

## ОЦЕНКА ЗНАЧИМОСТИ ПРИЗНАКОВ В ЗАДАЧЕ ДИАГНОСТИКИ ВИДА ТРАВМИРУЮЩЕГО ВОЗДЕЙСТВИЯ В РЕЗУЛЬТАТЕ ДТП

Н. В. Огаркова

*Воронежский государственный университет*

*Поступила в редакцию 20.06.2013 г.*

**Аннотация.** Статья посвящена некоторым вопросам автоматизации производства судебно-медицинских экспертиз. В работе описана задача классификации способа травмирующего воздействия, приведен полный список используемых диагностических признаков. Также описан программный комплекс, ориентированный на решение задачи диагностики и дана краткая характеристика результатов, полученных с его использованием. Кроме того приведен один из способов оценки значимости признаков, показаны и проанализированы полученные результаты.

**Ключевые слова:** дорожно-транспортное происшествие, автоматизация производства судебно-медицинских экспертиз, диагностика вида травмирующего воздействия, программный комплекс, оценка значимости признаков.

**Annotation.** The paper discusses issues related to the processes of automated forensic medical examination. Firstly, the paper describes the task of traumatic injuries classification and presents a complete list of diagnostic features. Secondly, a software package for automated diagnosis and the results of its application are presented. Finally, one of the methods of feature significance evaluation is offered along with the results of its application.

**Keywords:** road accidents, automated forensic medical examinations, diagnosis of traumatic injury, software package, evaluation of feature significance.

### ВВЕДЕНИЕ

Наиболее интересным и перспективным направлением в судебной медицине в настоящее время является разработка компьютерной модели эксперта, способной близко имитировать поведение специалиста при решении сложных практических задач [1].

Одним из актуальных вопросов, стоящих перед судебно-медицинским экспертом при производстве судебно-медицинских экспертиз в случае ДТП и требующих повышения качества и оперативности их решения, является диагностика механизмов и условий образования повреждений костей скелета. Решение вопроса о механизмах образования переломов основано на изучении характера и признаков разрушения костной ткани.

До настоящего времени для оценки признаков разрушения использовались визуально-описательные методики, не учитывающие их количественные характеристики. Кроме того отсутствовала единая систематизация диагностических критериев и единая методика после-

довательного анализа и оценки переломов для определения вида внешнего воздействия, в частности, удара и давления [2].

В соответствии с требованиями современной судебной медицины для достижения высокой точности результатов необходим систематизированный подход к оценке признаков, используемых при диагностике механизма травмы, и, в первую очередь, учет их количественных характеристик. Внедрение компьютерных технологий обработки и анализа информации в судебно-медицинскую деятельность – принципиально новый шаг к повышению эффективности, качества и достоверности исследований.

### ЗАДАЧА КЛАССИФИКАЦИИ ВИДА ТРАВМИРУЮЩЕГО ВОЗДЕЙСТВИЯ

В судебной медицине травмирующее воздействие, определение которого является первоочередной задачей эксперта-криминалиста в случае дорожно-транспортного происшествия, делится на два вида: «Удар» и «Давление», следовательно, она может рассматриваться как задача классификации, так как экспертное знание состоит в отнесении объекта к одному из

двух классов решений. Объекты, подлежащие классификации, описываются, как правило, оценками по различным критериям, которые могут иметь как качественный, так и количественный характер.

Материалом для исследования служили бедренные и большеберцовые кости лиц обоего пола, умерших в возрасте от 20 до 60 лет, поврежденные при экспериментальном моделировании переломов на биоманекенах. Для исследования зоны разрушения были изготовлены препараты в виде продольно-профильных шлифов, при анализе которых были выявлены различия, отражающие закономерности разрушения при различных условиях травматизации, а также был определен набор диагностических признаков [2, 3]:

1. Угол траектории разрушения (угол разруш).
2. Длина зоны разрыва (L ЗР).
3. Коэффициент разрыва относительно диаметра кости (коэф.разр/д).
4. Коэффициент разрыва относительно толщины компактного вещества (коэф.разр/т).
5. Протяженность зоны пластической деформации на стороне растяжения (зона п.д.р.).
6. Количество зубцов в зоне разрыва (кол.зуб).
7. Количество «ступенеобразных» изменений траектории магистрального разрушения на границе зон разрыва и сдвига (кол. ступ).
8. Количество поперечных микротрещин на стороне растяжения (кол.поп.тр).
9. Средняя длина поперечных микротрещин на стороне растяжения (ср.дл.поп.тр).
10. Количество микротрещин древовидной формы на стороне растяжения (кол.др.тр).
11. Количество продольных изолированных трещин в толще компактного вещества на стороне растяжения (кол.пр.тр.р).
12. Количество трещин, отходящих от костно-мозговой полости на стороне растяжения (кол.тр.кмп.р).
13. Длина зоны долома (L ЗД).
14. Коэффициент долома относительно диаметра кости (коэф.дол/д).
15. Коэффициент долома относительно толщины компактного вещества (коэф.дол/т).
16. Протяженность зоны пластической деформации на стороне сжатия (зона п.д.сж).
17. Высота наибольшего гребня в зоне долома (max.выс.гр).

