

Исследование методов организации виртуального эксперимента для задачи поиска эффективной связности функциональной магнитно-резонансной томографии действия человека

© Д.С. Ендеева

Московский государственный университет имени М.В. Ломоносова,
Москва, Россия

endeevavmk@gmail

Аннотация. Быстрое развитие информационных технологий позволяет резко увеличить как объем накопленных данных, так и способствует продвижению новых методов анализа этих данных. Работа посвящена исследованию техники проведения виртуального эксперимента на примере задачи анализа когнитивных функций человеческого мозга. Задача происходит из одной из самых динамично развивающихся областей с интенсивным использованием данных – нейрофизиологии. В процессе исследования построена онтология предметной области, описаны характерные особенности данных, явно специфицированы используемые априорные гипотезы. На примере одного из исследований формально описан поток работ поиска эффективной связности данных функциональной магнитно-резонансной томографии человека. Предполагается, что формальное описание всех этапов проведения виртуального эксперимента может быть полезным как исследователям из названной области, так и при проведении анализа данных из других мультидисциплинарных областей.

Ключевые слова: аналитика и управление данными, интенсивное использование данных, виртуальный эксперимент, когнитивные функции мозга.

Organizing a Virtual Experiment for the Analysis of Effective Connectivity of Human Task Functional Magnetic Resonance Imaging

© Darya Endeeva

Lomonosov Moscow State University,
Moscow, Russia

endeevavmk@gmail.com

Abstract. The rapid development of information technologies allows us to dramatically increase both the volume of accumulated data and promote new methods for analyzing this data. The article is devoted to the study of the technique of managing virtual experiment with the example of the problem of analyzing the cognitive functions of the human brain. The task comes from one of the most dynamically developing areas with the intensive use of data – neurophysiology. In the process of research, the ontology of the domain is constructed, the characteristic features of the data are described, the a priori hypotheses used are explicitly specified. Formally specified workflow for searching the effective connectivity of human fMRI is extracted. It is assumed that a formal description of all stages of the virtual experiment can be useful both for researchers from the given area and for analyzing data from other multidisciplinary areas.

Keywords: analytics and data management, data intensive domains, virtual experiment, cognitive functions.

1 Введение

Развитие информационных технологий оказывает огромное влияние на экспериментальную, теоретическую и вычислительную науку. Развитие технологий сильным образом повлияло на нейроинформатику, в которой происходит накопление огромного массива данных, развиваются новые методы, требующие больших

Труды XIX Международной конференции «Аналитика и управление данными в областях с интенсивным использованием данных» (DAMDID/ RCDL'2017), Москва, Россия, 10–13 октября 2017 года

вычислительных мощностей. Данные, в свою очередь, становятся основным источником новых научных открытий. Примерами крупных проектов в области нейрофизиологии являются проекты Human Connectome Project (HCP), NITRC: 1000 Functional Connectomes. Объем накопленных данных в этих проектах достигает петабайтов. Например, проект HCP содержит данные функциональной магнитно-резонансной томографии (фМРТ) около 1200 человек, полученные на сканерах 3T MR и 7T MR, а также данные магнитоэнцефалографии (МЭГ) 95 человек (среди которых структурные изображения; изображения, полученные в состоянии покоя субъекта; изображения, полученные при выполнении субъектом заданий; диффузные данные).

Появляется потребность в семантической методологии, которая упрощает моделирование научных знаний, проверку гипотез, семантическую интеграцию данных, анализ данных для различных дисциплин для использования неспециалистами. Важной задачей исследователей является обеспечение повторного использования методов другими исследователями и воспроизводимости результатов эксперимента, для чего процессы работ должны снабжаться метаданными и спецификациями методов.

Для организации обработки данных и систематического, повторяемого и воспроизводимого выполнения всех процедур, являющихся частью процесса исследования, становится целесообразно описывать и разрабатывать потоки работ, которые представляют собой точное описание научной процедуры, состоящей из множества последовательных шагов. Продуманные повторения потоков работ автоматически дают новые результаты при доступности новых исходных данных и новых результатов, а также новых методов. Сами рабочие процессы, как важная часть науки, ориентированной на данные, могут создаваться и динамически трансформироваться, исходя из текущих потребностей. Объединение рабочих процессов с опубликованными результатами обеспечивает прозрачность и сравнимость исследований, что может ускорить процесс научных открытий.

