

2. Мелкоян Д. С. Биолог. ж. Армении, 35, 5, 480, 1982
3. Мелкоян Д. С., Арсисян Т. Г. Изв. АН АрмССР, сер. тех. наук, 5, 23, 1985.
4. Адамян С. Г., Барсегян Л. Г., Мелкоян Д. С., Роолайд Х. А. Журн. экспер. и клин. мед., 20, 2, 628, 1980.
5. Мелкоян Д. С., Мелкоян А. А., Мкртчян О. А., Саркисян Д. С., Хондкарян Н. С. В кн.: Нейронные механизмы интегративной деятельности мозжечка, Ереван, 242—247, 1979.

Поступило 30.I 1986 г.

Биолог. ж. Армении, т. 39, № 9, с. 752—757, 1986

УДК 577.151.64:577.152.111

ЭКСПРЕСС—МЕТОД СМЕШАННОГО ПРОДОЛЬНО-ПОПЕРЕЧНОГО ИССЛЕДОВАНИЯ РОСТОВЫХ ПРОЦЕССОВ У ДЕТЕЙ

Г. Н. ХАЧАТРЯН

РИИЦ МЗ Армянской ССР, Ереван

Аннотация — Проведены продольно-поперечные исследования, в определенной степени решающие задачу классификации кривых роста путем разбиения на отдельные деревья. При помощи дискриминантного анализа выявлен информативный набор признаков. Получена возможность прогнозирования типа кривой роста конкретного ребенка и дефинитивной картины его морфофункционального статуса.

Անոտացիա — Անց են կացված երկայնական-լայնական ուղղմամբ ուսումնասիրություններ, որոնք որոշ առումով լուծում են կորերի աճման դասակարգման խնդիրը առանձին ծառերի բաժանման միջոցով: Դիսկրիմինանտ անալիզի միջոցով բացահայտված է տեսականիչների ինֆորմատիվ հավաքածուն: Ստացված է կոնկրետ երկայնի աճման կորի տեսակի և երա մորֆոֆունկցիոնալ ստատուսի դեֆինիտիվ պատկերի կանխատեսման նախադրություն:

Abstract — The classification of growth curves by portioning them on different clusters was obtained using cross-longitudinal investigations. By discriminant analysis the informative set of signs was revealed. The possibility of prediction of the form of growth curve of a concrete child and the definitive picture of its morphofunctional status was obtained.

Ключевые слова: модель смешанная, продольно-поперечная, ближайший аналог, дублер, вырождение модели в дерево, дискриминантный анализ.

Изучение процессов роста и развития занимает одно из центральных мест в антропологии [5, 6]. Естественно, наиболее приемлемой с методологической точки зрения для подобных исследований является продольная схема [5]. Однако в целом эта схема практически трудно реализуема. Например, при исследовании перинубертатного периода онтогенеза человека (7—17 лет) данный подход требует 11-летних исследований. Поперечные же наблюдения, давая популяционную картину, непосредственно не приемлемы для построения кривых роста.

Значительный прогресс в исследовании роста и развития, наблю-

даемый в последние годы, большинство исследователей связывают с появлением смешанных продольно-поперечных моделей [1, 8]. Стыковка когорт в этих моделях представляет одну из основных проблем. В некоторых работах [8] эту проблему пытаются решить с помощью долгосрочного перекрывания исследуемых когорт (3—4-летних). Очевидно, что этот подход сближает смешанную схему с продольной, т. е. увеличивает сроки исследований.

Ранее [1] нами была предложена смешанная модель, основанная на одной точке перекрывания, что предполагает построение кривых роста в интервале 7—17 лет с помощью двухлетних поперечных исследований с наименьшими, в определенном смысле, потерями. Опишем предложенную модель.

