

На правах рукописи

ЖВАНСКИЙ
Евгений Сергеевич

**МАСС-СПЕКТРОМЕТРИЯ В ЗАДАЧЕ ПРОФИЛИРОВАНИЯ
БИОЛОГИЧЕСКИХ ТКАНЕЙ ЧЕЛОВЕКА ДЛЯ ИДЕНТИФИКАЦИИ
ОПУХОЛЕЙ ГОЛОВНОГО МОЗГА**

01.04.01 – Приборы и методы экспериментальной физики

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва 2020

Работа прошла апробацию в департаменте молекулярной и биологической физики федерального государственного автономного образовательного учреждения высшего образования “Московский физико-технический институт (национальный исследовательский университет)” (МФТИ)

Научный руководитель: кандидат физико-математических наук, доцент,
Попов Игорь Алексеевич

Ведущая организация: Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Московский государственный университет имени
М.В.Ломоносова»

Защита диссертации состоится «25» декабря 2020 года в 13:00 на заседании диссертационного совета ФБМФ.01.04.01.002 по адресу: 141701, Московская обл., г. Долгопрудный, Институтский пер., д. 9.

С диссертацией можно ознакомиться в научно-технической библиотеке МФТИ и на сайте <https://mipt.ru/education/post-graduate/soiskateli-fiziko-matematicheskie-nauki.php>.

Работа представлена 15 октября 2020 года в Аттестационную комиссию федерального государственного автономного образовательного учреждения высшего образования «Московский физико-технический институт (национальный исследовательский университет)» для рассмотрения советом по защите диссертаций на соискание ученой степени кандидата наук, доктора наук в соответствии с п. 3.1 ст. 4 Федерального закона «О науке и государственной научно-технической политике».

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность

В ходе операции по удалению опухоли мозга хирургу важно правильно определить границу опухоли перед резекцией. От успешного решения этой задачи зависит успех операции в терминах полноты резекции, восстановления и послеоперационной выживаемости пациентов.

Стандартные методы определения границы опухоли обладают следующими недостатками: требуют заметного времени, обладают ограниченной чувствительностью, дорогие или плохо встраиваются в процедуру операции. Масс-спектрометрия (МС) является перспективным для интраоперационной диагностики методом. МС позволяет составлять молекулярные профили тканей, в частности липидные профили опухолевых тканей головного мозга.

В клетках здоровой мозговой ткани и опухоли разный метаболизм. Для постоянного роста опухолевых клеток им нужны источники энергии и строительного материала. Исследования показывают, что клетки глиомы для этой цели активно используют жирные кислоты. Небольшая часть этих жирных кислот поступает из крови, однако большая часть синтезируется самими опухолевыми клетками. Жирные кислоты участвуют в формировании липидов для строительства клеточных мембран. Соответственно, в активно растущей опухолевой ткани липидный профиль отличается от профиля здоровой ткани.

МС с использованием прямой атмосферной ионизации — быстрый, стабильный и воспроизводимый метод, допускающий автоматическую обработку данных. Благодаря этим свойствам метод можно адаптировать к клиническим условиям.

Развитие метода для анализа биологических образцов на основе молекулярного профилирования происходит одновременно в нескольких направлениях:

- оптимизация параметров и конструкции ионных источников для сокращения времени исследования (от десятков минут в стандартной процедуре до секунд при использовании МС)
- уменьшение объема анализируемого образца, требуемого для исследования
- повышение точности идентификации типа патологических изменений
- снижение зависимости результата идентификации от человеческого фактора за счет автоматизации обработки результатов

В данной работе уделяется внимание всем обозначенным направлениям.

Помимо хирургии, быстрый и воспроизводимый метод исследования образцов ткани человека может быть востребован в медицинской диагностике для выявления болезней на ранней стадии, или в терапии для уточнения диагноза и наблюдения за прогрессом лечения.

Цель работы

Целями настоящей работы было:

- Комплексное исследование взаимосвязи между наблюдаемым масс-спектрометрическим профилем образца ткани опухоли мозга человека и физико-химическими процессами, происходящими при экстракции биомолекул из данного образца в комбинированном ионном источнике на основе электрораспылительной ионизации. Демонстрация возможности использования такого источника для получения молекулярного профиля биологических образцов.
- Исследование возможностей применения подходов мультипараметрической классификации данных масс-спектрометрического профилирования к задачам определения границ опухоли во время нейрохирургической операции и классификации опухолевых тканей, а также создание собственных подходов и алгоритмов, оптимизированных для решения задач идентификации патологических изменений в тканях головного мозга.

Основные задачи исследования

Для достижения основной цели исследования были поставлены следующие задачи:

1. Исследование процессов микроэкстракции в потоке жидкости и создание на их основе атмосферного ионного источника, позволяющего проводить экспресс-анализ образцов опухолей головного мозга для их последующей идентификации.
2. Исследование стабильности масс-спектрометрических данных, получаемых при исследовании образцов опухолей головного мозга при помощи прямых методов ионизации и создание системы контроля качества измерений.
3. Поиск оптимального подхода к предобработке сигналов для получения стабильных, информативных и воспроизводимых данных для длительных масс-спектрометрических измерений, состоящих из большого числа последовательных сканов, при соблюдении условия обеспечения оптимальных технических параметров экспорта данных — размера сохраняемых данных и времени предобработки данных, не превышающего времени эксперимента.
4. Накопление базы данных для решения задач определения типа патологических изменений в головном мозге, включая:
 - а. Сбор масс-спектрометрических профилей образцов тканей опухолей головного мозга, различных по типу и локализации, у одного пациента.
 - б. Сбор масс-спектрометрических профилей образцов тканей опухолей головного мозга разных типов от разных пациентов, а также формирование референсных контрольных масс-спектрометрических профилей тканей головного мозга у пациентов с патологическими изменениями головного мозга неопухолевой природы.
5. Разработка алгоритмов классификации масс-спектрометрических профилей, включая:

а. Проверку возможности определения типа тканей, содержащихся в образцах, извлеченных во время нейрохирургической операции по удалению опухоли головного мозга у одного пациента, с учётом того, что образцы тканей могут минимально отличаться друг от друга на молекулярном уровне.

