



Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа: Инженерная школа ядерных технологий

Направление подготовки: 01.03.02 Прикладная математика и информатика

ООП/ОПОП: Прикладная математика в инженерии

Отделение школы (НОЦ): Экспериментальной физики

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

Тема работы
Оценка эффективности процедуры реабилитации после COVID-19

УДК 332.122(1-77)

Обучающийся

Группа	ФИО	Подпись	Дата
0В91	Скупченко Диана Денисовна		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Шинкеев Михаил Леонидович	канд. физ.- мат. наук		

Консультант (при наличии)

Должность	ФИО	Ученая степень, звание	Подпись	Дата

КОНСУЛЬТАНТЫ ПО РАЗДЕЛАМ:

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Кашук Ирина Вадимовна	к.т.н доцент		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Антоневич Ольга Алексеевна	к.б.н.		

Нормоконтроль (при наличии)

Должность	ФИО	Ученая степень, звание	Подпись	Дата

ДОПУСТИТЬ К ЗАЩИТЕ:

Руководитель ООП/ОПОП, должность	ФИО	Ученая степень, звание	Подпись	Дата
Руководитель ООП 01.03.02 «Прикладная математика и информатика»	Крицкий Олег Леонидович	канд. физ.- мат. наук		

Томск – 2023 г.



Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)

Школа: Инженерная школа ядерных технологий

Направление подготовки: 01.03.02 Прикладная математика и информатика

ООП/ОПОП: Прикладная математика в инженерии

Отделение школы (НОЦ): Экспериментальной физики

УТВЕРЖДАЮ:

Руководитель ООП/ОПОП

_____ Крицкий О.Л.
(Подпись) (Дата) (ФИО)

ЗАДАНИЕ на выполнение выпускной квалификационной работы

Обучающийся:

Группа	ФИО
0B91	Скупченко Диана Денисовна

Тема работы:

Оценка эффективности процедуры реабилитации после COVID-19
Утверждена приказом директора (дата, номер)

Срок сдачи обучающимся выполненной работы:	31.05.2023
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе (наименование объекта исследования или проектирования; производительность или нагрузка; режим работы (непрерывный, периодический, циклический и т. д.); вид сырья или материал изделия; требования к продукту, изделию или процессу; особые требования к функционированию (эксплуатации) объекта или изделия в плане безопасности эксплуатации, влияния на окружающую среду, энергозатратам; экономический анализ и т. д.)	База данных по COVID-19 реабилитации ФГБУ ФНКЦ МРиК ФМБА России
Перечень разделов пояснительной записки подлежащих исследованию, проектированию и разработке (аналитический обзор литературных источников с целью выяснения достижений мировой науки техники в рассматриваемой области; постановка задачи исследования, проектирования, конструирования; содержание процедуры исследования, проектирования, конструирования; обсуждение результатов выполненной работы; наименование дополнительных разделов, подлежащих разработке; заключение по работе)	1. Аналитический обзор литературных источников. 2. Отбор клинических, лабораторных и функциональных показателей, наиболее информативно характеризующих состояние здоровья пациентов. 3. Статистическое исследование значимости различия показателей до и после проведения различных процедур реабилитации. 4. Исследование влияния методик реабилитации на результаты.

	5. Кластеризация пациентов по основным информативным параметрам с целью прогнозирования реабилитационного эффекта.
Перечень графического материала (с точным указанием обязательных чертежей)	
Консультанты по разделам выпускной квалификационной работы (с указанием разделов)	
Раздел	Консультант
Социальная ответственность	Антоневич Ольга Алексеевна
Финансовый менеджмент	Кащук Ирина Вадимовна

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	
---	--

Задание выдал руководитель / консультант (при наличии):

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Шинкеев М.Л.	Канд. физ.-мат. наук		

Задание принял к исполнению обучающийся:

Группа	ФИО	Подпись	Дата
0В91	Скупченко Диана Денисовна		

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа		ФИО	
0B91		Скупченко Диана Денисовна	
Школа	ИЯТШ	Отделение (НОЦ)	ОЭФ
Уровень образования	Бакалавриат	Направление/специальность	01.03.02 Прикладная математика и информатика

Тема ВКР:

Оценка эффективности процедуры реабилитации после COVID-19	
Исходные данные к разделу «Социальная ответственность»:	
Введение – Характеристика объекта исследования (вещество, материал, прибор, алгоритм, методика) и области его применения. – Описание рабочей зоны (рабочего места) при разработке проектного решения/при эксплуатации	Объект исследования – пациенты, прошедшие реабилитацию после COVID-19 Область применения - медицина Рабочая зона: <u>офис</u> Размеры помещения - 24м² Количество и наименование оборудования рабочей зоны – ПК – 1шт. Рабочие процессы, связанные с объектом исследования, осуществляющиеся в рабочей зоне – программная разработка с использованием ПК
Перечень вопросов, подлежащих исследованию, проектированию и разработке:	
1. Правовые и организационные вопросы обеспечения безопасности при разработке проектного решения: – специальные (характерные при эксплуатации объекта исследования, проектируемой рабочей зоны) правовые нормы трудового законодательства; – организационные мероприятия при компоновке рабочей зоны.	–СП 2.2.3670-20 "Санитарно-эпидемиологические требования к условиям труда"; - ГОСТ 12.2.032-78 Рабочее место при выполнении работ сидя; – Трудовой кодекс Российской Федерации: федер. Закон от 30 дек. 2001 г. №197-ФЗ Раздел 10.
2. Производственная безопасность при разработке проектного решения: – Анализ выявленных вредных и опасных производственных факторов	Вредные факторы: - Отклонение показателей микроклимата; - Недостаточная освещённость рабочей зоны; - Повышенный уровень шума на рабочем месте; Опасные факторы: - Опасность поражения электрическим током;

	Требуемые средства коллективной и индивидуальной защиты от выявленных факторов: изоляция электрических проводов
3. Экологическая безопасность при разработке при эксплуатации:	Воздействие на селитебную зону: нет воздействия Воздействие на литосферу: утилизация ПК, макулатуры Воздействие на гидросферу: производство компьютерной техники Воздействие на атмосферу: выбросы углекислого газа при производстве компьютеров и его комплектующих
4. Безопасность в чрезвычайных ситуациях при разработке проектного решения:	Возможные ЧС: морозы; обрушение здания; пандемия; возникновение пожара Наиболее типичная ЧС: возникновение пожара
Дата выдачи задания для раздела по линейному графику	

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент	Антоневич Ольга Алексеевна	к.б.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0B91	Скупченко Диана Денисовна		

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

Группа	ФИО
0В91	Скупченко Диана Денисовна

Школа	ИЯТШ	Отделение школы (НОЦ)	ОЭФ
Уровень образования	Бакалавриат	Направление/специальность	01.02.03 Прикладная математика и информатика

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Стоимость материальных ресурсов и специального оборудования определены в соответствии с рыночными ценами г. Томска Тарифные ставки исполнителей определены штатным расписанием НИ ТПУ
2. Нормы и нормативы расходования ресурсов	Норма амортизационных отчислений на специальное оборудование
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Отчисления во внебюджетные фонды 30 %

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого потенциала инженерных решений (ИР)	Расчет конкурентоспособности SWOT-анализ
2. Формирование плана и графика разработки и внедрения ИР	Структура работ. Определение трудоемкости. Разработка графика проведения исследования
3. Обоснование необходимых инвестиций для разработки и внедрения ИР	Расчет бюджетной стоимости НИ
4. Составление бюджета инженерного проекта (ИП)	Интегральный финансовый показатель. Интегральный показатель ресурсоэффективности. Интегральный показатель эффективности.
5. Оценка ресурсной, финансовой, социальной, бюджетной эффективности ИР и потенциальных рисков	Расчет конкурентоспособности SWOT-анализ

Перечень графического материала (с точным указанием обязательных чертежей)

1. Оценка конкурентоспособности ИР
2. Матрица SWOT
3. Диаграмма Ганта
4. Бюджет НИ
5. Основные показатели эффективности НИ

Дата выдачи задания для раздела по линейному графику

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Доцент ОСГН ШБИП	Кащук Ирина Вадимовна	к.т.н доцент		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
0В91	Скупченко Диана Денисовна		

РЕФЕРАТ

Выпускная квалификационная работа «Оценка эффективности процедуры реабилитации после COVID-19» выполнена на 85 страницах, содержит 27 рисунков, 26 таблиц, 21 источник.

Ключевые слова: дисперсионный анализ, тест Стьюдента, критерий Вилкоксона, кластеризация, метод главных компонент.

Объектом исследования является база данных из 80-ти пациентов, переболевших COVID-19.

Цель работы – оценка эффективности процедур реабилитации после COVID-19.

В ходе работы проводился отбор признаков, наиболее информативно описывающих группы пациентов; исследование значимого различия показателей до и после, влияние методики реабилитации на результаты этих показателей; кластеризация испытуемых по основным параметрам.

Работа написана в Microsoft Word 2018. Для анализа в ходе исследования использовался программный пакет Statistica.

Оглавление

Введение	11
1 Основной раздел	12
Теоретическая часть	12
1.1.1 Дисперсионный анализ	12
1.1.2 Непараметрический дисперсионный анализ Краскала-Уоллиса.....	12
1.1.3 Многофакторный дисперсионный анализ	17
1.1.4 Дисперсионный анализ зависимых наблюдений	20
1.1.5 Метод главных компонент	23
1.1.6 Кластерный анализ.....	26
Практическая часть.....	31
1.2.1 Объект исследования	31
1.2.2 Отбор признаков для оценки реабилитации.....	33
1.2.3 Исследование эффективности процедур реабилитации по выбранным признакам и для различных групп пациентов.....	36
1.2.4 Анализ динамики физиологических функций (пробы Генчи и Штанге)	37
1.2.5 Анализ динамики функциональных функций (тест ТШХ).....	39
1.2.6 Анализ динамики физических нагрузок (шкала Борга)	41
1.2.7 Анализ динамики эмоционального интеллекта (тест EQ)	42
1.2.8 Анализ динамики функциональное состояние организма (индекс ИФИ).....	43
1.2.9 Анализ динамики физиологических функций (пробы Генчи и Штанге) по возрастным группам	44
1.2.10 Анализ динамики функциональных функций (тест ТШХ) по возрастным группам	46
1.2.11 Анализ динамики физических нагрузок (шкала Борга) по возрастным группам	46
1.2.12 Кластерный анализ.....	47
2. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	53
Введение	53
2.1 Оценка коммерческого потенциала и перспективности проведения исследований с позиции ресурсоэффективности и ресурсосбережения	53
2.1.1 Анализ конкурентных технических решений	53
2.1.2 SWOT-анализ.....	55
2.2. Планирование научно-исследовательских работ	58
2.2.1 Определение трудоемкости выполнения работ и разработка графика проведения	59
2.3 Бюджет научно-технического исследования	62
2.3.1 Расчет материальных затрат научно-технического исследования	62

2.3.2 Расчет амортизации специального оборудования	62
2.3.3 Основная заработная плата исполнителей темы.....	63
2.3.4 Отчисления во внебюджетные фонды (страховые отчисления)	65
2.3.5 Накладные расходы.....	65
2.3.6 Бюджет НИР	66
2.4 Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования.....	67
Заключение по разделу	69
3. Социальная ответственность	71
Введение	71
3.1 Правовые и организационные вопросы обеспечения безопасности.....	71
3.2 Производственная безопасность	72
3.2.1 Отклонение показателей микроклимата	73
3.2.2 Недостаточная освещенность рабочей зоны	75
3.2.3 Повышенный уровень шума на рабочем месте.....	76
3.2.4 Опасность поражения электрическим током	77
3.3 Экологическая безопасность	77
3.4 Безопасность в чрезвычайных ситуациях.....	78
3.4.1 Возникновение пожара	78
Заключение по разделу	79
Заключение.....	81
Список используемых источников	82

Введение

В недавнем времени не только Россия, но и весь мир пережил пандемию COVID-19, которая оказала колоссальное негативное влияние на здоровье многих людей. Даже после прохождения заболевания и выписки из больницы, некоторые пациенты продолжали испытывать проблемы со здоровьем и нуждались в реабилитации.

Данная работа посвящена оценке эффективности реабилитации пациентов после COVID-19 и определению лучших методов для достижения самых эффективных результатов в борьбе с последствиями заболевания. Результаты этого исследования могут быть полезны как для практикующих врачей, так и для специалистов в области здравоохранения, работающих с пациентами после COVID-19.

Целью выпускной квалификационной работы является оценка эффективности процедур реабилитации после COVID-19. Для ее достижения необходимо выполнить следующие задачи:

- 1) Провести анализ признаков, описывающих клинико-функциональную характеристику лиц, перенесших COVID-19;
- 2) Отобрать наиболее информативные признаки (или комбинации признаков) для оценки эффективности реабилитации пациентов после COVID-19;
- 3) Исследовать эффективность процедур реабилитации для различных групп пациентов;
- 4) Используя методы кластеризации, спрогнозировать эффективность процедуры реабилитации для различных групп пациентов.

1 Основной раздел

Теоретическая часть

1.1.1 Дисперсионный анализ

Дисперсионный анализ — статистический метод, позволяющий анализировать влияние различных факторов на результаты эксперимента путем исследования значимости различий в средних значениях. Суть дисперсионного анализа (analysis of variance — сокращенно ANOVA) заключается в разложении дисперсии измеряемого признака на независимые слагаемые, каждое из которых характеризует влияние того или иного фактора или их взаимодействия [1].

В терминах метода дисперсионного анализа та переменная, которая, должна оказывать влияние на конечный результат, называется фактором. Конкретное значение фактора называют уровнем фактора. Значение измеряемого признака называют откликом.

1.1.2 Непараметрический дисперсионный анализ Краскала-Уоллиса

В *однофакторной модели дисперсионного анализа* исходят из следующей модели порождения данных:

$$y_{ij} = \mu_j + \varepsilon_{ij} = \mu + \alpha_j + \varepsilon_{ij}, \quad i = \overline{1, n_j}, \quad j = \overline{1, k}, \quad (1.1)$$

где: y_{ij} - i -ое наблюдаемое значение отклика в j -ой группе (для j -го уровня фактора); μ - среднее значение отклика по всем уровням фактора (среднее по всей совокупности); μ_j - среднее значение отклика для j -го уровня фактора; $\alpha_j = \mu_j - \mu$ - дифференциальный эффект среднего, соответствующий j -му уровню фактора; ε_{ij} - независимые случайные величины с математическим ожиданием равным нулю и одинаковой дисперсией σ^2 .

При заданных величинах μ_j , величины α_j и μ определяются неоднозначно, поэтому необходимо наложить дополнительное условие, устанавливающее связь между ними.

Обычно используют одно из следующих условий:

$$\sum_{i=1}^k \alpha_i = 0, \sum_{i=1}^k n_i \alpha_i = 0, \alpha_k = 0 \quad (1.2)$$

Выражение $y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$ можно представить в виде

$$y_{ij} = \mu + (\mu_j - \mu) + (y_{ij} - \mu_j), \quad (1.3)$$

или:

$$y_{ij} - \mu = (\mu_j - \mu) + (y_{ij} - \mu_j) \quad (1.4)$$

Данное соотношение говорит о том, что отклонение наблюдаемого значения отклика для j -ой группы складывается из суммы двух слагаемых: отклонения отклика от среднего значения j -ой группы - $(y_{ij} - \mu_j)$, и отклонения среднего значения j -ой группы от среднего значения всей совокупности - $(\mu_j - \mu)$. Дисперсия отклика может быть представлена в виде суммы двух дисперсий, одна из которых характеризует внутригрупповую изменчивость, а вторая межгрупповую [1].

Разложение общей дисперсии на составляющие для выборочных данных обычно записывается в виде равенства сумм квадратов соответствующих отклонений:

$$SS_T = SS_B + SS_R, \quad (1.5)$$

где: $SS_T = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$ – общая, или полная, сумма квадратов

отклонений;

$$SS_B = \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y})^2 = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2 \quad - \text{ сумма квадратов отклонений}$$

групповых средних от общего среднего, или межгрупповая (межуровневая факторная) сумма квадратов отклонений (сумма квадратов эффекта фактора или просто эффект фактора);

$$SS_R = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \quad - \text{ сумма квадратов отклонений наблюдений от}$$

групповых средних, или внутригрупповая (остаточная) сумма квадратов отклонений, также называемая остаточным эффектом или эффектом ошибок;

k – число уровней фактора;

n_j – число наблюдений для j -го уровня фактора;

$$n = \sum_{j=1}^k n_j \quad - \text{ общее число наблюдений.}$$

В разложении дисперсии на составляющие заключена основная идея дисперсионного анализа: общая вариация переменной, порожденная влиянием фактора и измеренная суммой SS_T , складывается из двух компонент: SS_B и SS_R , характеризующих изменчивость этой переменной между уровнями фактора (SS_B) и внутри уровней фактора (SS_R) [1].

В дисперсионном анализе анализируются не сами суммы квадратов отклонений, а так называемые средние квадраты, которые получаются делением сумм квадратов отклонений на соответствующее число степеней свободы. Число степеней свободы для суммы квадратов случайных величин определяется как общее число линейно независимых слагаемых.

$$\text{Для полной суммы квадратов } SS_T = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 \quad \text{число степеней}$$

свободы $v_T = n - 1$, так как при ее расчете используются n наблюдений,

связанных между собой одним уравнением для общего выборочного среднего всей совокупности.

Для суммы квадратов эффекта фактора $SS_B = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2$ число степеней свободы $v_B = k - 1$, так как при ее расчете используются k групповых средних, связанных между собой также одним уравнением для общего выборочного среднего всей совокупности [1].

Для суммы квадратов ошибок $SS_R = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$ число степеней свободы $v_R = n - k$, ибо при его расчете используются n наблюдений, связанных между собой k уравнениями для выборочных средних k групп.

Соответственно выражения для средних квадратов отклонений, которые являются оценками соответствующих дисперсий, имеют вид:

1. $MS_T = \frac{1}{n-1} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$ - оценка общей дисперсии;
2. $MS_B = \frac{1}{k-1} \sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2$ - оценка межгрупповой дисперсии;
3. $MS_R = \frac{1}{n-k} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$ - оценка остаточной дисперсии σ^2 .

При условии истинности $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$, статистики MS_B и MS_R являются несмещенными оценками одной и той же дисперсии σ^2 .

