

УДК 519.25

ИНДИВИДУАЛЬНОЕ ПРОГНОЗИРОВАНИЕ ДВИГАТЕЛЬНОГО РАЗВИТИЯ У БОЛЬНЫХ ДЕТСКИМ ЦЕРЕБРАЛЬНЫМ ПАРАЛИЧОМ НА ОСНОВЕ ПОДХОДОВ СТАТИСТИЧЕСКОГО АНАЛИЗА

А.А. Ташкинов¹, А.В. Вильдеман¹, В.А. Бронников²

¹ Кафедра механики композиционных материалов и конструкций Пермского государственного технического университета, Россия, 614990, Пермь, Комсомольский проспект, 29, e-mail: wildemann@mail.ru

² Краевое государственное автономное учреждение «Центр комплексной реабилитации инвалидов», Россия, 614090, Пермь, ул. Лодыгина, 39

Аннотация. На основе подходов статистического анализа рассматриваются задачи прогнозирования двигательного развития у детей с церебральным параличом. В качестве прогнозируемого показателя используется индекс моторики. Предлагается метод индивидуального прогнозирования, основанный на переходе от уравнений, описывающих состояние группы пациентов, к уравнениям, характеризующим развитие конкретного больного. Получены математические модели, демонстрирующие приемлемую точность прогнозов двигательного развития на основе данных о средней зависимости и индивидуальных замерах индекса моторики на начальной стадии развития заболевания.

Ключевые слова: математические модели индивидуального прогнозирования, статистический анализ, детский церебральный паралич.

ВВЕДЕНИЕ

Актуальной проблемой математического моделирования является применение методов статистического анализа при исследовании медико-социальных систем.

Одна из важных задач междисциплинарного взаимодействия современной медицины, механико-математических методов и информационных технологий заключается в достоверном прогнозировании динамики состояния человека на основе данных начального этапа формирования болезни.

При разработке прогнозных моделей, наряду с определением общих закономерностей развития заболевания у выборочной группы пациентов, важным представляется учет индивидуальных характеристик конкретного больного для получения более точных индивидуальных прогнозов.

Методы многомерного статистического анализа широко применяются в различных областях медицины [2, 4, 6, 8, 12]. Вместе с тем одним из малоизученных направлений исследования является прогнозирование состояния больных с врожденными нарушениями двигательных функций, на текущее развитие которых во многом влияют родовые и дородовые факторы риска. К таким больным, в частности, относятся дети с церебральными параличами.

Целью настоящей работы является создание математических моделей прогнозирования двигательного развития у больных детским церебральным параличом (ДЦП) на основе подходов статистического анализа.

АНАЛИЗ ИСХОДНЫХ ДАННЫХ

В качестве комплексного показателя двигательного развития рассматривается индекс моторики – величина, характеризующая интегральный уровень развития моторных навыков [13].

Объектом исследования является ребенок с детским церебральным параличом, в результате обследования которого определены значения множества факторов, характеризующих его состояние. В работе исследовалась выборка, состоящая из 157 наблюдений, соответствующих 73 больным. Данные были получены в результате совместной работы с Пермским Центром комплексной реабилитации инвалидов. Выборка формировалась в результате случайных обращений пациентов в медико-социальное учреждение в течение 10 лет. При этом, несмотря на достаточно большое количество людей с диагнозом ДЦП, обследованных на протяжении данного промежутка времени, удалось выделить только 5 больных, регулярно наблюдаемых в течение всего рассматриваемого периода. В связи с этим исследуемые в настоящей работе модели прогнозирования будут иллюстрироваться на примере группы из пяти человек.

Во время обследования у каждого больного экспертным путем оценивались 12 различных групп двигательных навыков по пятибалльной шкале. В результате суммирования всех навыков определялось значение индекса моторики, лежащее в интервале от 0 до 60.

Изменение индекса моторики с возрастом у группы пациентов рассматривалось как случайный процесс $Y(t)$ с непрерывным временем и непрерывными состояниями [3, 11], где t – возраст в месяцах. Динамика изменения индекса моторики у j больного в группе в этом случае является реализацией $y^{(j)}(t)$ случайного процесса $Y(t)$.