Целью работы является исследование методов и средств для организации виртуальных экспериментов по анализу эффективной связности функциональной магнитно-резонансной томографии действия здоровых людей. Достижение цели предполагает решение следующих задач: разработка онтологии и концептуальной схемы предметной области, которая требуется для понимания процессов, происходящих в мозге, и последующего моделирования этих процессов; формирования гипотез; построение виртуального эксперимента; формализация потока работ виртуального эксперимента; обзор существующих решений по обработке и анализу данных функциональной магнитно-резонансной томографии, требующихся на разных этапах потока работ; исследование методов

анализа эффективной связности функциональной магнитно-резонансной томографии. Под виртуальным экспериментом подразумевается относительно новое направление в научно-исследовательском процессе, обусловленное реализацией моделей средствами вычислительной техники.

Основой для построения потока работ послужил виртуальный эксперимент, посвященный поиску эффективной связности между двумя областями мозга – V5 и pSTS – на основе фМРТ данных на задаче восприятия одушевленного движения методом Динамического Каузального Моделирования (Dynamic Causal Modelling, DCM) [1].

Сенсорная зрительная зона V5 (или MT, от англ. Middle Temporal) считается чувствительной к любому типу визуального движения [3]. Регион pSTS (от англ. posterior superior temporal sulcus) – задняя область верхней височной борозды – во многих исследованиях (например, в [4]) связывается с социальной перцепцией, анализом и восприятием биологических движений и анализом намерений человеческих субъектов.

Участникам показывали серию коротких видеоклипов, в которых простые фигуры (треугольники, квадраты, круги) либо взаимодействовали друг с другом (одушевленное движение), либо двигались механически (неодушевленное движение). Одушевленное движение обычно более сложное, нелинейное и, таким образом, менее предсказуемое по сравнению с неодушевленным движением, так как основано на скрытых намерениях агента, производящего движение. Неодушевленное движение, наоборот, происходит благодаря действиям относительно неизменных сил (физических законов) и поэтому считается более предсказуемым. Умение предсказывать одушевленное движение является ключевым для выживания, что подчеркивает актуальность исследования восприятия движения.

2 Задача исследования когнитивных функций мозга

2.1 Онтология предметной области

Для корректной совместной работы с экспертами из области была разработана онтология предметной области (см. Рис. 1).

С помощью визуализации методом магнитно-резонансной томографии (МРТ) можно изучать структуру (анатомию) мозга, а с помощью визуализации методом функциональной магнитно-резонансной томографии – его функциональность. Изображения, полученные методами МРТ и фМРТ, собирают с одного сканера. Единицей объема получаемого со сканера трехмерного изображения является воксель. Изображение, полученное методом МРТ, представляет собой трехмерный массив вокселей. Слайсом называется двумерный массив вокселей нейроиображения.