Поэтому проверка гипотезы H_0 сводится к проверке гипотезы о равенстве двух дисперсий на основе F - отношения:

$$F = \frac{MS_B}{MS_R} = \frac{n-k}{k-1} \frac{SS_B}{SS_R} \quad (1.6)$$

При истинности H_0 статистика F в случае нормального распределения величин ε_{ij} имеет распределение Фишера с $\nu_1 = k - 1$ и $\nu_2 = n - k$ числом степеней свободы. Если наблюдаемое значение статистики $F_{набл} \geq F_{кр}$, где $F_{кр}$ - критическая точка распределения Фишера уровня α (или квантиль уровня $1 - \alpha$) с числом степеней свободы $\nu_1 = k - 1$ и $\nu_2 = n - k$, то нулевая гипотеза отклоняется и считается, что средние для различных уровней фактора значимо различаются [1].

Для порядковых данных непараметрической альтернативой однофакторного дисперсионного анализа является ранговый дисперсионный анализ Краскела–Уоллиса. В данном методе вместо значений переменных используется ранг переменных.

Если обозначить через R_{ij} ранг элемента x_{ij} , в общем вариационном ряду значений отклика, то величины $\bar{R}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} R_{ij}$ будут определять средние ранги для элементов j -ой группы, а величина $\bar{R} = \frac{1}{n} \sum_{i=1}^{n_j} \sum_{j=1}^k R_{ij} = \frac{n+1}{2}$ средний ранг всей совокупности. Величина $\sum_{j=1}^k n_j (\bar{R}_j - \bar{R})^2$ будет характеризовать межгрупповой разброс рангов.

При условии истинности гипотезы H_0 равенства средних рангов групп, статистика (1.7) будет иметь приближенно распределение Хи-квадрат с $k - 1$ степенью свободы:

$$H = \frac{12}{n(n+1)} \sum_{j=1}^k n_j (\bar{R}_j - \bar{R})^2 \quad (1.7)$$

Если наблюдаемое значение статистики $H_{набл} \geq H_{кр}$, где $H_{кр}$ - критическая точка распределения Хи-квадрат с числом степеней свободы $k - 1$

уровня α (или квантиль уровня $1-\alpha$), то нулевая гипотеза отклоняется и считается, что средние ранги для различных уровней фактора значимо различаются.

1.1.3 Многофакторный дисперсионный анализ

Если анализируется одновременное влияние двух и более различных факторов на результаты наблюдений, то используется *многофакторный дисперсионный анализ*.

Исследуем влияние на величину X двух факторов А и В, имеющих, соответственно k и m уровней. В двухфакторной модели дисперсионного анализа обычно исходят из следующей модели порождения данных:

$$y_{ijl} = \mu_{ij} + \varepsilon_{ijl} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijl}, \quad l = \overline{1, n_{ij}}, \quad i = \overline{1, k}, \quad j = \overline{1, m}, \quad (1.8)$$

где: y_{ijl} - l -ое наблюдаемое значение отклика для i -го уровня фактора А и j -го уровня фактора В;

μ - среднее значение отклика по всей совокупности (генеральное среднее);

μ_{ij} - среднее значение отклика для i -го уровня фактора А и j -го уровня фактора В;

$\alpha_i = \mu_{i*} - \mu$ - главный эффект i -го уровня фактора А (μ_{i*} - среднее значение отклика для i -го уровня фактора А);

$\beta_j = \mu_{*j} - \mu$ - главный эффект j -го уровня фактора В (μ_{*j} - среднее значение отклика для j -го уровня фактора В);

$\gamma_{ij} = \mu_{ij} - \mu_{i*} - \mu_{*j} + \mu$ - эффект взаимодействия i -го уровня фактора А и j -го уровня фактора В;

ε_{ijl} - независимые случайные величины с математическим ожиданием равным нулю и одинаковой дисперсией σ^2 .

Средние могут определяться через величины μ_{ij} , например, как, взвешенные средние:

$$\mu = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^m \sum_{l=1}^{n_{ij}} M(y_{ijl}) = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^m n_{ij} \mu_{ij} \quad (1.9)$$

$$\mu_{i*} = \frac{1}{n_{i*}} \sum_{j=1}^m \sum_{l=1}^{n_{ij}} M(y_{ijl}) = \frac{1}{n_{i*}} \sum_{j=1}^m n_{ij} \mu_{ij} \quad (1.10)$$

$$\mu_{*j} = \frac{1}{n_{*j}} \sum_{i=1}^k \sum_{l=1}^{n_{ij}} M(y_{ijl}) = \frac{1}{n_{*j}} \sum_{i=1}^k n_{ij} \mu_{ij} \quad (1.11)$$

$$n = \sum_{i=1}^k \sum_{j=1}^m n_{ij}, \quad n_{i*} = \sum_{j=1}^m n_{ij}, \quad n_{*j} = \sum_{i=1}^k n_{ij} \quad (1.12)$$

Возможны и другие варианты определения средних величин.

Выражение $y_{ijl} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijl}$ можно представить в виде:

$$y_{ijl} - \mu = (\mu_{i*} - \mu) + (\mu_{*j} - \mu) + (\mu_{ij} - \mu_{i*} - \mu_{*j} + \mu) + (x_{ijl} - \mu_{ij}) \quad (1.13)$$

Данное соотношение говорит о том, что отклонение наблюдаемого значения отклика складывается из суммы четырех слагаемых: отклонения отклика от среднего значения для i, j -го набора уровней факторов А и В ($y_{ijl} - \mu_{ij}$), главных эффектов i -го уровня фактора А и j -го уровня фактора В и эффекта взаимодействия [1].

Дисперсия отклика может быть представлена в виде суммы четырех дисперсий, одна из которых характеризует внутригрупповую изменчивость для i, j -го набора уровней факторов А и В, а остальные соответствующие эффекты.

МНК оценки параметров модели двухфакторного дисперсионного анализа достаточно просто получить только в случае пропорциональных частот, то есть при условии, что:

$$\frac{n_{ij}}{n} = \frac{n_{i*}}{n} \cdot \frac{n_{*j}}{n} \quad \forall i = \overline{1, k}, j = \overline{1, m} \quad (1.14)$$

Данные условия выполняются, например, если количества наблюдений n_{ij} для каждого сочетания уровней факторов совпадают. План эксперимента подобного вида (или сам эксперимент) называют сбалансированным. В случае

пропорциональных частот получаем “интуитивно” понятные МНК-оценки для параметров модели вне зависимости от условий, накладываемых на параметры модели:

$$\hat{\mu} = \bar{y}, \hat{\alpha}_i = \bar{y}_{i*} - \bar{y}, \hat{\beta}_j = \bar{y}_{*j} - \bar{y}, \hat{\gamma}_{ij} = \bar{y}_{ij} - (\bar{y}_{i*} + \bar{y}_{*j}) + \bar{y} \quad (1.15)$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^m \sum_{l=1}^{n_{ij}} y_{ijl}, \bar{y}_{i*} = \frac{1}{n_{i*}} \sum_{j=1}^m \sum_{l=1}^{n_{ij}} y_{ijl}, \bar{y}_{*j} = \frac{1}{n_{*j}} \sum_{i=1}^k \sum_{l=1}^{n_{ij}} y_{ijl}, \bar{y}_{ij} = \frac{1}{n_{ij}} \sum_{l=1}^{n_{ij}} y_{ijl} \quad (1.16)$$

В случае пропорциональных также справедливо следующее разложение общей суммы квадратов отклонений на составляющие:

$$SS_T = SS_A + SS_B + SS_{AB} + SS_R, \quad (1.17)$$

где: $SS_T = \sum_{i=1}^k \sum_{j=1}^m \sum_{l=1}^{n_{ij}} (y_{ijl} - \bar{y})^2$ – общая, или полная, сумма квадратов

отклонений;

$$SS_A = \sum_{i=1}^k \sum_{j=1}^m \sum_{l=1}^{n_{ij}} (\bar{y}_{i*} - \bar{y})^2 = \sum_{i=1}^k n_{i*} (\bar{y}_{i*} - \bar{y})^2$$
 – сумма квадратов отклонений

средних по уровням фактора А от общего среднего, или сумма квадратов главных эффектов А (можно и так: сумма квадратов, соответствующих эффекту фактора А);

$$SS_B = \sum_{i=1}^k \sum_{j=1}^m \sum_{l=1}^{n_{ij}} (\bar{y}_{*j} - \bar{y})^2 = \sum_{j=1}^m n_{*j} (\bar{y}_{*j} - \bar{y})^2$$
 – сумма квадратов отклонений

средних по уровням фактора В от общего среднего, или сумма квадратов главных эффектов В;

$$SS_{AB} = \sum_{i=1}^k \sum_{j=1}^m \sum_{l=1}^{n_{ij}} (\bar{y}_{ij} - \bar{y}_{*j} - \bar{y}_{i*} + \bar{y})^2 = \sum_{i=1}^k \sum_{j=1}^m n_{ij} (\bar{y}_{ij} - \bar{y}_{*j} - \bar{y}_{i*} + \bar{y})^2$$
 – сумма

квадратов взаимодействия эффектов А и В;

$$SS_R = \sum_{i=1}^k \sum_{j=1}^m \sum_{l=1}^{n_{ij}} (y_{ijl} - \bar{y}_{ij})^2$$
 – остаточная сумма квадратов отклонений [1].

Число степеней свободы для сумм квадратов SS_A и SS_B равно соответственно $\nu_A = k - 1$ и $\nu_B = m - 1$. Число степеней свободы для суммы

квадратов взаимодействия эффектов SS_{AB} равно $\nu_{AB} = (k-1)(m-1)$. Число степеней свободы суммы квадратов остатков SS_R равно $\nu_R = n - km$. Соответственно средние суммы квадратов будут равны:

$$MS_A = \frac{SS_A}{k-1}, MS_B = \frac{SS_B}{m-1}, MS_{AB} = \frac{SS_{AB}}{(k-1)(m-1)}, MS_R = \frac{SS_R}{n - km} \quad (1.18)$$

Поскольку двухфакторная модель учитывает различные эффекты влияния факторов, то и статистический анализ для двухфакторной модели предполагает проверку гипотез о значимости различных эффектов. При условии истинности H_0 : «эффект незначим» средний квадрат эффекта является несмещенной оценкой дисперсии σ^2 , так же как и величина MS_R . Поэтому в качестве статистик критериев проверки гипотез о значимости соответствующих эффектов можно использовать отношения средней суммы квадратов эффектов к средней сумме квадратов остатков. При нормальном распределении остатков данные статистики имеют распределение Фишера с параметрами, определяемыми числами степеней свободы соответствующих сумм, участвующих в отношении [1].

Если наблюдаемое значение статистики $F_{набл} \geq F_{кр}$, где $F_{кр}$ - критическая точка распределения Фишера уровня α (или квантиль уровня $1 - \alpha$) с числом степеней свободы ν_1 и ν_2 , то нулевая гипотеза отклоняется и считается, что средние для различных уровней фактора значимо различаются.

1.1.4 Дисперсионный анализ зависимых наблюдений

Часто анализируются данные, в которых значения, соответствующие различным уровням фактора, соответствуют одним и тем же объектам. Данные повторных наблюдений уменьшают общую дисперсию данных за счет исключения составляющей индивидуальных различий из остаточной дисперсии, что позволяет повысить мощность критерия.

Пусть имеется k выборок, объемом n каждая, соответствующих k различным уровням фактора W , влияние которого на значения наблюдаемой переменной мы исследуем. Предполагается, что j -ое наблюдаемое значение

отклика в каждой группе соответствует одному и тому же объекту наблюдения. В простейшей однофакторной модели дисперсионного анализа повторных измерений исходят из следующей модели порождения данных:

$$y_{ij} = \mu_{ij} + \varepsilon_{ij} = \mu + w_i + \tau_j + \varepsilon_{ij} \quad i = \overline{1, k}, \quad j = \overline{1, n} \quad (1.19)$$

где: y_{ij} - j -ое наблюдаемое значение отклика в i -ой группе;

μ_{ij} - среднее значение отклика для i -го наблюдения в j -ой группе;

ε_{ij} - независимые случайные величины с математическим ожиданием равным нулю и одинаковой дисперсией σ^2 ;

μ - среднее значение отклика;

$w_i = \mu_{i*} - \mu$ - эффект i -го уровня фактора W ;

$\mu_{i*} = \frac{1}{n} \sum_{j=1}^n \mu_{ij}$ - среднее для i -го уровня фактора;

$\tau_j = \mu_{*j} - \mu$ - эффект индивидуальности j -го наблюдения,;

$\mu_{*j} = \frac{1}{k} \sum_{i=1}^k \mu_{ij}$ - среднее для j -го объекта наблюдения;

ε_{ij} - независимые случайные величины с математическим ожиданием равным нулю и одинаковой дисперсией σ^2 .

Уравнения (1.19) можно рассматривать как уравнения, соответствующие классической модели двухфакторного дисперсионного анализа, без компоненты, учитывающей взаимодействие факторов и числом наблюдений для каждой ячейки $n_{ij} = 1$ (что соответствует сбалансированному плану) [1].

Уравнения (1.19) можно переписать в виде:

$$y_{ij} - \mu_{*j} = (\mu_{i*} - \mu) + \varepsilon_{ij} = w_i + \varepsilon_{ij}, \quad i = \overline{1, k}, \quad j = \overline{1, n} \quad (1.20)$$

Таким образом, чтобы исключить эффект индивидуальности наблюдений и анализировать только влияние фактора W на результаты наблюдений, необходимо перейти к рассмотрению величин $y_{ij} - \mu_{*j}$.

Оценкой параметра μ_{*j} является выборочное среднее $\bar{y}_{*j} = \frac{1}{k} \sum_{i=1}^k y_{ij}$. Соответственно, сумма квадратов наблюдений за вычетом суммы квадратов, соответствующих индивидуальным различиям, будет равна $\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{*j})^2$.

Если обозначить $z_{ij} = y_{ij} - \bar{y}_{*j}$, то разложение для суммы квадратов будет иметь вид:

$$\sum_{i=1}^k \sum_{j=1}^n z_{ij}^2 = \sum_{i=1}^k n \bar{z}_i^2 + \sum_{i=1}^k \sum_{j=1}^n (z_{ij} - \bar{z}_i)^2 \quad (1.21)$$

Число степеней свободы для суммы $SS_T = \sum_{i=1}^k \sum_{j=1}^n z_{ij}^2$ равно $nk - n$, для суммы $SS_W = \sum_{i=1}^k n(\bar{z}_i - \bar{z})^2 = \sum_{i=1}^k n \bar{z}_i^2$ равно $k - 1$, для суммы $SS_R = \sum_{i=1}^k \sum_{j=1}^n (z_{ij} - \bar{z}_i)^2$ равно $nk - n - k + 1 = (n - 1)(k - 1)$.

Средние суммы квадратов отклонений будут равны:

$$MS_W = \frac{SS_W}{k - 1} = \frac{1}{k - 1} \sum_{i=1}^k n \bar{z}_i^2, \quad MS_R = \frac{SS_R}{(n - 1)(k - 1)} = \frac{1}{(n - 1)(k - 1)} \sum_{i=1}^k \sum_{j=1}^n (z_{ij} - \bar{z}_i)^2 \quad (1.22)$$

F-статистика для проверки гипотезы $H_0 : w_1 = w_2 = \dots = w_k = 0$ будет

иметь вид: $F = \frac{MS_W}{MS_R}$. Если наблюдаемое значение статистики $F_{набл} \geq F_{кр}$, где $F_{кр}$ - критическая точка распределения Фишера уровня α (или квантиль уровня $1 - \alpha$) с числом степеней свободы $\nu_1 = k - 1$ и $\nu_2 = (n - 1)(k - 1)$, то

нулевая гипотеза отклоняется и считается, что средние для различных уровней фактора W значимо различаются [1].

1.1.5 Метод главных компонент

Метод главных компонент осуществляет переход к новой совокупности некоррелированных признаков $\eta_1, \dots, \eta_k$, каждый из которых является линейной комбинацией исходных признаков $\xi_1, \xi_2, \dots, \xi_k$. При этом линейные комбинации выбираются таким образом, что среди всех возможных линейных нормированных комбинаций исходных признаков первая главная компонента η_1 обладает наибольшей дисперсией [3].

В дальнейшем из полученных величин можно оставить только m ($m < k$) наиболее значимых факторов, вносящих максимальный вклад в суммарную дисперсию и использовать эти величины как некие интегральные факторы, характеризующие всю совокупность признаков.

Пусть $\xi = (\xi_1, \xi_2, \dots, \xi_k)$ – центрированная многомерная случайная величина с матрицей ковариаций A .

Положим:

$$\eta_i = \beta^{(i)T} \xi = \beta_1^{(i)} \xi_1 + \beta_2^{(i)} \xi_2 + \dots + \beta_k^{(i)} \xi_k, \quad (1.23)$$

где $\beta^{(i)}$ – векторы неизвестных коэффициентов преобразования, $i = \overline{1, m}$, ($m \leq k$).

Будем называть величины η_i главными компонентами (факторами). Так как величины ξ_i центрированы, то $M(\eta_i) = 0, i = \overline{1, m}$, а дисперсия главных компонент определяется как:

$$D(\eta_i) = E(\eta_i \eta_i^T) = \beta^{(i)T} E(\xi \xi^T) \beta^{(i)} = \beta^{(i)T} A \beta^{(i)} \quad (1.24)$$

На коэффициенты преобразования $\beta^{(i)}$ накладываем условиям нормировки:

$$\|\beta^{(i)}\|^2 = \beta^{(i)T} \beta^{(i)} = 1 \quad (1.25)$$

Задача отыскания главных компонент сводится к задаче на собственные значения и собственные векторы матрицы ковариаций A . Коэффициенты преобразования $\beta^{(i)}$, $i = \overline{1, k}$ удовлетворяющие условиям задачи поиска главных компонент являются собственными векторами матрицы ковариаций A , соответствующими собственными значения $\lambda_1, \lambda_2, \dots, \lambda_k$ упорядоченным по убыванию значений, причем $D(\eta_i) = \lambda_i$, а $\sum_{i=1}^k D(\xi_i) = \sum_{i=1}^k D(\eta_i) = \sum_{i=1}^k \lambda_i$.

Оценка главных компонент на основе выборочных данных строится на основе выборочной матрицы ковариаций. Оценки собственных значений, являющиеся собственными числами выборочной матрицы ковариаций, в случае нормального распределения генеральной совокупности являются оценками максимального правдоподобия. Если единицы измерения исходных признаков различаются или их значения сильно различаются, то лучше использовать при нахождении оценок главных компонент вместо выборочной матрицы ковариаций выборочную корреляционную матрицу [3].

Математически строгих критериев отбора компонент не существует.

Обычно используют один из следующих эвристических методов.

