

Какому методу статистической обработки результатов следует отдавать предпочтение при оценке эффективности нового метода лечения?

S.L. Kabak, Yu.S. Kabak

Which method of statistical processing should be considered preferable while assessing effectiveness of treatment method?

В статьях клинической направленности, публикуемых в сборниках научных работ и на страницах республиканских периодических медицинских изданий, часто приводятся данные по оценке эффективности новых оригинальных методов лечения или содержатся сведения о факторах риска возникновения некоторых социально значимых болезней. При этом алгоритм проведения подобных исследований, как правило, сводится к следующей последовательности действий. Сначала получают количественные данные о динамике изменения биохимических показателей – маркеров тяжести патологического процесса или его активности у пациентов опытной группы, которым назначается испытуемый метод лечения. Рассчитываются средние арифметические значения показателей и стандартная ошибка среднего значения. Затем полученные результаты сравниваются со средним значением соответствующего показателя в контрольной группе, в которую входят пациенты, лечившиеся стандартным методом. При наличии статистически подтвержденного различия между двумя средними величинами (для этого используется критерий Стьюдента) делается вывод об эффективности нового метода лечения. При этом степень эффективности оценивается по тому, как близко (в процентном выражении) изучаемый показатель приблизился к нормальным значениям. Подобная аргументация имеет один существенный изъян. Совершенно не учитывает тот факт, что, во-первых, среди опытной группы, безусловно, имеются пациенты, которым лечение не помогло вовсе (их биохимические показатели активности процесса не изменились, а даже ухудшились). Во-вторых, в контрольной группе больных (даже не получавших лечения) без сомнения имеются пациенты, у которых объективно регистрируется улучшение состояния. Более того, авторами не производится количественная оценка индивидуальной выраженности изменений изучаемых показателей в процессе лечения.

Хотелось бы обратить внимание и еще на одно обстоятельство. Так, авторы приводят возраст пациентов, или продолжительность их болезни с указанием доверительного интервала (например, $40,1 \pm 3,5$ года), тогда как при оценке эффекта при альтернативной форме учета реакций (например, при указании процента больных, имеющих «хороший» результат лечения) сведения о границах доверительного интервала, а также оценка разности между долями отсутствуют. Это, на мой взгляд, с учетом существующих современных методов статистической обработки результатов научного исследования, не позволяет должным образом оценить эффективность нового метода лечения.

Настоящая публикация призвана оказать авторам методическую помощь в проведении адекватной статистической обработки клинического материала и на конкретном примере разобрать некоторые простейшие методы статистической обработки клинических данных, в том числе расчет показателей значимости полученных результатов, статистической достоверности, а также интерпретации силы ассоциации между элементами двух совокупностей.

Для решения задач по оценке эффективности нового метода лечения или поиска связи между клиническими исходами (обычно неблагоприятными) и действием этиологических факторов (обычно факторов риска) рекомендуется организовать исследование случай-контроль. Суть этого исследования состоит в том, что подбирается две группы пациентов с определенной патологией. Больные одной из этих групп (опытная группа) лечатся по новой технологии; больные второй группы (контрольная группа) получают общепринятое лечение. В каждой группе по объективным критериям (в том числе по положительной динамике ряда инструментальных и биохимических показателей – маркеров тяжести болезни) определяют количество пациентов с положительной динамикой и с отсутствием таковой. Данные заносят в таблицу сопряженности (табл.1).

Таблица

1

Таблица сопряженности 2x2

Исходы	Благоприятный	Неблагоприятный (нет положительной динамики + динамика отрицательная)	Всего
Опытная группа	a	b	a+b
Контрольная группа	c	d	c+d
Итого:	a+c	b+d	

При проведении эпидемиологического исследования с целью поиска факторов риска болезни может быть использована аналогичная таблица. В этом случае в опытную группу войдут пациенты, подверженные действию предполагаемого этиологического фактора, в контрольной группе окажутся пациенты, на которых этот фактор не действует. В графах благоприятный и неблагоприятный исход будет указано количество здоровых и больных людей соответственно.

Методика статистической обработки результатов по оценке эффективности нового метода лечения и при поиске связи между клиническими исходами и действием этиологических факторов одинаковая. В качестве иллюстрации для расчета конкретных показателей будут использованы данные эпидемиологического исследования, содержащиеся в статье De Moor and all. (2000). В ней автором делается попытка установить связь между качеством эндодонтического лечения зубов и наличием апикального периодонтита (табл.2).