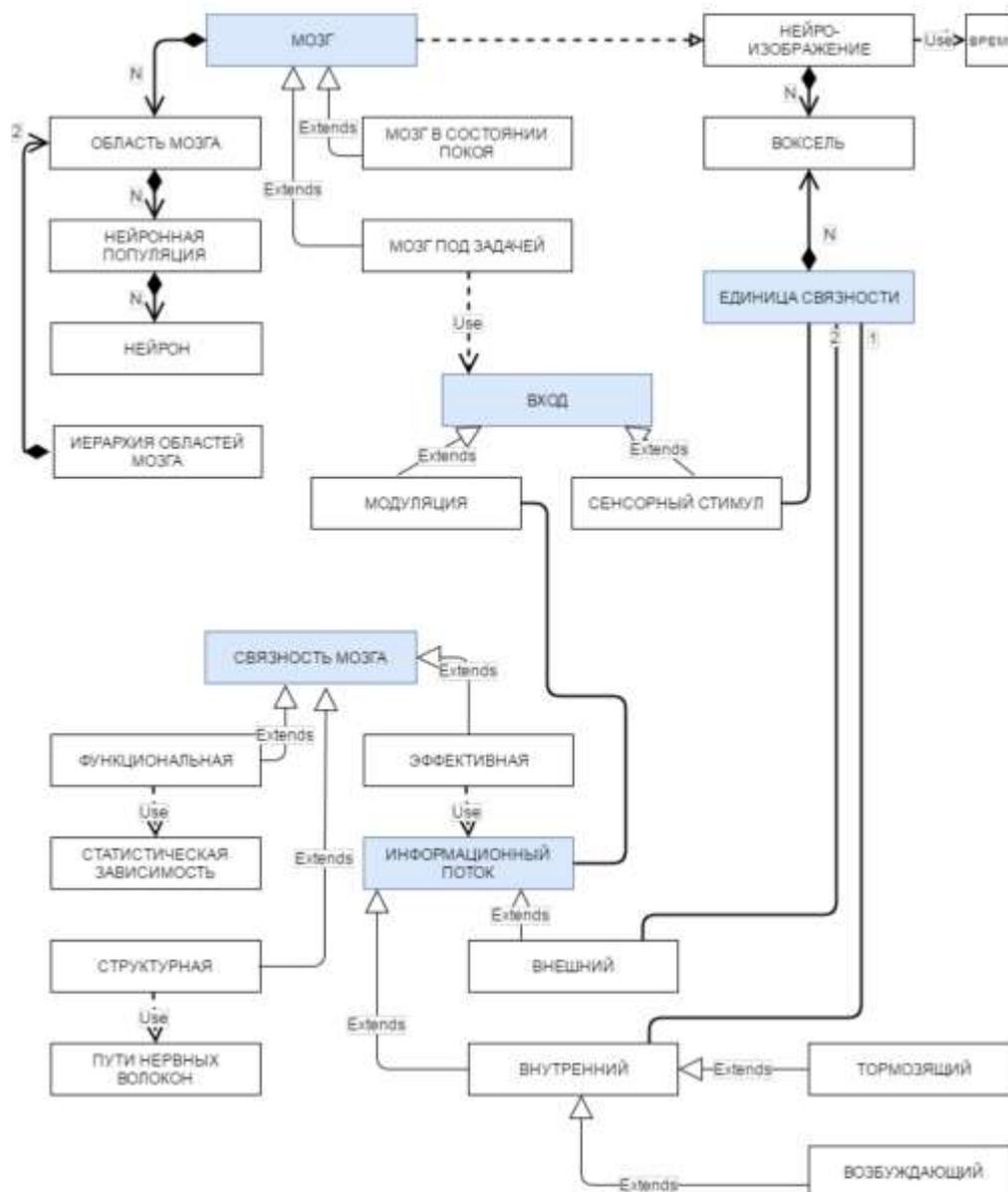


Рисунок 1

В каждом элементе массива (вокселе) содержится некоторое значение, отражающее силу сигнала в данном вокселе (чем больше значение, тем светлее пиксель на изображении). К определенному вокселю можно обратиться как к элементу массива по индексу.

Изображение, полученное методом фМРТ, состоит из последовательности трехмерных изображений и, таким образом, имеет четвертое измерение (временное) и содержит временные ряды для каждого вокселя.

Основным форматом данных в нейровизуализации методом фМРТ является формат NIFTI, содержащий в себе сами данные, а также информацию о том, как данные собирались со сканера.

Одной из главных причин сложности нервной системы является её сложная морфология, особенно взаимосвязность обрабатывающих нейронных

элементов. Паттерны нейронной связности играют важную роль в определении функциональных свойств нейронов и нейронных систем. В более развитых нервных системах связность мозга можно рассматривать на разных масштабных уровнях. На микроуровне отдельные нейроны связываются друг с другом при помощи синаптических связей, на мезоуровне нейронные популяции объединяются в сети, а на макроуровне области мозга связываются нейронными проводящими путями, которые часто представляют собой пучки аксонов и служат для соединения относительно удаленных участков мозга или нервной системы.

Выделяются три вида связности мозга – структурная (или анатомическая связность), функциональная связность и эффективная связность.

Структурная связность описывает анатомические связи между нейронами посредством аксонов. Структурные связи специфичны и непостоянны. Структурная связность различна не только для

разных особей одного и того же вида, но и изменяется со временем у одного индивидуума. Она является относительно стабильной на коротких временных интервалах.