Во первых, зная $sp \bar{A} = \sum_{i=1}^k \lambda_i$, можно выбрать те компоненты $\eta_i, i = \overline{1, m}$

из общего набора, которые бы объясняли не менее некоторой заданной доли q суммарной доли дисперсии признаков. Обычно q полагают не менее 0,7.

Другим критерием отбора является критерий Кайзера, который предполагает использование для нахождения оценок собственных значений выборочной матрицы корреляций (можно применять и в случае использования матрицы ковариаций, следует лишь в этом случае нормировать (умножить) все собственные значения на величину $k / sp(\bar{A})$). Согласно данному критерию оставляют только те главные компоненты, дисперсия которых больше 1. Это означает, что если фактор не выделяет дисперсию, эквивалентную, по крайней мере, дисперсии одной переменной, то он опускается [3].

Подход к оценке числа главных компонент по необходимой доле объяснённой дисперсии формально применим всегда, однако неявно он предполагает, что нет разделения на «сигнал» и «шум», и любая заранее заданная точность имеет смысл. Поэтому часто более продуктивна иная эвристика, основывающаяся на гипотезе о наличии «сигнала» (сравнительно малая размерность, относительно большая амплитуда) и «шума» (большая размерность, относительно малая амплитуда). С этой точки зрения метод главных компонент работает как фильтр: сигнал содержится, в основном, в проекции на первые главные компоненты, а в остальных компонентах пропорция шума намного выше.

Одним из наиболее популярных подходов оценки числа необходимых главных компонент является правило сломанной трости (англ. Broken stick model). Набор нормированных собственных чисел $(\lambda_i^* / \text{tr} \bar{A}, i = \overline{1, k})$ сравнивается с распределением длин обломков трости единичной длины, сломанной в $k-1$ случайно выбранной точке (точки разлома выбираются независимо и равномерно распределенными по длине трости) [3].

Пусть $l_i (i = \overline{1, k})$ - длины полученных кусков трости, занумерованные в порядке убывания длины: $l_1 \geq l_2 \geq \dots \geq l_k$.

Тогда математическое ожидание l_i :

$$L_i = M(l_i) = \frac{1}{k} \sum_{j=i}^k \frac{1}{j} \quad (1.26)$$

По правилу сломанной трости i -й собственный вектор (в порядке убывания собственных чисел λ_i) сохраняется в списке главных компонент, если $\frac{\lambda_1^*}{\text{tr} \bar{A}} > L_1, \frac{\lambda_2^*}{\text{tr} \bar{A}} > L_2, \dots, \frac{\lambda_i^*}{\text{tr} \bar{A}} > L_i$. К подобным критериям относится также графический критерий каменистой осыпи Кэттелла. Он состоит в поиске точки, где убывание собственных значений замедляется наиболее сильно. Число выделенных факторов не должно превышать количество факторов слева от этой точки.

1.1.6 Кластерный анализ

Кластерный анализ – это анализ, позволяющий получить разбиение большого объема данных на классы или группы согласно некоторому критерию или их совокупности. При этом объединяются в классы объекты близкие между собой в определенном смысле, причем предполагается отсутствие априорной информации о характеристиках объектов внутри классов [8].

В качестве основных задач, решаемых с помощью кластерного анализа, можно выделить:

1. Классификацию объектов с учетом большого числа признаков, отражающих их сущность;
2. Выявление скрытой структуры изучаемой совокупности объектов;
3. Снижение размерности пространства изучаемых объектов.

МЕРЫ БЛИЗОСТИ ОБЪЕКТОВ

Пусть дана совокупность из n объектов, каждый из которых характеризуется k признаками $\xi_1, \xi_2, \dots, \xi_k$. Обозначим: $X^{(i)} = \{X_1^{(i)}, X_2^{(i)}, \dots, X_k^{(i)}\}$ – вектор значений признаков для i -ого объекта, $X_j^{(i)}$ – значение j -го признака для i -ого объекта. Определим меру близости (или меру однородности) объектов. В качестве меры близости будем использовать расстояние между объектами. Расстоянием, или метрикой, определенной на множестве M называется числовая функция $\rho(X, Y)$, $X, Y \in M$, удовлетворяющая следующим условиям:

- 1) $\rho(X, Y) \geq 0$, $\rho(X, Y) = 0 \Leftrightarrow X = Y$;
- 2) $\rho(X, Y) = \rho(Y, X)$;
- 3) $\rho(X, Y) \leq \rho(X, Z) + \rho(Z, Y)$.

В теории кластерного анализа используются следующие метрики для измерения расстояния между отдельными объектами (точками в k -мерном пространстве):

Евклидово расстояние:

$$\rho(X^{(i)}, X^{(j)}) = \sqrt{\sum_{s=1}^k (X_s^{(i)} - X_s^{(j)})^2} \quad (1.27)$$

Используется, если признаки однородны по физическому смыслу и их значения одинаково “весомы”.

Квадрат евклидова расстояния:

$$\rho(X^{(i)}, X^{(j)}) = \sum_{s=1}^k (X_s^{(i)} - X_s^{(j)})^2 \quad (1.28)$$

Применяется для придания большего веса более отдаленным друг от друга объектам.

Оценка сходства между объектами сильно зависит от абсолютного значения признака и от степени его вариации в совокупности.

ИЕРАРХИЧЕСКИЕ АГЛОМЕРАТИВНЫЕ МЕТОДЫ КЛАСТЕРНОГО АНАЛИЗА

Иерархические (т.е. деревообразующие) методы являются наиболее простыми и распространенными среди алгоритмов кластерного анализа. На первом шаге каждый из объектов $X^{(1)}, \dots, X^{(n)}$ исходных данных рассматривается как отдельный кластер. По вычисленной матрице расстояний $R = \|\rho(X^{(i)}, X^{(j)})\| = \|\rho_{ij}\|$, где ρ – некоторая метрика, на каждом шаге объединяются наиболее близкие друг к другу кластеры. При этом, каждый раз пересчитывается матрица расстояний. Очевидно, что процесс завершится через $(n-1)$ шаг, когда в результате все объекты будут объединены в один кластер. Процесс объединения может быть прерван на каком-либо шаге, если изначально задать пороговое максимальное значение расстояния для объединения кластеров [8].

В зависимости от того, как определяется расстояние между кластерами (как происходит объединение кластеров), выделяют методы одиночной и полной связи, методы взвешенных и невзвешенных попарных средних, взвешенный и невзвешенный центроидные методы, метод Уорда.

МЕТОД ОДИНОЧНОЙ СВЯЗИ

В методе одиночной связи расстояние между кластерами определяется по методу “ближнего соседа”, то есть за расстояние между кластерами принимается расстояние между самыми близкими элементами двух кластеров:

$$\rho(K_i, K_j) = \min_{\substack{X_s \in K_i \\ X_k \in K_j}} \rho(X^{(s)}, X^{(k)}) \quad (1.29)$$

К достоинствам данного метода следует отнести то, что он не чувствителен к монотонному преобразованию матрицы расстояний между исходными объектами. К недостаткам следует отнести то, что зачастую данный метод приводит к появлению длинных цепочек кластеров [8].

МЕТОД ПОЛНОЙ СВЯЗИ

В методе полной связи расстояние между кластерами вычисляется по принципу «дальнего соседа», то есть за расстояние между кластерами принимается расстояние между самыми удаленными элементами двух кластеров:

$$\rho(K_i, K_j) = \max_{\substack{X_s \in K_i \\ X_k \in K_j}} \rho(X^{(s)}, X^{(k)}) \quad (1.30)$$

МЕТОД УОРДА

В методе Уорда объединение кластеров происходит таким образом, что сумма квадратов расстояний объектов до центров кластеров получает наименьшее приращение. То есть, для каждого кластера K_s , состоящего из объектов $X^{(1)}, \dots, X^{(n_s)}$, в соответствии с выбранной метрикой, вычисляется величина $D_s = \sum_{i=1}^{n_s} \rho^2(X^{(i)}, \bar{X})$, где $\bar{X}$ - центр кластера K_s . Данная величина, по сути, характеризует внутрикластерную дисперсию для кластера K_s . Для совокупности кластеров $\{K_1, K_2, \dots, K_q\}$, соответственно, вычисляется суммарная сумма квадратов отклонений $V_q = \sum_{s=1}^q D_s$. На первом шаге все величины $D_s = 0, s = \overline{1, n}$, и на каждом шаге объединяются те кластеры, которые приводят к наименьшему изменению величины V_q после объединения (для этого потребуется рассмотреть все возможные объединения кластеров).

Метод Уорда на каждом шаге приводит к формированию кластеров с наименьшей внутрикластерной дисперсией.

ФОРМУЛА ЛАНСА-УИЛЬЯМСА

В общем случае процесс кластеризации – весьма трудоемкий процесс. Число необходимых операций можно уменьшить, если на каждом шаге использовать формулы, позволяющие подсчитывать расстояние от образовавшегося кластера $w = u \cup v$ до любого другого кластера s , используя известные расстояния с предыдущих шагов. В частности, это можно сделать, с помощью формулы, предложенной Лансом и Уильямсом в 1967 году:

$$\rho(w, s) = \alpha_u \rho(u, s) + \alpha_v \rho(v, s) + \beta \rho(u, v) + \gamma |\rho(u, s) - \rho(v, s)| \quad (1.31)$$

где $\alpha_u, \alpha_v, \beta, \gamma$ - числовые параметры.

Расстояние между кластерами для рассмотренных выше методов кластеризации удовлетворяет формуле Ланса-Уильямса со следующими коэффициентами:

Метод одиночной связи (англ. *single linkage*)

$$\alpha_u = 1/2, \alpha_v = 1/2, \beta = 0, \gamma = -1/2$$

Метод полной связи (англ. *complete linkage*)

$$\alpha_u = 1/2, \alpha_v = 1/2, \beta = 0, \gamma = 1/2$$

Метод Уорда (англ. *Ward's method*)

$$\alpha_u = \frac{n_u + n_s}{n_u + n_v + n_s}, \alpha_v = \frac{n_v + n_s}{n_u + n_v + n_s}, \beta = -\frac{n_s}{(n_u + n_v + n_s)^2}, \gamma = 0$$

ИТЕРАТИВНЫЕ МЕТОДЫ КЛАСТЕРИЗАЦИИ

В итеративных методах кластерного анализа на каждой итерации каждый из исходных объектов относится к одному из кластеров в соответствии с заданными условиями кластеризации, при этом одновременно происходит перерасчет параметров кластеров. Затем процесс кластеризации повторяется до тех пор, пока не получатся устойчивые образования кластеров, не меняющиеся от итерации к итерации. В качестве условий кластеризации используются, например, число кластеров, которые необходимо получить,

предельный радиус кластеров и т.д. Данные условия выбираются обычно по результатам иерархических методов [8].

МЕТОД К – СРЕДНИХ

Пусть необходимо совокупность объектов $X^{(1)}, \dots, X^{(n)}$ разбить на k кластеров.

1) На первом шаге из n векторов случайным образом (или не случайным) выбираем k из них, считая, что каждый из них образует центр кластера.

2) На следующем шаге каждый вектор из оставшегося набора данных относится к тому из кластеров, расстояние до центра, которого от данного вектора минимально.

3) Пересчитываем центры кластеров.

4) Повторяем пункты 2 и 3 до тех пор, пока не получится устойчивое разбиение, то есть составы кластеров на i -ой и $i+1$ -ой итерации не будут различаться.

Представленный алгоритм называется алгоритмом Болла-Холла. Существует разновидность алгоритма, при которой центр кластера пересчитывается при каждом добавлении в кластер нового вектора (вариант Маккуина) [8].

Можно показать, что при большом n и малом k алгоритм быстро сходится к устойчивому решению, т.е. к решению, в котором кластеры, полученные после применения алгоритма на i -ой итерации, совпадают по количеству и составу с кластерами, найденными на $i+1$ -ой итерации. Тем не менее, алгоритм применяют несколько раз, используя различные начальные векторы в качестве центров кластеров и, если результаты различаются, выбирают оптимальное разбиение, используя один из критериев качества разбиения. Также при определении порядка выбора объектов на каждой итерации, рекомендуется этот порядок определять случайным образом [8].

Для проверки значимости различия используем критерий Фишера. В предположении нормальности исходных данных статистика критерия:

$F = \frac{SS_B}{q-1} \frac{n-q}{SS_W}$, при истинности $H_0 : SS_B = SS_W$ приближенно имеет распределение

Фишера с $q-1, n-q$ степенями свободы. Если наблюдаемое значение статистики превышает критическое для заданного уровня значимости, то H_0 отвергается и различие признаков для классов признается значимым. Чем больше признаков, значения которых значимо различаются для различных классов, тем лучше кластеризация.

Практическая часть

1.2.1 Объект исследования

В процессе выполнения Государственного задания Филиалом ТНИИКИФ ФГБУ СибФНКЦ ФМБА России проведено обследование 80 пациентов, перенесших COVID -19, находящихся на стационарном этапе медицинской реабилитации (30-40 дней после перенесенной инфекции). Из них 71,87% женщины, 28,13% мужчин, средний возраст $56,92 \pm 9,29$ лет.

Исследовательские процедуры пациентам осуществлены в полном объеме. Отклонений от утвержденных протоколов обследования и реабилитации не зарегистрировано.

Среди обследованных пациентов, перенесших COVID-19, наиболее многочисленной была группа лиц в возрасте 44-60 и 60-75 (таблица 1), при этом, согласно новой пенсионной реформе, на трудоспособный возраст (от 16 до 65 – для мужчин и от 16 до 60 – для женщин) приходилось 42 (65,63%) человек, для которых восстановление трудоспособности является важнейшей задачей.

Таблица 1 – Распределение обследованных больных в зависимости от возраста

Возраст, лет	(n= 80)	
	абс.	%

Молодой 18-44	10	12,5
Средний 45-60	40	50
Пожилой 60-75	30	37,5

При первичном ЭКГ исследовании различные патологические изменения были выявлены у 42 (65,63%) пациентов, поступивших на стационарный этап медицинской реабилитации (таблица 2).

Таблица 2 – Результаты ЭКГ исследования лиц, перенесших COVID -19, при поступлении на стационарный этап медицинской реабилитации, абс. (%)

ЭКГ - паттерн	n=80
Нарушение процессов реполяризации	10 (12,5)
Неспецифические нарушения внутрижелудочковой проводимости	15 (18,75)
Нарушения ритма (брадикардия, тахикардия, единичные наджелудочковые экстрасистолы)	27 (33,75)
Нарушения внутрипредсердной проводимости	15 (18,75)
Признаки гипертрофии ЛЖ	15 (18,75)

Всем больным, поступившим на стационарный этап медицинской реабилитации, проводился ТШХ. На момент поступления среднее значение пройденной дистанции ТШХ составило $347,93 \pm 57,52$ м., что свидетельствует о снижении толерантности к физической нагрузке, при этом основной причиной прекращения проведения теста стали жалобы на выраженную слабость, усталость (n= 38, 59,38%) и одышку (n= 26, 40,6%).

В структуре жалоб обследованных пациентов значительную часть занимали жалобы на общую слабость, быструю утомляемость (100%, n=80), при степени выраженности 2,00 [2,00; 2,00] балла. Одышка при обычной

физической нагрузке, оцениваемая по шкале mMRC, беспокоила абсолютное большинство пациентов (96,8%), при этом средняя степень тяжести одышки диагностирована у 70,3% обследованных, легкая степень тяжести – у 17,2%, тяжелая одышка у 9,4%. Не предъявляли жалобы на одышку только 2 пациента.

1.2.2 Отбор признаков для оценки реабилитации

Осуществление критического анализа литературных источников, в том числе нормативных документов Минздрава и Минтруда Российской Федерации, методических документов международных медицинских организаций, научных публикаций [30,31,73,74,75] позволило определить наиболее показательные категории МКФ (Международной классификации функционирования), позволяющих дать объективную оценку состояния здоровья пациентов, перенесших COVID-19. Были сформированы критерии оценки физиологических функций по МКФ, с добавлением МКФ определителя. Использование такого подхода позволит осуществлять не только синдромальную, но и интегральную оценку РД и результативности реабилитационных мероприятий. При этом для всех измеряемых параметров были выбраны тесты, пробы, оценочные шкалы, опросники и т.п., которые являются качественными и доступными методами оперативного контроля эффективности восстановительного лечения и которые реально можно использовать на стационарном этапе медицинской реабилитации.

Была построена матрица корреляций информативных признаков (рис. 1), чтобы действительно убедиться в корректности информативности параметров и эффективности их результатов.

По матрице можем заметить, что есть признаки, совсем не коррелирующие между собой (например, шкала Борга и тест на эмоциональный интеллект), есть признаки, слабо положительно и отрицательно коррелирующие между собой (тест ТШХ и проба Генчи или ИФИ и проба Штанге).

	Correlations					
	Marked correlations are significant at $p < ,05000$					
	N=80					
Variable	Генчи1	Штанге1	ТШХ1	Боргра1	EQ1	ИФИ1
Генчи1	1,000000	0,568238	0,075323	-0,209067	-0,232619	-0,110357
Штанге1	0,568238	1,000000	0,092181	-0,189058	-0,044495	-0,045302
ТШХ1	0,075323	0,092181	1,000000	-0,228146	-0,222139	-0,297505
Боргра1	-0,209067	-0,189058	-0,228146	1,000000	0,101155	0,309403
EQ1	-0,232619	-0,044495	-0,222139	0,101155	1,000000	0,232438
ИФИ1	-0,110357	-0,045302	-0,297505	0,309403	0,232438	1,000000

Рис. 1 Матрица корреляций информативных признаков

Оценка РД произведена конструктивно с использованием следующих критериев:

1. Информативные показатели на основе значений дистанции ТШХ, как ведущего критерия шкалы реабилитационной маршрутизации (ШРМ);

МЕТОДИКА ПРОВЕДЕНИЯ ТЕСТА

При проведении пробы с 6-минутной ходьбой ставится задача пройти как можно большую дистанцию за 6 минут в собственном темпе, после чего пройденное расстояние регистрируется.

Хорошо известный в медицинской практике нагрузочный тест, используемый для оценки функционального статуса пациентов с различными хроническими заболеваниями сердечно-сосудистой и бронхолёгочной систем [11].

2. Оценки физиологических функций

Проба Штанге позволяет оценить устойчивость организма человека к болезням, отражающим общее состояние кислородобеспечивающих систем организма при выполнении задержки дыхания на фоне глубокого вдоха, а проба Генчи – на фоне глубоко выдоха.