Пусть $\{X_j^k\}$ ($k = 7, 8, \dots, 17$), ($j = 1, 2, \dots, M$), ($i = 1, 2, \dots, N_k$) матрица данных M соматических измерений $N = \sum N_k$ детей в K возрастных когортах в первый год исследований, а $\{Y_j^k\}$ — во второй год, т. е. область продольных исследований в предлагаемой смешанной модели ограничивается одним годом. Задача построения кривых роста в такой модели сводится к поиску наиболее схожих (в определенном смысле) объектов в $\{Y_j^k\}$ и $\{X_j^k\}$ наборах. Действительно, возьмем i -ого ребенка в 7-летней когорте— X_j^7 . Рост его в рамках продольной модели можно проследить лишь через год— Y_j^8 . Теперь, если в массиве $\{X_j^8\}$ отыскать наиболее близкий к Y_j^8 объект, например X_j^8 , то его дальнейшее развитие также можно проследить через год. Повторяя предыдущие рассуждения, можно построить последовательность X_j^7, X_j^8 и т. д., имитирующую наилучшим образом продольные исследования.

Поясним понятие схожести или «ближайшего аналога». Соматометрически каждый ребенок представляется точкой в 30-мерном пространстве своих параметров. Очевидно, чем ближе точки в соматометрическом пространстве этих координат, тем более схож морфологический статус детей. Под «морфологическим статусом» подразумевается как общий размер (крупность), так и пропорции (форма). Отметим, что существующие схемы анализа возрастной изменчивости длины тела обладают тем недостатком, что используемый признак характеризует в определенной мере лишь интегральный показатель размера, тогда как такой существенный показатель, как форма тела (пропорции), остается вне анализа. При многомерном описании объекта, при последующем анализе, учитываются как компоненты его размера, так и формы тела. Именно этот факт позволяет считать, что в анализируемом множестве соматометрических параметров заключена достаточная информация об интегральных показателях телосложения, обладающих высокой степенью генетической детерминации [2]. Поскольку сам процесс возрастных преобразований в организме ребенка (включая и хронологические аспекты этих явлений) наследственно обусловлен [3], можно предположить, что определенная часть генетической информации о ростовых процессах заложена в совокупности исследуемых признаков фиксированного возраста (статистические данные). Именно поэтому предлагаемая схема имитации приобретает биологический смысл.

Введем строгую меру схожести: объект i во множестве $\{X^k\}$ будет считаться «ближайшим аналогом» объекта j во множестве $\{Y^k\}$ если:

$$\varphi_{ij}^k = |X_i^k - Y_j^k| = \sum_{l=1}^n (X_{il}^k - Y_{jl}^k)^2 = \min_{\substack{l \in X^k \\ q \in Y^k}} \varphi_{lq}^k.$$

Очевидно, что определение не совсем отвечает интуитивно ожидаемому. Действительно, возможен случай, когда почти все X_{il}^k и Y_{jl}^k близки. Тогда φ_{ij}^k будет минимальной, однако по какому-то параметру разница может достичь достаточно большой величины. Для исключения подобных ситуаций вводится дополнительное ограничение:

$$\max_i \left| \frac{X_{il}^k - Y_{jl}^k}{X_{il}^k + Y_{jl}^k} \right| < A;$$

Нормировка осуществляется для нивелирования масштаба различных показателей. Кроме того, возможен случай, когда даже «ближайший аналог» оказывается недостаточно близким. Для исключения большой ошибки вводится ограничение на «ближайшего аналога».

$$\varphi_{ij}^k = |X_i^k - Y_j^k| < B.$$

Параметры модели A и B подбираются эмпирически из машинных экспериментов (в настоящей работе оптимальным считается $A=0.4$ и $B=116$).

Необходимо отметить еще одну немаловажную деталь алгоритма. «Ближайший аналог» в когорте $\{X\}^k$ отбирается последовательно для объектов из $\{Y\}^k$, т. е. не исключается случай, когда тот же объект из $\{X\}^k$ будет ближайшим для нескольких из $\{Y\}^k$. Именно это соображение вынудило нас отказаться от строгой идеи взаимнооднозначного соответствия между $\{X\}^k$ и $\{Y\}^k$. Таким образом, в одну сторону соответствие неоднозначно. Это означает, что модель вырождается в дерево, т. е. один и тот же объект 7-летней когорты в рамках модели будет соответствовать нескольким 17-летним.

