

Математико-компьютерная обработка КВВ-экспериментов по распознаванию легочных заболеваний

А. Г. Чучалин, В. Б. Кудрявцев, Д. В. Алексеев, Э. Х. Анаев,
Т. Н. Анохина, М. В. Носов, А. И. Ревельский,
И. А. Ревельский, А. А. Родионов

Статья посвящена исследованию возможности диагностики хронической обструктивной болезни легких (ХОБЛ) и бронхиальной астмы, основанной на изучении состава среднелетучих органических соединений на ультранизком уровне в конденсате выдыхаемого воздуха (КВВ), с использованием методов выделения примесей и хромато-масс-спектрометрическом анализе всего концентрата аналитов КВВ-экспериментов. На основании обработки результатов анализа 70 образцов КВВ, собранных у здоровых, больных ХОБЛ и бронхиальной астмой людей, с использованием специально разработанного алгоритма, основанного на линейных методах теории распознавания образов, установлено, что можно различить здоровых и больных бронхиальной астмой с надежностью 75%, здоровых и больных ХОБЛ — с надежностью 85%, больных бронхиальной астмой и ХОБЛ — с надежностью 83%. Установлен ряд веществ, которые могут быть потенциальными маркерами рассматриваемых заболеваний.

Ключевые слова: распознавание образов, хроническая обструктивная болезнь легких, бронхиальная астма, конденсат выдыхаемого воздуха, хромато-масс-спектрометрический анализ.

Введение

Исследованию химического состава выдыхаемого воздуха с целью изучения возможности диагностики различных лёгочных заболеваний посвящено большое число работ. Это связано с тем, что применяемые в медицине инструментальные методы диагностики не обеспечивают высокой степени достоверности на ранних стадиях заболеваний, что приводит к неправильному лечению и связанным с ним негативным последствиям. Подавляющее большинство публикаций по выдыхаемому воздуху посвящено исследованию состава газообразных, легколетучих и, в меньшей степени, среднелетучих органических соединений, которые концентрируют с использованием адсорбции и, реже, абсорбции с последующим анализом малой части абсорбата. Наиболее эффективный метод анализа, используемый в данной области, основан на газохроматографическом разделении в комбинации с различными методами детектирования. Масс-спектрометрическое детектирование позволяет производить одновременное определение веществ разных классов. Однако результаты анализа газообразной части выдыхаемого воздуха неадекватно отражают состав смеси выдыхаемых эндогенных соединений в связи с тем, что многие летучие и среднелетучие соединения могут поступать в лёгкие из окружающей среды.

Отдельно выделяют такой объект, как водный конденсат выдыхаемого воздуха (КВВ), получаемый путём охлаждения выдыхаемого воздуха до температуры -10°C . В настоящее время существует множество (сотни) публикаций по анализу КВВ. В подавляющем большинстве из них речь идёт об определении нелетучих органических соединений различными методами. В литературе известны лишь единичные публикации по исследованию состава среднелетучих примесей в КВВ. Концентрации присутствующих в КВВ летучих и среднелетучих органических соединений очень низкие ($10^{-7}\%$ по массе и ниже), и состав их практически неизвестен.

Изучение состава смеси этих примесей в КВВ представляет несомненный интерес в связи с возможностью использования полученных данных в целях диагностики лёгочных заболеваний. Анализ среднелетучих примесей, сконцентрированных в КВВ, может лучше от-

ражать состав выдыхаемых органических соединений, по сравнению с анализом адсорбата или абсорбата выдыхаемого воздуха. Продукты жизнедеятельности организма попадают из различных органов в кровь, из крови в межклеточную жидкость лёгких, а оттуда на поверхность слизистой и в аэрозоль, который затем собирают в конденсоре. Для исследования состава смесей соединений, содержащихся в КВВ, актуальна разработка способа их одновременного определения на ультраследовом уровне, так как существующие способы не обеспечивают такой возможности.

В связи с этим, целью настоящей работы являлась разработка способа определения среднелетучих органических соединений различной полярности на ультраследовом уровне в КВВ методом ГХ/МС и изучение состава смесей зарегистрированных этим способом соединений, выделенных из образцов КВВ, собранных у здоровых, больных бронхиальной астмой и хронической обструктивной болезнью лёгких (ХОБЛ) людей с использованием компьютерной математики.

1. Оборудование, исходные вещества и методика эксперимента

В работе было использовано перечисленное ниже оборудование.

- Установка для сбора конденсата выдыхаемого воздуха ECoScreen (производитель Erich Jaeger).
- Хроматограф Trace GC ultra (Thermo Electron corp.).
- Масс-спектрометр Trace DSQ (Thermo Electron corp.) с электронной и химической ионизацией.
- Микрошприц на 10 мкл (SGE Ltd.).
- Среднеполярная капиллярная колонка RTX-200MS (Restek corp.): неподвижная фаза — трифторпропилметилполисилоксан, толщина плёнки неподвижной фазы — 0,25 мкм, диаметр колонки — 0,32 мм, длина — 30 м.
- Самодельное устройство для охлаждения участка капиллярной хроматографической колонки длиной 6 см (элемент Пельтье), обеспечивающее охлаждение до -20°C .

- Устройство для сорбционного концентрирования самодельное. Устройство состоит из стального картриджа, внутри которого помещён кварцевый вкладыш с сорбентом. В качестве сорбента была использована комбинация из сверхтонкого кварцевого волокна (СКВ) и сорбента Tenax GC. Концентрирование осуществляли в потоке газа (азота, 5–100 мл/мин) при подаче в него пробы органического раствора шприцевым дозатором со скоростью 5–25 мкл/мин. Во время ввода пробы в инжектор хроматографа картридж соединяли с инжектором посредством иглы. Перенос концентрата осуществляли в потоке гелия (1–5 мл/мин) при нагреве картриджа потоком горячего воздуха.

В работе были использованы следующие модельные соединения для газовой хроматографии производства SIGMA-ALDRICH: деканол-1, ундеканол-1, додеканол-1, тетрадеканол-1, додекановая кислота, тетрадекановая кислота, гексадекановая кислота, додекан, докозан, п-цимен, дифенилметан, нафталин, гептанон-2, гептанон-4, 5-метилгексанон-2, 5-метилгептанон-3, 2,6-диметилгептанон-4, нонанон-2, додеканон-2, октаналь, нонаналь, деканаль, ундеканаль, додеканаль. Данные вещества были использованы при составлении модельных смесей.