МЕТОДИКА ПРОВЕДЕНИЯ ПРОБ

Проба Штанге:

Исследуемому необходимо сделать 2–3 спокойных вдоха и выдоха, затем сделать обычный вдох и задержать дыхание. Засечь время.

Регистрируется время задержки дыхания до вдоха.

Проба Генчи:

Исследуемому предлагают сделать глубокий вдох, затем максимальный выдох. Исследуемый задерживает дыхание при зажатом пальцами носе и плотно закрытом рте.

Регистрируется время задержки дыхания между вдохом и выдохом.

3. Индекс функциональных изменений ИФИ

Показатель, интегрально отражающий функциональное состояние организма, учитывает частоту пульса, артериальное давление, возраст, физическое состояние, включая массу тела и рост.

Для оценки функционирования сердечно-сосудистой системы и оценки её адаптационного потенциала также определяют индекс функциональных изменений (коэффициент здоровья) по формуле:

$$\text{ИФИ} = 0,011 \cdot \text{ЧСС} + 0,014 \cdot \text{САД} + 0,008 \cdot \text{ДАД} + 0,009 \cdot m - 0,009 \cdot P - 0,27, \quad (1.32)$$

где: ЧСС – частота сердечных сокращений;

САД – систолическое артериальное давление;

ДАД – диастолическое артериальное давление;

P – рост; m – масса тела; В – возраст.

В зависимости от значения ИФИ по таблице находят соответствующий уровень функционирования ССС [10].

Величина ИФИ измеряется в баллах. Адаптация считалась удовлетворительной при ИФИ до 2,60 балла. Напряжение механизмов адаптации регистрировалось при ИФИ от 2,60 до 3,10 балла. Неудовлетворительной адаптация считалась при ИФИ от 3,10 до 3,50 балла. Срыв адаптации регистрировался при ИФИ, равном 3,50 балла и выше.

4. Тест на EQ

Предназначен для измерения эмоционального интеллекта (EQ) в соответствии с теоретическими представлениями автора.

5. Шкала Борга

Шкала Борга – субъективный способ определения уровня нагрузки во время занятий лечебной физкультурой. Доктор Гуннар Борг, создатель шкалы, разбил ее от 6 до 20 баллов, как ориентир по сердечному ритму: при умножении балла Борга на 10, полученное значение приблизительно соответствует частоте сердечных сокращений для соответствующего уровня активности.

Формула для расчета частоты сердечных сокращений (ЧСС):

$$\text{ЧСС}_{\text{max}} = 220 - \text{«Возраст»}, \quad (1.33)$$

При выполнении физических упражнений выделяют пять уровней интенсивности тренировок или пульсовых зон:

1. умеренная нагрузка (пульс 50–60 % от максимальной ЧСС)
2. «контроль веса» (пульс 60–70 % от максимальной ЧСС)
3. аэробная зона (пульс 70–80 % от максимальной ЧСС)
4. анаэробная тренировка (пульс 80–90 % от максимальной ЧСС)
5. максимальная нагрузка (пульс 90–100 % от максимальной ЧСС) [9]

1.2.3 Исследование эффективности процедур реабилитации по выбранным признакам и для различных групп пациентов

В исходных данных была представлена база данных из пациентов трех различных групп по реабилитационному лечению. К первой применялось только физиолечение. Причем каждая следующая группа включает в себя помимо своего метода лечения методы предыдущих групп. Ко второй отнесли физиолечение и пробиотические продукты. Пациенты третьей группы в комплексе применяли 2 предыдущих метода, а также продукты пантового оленеводства.

Для анализа данных в работе были использованы:

1. (для выявления значимого различия показателей до и после реабилитации) тест Стьюдента для зависимых наблюдений и критерий Вилкоксона (для категориальных данных);

2. (для выявления значимого различия показателей для различных групп) тесты Стьюдента и Манна-Уитни, дисперсионный анализ, непараметрический дисперсионный анализ Краскала-Уоллиса (для категориальных данных);
3. (для выявления значимого различия показателей до и после реабилитации для различных групп) многофакторный дисперсионный анализ зависимых наблюдений.

1.2.4 Анализ динамики физиологических функций (пробы Генчи и Штанге)

Для начала был проведен многофакторный дисперсионный анализ (ДА) зависимых наблюдений для пробы Генчи (рис. 2).

В данном анализе 3 фактора – группы пациентов, описанные выше.

По пробе Генчи есть значимый положительный эффект реабилитации по каждой группе (рис. 4).

При этом группы изначально были неоднородны – наблюдается значимое различие средних по пробе Генчи до процедур реабилитации (что позволяет сделать вывод о неудачном изначальном формировании групп пациентов).

Поскольку эффект взаимодействия незначим, эффективность реабилитации для какой-то из групп не может выделяться.

Repeated Measures Analysis of Variance (Data1.sta)					
Sigma-restricted parameterization					
Effective hypothesis decomposition					
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	64293,94	1	64293,94	953,6295	0,000000
Группа	733,23	2	366,62	5,4378	0,006180
Error	5191,36	77	67,42		
R1	417,69	1	417,69	66,6424	0,000000
R1*Группа	14,08	2	7,04	1,1236	0,330386
Error	482,61	77	6,27		

Рис. 2 Результаты дисперсионного анализа для пробы Генчи

		LSD test; variable DV_1 (Data1.sta)						
		Probabilities for Post Hoc Tests						
		Error: Between: Within: Pooled MS = 36,844, df = 91,194						
Cell No.	Группа	R1	{1}	{2}	{3}	{4}	{5}	{6}
			22,000	26,125	19,031	22,625	17,250	19,750
1	1	Генчи1		0,000000	0,053482	0,681403	0,012249	0,229163
2	1	Генчи2	0,000000		0,000010	0,023351	0,000007	0,000908
3	2	Генчи1	0,053482	0,000010		0,000000	0,340385	0,699856
4	2	Генчи2	0,681403	0,023351	0,000000		0,004785	0,125345
5	3	Генчи1	0,012249	0,000007	0,340385	0,004785		0,006028
6	3	Генчи2	0,229163	0,000908	0,699856	0,125345	0,006028	

Рис. 3 Множественные сравнения средних по результатам ДА для пробы Генчи (тест LSD Фишера)

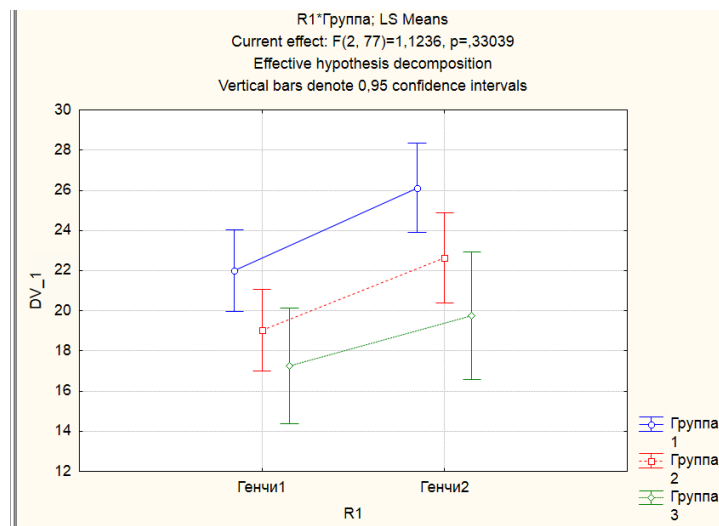


Рис. 4 Графическая иллюстрация результатов реабилитации в различных группах для пробы Генчи

Затем тоже самое проделывалось с пробой Штанге.

В результате в результате исследования значимый положительный эффект реабилитации по каждой группе тоже есть. Однако, эффект взаимодействия между группами незначим и нельзя сделать вывода о том, что эффективность реабилитации для какой-то из групп выше или ниже.

Find/Replace Fill Standardize Links					
Repeated Measures Analysis of Variance (Data1.sta)					
Sigma-restricted parameterization					
Effective hypothesis decomposition					
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	118336,0	1	118336,0	1099,820	0,000000
Группа	1863,2	2	931,6	8,658	0,000406
Error	8284,9	77	107,6		
R1	663,1	1	663,1	122,509	0,000000
R1*Группа	33,2	2	16,6	3,069	0,052174
Error	416,7	77	5,4		

Рис. 5 Результаты дисперсионного анализа для пробы Штанге

Find/Replace		Fill	Standardize	Links					
		LSD test; variable DV_1 (Data1.sta) Probabilities for Post Hoc Tests Error: Between; Within; Pooled MS = 56,504, df = 84,727							
ta1.sta)	Cell No.	Группа	R1	{1}	{2}	{3}	{4}	{5}	{6}
	1	1	Штанге1	30,625	36,188	25,750	29,687	23,188	26,563
	2	1	Штанге2	0,000000	0,000000	0,011171	0,619160	0,001755	0,081151
ta1.sta)	3	2	Штанге1	0,011171	0,000000	0,000000	0,000851	0,000000	0,000070
	4	2	Штанге2	0,619160	0,000851	0,000000	0,000000	0,268699	0,724951
	5	3	Штанге1	0,001755	0,000000	0,268699	0,005909	0,005909	0,178143
	6	3	Штанге2	0,081151	0,000070	0,724951	0,178143	0,000100	0,000100

Рис. 6 Множественные сравнения средних по результатам ДА для пробы Штанге

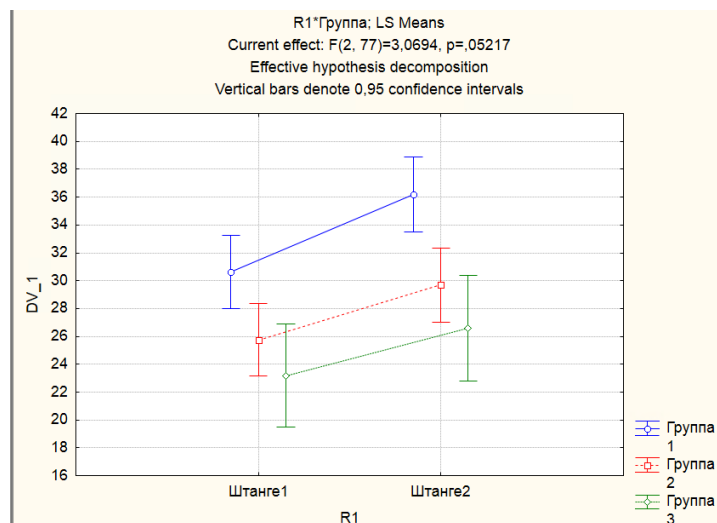


Рис. 7 Графическая иллюстрация результатов реабилитации в различных группах для пробы Штанге

1.2.5 Анализ динамики функциональных функций (тест ТШХ)

Далее откликом исследования в дисперсионном анализе стал тест ТШХ.

По полученному результату (рис. 9) можно предположить, что эффективность в 2,3 группе выше (по изгибу графика), но эффект взаимодействия все же незначим.

/Replace		Fill	Standardize	Links	
Repeated Measures Analysis of Variance (Data1.sta)					
Sigma-restricted parameterization					
Effective hypothesis decomposition					
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	20876903	1	20876903	2882.056	0.000000
Группа	74678	2	37339	5.155	0.007926
Error	557769	77	7244		
R1	244160	1	244160	194.839	0.000000
R1*Группа	2791	2	1396	1.114	0.333586
Error	96492	77	1253		

Рис. 8 Результаты дисперсионного анализа для теста ТШХ

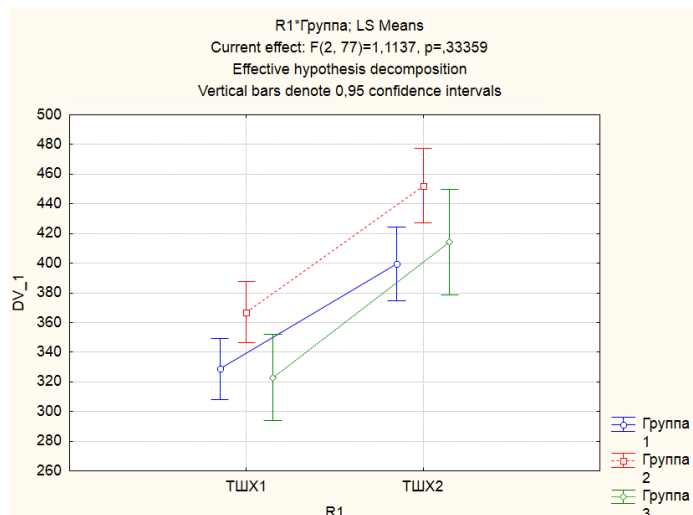


Рис. 9 Графическая иллюстрация результатов реабилитации в различных группах для теста ТШХ

LSD test; variable DV_1 (Data1.sta)
Probabilities for Post Hoc Tests
Error: Between; Within; Pooled MS = 4248.4, df = 102.87

Группа	R1	{1}	{2}	{3}	{4}	{5}	{6}
1	TШХ1	328.84	0,000000	0,021026	0,000000	0,765485	0,000044
1	TШХ2	0,000000	0,048534	0,001665	0,000211	0,469150	
2	TШХ1	0,021026	0,048534		0,000000	0,029140	0,020337
2	TШХ2	0,000000	0,001665	0,000000		0,000000	0,058875
3	TШХ1	0,765485	0,000211	0,029140	0,000000		0,000000
3	TШХ2	0,000044	0,469150	0,020337	0,058875	0,000000	

Рис. 10 Множественные сравнения средних по результатам ДА для теста ТШХ

Рассмотрим относительные изменения ТШХ для трех групп. В результате дисперсионного анализа - эффект группы не значим, хоть и для третьей группы процент изменения больше, но он не значимо больше.

Univariate Tests of Significance for ТШХ% (Data1.sta)
Sigma-restricted parameterization
Effective hypothesis decomposition

Effect	SS	Degr. of Freedom	MS	F	p
Intercept	2,604955	1	2,604955	282,5099	0,000000
Группа	0,021661	2	0,010831	1,1746	0,314421
Error	0,709998	77	0,009221		

Рис. 11 Относительные изменения для теста ТШХ

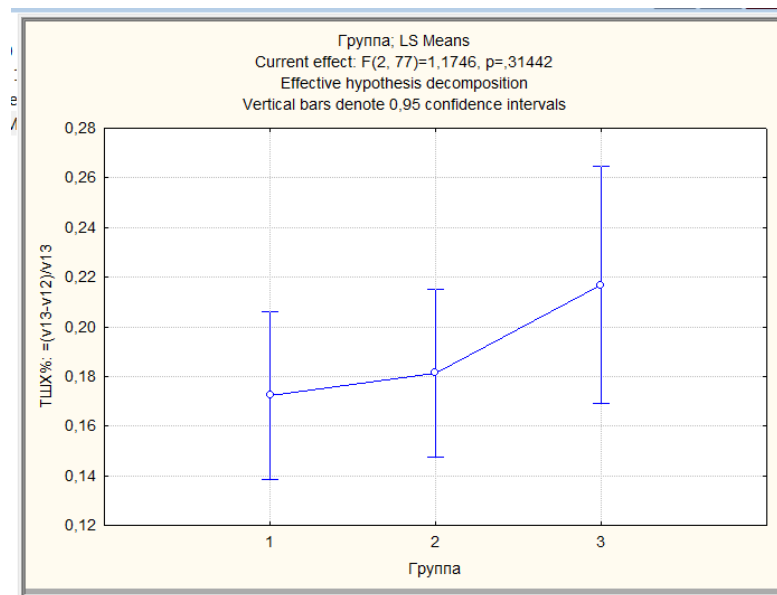


Рис. 12 Графическая иллюстрация относительных изменений для теста ТШХ

1.2.6 Анализ динамики физических нагрузок (шкала Борга)

Анализируя результаты пациентов по шкале Борга есть значимый положительный эффект реабилитации в каждой группе. Но эффект взаимодействия незначим – нельзя выделить эффективность реабилитации для какой-то из групп.

Графически результаты реабилитации очень близки меж собой (рис. 14).

Repeated Measures Analysis of Variance (Data1.sta)					
Sigma-restricted parameterization					
Effective hypothesis decomposition					
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	628,1289	1	628,1289	423,1605	0,000000
Группа	2,8969	2	1,4484	0,9758	0,381506
Error	114,2969	77	1,4844		
R1	43,0664	1	43,0664	272,4406	0,000000
R1*Группа	0,0219	2	0,0109	0,0692	0,933206
Error	12,1719	77	0,1581		

Рис. 13 Дисперсионный анализ для шкалы Борга

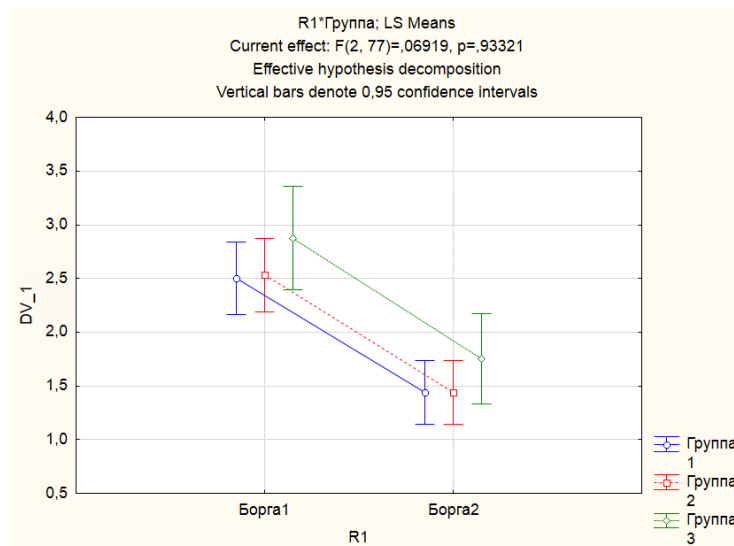


Рис. 14 Графическая иллюстрация относительных изменений для шкалы Борга

1.2.7 Анализ динамики эмоционального интеллекта (тест EQ)

По дисперсионному анализу теста эмоционального интеллекта, опираясь на угол представленных графиков для 3-ех групп можно заметить, что наибольшее изменение есть у 1-ой группы, и к тому же есть значимый эффект взаимодействия эффекта реабилитации. В данном случае он выражается в том, что изменения в 3 группе менее значимы, чем в 1 и 2.

