УДК 51-76+57.087+ 577.3

А.Ю. Суковатова, А.Н. Романов, А.О. Ковригин

Использование регрессионного анализа для моделирования диэлектрических свойств биологических жидкостей на примере сыворотки крови

A.Yu. Sukovatova, A.N. Romanov, A.O. Kovrigin

Application of Regression Analysis for Modeling Dielectric Characteristics of Biological Liquids on the Example of Blood Serosity

Установлены зависимости диэлектрических параметров сыворотки крови от совокупности биохимических показателей. Предложен новый подход к разработке методики экспресс-анализа биологических жидкостей на основе использования зависимостей диэлектрических характеристик и биохимических параметров биологических жидкостей человека.

Ключевые слова: диэлектрические характеристики, многомерный анализ данных, регрессия на главные компоненты, проекция на латентные структуры, биологические жидкости.

The dependence of dielectric parameters of blood serosity on group of biochemical factors is studied. New approach to the developing express-analysis method of biological liquids basing on interconnection of dielectric characteristics and biochemical parameters of human biological liquids is suggested.

Key words: dielectric characteristics, data multivariate analysis, PCR, PLS, biological liquids.

Эффективность лечения многих заболеваний, в том числе представляющих реальную угрозу для жизни человека, в значительной мере зависит от того, на какой стадии развития болезни поставлен диагноз. В связи с этим важной проблемой профилактики заболеваний является их ранняя диагностика, основанная на оценке функционального состояния человеческого организма и выявлении различных видов патологий, связанных с нарушениями водно-солевого баланса, углеводного обмена и т.д.

Одной из распространенных и доступных для анализа жидкостей человека является кровь, состоящая на 55% из жидкой части (плазмы) и на 45% из клеточных форменных элементов (эритроцитов, лейкоцитов и тромбоцитов).

Кровь имеет определенный физико-химический состав, в значительной мере зависящий от естественно-обменных (в том числе возрастных) и патологических процессов.

Появление и развитие в организме человека злокачественных новообразований вызывает качественные и количественные изменения в фазовом составе крови и в молекулярных параметрах входящих в нее белков, что может быть использовано для диагностических целей. Основной проблемой при разработке методов диагностики различных заболеваний является необходимость выбора таких параметров крови, которые с наибольшей вероятностью позволяют различить здорового и больного данным заболеванием пациента.

Изменения состава крови, наблюдаемые при опухолевых заболеваниях, могут быть использованы для диагностических целей [1]. Учитывая биологическую доступность крови для анализов, исследование ее состава и его изменений является наиболее информативным при разработке эффективных методов ранней диагностики заболеваний, основанных на изучении электрофизических параметров [2].

Изучение диэлектрических свойств биологических жидкостей ведется достаточно давно. Разными авторами исследовались электрофизические и диэлектрические свойства биологических клеток, а также составляющих их отдельных компонент: протеинов, ДНК, лизосом, мембранных комплексов, а также различных фаз воды, присутствующей в клеточной мембране. Исследовались диэлектрические свойства плазмы крови [3–8], состояние гидратного окружения эритроцитов человека при моделировании различных путей трансмембранной передачи сигнала в клетку. Известны работы по определению диэлектрических характеристик кровяных клеток, находящихся в стадии апоптоза [9]. Появление подобных публикаций в международной печати указывает на возрастающий интерес к этой проблеме.

Для изучения структуры, фазового состава и диэлектрических параметров крови целесообразно использовать дециметровый диапазон, в котором наиболее ярко должно проявляться влияние белков, аминокислот и других растворенных веществ,

а также наблюдаются заметные различия в диэлектрических свойствах разных фаз воды. В данной работе приведены результаты исследований диэлектрических характеристик крови и ее составляющих в частотном диапазоне 0,4…1,2 ГГц.

Для ранней диагностики онкологических заболеваний перспективен метод, основанный на измерении электрофизических параметров крови, зависящих от заболеваемости [10–11].