б. Проверку работы классификаторов на расширенной базе данных, включающей образцы многих пациентов.

6. Разработка методов совместного анализа масс-спектров, полученных на разных типах приборов, с разным разрешением. Исследование того, существуют ли параметры, при которых масс-спектры, измеренные при помощи разных масс-анализаторов с различной разрешающей способностью, считать одинаково информативными.

Научная новизна

1. Разработан новый ионный источник, позволяющий осуществлять экспресс-анализ тканей опухолей головного мозга. В предложенном источнике постоянный поток растворителя омывает образец и экстрагирует молекулы для последующего анализа с помощью МС.

2. Впервые показано, что измеренные при помощи нового ионного источника липидные профили разных гистологических типов тканей отличаются друг от друга и могут быть использованы для их идентификации.

3. Разработан новый подход к анализу многомерных (многопризнаковых) данных в условиях недостаточности выборки образцов. Показано, что для классификации можно не учитывать интенсивности пиков, а лишь их присутствие в спектре, что очень важно для использования различных масс-спектрометров, обладающих разными аналитическими характеристиками.

4. Экспериментально показано, что при формировании профилей тканей, содержащих в себе элементы нескольких гистологических типов, выполняется свойство аддитивности — профиль такой ткани можно представить как линейную комбинацию профилей тканей разных классов.

5. Разработан и валидирован новый алгоритм предварительной обработки данных для выделения признаков (ионов) в масс-спектре, позволяющий работать при низких значениях отношения сигнал/шум.

6. На основе исследования различных метрик в матрицах подобия масс-спектров разработан метод оценки стабильности и воспроизводимости масс-спектрометрических данных, полученных в разных экспериментах по общему протоколу в условиях прямой атмосферной ионизации.

7. Предложено унифицированное представление липидных масс-спектрометрических профилей, позволяющее одновременно обрабатывать масс-спектры, полученные при помощи различных устройств или при различных параметрах измерения.

Практическая и теоретическая значимость работы

Результаты работы могут быть использованы для создания автоматизированных измерительных комплексов для прямого масс-спектрометрического профилирования опухолевых тканей. Такие системы могут применяться как альтернативные аналитические методы интраоперационного анализа типов тканей для быстрого, точного и автоматического определения границы опухоли.

Также подход к предобработке, анализу и классификации многомерных данных может быть использован не только в медицине и диагностике, а может быть обобщен и применен для создания систем автоматической обработки данных в других приложениях, где необходим анализ молекулярных профилей образцов. Например, в исследовании растений или продуктов питания.

Основные положения, выносимые на защиту

1. Новый ионный источник для прямого масс-спектрометрического профилирования тканей опухолей головного мозга, позволяющий получать стабильные масс-спектрометрические профили, характеризующие тип ткани, анализировать липидный и метаболический состав исследуемых тканей и осуществлять их идентификацию в целях интраоперационного контроля положения границы опухоли.

2. Липидные профили различных тканей и опухолей могут служить идентификатором типа ткани и опухоли.

3. Алгоритм предобработки данных, позволяющий выделить в масс-спектрометрическом профиле пики (группы пиков), однозначно характеризующие анализируемую ткань, даже если их интенсивность в отдельных масс-спектрах ниже уровня шума, и если наблюдается постепенное отклонение измеренной массы молекул в силу аппаратной ошибки.

4. Метод для оценки качества масс-спектрометрических профилей, получаемых в условиях атмосферной ионизации образцов биологических тканей.

5. Способ представления масс-спектрометрических профилей как многомерных данных и подход к классификации, который может работать в условиях малого количества образцов и позволяет делать вывод о доле клеток с различными патологическими изменениями (некротическая ткань с некротизированными сосудами, некротическая ткань с тенями опухолевых клеток, опухоль с некрозом, опухоль, граница опухоли, некротизированная опухоль, крошки опухолевых клеток).

6. Способ различия здоровой и опухолевой ткани в пределах ошибки среднего для всей базы данных путём классификации масс-спектрометрических профилей базы данных пациентов с различными диагнозами, с использованием сигнатур опухолевой, некротической и здоровой тканей одного пациента.

7. Способ унифицированного представления липидных масс-спектрометрических профилей, который позволяет совместно обрабатывать данные, полученные при помощи разных устройств при различных параметрах измерения.

Публикации по теме диссертации и аprobация результатов:

По материалам диссертации опубликовано 5 печатных статей в журналах, включенных в рекомендованный ВАК РФ список. Основные результаты диссертации были представлены на российских и международных конференциях: I Всероссийская конференция с международным участием «Химический анализ и медицина» (Москва, 2015), VI-ая международная конференция-школа для молодежи «Фундаментальные вопросы масс-спектрометрии и ее аналитические применения» (Долгопрудный, 2016), Second International Conference on Mass Spectrometry (ICMS, Kottayam, Kerala, India, 2015), 68th ASMS Conference on Mass Spectrometry and Allied Topics (Online Meeting, 2020), 67th ASMS Conference on Mass Spectrometry and Allied Topics (Atlanta, 2019), 66th ASMS Conference on Mass Spectrometry and Allied Topics (San Diego, 2018), 65th ASMS Conference on Mass Spectrometry and Allied Topics (Indianapolis, 2017), 64th ASMS Conference on Mass Spectrometry and Allied Topics (San Antonio, USA, 2016), 63rd ASMS Conference on Mass Spectrometry and Allied Topics (St. Louis, USA, 2015), The 25th Intelligent Systems for Molecular Biology and European Conference on Computational Biology (16th Annual Conference, Prague, Czech Republic, 2017), 2018 Congress of Neurological Surgeons (CNS) Annual Meeting (Houston, Texas, USA, 2018), 2019 Congress of Neurological Surgeons (CNS) Annual Meeting (Сан-Франциско, Соединённые Штаты Америки, 2019).