Согласованность работы отдельных областей оценить труднее, чем локализовать их активность. Один из возможных подходов – оценить функциональные связи, то есть статистические зависимости между активностью в различных областях мозга. Функциональная связность часто вычисляется на всех элементах системы, независимо от того, соединены ли данные элементы напрямую структурными связями. Функциональная связность также очень зависима от времени (может изменяться за десятые и сотые доли секунды). Также функциональная связность не связана очевидным образом с внутренней структурой модели или с влиянием одних областей мозга на другие.

Лучше всего интеграция оценивается в терминах эффективных связей, которые показывают, каким образом одна нейрональная система влияет (оказывает эффект) на другую. В некотором смысле эффективные связи – отражение истинных нейрональных процессов, в то время как функциональные связи – лишь статистически значимое сходство активности в различных областях мозга. Применительно к взаимодействиям между нейрональными популяциями эффективные связи – это каузальное взаимодействие между возбуждающими популяциями (являющимися причиной возбуждения или источником информации) и зависимыми от них или возбуждаемыми популяциями [8].

2.2 Пример множества априорных гипотез

Гипотезы являются ключевым элементом эксперимента. Формализация априорных гипотез может быть полезна, если потом потребуется заменить одну из них конкурирующей.

Исследование эффективной связности на задаче восприятия одушевленного движения опирается на базовые априорные гипотезы, которые в последующем сокращают пространство соревнующихся моделей:

H1: Область V5 – область зрительной системы, которая по предположению играет важную роль в восприятии любого типа движения;

H2: Область pSTS – область, которая по предположению больше активируется при восприятии одушевленного движения;

H3: По теории предсказательного кодирования по прямым связям передается ошибка предсказания, а по обратным – предсказание. Иерархически высший регион стремится уменьшить ошибку предсказания;

H4: Одушевленное движение менее предсказуемое, чем механическое;

H5: Внутренняя структура моделей не различается среди участников эксперимента;

H6: Есть разница в связности мозга (количестве информации, передаваемой между регионами) при восприятии одушевленного и неодушевленного движения;

H7: Входные стимулы по-разному влияют на исследуемые регионы.

2.3 Анализ методов эффективной связности фМРТ человека

Существует два разных подхода к поиску каузальных взаимодействий: моделирование и статистические оценки асимметричных метрик связности двух и более процессов (временных рядов).

Известны следующие метрики для линейного случая: кросс-корреляция (англ. cross-correlation), условная взаимная информация (англ. conditional mutual information), трансферная энтропия (англ. transfer entropy, TE [14]), моментальный перенос информации (англ. momentary information transfer) и др.

Перечисленные метрики асимметричны при перестановке процессов местами и при обращении времени. Для их использования, как правило, не требуется никакого знания о системе и составляющих ее подсистемах, а также нет необходимости в априорных гипотезах и характере каузальных взаимодействий. Метрики теории информации могут быть применены к данным любой природы: от анализа финансовых рынков до геофизики и нейронаук. Отсюда вытекают основные преимущества и недостатки подобных метрик: универсальность; возможность первоначального анализа сложных систем, о которых ничего не известно; способны показать лишь наличие или отсутствие влияния одних процессов на другие, не предполагая под временными рядами никакого внутреннего процесса.

Моделирование представлено большим количеством методов, таких, как причинность по Грейнджеру (англ. Granger Causality, GC [15]), моделирование структурными уравнениями (англ. Structural Equation Modelling, SEM [16, 17]), динамическое моделирование причинности (динамическое каузальное моделирование, англ. Dynamic Causal Modelling, DCM [18]) и др.

В нейронауках наблюдаемые экспериментальные данные неинвазивных методов (электрическая активность на скальпе в ЭЭГ или изменение концентрации дезоксигемоглобина в фМРТ) являются лишь коррелятами истинной, интересующей экспериментатора активности нейронов, поэтому моделирование может быть очень полезным для понимания скрытых состояний и процессов.

Недостатки моделирования: зависимость от априорных знаний и предположений; методы плохо подходят для первоначального анализа больших массивов экспериментальных данных, так как нуждаются в проверяемых гипотезах.

3 Формализация потока работ

Схема потока работ представлена на Рис. 2.



Рисунок 2

Формирование и вывод априорных гипотез, цели и задачи исследования, предположение о методе, который будет использован для анализа. На этом этапе исследователи формируют задачу исследования, обозначают априорные гипотезы, если это необходимо, формируют представления о том, какими методами будет осуществляться анализ, какие инструменты будут использоваться и т. д.

