

2009



*Information Technologies in Science,
Sociology, Economics
and Business*

**Информационные
технологии в науке,
социологии, экономике
и бизнесе**

САТЕЛЛИТНЫЙ СИМПОЗИУМ 3: НОВЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В НАУЧНЫХ ИССЛЕДОВАНИЯХ

УДК – 519.24

МЕТОД ГЛАВНЫХ КОМПОНЕНТ КАК СРЕДСТВО ОБРАБОТКИ ДАННЫХ ПЕРОРАЛЬНОГО ГЛЮКОЗО-ТОЛЕРАНТНОГО ТЕСТА

Авшалумов А. Ш., Марковский В. Б., Синицына Е. Н., Филаретов Г. Ф.

Московский Институт Кибернетической Медицины (МИКМ)

THE METHOD OF PRINCIPAL COMPONENTS AS AN INSTRUMENT FOR ANALYSIS OF ORAL GLUCOSE TOLERANCE TEST DATES

Avshalumov A. Sh., Filaretov G. F., Markovsky V. B., Sinityna E. N.

Moscow Research Institute for Cybernetic Medicine (MRICM)

It is proposed to use the method of principal components (MPC) as an instrument for data's analysis of oral glucose tolerance test (OGTT), which plays very important role in diabetic diagnoses. It is shown that with help of MPC we can get some new useful information about OGTT results.

Введение. Одной из важнейших проблем в современной медицине является, так называемая, проблема инсулинорезистентности. Под термином «инсулинорезистентность» принято понимать состояние организма, характеризующееся снижением чувствительности инсулинзависимых тканей к воздействию гормона инсулина, вырабатываемого в поджелудочной железе, что приводит к хронической гиперинсулинемии. Инсулинорезистентность и гиперинсулинемия – два фактора, определяющих развитие сахарного диабета и его осложнений, а также целого ряда патологических изменений обмена веществ и его регуляции объединенных под термином «метаболический синдром».

Как известно, в последние 30 лет отмечается резкий рост заболеваемости сахарным диабетом во всем мире, особенно в промышленно развитых странах. Каждые 10-15 лет число больных сахарным диабетом удваивается. Это заболевание ведет к увеличению летальности пациентов с сердечно-сосудистой патологией в 3 раза, риск развития инфаркта миокарда возрастает в 2 раза, патологии почек в 17 раз, гангрены нижних конечностей в 20 раз, гипертонической болезни – более чем в 3 раза.

Исследование инсулинорезистентности и закономерностей развития патологических состояний связанных с ней может привести к разработке эффективных способов профилактики развития не только сахарного диабета, но и всего комплекса признаков метаболического синдрома.

К настоящему времени предложено достаточно много способов оценки инсулинорезистентности. Наибольшее распространение среди них получил пероральный глюкозо-толерантный тест (ПГТТ) [1]. Тест реализуется путем приема внутрь натощак 75г сухой глюкозы, растворенной в 250 мл воды, и измерения концентрации иммунореактивного инсулина (ИРИ), а в расширенном варианте также глюкозы и С-пептида в плазме крови в начальный (нулевой) момент времени, т.е. до приема, и через 30, 60, 120 минут после приема. Стандартная обработка результатов теста достаточно проста и сводится к фиксации начального уровня ИРИ и его сопоставлению с конечным [2], оценке индекса CARO [3], равному отношению концентрации глюкозы к уровню ИРИ в начальный момент и расчету, так называемой, «площади под инсулиновой кривой» [1]. Последний показатель считается наиболее информативным и сводится по сути к вычислению суммы концентраций ИРИ плазмы до начала теста и через 30, 60 и 120 минут.

Как показывает практика, данные, полученные с помощью ПГТТ, в каждом отдельном случае достаточно хорошо отражают степень и характер нарушений секреции инсулина и гиперинсулинемии. Однако ясно, что используемая обработка измерений не позволяет извлечь из них всю полезную информацию в особенности, если проводится одновременное определение ИРИ, глюкозы и С-пептида. В этом случае желательно получить интегрированное представление о степени изменчивости всех показателей, что, возможно, позволит обнаружить их связь с клиническими проявлениями инсулинорезистентности. В свою очередь это даст возможность по-новому оценить и объективизировать данные полученные при диагностике состояния инсулинорезистентности и об эффективности методов ее коррекции. Далее рассматривается возможность использования для анализа результатов ПГТТ одного из известных статистических методов, а именно *метода главных компонент* (МГК) [4].

МГК эффективно решает задачу оптимального уменьшения размерности исходных многомерных данных путем перехода к новым переменным (главным компонентам или факторам), являющимся некоррелированными нормированными линейными комбинациями исходных наблюдаемых переменных. Эти

комбинации строятся таким образом, что их характеристики рассеяния (дисперсии) располагаются в убывающем порядке. Предполагается, что небольшое число главных компонент (обычно две – три) могут взять на себя (объяснить) большую часть общего рассеяния исходных данных.

