

К вопросу о первичной профилактике гипертонической болезни

Г.И. Образцова, В.М. Кочетков

Санкт-Петербургская педиатрическая академия, кафедра терапии

Санкт-Петербургская медицинская академия последипломного образования, кафедра педиатрии № 3

Резюме

В качестве одного из инструментов первичной профилактики рассматриваются возможные подходы к определению риска развития заболевания на основе анализа неблагоприятных сопутствующих факторов. Описано применение метода Байеса, критерия значимости и метода бинарной логистической регрессии и дано их сравнение. Приведены результаты оценки риска развития гипертонической болезни у детей для различных комбинаций неблагоприятных факторов.

Ключевые слова: первичная профилактика, теорема Байеса, критерий значимости, бинарная логистическая регрессия.

On the problem of primary prevention of hypertension

G.I. Obratsova, V.M. Kochetkov

St Petersburg State Pediatric Medical Academy, Department of Therapeutics

St Petersburg Medical Academy of Postgraduate Studies, Department of Pediatrics № 3

Resume

Some approaches to estimation of disease risk by means of unfavorable attendant factors analysis are considered as one of initial preventive methods. Bayesian theorem, significance test and binary logistic regression are shortly described and compared. Calculation results concerning hypertension in children for various combinations of unfavorable factors are presented.

Key words: initial prevention, Bayesian theorem, significance test, binary logistic regression.

Статья поступила в редакцию: 03.11.08. и принята к печати: 08.12.08.

Введение

Значительная часть взрослого контингента больных эссенциальной гипертензией, как известно, формируется из детей и подростков с повышенным артериальным давлением (АД) [1], в то же время распространенность артериальной гипертензии (АГ) в детском и подростковом возрасте составляет, по данным разных авторов [1–3], значительную величину — от 5 % до 20 %.

Известно также, что наличие АГ у детей и взрослых характеризуется появлением факторов риска, имеющих наследственную природу либо являющихся результатом средового влияния. Число таких факторов ограничено (15–20), и принято считать, что чем больше их у наблюдаемого человека, тем выше риск развития заболевания. Однако методы оценки индивидуального риска развития заболевания на основе интегральной совокупности факторов риска разработаны недостаточно. Практическому врачу важно иметь в распоряжении показатель, вычисленный на основе одного из статистических методов и дающий возможность оценить суммарный вероятностный риск развития заболевания. Наличие подобного показателя позволило бы выделить среди пациентов тех лиц, у которых высок риск развития заболевания. Такая процедура способствовала бы разработке систем первичной про-

филактики перечисленных заболеваний, являющихся основной причиной инвалидности и преждевременной смертности населения.

Ниже предлагаются несколько подходов, позволяющих методами математической статистики определить риск заболевания на основе учета совокупности неблагоприятных факторов. Указанные подходы могут иметь широкую область применения в первичной профилактике различных заболеваний, в настоящей работе один из таких подходов проиллюстрирован, в частности, на практическом примере анализа факторов риска АГ у детей.

Перейдем к формулированию задачи в статистических терминах. Пусть для двух обследуемых групп — больных, имеющих АГ, и контрольной группы — проведена оценка наличия или отсутствия учитываемых факторов риска. Если объем выборки достаточно велик (150–200 или более), то результаты проведенного обследования позволяют найти количественную оценку риска развития заболевания у любого пациента, у которого наличие этих факторов известно. В отношении самих факторов будем предполагать, что они имеют бинарную природу: фактор у пациента либо проявляется, либо нет. К такому виду данных относится, например, наличие или отсутствие определенных признаков в анамнезе (семейном, аллергологическом, акушерском, гинекологическом

и т. п.). В дальнейшем для краткости вероятность риска развития заболевания будем называть просто риском.