Были приготовлены 2 типа растворов модельных смесей: 17-ти и 24-х компонентные. Отличие в составе модельных смесей заключается в наличии в 24-х компонентной последних 7 стандартов, указанных в списке. Исходные модельные растворы готовили в ацетоне, рабочие растворы — в МТВЭ, которые получали последовательным разбавлением исходных. Концентрации веществ в исходных смесях модельных соединений составляли от $9,5 \times 10^{-10}$ до $7,7 \times 10^{-8}$ г/мкл и от $5,1 \times 10^{-9}$ до $9,2 \times 10^{-8}$ г/мкл, соответственно. Для решения различных задач использовали модельные смеси, полученные разбавлением исходных в 10, 100 и 1000 раз. Для установления степеней извлечения модельных соединений при жидкость-жидкостной экстракции из водных растворов метилтретбутиловым эфиром с предварительным высаливанием были приготовлены водные растворы модельных соединений (диапазон концентраций 3×10^{-9} – 1×10^{-7} г/мкл).

Для приготовления растворов смесей модельных соединений использовали ацетон и МТВЭ марки «для ВЭЖХ».

В качестве газа-носителя использовали гелий (марки А). В процессе сорбционного концентрирования использовали азот (ОСЧ, 99,995%). Для химической ионизации использовали газ-реагент изобутан (95%).

Скорость потока газа-носителя меняли в пределах от 1 до 2 мл/мин. Начальную температуру термостата хроматографа варьировали в районе 40–45°C, время начальной изотермы — 3–5 мин. Скорость нагрева термостата хроматографа меняли в пределах от 5 до 50°C, разделив температурную программу на 2 участка в точке 200°C. Конечную температуру термостата меняли в пределах 250–300°C. Время конечной изотермы меняли от 1 до 5 мин. Температуру инжектора хроматографа и температуру соединительной линии меняли вместе с конечной температурой термостата от 250 до 300°C.

Масс-спектрометрическое детектирование с химической ионизацией осуществляли в режиме регистрации полного ионного тока. Сканирование осуществляли в диапазоне значений m/z от 35 до 350. Температура источника ионов составляла 250°C. Энергия электронов — 70 эВ.

При масс-спектрометрическом детектировании поток газа-реагента (изобутан) меняли в пределах от 0,5 до 2,5 мл/мин. Температуру источника ионов меняли в диапазоне от 120 до 150°C. Энергия электронов, испускаемых термокатодом, составляла 80 эВ. Сканирование проводили в диапазоне масс от 65 до 350 m/z . Детектирование проводили в режимах сбора данных по полному ионному току (TIC) и селективного ионного мониторинга (SIM) по двум основным массам в масс-спектре каждого вещества.

При изучении сорбции и термодесорбции модельных соединений проводили изучение влияния таких параметров, как температура десорбции, объёмная скорость потока газа-носителя, температура колонки хроматографа, температура начального участка колонки длиной 6 см и продолжительность термодесорбции, на процесс переноса вещества с сорбента в инжектор газового хроматографа. Температуру термодесорбции и объёмную скорость потока газа-носителя ва-

рьировали совместно в диапазоне от 250 до 300°C и от 1 мл/мин до 5 мл/мин. Гелий подавали в картридж с сорбентом за минуту до его соединения посредством иглы с инжектором хроматографа.

Значения начальной температуры термостата составляли 40, 37, 35°C. Температуру начального участка колонки меняли в пределах от 10 до -20°C с интервалом 10°C. Продолжительность термодесорбции — от 1 до 5 мин с интервалом 1 мин.

Эксперимент по вводу больших проб

Сорбционное концентрирование модельных соединений из растворов (17-ти и 24-х компонентные модельные смеси) изучали с использованием в качестве сорбентов СКВ и Tenax GC в комбинации с СКВ, количества которых составляли 1,5–6 мг для Tenax GC и 0,5–2 мг для СКВ. Сорбент располагался в кварцевом лайнере, через который пропускали поток газа (азота). Посредством шприцевого дозатора с постоянной скоростью подавали в лайнер анализируемый раствор. Объёмная скорость потока анализируемого раствора — 5–25 мкл/мин, объёмная скорость потока газа — 5–100 мл/мин.

Термодесорбцию осуществляли в условиях, выбранных в результате исследования сорбции/десорбции смесей модельных соединений.

2. Результаты опытов и их обсуждение

Исследование условий ГХ/МС определения следовых количеств модельных соединений

Первоначально выбор условий газохроматографического разделения проводили для компонентов, входивших в состав 17-ти компонентной модельной смеси ($9,5 \times 10^{-11}$ – $7,7 \times 10^{-9}$ г/мкл). Варьировали скорость нагрева хроматографической колонки, а также значение конечной температуры нагрева колонки. В результате были выбраны следующие условия хроматографического разделения, которые были оптимальными для 17-ти и 24-х компонентных смесей. Температурная программа термостата: изотерма при 40°C — 3 мин; нагрев до 200°C со скоростью 5°C/мин; нагрев до 250°C со скоростью

40°C/мин; изотерма при 250°C — 1 мин. Температура инжектора хроматографа — 250°C, температура соединительной линии между хроматографом и масс-спектрометром — 250°C.

Масс-спектрометрическое детектирование осуществляли в режиме регистрации полного ионного тока. Сканирование осуществляли в диапазоне значений m/z от 35 до 350 (1,54 скан/с). Температура источника ионов составляла 250°C. Энергия электронов — 70 эВ. В случае химической ионизации энергия электронов составляла 80 эВ, диапазон масс — 120°C и поток изобутана — около 1 мл/мин.

Выбор оптимальных условий хроматографического разделения позволил изучить масс-спектры 24-х модельных соединений ($5,1 \times 10^{-10}$ – $9,2 \times 10^{-9}$ г/мкл). Для компонентов модельных смесей проведено сравнение экспериментальных масс-спектров с масс-спектрами из библиотеки масс-спектров NIST за 2005 год. Сравнение масс-спектров производилось по двум параметрам совпадения, которые рассчитывает поисковая программа.

Коэффициенты совпадения для всех изученных модельных соединений при прямом и обратном поиске находились в диапазоне от 830 до 960 в зависимости от соединения, что соответствует хорошему и отличному совпадению.