18. Количество в зоне долома гребней с остроугольной вершиной (кол.гр.о).

19. Количество в зоне долома гребней с закругленной или П-образной вершиной (кол.гр.з).

20. Количество трещин «расклинивающего» типа, отходящих от концевых отделов воронкообразных углублений в зоне долома (кол.тр.конц).

21. Количество трещин «расклинивающего» типа, отходящих от боковых отделов воронкообразных углублений в зоне долома (кол.тр.бок).

22. Длина наибольшей трещины «расклинивающего» типа (L тр.раскл).

23. Угол отхождения микротрещины «расклинивающего» типа (уг.тр.раскл).

24. Количество трещин, отходящих от костно-мозговой полости на стороне сжатия (кол.тр.кмп.сж).

25. Количество изолированных микротрещин X-, Y- и #-образной формы на стороне сжатия (кол.XYтр).

26. Количество изолированных продольных микротрещин на стороне сжатия (кол.пр.тр.сж).

Программный комплекс и результаты его применения

С учетом потребностей экспертной практики был разработан программный комплекс, ориентированный на получение обоснованного вывода о виде внешнего воздействия при производстве судебно-медицинской экспертизы [5, 6, 7].

Для построения математических моделей использовалась логистическая регрессия – разновидность множественной регрессии, общее назначение которой состоит в анализе связи между несколькими независимыми переменными (факторными признаками) и бинарной зависимой переменной.

Для регрессионной модели с биномиальной зависимой переменной и логистическим распределением вероятность того, что случайная величина  $Y$ , значения которой определяются формулой  $Y = \Theta \tilde{X}$ , примет единичное значение равна:

$$P(Y = 1 | \tilde{X}) = F(\Theta^T \tilde{X}) = \frac{e^{\Theta^T \tilde{X}}}{1 + e^{\Theta^T \tilde{X}}} = \frac{1}{1 + e^{-\Theta^T \tilde{X}}}, \quad (1)$$

где  $\tilde{X} = (1, x^{(1)}, x^{(2)}, \dots, x^{(p)})^T$  – вектор значений независимых переменных, соответствующий

множеству анализируемых признаков, а  $\Theta = (\theta_0, \theta_1, \theta_2, \dots, \theta_p)$  – вектор коэффициентов регрессии.

Для расчета коэффициентов логистической регрессии в программном комплексе были реализованы метод наименьших квадратов и метод максимального правдоподобия.

Для оценки качества построенных моделей использовались следующие показатели:

- количество правильно/неправильно классифицированных случаев;
- значение среднеквадратической ошибки;
- значение множественного коэффициента корреляции  $R$ , определяемого формулой (2).

$$R = \sqrt{1 - \frac{Q_{\text{ост}}}{Q_{\text{общ}}}} = \sqrt{\frac{Q_{\text{объясн}}}{Q_{\text{общ}}}}, \quad (2)$$

где  $Q_{\text{ост}}$  – остаточная сумма квадратов (сумма квадратов отклонений фактических значений от расчетных)

$$Q_{\text{ост}} = \sum_{i=1}^n (y_i - \hat{y}_i)^2,$$

$Q_{\text{объясн}}$  – объясненная сумма квадратов (сумма квадратов отклонений расчетных значений от среднего)

$$Q_{\text{объясн}} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2,$$

а  $Q_{\text{общ}}$  – общая сумма квадратов

$$Q_{\text{общ}} = Q_{\text{ост}} + Q_{\text{объясн}} = \sum_{i=1}^n (y_i - \bar{y})^2,$$

где  $\bar{y} = \sum_{i=1}^n y_i / n$  – среднее значение.

Для выбора наилучшего подмножества факторных признаков в программном комплексе были реализованы: пошаговый регрессионный анализ и метод всех возможных регрессий.

Программный комплекс был применен для анализа признаков разрушения, полученных при исследовании переломов костей на продольно-профильных распилах.

Использование метода пошагового включения признаков в модель логистической регрессии показало, что для получения оптимальной модели достаточно использовать семь признаков. Поскольку количество использованных в модели признаков оказалось невелико, то было сделано предположение, что существуют и другие наборы, на основе которых могут быть построены оптимальные модели [3].