Таблица

2

Таблица сопряженности для данных из статьи De Moor and all. (2000)

Уровень заполнения корневого канала	Норма (благоприятный исход)	Апикальный периодонтит (неблагоприятный исход)	Всего
0-2 мм от радиологической вершины (контрольная группа)	77 (а)	32 (b)	109
>2 мм от радиологической вершины ¹ (основная группа)	41 (с)	74 (d)	115
Итого:	118	106	224

Для установления степени различия между двумя группами (контрольной и опытной), приведенными в таблице сопряженности, можно использовать критерий χ^2 (В.А.Медик и соавт., 2000). По этому критерию обычно определяют, имеется ли согласие или налицо различия в эмпирическом (полученном исследователем), и теоретическом распределениях. Если полученный в ходе расчетов эмпирический показатель равен или меньше теоретического показателя (находится по соответствующей таблице), так называемая «нулевая» гипотеза не отвергается, т.е. в таком случае эмпирические и теоретические распределения не различаются. Наоборот, если вычисленный показатель больше полученного из таблицы, нулевая «гипотеза» опровергается и считается, что эмпирические статистики достоверно отличаются от теоретических, т.е. изучаемый фактор (в рассматриваемом примере – качество пломбировки корневого канала) статистически достоверно влияет на изучаемый результат (способствует возникновению апикального периодонтита).

Вычисление критерия проводится по формуле:

$$\chi^2 = \frac{(ad-bc)^2 \cdot n}{(a+b) \cdot (c+d) \cdot (a+c) \cdot (b+d)}, \text{ где } n=a+b+c+d$$

В рассматриваемом примере:

$$\chi^2 = 43862 \cdot 224 / 109 \cdot 115 \cdot 118 \cdot 106 = 4309087104 / 156787780 = 27,4.$$

Сравниваем полученный результат с табличными данными. При достоверном различии ($p < 0,05$) критерий χ^2 должен быть больше 3,84. Таким образом, рассчитанная величина коэффициента доказывает наличие статистически значимого различия между двумя распределениями (контрольной и основной группами).

Однако, критерий χ^2 не может служить мерой различий между контрольной и опытной группами. Для количественной оценки результата проведенных исследований следует использовать такой показатель из доказательной медицины, как отношение шансов (ОШ), который равен отношению шансов того, что событие произойдет (при не качественной пломбировке канала разовьется апикальный периодонтит), к шансам, что это событие не произойдет (В.В.Власов, 2001). Показатель рассчитывается по формуле:

$$\text{ОШ} = (a/c)/(b/d) = a \cdot d / b \cdot c, \text{ или } \text{ОШ} = 77 \cdot 74 / 32 \cdot 41 = 5698 / 1312 = 4,34$$

Отношение шансов используется для оценки относительного риска развития неблагоприятного исхода какой-либо медицинской манипуляции. При величине показателя больше 1, вероятность события в основной группе выше, чем в контрольной. Таким образом, ОШ, рассчитанное в конкретном исследовании, позволяет утверждать, что вероятность развития апикального периодонтита при некачественной пломбировке корневого канала в 4,3 раза выше, чем в случаях, когда уровень заполнения канала произведена правильно.

Как для любого статистического показателя для отношения шансов необходимо рассчитать его доверительные границы, чтобы с определенной долей вероятности утверждать, что истинное (популяционное) значение показателя находится в границах этого интервала.

Самый простой способ – воспользоваться калькулятором в Интернете (например, расположенном по адресу www.hutchon.freeserve.co.uk/ConfidOK.htm).

Можно провести расчет самостоятельно. Для этого необходимо осуществить следующую последовательность математических действий:

1. Найти натуральный логарифм числа, выражающего отношение шансов:
 $\ln OШ = \ln 4.34 = 1,46787$;

2. По формуле: $D \ln OШ = 1/a + 1/b + 1/c + 1/d$ вычислить дисперсию натурального логарифма отношения шансов:
 $D \ln OШ = 1/77 + 1/32 + 1/41 + 1/74 = 0,0815$;

3. По формуле: $SE OШ = 2 \sqrt{D \ln OШ}$ произвести расчет стандартной ошибки отношения шансов:
 $SEOШ = 2 \sqrt{D \ln OШ} = 2 \sqrt{0,0815} = 0,285$