Проектирование эксперимента. На данном этапе исследователи обдумывают и подготавливают эксперимент, опирающийся на априорные гипотезы и цели исследования. Например, эксперименты для проекта НСР спроектированы с помощью системы E-Prime.

Сбор данных. На данном этапе проводится эксперимент, участники эксперимента выполняют задания, в то время как исследователи собирают нейроданные с МРТ сканера.

Предобработка данных. На данном этапе данные, полученные на предыдущем этапе, подвергаются низкоуровневой предобработке, включающей в себя коррекцию движения, удаление пространственных артефактов и искажений, нелинейную регистрацию вокселей в пространстве MNI152, выделение масок и т. д.

Ядерное сглаживание. Гауссовское ядерное сглаживание – техника сглаживания изображений, наиболее часто используемая в нейровизуализации.

ФМРТ данные содержат много шума, но исследования показывают, что в основном этот шум гауссовский, т. е. случайный, независимый от вокселя к вокселю и ориентированный вокруг нуля. Таким образом, если усреднить интенсивности соседних вокселей, то шум будет стремиться к нулю, а сигнал – к некоторому среднему значению, отличному от нуля.

Таким образом, применяя пространственное сглаживание некоторой ядерной функцией, исследователи стремятся «размыть» изображение, смягчив жесткие края, снизив общую пространственную частоту, и улучшить отношение сигнал/шум:

$$Y(t) = \int K(t, s)X(s)ds,$$

где K – ядро интеграла, X – входной сигнал, Y – выходной сигнал.

Создание контрастов и регрессоров, подготовка дизайн-матрицы. Различные программы имеют разные методы настройки дизайн-матриц эксперимента, но все они основываются на том, что исследователь описывает набор различных экспериментальных условий (или эффектов) и указывает для каждого из них время начала и время окончания (либо время начала и продолжительность). Эксперименты могут иметь разные формы, от простейших блочных конструкций до сложных случайных событий со множеством условий. Основная гипотеза заключается в том, что некоторые воксели в мозге имеют статистически значимую активацию при некоторой комбинации экспериментальных условий и параметров.

Дизайн-матрица представляет собой матрицу, в которой строки представляют собой временные точки, а столбцы – смоделированные экспериментальные эффекты. Обычно исследователь модифицирует дизайн-матрицу, чтобы сделать модель более точно характеризующей активность мозга. Часто в матрицу добавляется константный регрессор для учета среднего значения сессии. Иногда к матрице добавляются линейные или полиномиальные дрейфы.

Иногда столбцы матрицы свертываются с некоторой функцией гемодинамического ответа для отражения размытости в сигнале [9].

Первый уровень анализа (по каждому субъекту). Большинство стандартных способов анализа фМРТ предполагает использование общей линейной модели (англ. General Linear Model, GLM), которая

по сути представляет собой расширение линейной множественной регрессии для случая одной зависимой переменной.

В ОЛМ исследователь задает предполагаемую модель ответов мозга при изменении некоторых экспериментальных условий, а затем проверяет истинность этой гипотезы. (Альтернативой ОЛМ может считаться метод главных компонент (англ. PCA)). Оценка общей линейной модели является основным статистическим этапом процесса анализа нейроданных, в котором исследуется, насколько заданная модель правильно интерпретирует реальные данные для каждого вокселя.

Стандартное уравнение GLM:

$$Y(t) = bX + \varepsilon,$$

где Y – изменяющиеся во времени интенсивности от одного вокселя, X – дизайн-матрица, ε – вектор нормально распределенных ошибок, а b – «бета-веса» – искомый параметр, вектор значений, показывающий, насколько значительным был вклад каждого эффекта экспериментального условия в объяснение значений в данном вокселе. Уравнение решается методом наименьших квадратов (англ. Ordinary Least Squares, OLS):

$$b = (X^T X)^{-1} X^T Y.$$

Если b -веса регрессора А значительно выше, чем b -веса регрессора В в данном вокселе, подтверждается гипотеза о том, что А обладает большим эффектом, чем В в этом вокселе.

После получения оценок эффектов и ошибок для каждого вокселя можно вычислить статистическую значимость для эффектов с использованием контрастов, определенных на предыдущем шаге, и получить контрастное изображение (англ. contrast image, beta image), содержащее информацию о величине интересующего эффекта: $t = \frac{c^T b}{Std(c^T b)}$.