Метод всех возможных регрессий применялся для случая, когда максимальное количество признаков, включаемых в уравнение, изменялось от одного до семи. При этом в базу вносились уравнения, коэффициент множественной корреляции которых был более 0,8. Из полученных уравнений регрессии 42 имели множественный коэффициент корреляции, равный 0,9999(9) и среднеквадратическую ошибку, равную 0 [2, 4], а всего таких уравнений было построено 21529 шт.

### ОЦЕНКА СТЕПЕНИ ЗНАЧИМОСТИ ПРИЗНАКОВ

Помимо разработки моделей и алгоритмов, позволяющих автоматизировать задачу диагностики вида внешнего воздействия в результате ДТП, одним из важнейших вопросов для экспертов в области производства судебно-медицинских экспертиз остается вопрос о степени значимости диагностических признаков.

Для того, чтобы оценить значимость того или иного факторного признака, включенного в модель, можно использовать разность между оценками качества двух моделей: модели, построенной с использованием данного признака, и без него. В качестве оценки качества модели можно использовать коэффициент множественной корреляции  $R$ . Способ оценки значимости признака приведен ниже.

Пусть некоторая  $i$ -ая модель построена с использованием  $k$  заданных признаков  $\tilde{X} = (1, x^{(2)}, x^{(2)}, \dots, x^{(k)})^T$ , при этом для этой модели был посчитан коэффициент множественной корреляции  $R_i$ . Тогда для оценки значимости какого-либо признака из этой модели надо, во-первых, рассчитать первичные оценки значимости всех признаков из этой модели  $t_i^{(j)}$  ( $j = 1..k$ ), то есть для каждого из признаков построить модель  $\tilde{X} = (1, x^{(1)}, x^{(2)}, \dots, x^{(j-1)}, x^{(j+1)}, \dots, x^{(k)})^T$  и рассчитать множественный коэффициент корреляции  $R_i^{(j)}$ . Тогда первичная оценка значимости будет определяться формулой (3)

$$t_i^{(j)} = R_i - R_i^{(j)}. \quad (3)$$

Поскольку факторные признаки, используемые при построении модели, не являются независимыми (парные, а тем более множественные коэффициенты корреляции не равны нулю), то сумма первичных оценок будет отлична от  $R_i$ . Для того чтобы получить окончательную оценку  $\hat{t}_i^j$  значимости признака  $x_i^{(j)}$  в

$i$ -ой модели величину  $t_i^{(j)}$  надо нормировать (4).

$$\hat{t}_i^{(j)} = t_{ii}^{(j)} / T, \quad (4)$$

где  $T = \sum_{j=1}^k t_i^{(j)}$ .

Если в результате использования метода всех возможных регрессий было построено  $M$  моделей, при этом для каждой модели была определена характеристика  $R_i$ , то для оценки степени значимости признака  $x^{(j)}$  необходимо определить две матрицы:

$$(a_{ij}) = \begin{cases} 0, & \text{если признак с номером } j \\ & \text{не вошел в модель } i, \\ 1, & \text{если признак с номером } j \\ & \text{вошел в модель } i; \end{cases}$$

$$(b_{ij}) = \begin{cases} 0, & \text{если признак с номером } j \\ & \text{не вошел в модель } i, \\ \hat{t}_i^{(j)}, & \text{если признак с номером } j \\ & \text{вошел в модель } i. \end{cases}$$

Тогда в качестве показателей значимости признака могут быть использованы следующие характеристики:

- Абсолютная и относительная частота использования признаков в моделях

$$Q_j = \sum_{i=1}^M a_{ij},$$

$$\tilde{Q}_j = Q_j / M.$$

- Взвешенная оценка частоты использования признака

$$Q_j^* = \sum_{i=1}^M R_i a_{ij} / R,$$

где  $R = \sum_{i=1}^M R_i$ .

- Абсолютная и относительная значимость признаков в моделях

$$V_j = \sum_{i=1}^M b_{ij},$$

$$\tilde{V}_j = V_j / M.$$

- Взвешенная значимость признака в моделях

$$V_j^* = \sum_{i=1}^M R_i b_{ij} / R.$$

## РЕЗУЛЬТАТЫ ОЦЕНКИ ЗНАЧИМОСТИ ПРИЗНАКОВ

Для оценки значимости признаков было решено проанализировать 42-е наилучших модели, которые имели множественный коэффициент корреляции, равный 0,9999(9). В таких условиях рассчитывать относительные и взвешенные характеристики не было смысла. При оценке степени значимости факторных признаков были получены следующие результаты (см. табл. 1).