В силу случайности обращений пациента в лечебные учреждения $y^{(j)}(t)$ представляет собой неэквидистантный временной ряд, т.е. характеризуется случайными интервалами времени между отсчетами. Для регуляризации временного ряда в работе использовался метод скользящего среднего, основанный на представлении $y^{(j)}(t)$ в некоторой точке посредством взвешенного среднего значений, наблюдаемых в окрестности этой точки [1].

С целью исследования структуры случайного процесса была построена корреляционная функция

$$r_y(t, t') = \frac{M[\dot{Y}(t)\dot{Y}(t')]}{\sqrt{D_y(t)D_y(t')}} \quad (1)$$

где $\dot{Y}(t)$ и $\dot{Y}(t')$ – сечения центрированного случайного процесса $\dot{Y}(t)$ в моменты времени t и t' , $D_y(t)$ и $D_y(t')$ – дисперсии $Y(t)$ в моменты t и t' .

На рис. 1 приведена диаграмма рассеивания для пяти реализаций случайного процесса $Y(t)$, на рис. 2 изображен график соответствующей корреляционной функции при фиксированном $t = 15$.

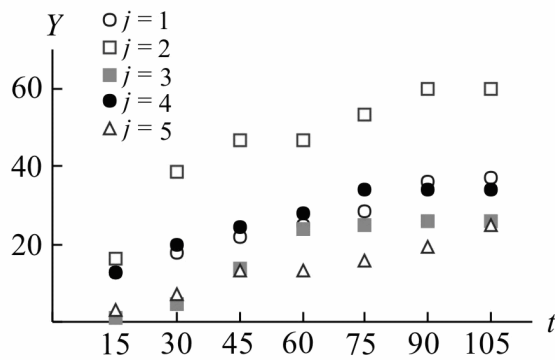


Рис. 1. Реализации индекса моторики на примере пяти больных

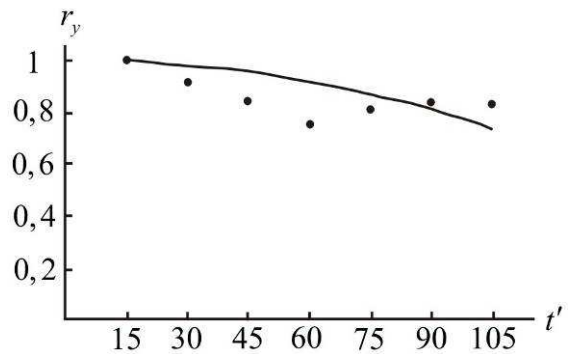


Рис. 2. Корреляционная функция индекса моторики

Анализ корреляционной функции данного процесса показал, что для него характерна достаточно сильная вероятностная зависимость между сечениями, что говорит о простой структуре процесса с плавно меняющимися реализациями.

ОПРЕДЕЛЕНИЕ СРЕДНЕГРУППОВОЙ ЗАВИСИМОСТИ ИНДЕКСА МОТОРИКИ ОТ ВОЗРАСТА

Для получения достоверных прогнозов относительно будущего состояния больного необходимо учитывать при построении прогнозных моделей индивидуальные характеристики конкретного пациента.

Вопросы индивидуального прогнозирования подробно рассматриваются в работах Ю.П. Самарина и его учеников [5, 9] применительно к прогнозированию состояния механических объектов.

Суть методов индивидуального прогнозирования заключается в переходе от уравнений, описывающих осредненное состояние группы пациентов, к уравнениям, характеризующим конкретного больного.

В соответствии с подходом индивидуального прогнозирования на первом этапе вычисляется средний индекс моторики для группы больных.

Для определения зависимости среднего индекса моторики от времени было проведено осреднение процесса $Y(t)$ по реализациям. Задача аппроксимации полученных таким образом средних значений состояла в выделении тренда временного ряда $f(t)$.

При определении тренда использовался параметрический метод, основанный на предположении о наличии математической модели зависимости, заданной с точностью до параметров:

$$f(t, \theta) = a_0 + \sum_{q=1}^Q a_q \phi_q(t, \lambda_q), \quad (2)$$

где $\theta = (a_0 \ a_1 \ \dots \ a_q \ \dots \ a_Q \ \lambda_1 \ \dots \ \lambda_q \ \dots \ \lambda_Q)^T$ – вектор неизвестных параметров, $\phi_q(t, \lambda_q)$ – базисные функции, заданные с точностью до параметра λ_q .