При задании контрастов, призванных выделить наиболее активные воксели при определенном условии, тестируют гипотезу: $H_0: cb=0$, где c – веса для b .

Второй уровень анализа (групповой). Главный вопрос группового анализа заключается в том, повторяется ли паттерн активности вокселей среди участников эксперимента. Главная цель состоит в том, чтобы найти групповую маску, которая бы выделяла воксели, активность которых была статистически значима среди всех участников для некоторого экспериментального условия.

На вход ОЛМ второго уровня, в отличие от ОЛМ первого уровня, в качестве вектора Y подаются контрастные изображения, полученные на предыдущем шаге. Дизайн-матрица на этом уровне анализа общая (групповая).

Есть несколько вариантов статистических тестов для группового анализа в зависимости от того, учитывается ли различие между субъектами (Random Effects vs Fixed Effects); относятся ли все субъекты к

одной группе либо их нужно анализировать, разделив перед этим на несколько групп, например, на мужчин и женщин (one-sample t-test, two-sample t-test); какие и какое количество контрастов будет использоваться (t-test, F-test, ANOVA).

Выбор вокселей областей интереса, показавших наибольшую активацию, извлечение временных рядов для этих областей (локализация нейронной активности). На данном шаге воксели сопоставляются с атласом регионов мозга. Если гипотеза касалась определенных регионов мозга, то определяются координаты областей, интересующих исследователей. По статистическим картам предыдущего этапа выделяются пиковые воксели, находящиеся ближе всего к интересующим областям, после чего единицей связности (областью интереса) помечаются либо пиковый воксель, либо еще несколько окружающих его вокселей. Далее извлекаются временные ряды для этих областей интереса и усредняются, если вокселей в единице связности несколько.

Спецификация моделей. Спецификация моделей зависит от выбранных априорных гипотез и метода анализа. Каждый метод анализа сокращает пространство моделей [10].

Моделирование В DCM разделяется на два уровня – нейронный и BOLD. МРТ измеряет активность исследуемых областей не напрямую (нейронный уровень), а посредством измерения гемодинамических реакций, являющихся следствием нейронной активации (BOLD уровень), и основной идеей метода DCM является оценка параметров нейронного уровня с использованием априорно известной биофизической информации так, чтобы модель нейронной системы наилучшим образом предсказывала наблюдаемый BOLD-сигнал.

Далее под системой подразумевается набор элементов (нейронных популяций, областей), взаимодействующих друг с другом во времени и в пространстве.

DCM моделирует для каждой исследуемой популяции изменение абстрактного нейронного состояния (англ. state) во времени. Важны не сами значения этих состояний, а их динамика во времени.

Уравнение состояний для билинейной модели имеет вид:

$$\dot{z} = (A + \sum_{j=1}^m u_j B^j)z + Cu.$$

Таким образом, в DCM нейронная динамика моделируемой системы зависит от 4х параметров: A – внутренняя связность элементов системы (влияние нейронных популяций друг на друга в терминах причинности), B – контекстуальные входы (изменения силы связности), C – управляемые входы (сенсорные стимулы, источники активации модели), σ – задержка вызванной активности.

Моделируемая динамика нейронных популяций z (нейронный уровень) трансформируется в BOLD-сигнал y (BOLD уровень) с помощью гемодинамической модели, называемой моделью

«Воздушного шара». Комбинируя нейрональные z и гемодинамические y состояния в совместный вектор состояния x , а нейрональные и гемодинамические параметры – в совместный вектор параметров θ , получим полную прямую (порождающую) модель.

Важно отметить, что с помощью DCM можно получить два разных вывода [10].

Если интересны не конкретные параметры модели, но сама ее структура, то требуется заключение о пространстве моделей, например, если хочется узнать, имеет ли конкретная нейронная система последовательную или параллельную архитектуру, и при этом не требуется знать, касается ли контекстно-зависимая модуляция прямых или обратных связей или является ли механизм модуляции линейным или нелинейным. Или наоборот, если интересны нейрофизиологические механизмы, закодированные в конкретных параметрах данной модели, то требуется заключение о параметрах модели. Например, для некоторой модели хочется знать, будет ли конкретная связь оказывать с большей вероятностью возбуждающее или тормозящее воздействие на целевую область.