Таблица 1  
Оценка значимости признаков

Название	Частота ( $Q_j$ )	Значимость ( $V_j$ )
Л ЗР	7	0,251
Коэф.разр/т	28	0,241
Коэф.разр/д	7	0,216
кол.ХУтр	42	0,201
кол.зуб	37	0,163
уг.тр.раскл	5	0,133
кол.тр.кмп.р	42	0,132
тах.выс.гр	18	0,11
Л ЗД	13	0,104
Коэф.дол/д	9	0,099
Зона п.д.р	22	0,094
Л тр.раскл	11	0,088
кол.тр.конц	21	0,086
кол.тр.бок	1	0,085
Коэф.дол/т	4	0,077
кол.пр.тр.р	9	0,074
кол.пр.тр.сж	5	0,074
Угол разруш	2	0,064
кол.гр.з	2	0,06
Зона п.д.сж	2	0,059
кол.ступ	2	0,058
ср.дл.поп.тр	1	0,052
кол.гр.о	1	0,05

Анализ полученных результатов показал, что наиболее значимыми для диагностики способа травмирующего воздействия являются признаки, характеризующие зону разрыва («Длина зоны разрыва», «Коэффициент разрыва относительно диаметра кости» и «Коэффициент разрыва относительно толщины компактного вещества»). При этом эти признаки являются зависимыми и взаимно исключают друг друга (их суммарная частота использования равна количеству рассмотренных моделей).

Таким же значимым признаком является «Количество изолированных микротрещин X-, Y- и #-образной формы на стороне сжатия» (у него высокая степень значимости, а кроме того он фигурирует во всех рассмотренных моделях).

Следующими по степени значимости являются признаки «Количество зубцов в зоне разрыва» и «Угол отхождения микротрещины «расклинивающего» типа», причем они также взаимно исключают друг друга. Примерно таким же по степени значимости является признак «Количество трещин, отходящих от костномозговой полости на стороне растяжения», но он присутствует во всех моделях. Остальные признаки играют скорее уточняющую (корректирующую) роль.

#### СПИСОК ЛИТЕРАТУРЫ

1. Диагностикум механизмов и морфологии переломов при тупой травме скелета. Т. 1. Механизмы и морфология переломов длинных трубчатых костей / В.И. Бахметьев, В.Н. Крюков, В.П. Новоселов и др. – 2-е изд. – Новосибирск: Наука. Сибирская издательская фирма РАН, 2002. – 166 с.

2. Кирилов В.А. Судебно-медицинская диагностика вида внешнего воздействия на основе математического моделирования взаимосвязей параметров разрушения диафиза длинных трубчатых костей нижних конечностей / В.А. Кирилов, В.И. Бахметьев, Н.В. Огаркова // Системный анализ и управление в биомедицинских системах. – М, 2009. – Т. 8, № 2. – С. 339–343.

**Огаркова Наталья Владимировна** – преподаватель кафедры программного обеспечения и администрирования информационных систем Воронежского государственного университета. E-mail: Nataly.Ogarkova@gmail.com).

3. Применение логистической регрессии для анализа морфологических признаков разрушения на продольно-профильных распилах костей / Н.В. Огаркова, В.А. Кирилов // Информатика : проблемы, методология, технологии : материалы 8 междунар. науч.-метод. конф., 7–8 февр. 2008 г., г. Воронеж. – Воронеж, 2008. – Т. 2. – С. 127–131.

4. Огаркова Н.В. Применение метода всех возможных регрессий для анализа признаков разрушения кости в судебно-медицинском отношении / Н.В. Огаркова, В.А. Кирилов // Информатика : проблемы, методология, технологии : материалы 9-й междунар. науч.-метод. конф., 12–13 февр. 2009 г., г. Воронеж. – Воронеж, 2009. – Т. 2. – С. 151–154.

5. Огаркова Н.В. Программный комплекс для распознавания способа травмирующего воздействия в результате ДТП. / Н.В. Огаркова, Е.Ю. Ряжских // Вестн. Воронеж. гос. ун-та. Сер. Системный анализ и информационные технологии. – Воронеж, 2009. – № 2. – С. 101–106.

6. Огаркова Н.В. Автоматизация производства судебно-медицинских экспертиз в случае возникновения дорожно-транспортных происшествий с летальным исходом. / Н.В. Огаркова, В.А. Кирилов // Вестн. Воронеж. гос. ун-та. Сер. Системный анализ и информационные технологии. – Воронеж, 2010. – № 1. – С. 68–73.

7. Голуб В.А. Программа построения и анализа регрессионной модели для распознавания способа травмирующего воздействия в результате ДТП / В.А. Голуб, Н.В. Огаркова, А.С. Попов // Вестн. Воронеж. гос. ун-та. Сер. Системный анализ и информационные технологии. – Воронеж, 2007. – № 1. – С. 16–20.

**Ogarkova Natalia Vladimirovna** – lecturer, Department of Programming and System Administration, Voronezh State University. E-mail: Nataly.Ogarkova@gmail.com).