В настоящей работе для определения зависимости $f(t)$ применялся метод последовательного выделения экспоненциальных слагаемых, предложенный впервые Ю.П. Самариним [10].

В соответствии с данным методом в качестве базисных используются функции вида

$$\phi_q(t, \lambda_q) = 1 - e^{-\lambda_q t}. \quad (3)$$

Применение базисных функций (3) в настоящей задаче обосновано тем, что всякую непрерывную на конечном отрезке функцию можно с любой степенью точности приблизить степенным полиномом (теорема Вейерштрасса), а разложение (3) сводится к степенному полиному подстановкой $t = -\ln x$.

К числу особенностей метода относится то, что необходимое количество оцениваемых параметров λ_q не задается заранее, а определяется в процессе расчета. Кроме того, предполагается, что данные параметры значительно отличаются друг от друга, это позволяет с достаточной степенью точности описывать отдельную дугу аппроксимирующей кривой с использованием только одного из параметров.

В результате применения метода нелинейная зависимость среднего индекса моторики от возраста была аппроксимирована с помощью двух экспоненциальных слагаемых (рис. 3, 4):

$$f(t) = -16,49 + 66,87(1 - e^{-0,004t}) + 28,39(1 - e^{-0,09t}). \quad (4)$$

Коэффициент детерминированности $R^2 = 1 - \sum_i (\tilde{f}_i - f_i)^2 / \sum_i (\tilde{f}_i - \langle \tilde{f}_i \rangle)^2$ для аппроксимации (4) составил 0,99.

При этом было получено, что параметр $\lambda_2 = 0,09$ в выражении (4) характеризует динамику индекса моторики на интервале $t \in [0; 45)$, а $\lambda_1 = 0,004$ – на интервале $t \in [45; 105]$.

Отличительной чертой метода последовательного выделения экспоненциальных слагаемых является то, что он позволяет получить аппроксимирующее выражение, описывающее различные участки кривой индекса моторики с помощью различных параметров λ_q , не прибегая при этом к использованию кусочно-гладких функций.

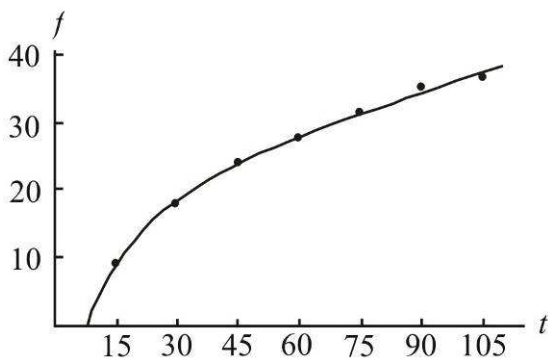


Рис. 3. Тренд среднего индекса моторики, полученный с использованием двух экспоненциальных слагаемых

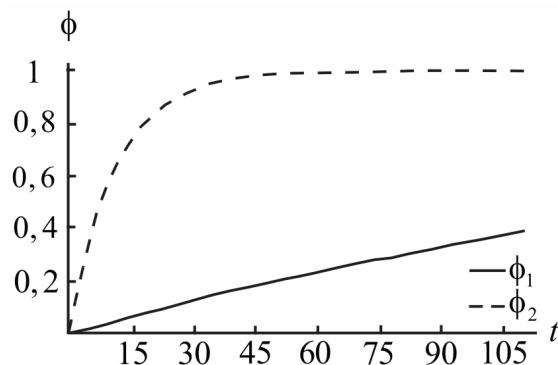


Рис. 4. Базисные функции, используемые при аппроксимации

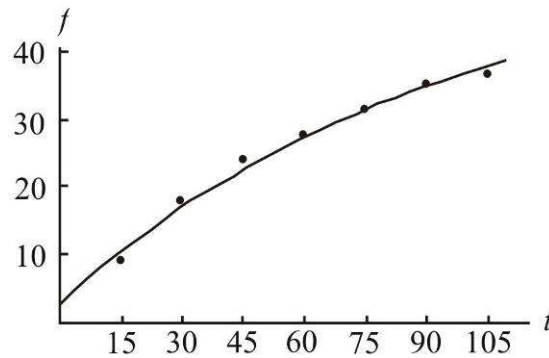


Рис. 5. Тренд среднего индекса моторики, полученный с использованием одного экспоненциального слагаемого

Для сравнения зависимость среднего индекса моторики от возраста может быть аппроксимирована с использованием одного экспоненциального слагаемого (рис. 5). Однако коэффициент детерминированности при этом будет меньше: $R^2 = 0,98$.