Для описания процедуры расчета риска введем следующие обозначения: n_3 и n_B — число обследованных пациентов соответственно в группе здоровых и больных (объем выборки), K — число факторов. Пусть далее i — списочный номер пациента и m — номер фактора. Каждый пациент с номером i характеризуется присущим ему набором различных факторов m , который описывается рядом индикаторов δ_{im} : будем считать, что $\delta_{im} = 1$, если фактор m проявлен, и $\delta_{im} = 0$ в противном случае.

Результатом обследования обеих групп оказываются две таблицы, состоящие из наборов индикаторов для здоровых и больных и имеющие количество столбцов K (номера факторов) и количество строк, соответствующих объемам выборки n_3 и n_B . В дальнейшем указанные таблицы будем называть соответственно T_3 и T_B .

Пусть теперь обследуется пациент с набором индикаторов δ_m , $1 \leq m \leq K$ (для обследуемого пациента набор индикаторов имеет лишь один индекс, указывающий номер фактора). Задача состоит в том, чтобы на основе таблиц T_3 и T_B определить для такого пациента риск заболевания.

Оценка искомого значения риска может быть произведена несколькими способами. Мы предлагаем для этой цели применить один из трех подходов: метод, основанный на теореме Байеса, критерий значимости и метод логистической регрессии. Перейдем к их краткому описанию.

Метод, использующий теорему Байеса

Если проявления факторов считаются статистически независимыми, то из таблицы T_B для больных пациентов можно найти частоту проявления фактора с номером m :

$$P(m, B) = \frac{1}{n_B} \sum_{i=1}^{n_B} d_{im}, \quad 1 \leq m \leq K, \quad (1)$$

а по таблице T_3 — соответствующую величину для контрольной группы (здоровых):

$$P(m, 3) = \frac{1}{n_3} \sum_{i=1}^{n_3} d_{im}, \quad 1 \leq m \leq K. \quad (2)$$

Если частоты $P(m, B)$ и $P(m, 3)$ рассчитаны для всех факторов, то, согласно теореме Байеса [5], величина риска может быть рассчитана по формуле:

$$R = \frac{z \cdot P(B)}{z \cdot P(B) + (1 - z)P(3)}, \quad (3)$$

где z — априорная вероятность заболевания, характеризующая общую его распространенность в популяции¹, а величины $P(B)$ и $P(3)$ определяются формулами:

$$P(B) = \prod_{m=1}^K \{P(m, B)d_m + [1 - P(m, B)](1 - d_m)\}, \quad (4)$$

$$P(3) = \prod_{m=1}^K \{P(m, 3)d_m + [1 - P(m, 3)](1 - d_m)\}, \quad (5)$$

учитывающими факторы риска обследуемого пациента.

Формулы (4) и (5) справедливы, если факторы статистически независимы. Наличие или отсутствие такой независимости определяется расчетом коэффициентов взаимной корреляции на основе таблиц T_B и T_3 либо с помощью известных компьютерных статистических программ, например, популярного пакета SPSS (Statistical Package for the Social Sciences).

Если статистической независимости нет, то формула (3) для риска R остается в силе, но величины $P(B)$ и $P(3)$ уже не могут рассчитываться по формулам (4) и (5). В этом случае вероятность комбинации наборов индикаторов δ_m должна оцениваться непосредственно по таблицам T_B и T_3 без использования формул (1) и (2), определяющих частоты факторов. При этом целесообразно учитывать статистическую зависимость лишь для тех факторов, которым в соответствии с табличными данными отвечает высокая вероятность реализации. Для прочих факторов можно подобную зависимость не учитывать, что резко упрощает расчеты.

Для риска R , оцениваемого по формуле (3), можно найти границы доверительного интервала, который зависит прежде всего от априорной вероятности z и объемов выборки n_3 и n_B . Как показывают расчеты, при объемах выборки порядка 200 и выше расчет по формуле (3) дает достаточно точные результаты.

Пример практического использования описанного подхода рассматривается ниже.