4. Размеры доверительного интервала для отношения шансов определяются исходя из стандартной ошибки и коэффициента, который при 90% доверительном интервале равен 1,645, 95% интервале – 1,96, а при значении доверительной вероятности в 99% – 2,58. Задавая 95% доверительную вероятность, получаем границы доверительного интервала для значения натурального логарифма в виде натуральных логарифмов:
 $\ln OШ \pm 1,96 SEOШ = 1,46787 \pm 0,559$, или $0,909 < 1,468 < 2,027$

5. Доверительный интервал необходимо получить для значения параметра OШ, равного 4,34. Из определения натурального логарифма $\ln N = A$, получаем значение $N = e^A$. Или в нашем случае: $4,34 = e^{1,468}$. Используя свойство сохранения неравенства для степеней с одинаковыми основаниями, переходим к выражению:

$$e^{0,909} < e^{1,468} < e^{2,027}$$

Подставив численное значение экспоненты: $e = 2,718282$, получаем доверительные границы интервала:
 $2,48 < 4,34 < 7,59$

Наконец, третий способ расчета отношения шансов и 95% доверительного интервала – использование представленной ниже программы для персонального компьютера на языке C++:

```
#include "math.h"
int main()
{
int a,b,c,d;
printf(«Введите через запятую коэффициенты таблицы сопряженности
(a,b,c,d):»); //ввод результатов
scanf(«%i,%i,%i,%i»,&a,&b,&c,&d);
double or = (a*d)/(double)(b*c); //отношение шансов
double lor=log(or); //натуральный логарифм отношения шансов
double var=1./a + 1./b + 1./c + 1./d; //дисперсия натурального логарифма
отношения шансов:
```

```

double se=sqrt(var);//стандартная ошибка отношения шансов
double ll=exp(lor - (1.96*se));//нижняя граница отношения шансов
double ul=exp(lor + (1.96*se));//верхняя граница отношения шансов
printf(«\nОтношение шансов: %.2f < %.2f < %.2f\n»,ll,or,ul);//вывод результатов
return 0;
}

```

В той же публикации De Moor and all. (2000) приводятся сведения о том, что из 4617 исследованных зубов предварительное эндодонтическое лечение корневых каналов было зарегистрировано в 312 зубах. Используя приведенные данные, можно проиллюстрировать какие статистические методы используются при альтернативной форме учета реакций, когда элементы совокупности делятся на две части по одному, чаще всего качественному, признаку. При этом единственным способом описания качественных признаков является подсчет числа объектов, имеющих одно и то же значение, или долю (p) от общего числа объектов, которая приходится на то или иное значение (В.И. Сергиенко, И.Б. Бондарева, 2001). В конкретном примере речь идет о том, какую долю от общего количества исследованных зубов составляют эндодонтически леченные зубы. Для расчета используется формула:

$p = n_1 \cdot 100\% / n_1 + n_2$, где n_1, n_2 – численности альтернатив.

Расчетная доля леченных зубов составляет: $p = 312 \cdot 100\% / 4617 = 6,8\%$.

Доверительный интервал для доли леченных зубов определяются как $p \pm Z_a \cdot SE_p$, где SE_p – стандартная ошибка доли, которая задается соотношением: $SE_p = 2Ц_p(1-p)/n$, а Z_a – определяется по таблице (при 95% доверительном интервале, $Z_a = 1,96$).

Доверительный интервал генеральной доли находится в пределах $6,8 \pm 1,96 \cdot 2Ц_0,068(1-0,068) / 4617 = 6,8 \pm 0,72\%$.

Таким образом, приведенные выше несложные методы статистического анализа могут оказаться полезными исследователю для обработки собственных результатов научных исследований, а также позволят специалисту любого профиля оперативно проводить экспертную оценку достоверности результатов любых клинических публикаций.

Литература

1. Власов В.В. Введение в доказательную медицину. – М: Медиа Сфера, 2001. – 282с.
2. Медик В.А., Токмачев М.С., Фишман Б.Б. Статистика в медицине и биологии: Руководство. В 2-х томах / Под ред. Ю.М.Комарова. Т. 1. Теоретическая статистика. – Медицина: 2000. – 412 с.
3. Сергиенко В.И., Бондарева И.Б. Математическая статистика в клинических исследованиях. – М. Гэотар-мед, 2001. – 256 с.
4. De Moor R.J.G., Homme G.M.G., De Boever J.G., Delme K.I.M., Martens G.E.I. Periapical health related to the quality of root canal treatment in Belgian population. International Endodontic Journal. 2000, 33, p.113 – 120.

1. Показатель некачественного заполнения корневого канала
 2. Натуральный логарифм – логарифм по основанию e. Константа e приблизительно равняется 2,718282.