Для определения пределов обнаружения последовательно проводили анализ модельных смесей, полученных разбавлением исходных в 1000, 100 и 10 раз (в МТБЭ). Каждый анализ повторяли 3 раза. При переходе от смеси к смеси осуществляли холостой опыт (анализ чистого растворителя). Значения пределов обнаружения, рассчитанные по масс-хроматограммам, приведены в таблице 1.

Таблица 1. Пределы обнаружения определяемых веществ (регистрацию масс-спектров проводили в режиме полного ионного тока, расчёт проводили по масс-хроматограммам наиболее интенсивных ионов).

Вещество	Предел обнаружения, г/мкл	m/z , наиболее интенсивного иона
Деканол-1	$8,5 \times 10^{-11}$	70
Ундеканол-1	$7,6 \times 10^{-12}$	83
Додеканол-1	$3,7 \times 10^{-11}$	69

Тетрадеканол-1	$3,2 \times 10^{-11}$	83
Додекановая кислота	$6,0 \times 10^{-10}$	157
Тетрадекановая кислота	$5,5 \times 10^{-10}$	73
Гексадекановая кислота	$1,4 \times 10^{-10}$	60
Додекан	$1,6 \times 10^{-11}$	57
Докозан	$3,8 \times 10^{-10}$	57
n-Цимен	$7,2 \times 10^{-11}$	119
Дифенилметан	$7,0 \times 10^{-11}$	168
Нафталин	$1,1 \times 10^{-11}$	128
Гептанон-2	$1,5 \times 10^{-11}$	58
Гептанон-4	$1,0 \times 10^{-11}$	71
5-Метилгексанон-2	$5,9 \times 10^{-12}$	58
5-Метилгептанон-3	$1,3 \times 10^{-11}$	99
2,6-Диметилгептанон-4	$8,2 \times 10^{-12}$	85
Нонанон-2	$2,2 \times 10^{-11}$	58
Додеканон-2	$4,5 \times 10^{-11}$	58
Октаналь	$1,7 \times 10^{-11}$	57
Нонаналь	$2,5 \times 10^{-11}$	57
Деканаль	$1,1 \times 10^{-10}$	57
Ундеканаль	$2,8 \times 10^{-10}$	57
Додеканаль	$4,9 \times 10^{-10}$	57

Как видно из таблицы 1, пределы обнаружения сильно различаются для веществ разных классов. Наименьшие значения пределов обнаружения у кетонов, наибольшие — у кислот.

Исследование возможности снижения пределов обнаружения исследуемых модельных соединений за счёт ввода больших по объёму проб off-line с предварительным удалением растворителя в присутствии сорбента вне аналитической системы

Первым этапом этого исследования был выбор сорбента, обеспечивающего сорбцию и термодесорбцию ультраследовых количеств модельных среднелетучих соединений различной полярности. Были

изучены два варианта: сорбция на СКВ (сверхтонкое кварцевое волокно) и на СКВ в комбинации с сорбентом Tenax GC. В обоих случаях проводилось сравнение результатов анализа модельной смеси (17 компонентов, от $9,5 \times 10^{-11}$ до $7,7 \times 10^{-9}$ г/мкл) при стандартном вводе 1 мкл раствора в инжектор и при нанесении 1 мкл того же раствора на сорбент с последующей термодесорбцией. Оптимальные условия переноса пробы с сорбента в инжектор хроматографа путём термодесорбции были следующие: поток газа-носителя (2 мл/мин), температура десорбции — 250°C , продолжительность термодесорбции — 5 мин; температура термостата хроматографа при термодесорбции — 40°C .

В результате проведённого эксперимента были выбраны оптимальные значения параметров десорбции веществ модельной смеси. Температура десорбции — 300°C , скорость потока газа-носителя — 1,2 мл/мин, продолжительность — 3 мин, температура колонки хроматографа — 35°C . Масса сорбента Tenax GC — 2 мг, масса СКВ — 1 мг.

Применение комбинированного сорбента СКВ/Tenax GC позволило успешно провести сорбцию/десорбцию всех 17 компонентов модельной смеси. Отношения средних значений аналитического сигнала (площадь пика на хроматограмме) компонентов модельной смеси в случае с термодесорбцией с комбинированного сорбента к соответствующим значениям аналитического сигнала в случае стандартного ввода 1 мкл в инжектор шприцем, при одинаковом количестве определяемых веществ, находились в диапазоне от 54 до 100%.

Следующим этапом было изучение возможности ввода больших проб органических растворов модельных соединений в хромато-масс-спектрометр с предварительным удалением растворителя вне аналитической системы. Концентрирование проводили путём сорбции определяемых веществ с одновременным удалением растворителя в потоке парогазовой смеси, образующейся в результате одновременной подачи потоков газа и анализируемого раствора в картридж с сорбентом. Были определены площади пиков веществ при стандартном вводе 1 мкл раствора 24-х компонентной модельной смеси ($5,1 \times 10^{-10}$ – $9,2 \times 10^{-9}$ г/мкл) в инжектор, при вводе 1 мкл этой же смеси с использованием картриджа путём сорбции/десорбции и при

вводе 100 мкл 24-х компонентной модельной смеси ($5,1 \times 10^{-12}$ – $9,2 \times 10^{-11}$ г/мкл) путём концентрирования аналитов в картридже с сорбцией с одновременным удалением растворителя и последующей термодесорбции концентрата в инжектор хроматографа.

Проведённые эксперименты показали, что для летучих соединений, таких как гептанон-2, гептанон-4, 5-метилгексанон-3, 5-метилгептанон-3 и п-цимен отношение аналитического сигнала при вводе больших проб органического раствора модельных соединений к аналитическому сигналу при вводе 1 мкл стандартным способом в инжектор было менее 10% (при условии, что пробы содержат одинаковые количества компонентов). В дальнейшем было решено исключить данные соединения из рассмотрения (при изучении результатов ввода больших проб), так как разрабатываемый нами способ предназначен, прежде всего, для определения среднелетучих органических соединений. Для остальных веществ данный показатель полноты переноса с сорбента в инжектор составил в среднем 60%.

Таким образом, в результате проведённых исследований был разработан способ ГХ/МС определения ультраследовых содержаний среднелетучих органических соединений различной полярности в органических растворах, основанный на удалении растворителя из больших проб растворов в присутствии сорбента, термодесорбции и ГХ/МС анализе всего концентрата аналитов. Пределы обнаружения способа составляли 10^{-8} – 10^{-6} %.