Машинные эксперименты показали, что в рамках описанной модели практически невозможно завершить имитацию. Это в основном связано с недостатками исходного материала (часть детей, переходящих в другие школы, выпадала из анализа второго года). Для преодоления этих трудностей вводится дополнительное понятие — «дублер», каковым является следующий по близости объект из $\{X\}^k$ когорты. Таким образом, если имитационное дерево обрывается на каком-то шаге (ввиду отсутствия данных второго года по «ближайшему соседу»), то имеется возможность его продолжения за счет «дублера». На рис. 1 представлены результаты проведенной имитации.

Очевидно, что предлагаемая модель должна быть весьма чувствительна к сензитивным периодам роста и развития, а количественной оценкой сензитивности может служить относительная суммарная ошибка имитации для данной когорты (рис. 2).

Заранее в модели не была заложена идея классификации кривых роста, однако, как видно из рис. 1, модель сама в процессе работы

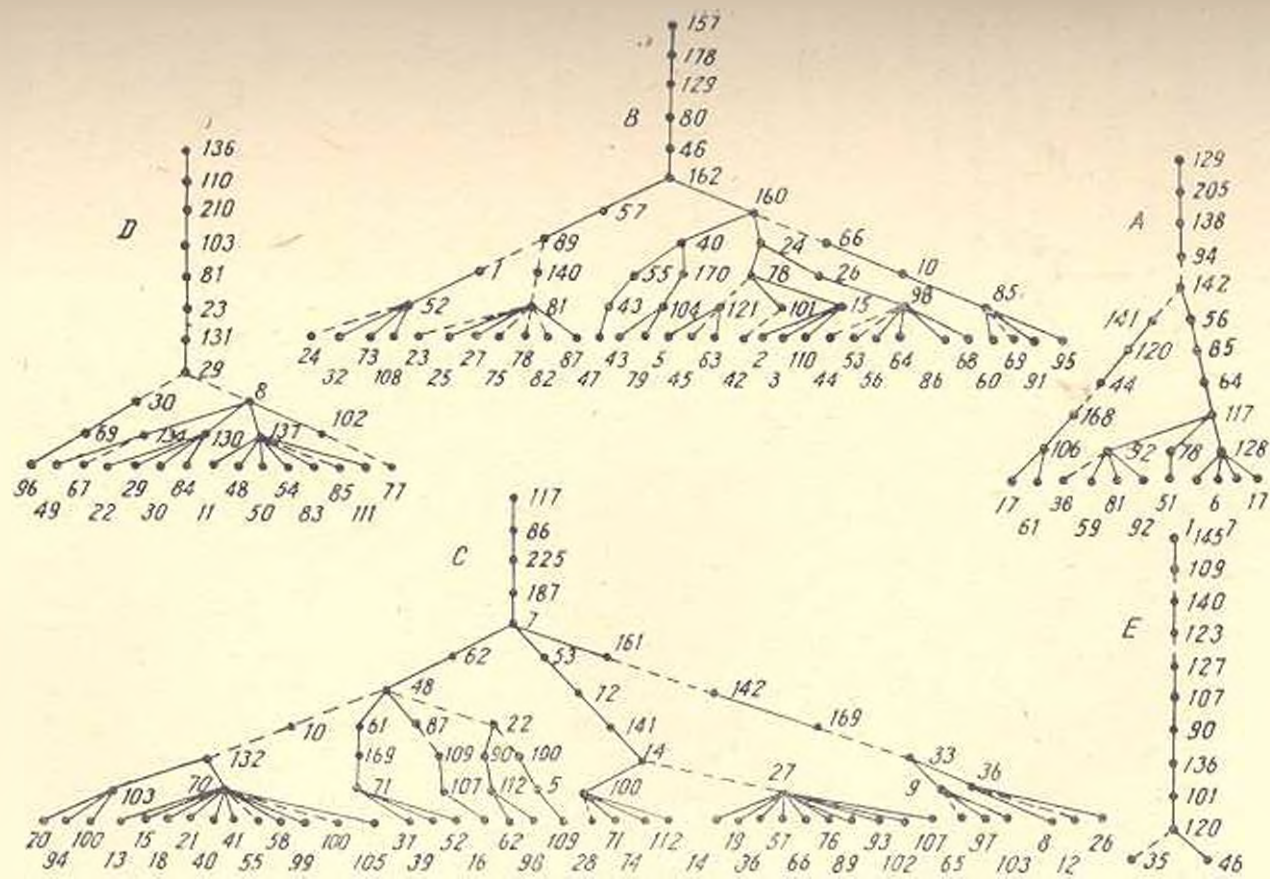


Рис. 1 Классификация кривых роста путем разбиения на деревья.

одновременно в определенной степени решала и задачу классификации, путем разбиения кривых на отдельные деревья. Общим для большинства полученных деревьев является тот факт, что до определенного возраста (11—13 лет) процесс построения кривой роста однозначен, дискретность появляется на следующих этапах, в период полового созре-

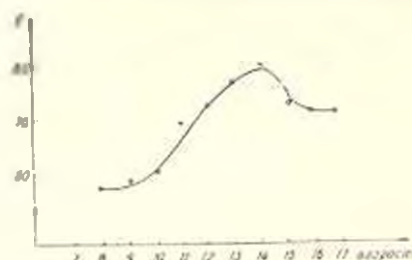


Рис. 2. Возрастная ошибка суммарной динамики.

вания. Из полученных типов кривых 2, по всей вероятности, являются крайними вариантами, соответствующими росту наиболее крупных (класс E) и, наоборот, наиболее низкорослых, но с относительно большой массой (класс D) детей.