Основной результат. Возможности МГК применительно к рассматриваемой задаче будем анализировать с помощью экспериментальных данных, полученных в Клинике Московского института кибернетической медицины при обследовании с помощью ПГТТ 32 пациентов с признаками инсулинорезистентности (исследуемая группа) и 11 практически здоровых людей, образующих контрольную группу. Ставилась задача прежде всего выявить принципиальную возможность и полезность использования МГК при обработке данных ПГТТ. Для этого важно определить, можно ли с помощью МГК четко классифицировать измерения на относящиеся к здоровым или больным и какова при этом информативность отдельных показателей (ИРИ, С-пептида и глюкозы и всей совокупности в целом).

На рис. 1 на плоскости первых двух главных компонент представлены результаты обработки данных ПГТТ для ИРИ (а), С-пептида (б), глюкозы (в) и всей совокупности (г), где квадратиками обозначены представители первой группы, а кружочками – второй.

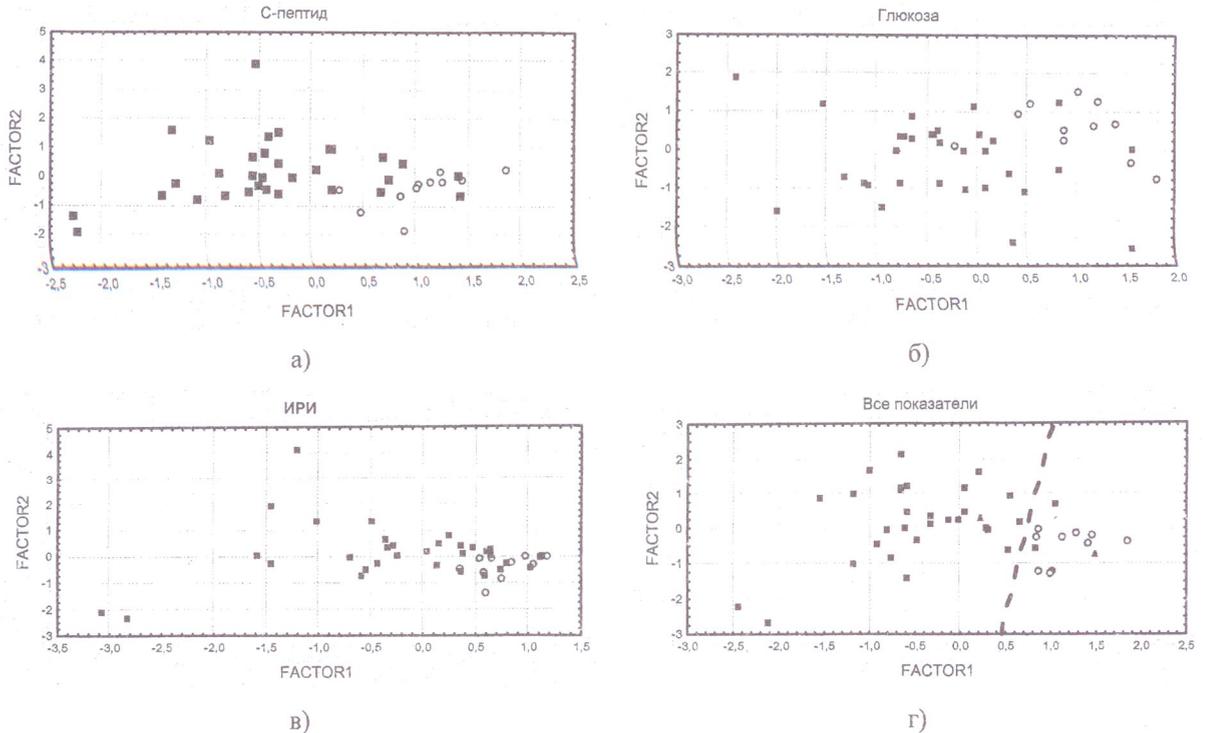


Рис. 1

Ясно, что, хотя точки различных групп в основном разделяются, однако в целом границу (например, линейную) провести достаточно сложно. Если, к примеру, использовать в качестве разграничивающей прямую, изображенную на рис. 1г), то ошибочно классифицированы только две точки (из первой группы они попали во вторую); две точки, обозначенные на этом рисунке треугольниками, также попали в чужие группы, но, как показал более тщательный анализ, лишь из-за неточности в исходной информации.

Несмотря на достаточно благоприятный результат, попытаемся улучшить его путем модификации исходных измерений. Предлагается в качестве исходных переменных использовать начальные показатели ИРИ, глюкозы и С-пептида и разности между соответствующими значениями в два момента времени: P1- разность между значениями, полученными в точке 30 мин и в начальной точке, P2 – между значениями в точках 60 и 30 мин, P3 – аналогично для точек 120 и 60 мин и, наконец P4 – разность между конечным (на 120 минуте) и начальным значениями. Поскольку из четырех перечисленных разностей только три независимы (любая может быть найдена по трем другим), встает вопрос, какую из них целесообразно исключить. На рис. 2 представлены результаты обработки данных с помощью МГК, когда исключена разность P4 (рис.2а), P3 (рис.2б), P2 (рис. 2в) и P1 (рис.2г).