Repeated Measures Analysis of Variance (Data1.sta)					
Sigma-restricted parameterization					
Effective hypothesis decomposition					
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	18056.64	1	18056.64	3398.637	0.000000
Группа	46.90	2	23.45	4.414	0.015320
Error	409.09	77	5.31		
R1	150.06	1	150.06	181.519	0.000000
R1*Группа	7.79	2	3.89	4.710	0.011756
Error	63.66	77	0.83		

Рис. 15 Дисперсионный анализ для теста EQ

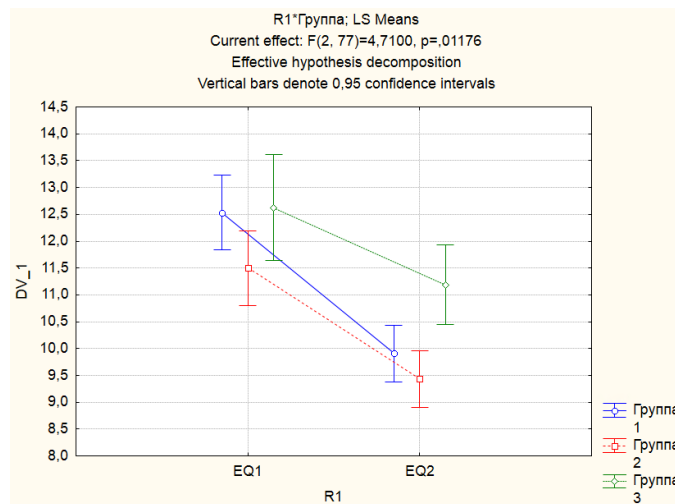


Рис. 15 Графическая иллюстрация относительных изменений теста EQ

LSD test; variable DV_1 (Data1.sta)								
Probabilities for Post Hoc Tests								
Error: Between; Within; Pooled MS = 3,0698, df = 100,40								
Cell No.	Группа	R1	{1}	{2}	{3}	{4}	{5}	{6}
1	1	EQ1	12,531	9,9063	11,500	9,4375	12,625	11,188
2	1	EQ2	0,000000	0,000000	0,020502	0,000000	0,861624	0,013860
3	2	EQ1	0,020502	0,000435		0,000000	0,038500	0,561524
4	2	EQ2	0,000000	0,287117	0,000000		0,000000	0,001512
5	3	EQ1	0,861624	0,000002	0,038500	0,000000		0,000026
6	3	EQ2	0,013860	0,018795	0,561524	0,001512	0,000026	

Рис. 16 Множественные сравнения средних по результатам ДА для теста EQ

1.2.8 Анализ динамики функциональное состояние организма (индекс ИФИ)

По оценке индекса ИФИ для трех групп пациентов также есть значимый положительный эффект реабилитации для каждой из них. Поскольку эффект взаимодействия незначим, нельзя сделать вывода о том, что эффективность реабилитации для какой-то из групп может выделяться.

Лишь графически можно сделать вывод о лучшей позиции 3-ей группы, в то время как 1-ая и 2-ая имеют достаточно близкий результат.

Repeated Measures Analysis of Variance (Data1.sta)					
Sigma-restricted parameterization					
Effective hypothesis decomposition					
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	1255,307	1	1255,307	3830,077	0,000000
Группа	0,189	2	0,095	0,289	0,749960
Error	25,237	77	0,328		
R1	2,067	1	2,067	89,011	0,000000
R1*Группа	0,086	2	0,043	1,854	0,163550
Error	1,788	77	0,023		

Рис. 17 Дисперсионный анализ для ИФИ

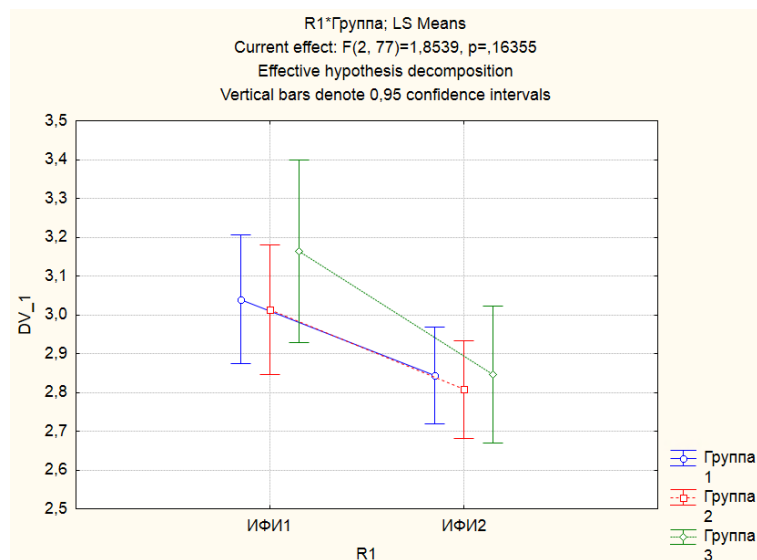


Рис. 18 Графическая иллюстрация относительных изменений для ИФИ

LSD test; variable DV_1 (Data1.sta)								
Probabilities for Post Hoc Tests								
Error: Between; Within; Pooled MS = ,17549, df = 87,858								
Cell No.	Группа	R1	{1}	{2}	{3}	{4}	{5}	{6}
			3,0400	2,8442	3,0139	2,8074	3,1631	2,8466
1	1	ИФИ1		0,000002	0,803361	0,028906	0,339849	0,135053
2	1	ИФИ2	0,000002		0,108695	0,726497	0,014780	0,985075
3	2	ИФИ1	0,803361	0,108695		0,000001	0,247734	0,195495
4	2	ИФИ2	0,028906	0,726497	0,000001		0,006776	0,760878
5	3	ИФИ1	0,339849	0,014780	0,247734	0,006776		0,000000
6	3	ИФИ2	0,135053	0,985075	0,195495	0,760878	0,000000	

Рис. 19 Множественные сравнения средних по результатам ДА для ИФИ

1.2.9 Анализ динамики физиологических функций (пробы Генчи и Штанге) по возрастным группам

Графически немного эффективнее 3ья группа, но эффект взаимодействия также незначим.

Касаясь влияния реабилитации по параметру шкалы Генчи по группе возраст можно сказать, что есть значимый положительный эффект.

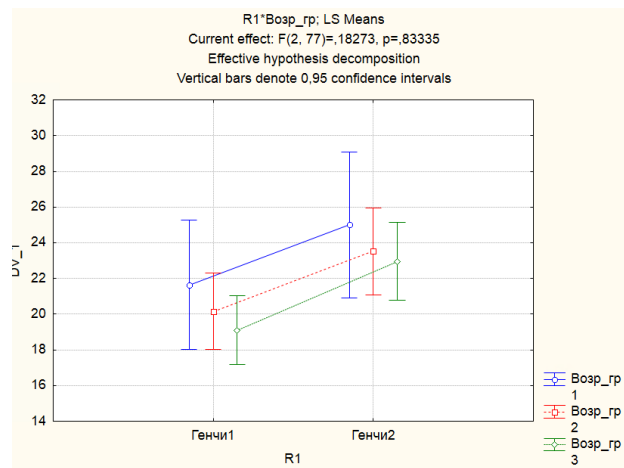


Рис. 20 Графическая иллюстрация относительных изменений для пробы Генчи (фактор – возраст)

		LSD test; variable DV_1 (Data1.sta)							
		Probabilities for Post Hoc Tests							
		Error: Between; Within; Pooled MS = 41.078, df = 89,961							
Cell No.	Возр_гр	R1	{1}	{2}	{3}	{4}	{5}	{6}	
			21.636	25.000	20.161	23.516	19.105	22.947	
1	1	Генчи1		0.002597	0.513634	0.405534	0.251783	0.551717	
2	1	Генчи2	0.002597		0.034141	0.511132	0.008605	0.352085	
3	2	Генчи1	0.513634	0.034141		0.000002	0.497746	0.075830	
4	2	Генчи2	0.405534	0.511132	0.000002		0.005521	0.714726	
5	3	Генчи1	0.251783	0.008605	0.497746	0.005521		0.000000	
6	3	Генчи2	0.551717	0.352085	0.075830	0.714726	0.000000		

Рис. 21 Множественные сравнения средних по возрастным группам для пробы Генчи (тест LSD Фишера)

Графически немного эффективнее 1ья группа, но эффект взаимодействия также незначим. По влиянию реабилитации по параметру шкалы Штанге по группе возраст можно сказать, что есть значимый положительный эффект.

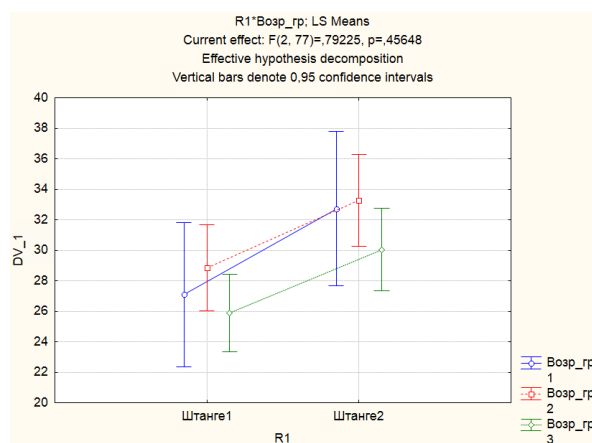


Рис. 22 Графическая иллюстрация относительных изменений для пробы Штанге (фактор – возраст)

1.2.10 Анализ динамики функциональных функций (тест ТШХ) по возрастным группам

Графически немного эффективнее 1-ья группа, но эффект взаимодействия также незначим.

По влиянию реабилитации по параметру теста ТШХ по группе возраст можно сказать, что есть значимый положительный эффект и возраст действительно значимо влияет на результаты ТШХ. Графически немного эффективнее 2-ая и 3-ья группы.

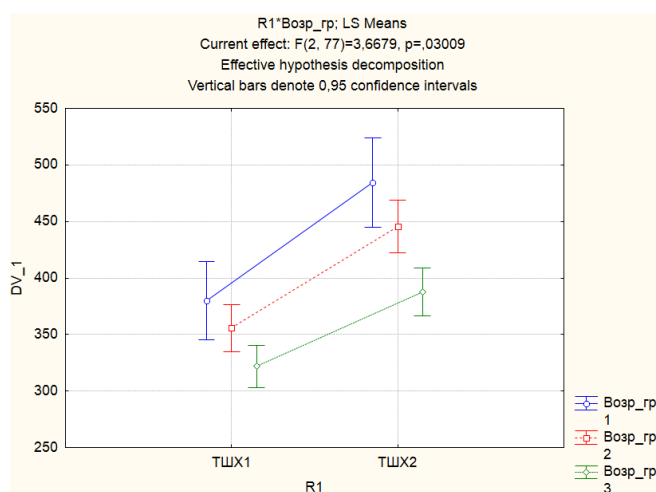


Рис. 23 Графическая иллюстрация относительных изменений для теста ТШХ (фактор – возраст)

1.2.11 Анализ динамики физических нагрузок (шкала Борга) по возрастным группам

По влиянию реабилитации по параметру теста Борга по группе возраст можно сказать, что есть значимый положительный эффект.

В данном случае можно утверждать, что есть значимое влияние взаимодействия эффекта реабилитации и возрастной группы для теста Борга в силу того, что эффект в старшей возрастной группе менее выражен чем в средней.

Repeated Measures Analysis of Variance (Data1.sta)					
Sigma-restricted parameterization					
Effective hypothesis decomposition					
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	18874316	1	18874316	2911,719	0,000000
Возр_гр	133318	2	66659	10,283	0,000110
Error	499129	77	6482		
R1	227037	1	227037	192,857	0,000000
R1*Возр_гр	8636	2	4318	3,668	0,030090
Error	90647	77	1177		

Рис. 24 Дисперсионный анализ для шкалы Борга (фактор – возраст)

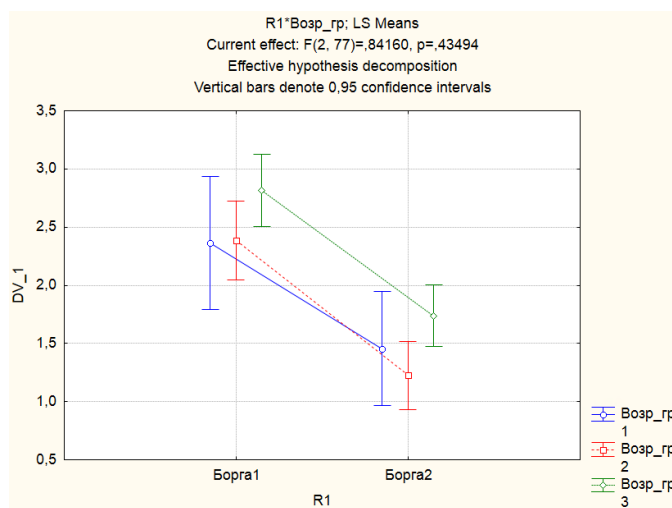


Рис. 25 Графическая иллюстрация относительных изменений для шкалы Борга (фактор – возраст)

1.2.12 Кластерный анализ

Для установления реабилитационного диагноза и оценки результативности методов реабилитации была проведена кластеризация пациентов по основным информационным признакам (проба Генчи, проба Штанге, тест ТШХ, ИФИ (до и после процедур реабилитации) – всего 8 признаков).

Предварительно, используя метод главных компонент, был осуществлен переход от пространства информационных признаков к пространству главных компонент.

Собственные значения матрицы корреляций выбранных признаков, упорядоченные по убыванию значений, приведены в таблице на рис. 26, а на рис. 27 приведен график собственных значений в зависимости от порядкового номера.

Value number	Собственные значения корреляционной матрицы			
	Eigenvalue	% Total variance	Cumulative Eigenvalue	Cumulative %
1	3,315798	41,44747	3,315798	41,4475
2	2,192138	27,40172	5,507936	68,8492
3	1,230417	15,38021	6,738353	84,2294
4	0,706030	8,82537	7,444382	93,0548
5	0,239840	2,99800	7,684222	96,0528
6	0,153437	1,91796	7,837659	97,9707
7	0,094163	1,17704	7,931822	99,1478
8	0,068178	0,85222	8,000000	100,0000

Рис. 26 Таблица собственных значений матрицы корреляций

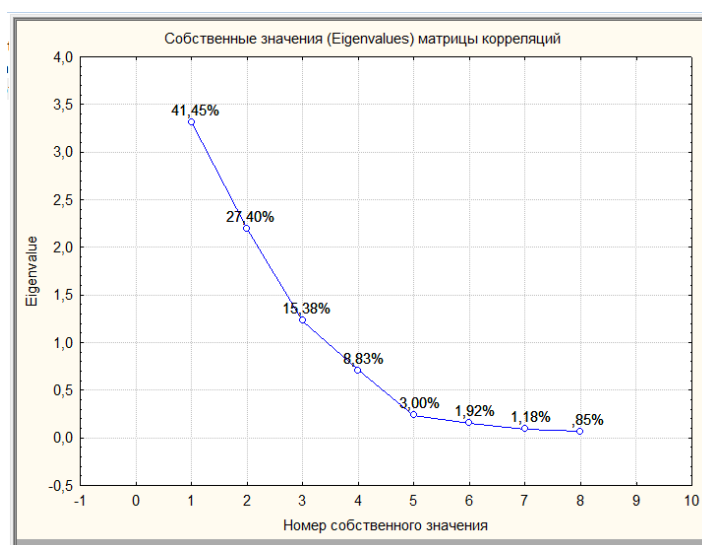


Рис. 27 График собственных значений матрицы корреляции

Первые 3 компоненты объясняют 84,2 % дисперсии исходных признаков, причем собственные значения для этих компонент больше единицы, что по критерию Кайзера позволяет выбрать именно 3 компоненты. Если ориентироваться на метод каменистой осыпи, то график имеет явный излом на 5 компоненте, что говорит о том, что мы должны выбрать в качестве главных 4 первые компоненты.

Таким образом, следует выбрать либо 3, либо 4 главные компоненты. Последующий кластерный анализ показал, что использование 4 компонент избыточно, полученные кластеры не различались в итоге значимо по 4-ой компоненте.

Отталкиваясь от этого, выбираем 3 компоненты и проводим разведочный иерархический кластерный анализ.

Полученная дендрограмма изображена на рис. 28:

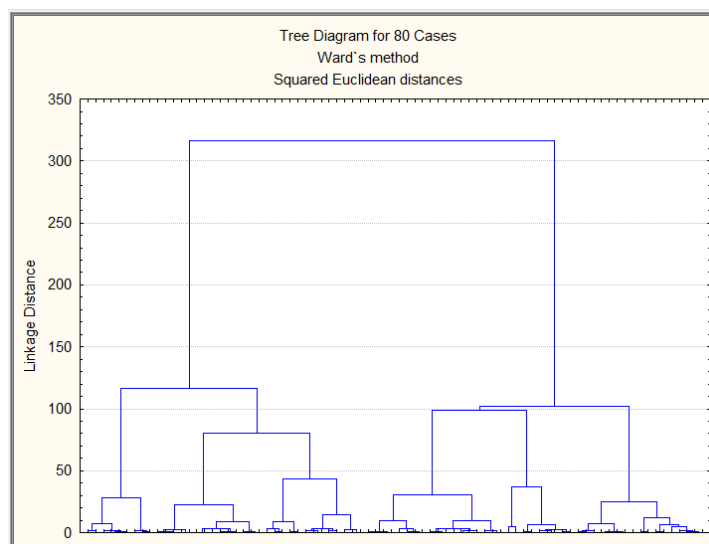


Рис. 28 Дендрограмма иерархической кластеризации

На основе анализа дендрограммы разумно выбрать разбиение на 6 кластеров. Окончательное разбиение осуществляем методом к-средних.

Состав полученных кластеров по пациентам различных трех групп приведен в следующей таблице (рис. 29):

CLUSTER	Группа 1	Группа 2	Группа 3	Row Totals
1	0	10	3	13
	0,00%	76,92%	23,08%	
2	3	10	3	16
	18,75%	62,50%	18,75%	
3	5	3	2	10
	50,00%	30,00%	20,00%	
4	6	2	6	14
	42,86%	14,29%	42,86%	
5	7	5	2	14
	50,00%	35,71%	14,29%	
6	11	2	0	13
	84,62%	15,38%	0,00%	
All Grps	32	32	16	80

Рис. 29 Таблица состава полученных кластеров по основным группам

В следующей таблице приведен состав кластеров по возрастным группам (рис. 30):

CLUSTER	Возр_гр 1	Возр_гр 2	Возр_гр 3	Row Totals
1	5	6	2	13
	38,46%	46,15%	15,38%	
2	1	6	9	16
	6,25%	37,50%	56,25%	
3	0	1	9	10
	0,00%	10,00%	90,00%	
4	0	4	10	14
	0,00%	28,57%	71,43%	
5	3	8	3	14
	21,43%	57,14%	21,43%	
6	2	6	5	13
	15,38%	46,15%	38,46%	
All Grps	11	31	38	80

Рис. 30 Таблица состава полученных кластеров по возрастным группам

Исследуем, чем различаются кластеры по основным информативным признакам (используем метод многофакторного дисперсионного анализа зависимых наблюдений).