В данной работе ставится задача использования регрессионного анализа для моделирования диэлектрических свойств биологических жидкостей на примере сыворотки крови с целью создания методики экспресс-анализа биологических жидкостей. Исследованы зависимости показателей преломления и поглощения электромагнитной волны, проходящей через образец сыворотки, от следующих биохимических показателей: билирубина, общего белка, мочевины, глюкозы, калия, натрия.

Методика измерений. Для изучения изменений диэлектрических характеристик крови, связанных с развитием злокачественных новообразований, был выбран дециметровый диапазон, в котором наблюдаются заметные различия в диэлектрических свойствах разных водных фракций, образующихся в результате взаимодействия с белками, аминокислотами и другими присутствующими в крови веществами. Диэлектрические характеристики исследуемых биологических объектов измеряли методом мостовых схем, обладающим высокой точностью определения диэлектрических параметров дисперсных гетерогенных смесей и жидкостей, характеризующихся высокой поглощательной способностью. Лабораторная установка, выполненная с использованием измерительного и опорного каналов промышленного фазометра ФК2-18, позволяла проводить измерения диэлектрических параметров жидкостей в диапазоне частот от 300 МГц до 5,0 ГГц и имела следующие технические характеристики: диапазон однозначных измерений фазы ±180; диапазон измерения затуханий от 0 до 60 дБ; погрешность измерений ослабления 0,5 дБ, фазы – 2°.

Отбор образцов крови и ее анализ проводили согласно существующим методикам в Алтайском филиале Российского онкологического научного центра им. Н.Н. Блохина. Для изучения влияния компонентного состава и отдельных элементов на диэлектрические свойства исследовали цельную кровь (с добавлением в нее цитрата натрия для предотвращения свертываемости), плазму (после удаления из цельной крови форменных элементов путем центрифугирования) и сыворотку, полученную из плазмы в результате удаления из нее фибриногена.

В качестве основной величины, количественно характеризующей содержание веществ в исследуемом образце, использовали массовую долю S растворенного вещества, которую определяли по формуле $S = M_c / M$, где M, M_c – массы образца и сухого остатка, измеренные путем взвешивания на аналитических весах с точностью до 0,001 г до и после выпаривания воды из образца в термостате при температуре 215 °С. Образцы с требуемой S получали путем добавления исследуемых биологических жидкостей, содержащих растворенные вещества, в дистиллированную воду.

В эксперименте измеряли на разных частотах ослабление A и сдвиг фазы φ электромагнитной волны в исследуемом образце, используемые для расчета действительной (ε') и мнимой (ε'') частей комплексной диэлектрической проницаемости $\varepsilon = \varepsilon' + i\varepsilon''$, а также показателей преломления n и поглощения k, связанных с ε следующей формулой: $(\varepsilon)^{1/2} = n + ik$. Измерения диэлектрических параметров водных растворов проводили при температуре 25±1 °С.

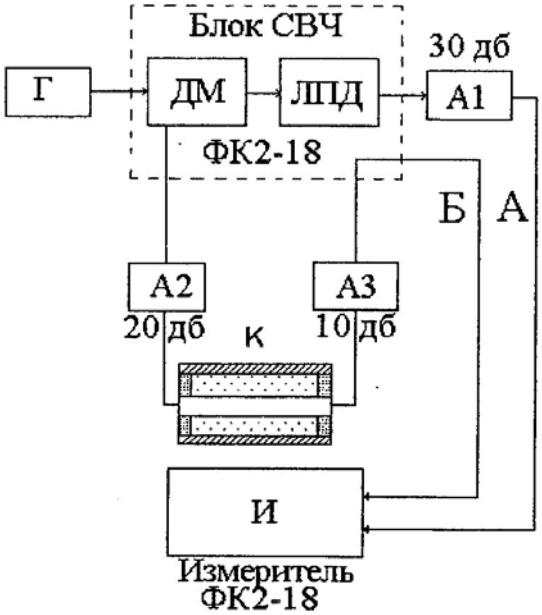

Схема установки мостового типа
на основе измерителя фаз

Измерительные контейнеры были выполнены в виде отрезков коаксиальной линии. Соединительные высокочастотные тракты в установке были выполнены в виде коаксиальных отрезков линий.