Кроме того, сам характер кривой индекса моторики, полученной с помощью аппроксимации двумя экспоненциальными слагаемыми (см. рис. 3), более точно соответствует развитию моторных навыков на начальном этапе у больных ДЦП, для которых, как правило, характерна задержка начала двигательного развития.

ИНДИВИДУАЛЬНОЕ ПРОГНОЗИРОВАНИЕ ИНДЕКСА МОТОРИКИ

При построении моделей индивидуального прогнозирования вместо среднего индекса моторики рассматривается случайная функция:

$$Y(t) = A_0 + \sum_{q=1}^Q A_q(t)(1 - e^{-\lambda_q t}), \quad (5)$$

где $A_q(t) = A_q + A'_q(t)$;

A_q ($q = \overline{0, Q}$) – случайные величины, для которых $\langle A_q \rangle = a_q$;

$A'_q(t)$ – центрированные случайные функции, для которых $\langle A'_q(t) \rangle = 0$;

λ_q ($q = \overline{1, Q}$) – детерминированные величины.

Таким образом, функция (5) представима в виде

$$Y(t) = F(t) + \xi(t), \quad (6)$$

где

$$F(t) = A_0 + \sum_{q=1}^Q A_q(1 - e^{-\lambda_q t}), \quad (7)$$

$$\xi(t) = \sum_{q=1}^Q A'_q(t)(1 - e^{-\lambda_q t}). \quad (8)$$

$F(t)$ будем называть главной составляющей индекса моторики, $\xi(t)$ будем рассматривать как наложенный «шум».

Считая, что значения шума существенно меньше главной части, будем искать индивидуальную зависимость индекса моторики от времени в виде аппроксимации (7).

Для j пациента выражение (7) примет вид

$$F^{(j)}(t) = A_0^{(j)} + \sum_{q=1}^Q A_q^{(j)} (1 - e^{-\lambda_q t}), \quad (9)$$

где $A_q^{(j)}$ ($q = \overline{0, Q}; j = \overline{1, N}$) – реализации случайных величин A_q .

Выражение (9) содержит $Q+1$ неизвестную $A_q^{(j)}$. λ_q считаются известными, определенными при выделении группового тренда $f(t)$.

Задача индивидуального прогнозирования заключается в том, чтобы по ряду измерений состояния больного на начальном интервале протекания заболевания τ_0 спрогнозировать развитие его состояния в будущем.

Для подбора неизвестных $A_q^{(j)}$ воспользуемся методом наименьших квадратов на интервале τ_0 . Исходя из анализа структуры $Y(t)$, $F^{(j)}(t)$ будем искать в виде монотонно возрастающей функции, что соответствует выполнению условий $A_q^{(j)} \geq 0$ ($q = \overline{1, Q}$).

В результате получаем следующую постановку задачи оптимизации с ограничениями типа неравенств:

$$\begin{cases} \sum_{i=1}^{n_0} (F^{(j)}(t_i) - y_i^{(j)})^2 \rightarrow \min, \\ A_q^{(j)} \geq 0 \quad (q = \overline{1, Q}), \end{cases} \quad (10)$$

где n_0 – количество измерений на интервале τ_0 ;

$j = \overline{1, N}$ – условный номер больного, в настоящей работе рассматривалось $N = 5$ больных;

$y_i^{(j)}$ – измерения индекса моторики у j пациента в i момент времени; $t_i \in \tau_0$.

Решение системы (10) может быть найдено численно, например одним из градиентных методов, либо аналитически по теореме Куна–Таккера [7].

В качестве примера продемонстрируем возможности метода индивидуального прогнозирования для больного ($j = 2$), динамика двигательного развития которого наиболее сильно отличается от среднegrupповой (рис. 10).

Предположим, что в качестве исходных данных имеются три замера индекса моторики в течение первых четырех лет жизни пациента, а также тренд среднего индекса моторики, определенный ранее для группы больных (рис. 6). Задача состоит в прогнозировании индивидуального двигательного развития на ближайшие пять лет.