Личный вклад:

Автор внес основной вклад в работу. При его участии были проведены все эксперименты. Автором работы была проведена обработка данных и обобщение результатов.

Содержание работы:

Диссертация изложена на 147 страницах машинописного текста, включает 6 таблиц, 44 рисунка. Работа состоит из введения, обзора литературы, 4 глав описания методов и собственных результатов исследований, каждая из этих глав содержит обсуждение результатов, заключение и выводы. Список цитируемых работ содержит 137 источников.

Введение:

В этой части диссертации рассмотрена актуальность поставленных проблем, сформулированы цели, задачи, научная новизна работы и ее практическая значимость,

Глава 1: Обзор литературы

В первой главе представлен обзор литературы.

В первой части главы описаны методы диагностики и лечения опухолей головного мозга и рассмотрены различные существующие подходы к решению задачи определения границы опухоли во время операции и их недостатки.

В следующей части описаны метаболические пути в тканях мозга и их изменение в процессе развития опухоли опухолей, так как это фундаментальная основа результатов диссертационной работы.

В третьей части дан обзор ранее описанных масс-спектрометрических подходов для изучения образцов биологических тканей.

Глава 2: Метод ионизации для тканей мозга

Во второй главе описан разработанный в рамках диссертационной работы источник прямой атмосферной ионизации электрораспылением ткани с иглы.

Описанные ранее методы исследования биологических образцов с помощью масс-спектрометрии не полностью удовлетворяли требованиям интраоперационной диагностики. Нужен был простой в пробоподготовке, быстрый и легко автоматизируемый метод определения границы опухоли во время операции. Необходимо было совершенствовать предложенные ранее методы с целью их дальнейшей адаптации к реальным клиническим условиям. Для этого мы разработали модификацию метода ионизации с помощью электрораспыления напрямую из ткани (spray-from-tissue ionization). Впоследствии метод был назван ионизацией электрораспылением с иглы, в литературе известен как NESI (Needle ElectroSpray Ionization).

В этом методе небольшой фрагмент ткани помещается на стандартную медицинскую инъекционную иглу, через которую подается растворитель. Приложение высокого напряжения к игле создает электрическое поле, достаточно сильное, чтобы инициировать электрораспылительную ионизацию непосредственно из образца. Техническая новизна предлагаемого метода заключается в том, что в данном методе образец омывается постоянным внешним потоком растворителя. В предложенной конструкции решаются такие проблемы других атмосферных ионных источников, как пересушивание образца, сложность формирования острия из мягкой ткани головного мозга и невозможность пропитки бумаги аналитом в случае исследования тканей.

Образцы для исследования были предоставлены ФГАУ НМИЦ нейрохирургии им. академика Н.Н. Бурденко Минздрава России. Исследование было проведено в соответствии с рекомендациями этического комитета НМИЦ им. акад. Н.Н. Бурденко (Приказ 40 от 12.04.2016 в редакции приказа 131 от 17.07.2018). У всех пациентов до операции было получено согласие на использование удаленных тканей в научных исследованиях в соответствии с законодательством и требованиями этического комитета. Исследование проводилось в соответствии с Хельсинкской декларацией, пересмотренной в 2013 году. Все протоколы и процедуры были выполнены в соответствии с руководящими принципами и правилами, утвержденными контролирующим советом НМИЦ им. акад. Н.Н. Бурденко.

Была проведена серия экспериментов с лабораторным прототипом NESI (см. рисунок 1). В результате этой работы были определены оптимальные параметры

измерения масс-спектров образцов опухолей мозга. Оптимальный расход растворителя составлял 4 ± 0.5 мкл/мин; точное значение зависело от формы образца. Более интенсивный поток растворителя приводил к нестационарному конусу

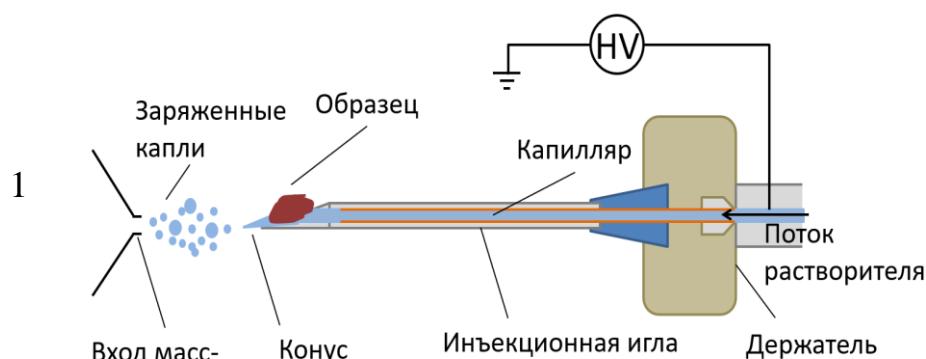


Рисунок 1. Источник ионизации NESI

Тейлора и периодическому срыву больших капель растворителя. Более низкие потоки приводили к исчезновению видимого конуса Тейлора и неустойчивым брызгам.

Также следили за ионизацией при различных значениях подаваемого напряжения. Оптимальными оказались значения HV 5,1 кВ в положительном режиме и 6,0 кВ в отрицательном режиме. Напряжения меньше оптимальных приводили к нестабильности процесса образования ионов, а более высокие — к коронному разряду и значительным изменениям масс-спектра.

Исследование проводили на масс-спектрометре высокого разрешения на основе ионного циклотронного резонанса (ICR) Thermo Finnigan LTQ FT Ultra. С помощью этого прибора получали tandemные масс-спектры методом столкновительной диссоциации (CID). Образцы анализировали как в положительном, так и в отрицательном режимах (m/z 200–2000, разрешение 56 000 при m/z 800).

Для доказательства устойчивости предложенного метода ионизации была проведена серия экспериментов и в спектрах каждого образца ткани сравнивались основные идентифицированные соединения.