Процесс измерений заключался в следующем. С помощью делителя мощности (ДМ) производили балансировку мостовой схемы, в плечо которой устанавливали пустой измерительный контейнер. Сигнал от генератора подается на ДМ и делится поровну между опорным (А) и измерительным (Б) каналами. В начале измерений, при отсутствии в контейнере исследуемого образца, устанавливается нулевое значение разности фаз и амплитуд на фазометре. Затем в контейнер помещали образец и по индикатору измерителя ФК2-18 отсчитывали значения разности фаз и затухание.

Описание методики обработки данных. С целью учета происходящих изменений были определены регрессионные зависимости n и к сыворотки крови от одновременного изменения нескольких биохимических параметров сыворотки крови.

Для расчета использовали известные методы регрессионного анализа данных: регрессию на главные компоненты и проекцию на латентные структуры [12]. Проводили многомерную калибровку данных по 73 образцам сыворотки крови (использованные параметры: показатели преломления и поглощения, группа крови, частота сигнала, возраст, температура, данные биохимического анализа). Строили многомерную регрессионную модель (X, Y) для экспериментальных данных, которую затем использовали для определения новых значений Y по новым измерениям X. Смысл этой операции заключался в том, чтобы получить возможность в дальнейшем не измерять значения Y, а определять их по значениям X.

Регрессию на главные компоненты можно рассматривать как двухэтапную процедуру: сначала с помощью метода главных компонент преобразуется матрица X, а затем получившаяся матрица непосредственно используется в модели множественной линейной регрессии [12].

Работу алгоритма проекции на латентные структуры удобнее всего представить как два анализа методом главных компонент, производимых одновременно – для X и для Y. Непосредственно вовлекая структуры данных Y в декомпозицию X, можно получить такие же результаты прогноза, как и при регрессии на главные компоненты, но основываясь на меньшем количестве компонент [12].

1. Исследовали методами регрессии на главные компоненты и проекции на латентные структуры зависимости n и k от следующих параметров: возраст z (диапазон изменений от 32 до 78 лет), частота сигнала f (от 0,4 до 1,2 ГГц), группа крови g (g = 1…4), резус-фактор r (0, 1), температура крови t (0…25 °С).

Получены следующие регрессионные зависимости:

Метод проекции на латентные структуры

$n = 9{,}478 - 0{,}021t - 0{,}636f - 0{,}614r + 0{,}141g + 0{,}008z$,

$k = 4{,}858 - 0{,}012t - 2{,}633f + 0{,}254r - 0{,}129g - 0{,}006z$.

Метод регрессии на главные компоненты

$n = 8{,}655 - 0{,}02t - 0{,}022f - 0{,}455r + 0{,}198g + 0{,}006z$,

$k = 1{,}644 - 0{,}003t - 0{,}011f + 0{,}118r + 0{,}139g + 0{,}008z$.

2. Исследовали методом главных компонент зависимости n и k от следующих биохимических параметров: билирубин общий (Bi), общий белок (Be), мочевина (M), глюкоза (G), калий (K), натрий (N).

Получены следующие регрессионные зависимости:

Метод проекции на латентные структуры

$n = 13{,}73 - 0{,}006Bi - 0{,}004Be + 0{,}015M - 0{,}015G - 0{,}114K - 0{,}028N$,

$k = 9{,}308 - 0{,}083Bi - 0{,}019Be + 0{,}108M + 0{,}081G + 0{,}123K - 0{,}052N$.

Метод регрессии на главные компоненты

$n = 13{,}289 - 0{,}007Bi - 0{,}002Be + 0{,}028M - 0{,}008G - 0{,}09K - 0{,}028N$,

$k = 7{,}608 - 0{,}087Bi - 0{,}017Be + 0{,}116M + 0{,}085G + 0{,}09K - 0{,}04N$.