Как видно из рис. 6, индивидуальные замеры на начальном этапе развития заболевания очень плохо согласуются со средней зависимостью индекса моторики для группы больных. Поэтому прогнозирование на основе только группового тренда не приведет к получению результатов приемлемой точности.

С другой стороны, прогнозирование на основе только индивидуальных данных также не будет точным в связи с малым количеством замеров индекса моторики, а также в связи с тем, что период основания прогноза меньше периода, на который строится прогноз.

Для решения поставленной задачи целесообразно применить описанный ранее метод индивидуального прогнозирования, позволяющий скорректировать значения параметров групповой модели $f(t)$ с использованием индивидуальных данных конкретного больного.

Решая систему (10) аналитически, используя теорему Куна–Таккера, находим реализации A_q , доставляющие минимум в задаче оптимизации: $A_0^{(2)} = -64,17$, $A_1^{(2)} = 49,48$, $A_2^{(2)} = 104,34$.

На рис. 7 построена полученная таким образом индивидуальная зависимость. Коэффициент детерминированности для всей кривой составил при этом $R^2 = 0,96$.

Как видно из рис. 8, модель индивидуального прогнозирования существенно точнее описывает контрольные данные (серые точки) по сравнению с групповым трендом.

Для оценки точности индивидуального прогноза был построен 95%-й доверительный интервал $\Delta = t_\beta S_p$ (рис. 9), где S_p – средняя квадратическая ошибка прогноза, t_β – значение t -статистики Стьюдента для доверительной вероятности β .

Увеличение интервала основания прогноза τ_0 приводит к улучшению коэффициента детерминированности R^2 , что говорит об адекватности построенной модели. Так, для $\tau_0 = 30$ ($n_0 = 2$) – $R^2 = 0,76$; для $\tau_0 = 45$ ($n_0 = 3$) – $R^2 = 0,96$; $\tau_0 = 105$ ($n_0 = 7$) – $R^2 = 0,98$.

Применяя изложенный метод к прогнозированию индекса моторики для оставшихся пациентов из группы, получаем индивидуальные кривые, изображенные на рис. 10. В таблице приведены соответствующие коэффициенты детерминированности.

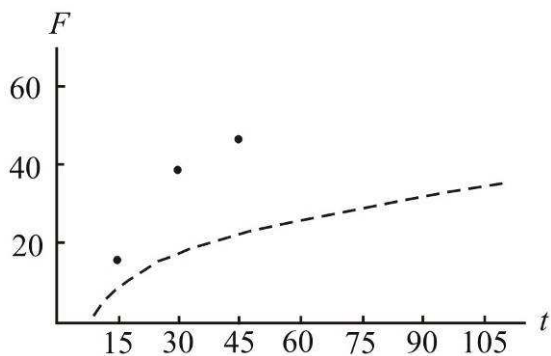


Рис. 6. Индивидуальные замеры (точки) и групповой тренд (пунктирная линия) индекса моторики

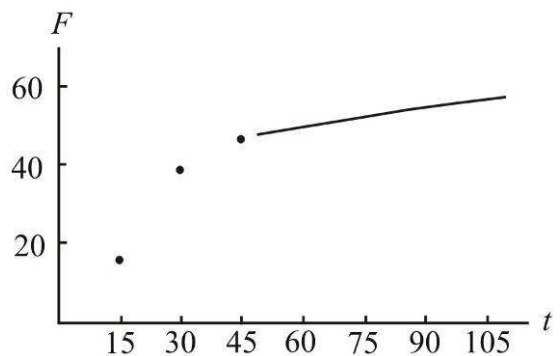


Рис. 7. Индивидуальные замеры (точки) и индивидуальный прогноз (сплошная линия) индекса моторики

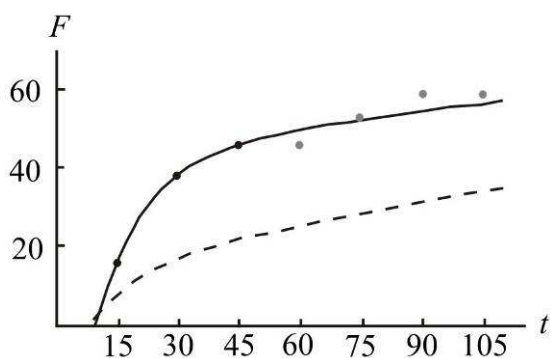


Рис. 8. Индивидуальные замеры (точки), индивидуальный прогноз (сплошная линия) и групповой тренд (пунктирная линия) индекса моторики