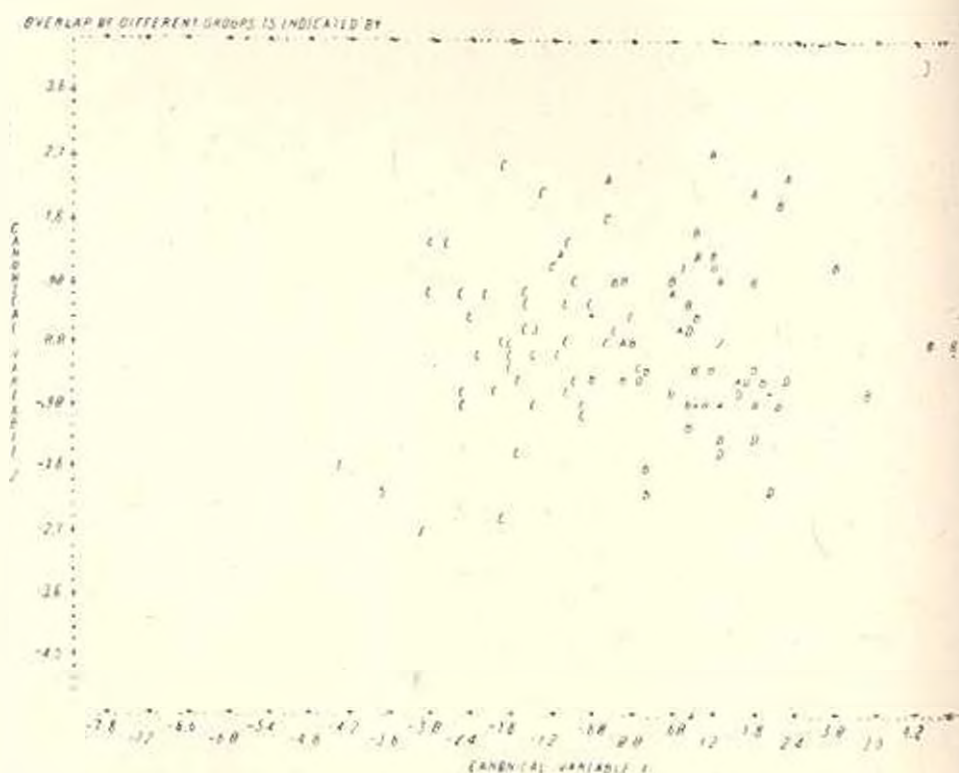


Рис. 3. Проекция анализируемого множества на плоскость Фишера.

Для оценки значимости различий между группами детей, принадлежащих к одному дереву, по сравнению с группами, относящимися к разным деревьям, был проведен ДА с пошаговым выявлением информативных признаков. Оценка разделения с помощью F-аппроксимации U-статистики Уилкса дает достоверность более чем на 99,9% по-

уровне. В информативный набор входит 6 признаков, описывающих основные линейные (длина тела и предплечья), широтные (плечевой диаметр, эпифиз предплечья) и обхватные (обхват плеча и предплечья) размеры. Полученные ЛДФ обеспечивают в среднем около 80% реклассификации и свыше 70% верного прогноза. На рис. 3 представлена проекция анализируемого множества на плоскость Физнера. На основании одномерных статистик и анализа этого рисунка можно считать, что, согласно кривым роста, выявлено 3 основных варианта типизации 17-летних детей и 2 промежуточных. Отметим, что основные типы распознаются достаточно хорошо (100% — С, 97% — В), тогда как ошибки в основном приходится на промежуточные классы (40% — А, 47% — D).

Практический выход полученных результатов заключается в принципиальной возможности прогнозирования типа кривой роста конкретного ребенка с получением дефинитивной картины его морфофункционального статуса. Очевидно, что на основании этих данных можно прогнозировать также популяционную картину на заданный промежуток времени.

Практическая ценность полученных результатов, на наш взгляд, заключается в следующем.