Перед проведением исследования методом DCM следует уточнить тип вывода, требуемый для данного вопроса. Этот выбор определяет последовательности шагов анализа данных [11].

Даже если исследователь заинтересован в выводе параметров модели, первым шагом обычно является Байесовский выбор модели (БВМ, англ. Bayesian Model Selection, BMS). BMS – это статистическая процедура по вычислению приближения доказательства модели (англ. model evidence).

Доказательство модели (или функцию маргинального правдоподобия) можно рассматривать как «святой грааль» сравнения моделей, которое количественно определяет свойства хорошей модели. В BMS модели обычно сравниваются по их байесовскому фактору, т. е. по соотношению их функций маргинального правдоподобия или, что эквивалентно, их разнице логарифмических доказательств.

Для вывода о пространстве модели BMS достаточно, но его можно применять по-разному. Можно, например, определить одну оптимальную модель либо разбить пространство моделей и сравнивать множества или семейства моделей, которые отличаются одним или несколькими структурными аспектами.

Выбор лучшей модели. Обычно на первом этапе специфицируются все вероятные модели, затем используется BMS для выбора оптимальной модели из всех альтернатив и, наконец, происходит оценка апостериорных и условных вероятностей параметров этой оптимальной модели. Для анализа данных одного субъекта вывод о каком-либо определенном параметре (или о линейных комбинациях параметров) довольно прост: нужно оценить апостериорную плотность интересующего параметра и количественно оценить вероятность того, что

значение параметра больше или меньше некоторого порога. Для групповых исследований существует два варианта в зависимости от того, предполагается ли, что интересующие параметры имеют фиксированное распределение среди всех участников эксперимента (фиксированные эффекты, англ. FFX, Fixed Effects) или сами вероятностно распределены в популяции (случайные эффекты, RFX, Random Effects) (используется для поиска патологий и заболеваний среди исследуемых).

Цель выбора модели состоит в том, чтобы определить оптимальную модель из набора правдоподобных альтернатив, которая является наиболее полезной, то есть представляет наилучший баланс между точностью и сложностью и, таким образом, обеспечивает максимальную обобщаемость. Чтобы не потеряться в пространстве бесконечных возможных моделей, требуется тщательно определять размер пространства моделей.

При FFX предположении полезной метрикой является групповой Байесовский фактор (англ. Group Bayes Factor, GBF), который выражает функцию маргинального правдоподобия одной модели относительно функции маргинального правдоподобия другой модели, рассматривая группу в целом. У GBF есть простое определение: поскольку коэффициенты Байеса являются вероятностями, которые независимы между субъектами, GBF является продуктом отдельных факторов Байеса. При сравнении более двух моделей проще для каждой модели определить групповую логарифмическую функцию маргинального правдоподобия, что является суммой логарифмических доказательств по всем субъектам. На практике самым простым и информативным способом отображения группового логарифмического доказательства является построение гистограммы доказательств для каждой модели после вычитания лог-доказательства для модели с наименьшими доказательствами.

Нужно помнить, что доказательство модели определено в отношении одного конкретного набора данных. Это означает, что BMS нельзя применять к моделям, которые относятся к разным наборам данным. Важно иметь в виду, что любой результат, полученный BMS или любой другой процедурой выбора модели, выражает относительное утверждение о добротности модели, которое обусловлено рассмотренным пространством моделей [12].

Оценка параметров лучшей модели. Для вывода о параметрах модели необходимо оценить их апостериорные плотности распределения. Однако эти апостериорные оценки обусловлены выбранной моделью. По этой причине BMS обычно является обязательным этапом, даже если гипотеза касается значений параметров модели, а не структуры модели как таковой.

Для анализа групповых оценок параметров при FFX предположении обычно используют метод усреднения байесовских параметров (англ. Bayesian parameter averaging, BPA). Он эффективно вычисляет

совместную апостериорную плотность для всей группы, рассматривая апостериорную плотность одного субъекта как априорную для следующего. Метод производит единственную апостериорную плотность для всей группы, которая может быть использована для байесовского вывода. Альтернативами ВРА считаются одномерный вариант ВРА и простое усреднение временных рядов испытуемых в качестве этапа предварительной обработки (что возможно только в том случае, если стимулы появляются в одно и то же время для всех субъектов) [13].