Различие в кластерах по пробе Генчи (рис. 31):

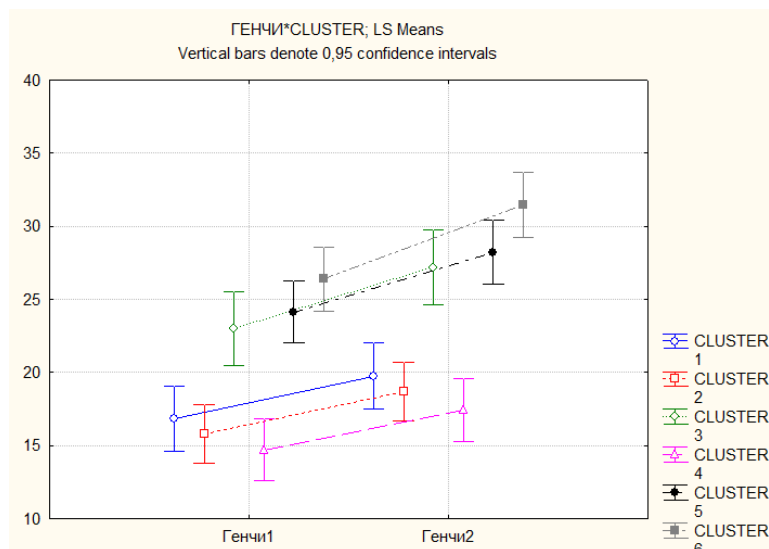


Рис. 31 Графическая иллюстрация различий кластеров для пробы Генчи

Видим, что кластеры можно разбить на 2 группы с низким уровнем пробы Генчи у пациентов (1, 2, 4 кластеры) и высоким уровнем (3, 5, 6 кластеры). Внутри этих групп значения пробы Генчи у пациентов до (после) реабилитации значимо не различаются, в то же время есть значимое различие значений для любой пары из двух групп.

В то же время значимых различий в процентном изменении среднего значения пробы Генчи для этих кластеров нет (рис. 32):

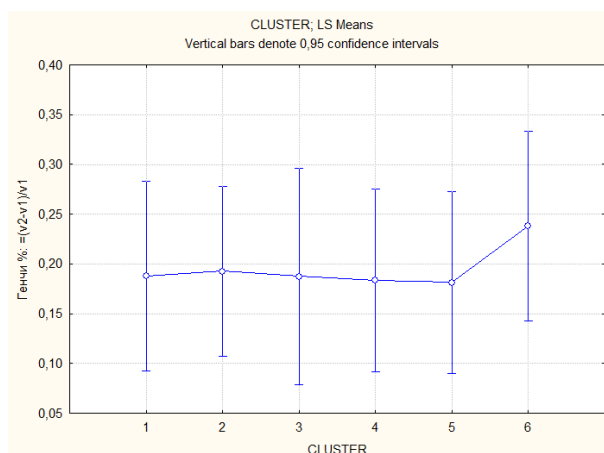


Рис. 32 Графическая иллюстрация для пробы Генчи в процентном изменении

Различие в кластерах по пробе Штанге (рис. 33):

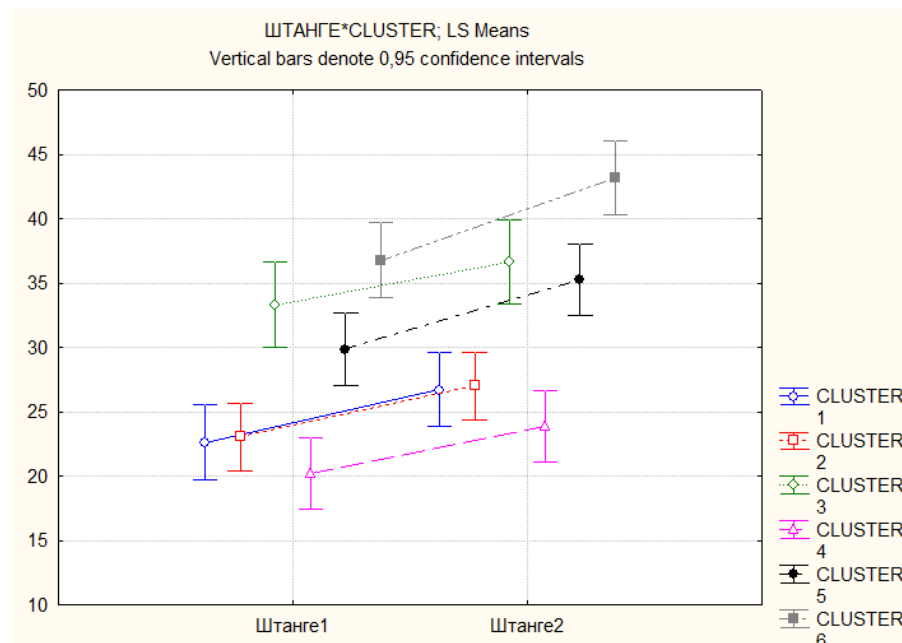


Рис. 33 Графическая иллюстрация различий кластеров для пробы Штанге

Видим, что результат практически аналогичен предыдущему: кластеры можно разбить на 2 группы с низким уровнем пробы Штанге у пациентов (1, 2, 4 кластеры) и высоким уровнем (3, 5, 6 кластеры). Отличие лишь в том, что для кластеров 5 и 6 из второй группы существует слабо значимое различие между значениями проб (как до, так и после реабилитации).

Различие кластеров по ТШХ (рис. 34):

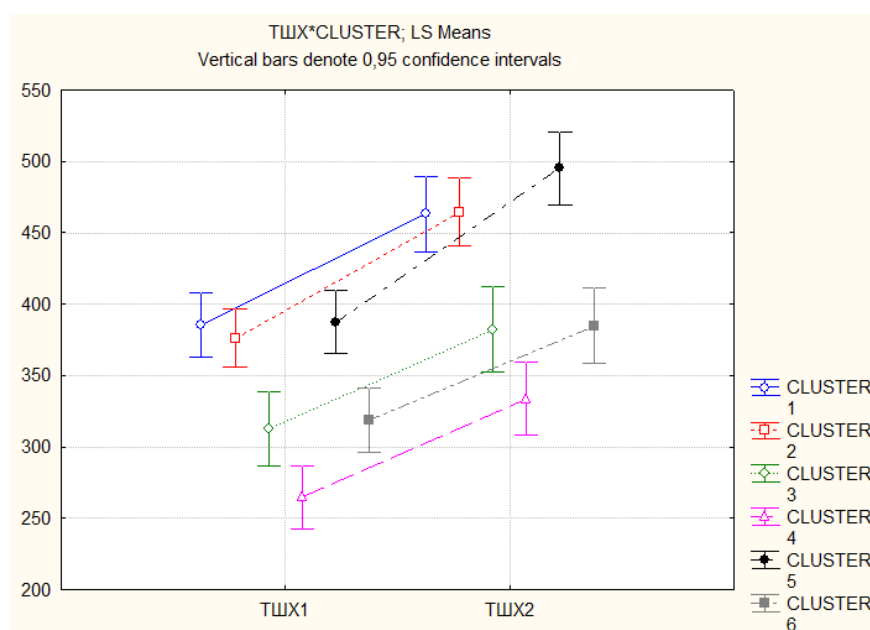


Рис. 34 Графическая иллюстрация различий кластеров по тесту ТШХ

В данном случае кластеры можно также разбить на 2 группы с низким и высоким уровнем ТШХ. В первую группу вошли 3,4 и 6 кластеры, во вторую 1,2 и 5 кластеры. Внутри этих групп значения ТШХ у пациентов до (после) реабилитации значимо не различаются, в то же время есть значимое различие значений для любой пары из двух групп. Значимых различий в процентом изменении среднего значения ТШХ для любых кластеров нет.

Различие кластеров по ИФИ (рис. 35).

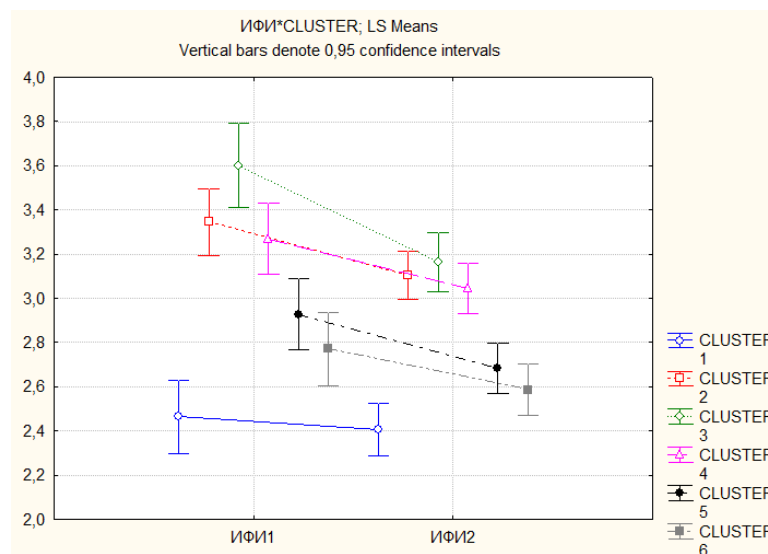


Рис. 35 Графическая иллюстрация различий кластеров по индексу ИФИ

Можем выделить три группы кластеров: с высоким уровнем ИФИ (2, 3, 4 кластеры), со средним уровнем ИФИ (5, 6 кластеры) и с низким уровнем ИФИ – 1 кластер. Причем для первого кластера нет значимого эффекта реабилитации, однако это связано с тем, что пациенты данного кластера изначально обладают удовлетворительным адаптационным потенциалом.

2. Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

Введение

Главная цель данного раздела – оценка перспективности развития и планирование финансовой и коммерческой ценности конечного продукта, представленного в рамках исследовательской программы.

Коммерческая ценность определяется не только наличием более высоких технических характеристик над конкурентными разработками, но и тем, насколько быстро разработчик сможет ответить на следующие вопросы – будет ли продукт востребован на рынке, какова будет его цена, каков бюджет научного исследования, какое время будет необходимо для продвижения разработанного продукта на рынок.

Данный раздел предусматривает рассмотрение следующих задач:

1. Оценка коммерческого потенциала разработки;
2. Планирование научно-исследовательской работы;
3. Расчет бюджета научно-исследовательской работы;
4. Определение ресурсной, финансовой, бюджетной эффективности исследования.

Цель данной НИР – оценка эффективности процедуры реабилитации для пациентов, перенесших COVID-19.

2.1 Оценка коммерческого потенциала и перспективности проведения исследований с позиции ресурсоэффективности и ресурсосбережения

2.1.1 Анализ конкурентных технических решений

В ходе исследования была рассмотрена конкурирующая программа Microsoft Excel, способная также проводить дисперсионный и кластерный анализы.

В следствие того, что каждая программа имеет свои достоинства и недостатки необходим их детальный анализ. В таблице 2.1 показано сравнение

разработок конкурентов и разработки данной ВКР с точки зрения технических и экономических критериев оценки эффективности.

Таблица 2.1 – Сравнение конкурентных приложений

Критерии оценки	Вес критерия	Баллы		Конкурентоспособность	
		Б _ф	Б _к	К _ф	К _к
1	2	3	4	5	6
Технические критерии оценки ресурсоэффективности					
1. Простота эксплуатации	0.12	4	5	0.48	0.6
2. Потребность в ресурсах памяти	0.06	4	3	0.24	0.18
3. Точность результатов	0.15	4	2	0.6	0.3
4. Предоставляемые возможности	0.21	5	3	1.05	0.63
5. Скорость работы	0.05	5	4	0.25	0.2
Экономические критерии оценки эффективности					
1. Конкурентоспособность продукта	0,09	4	3	0.36	0.27
2. Уровень проникновения на рынок	0,08	3	5	0.24	0.4
3. Цена	0,05	4	5	0.2	0.25
4. Предполагаемый срок эксплуатации	0,09	4	4	0.36	0.36
5. Послепродажное обслуживание	0,08	3	4	0.24	0.32
6. Срок выхода на рынок	0,02	5	4	0.1	0.08
Итого	1	46	42	3.96	3.59

Расчет конкурентоспособности определяется по формуле:

$$K = \sum B_i * B_i = 3.96, \quad (2.1)$$

где K – конкурентоспособность проекта;

B_i – вес показателя (в долях единицы);

B_i – балл показателя.

Согласно проведенному анализу, можно сделать вывод, что используемая программа (Statistica) близка по многим показателям со своим конкурентом. Сильными сторонами разрабатываемого проекта стали такие показатели как: предоставляемые возможности, скорость работы и срок выхода на рынок.

Слабыми сторонами выделились послепродажное обслуживание и уровень проникновения на рынок, так как Microsoft Excel считается офисной программой, ее использование актуальнее.

2.1.2 SWOT-анализ

Для исследования внешней и внутренней среды проекта, в этой работе проведен SWOT-анализ с детальной оценкой сильных и слабых сторон исследовательского проекта, а также его возможностей и угроз.

Первым этапом создается матрица SWOT, где описаны слабые и сильные стороны проекта, выявленные возможности и угрозы для реализации проектной работы (таблица 2.2).

Таблица 2.2 – Матрица SWOT-анализа

Сильные стороны	Слабые стороны
С1. Выявление эффективности различных методик лечения пациентов после COVID-9	Сл1. Отсутствие ссылок и материалов для соответствующих научных исследований.
С2. Использование разрабатываемого анализа как методику выявления эффективности каких-либо процедур	Сл2. Трудности в «считывании» результатов программы после анализов
С3. Низкая цена проекта	Сл3. Низкая скорость обработки больших данных

Возможности	Угрозы
В1. Возникновение потенциального спроса в сфере реабилитационного анализа после COVID-9 на новые разработки.	У1. Снижение стоимости разработок конкурентов.
В2. Увеличение выводов после анализа благодаря изучению большего набора исходных данных.	У2. Увеличение сложности проекта

На втором этапе на основании матрицы SWOT строятся интерактивные матрицы возможностей и угроз, позволяющие оценить эффективность проекта, а также надежность его реализации.

Соотношения параметров представлены в таблицах 2.3–2.6.

Таблица 2.3 – Интерактивная матрица проекта «Возможности проекта и сильные стороны»

Сильные стороны проекта				
Возможности проекта		C1	C2	C3
	B1	+	+	-
	B2	-	+	-

Таблица 2.4 – Интерактивная матрица проекта «Возможности проекта и слабые стороны»

Слабые стороны проекта				
Возможности проекта		Сл1	Сл2	Сл3
	B1	-	-	-
	B2	-	+	+

Таблица 2.5 – Интерактивная матрица проекта «Угрозы проекта и сильные стороны»

Сильные стороны проекта				
Угрозы проекта		C1	C2	C3
	У1	-	-	-
	У2	-	+	-

Таблица 2.6 – Интерактивная матрица проекта «Угрозы проекта и слабые стороны»

Слабые стороны проекта				
Угрозы проекта		Сл1	Сл2	С3
	У1	-	-	-
	У2	-	+	+

Результаты анализа представлены в итоговой таблице 4.7.

Таблица 2.7 – Итоговая таблица SWOT-анализа

	Сильные стороны научно-исследовательского проекта С1. Выявление эффективности различных методик лечения пациентов после COVID-9 С2. Использование разрабатываемого анализа как методику выявления эффективности каких-либо процедур С3. Низкая цена проекта	Слабые стороны научно-исследовательского проекта Сл1. Отсутствие ссылок и материалов для соответствующих научных исследований. Сл2. Трудности в «считывании» результатов программы после анализов Сл3. Низкая скорость обработки больших данных
Возможности В1. Возникновение потенциального спроса в сфере реабилитационного анализа после COVID-9 на новые разработки. на новые разработки. В2. Увеличение выводов после анализа благодаря изучению большего набора исходных данных.	Направления развития В1С1С2. Благодаря используемой эффективной концепции для проекта появится спрос на новые разработки в этой области. В2С2. Увеличение возможностей проделанного анализа за счёт его дальнейшего развития и применения для более универсальных случаев	Сдерживающие факторы В2Сл2Сл3. Возможные трудности с улучшением анализа из-за не корректного считывания результатов и долгой обработки больших данных
Угрозы У1. Снижение стоимости разработок конкурентов. У2. Увеличение сложности проекта	Угрозы развития У2С2. Сложность в разработке более универсальной версии проекта принесет значимые трудозатраты	Уязвимости: У2Сл2Сл3. Сложность в модернизации проекта связана с низкой скоростью обработки больших данных и считывании результатов программы

В результате SWOT-анализа можно сделать вывод, что у разрабатываемого проекта есть свои преимущества и недостатки, которые влияют на реализацию работы. Однако по результирующей таблице все угрозы несут лишь предполагаемый, а не потенциальный характер. В реальности проект в своем виде совмещает в себе достаточно сильных сторон.

2.2. Планирование научно-исследовательских работ

Планирование комплекса научно-исследовательских работ осуществляется в порядке:

1. определение структуры работ в рамках научного исследования;
2. определение количества исполнителей для каждой из работ;
3. установление продолжительности работ;
4. построение графика проведения научных исследований.

Для оптимизации работ удобно использовать классический метод линейного планирования и управления.

Результатом такого планирования является составление линейного графика выполнения всех работ.

Порядок этапов работ и распределение исполнителей для данной научно-исследовательской работы, приведен в таблице 2.8.

Таблица 2.8 – Перечень этапов, работ и распределение исполнителей

Основные этапы	№ раб	Содержание работ	Должность исполнителя
Разработка технического задания	1	Составление и утверждение технического задания, утверждение плана-графика	Научный руководитель
	2	Календарное планирование выполнения работ	Инженер, научный руководитель
Выбор способа решения поставленной задачи	3	Обзор научной литературы	Инженер
	4	Выбор методов исследования	Инженер
Теоретические и программные исследования	5	Исследование теоретической документации об анализируемом объекте	Инженер, научный руководитель
	6	Обработка исходных данных	Инженер
	7	Проведение анализа для обработанных данных	Инженер

Обобщение и оценка результатов	8	Обработка полученных данных	Инженер
	9	Оценка правильности полученных результатов	Инженер, Научный руководитель
Оформление отчета по НИР	10	Составление пояснительной записки	Инженер

2.2.1 Определение трудоемкости выполнения работ и разработка графика проведения

При проведении научных исследований основную часть стоимости разработки составляют трудовые затраты, поэтому определение трудоемкости проводимых работ является важным этапом составления сметы.