В литературе отражен ряд безуспешных попыток аналитического описания кривых роста в рамках единой модели, для которой оценка параметров проводится либо на всем материале (регрессионная модель), либо же осуществляется индивидуальное сглаживание, результаты которого непосредственно не могут быть перенесены на другой объект. Одной из основных причин этих неудач, по-видимому, является большая гетерогенность кривых для анализа их в рамках одной модели и практическая нецелесообразность индивидуального расчета. Полученный в данной работе результат занимает фактически промежуточное положение в терминах вышеописанных подходов анализа роста кривых.

Л И Т Е Р А Т У Р А

1. Ордуханян А. А., Хачатрян Г. Н. Тез. III Всесоюз. конф. «Физиология развития человека», 260, М., 1985.
2. Bergman P., Orczykowska Scortkowska Z. *Stud Phys. Anthropology*, 8, 61—78, 1976.
3. Geddal. 4th Congr. Human Genetics. Paris, 6—11, sept, 1971. *Experta Medical Congr. Ser.* № 2 3, 72—73, Amsterdam, 1971.
4. Geddal, Branci G. *Acta genetic Med. gemehol*, 24, 15—30, 1975.
5. Kowalski C. G., Gutre K. E. *Longitudinal data analysis Growth*, 131, 169, 38, 1974.
6. Preece M. A., Baines M. J. *Annales of Human Biology*, 1, 5, 1—24, 1978.
7. Relthford H. J., Lees F. C., Byard P. J. *Human Biol.*, 4, 50, 461—475, 1978.
8. Van't Hoff M. A., Ronde M. J., Kowalski C. J. *Human Biolog.*, 4, 49, 165—179, 1977.
9. Wright S. *Genetical group and special size factors genetics*, 12, 5, 603—619, 1932.

Поступило 15.IV 1986 г.