical data analysis. Edinburgh. Scotland. 2012.
66. Radcliffe N.J., Surry P.D. Differential analysis: modeling true response by isolating the effect of a single action. // Proceedings of credit scoring and credit control VI. 1999.
67. Radcliffe N.J., Surry P.D. Real-world uplift modeling with significance-based uplift trees., Portrait Technical Report TR-2011-1, stochastic solutions, Tech. rep. 2011.
68. Hansotia B., Rukstales B. Incremental Value Modeling // Journal of Interactive Marketing, Vol. 16, No. 3, 2002. pp. 35-46.
69. Chickering D.M., Heckerman D. A decision theoretic approach to targeted advertising. // Proceedings of the 16th conference in uncertainty in artificail intelligence (UAI'00). 2000. pp. 82-88.
70. Rzepakowski P., Jaroszewicz S. Decision trees for uplift modeling // Proceedings of the 10th IEEE International conference on data mining (ICDM). Sydney. Australia. 2010. pp. 441-450.
71. Rzepakowski P., Jaroszewicz S. Decision trees for uplift modeling wth single and multiple treatments // Knowledge and Information Systems, Vol. 32, No. 2, 2012. pp. 303-327.
72. Kuusisto F., Costa V.S., Nassif H., Burnside E., Page D., Shavlik J. Support vector machines for differenctial prediction // Proceedings of the ECML-PKDD. 2014.
73. Jaroszewicz S., L. Zaniewicz Ł. Székely regularization for uplift modeling. // In: Challenges in computational statistics and data mining. Springer International Publishing, 2016. pp. 135-154.
74. Zaniewicz L., Jaroszewicz S. Lp - Support vector machines for uplift modeling // Knowledge and Information Systems, Vol. 53, No. 1, 2017. pp. 269-296.
75. Guelman L., Guillen M., Perez-Marin A.M. Random forests for uplift modeling: an insurance customer retention case. Vol 115. // In: Modeling and simulation in engingeering, economics and management. Springer, Berlin, 2012. pp. 123-133.
76. Soltys M., Jaroszewicz S., Rzepakowski P. Ensemble methods for uplift modeling // Data Mining and Knowledge Discovery, Vol. 29, No. 6, 2015. pp. 1531-1559.
77. Chen G., Zhong H., Belousov A., Devanarayan V. A PRIM approach to predictive-signature development for patient stratification // Statistics in Medicine, Vol. 34, No. 2, 2015. pp. 317-342.
78. Lipkovich I., Dmitrienko A., Denne J., Enas G. Subgroup identification based on differential effect search—A recursive partitioning method for establishing response to treatment in patient subpopulations // Statistics in Medicine, Vol. 30, No. 21, 2011. pp. 2601-2621.
79. Friedman J.H., Fisher N.I. Bump Hunting in High-Dimensional Data // Statistics and Computing, Vol. 9, No. 2, 1999. pp. 123-243.
80. Sivaganesan S., Laudb P.W., Müller P. A Bayesian subgroup analysis with a zero-enriched Polya Urn scheme // Statistics in Medicine, Vol. 30, No. 4, 2010. pp. 312-323.
81. Korepanova N., Kuznetsov S.O., Karachunskiy A.I. Matchings and Decision Trees for Determining Optimal Therapy // In: Analysis of Images, Social Networks and Texts Third International Conference, AIST 2014, Yekaterinburg, Russia, April 10-12, 2014, Revised Selected Papers. Springer International Publishing, 2014. pp. 101-110.
82. Gale D., Shapley L.S. College Admissions and the Stability of Marriage // The American Mathematical Monthly, Vol. 69, No. 1, 1962. pp. 9-15.
83. Roth A.E. Differed acceptance algorithm: history, theory, practice, and open questions, Harvard University, 2007.

84. Alkan A., Gale D. Stable schedule matching under revealed preference // Journal of Economic Theory, Vol. 112, 2003. pp. 289-306.
85. Ganter B., Kuznetsov S.O. Pattern Structures and Their Projections // 9th International Conference on Conceptual Structures (ICCS 2001). 2001. Vol. 2120. pp. 129-142.
86. Kuznetsov S.O. Pattern Structures for Analyzing Complex Data // 12th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing (RSFDGrC 2009). 2009. Vol. 5908. pp. 33-44.
87. Korepanova N., Kuznetsov S.O. Pattern Structures for Treatment Optimization // In: CLA 2016: Proceedings of the Thirteenth International Conference on Concept Lattices and Their Applications. CEUR Workshop Proceedings. Moscow: Higher School of Economics, National Research University, 2016. pp. 217-228.
88. Korepanova N.V., Kuznetov S.O. Vybor terapii onkologicheskogo zabolevaniya v podgruppakh patsientov na osnove analiza zamknutykh opisaniy [The choice of therapy for oncological disease in subgroup of patient on the basis of the analysis of close descriptions]. // In: Pyatnadtsataya natsionalnaya conferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem KII-2016 [15th national conference on artificial intelligence with international participation CAI-2016] (October 3-7, 2016, Smolensk): Vol 1. Universum, Smolensk. pp. 352-359.

Korepanova N.V. International Laboratory for Intelligent Systems and Structural Analysis, National Research University Higher School of Economics, Moscow, Russia. E-mail: nkorepanova@hse.ru