Наиболее интенсивные пики были определены по их точной массе с помощью инструмента Lipid MS Predict, CID MS/MS и по их изотопному распределению из моделирования в XCalibur 2.0.7 QualBrowser (в основном для различения ионов $[M + Cl]$).

Предлагаемый метод ионизации позволил выявить не только основные компоненты масс-спектров, но и пики низкой интенсивности малых компонентов (только краткий анализ дал нам информацию о примерно 30 соединениях в m/z диапазоне 100–1000). Списки липидов, которые наблюдались в нашем эксперименте, мы сравнили с другими работами. Около половины идентифицированных нами липидов (плазменил PE, PS, GalCer, PI) были ранее упомянуты Эберлином и др. в своих исследованиях [1]. С результатами Балога и коллег [2] совпал только основной пик при

$m/z = 885,56$, что может быть связано с совершенно разными методами ионизации, используемыми в этих исследованиях.

Предложенный метод ионизации дает воспроизводимые результаты: спектры различных образцов ткани состояли из одних и тех же основных пиков с одинаковыми изотопными распределениями в течение всего времени измерения. Представленный метод позволяет достаточно просто получить стабильный спрей непосредственно из мягких тканей.

Полученные результаты являются доказательством применимости масс-спектрометрии в качестве метода экспресс-анализа тканей. Исследование также дало предварительные результаты, подтверждающие возможность наблюдения различий в масс-спектрах разных опухолей.

Глава 3: Система контроля стабильности и воспроизводимости масс-спектров при анализе опухолей головного мозга.

В третьей главе представлен метод визуального анализа стабильности измерений, использующий матрицу подобия спектров в своей основе.

Предлагаемый метод подходит как для сравнения масс-спектров различных образцов, так и для оценки стабильности масс-спектров одного образца, собранных за определенный период времени.

Подход основан на визуализации различий между измерениями, и анализе паттернов этих различий. В подходе был использован коэффициент Пирсона и мера косинуса для сравнения спектров. Метод позволяет обнаружить и удалить спектральные выбросы, заметно отличающиеся от соседей, что значительно повышает качество анализа образца. Применимость метода продемонстрирована на множестве образцов опухоли головного мозга. Разработанный метод может быть использован в нейрохирургии, где масс-спектрометрия используется для интраоперационного мониторинга границы опухоли.

Каждый масс-спектр интерпретировался как N -мерный вектор по следующему алгоритму: диапазон от 100 до 1300 m/z разбивался на ячейки размером 0,01 m/z . Каждому значению m/z в полученном разбиении присваивалось значение интенсивности, равное сумме интенсивности значений из исходных данных, попавших в эту ячейку. Для сравнения спектров были выбраны две характеристики. Первая — коэффициент корреляции Пирсона. Вторая характеристика — мера косинуса угла между двумя векторами, полученными из масс-спектров.

Ко всем полученным данным применялись фильтрация и сглаживание с помощью алгоритма скользящей медианы. В этой работе было использовано сглаживание окном размером $N = 5, 7, 21$ и 51 спектр с шагом в единицу, в половину окна и размером величины окна. Коэффициент Пирсона и мера косинуса были рассчитаны для сглаженных спектров. Все расчеты и визуализация были сделаны с

использованием самостоятельно написанного кода в программном обеспечении MATLAB.

Результаты были визуализированы с помощью матриц подобия на основе меры коэффициента Пирсона или меры косинуса. Каждый пиксель матрицы корреляции показывает значение метрики для двух спектров, соответствующих координатам по горизонтальной и вертикальной осям (рисунок 2).

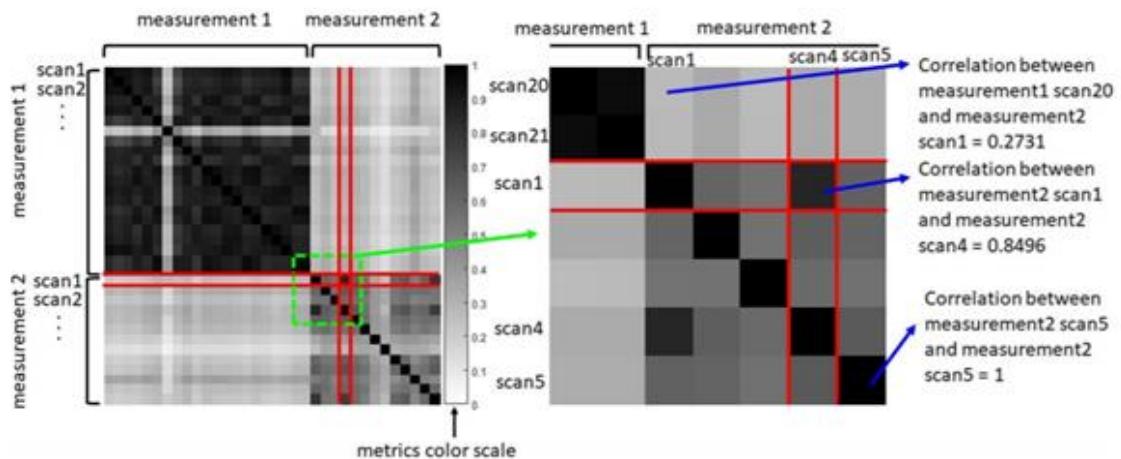


Рисунок 2. Матрица подобия спектров от двух разных измерений без сглаживания. Метрика различий спектров двух измерений и внутренних различий в каждом измерении показаны на шкале справа от матрицы. Сканы сгруппированы по измерениям и упорядочены по времени слева направо и сверху вниз в пределах измерения. Одна и та же цветовая шкала используется как для коэффициента Пирсона, так и для косинусной меры. Фрагмент матрицы увеличен до масштаба отдельных пикселей и показан справа

Глава 4. Классификация патологических типов тканей опухолей головного мозга с помощью масс-спектрометрии высокого разрешения

Эта глава посвящена исследованию возможности классификации различных типов патологической ткани с помощью методов прямой ионизации непосредственно из ткани. В описываемом исследовании изучались образцы опухолевой ткани, полученные во время удаления глиобластомы. Образцы делились на две части. Первая часть фрагментов была исследована методом иммуногистохимии и разделены на восемь классов: (1) некротическая ткань с некротизированными сосудами, (2) некротическая ткань с тенями опухолевых клеток, (3) опухоль с некрозом (опухолевая ткань как основная), (4) опухоль, (5) некротизированная опухоль (некротические ткани как основные), (6) части опухолевых клеток, (7) пограничная ткань мозга и (8) гиперплазия мозговой ткани. Вторая часть каждого образца исследовалась с помощью масс-спектрометрии с помощью описанного во второй главе метода прямой атмосферной ионизации электрораспылением с иглы (NESI).