Следующим этапом нашего исследования являлось изучение возможности концентрирования модельных веществ из больших проб водных растворов.

Были проведёны эксперименты по экстракции веществ, входящих в состав модельных смесей (в диапазоне концентраций 3×10^{-9} – 1×10^{-7} г/мкл), из водных растворов метилтретбутиловым эфиром с целью установить степени извлечения всех экстрагируемых соединений (Таблица 2). Соотношение водной и органической фаз при экстракции было 3/1. Высаливание проводили хлоридом натрия (0,1 от массы водной фазы).

Как видно из таблицы 2, степени извлечения в большинстве случаев превышали 80%. Степени извлечения наиболее полярных ве-

Таблица 2. Степени извлечения веществ модельной смеси при жидкость-жидкостной экстракции из воды с предварительным высаливанием ($n = 3$, $P = 0,95$).

Соединение	Степень извлечения при экстракции
Деканол-1	$41 \pm 3\%$
Ундеканол-1	$33 \pm 2\%$
Додеканол-1	$59 \pm 3\%$
Тетрадеканол-1	$52 \pm 6\%$
Додекановая кислота	$43 \pm 3\%$
Тетрадекановая кислота	$40 \pm 2\%$
Гексадекановая кислота	$63 \pm 6\%$
Додекан	$100 \pm 8\%$
Докозан	$100 \pm 9\%$
Дифенилметан	$95 \pm 4\%$
Нафталин	$100 \pm 5\%$
2,6-Диметилгептанон-4	$81 \pm 4\%$
Нонанон-2	$82 \pm 5\%$
Додеканон-2	$100 \pm 8\%$
Октаналь	$83 \pm 5\%$
Нонаналь	$82 \pm 4\%$
Деканаль	$89 \pm 7\%$
Ундеканаль	$100 \pm 8\%$
Додеканаль	$90 \pm 5\%$

ществ составили 30–60%, что можно объяснить их растворимостью в воде.

Проведённые исследования по вводу больших проб органических растворов и по экстракции среднелетучих органических соединений различной полярности из водных растворов позволили перейти к изучению возможности анализа больших проб экстрактов из водных растворов.

В таблице 3 представлены результаты ввода больших проб органических растворов и экстрактов из водных растворов. Из данных, приведённых в таблице, видно, что при использовании разработан-

Таблица 3. Отношения значений аналитического сигнала компонентов модельной смеси в случаях с термодесорбцией (1 и 100 мкл модельного органического раствора и 100 мкл экстракта из водного раствора) с комбинированного сорбента (S_1) к соответствующим значениям аналитического сигнала в случае стандартного ввода 1 мкл (S_2) при одинаковом количестве определяемых веществ ($n = 3$, $P = 0,95$).

Соединение	Отношение S_1/S_2 , %		
	1 мкл	100 мкл	
		Орг. раствор	Экстракт
Деканол	78 ± 13%	60 ± 30%	40 ± 20%
Ундеканол	79 ± 4%	70 ± 20%	40 ± 15%
Додеканол	79 ± 5%	61 ± 17%	50 ± 30%
Тетрадеканол	79 ± 8%	77 ± 18%	40 ± 20%
Додекановая кислота	56 ± 7%	62 ± 15%	30 ± 20%
Тетрадекановая кислота	62 ± 19%	73 ± 9%	40 ± 20%
Гексадекановая кислота	88 ± 40%	100 ± 30%	60 ± 30%
Додекан	75 ± 6%	22 ± 10%	30 ± 10%
Докозан	67 ± 14%	70 ± 30%	70 ± 30%
Дифенилметан	69 ± 6%	50 ± 30%	80 ± 40%
Нафталин	69 ± 5%	44 ± 14%	50 ± 20%
2,6-диметил-4-гептанон	77 ± 6%	20 ± 10%	20 ± 7%
Нонанон-2	74 ± 4%	40 ± 11%	70 ± 20%
Додеканон-2	76 ± 5%	70 ± 30%	70 ± 20%
Октаналь	76 ± 6%	18 ± 4%	30 ± 15%
Нонаналь	74 ± 5%	37 ± 9%	50 ± 20%
Деканаль	75 ± 5%	57 ± 17%	60 ± 20%
Ундеканаль	87 ± 6%	80 ± 30%	80 ± 30%
Додеканаль	100 ± 15%	100 ± 30%	90 ± 30%

ного способа ввода больших проб происходит увеличение значений аналитического сигнала в 20–90 раз, по сравнению со стандартным вводом 1 мкл в инжектор. Наиболее эффективен данный способ для определения среднелетучих органических соединений. Пределы обна-

ружения разработанного способа составляют $10^{-10} - 10^{-8}\%$ (по массе) в зависимости от определяемого вещества.

Исследование состава среднелетучих примесей, присутствующих в образцах конденсата выдыхаемого воздуха

Было проведено предварительное исследование состава смесей среднелетучих органических веществ в двух образцах КВВ без предварительного концентрирования. Проводили жидкость-жидкостную экстракцию с предварительным высаливанием с последующим вводом 1 мкл экстракта. Затем проводили смыв с контейнера, в который собирали КВВ, тем же количеством МТБЭ, что использовали для экстракции. Хроматограммы экстракта и смыва снимали по 3 раза для каждого образца. Проведённые ГХ/МС анализы экстрактов и смывов с контейнеров показали отсутствие на хроматограммах каких-либо пиков, помимо пиков холостого опыта.

С целью определения неизвестных среднелетучих органических веществ в экстракте из КВВ был проведён анализ двух образцов КВВ с использованием разработанного способа анализа водных растворов (объём пробы 100 мкл). Параллельно с анализом экстракта КВВ проводили анализ образцов смыва растворителем (МТБЭ) с контейнера и чистого растворителя.

После сравнения хроматограмм экстрактов с хроматограммами растворителя было установлено, что в экстракте из КВВ содержится свыше 100 веществ, сигнал которых отсутствует в холостом опыте. Большинство из них не подлежало достоверной идентификации по масс-спектрам, ввиду малого содержания.