Очевидными преимуществами обладает последний вариант, где точки различных групп подразделяются наиболее просто и четко с помощью только одного фактора (FACTOR1).

Представляет интерес вопрос о том, какие же исходные переменные наиболее тесно связаны с этим фактором. Эту связь можно оценить с помощью соответствующих коэффициентов парной корреляции. Выясняется, что наиболее информативными в указанном смысле являются начальные значения ИРИ и С-пептида (коэффициенты корреляции между этими переменными и первой главной компонентой равны -0,810 и

-0,829), разности P4, т.е. разности между значениями на 120 минуте и в начальный момент для всех показателей (коэффициенты корреляции равны для ИРИ -0,733, для глюкозы -0,723, для С-пептида -0,768) и разность P2 для глюкозы (коэффициент корреляции -0,713).

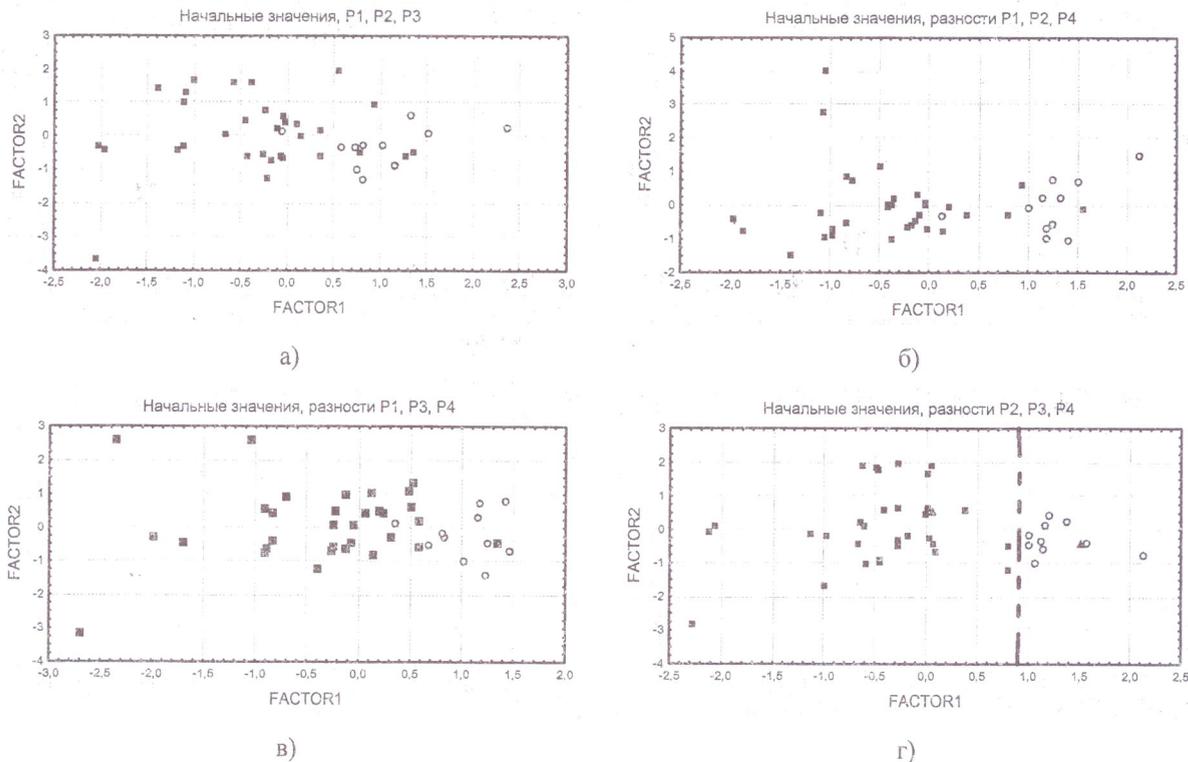


Рис. 2

Выводы.

1) Полученные результаты свидетельствуют о перспективности использования метода главных компонент при обработке данных ПГТТ.

2) Необходимо продолжить исследование возможностей данного метода по крайней мере по двум направлениям:

- определение эффективности используемых терапевтических процедур путем отслеживания динамики перемещения соответствующей отображающей точки данного пациента на плоскости главных компонент по отношению к выявленной разграничивающей линии;

- анализ взаимосвязи расположения точек на плоскости главных компонент и их явно выраженного группирования в некоторые кластеры с используемыми в настоящее время характеристиками инсулинорезистентности и метаболического синдрома.

ЛИТЕРАТУРА

1. Творогова М.Г., Яськова К.Н., Мычка В.Б., Чазова И.Е. Инсулинорезистентность и методы ее диагностики // Лабораторная медицина, 2003, №6.
2. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. // Diabetes. -1979: 28: 1039-1057.
3. Caro F. Insulin resistance in obese and nonobese man. // J. Clin. Endocrinol. Metab. 1991; vol. 73: 691-695.
4. Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы экономики. Учебник для вузов. – М.: ЮНИТИ, 1998. 1022 с.