В главе представлена методика классификации опухолевых тканей на основе обработки данных масс-спектрометрических профилей, подробно описано, создание обучающей выборки и характеристические таблицы различных типов опухолей, а также представлены результаты работы созданных классификаторов, разделяющих исследуемые образцы в соответствии с типом ткани.

На рисунке 3 показаны результаты расчета подобия между тестовыми образцами разных классов и характеристическими таблицами классов опухолевой, некрозной и ткани границы опухоли. Эллипсы построены по данным, полученным после применения классификаторов и нарисованы в трех координатах — опухолевой ткани (tumor), некрозной (necrosis) и ткани границы опухоли (unmodified). Видно, что образцы всех групп тканей с разными диагнозами хорошо разделяются.

Также в работе показано, что классификаторы некротических и опухолевых тканей дают комбинированный результат, когда ткань неоднородна и состоит из обоих типов клеток.

В главе описан еще один подход к обработке данных — выделение низкоинтенсивных пиков характерных для отдельных типов тканей соединений (feature selection). Подход позволяет выделять пики соединений, интенсивность которых сравнима или даже ниже уровня шума из-за особенностей процесса экстракции в источнике NESI, а не из-за того, что соединение присутствует в

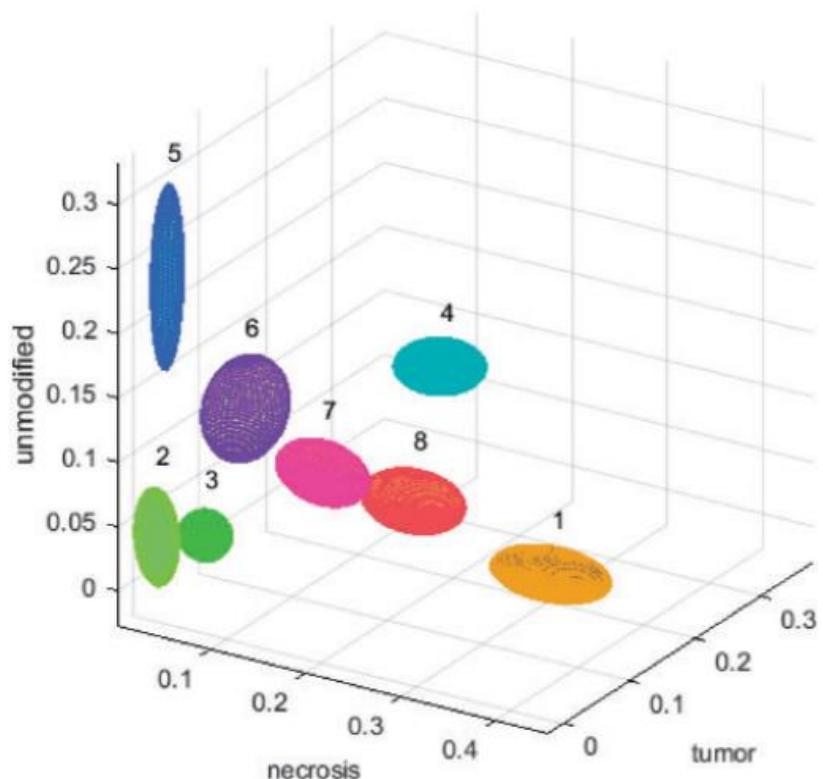


Рисунок 3. Распределение результатов схожести тестовых образцов с неопухолевой тканью (unmodified), опухолевой (tumor) и некрозом (necrosis).

Тестовая группа тканей:

- 1 — Некротическая ткань с некротизированными сосудами, 2 — Некротическая ткань с тенями опухолевых клеток, 3 — Опухоль с некрозом, 4 — Опухоль, 5 — Ткань границы опухоли, 6 — Ткань границы опухоли с гиперплазией, 7 — Некротизированная опухоль, 8 — Крошки опухолевых клеток

образце в малом количестве. Учет таких пиков позволяет строить более точные классификаторы для автоматического определения типов тканей опухолей головного мозга по масс-спектрометрическим профилям.

Глава 5. Унифицированное представление спектров высокого и низкого разрешения

Большая часть масс-спектрометрических исследований в области биомедицинских наук проводится с максимально высоким разрешением, доступным ученому. В клинике из-за ограничений по стоимости как правило используют устройства с низким разрешением. В данной главе рассказывается о методе предварительной обработки масс-спектров, который делает данные высокого разрешения пригодными для разработки методов классификации и регрессии применимых к данным низкого разрешения, при сохранении точности анализа.

Эта глава демонстрирует подход к подавлению шумов в спектрах, чтобы унифицировать представление спектров как высокого, так и низкого разрешения. В ходе исследования было показано, более 50% пиков теряются при преобразовании из высокого разрешения в низкое и наоборот (Рисунок 4). Применение одинаковой ширины биннинга к спектрам как с высоким, так и с низким разрешением уменьшает количество уникальных пиков в спектрах с высоким разрешением, связанных с неточностями разбиения.