Всего по результатам опытов по анализу экстрактов из образцов КВВ было идентифицировано 33 индивидуальных соединения, ещё 11 веществ характеризовались достаточно интенсивным сигналом, стабильным масс-спектром и временем удерживания, однако идентификация по библиотеке масс-спектров оказалась невозможной.

Результаты прямого и обратного поиска для идентифицированных в КВВ веществ по библиотеке масс-спектров NIST приведены в таблице 4.

Таблица 4. Результаты прямого и обратного поиска при идентификации среднелетучих органических веществ по библиотеке масс-спектров.

Название вещества	Результаты сравнения с масс-спектрами из библиотеки NIST-2005	
	Прямой поиск	Обратный поиск
Фенол	906	912
Гексановая кислота	805	861
Бензиловый спирт	939	946
Гептановая кислота	766	825
2-этилгексановая кислота	821	852
Нонаналь	671	815
Ментол	919	920
Октановая кислота	855	889
Деканол-1	796	828
Деканаль	813	847
2-феноксиэтанол	905	927
Хинолин	682	876
Нонановая кислота	842	843
Индол	801	871
Никотин	821	854
Декановая кислота	779	806
2,3-дигидро-1-Н-инден-1-он	574	805
Бутоксизтоксизтилацетат	714	815
Циклододекан	898	913
3-гидрокси-2,4,4-триметилпентиловый эфир 2-метилпропановой кислоты	665	739
Ундекановая кислота	774	770
Додекановая кислота	837	899
Триацетин	781	894
Бензофенон	792	892
Тетрадекановая кислота	726	776
Трибутилфосфат	720	853
Метил дигидрожасмонат	798	846

Пентадекановая кислота	670	800
Этил цитрат	898	903
Гексадекановая кислота	851	856
Октадекановая кислота	866	869
Этилен брассилат	816	850
7,9-ди-tert-бутил-1-окса- спиро(4,5)дека-6,9-диен- 2,8-дион	869	875

Два образца экстрактов из КВВ были проанализированы методом ГХ/МС в режиме ХИ с регистрацией положительных ионов. Объём анализируемой пробы — 100 мкл. Были установлены значения отношений массы к заряду для основных ионов в масс-спектрах химической ионизации для 22-х веществ, идентифицированных в экстракте из КВВ. Сравнение полученных значений m/z и значений молекулярных масс подтверждает правильность идентификации данных соединений по их масс-спектрам электронной ионизации, так как в условиях химической ионизации изобутаном этих веществ согласно теории преимущественно образуются псевдомолекулярные ионы $[M+H]^+$.

Исследование состава конденсата выдыхаемого воздуха здоровых, больных хронической обструктивной болезнью лёгких и больных бронхиальной астмой

Были собраны образцы КВВ здоровых людей (30 проб), больных хронической обструктивной болезнью лёгких (20 проб) и больных бронхиальной астмой (20 проб). Каждый образец проанализировали с использованием разработанного способа ГХ/МС анализа водных растворов. По изучении хроматограмм было выделено 40 веществ (из числа тех 44-х, что были обнаружены в КВВ).

Было произведено сравнение частот обнаружения того или иного вещества в КВВ и диапазонов значений аналитического сигнала по каждому веществу по трём группам: здоровые, больные ХОБЛ, больные бронхиальной астмой.

Большинство веществ из выделенного набора присутствовало практически во всех образцах КВВ каждой группы. Исключение составляют деканол-1, 2-феноксиэтанол, никотин, циклододекан, 3-гидрокси-2,4,4-триметилпентиловый эфир 2-метилпропановой кислоты, ундекановая кислота. В случае с 2-феноксиэтанолом заметно резкое снижение частоты его появления в КВВ, собранном у больных ХОБЛ, по сравнению с больными бронхиальной астмой и здоровыми. Также следует отметить, что присутствие деканола-1 в образцах КВВ, собранных у больных бронхиальной астмой, наблюдается заметно реже, чем у здоровых и больных ХОБЛ.

Исходя из диапазонов значений аналитического сигнала, можно сделать вывод о том, что концентрация каждого вещества от образца к образцу меняется в пределах 1–3-х порядков. Наименьшие диапазоны значений аналитического сигнала по всем трём группам у деканола, циклододекана, и трибутилфосфата. Наибольшие диапазоны — у фенола, индола и этил цитрата.

Наиболее представительную группу, среди обнаруженных в КВВ соединений, составляют насыщенные жирные кислоты (12 штук).

После обработки результатов анализа 70 образцов КВВ, собранных у здоровых, больных ХОБЛ и больных бронхиальной астмой, с использованием специально разработанного алгоритма, основанного на линейных методах теории распознавания образов, было установлено, что на основании полученных данных можно различить здоровых и больных бронхиальной астмой с надёжностью 75%, здоровых и больных ХОБЛ — с надёжностью 85%, больных бронхиальной астмой и ХОБЛ — с надёжностью 83%. Из 40 соединений, по которым проводилось сравнение, 9 имеют наибольший информационный вес. В случае «здоровые — ХОБЛ» это были два неидентифицированных вещества с известными временами удерживания и масс-спектрами (ЭИ), этил цитрат и 2,3-дигидро-1-Н-инден-1-он. В случае «здоровые — бронхиальная астма» — другие три из неидентифицированных веществ и деканол-1. В случае «бронхиальная астма — ХОБЛ» — одно неидентифицированное вещество, деканол-1 и 2-феноксиэтанол. Эти вещества могут быть потенциальными маркерами рассмотренных заболеваний.

Данные, химического анализа были организованы в таблицы следующего вида (табл. 5)¹.

Таблица 5. Структура таблиц данных. В данном случае $n(z) = 60$; $n(a) = n(b) = 40$, $s = 40$.