Применение критерия значимости

Оценку риска можно производить и другим методом, который мы кратко опишем для случая статистической независимости факторов. Метод реализуется наиболее просто, если используются данные лишь таблицы T_B . В этом случае знание частот проявления факторов (1) позволяет оценить по формуле (4) вероятность реализации для любого набора индикаторов. Общее число возможных наборов составляет 2^K , а сумма вероятностей реализации всех наборов равна единице.

Применение критерия значимости предполагает выбор некоторой вероятности α (называемой уровнем значимости и равной обычно 0,1 или 0,05), с которой гипотеза риска может быть ошибочно отвергнута. Обычно критерий значимости используется для случайных величин, имеющих непрерывное распределение, и подобному его применению посвящена обширная литература [6]. Попытка использовать критерий значимости для определения риска заболевания наталкивается на два осложнения: во-первых, число комбинаций факторов риска конечно, и, во-вторых, множество этих комбинаций нельзя ранжировать по величине в силу того, что это множество является неупорядоченным. Тем не менее,

¹ Эта величина определяется по статистическим данным поликлиник и больниц.

можно предложить вариант использования критерия значимости и в рассматриваемом случае.

Для реализации этого метода по формуле (4) рассчитываются вероятности $P(B)$, относящиеся ко всем возможным комбинациям наборов индикаторов, после чего отбирается их часть, отвечающая наименее вероятным комбинациям, для которой суммарная вероятность равна заданному уровню значимости α . Если набор индикаторов для обследуемого пациента не попадает в эту часть, то гипотеза относительно риска заболевания принимается при уровне значимости α . В противном случае гипотеза отвергается.

Описанный метод можно обобщить как в направлении учета данных обеих таблиц T_B и T_3 , так и учета статистической зависимости факторов.

Практическому применению описанного метода способствует то обстоятельство, что при большом числе факторов (более пяти) функция распределения для вероятностей проявления факторов оказывается весьма близкой к нормальному закону. Это упрощает оценку критической области, поскольку квантили нормального распределения табулированы и приводятся во многих руководствах по математической статистике [5].

Использование логистической регрессии

Идея метода состоит в том, что для учитываемых факторов на основе анализа строк таблиц T_B и T_3 находят области, в которых проявления факторов характерны соответственно для больных или здоровых. С этой целью строится логистическая зависимость вида

$$F(q_1, q_2, \dots, q_K) = \frac{1}{1 + \exp[-\sum_{m=1}^K q_m d_m]},$$

где $q_1, q_2, \dots, q_K$ — неизвестные коэффициенты, подлежащие определению. Эти коэффициенты на основе таблиц T_B и T_3 подбираются так, чтобы для наборов факторов, характерных для больных, величина F была близка к единице, а для здоровых соответственно близка к нулю.

Процедура поиска коэффициентов $q_1, q_2, \dots, q_K$ использует стандартные математические алгоритмы [7], а также реализуется в ряде компьютерных статистических пакетов, например, SPSS.

После того как искомые коэффициенты найдены, величина F рассчитывается для обследуемого пациента и далее определяется, к какому из значений — нулю или единице — она ближе. В итоге принимается решение о наличии или отсутствии риска.

В связи с тем, что величина F по существу рассматривается как бинарная (поскольку принимает лишь два значения), метод обычно называют методом бинарной логистической регрессии.

Преимущество описанного метода состоит в том, что статистическая зависимость факторов, если она есть, учитывается автоматически, а недостаток — в невозможности оценить точность и количественную достоверность

результата. Примеры использования метода бинарной логистической регрессии приведены в книге «Математическая статистика» Бикела П. и соавт. (1983).

Сравнение расчетных методов

Если априорная вероятность заболевания z известна, то метод Байеса имеет неоспоримые преимущества: он сравнительно прост в отношении объема проводимых расчетов, обеспечивает оценку доверительного интервала для риска и позволяет учитывать взаимовлияние факторов. Метод Байеса несложно обобщить также на случай оценки риска для нескольких видов заболеваний.