Также было показано, что основные различия между спектрами низкого и высокого разрешения связаны с низкоинтенсивным шумом, который специфичен для конкретного прибора и не связан с измеряемым образцом. Для того, чтобы уменьшить “шумовое” различие между спектрами высокого и низкого разрешения, был применен поиск оптимальных порогов фильтрации на основе индексов Дайса[3] и Тверски[4]. Было показано, что применение пороговых значений, полученных путем максимизации коэффициента Дайса, уменьшает количество уникальных пиков в обеих настройках разрешения, увеличивая при этом долю общих пиков, соответственно, до уникальных. Однако, поскольку коэффициент Дайса обрабатывает оба набора пиков одинаково, невозможно настроить процедуру оптимизации, чтобы учесть разницу в важности уникальных пиков для спектров высокого и низкого разрешения.

Использование индекса Тверски было введено для уменьшения веса уникальных пиков с высоким разрешением и увеличения количества несогласованных пиков в спектрах с низким разрешением. Оптимальное значение параметра асимметрии α составило 0,2. Применение индекса Тверски приводит к более низким пороговым

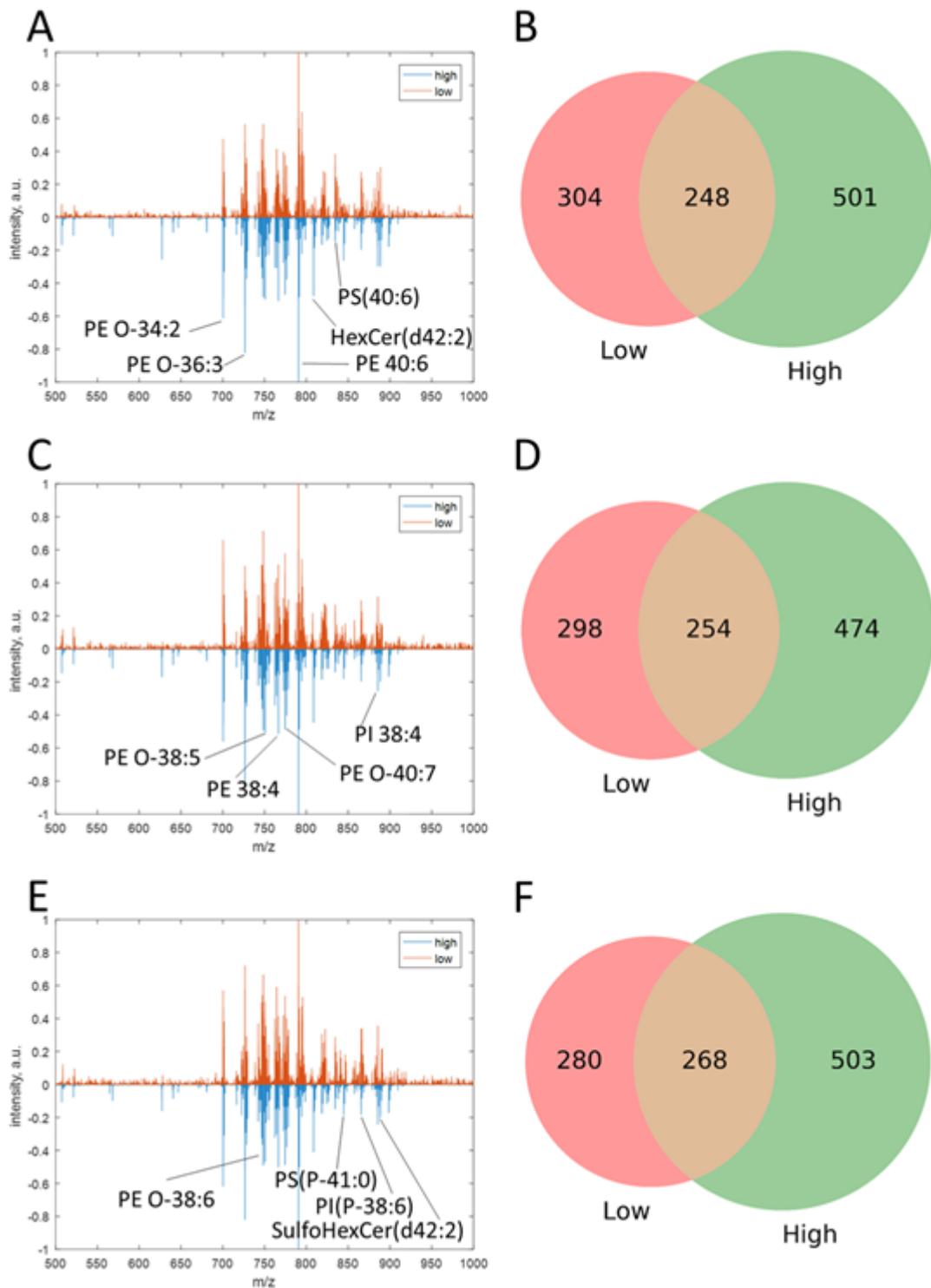


Рисунок 4. Спектры высокого разрешения (нижний спектр) и низкого разрешения (верхний) после применения разбиения и peak-peaking процесса. Спектры A, C, E – спектры первого, второго и третьего образца соответственно. B, D, F – соответствующие диаграммы Венна.

значениям и сохраняет больше уникальных пиков с высоким разрешением по сравнению с методом максимизации коэффициента Дайса.

В результате был предложен подход, который использует обнаружение порога шума на основе индекса Тверски, и сравнивает спектры с различным разрешением, а также минимизирует процент низкоинтенсивных пиков, которые попадают в результирующий спектр только из данных высокого разрешения. При этом отброшенные пики не должны играть заметную роль в характеризации образца и применении к нему алгоритмов классификации. Представленный метод дает возможность разрабатывать аналитические алгоритмы с использованием данных масс-спектрометрии высокого разрешения и применять эти алгоритмы в клинике к данным, полученным при помощи масс-спектрометров низкого разрешения, без заметного снижения точности.

Результатом исследования стал алгоритм, который позволяет преобразовывать данные высокого разрешения в представление, наиболее близкое к структуре спектра с низким разрешением и подходящее для применения методов машинного обучения, что значительно снижает потребность в экспериментальных данных с низким разрешением при внедрении полученных результатов в клиническую практику.