Для определения ожидаемого (среднего) значения трудоемкости использована следующая формула:

$$t_{ож\ i} = \frac{3t_{min\ i} + 2t_{max\ i}}{5}, \quad (2.2)$$

где $t_{ож\ i}$ – ожидаемая трудоемкость выполнения i -ой работы, человеко-дни; $t_{min\ i}$ – минимально возможная трудоемкость выполнения заданной i -ой работы, человеко-дни; $t_{max\ i}$ – максимально возможная трудоемкость выполнения заданной i -ой работы, человеко-дни.

Зная величину ожидаемой трудоемкости, можно определить продолжительность каждой i -ой работы в рабочих днях T_{pi} , при этом учитывается параллельность выполнения работ разными исполнителями. Данный расчёт позволяет определить величину заработной платы.

$$T_{pi} = \frac{t_{ож\ i}}{Ч_i}, \quad (2.3)$$

где T_{pi} – продолжительность одной работы, рабочие дни; $t_{ож\ i}$ – ожидаемая трудоемкость выполнения одной работы, человеко-дни; $Ч_i$ – численность исполнителей, выполняющих одновременно одну и ту же работу на данном этапе, чел.

Для перевода длительности каждого этапа из рабочих в календарные дни, необходимо воспользоваться формулой (2.4):

$$Тки = Тpi * k_{кал}, \quad (2.4)$$

где T_{ki} – продолжительность выполнения i -й работы в календарных днях; T_{pi} – продолжительность выполнения i -й работы в рабочих днях; $k_{\text{кал}}$ – календарный коэффициент.

Календарный коэффициент определяется по формуле:

$$K_{\text{кал}} = \frac{T_{\text{кал}}}{T_{\text{кал}} - T_{\text{вых}} - T_{\text{пр}}} = \frac{365}{365 - 104 - 14} = 1,48, \quad (4.5)$$

где $T_{\text{кал}}$ – общее количество календарных дней в году; $T_{\text{вых}}$ – общее количество выходных дней в году; $T_{\text{пр}}$ – общее количество праздничных дней в году (все значения приняты за 2023 год).

Расчеты временных показателей проведения научного исследования обобщены в таблице 2.9.

Таблица 2.9 – Временные показатели проведения научного исследования

Название работы	Трудоёмкость работ						Длительность работ в рабочих днях T_{pi}	Длительность работ в календарных днях T_{ki}
	$tmin$, чел-дни		$tmax$, чел-дни		$t_{ож}i$, чел-дни			
	Исп.1	Исп.2	Исп.1	Исп.2	Исп.1	Исп.2		
1	2	3	4	5	6	7	8	9
1. Составление и утверждение технического задания, утверждение плана графика	2	-	4	-	2,8	-	2,8	4
2. Календарное планирование выполнения работ	1	3	3	4	1,8	3,4	2,6	4
3. Обзор научной литературы	-	6	-	10	-	7,6	7,6	11
4. Выбор методов исследования	-	3	-	5	-	3,8	3,8	6
5. Исследование теоретической документации об анализируемом объекте	2	6	4	8	2,8	6,8	4,8	7
6. Обработка исходных данных	-	5	-	7	-	5,8	5,8	9

7. Проведение анализа для обработанных данных	-	15	-	20	-	17	17	25
8. Обработка полученных данных	-	10	-	15	-	12	12	18
9. Оценка правильности полученных результатов	2	3	4	5	2,8	3,8	3,3	5
10. Составление пояснительной записки		8		10	-	8,8	8,8	13
Итого:	7	59	15	84	10,2	69	68,5	102

Примечание: Исп. 1 – научный руководитель, Исп. 2 – инженер.

На основе таблицы составлен календарный план-график выполнения проекта с использованием диаграммы Ганта (таблица 2.10).

Таблица 2.10 – Диаграмма Ганта

№	Вид работ	Исп	T k i	Продолжительность работ											
				февр			март			апрель			май		
1	Составление и утверждение технического задания, утверждение плана графика	Исп1	4	■											
2	Календарное планирование выполнения работ	Исп1 Исп2	4	■	□										
3	Обзор научной литературы	Исп2	1 1			□									
4	Выбор методов исследования	Исп2	6				□								
5	Исследование теоретической документации об анализируемом объекте	Исп1 Исп2	7				■	□							
6	Обработка исходных данных	Исп2	9						□						
7	Проведение анализа для обработанных данных	Исп2	2 5							□					
8	Обработка полученных данных	Исп2	1 8								□		■	□	
9	Оценка правильности полученных результатов	Исп1 Исп2	5											□	
10	Составление пояснительной записки	Исп2	1 3												

Примечание: □ - Исп. 2 (Инженер); ■ - Исп. 1 (Научный руководитель)

2.3 Бюджет научно-технического исследования

При планировании бюджета научно-технического исследования учитывались все виды расходов, связанных с его выполнением. В этой работе использовать следующую группировку затрат по следующим статьям:

- материальные затраты выпускной квалификационной работы (ВКР);
- основная заработная плата исполнителей темы;
- дополнительная заработная плата исполнителей темы;
- отчисления во внебюджетные фонды (страховые отчисления);
- накладные расходы ВКР.

2.3.1 Расчет материальных затрат научно-технического исследования

Материальные затраты — это затраты организации на приобретение сырья и материалов для создания готовой продукции.

В связи с тем, что данное исследование связано с теоретическими разработками - в качестве материальных затрат будут взяты следующие материалы:

Таблица 2.11 – материальные затраты

Наименование материала	Цена за ед., руб	Кол-во, ед.	Сумма, руб.
Набор канцелярских принадлежностей	550	2	1100
Картридж для принтера	1500	2	3000
Итого			4100

Цены приняты на основании прайс-листа поставщика: <https://www.dns-shop.ru/>.

2.3.2 Расчет амортизации специального оборудования

Расчет сводится к определению амортизационных отчислений, так как оборудование эксплуатировалось гораздо ранее начала данной работы, поэтому при расчете будем учитывать только рабочие дни.

Норма амортизации рассчитывается по формуле:

$$НА = \frac{1}{n}, \quad (2.6)$$

где n - срок полезного использования в количестве лет.

Амортизация оборудования рассчитывается по формуле:

$$A = \frac{H_A * I}{12} * m, \quad (2.7)$$

где B – итоговая сумма, тыс. руб.; m – время использования, мес.

При выполнении научно-исследовательского проекта использовался ПК - Asus. Срок полезного использования данного ноутбука по паспорту составляет 3 года.

Таблица 2.12 - Затраты на оборудование

№	Наименование дования	Кол-во, шт.	СПИ, лет	Цена единицы дования, тыс.	Общая стоимость оборудования, тыс.руб.
1	ПК Asus	1	3	40,5	40,5
ИТОГО		40,5 тыс. руб.			

Рассчитаем норму амортизации для ПК: $H_A = \frac{1}{3} = 0,33$

И общую сумму амортизационных отчислений находим следующим образом: $A = \frac{0,33 * 40500}{12} * 3 = 3341,25$ руб

2.3.3 Основная заработная плата исполнителей темы

В данном разделе рассчитывается заработная плата инженера и руководителя, помимо этого, необходимо рассчитать расходы по заработной плате, определяемые трудоемкостью проекта и действующей системой оклада. Основная заработная плата $Z_{осн}$ одного работника рассчитывается по следующей формуле:

$$Z_{осн} = Z_{дн} * T_p, \quad (2.8)$$

где $Z_{дн}$ – среднедневная заработная плата, руб.; T_p – продолжительность работ, выполняемых работником, раб.дн. (таблица 2.9).

Среднедневная заработная плата рассчитывается по формуле (для шестидневной рабочей недели (рабочая неделя руководителя)):

$$З_{дн} = \frac{З_м * М}{F_д} = \frac{51285 * 10,3}{246} = 2147,3 \quad (2.9)$$

где $З_м$ – месячный должностной оклад работника, руб.; $F_д$ – действительный годовой фонд рабочего времени научно-технического персонала, раб. дней;

$М$ – количество месяцев работы без отпуска в течение года: при отпуске в 28 раб. дня: $М = 11,2$ месяца, 5-дневная рабочая неделя; при отпуске в 56 раб. дней – $М = 10,3$ месяца, 6-дневная рабочая неделя.

Для пятидневной рабочей недели (рабочая неделя инженера):

$$З_{дн} = \frac{З_м * М}{F_д} = \frac{33150 * 11,2}{213} = 1743,1 \text{ руб} \quad (2.10)$$

Должностной оклад работника за месяц:

для руководителя:

$$З_м = З_{мс} * (1 + k_{np} + k_d) * k_p, \quad (2.11)$$

$$З_м = 26300 * (1 + 0.3 + 0.2) * 1,3 = 51285 \text{ руб},$$

где $З_{мс}$ – заработная плата, согласно тарифной ставке, руб.; k_{np} – премиальный коэффициент (0,3); k_d – коэффициент доплат и надбавок (0,2); k_p – районный коэффициент (1,3) для г. Томска.

для инженера:

$$З_м = З_{мс} * (1 + k_{np} + k_d) * k_p, \quad (2.12)$$

$$З_м = 17000 * (1 + 0.3 + 0.2) * 1,3 = 33150 \text{ руб}.$$

Таблица 2.13 – Баланс рабочего времени исполнителей

Показатели рабочего времени	Руководитель	Инженер
Календарное число дней	365	365
Количество нерабочих дней выходные дни праздничные дни	52/14	104/14
Потери рабочего времени - отпуск - невыходы по болезни	48/5	24/10
Действительный годовой фонд рабочего времени	246	213

Таблица 2.14 – Расчет основной заработной платы исполнителей

Исполнители НИ	$З_{мс}, руб$	$k_{пр}$	k_{∂}	k_p	$З_{руб.м},$	$З_{\partialн},$ $руб$	$T_p, \text{раб.дн.}$	$З_{осн}, руб$
Руководитель	26300	0,3	0, 2	1,3	51285	2147,3	13,5	28988,6
Инженер	17000	0,3	0, 2	1,3	33150	1743,1	68,5	119402,4
Итого:								148391

Дополнительная заработная плата определяется по формуле:

для руководителя:

$$З_{доп} = k_{доп} * З_{осн} = 0.15 * 28988.6 = 4348.3 \text{ руб.}, \quad (2.12)$$

где $k_{доп}$ – коэффициент дополнительной заработной платы (на стадии проектирования принимаем равным 0,15).

для инженера:

$$З_{доп} = k_{доп} * З_{осн} = 0.15 * 119402.4 = 17910.4 \text{ руб.} \quad (2.13)$$

2.3.4 Отчисления во внебюджетные фонды (страховые отчисления)

Отчисления во внебюджетные фонды определяется по формуле:

для руководителя:

$$З_{внеб} = k_{внеб} * (З_{осн} + З_{доп}) = 0.3 * (28988.6 + 4348.3) = 10001.1 \text{ руб.}, \quad (2.14)$$

где $k_{внеб}$ – коэффициент отчислений на уплату во внебюджетные фонды (пенсионный фонд, фонд ОМС и социальное страхование).

Общая ставка взносов составляет в 2023 году – 30% (ст. 425, 426 НК РФ).

для инженера:

$$З_{внеб} = k_{внеб} * (З_{осн} + З_{доп}) = 0.3 * (119402.4 + 17910.4) = 41193.8 \text{ руб.} \quad (2.15)$$

2.3.5 Накладные расходы

Накладные расходы включают в себя расходы на:

- Печать, ксерокопирование материалов исследования;
- Оплату услуг связи
- Сумма 5 статьи затрат, рассчитанных выше, приведена в таблице ниже

и используются для расчета накладных расходов.

Таблица 2.15 – Группировка затрат по статьям

Статьи					
1	2	3	4	5	6
Амортизация	Сырье, материалы	Основная заработная плата	Дополнительная заработная плата	Отчисления на социальные нужды	Итого без накладных расходов
3341,25	4100	148391	22258.7	51194.9	229285,85

Величина накладных расходов определяется по формуле:

$$Z_{\text{накл}} = (\text{сумма статей}) * k_{\text{нр}} = (229285,85) * 0,2 = 45857,17 \text{ руб.} \quad (2.16)$$

где $k_{\text{нр}}$ – коэффициент, учитывающий накладные расходы.

Величина коэффициента принимается равной 0,2.

2.3.6 Бюджет НИР

На основании полученных данных по отдельным статьям затрат составляется калькуляция плановой себестоимости ВКР «Оценка эффективности процедуры реабилитации после COVID-19» по форме, приведенной в таблице 2.16.

В таблице также представлено определение бюджета затрат конкурирующего проекта - Microsoft Excel.

Различия будут отражены в ЗП сотрудников и количестве оборудования (взято 2 ПК). Средняя заработная плата программиста была принята согласно источнику: <https://www.avito.ru/tomsk/vakansii?cd=1&q=программист>.

Таблица 2.16 – Группировка затрат по статьям

№	Наименование статьи	Сумма, руб.		Примечание
		Текущий Проект	Microsoft Excel	
1	Материальные затраты НИР	4100	4100	Пункт 4.2.3.1
2	Затраты на специальное оборудование	40500	81000	Пункт 4.2.3.2

3	Затраты по основной заработной плате исполнителей темы	148391	409779	Пункт 4.2.3.3
4	Затраты по дополнительной заработной плате исполнителей темы	22258,7	61466,9	Пункт 4.2.3.3
5	Отчисления во внебюджетные фонды	51194,9	141373,8	Пункт 4.2.3.4
6	Накладные расходы	46521,9	128469,4	Пункт 4.2.3.5
Бюджет затрат НИР		266444,6	826189,1	Сумма ст. 1- 6

2.4 Определение ресурсной (ресурсосберегающей), финансовой, бюджетной, социальной и экономической эффективности исследования

Для определения эффективности исследования рассчитан интегральный показатель эффективности научного исследования путем определения интегральных показателей финансовой эффективности и ресурсоэффективности.

Интегральный показатель финансовой эффективности научного исследования получен в процессе оценки бюджета затрат трех программных обеспечений. Для этого наибольший интегральный показатель реализации технической задачи принят за базу расчета (как знаменатель), с которым соотносятся финансовые значения по всем программным обеспечениям.

В качестве аналогов данной ВКР рассмотрены:

Microsoft Excel — программа для работы с экономико-статистическими расчетами, графическими инструментами;

Интегральный финансовый показатель разработки рассчитывается как:

$$I_{\text{финр}}^{\text{исп.}i} = \frac{\Phi_{pi}}{\Phi_{\text{max}}}, \quad (2.17)$$

где: $I_{\text{финр}}^{\text{исп.}i}$ — интегральный финансовый показатель разработки; Φ_{pi} — стоимость i -го варианта исполнения; Φ_{max} — максимальная стоимость исполнения.

$$\Phi_{\text{тек.проект}} = 266444.6; \Phi_{\text{Excel}} = 826189,1.$$

$$I_{\text{финр}}^{\text{тек.проект}} = \frac{266444.6}{826189,1} = 0,32$$

$$I_{\text{финр}}^{\text{Excel}} = \frac{826189,1}{826189,1} = 1,0$$

В результате расчета консолидированных финансовых показателей по двум вариантам разработки вариант 1 (текущий проект) с меньшим перевесом считается более приемлемым с точки зрения финансовой эффективности.

Интегральный показатель ресурсоэффективности вариантов выполнения ВКР (I_{pi}) определен путем сравнительной оценки их характеристик, распределенных с учетом весового коэффициента каждого параметра (таблица 4.17).

Таблица 2.17 – Сравнительная оценка характеристик конкурентов ВКР

Критерии/Объект исследования	Весовой коэффициент параметра	Текущий проект	Excel
Простота эксплуатации	0,2	9	9
Потребность в ресурсах памяти	0,09	9	8
Точность результатов	0,3	8	7
Предоставляемые возможности	0,28	8	5
Скорость работы	0,13	7	6
ИТОГО	1	6,0	3,11

Расчет интегрального показателя для разрабатываемого проекта:

$$I_{\text{тек.пр.}} = 0,2 * 9 + 0,09 * 9 + 0,3 * 8 + 0,28 * 8 + 0,13 * 7 = 6,0;$$

$$I_{\text{Excel}} = 0,2 * 9 + 0,09 * 8 + 0,3 * 7 + 0,28 * 5 + 0,1 * 6 = 3,11;$$

Интегральный показатель эффективности вариантов исполнения разработки вычисляется на основании показателя ресурсоэффективности и интегрального финансового показателя по формуле:

$$I_{\text{исп.}i} = \frac{I_{p-\text{исп.}i}}{I_{\text{финр}}^{\text{исп.}i}} \quad (2.18)$$

$$I_{\text{тек.пр.}} = \frac{6,0}{0,322} = 18,6$$

$$I_{\text{Excel}} = \frac{3,11}{1,0} = 3,11$$

Далее интегральные показатели эффективности данной ВКР сравнивались с интегральными показателями эффективности конкурента с целью определения сравнительной эффективности проекта (таблица 2.18).

Сравнительная эффективность проекта ($\mathcal{E}_{\text{ср}}$):

$$\mathcal{E}_{\text{ср}} = \frac{I_{\text{исп.}i}}{I_{\text{исп.тек.проекта}}} \quad (2.19)$$

Таблица 2.18 – Сравнительная эффективность разработки

№	Показатели	Текущий проект	Excel
1	Интегральный финансовый показатель разработки	0,322	1
2	Интегральный показатель ресурсоэффективности разработки	6,0	3,11
3	Интегральный показатель эффективности	18,6	3,11
4	Сравнительная эффективность вариантов исполнения	1	0,17

Сравнение среднего интегрального показателя показало, что наиболее финансово и ресурсоэффективным является разрабатываемый проект в сравнении с конкурентом.

Заключение по разделу

В результате выполнения целей данного раздела можно сделать следующие выводы:

Результатом анализа конкурентных технических решений является выявление сильных и слабых сторон программного обеспечения относительно конкурентов.

В ходе планирования для руководителя и инженера был разработан график реализации этапов работ, позволяющий оценивать и планировать рабочее время исполнителей. Определено следующее: общее количество календарных дней для выполнения работ - 102 дня; общее количество дней, в течение которых работал инженер - 98 дней, а руководитель - 20 дней.

Для оценки затрат на реализацию проекта разработан проектный бюджет, который составляет 266 444,6 руб.