Исследованные величины изменялись в следующих пределах:

– общий билирубин Bi = 6,8…17,60 мкмоль/л (норма – 8,5…20,5),

– общий белок Be = 44…75 г/л (норма – 65…85),

– мочевина M = 4,8…17,90 ммоль/л (норма – 2,5…8,3),

– глюкоза G = 4,1…8,6 ммоль/л (норма – 3,3…6,5),

– калий K = 2,9…5,1 ммоль/л (норма – 3,4…5,3),

– натрий N = 132…144 ммоль/л (норма – 130…150).

Использование установленных эмпирических зависимостей n и k от исследованных параметров позволяет рассчитать численные значения n и k по биохимическим параметрам сыворотки крови для разных случаев. Построение таких моделей для здоровых людей позволит контролировать происходящие в их организме изменения и, таким образом, осуществлять диагностику заболеваний (в принципе не только онкологических).

Для показателя k число главных компонент как в регрессию на главные компоненты, так и в проекцию на латентные структуры равно 1, для показателя n число главных компонент для регрессии на главные компоненты равно 6, для проекции на латентные структуры – 2.

В данном случае целесообразнее использовать модель, построенную с помощью метода проекции на латентные структуры, так как он использует меньшее число главных компонент и значение ошибки предсказания для этого метода меньше.

Библиографический список

1. Клиническая интерпретация лабораторных исследований / под ред. А.Б. Белевитина, С.Г. Щербака. – СПб., 2006.

2. Поляков В.М., Шмаленюк А.С. СВЧ-термография и перспективы ее развития. Применение в медицине и народном хозяйстве // Обзоры по электронной технике. Сер. 1. Электроника СВЧ. – М., 1991. – Вып. 8 (1640).

3. Pitera J.W., Falta M., van Gunsteren W.F. Dielectric properties of proteins from simulation: the effects of solvent, ligands, pH, and temperature // Biophys. – 2001. – J. 80.

4. Bonincontro A., Cinelli S., Onori G., Stravato A. Dielectric Behavior of Lysozyme and Ferricytochrome-c in Water/Ethylene-Glycol Solutions // Biophys. – 2004. – J. 86.

5. Cherepanov D.A., Feniouk B.A., Junge W., Mulkidjanian A.Y. Low dielectric permittivity of water at the membrane interface: effect on the energy coupling mechanism in biological membranes // Biophys. – 2003. – J. 85.

6. Takashima S., Gabriel C., Sheppard R.J., E.H. Grant. Dielectric behavior of DNA solution at radio and microwave frequencies // Biophys. – 1984. – J. 46.

7. Takashima S., Gasaleggio A., Giuliano F., Morando M., Arrigo P., Ribdella S. Dielectric relaxation time and structure of bound water in biological materials // Biophys. – 1986. – J. 49.

8. Kashyap S.C. Dielectric properties of blood plasma // Electronics Letters. – 1981. – V. 17, № 19.

9. Kanapitsas A., Vartzeli-Nikaki P., Konsta A.A., Visvardis E., Sideris E.G. Dielectric study of the programmed cell death, apoptosis, in human leukemia // IEEE Transactions on Dielectrics and Electrical Insulation. – 2006. – Vol. 13, №5.

10. Романов А.Н., Винокурова Е.Ю., Ковригин А.О., Лазарев А.Ф., Лубенников В.А., Романова Н.А., Комаров С.А. Диэлектрические характеристики биологических жидкостей человека при развитии онкологических заболеваний. – Барнаул, 2008.

11. Романов А.Н., Ковригин А.О., Винокурова Е.Ю., Лазарев А.Ф., Лубенников В.А., Романова Н.А. К разработке методов диагностики злокачественных новообразований в организме человека на основе измерения диэлектрических характеристик крови в микроволновом диапазоне // Наука – Алтайскому краю, 2008: сб. науч. ст. по результатам НИР, выполненных за счет средств краевого бюджета.– Барнаул, 2008. – Вып. 2.

12. Эсбенсен К. Анализ многомерных данных: избр. главы / пер. с англ. С.В. Кучерявского; под ред. О.Е. Родионовой. – Барнаул, 2002.