Однако применить метод Байеса невозможно, если априорная вероятность z заболевания неизвестна. В этом случае остаются два метода — критерий значимости и бинарная логистическая регрессия. Оба указанных метода имеют одинаковую область применения, однако преимуществом первого является четкое определение достоверности получаемого результата, а преимуществом второго — наличие хорошо разработанных программных средств. Именно эти обстоятельства должны играть решающую роль при выборе метода решения задачи.

Если число учитываемых факторов превышает три, то ручные расчеты становятся весьма громоздкими и для получения результата необходимы специализированные компьютерные программы.¹

Пример опытного изучения факторов риска АГ у детей

В настоящем разделе приводятся результаты обследования детей и подростков на предмет выявления и оценки факторов риска развития АГ.

В отношении АГ значение может иметь профилактическое обследование детей, так как не подлежит сомнению факт более высокой эффективности профилактики и лечения АГ на ранних этапах ее становления, в частности, в детстве, а не на стадии стабилизации и органических повреждений.

При исследовании проводилось анкетирование детей, разделенных на две группы. Первую группу составили 190 детей и подростков в возрасте от 7 до 18 лет (133 мальчика и 57 девочек) с первичной АГ. Во вторую группу (контрольную) вошли 193 учащихся средней школы 1–11 классов (93 мальчика и 100 девочек). Предложенная для заполнения анкета включала в себя вопросы по следующим разделам: данные о сердечно-сосудистой патологии у членов семьи (в том числе у родственников второй степени родства), акушерско-гинекологический анамнез матери, информация о родовом акте, данные периода новорожденности, анамнеза жизни детей, сведения об учебных, физических нагрузках и характеристиках сна.

Диагноз АГ ставился в тех случаях, когда при казуальных (случайных) измерениях АД у ребенка неоднократно (не менее трех раз при отдельных визитах к врачу) отмечался повышенный уровень АД. Уровень АД считался повышенным, если показатель

¹ Такие программы могут быть предоставлены авторами настоящей работы, при желании их использовать можно обращаться по телефону (812) 558–6414 или по адресу электронной почты v.kochetkov@bk.ru.

систолического или диастолического АД был равен или превышал значение 95-го перцентиля с учетом пола, возраста и длины тела ребенка в соответствии с рекомендациями по оценке казуальных измерений АД, сформулированных Американской рабочей группой по контролю АГ у детей и подростков [4]. В результате проведения комплексного клинического и инструментального обследования детей с повышенным уровнем АД был исключен вторичный (симптоматический) характер имеющейся у них АГ.

Достоверность разницы между средними значениями в оцениваемой и контрольной группах оценивалась по критерию Стьюдента, а достоверность разницы между долями (частотами) — по методу углового преобразования Фишера. При этом рассчитанная вероятность фактического наличия разницы в обоих случаях составила более 0,99 для всех учитываемых факторов, что делает частоту проявления факторов показательной для абсолютного факториального различия между больными и здоровыми пациентами.

В приведенной ниже табл. 1 представлены данные сравнительного анализа анкет и объективного обследования детей с повышенным уровнем АД и контрольной группы (приведены только статистически достоверные различия).

Из приведенных в таблице данных видно, что в группе детей с повышенным уровнем АД значительно чаще, чем в контрольной группе, встречается наследственная отягощенность по гипертонической болезни (лабильная и стабильная АГ у родителей, гипертоническая болезнь у дедушек и бабушек). Аналогичные различия встречаются и в отношении таких сердечно-сосудистых осложнений,

как инфаркт миокарда и инсульт у дедушек и бабушек. При сравнении данных, характеризующих особенности течения беременности у матерей обследованных детей, также были обнаружены существенные различия. Так, у матерей детей с АГ, чаще, по сравнению с матерями детей контрольной группы, отмечались следующие патологические состояния: хронические гинекологические заболевания, АГ, избыточная прибавка веса, протеинурия и сахарный диабет. В течение родового периода у матерей детей с диагностированной АГ чаще применялась лекарственная стимуляция родовой деятельности, использовались акушерские щипцы. При рождении таких детей чаще, по сравнению с контрольной группой, возникала родовая травма и отмечались признаки гипоксического поражения центральной нервной системы. Кроме того, у детей с АГ значительно чаще, чем у обследованных школьников, отмечались жалобы на нарушения сна, чаще встречались обморочные и предобморочные состояния. В анамнезе у таких детей чаще присутствовали указания на сотрясения головного мозга и применение общей анестезии при оперативных вмешательствах.