Z	P₁	P₂	...	P_S
Z₁	$z_{1,1}$	$z_{1,2}$...	$z_{1,s}$
Z₂	$z_{2,1}$	$z_{2,2}$...	$z_{2,s}$
...
Z_{n(z)}	$z_{n(z),1}$	$z_{n(z),2}$...	$z_{n(z),s}$

B	P₁	P₂	...	P_S
B₁	$b_{1,1}$	$b_{1,2}$...	$b_{1,s}$
B₂	$b_{2,1}$	$b_{2,2}$...	$b_{2,s}$
...
B_{n(b)}	$b_{n(b),1}$	$b_{n(b),2}$...	$b_{n(b),s}$

A	P₁	P₂	...	P_S
A₁	$a_{1,1}$	$a_{1,2}$...	$a_{1,s}$
A₂	$a_{2,1}$	$a_{2,2}$...	$a_{2,s}$
...
A_{n(a)}	$a_{n(a),1}$	$a_{n(a),2}$...	$a_{n(a),s}$

3. Математический метод распознавания (диагностики)

Итак, имеется три группы прямоугольных таблиц, которые обозначены Z, A, B . Первая группа содержала $2 \times 30 = 60$ строк, вторая — $2 \times 20 = 40$ строк, третья — $2 \times 20 = 40$ строк, число столбцов (признаков) в каждой таблице было равно 40 (по количеству веществ), элементами таблицы являются неотрицательные действительные числа, равные концентрации соответствующего вещества в отдельных образцах. Исследовались пары таблиц (Z, A) , (Z, B) , (A, B) . Для каждой

¹Сами данные приводятся в приложении № 1.

пары таблиц искались наборы признаков и такая линейная комбинация их числовых значений, что на строчках из первой таблицы каждой пары по соответствующим признакам получалось положительное число, а на строчках второй таблицы — отрицательное число; причем суммарное число, в которых не получался требуемый знак, то есть строк из первой таблицы, на которых линейная комбинация отрицательна, и строк из второй таблицы, на которых она положительна, должно быть минимально. Поиск линейной комбинации соответствует прохождению разделяющей плоскости в трехмерном пространстве через начало координат. Такое требование было связано с тем, что концентрации веществ, полученные в результате эксперимента были известны только в относительном (а не абсолютном выражении), то есть предоставленные данные представлены были известны с точностью до пропорциональности. Таким образом, точки соответствующие строкам первой таблицы пары, попадают в положительное полупространство (относительно найденной гиперплоскости), а точки второй таблицы — в отрицательное полупространство.

Условия на минимизацию числа признаков, по которым производилась бы диагностика (не более 3–4), ограничение на качество распознавания и возможности по быстройдействию вычислительных средств привели к тому, что были использованы выборки по трем признакам. Выборки по двум признакам давали неудовлетворительную точность распознавания (около 60%), а выборки по 4 и более признакам потребовали бы чрезмерно больших затрат машинного времени и, в случае применения на практике, увеличили бы стоимость проведения соответствующих анализов и сложность процедуры диагностики.

4. Формальная математическая постановка задачи

Представим строки таблиц как вектора s -мерного евклидова пространства. Даны множества векторов, обозначим их Z, A и $B \subset P^s$. Будем обозначать $\Pi_{k,l,m}(x_1, \dots, x_40) = (x_k, x_l, x_m)$ — проек-

ция вектора (x_1, \dots, x_s) на подпространство, образованное векторами e_k, e_l и e_m .

Аналогично, обозначим проекцию множества M как $\Pi_{k,l,m}(M) = \{\Pi_{k,l,m}(m) | m \in M\}$.

Мощность множества M будем обозначать $|M|$.

Было рассмотрено 3 случая.

- 1) $M_1 = Z, M_2 = B$;
- 2) $M_1 = Z, M_2 = A$;
- 3) $M_1 = B, M_2 = A$.

Для каждого случая решалась следующая задача:

Пусть заданы множества M_1 и M_2 , мощности которых равны $n_1 = |M_1|$ и $n_2 = |M_2|$. Будем рассматривать всевозможные их проекции $M_1 = M_1(k, l, m) = \Pi_{k,l,m}(M_1)$ и $M_2 = M_2(k, l, m) = \Pi_{k,l,m}(M_2)$. Для каждой построим «почти» отделяющую гиперплоскость. Дело в том, что множества могут не быть 0-линейно отделимы (как это было в случае численных данных, полученных из эксперимента), таким образом, рассматривались всевозможные гиперплоскости вида $a_1x_1 + a_2x_2 + a_3x_3 = 0$. Подсчитывались количества неправильно распознанных векторов. Пусть $E_1 = |\{(x_1, x_2, x_3) \in M_1 | a_1x_1 + a_2x_2 + a_3x_3 \leq 0\}|$ и $E_2 = |\{(x_1, x_2, x_3) \in M_2 | a_1x_1 + a_2x_2 + a_3x_3 \leq 0\}|$ — число ошибок распознавания первого и второго множеств. Далее находилась минимум $U_{k,l,m} = \min_{a_1, a_2, a_3} (k_1E_1 + k_2E_2)$, где весовые коэффициенты (в данном случае) выбирались обратно пропорционально мощности соответствующего множества, $k_i = 1/n_i, i = 1, 2$.

Далее находилась минимум $U = U(M_1, M_2) = \min_{1 \leq k < l < m \leq 40} U_{k,l,m}$.

Полученные в результате численного эксперимента величины приводятся в приложении № 2.

Далее рассматривался вопрос об «устойчивости» полученных решений.

Для этого фиксировалась пара чисел (q_1, q_2) , случайным образом выбирались подмножества $M_1^{(i)} \subset M_1$ и $M_2^{(i)} \subset M_2, i = 1, 2, \dots, N$ такие, что $|M_1^{(i)}| = |M_1| - q_1$ и $|M_2^{(i)}| = |M_2| - q_2$, и для них решалась та же задача минимизации и находилось $U(M_1^{(i)}, M_2^{(i)})$, после чего

находилось среднее значение $U_{\text{ср.}}(q_1, q_2)$ по $N = 200$ случайным парам $(M_1^{(i)}, M_2^{(i)})$.

Кроме того, подсчитывались частоты, с которыми различные параметры участвуют в оптимальном решении.

Более формально, для каждой пары $M_1^{(i)}, M_2^{(i)}$ строился $w(M_1^{(i)}, M_2^{(i)}) = (w_1, \dots, w_s)$, следующим образом:

$$w_j = \begin{cases} 1, & \text{при } j = k_{\min}, l_{\min}, m_{\min}, \\ 0, & \text{иначе,} \end{cases}$$

где $k_{\min}, l_{\min}, m_{\min}$ такие, что $U(M_1^{(i)}, M_2^{(i)}) = U_{k_{\min}, l_{\min}, m_{\min}}$.