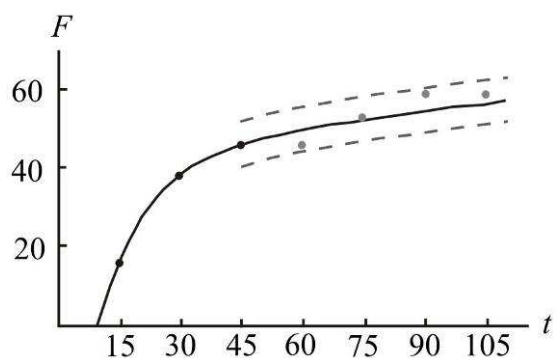


Рис. 9. Индивидуальные замеры (точки), индивидуальный прогноз (сплошная линия) индекса моторики и соответствующий ему 95%-й доверительный интервал (пунктирная линия)

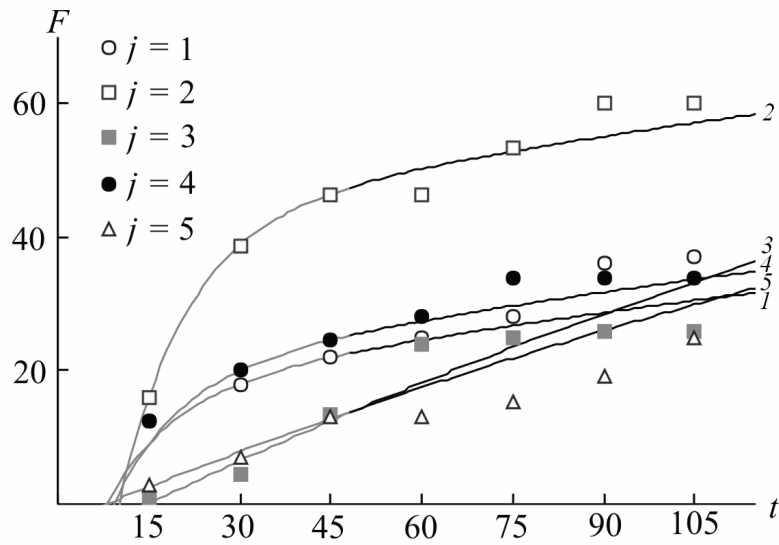


Рис. 10. Индивидуальные замеры и индивидуальные прогнозы индекса моторики

Коэффициенты детерминированности для кривых на рис. 10

Номер реализации j	1	2	3	4	5	Итого
R^2	0,84	0,96	0,85	0,95	0,60	0,84

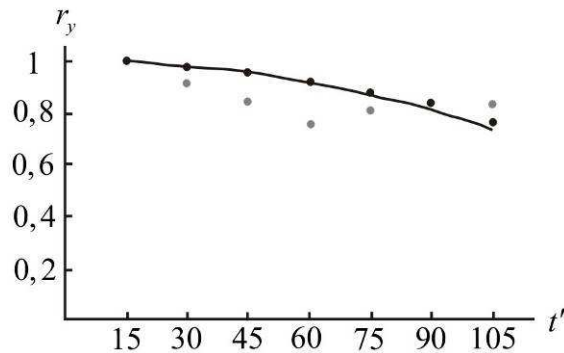


Рис. 11. Значения корреляционных функций, полученные для прогнозных (черные точки) и реальных (серые точки) данных, и кривая аппроксимации реальной корреляционной функции

С целью проверки достоверности построенных прогнозных моделей, в отсутствие независимой контрольной выборки больных, были получены значения нормированной корреляционной функции для модельных данных прогнозирования индекса моторики (рис. 11).

При этом значения модельной корреляционной функции практически точно легли на кривую, полученную в результате аппроксимации данных реальной корреляционной функции зависимостью вида $1 - \alpha t^2$, что свидетельствует об адекватном моделировании структуры случайного процесса.