При оценке антропометрических данных было установлено, что среди детей с АГ в сравнении с контрольной группой значительно большее количество детей имеет избыточную массу тела и страдает ожирением.

Таким образом, в проведенных исследованиях был получен ряд факторов, связанных с риском развития АГ у детей и подростков. На базе этих исследований оценка наличия неблагоприятных факторов у конкретного пациента — ребенка или подростка — позволяет по описанным выше методикам определить для него степень риска развития АГ.

Таблица 1

ЧАСТОТЫ ПРОЯВЛЕНИЯ РАЗЛИЧНЫХ ФАКТОРОВ РИСКА У ДЕТЕЙ С АРТЕРИАЛЬНОЙ ГИПЕРТЕНЗИЕЙ И В КОНТРОЛЬНОЙ ГРУППЕ ПО ДАННЫМ АНКЕТИРОВАНИЯ И КЛИНИЧЕСКОГО ОБСЛЕДОВАНИЯ

Изученные факторы	Дети с АГ (n _б = 190)	Контрольная группа (n _к = 193)
*Повышение АД у матери	59 %	24 %
*Повышение АД у отца	50 %	21 %
Гипертоническая болезнь у дедушек и бабушек	50 %	25 %
Наличие инфарктов миокарда у дедушек и бабушек	36 %	16 %
Инсульты у дедушек и бабушек	39 %	24 %
Хронические гинекологические заболевания у матери до и в период беременности	19 %	5 %
АГ у матери в период беременности	27 %	16 %
Избыточная прибавка веса у матери во время беременности	35 %	22 %
Протеинурия у матери во время беременности	19 %	8 %
Сахарный диабет у матери до и во время беременности	3,8 %	0,5 %
Лекарственная стимуляция родовой деятельности у матери	30 %	21 %
Применение акушерских щипцов в родах	3,5 %	0,5 %
Гипоксия и травма в родах	27 %	10 %
Шестидневное обучение (у детей 14–17 лет)	65 %	34 %
Нагрузка более 7 уроков в день (у детей 14–17 лет)	22 %	1 %
Обморочные и предобморочные состояния	26 %	9 %
Количество ночного сна менее 8 часов в сутки (у детей 14–17 лет)	58 %	28 %
**Нарушения сна	57 %	30 %
Сотрясение головного мозга	24 %	11 %
Оперативные вмешательства под наркозом	31 %	11 %
Избыточная масса тела	23 %	10,7 %
Ожирение	13 %	4 %

Примечания: АГ — артериальная гипертензия; АД — артериальное давление;

* — имеются в виду как стабильная АГ, так и эпизоды повышения АД у родителей;

** — трудное засыпание, чуткий поверхностный сон, снохождение, снохождение, страшные сновидения.

Поскольку априорная вероятность наличия АГ у детей известна и лежит в диапазоне $z = 0,05-0,20$, то оказалось возможным применить подход, основанный на методе Байеса. В связи с этим расчет риска R производился по формуле (3) с использованием соотношений (1), (2), (4) и (5).

Приведем результаты расчета риска АГ, относящиеся к различным сочетаниям обнаруженных неблагоприятных факторов у гипотетического пациента. При этом последовательно рассмотрим три варианта.

Вариант 1. Допустим, что у тестируемого ребенка выявлены следующие факторы: повышение АД у матери, повышение АД у отца, инсульты у дедушек и бабушек, хронические гинекологические заболевания у матери в период беременности.