Альтернативным подходом для вывода о параметрах модели является байесовское усреднение модели (англ. Bayesian model averaging, ВМА). Этот подход отказывается от зависимости вывода о параметрах от выбранной модели. Вместо этого он использует все рассматриваемое пространство моделей и вычисляет средневзвешенные значения каждого параметра моделей, где взвешивание задается апостериорной вероятностью для каждой модели. Он представляет собой полезную альтернативу, когда ни одна из моделей не считается явно превосходящей все остальные. ВМА также используется для сравнения оценок параметров между группами в тех случаях, когда ВМС указала групповую разницу в отношении оптимальной модели.

4 Заключение

Работа посвящена исследованию методов проведения виртуального эксперимента. В качестве примера использована задача поиска эффективной связности фМРТ действия. В процессе исследования построена онтология предметной области, описаны характерные особенности данных, явно специфицированы используемые априорные гипотезы, формально описан поток работ поиска эффективной связности фМРТ человека.

Поддержка

Работа выполнена при поддержке РФФИ (грант 16-07-01028).

Литература

- [1] Hillebrandt, H., Friston, K.J. et al.: Effective Connectivity During Animacy Perception – Dynamic Causal Modelling of Human Connectome Project Data. *Scientific Reports*, 9 (4:6240) (2014).
- [2] Penny, W., Friston, K.J.: *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, 656 p. (2006)

- [3] Born, R.T., Bradley, D.C.: Structure and Function of Visual Aarea Mt. *Annual Reviews Neuroscience*, 28, pp. 157-189 (2005)
- [4] Castelli, F., Happe, F. et al.: Movement and Mind: A Functional Imaging Study of Perception and Interpretation of Complex Intentional Movement Patterns. *NeuroImage*, 12, pp. 314-325 (2000)
- [5] Van Essen, D. et al. The Human Connectome Project: A Data Acquisition Perspective. *NeuroImage*, 62, pp. 2222-2231 (2012)
- [6] Reference Manual – Q2 Data Release (June 2013) | WU-Minn Consortium of the NIH Human Connectome Project
http://www.humanconnectome.org/documentation/Q2/Q2_Release_Reference_Manual.pdf
- [7] Glasser, M.F. et al.: The Minimal Preprocessing Pipelines for the Human Connectome Project. *NeuroImage*, 80, pp. 105-124 (2013).
- [8] Van Essen, D., Ugurbil, K.: The Future of the Human Connectome [Review]. *NeuroImage*, 62 (2), pp. 1299-1310 (2012).
- [9] <http://mindhive.mit.edu/node/46>
- [10] Stephan, K.E. et al.: Ten Simple Rules for Dynamic Causal Modeling. *NeuroImage*, 49, pp. 3099-3109 (2010)
- [11] Kasess, C.H. et al.: Multi-Subject Analyses with Dynamic Causal Modeling. *NeuroImage*, 49, pp. 3065-3074 (2010)
- [12] Stephan, K.E., Penny, W.D. et al.: Bayesian Model Selection for Group Studies. *NeuroImage*, 46, pp. 1004-1017 (2010)
- [13] Rosa, M. J., Friston, K., Penny, W.: Post-hoc Selection of Dynamic Causal Models. *J. of Neuroscience Methods*, 208, pp. 66-78 (2012)
- [14] Schreiber, T. Measuring Information Transfer. *Physical Review Letters*, 85, p. 461 (2000)
- [15] Friston, K. Dynamic Causal Modeling and Granger Causality Comments on: The Identification of Interacting Networks. *The Brain Using fMRI: Model Selection, Causality and Deconvolution*. *Neuroimage*, 58, pp. 303-305 (2011)
- [16] Zhuang, J., LaConte, S. et al.: Connectivity Exploration with Structural Equation Modeling: an fMRI Study of Bimanual Motor Coordination. *NeuroImage*, 25, pp. 462-470 (2005)
- [17] Pearl, J.: The Causal Foundations of Structural Equation Modeling. Chapter for R.H. Hoyle (Ed.), *Handbook of Structural Equation Modeling*, Guilford Press, 5, pp. 68-91 (2012)
- [18] Friston, K.J., Harrison, L., Penny, W.: Dynamic Causal Modeling. *Neuroimage*, 19 (4), pp. 1273-1302 (2003)