Интенсивности пиков, остающихся после применения алгоритма, довольно линейны в спектрах как низкого, так и высокого разрешения. Различия можно объяснить возможным агрегированием интенсивностей нескольких пиков высокого разрешения в один пик низкого разрешения.

Разработанный подход представляет собой оптимальную стратегию преобразования данных с высоким и низким разрешением в одно и то же унифицированное представление.

Выводы:

Все поставленные задачи выполнены и по результатам работы сделаны следующие выводы:

1. Метод экспресс-анализа биопсийных образцов опухолей головного мозга для их последующей идентификации, разработанный на основе комбинированного ионного источника с прямой атмосферной ионизацией, в котором реализуется жидкостная экстракция и электроспрейная ионизация, позволяет получать масс-спектрометрические профили, характеризующие липидный состав исследуемых образцов.

2. Метод профилирования биологических образцов на основе комбинации жидкостной экстракции с электроспрейной ионизацией позволяет получать стабильные и воспроизводимые результаты в течение длительного промежутка времени, достаточного для регистрации информативного молекулярного профиля исследуемого образца. Метод контроля качества масс-спектрометрических профилей, а также их стабильности и воспроизводимости, построенный на основе матриц

подобия спектров, позволяет визуально оценивать, насколько данные, полученные, в отдельных сканированиях, являются типичными для всего измерения в целом и позволяет отфильтровывать ошибочные сканирования, вызванные сбоями в процессе экстракции или измерения.

3. Подход к предварительной обработке данных масс-спектрометрического профилирования, заключающийся в кластеризации пиков в масс-спектре, в том числе находящихся ниже уровня шума по интенсивности, позволяет выделять стабильно детектируемые ионы, что уменьшает размерность сохраняемых данных и позволяет реализовать обработку на стандартной персональной рабочей станции.

4. Разработанный алгоритм классификации данных масс-спектрометрического профилирования на основе комбинации методов статистического анализа позволяет осуществлять:

— классификацию образцов различных опухолевых тканей пациента с помощью алгоритмов машинного обучения, построенных на масс-спектрометрических профилях тканей этого же пациента;

— классификацию образцов тканей разных пациентов с помощью алгоритмов машинного обучения, построенных на масс-спектрометрических профилях различных опухолевых тканей одного пациента. Так как вариабельность типов опухолевых тканей на молекулярном уровне выше, чем вариабельность одинаковых типов тканей у разных пациентов, алгоритм корректно справляется с классификацией образцов.

5. Конверсия данных высокого и низкого разрешения в универсальную шкалу масс с заданным шагом может осуществляться с максимальным сохранением информативности, что является путём для создания унифицированного подхода к хранению и последующей обработке данных.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в научных журналах

1. Kononikhin A, **Zhvansky E**, Shurkhay V, Popov I, Bormotov D, Kostyukevich Y, Karchugina S, Indeykina M, Bugrova A, Starodubtseva N, Potapov A, Nikolaev E. A novel direct spray-from-tissue ionization method for mass spectrometric analysis of human brain tumors. //Anal Bioanal Chem. 2015 Oct;407(25):7797-805.
2. **ES Zhvansky**, AA Sorokin, IA Popov, VA Shurkhay, AA Potapov, EN Nikolaev. High-resolution mass spectra processing for the identification of different pathological tissue types of brain tumors. // Eur J Mass Spectrom. 2017 Aug; 23(4): 213-6.
3. Anatoly Sorokin, **Evgeny Zhvansky**, Vsevolod Shurkhay, Konstantin Bocharov, Igor Popov, Nikita Levin, Dmitry Zubtsov, Denis Bormotov, Yury Kostyukevich, Alexander Potapov, Eugene Nikolaev. Feature selection algorithm for spray-from-tissue mass spectrometry. // Eur J Mass Spectrom. 2017 Aug; 23(4): 237-41.

4. **Zhvansky E. S.**, Pekov S. I., Sorokin A. A., Shurkhay V. A., Eliferov V. A., Potapov A. A., Nikolaev E. N., Popov, I. A. Metrics for Evaluating the Stability and Reproducibility of Mass Spectra // *Sci. Rep.* 2019, 9 (1), 914
5. **Zhvansky E. S.**, Sorokin A. A., Pekov S. I., Indeykina M. I., Ivanov D. G., Shurkhay V. A., Eliferov V. A., Zavorotnyuk D. S., Levin N. G., Bocharov K. V., Tkachenko S. I., Belenikin M. S., Potapov A. A., Nikolaev E. N., Popov I. A. Unified Representation of High- and Low-Resolution Spectra to Facilitate Application of Mass Spectrometric Techniques in Clinical Practice // *Clinical Mass Spectrometry* 2019, 12, 37–46