Далее, вектора $w(M_1^{(i)}, M_2^{(i)})$ усреднялись по $(M_1^{(i)}, M_2^{(i)})$ и делились на 3 (для нормировки). Полученные величины — вектор информационные веса признаков — отражают (в некоторой степени) значимость данного параметра для решения задачи разделения множеств. Таблица усредненных векторов приводится в приложении № 3.

Полученные разделяющие плоскости могут быть использованы в дальнейшем для выработки рекомендаций для диагностирования заболеваний на основе измеренных концентраций.

Пример инструкции содержится в приложении 4.

Полученные результаты показывают:

1. Задача разделения может быть решена с точностью:
 - а) на паре «Здоровые — Астма» правильное распознавание — 75%;
 - б) на паре «Здоровые — ХОБЛ» правильное распознавание — 85%;
 - в) на паре «ХОБЛ — Астма» правильное распознавание — 84%.
2. Лишь ограниченное количество параметров представляет интерес с точки зрения решения каждой задачи. При решении задачи а) основной информационный вклад дают параметры №№ 2, 6, 1 и 25. При решении задачи б) основной информационный вклад дают параметры №№ 8, 10, 4 и 31. При решении задачи в) основной информационный вклад дают параметры №№ 27, 25 и 1. Таким образом значения 9 параметров (из 40) позволяют решить задачу постановки диагноза.
3. Как показывает численный эксперимент, при уменьшении выборки происходит постепенное уменьшение точности распознавания. Представляется вероятным, что при увеличении объема выборки и точно-

сти или достоверности данных возможно увеличение точности распознавания.

5. Заключение

Проведено исследование возможности выделения в присутствии сорбента из больших проб растворов ультраследовых количеств (нг-пг) соединений, кипящих в широком диапазоне температур, принадлежащих к различным классам, перевода полученного концентрата в инжектор хроматографа, за счёт термодесорбции, и последующего ГХ/МС анализа всего концентрата.

На основании этого исследования разработан способ определения ультраследовых содержаний среднелетучих органических соединений, принадлежащих к различным классам, в органических растворах, основанный на выделении этих соединений в присутствии сорбента из больших проб растворов (100 мкл), переводе адсорбата за счёт термодесорбции в инжектор хроматографа и ГХ/МС анализе всего концентрата.

Проведено исследование жидкостной экстракции метилтретбутиловым эфиром изученных модельных соединений из водных растворов с предварительным высаливанием хлоридом натрия.

Разработан способ определения среднелетучих соединений различной полярности в водных растворах, основанный на жидкостной экстракции с высаливанием, сорбционным выделении примесей из большой части экстракта, переводе их за счёт термодесорбции в инжектор и ГХ/МС анализе всего концентрата.

С использованием разработанного способа анализа водных растворов исследован состав смесей органических соединений в 70-ти образцах КВВ, собранных у больных ХОБЛ, бронхиальной астмой и здоровых людей. В результате математической обработки полученных данных показана возможность различать с высокой надёжностью здоровых, больных ХОБЛ и больных бронхиальной астмой между собой. Полученные результаты открывают широкие перспективы для дальнейших исследований.

Приложения

Приложение № 1. Таблицы исходных данных (фрагмент).

Номер	Пол	Возраст	ИКЧ	Диагноз	дата сбора конденсата	объем образца, мл	1	2	3	4	5
51	Ж	48	27	астма	06.02.2007	0,82	7,62E+05	2,62E+06	4,52E+06	2,40E+05	1,47E+05
51	Ж	48	27	астма	06.02.2007	0,82	1,39E+05	4,28E+05	6,23E+05	5,68E+04	2,29E+06
52	Ж	56		астма	06.02.2007	2,5	5,65E+05	1,60E+06	3,61E+06	1,49E+05	1,01E+06
52	Ж	56		астма	06.02.2007	2,5	1,04E+06	2,94E+06	5,49E+06	2,36E+05	1,42E+06
53	Ж	61		астма	08.02.2007	1,15	3,98E+04	1,17E+05	2,70E+05	5,99E+04	1,09E+05
53	Ж	61		астма	08.02.2007	1,15	2,76E+04	7,39E+04	1,23E+05	2,63E+04	3,64E+04
54	Ж	51	15	астма	08.02.2007	1,7	7,10E+04	1,90E+05	4,43E+05	6,82E+04	1,65E+05
54	Ж	51	15	астма	08.02.2007	1,7	1,04E+05	2,75E+05	4,19E+05	4,89E+04	1,30E+05
55	Ж	47	5	астма	08.02.2007	1,7	5,80E+04	1,78E+05	3,73E+05	4,20E+04	1,39E+05
55	Ж	47	5	астма	08.02.2007	1,7	8,30E+04	2,76E+05	5,26E+05	4,67E+04	2,13E+05
56	Ж	57		астма	08.02.2007	1,55	1,05E+05	2,02E+05	2,93E+05	3,10E+04	8,18E+04
56	Ж	57		астма	08.02.2007	1,55	3,32E+04	8,16E+04	1,18E+05	3,12E+04	4,60E+04
57	Ж	49	30	астма	08.05.2007	1,9	5,09E+04	1,32E+05	3,00E+05	5,13E+04	1,41E+05
57	Ж	49	30	астма	08.05.2007	1,9	1,34E+04	4,52E+04	2,13E+05	4,79E+04	1,20E+05
58	Ж	51		астма	10.05.2007	1,5	3,80E+04	1,32E+05	4,54E+05	5,73E+04	3,17E+05
58	Ж	51		астма	10.05.2007	1,5	7,91E+04	2,32E+05	4,21E+05	7,41E+04	2,68E+05
59	М	51		астма	10.05.2007	2	1,35E+05	4,34E+05	1,04E+06	6,23E+04	2,70E+05
59	М	51		астма	10.05.2007	2	6,71E+04	3,04E+05	3,97E+05	7,00E+04	1,95E+05
60	Ж	73		астма	15.05.2007	1,25	1,17E+05	1,50E+05	3,57E+05	6,68E+04	1,97E+05
60	Ж	73		астма	15.05.2007	1,25	1,90E+05	2,54E+05	5,62E+05	8,65E+04	3,42E+05
61	Ж	67		астма	22.05.2007	1,1	6,51E+04	1,04E+05	2,19E+05	3,84E+04	1,23E+05
61	Ж	67		астма	22.05.2007	1,1	5,29E+04	7,16E+04	2,27E+05	5,47E+04	8,73E+04
62	Ж	51	5	астма	22.05.2007	2	1,06E+05	1,64E+05	6,12E+05	5,80E+04	2,84E+05

Приложение № 2. Разделяющие гиперплоскости².