ЗАКЛЮЧЕНИЕ

Предложенный метод индивидуального прогнозирования представляет собой эффективный инструмент обработки и анализа статистических данных, что было подтверждено применением данного метода при прогнозировании двигательного развития у больных детским церебральным параличом.

Полученные математические модели позволяют с достаточной степенью точности описывать динамику развития моторных навыков в условиях, когда интервал основания прогноза меньше периода, на который строится прогноз.

С целью дальнейшего улучшения достоверности индивидуальных прогнозов планируются исследования влияния, оказываемого на динамику индекса моторики со стороны родовых и дородовых факторов риска, определяющих начальные условия формирования двигательных навыков у больных детским церебральным параличом.

Применение предложенных методик прогнозирования в медицинской практике лечения детского церебрального паралича может служить дополнительным источником информации, характеризующей динамику состояния больного и используемой в процессе принятия решений специалистом в области реабилитации инвалидов.

БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке РФФИ (грант № 10-04-96096-р_урал_a).

СПИСОК ЛИТЕРАТУРЫ

1. Андерсон Т. Статистический анализ временных рядов. – М: Мир, 1976. – 760 с.
2. Безруков Н.С., Еремин Е.Л., Перельман Ю.М. Автоматизированная система диагностики заболеваний легких // Проблемы управления. – 2007. – № 5. – С. 75–80.
3. Вентцель Е.С., Овчаров Л.А. Теория случайных процессов и ее инженерные приложения: учеб. пособие для втузов. – М: Высш. шк., 2000. – 383 с.
4. Драгун И.А. Автоматизированная система количественной оценки операционного риска: автореф. дис. ... канд. техн. наук. – Барнаул, 2006. – 22 с.
5. Еремин Ю.А., Радченко В.П., Самарин Ю.П. Расчет индивидуальных деформационных свойств элементов конструкций в условиях ползучести // Машиноведение. – 1984. – № 1. – С. 67–72.
6. Колосов В.П., Перельман Ю.М., Ульянычев Н.В. Пути построения прогнозных моделей в пульмонологии // Информатика и системы управления. – 2005. – № 2. – С. 64–71.
7. Кузнецов А.В., Сакович В.А., Холод Н.И. Высшая математика: Математическое программирование. – Минск: Вышэйшая школа, 1994. – 286 с.
8. Реброва О.Ю. Применение методов интеллектуального анализа данных для решения задачи медицинской диагностики // Новости искусственного интеллекта. – 2004. – № 3. – С. 76–80.
9. Самарин Ю.П. О применении стохастических уравнений в теории ползучести материалов // Механика твердого тела. – 1974. – № 1. – С. 88–94.
10. Самарин Ю.П. Построение экспоненциальных аппроксимаций для кривых ползучести методом последовательного выделения экспоненциальных слагаемых // Проблемы прочности. – 1974. – № 9. – С. 24–27.
11. Свешников А.А. Прикладные методы теории случайных функций. – М: Наука, 1968. – 464 с.
12. Суфиянов В.Г. Разработка адаптивных статистических моделей классификации и прогнозирования: автореф. дис. ... канд. физ.-мат. наук. – Пермь, 2004. – 16 с.
13. Ташкинов А.А., Вильдеман А.В., Бронников В.А. Применение метода деревьев классификации к прогнозированию уровня развития моторики у больных с нарушениями двигательных функций // Российский журнал биомеханики. – 2008. – Т. 12, № 4. – С. 84–95.

INDIVIDUAL PREDICTION OF MOTOR ACTIVITY DEVELOPMENT OF PATIENTS WITH A CHILDREN'S CEREBRAL PARALYSIS ON THE BASIS OF STATISTICAL ANALYSIS APPROACHES

A.A. Tashkinov, A.V. Wildemann, V.A. Bronnikov (Perm, Russia)

On the basis of the statistical analysis approaches, the problems of motor activity development prediction of children with a cerebral paralysis are considered. The motility index is used as a predicted parameter. The method of individual prediction based on transition from the equations describing a patients group condition to the equations describing the concrete person condition is offered. The mathematical models showing acceptable prediction accuracy on the basis of the average dependence data and individual measurements of motility index at an initial stage of disease development are obtained.

Key words: mathematical models of individual prediction, statistical analysis, children's cerebral paralysis.

Получено 12 марта 2010