Тезисы на конференциях

6. **Жванский Е.**, Попов И., Кононихин А., Шурхай В., Бормотов Д., Бугрова А., Потапов А., Николаев Е. Масс-спектрометрическое профилирование тканей опухолей мозга как способ интраоперационного контроля хода нейрохирургической операции // I Всероссийская конференция с международным участием «Химический анализ и медицина». Москва, 2015
7. **Zhvansky E.** Methods for quantitative analysis of elemental composition of biological substances via ultra-high resolution mass spectrometry // VI-АЯ МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ-ШКОЛА ДЛЯ МОЛОДЕЖИ «ФУНДАМЕНТАЛЬНЫЕ ВОПРОСЫ МАСС-СПЕКТРОМЕТРИИ И ЕЕ АНАЛИТИЧЕСКИЕ ПРИМЕНЕНИЯ», Dolgoprudnyi, Russia, 2016
8. Popov I., **Zhvansky E.**, Bormotov D., Shurkhay V., Indeykina M., Kostyukevich Y., Kononikhin A., Starodubtseva N., Potapov A., Nikolaev E. // Second International Conference on Mass Spectrometry (ICMS), Kottayam, Kerala, India, 2015
9. Popov I., **Zhvansky E.**, Bormotov D., Shurkhay V., Indeykina M., Kostyukevich Y., Kononikhin A., Starodubtseva N., Potapov A., Nikolaev E. Novel spray-from-tissue ionization method for online and offline neurosurgery assistance: identification of profiles by FTICR and quadrupole ion trap MS // 63rd ASMS Conference on Mass Spectrometry and Allied Topics, St. Louis, USA, 2015
10. **Zhvansky E.**, Popov I., Sorokin A., Levin N., Shurkhay V., Bormotov D., Indeykina M., Kononikhin A., Kostyukevich Y., Kukaev E., Potapov A., Nikolaev E. Machine learning algorithms for advanced identification of tumor tissue and tumor borders // 64th ASMS Conference on Mass Spectrometry and Allied Topics, San Antonio, USA, 2016
11. Potapov A., Kukaev E., Levin N., Popov I., Kostyukevic Y., Kononikhin A., Nikolaev E., **Zhvansky E.**, Shurkhay V., Indeykina M. Ambient MS based analytical platform for rapid identification of brain cancer tumor tissues // 21ST INTERNATIONAL MASS SPECTROMETRY CONFERENCE Toronto, Ontario, Canada, 2016
12. Sorokin A., Shurkhay V., Popov I., **Zhvansky E.**, Levin N., Potapov A., Nikolaev E. Investigation of meningioma samples by FTICR mass spectrometry // The 25th Intelligent Systems for Molecular Biology and European Conference on Computational Biology (16th Annual Conference) ISMB/ECCB 2017, Prague, Czech Republic, 2017
13. **Zhvansky E.**, Sorokin A., Popov I., Shurkhay V., Levin N., Potapov A., Nikolaev E. Novel approach for chemical formula identification based on isotopic distribution observed with ultrahigh resolution FTICR mass spectrometer // 65th ASMS Conference on Mass Spectrometry and Allied Topics, Indianapolis, 2017

14. **Zhvansky E.**, Sorokin A., Popov I., Shurkhay V., Zubtsov D., Potapov A., Nikolaev E. The approach for brain tumor detection via mass-spectrometry-based classifier // The Annual Scientific Meeting CLINICAL PROTEOMICS. POSTGENOME MEDICINE, Moscow, Russia, 2017
15. Popov I., **Zhvansky E.**, Sorokin A., Pekov S., Eliferov V., Vorobyev A., Shurkhay V., Potapov A., Nikolaev E. Comparison of high- and low-resolution MS data for direct tissue profiling on a way from laboratory to clinic // 66th ASMS Conference on Mass Spectrometry and Allied Topics, San Diego, 2018
16. Sorokin A., Popov I., Kononikhin A., **Zhvansky E.**, Pekov S., Shurkhay V., Potapov A., Nikolaev E. The Database For Mining The Biomarkers Of the Brain Tumors From the Mass-Spectrometry Data // 2018 Congress of Neurological Surgeons (CNS) Annual Meeting, Houston, Texas, 2018
17. Popov A., Shurkhay V., Eliferov V., **Zhvansky E.**, Pekov S., Sorokin A., Potapov A., Nikolaev E. Online MS-based Workflow For Intraoperative Identification Of Brain Tumor Tissues // 2018 Congress of Neurological Surgeons (CNS) Annual Meeting, Houston, Texas, USA, 2018
18. **Zhvansky E.**, Sorokin A., Ivanov D., Eliferov V., Bugrova A., Pekov S., Popov I., Nikolaev E. Dissimilarity metrics mapping algorithm for assist region detection in mass spectrometry imaging // 67th ASMS Conference on Mass Spectrometry and Allied Topics, Atlanta, USA, 2019
19. Popov I., Sorokin A., Shurkhay V., Eliferov V., **Zhvansky E.**, Bormotov D., Pekov S., Potapov A., Nikolaev E. Mass spectrometry molecular profiling as an alternative to intraoperative histological examination // Congress of Neurological Surgeons (CNS) Annual Meeting, San Francisco, California, USA, 2019
20. **Zhvansky E.**, Sorokin A., Ivanov D., Zavorotnyuk D., Pekov S., Eliferov V., Nikolaev E., Popov I. Software tool for visual inspection of the stability and reproducibility of mass spectra // 68th ASMS Conference on Mass Spectrometry and Allied Topics, Online Meeting, 2020

Использованная литература:

1. Eberlin C. et al. Discrimination of Human Astrocytoma Subtypes by Lipid Analysis Using Desorption Electrospray Ionization Imaging Mass Spectrometry. P. 5953–5956.
2. Balog J. et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry // Science Translational Medicine. American Association for the Advancement of Science, 2013. Vol. 5, № 194. P. 194ra93-194ra93.
3. Dice L.R. Measures of the Amount of Ecologic Association Between Species // Ecology. Wiley, 1945. Vol. 26, № 3. P. 297–302.
4. Tversky A. Features of similarity // Psychological Review. 1977. Vol. 84, № 4. P. 327–352.

ЖВАНСКИЙ ЕВГЕНИЙ СЕРГЕЕВИЧ

МАСС-СПЕКТРОМЕТРИЯ В ЗАДАЧЕ ПРОФИЛИРОВАНИЯ
БИОЛОГИЧЕСКИХ ТКАНЕЙ ЧЕЛОВЕКА ДЛЯ ИДЕНТИФИКАЦИИ
ОПУХОЛЕЙ ГОЛОВНОГО МОГЗА

Автореферат диссертации на соискание ученой степени
кандидата физико-математических наук

Подписано в печать 22.10.2020. Формат 60x90 $1/16$. Усл. печ. л. 1.25. Тираж 50 экз.
Заказ № 18060. Печать трафаретная. Типография «11-й ФОРМАТ». ИНН
7726330900. 115230, Москва, Варшавское ш., 36.

Федеральное государственное автономное образовательное учреждение высшего образования “Московский физико-технический институт (национальный исследовательский университет)”
141701, Московская обл., г. Долгопрудный, Институтский пер., д. 9.