Число ошибок I рода		коэффициенты г/п			Номера образцов на которых возникает ошибка									
Число ошибок II рода	базис	a1	a2	a3	I рода			II рода						
7	6 3,7,9	0.34	-0.94	0.03	4 5 26 27 36 52 53	II рода								
8	6 5,7,9	0.34	-0.94	0.03	4 5 17 26 27 36 52 53	II рода								
10	6 7,9,30	-0.72	0.02	0.69	4 5 26 27 29 36 40 52 53 58	II рода								
10	6 7,9,31	-0.98	0.04	0.19	0 1 4 5 17 26 27 36 52 53	II рода								
11	6 7,9,33	-0.98	0.04	0.19	0 4 5 10 16 17 26 27 36 52 53	II рода								
7	6 7,9,34	-0.96	0.04	0.28	0 4 5 26 27 52 53	II рода								
11	7 7,19,26	-0.98	0.02	0.20	22 23 26 27 28 29 52 56 57 58 59	II рода								
Здоровые vs Астма														
Число ошибок I рода	базис	a1	a2	a3	I рода			II рода						
13	10 2,22,25	-0.84	0.22	-0.50	20 22 23 28 29 40 41 42 52 53 57 58 59	II рода								
14	9 10,5,24	-0.66	0.52	0.54	17 24 30 31 32 33 34 35 36 37 39 40 41 52	II рода								
14	9 1,32,33	-0.26	-0.10	0.96	6 7 8 9 20 21 30 31 36 37 56 57 58 59	II рода								
14	9 1,33,35	-0.60	0.62	0.51	4 5 20 21 22 23 26 27 28 29 32 33 36 41	II рода								
14	10 1,6,32	-0.64	0.76	-0.11	2 4 5 20 21 22 23 26 27 32 33 37 40 48	II рода								
14	10 1,27,33	-0.22	-0.22	0.95	6 7 8 9 36 40 41 42 50 54 56 57 58 59	II рода								
14	10 9,12,22	0.22	-0.82	0.53	10 20 34 40 42 43 45 48 52 53 56 57 58 59	II рода								

²Ошибка 1 рода — принять здорового за больного. 2-го рода — наоборот.

Приложение № 3. Информационные веса признаков.

Информационные веса					
Здоровые - ХОБЛ		Здоровые-Астма			
№	вес	№	вес	№	вес
1	0.006667	1	0.146667	1	0.153333
2	0.013333	2	0.216667	2	0.003333
3	0.000000	3	0.020000	3	0.000000
4	0.093333	4	0.063333	4	0.000000
5	0.010000	5	0.003333	5	0.003333
6	0.016667	6	0.200000	6	0.046667
7	0.023333	7	0.046667	7	0.026667
8	0.326667	8	0.000000	8	0.036667
9	0.000000	9	0.000000	9	0.000000
10	0.303333	10	0.013333	10	0.000000
11	0.003333	11	0.000000	11	0.003333
12	0.000000	12	0.010000	12	0.013333
13	0.000000	13	0.023333	13	0.000000
14	0.000000	14	0.000000	14	0.000000
15	0.000000	15	0.000000	15	0.000000
16	0.000000	16	0.000000	16	0.000000
17	0.000000	17	0.000000	17	0.000000
18	0.000000	18	0.000000	18	0.000000
19	0.026667	19	0.000000	19	0.000000
20	0.000000	20	0.000000	20	0.000000
21	0.006667	21	0.000000	21	0.000000
22	0.003333	22	0.000000	22	0.000000
23	0.000000	23	0.050000	23	0.000000
24	0.003333	24	0.000000	24	0.000000
25	0.006667	25	0.100000	25	0.330000
26	0.003333	26	0.010000	26	0.000000
27	0.003333	27	0.013333	27	0.333333
28	0.000000	28	0.000000	28	0.013333
29	0.000000	29	0.000000	29	0.000000
30	0.000000	30	0.000000	30	0.000000
31	0.090000	31	0.000000	31	0.000000
32	0.020000	32	0.020000	32	0.036667
33	0.000000	33	0.006667	33	0.000000
34	0.003333	34	0.020000	34	0.000000
35	0.006667	35	0.000000	35	0.000000
36	0.013333	36	0.003333	36	0.000000
37	0.000000	37	0.000000	37	0.000000
38	0.003333	38	0.000000	38	0.000000
39	0.013333	39	0.000000	39	0.000000
40	0.000000	40	0.033333	40	0.000000

Приложение № 4. Образец инструкции.

И Н С Т Р У К Ц И Я
по диагностике астмы и бронхитаПравило № 1 (диагностика астмы)

Пациент подозревается на заболевание астмой, если выполняется хотя бы одно из следующих условий:

а) $3,5 \cdot P_3 + 2,08 \cdot P_{35} > P_{23}$;

б) $0,79 \cdot P_6 + 0,82 \cdot P_{25} < P_1$.

Вероятность получения правильного результата составляет не менее 77% (на имеющихся данных).

Правило № 2 (диагностика бронхита)

Пациент подозревается на заболевание ХОБЛ, если выполняется хотя бы одно из следующих условий:

а) $0,04 \cdot P_{10} + 0,29 \cdot P_{35} < P_8$;

б) $0,36 \cdot P_4 + 0,03 \cdot P_{10} < P_8$.

Вероятность получения правильного результата составляет не менее 87% (на имеющихся данных).

Если после применения правил № 1 и № 2 установлено, что пациент болен необходимо подтвердить результат.

Правило № 3 (дифференциация астмы от бронхита при подозрении на их наличие)

Пациент подозревается на заболевание астмой, если выполняется следующее условие: $1,71 \cdot P_1 + 14,14 \cdot P_{25} < P_{27}$.

Пациент подозревается на заболевание бронхитом, если выполняется следующее условие: $1,71 \cdot P_1 + 14,14 \cdot P_{25} > P_{27}$.

Вероятность получения правильного результата составляет не менее 86% (на имеющихся данных).

Примечание. Метод, на основании которого разработана настоящая Инструкция, не позволяет качественно диагностировать наличие легочных заболеваний без дифференциации астмы и бронхита.

Все параметры измеряются в единицах указанных в экспертном материале.