Этому варианту отвечает следующая расчетная таблица, где в первом столбце даны возможные значения априорной вероятности заболевания или, иными словами, показателя распространенности заболевания в популяции (табл. 2).

Таблица 2

ОПРЕДЕЛЕНИЕ РИСКА РАЗВИТИЯ АРТЕРИАЛЬНОЙ ГИПЕРТЕНЗИИ У РЕБЕНКА ПРИ НАБОРЕ НЕБЛАГОПРИЯТНЫХ ФАКТОРОВ, ВАРИАНТ № 1

z	R (риск развития АГ)	$1-R$ (вероятность избежать АГ)
0,05	0,03	0,97
0,10	0,05	0,95
0,15	0,08	0,92
0,20	0,11	0,89

Примечания: АГ — артериальная гипертензия; z — априорная вероятность.

В последующих вариантах к упомянутым факторам добавляются новые.

Вариант 2. То же, что в предыдущем случае, но зафиксированы два новых фактора: сахарный диабет у матери, гипоксия и травма в родах. Результаты расчета для этого варианта приведены в табл. 3.

Таблица 3

ОПРЕДЕЛЕНИЕ РИСКА РАЗВИТИЯ АРТЕРИАЛЬНОЙ ГИПЕРТЕНЗИИ У РЕБЕНКА ПРИ НАБОРЕ НЕБЛАГОПРИЯТНЫХ ФАКТОРОВ, ВАРИАНТ № 2

z	R (риск развития АГ)	$1-R$ (вероятность избежать АГ)
0,05	0,41	0,59
0,10	0,59	0,41
0,15	0,70	0,30
0,20	0,76	0,24

Примечания: АГ — артериальная гипертензия; z — априорная вероятность.

Вариант 3. То же, что в варианте 2, но добавлены еще два фактора: шестидневное обучение школьника и отмечавшееся у ребенка сотрясение мозга. Полученные расчеты риска развития АГ представлены в табл. 4.

Как видно из табл. 2–4, увеличение количества зафиксированных неблагоприятных факторов приводит к весьма сильному росту риска развития АГ у ребенка.

Таблица 4

ОПРЕДЕЛЕНИЕ РИСКА РАЗВИТИЯ АРТЕРИАЛЬНОЙ ГИПЕРТЕНЗИИ У РЕБЕНКА ПРИ НАБОРЕ НЕБЛАГОПРИЯТНЫХ ФАКТОРОВ, ВАРИАНТ №3

z	R (риск развития АГ)	$1-R$ (вероятность избежать АГ)
0,05	0,86	0,14
0,10	0,93	0,07
0,15	0,95	0,05
0,20	0,97	0,03

Примечания: АГ — артериальная гипертензия; z — априорная вероятность.

Предложенные методы оценки риска при учете неблагоприятных факторов могут быть использованы для определения индивидуального риска развития первичной АГ у детей и подростков. Выявление среди детского контингента лиц с высоким риском развития АГ поможет целенаправленному проведению профилактики этого заболевания.

Литература

1. Доклад Комитета экспертов ВОЗ № 792, Профилактика в детском и юношеском возрасте сердечно-сосудистых заболеваний, проявляющихся в зрелые годы: время действовать; Женева. 1992.
2. Александров А.А. Повышенное артериальное давление в детском и подростковом возрасте (ювенильная артериальная гипертония). РМЖ. 1997;9:559–65.
3. Брызгунов И.П. Первичная артериальная гипертензия у детей и подростков. Вопросы современной педиатрии. 2003;2(3):68–71.
4. Updates on the 1987 task force report on high blood pressure in children and adolescents: a working group report from national high blood pressure education program. Pediatrics. 1996 Oct;98(4):649–50.
5. Агекян Т.А.. Теория вероятности для астрономов и физиков. М., Наука; 1974. с. 46.
6. Бикел П., Доксам К. Математическая статистика, вып. 2. М.: Финансы и статистика; 1983. 254 с.
7. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. М.: Наука; 1974. 832 с.