В результат оценки эффективности ИР можно сделать выводы:

значение интегрального финансового показателя ИР составляет 0,32 - значит ИР является финансово выгодной по сравнению с аналогами;

значение интегрального показателя ресурсоэффективности ИР составляет 6,0 (по сравнению с значением конкурента - 3,11);

значение интегрального показателя эффективности ИР -18,6 (по сравнению с значением конкурента – 3,11) является самым высоким, и, соответственно, техническое решение, рассматриваемое в ИР, является наиболее эффективным вариантом исполнения.

3.Социальная ответственность

Введение

Объектом исследования данной ВКР является база данных из пациентов, прошедших реабилитацию после COVID-19.

Для обеспечения должной эффективности реабилитации пациентов требуется создание общей технологии реабилитации, которая может рассматриваться в качестве базиса для построения персонифицированных программ. А также полученные результаты окажут значимое влияние на работу исследователей – врачей, занимающимися данной темой.

Работа выполняется на персональном компьютере, поэтому в данной главе представляется анализ опасных и вредных факторов при работе с ПК, влияние этих факторов на окружающую среду, а также меры по защите окружающей среды.

Предметом исследования является рабочая зона программиста: ПК, компьютерный стол, клавиатуру, компьютерную мышь и стул.

3.1 Правовые и организационные вопросы обеспечения безопасности

Написание программы, для анализа данных пациентов, прошедших реабилитацию после COVID-19, осуществляется за компьютерным столом, следовательно, рабочее место должно удовлетворять требованиям ГОСТ 12.2.032-78 «Система стандартов безопасности труда (ССБТ). Рабочее место при выполнении работ сидя» [1] и СП 2.2.3670-20 "Санитарно-эпидемиологические требования к условиям труда" [2].

Согласно ст.100 ТК РФ «Рабочее время» [3] режим рабочего времени устанавливается правилами внутреннего трудового распорядка, которые утверждаются работодателем, учитывая мнение представительного органа работников. Также в данном документе регламентируется продолжительность рабочего времени, которая не должна быть меньше указанного времени в трудовом договоре, но и не должна превышать 40 часов в неделю. Вдобавок

ко всему законодательством предусмотрено установление в течение рабочего дня перерывов для питания и отдыха.

Согласно ГОСТ 12.2.032-78 Конструкцией рабочего места должно быть обеспечено выполнение трудовых операций в пределах зоны досягаемости моторного поля, также должно быть обеспечено оптимальное положение работающего, при работе двумя руками органы управления размещают с таким расчетом, чтобы не было перекрещивания рук. В соответствии СП 2.2.3670-20 на рабочем месте, которое предназначено для работы в положении сидя, производственное оборудование и рабочие столы должны иметь пространство для размещения ног высотой не менее 600 мм, глубиной - не менее 450 мм на уровне колен и 600 мм на уровне стоп, шириной не менее 500 мм.

Во время разработки и написания программы было предоставлено рабочее место, где соблюдены все требования по организации труда с ЭВМ.

3.2 Производственная безопасность

При работе с прикладными математическими программами пользователь подвергается воздействию различных вредных и опасных факторов, представленными в таблице 1. А также в таблице приведены соответствующие нормативные документы и этапы работ, во время которых данные факторы могут повлиять на программиста.

Таблица 3.1 – Возможные опасные и вредные факторы

Факторы (ГОСТ 12.0.003-2015)	Нормативные документы
Отклонение показателей микроклимата	СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» [4]
Недостаточная освещённость рабочей зоны	СП 52.13330.2016 «Естественное и искусственное освещение» [5]

Повышенный уровень шума на рабочем месте	СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» [6]
Опасность поражения электрическим током	ГОСТ Р 58698-2019 «Защита от поражения электрическим током» [7]

Исходя из данной таблицы можно заметить, что на пользователя ПК не воздействуют химические и биологические факторы, а только физические и психологические.

3.2.1 Отклонение показателей микроклимата

Отклонение показателей микроклимата во время рабочего процесса от нормы оказывает влияние на здоровье пользователя ПК.

Влияние недостаточной влажности выражается в испарении влаги со слизистых оболочек, что может вызвать пересыхание, растрескивание, а также заражение микробами.

Низкая температура воздуха и повышенная скорость движения воздуха могут вызвать переохлаждение организма путем увеличения интенсивности теплообмена и процесса теплоотдачи во время испарения пота.

Нормы микроклимата производственных помещений регламентируются СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания», регулируют оптимальные и допустимые значения показателей в рабочей зоне, соответствующие физиологическим потребностям организма человека, для создания комфортных и безопасных условий труда.

В ходе анализа базы данных пациентов, используется ПК, который может влиять на микроклимат путем снижения относительной влажности и повышению температуры в рабочем помещении.

Рабочий процесс по энергозатратам относится к категории Ia (производится сидя). Для данной категории в таблицах 2 и 3 приводятся оптимальные и допустимые значения показателей микроклимата на рабочих местах.

Таблица 3.2 – Оптимальные величины показателей микроклимата на рабочих местах

Период года	Температура воздуха, °С	Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с
Холодный	22-24	21-25	60-40	0,1
Тёплый	23-25	22-26	60-40	0,1

Таблица 3.3 – Допустимые величины показателей микроклимата на рабочих местах

Период года	Температура воздуха, °С		Температура поверхностей, °С	Относительная влажность воздуха, %	Скорость движения воздуха, м/с	
	диапазон ниже оптимальных величин	диапазон выше оптимальных величин			для диапазона температур воздуха ниже оптимальных величин, не более	для диапазона температур воздуха выше оптимальных величин, не более
Холодный	20,0 – 21,9	24,1 – 25,0	19,0 – 26,0	15 – 75	0,1	0,1
Тёплый	21,0 – 22,9	25,1 – 28,0	20,0 – 29,0	15 – 75	0,1	0,1

Необходимо осуществлять мероприятия по защите работников от потенциального перегрева и охлаждения, если невозможно поддерживать допустимые величины локального микроклимата: регламентирование рабочего времени и времени для отдыха в помещении с микроклиматом, нормализующим тепловое состояние и уменьшение длительности рабочего дня.

3.2.2 Недостаточная освещенность рабочей зоны

Недостаточная освещенность рабочей зоны является вредным производственным фактором, который приводит к повышению утомляемости и снижению работоспособности человека на предприятии, также длительная работа в условиях низкой освещенности может привести к снижению остроты зрения. Все нормы естественного, искусственного и совместного освещения регламентируются СП 52.13330.2016 «Естественное и искусственное освещение».

Написание программного кода на ПК относится к категории работ высокой точности – Б (наименьший или эквивалентный объект различения 0,30 – 0,50 мм), подразряд 1 (относительная продолжительность зрительной работы при направлении зрения на рабочую поверхность не менее 70%).

В таблице 4 представлены требования к освещению рабочего помещения для разряда Б1.

Таблица 3.4 – Требования к освещению рабочего помещения

Искусственное освещение				Естественное освещение	
Освещенность на рабочей поверхности от системы общего освещения, лк	Цилиндрическая освещенность	Объединенный показатель дискомфорта, не более	Коэффициент пульсации освещенности, Кп, %, не более	Коэффициент естественной освещенности, %, при	
				верхнем или комбинированном	боковом
300	100	21	15	3	1

Повышенное глазное напряжение может быть вызвано ярким светом в зоне периферийного зрения. Чтобы снизить влияние фактора - недостаточную освещенность, необходимо, чтобы уровень естественного освещения рабочего пространства приблизительно совпадал с яркостью дисплея. Эту проблему можно решить с помощью установки дополнительных осветительных приборов.

3.2.3 Повышенный уровень шума на рабочем месте

Превышенный уровень шума на рабочем месте может создать психологический и физический стресс, который также может снизить производительность работника, его концентрацию и вызвать повышение уровня утомляемости. Данный фактор может быть вызван из-за шума, создаваемого работой ПК и кондиционером для воздуха.

Предельно допустимые показатели уровня звука, звукового давления регламентируются СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания». Для разработчика нейронной сети данные показатели отображены в таблице 6. Интенсивный шум, превышающий значение, равное 80 дБА, при длительном воздействии может привести к полной или частичной потере слуха, так что такой максимальный уровень является критическим для разработчика.

Шум приводит к быстрой утомляемости человека, что в свою очередь ведет к производственным ошибкам.

Таблица 3.5 – Предельно допустимые уровни звукового давления, уровни звука и эквивалентные уровни звука для инженера-программиста

Вид трудовой деятельности, рабочее место	Уровни звукового давления, дБ в октавных полосах со среднегеометрическими частотами, Гц						
	31,5	63	125	250	500	1000	2000
Творческая деятельность, руководящая работа с повышенными требованиями, научная деятельность, конструирование и проектирование, программирование, преподавание и обучение, врачебная деятельность. Рабочие места в помещениях дирекции, проектноконструкторских бюро, расчетчиков, программистов вычислительных машин, в лабораториях для теоретических	86	71	61	54	49	45	42

работ и обработки данных, приема больных в здравпунктах							
---	--	--	--	--	--	--	--

Исходя из табличных значений, ориентиром нормы определим уровень шума 50 дБА. При таком уровне шума всё отчётливо слышно.

Для того, чтобы снизить уровень шума на рабочем месте необходимо для всего оборудования систематично проводить техническое обслуживание, так как загрязнение (пыль) способно увеличивать производимый шум.

3.2.4 Опасность поражения электрическим током

Говоря об электробезопасности, подразумевается система мероприятий, направленная на защиту пользователей от опасного воздействия электрического тока, статического электричества и электромагнитного поля. Значения вышеперечисленных факторов регулируются ГОСТ Р 58698-2019.

Таблица 3.6 – Пороги напряжения прикосновения для реагирования

Характер реагирования	Пороги напряжения, В
Реакция испуга	2 (переменный ток)
	8 (постоянный ток)
Мышечная реакция	20 (переменный ток)
	40 (постоянный ток)

Чтобы избежать влияния подобных опасных факторов, необходимо принять следующие меры: использовать защитные ограждения или оболочки; настроить автоматическое отключение питания; размещать опасные для жизни и здоровья человека участки электропроводов и приборов вне зоны досягаемости рукой; осуществлять питание от безопасного источника.

3.3 Экологическая безопасность

Затем необходимо проанализировать, какое влияние на окружающую среду оказывает написанная программа по исследованию базы данных пациентов и возможные источники загрязнения окружающей среды, которые могут возникнуть в ходе написания ВКР.

Сама программа не создает вредные выбросы, которые могли бы повлиять на гидросферу, литосферу и атмосферу.

Повышения радиационного фона не было, но ПК оказывает влияние на окружающий воздух, а именно при его нагреве воздух становится сухим. Также неправильная утилизация компьютера и его составляющих может привести к загрязнению литосферы, поэтому необходимо извлекать его комплектующие, сортировать их и отправлять на повторное использование для уменьшения вреда окружающей среде.

При соблюдении всех норм во время использования и утилизации компьютерной техники вредное воздействие на окружающую среду будет минимизировано.

Одна из главных проблем при защите окружающей среды в компьютерных лабораториях - утилизация люминесцентных ламп, которые в себе содержат ртуть в дозах от 1 до 70 мг, ядовитое вещество, относящееся к 1-ому классу опасности. Такое количество ядовитого вещества может причинить вред здоровью при повреждении лампы.

Хранение и удаление отходов, именно люминесцентных ламп, происходят в соответствии с требованиями Постановления Правительства Российской Федерации от 28.12.2020 № 2314 «Об утверждении Правил обращения с отходами производства и потребления в части осветительных устройств, электрических ламп, ненадлежащие сбор, накопление, использование, обезвреживание, транспортирование и размещение которых может повлечь причинение вреда жизни, здоровью граждан, вреда животным, растениям и окружающей среде» [8].

Герметически стальной крышкой закрывают наполненную тару с отходами, если есть необходимость, то ее заваривают и передают по договору специализированным предприятиям, которые имеют на их утилизацию.

3.4 Безопасность в чрезвычайных ситуациях

3.4.1 Возникновение пожара

Пожар может возникнуть при работе с компьютером из-за разных причин: короткое замыкание, курение, сильный перегрев.

Для того, чтобы избежать этого, пользователю ПК необходимо не курить, иметь средства пожаротушения в рабочем состоянии, своевременно диагностировать и обслуживать надлежащим образом свою технику.

Рабочее здание должно удовлетворять всем требованиям пожарной безопасности: обязательное присутствие плана эвакуации, пожарной сигнализации, углекислотного огнетушителя с проверенным клеймом, эвакуационных выходов и схем, которые к ним ведут.

Согласно СП 12.13130.2009 [9] рабочее место разработчика программного обеспечения относится к категории «В», (пожароопасные), так как в данном помещении присутствует пыль, вещества и материалы, которые при взаимодействии с воздухом могут гореть.

В рабочем здании присутствуют план эвакуации, огнетушители, противопожарная сигнализация, датчики дыма, имеются должностные люди - ответственные за противопожарную безопасность.

Заключение по разделу

В ходе написания раздела «Социальная ответственность» были рассмотрены основные нормативные акты, регламентирующую безопасную работу за ПК, проанализированы опасные и вредные факторы, которые могут возникнуть в ходе рабочего процесса.

В ходе написания ВКР все меры безопасности были соблюдены, вредное воздействие на окружающую среду минимизированы, описаны меры безопасности для наиболее типичной ЧС.

В ходе написания раздела «Социальная ответственность» были рассмотрены правовые и организационные вопросы безопасности на рабочем месте, которые должны удовлетворять всем требованиям безопасности. Требуемое освещение обеспечивается за счет нескольких энергосберегающих ламп. Уровень шума находится в допустимом диапазоне. Микроклиматические условия соблюдаются за счет системы отопления в

холодное время и проветривания в теплое время. Защита от воздействия электрического тока обеспечивается путем проверки состояния ПК и соблюдения правил безопасности при работе с ним. Рабочее место имеет все необходимые средства для обеспечения пожарной безопасности.

В соответствии ПУЭ, рабочее место относится к 1 категории по электробезопасности: помещение без повышенной опасности.

Для работы в офисе персоналу присваивается 1 группа по электробезопасности.

В соответствии СанПиН 1.2.3685-21 работа пользователя-разработчика относится к категории 1б, так как он проводит большую часть времени за компьютером.

В соответствии СП 12.13130.2009 рабочее место относится к категории «В», (пожароопасные), потому что в нем присутствует вещества и материалы, способные гореть при взаимодействии с воздухом.

В соответствии Постановлению от 31 декабря 2020 года №2398 [10] категория объекта, то есть ПЭВМ, оказывающего негативное воздействие на окружающую среду, относится к IV категории.

В заключение также были рассмотрены пункты, связанные с экологической безопасностью, уменьшением вредного влияния на окружающую среду и человека, безопасностью в чрезвычайных ситуациях и меры по их предупреждению.

Заключение

В результате проделанной работы были получены следующие выводы:

1. Отобраны информативные признаки для оценки эффективности процедур реабилитации - пробы Генчи и Штанге, шкала Борга, ТШХ, ИФИ, тест на эмоциональный интеллект.
2. Показано, что по каждому из информативных параметров есть высоко-значимый эффект процедур реабилитации ($p < 0,000001$).
3. Было выявлено, что нет значимого различия в эффекте реабилитации при добавлении к физиолечению пробиотических продуктов и продуктов пантового оленеводства. Все различия результатов в данных группах обусловлены лишь изначальным различием средних значений информативных параметров в данных группах.
4. Показано, что по ряду показателей (ТШХ, шкала Борга) на эффективность реабилитации влияет возраст пациентов.
5. Была проведена кластеризация всей совокупности пациентов по основным информативным признакам, по результатам которой пациенты были разделены на группы с различными реабилитационными диагнозами.

Список используемых источников

1. Дисперсионный анализ [Электронный ресурс] // URL: [https://portal.tpu.ru/SHARED/s/SCHINKEEV/UMM/Tab5/MU_4_0\(DA\).docx](https://portal.tpu.ru/SHARED/s/SCHINKEEV/UMM/Tab5/MU_4_0(DA).docx) (дата обращения: 20.04.2023 г.)
2. ANOVA Articles - Statistics By Jim [Электронный ресурс] // URL: <https://statisticsbyjim.com/anova/> (дата обращения: 02.05.2023 г.)
3. Метод главных компонент [Электронный ресурс] // URL: [https://portal.tpu.ru/SHARED/s/SCHINKEEV/UMM/Tab5/MU_5\(MGK\).docx](https://portal.tpu.ru/SHARED/s/SCHINKEEV/UMM/Tab5/MU_5(MGK).docx) (дата обращения: 25.04.2023 г.)
4. Статистические методы обработки и планирования эксперимента : учебное пособие [Электронный ресурс] // URL: <https://e-univers.ru/upload/iblock/27b/3gp9agb1utoekedgth43qp78gy44p0br.pdf> (дата обращения: 27.04.2023 г.)
5. Дисперсионный анализ [Электронный ресурс] // URL: <https://studfile.net/preview/9967790/page:14/> (дата обращения: 30.04.2023 г.)
6. Анализ главных компонент (PCA) [Электронный ресурс] // URL: <https://www.helenkapatsa.ru/mietod-ghlavnykh-komponent/> (дата обращения: 02.05.2023 г.)
7. Кластерный анализ [Электронный ресурс] // URL: <https://studfile.net/preview/6327567/page:35/> (дата обращения: 05.05.2023 г.)
8. Кластерный анализ [Электронный ресурс] // URL: [https://portal.tpu.ru/SHARED/s/SCHINKEEV/UMM/Tab5/MU_9\(CLA\).docx](https://portal.tpu.ru/SHARED/s/SCHINKEEV/UMM/Tab5/MU_9(CLA).docx) (дата обращения: 06.05.2023 г.)
9. Шкала Борга. Оценка переносимости физической нагрузки [Электронный ресурс] // URL: <https://shopdon.ru/blog/shkala-borga-otsenka-perenosimosti-fizicheskoy-nagruzki/> (дата обращения: 07.05.2023 г.)

- 10.Проведение пробы с физической нагрузкой [Электронный ресурс] // URL: <https://studfile.net/preview/1730682/page:17/>(дата обращения: 05.05.2023 г.)
- 11.Тест с 6-минутной ходьбой Методическое пособие по проведению профилактического медицинского осмотра и диспансеризации определенных групп взрослого населения, углубленной диспансеризации для граждан, перенесших новую коронавирусную инфекцию (COVID-19) (утв. Минздравом России) [Электронный ресурс] // URL: <https://sudact.ru/law/metodicheskoe-posobie-po-provedeniiu-profilakticheskogo-meditsinskogo-osmotra/8/8.1-test-s-6-minutnoi-khodboi/>(дата обращения: 02.05.2023 г.)
- 12.ГОСТ 12.2.032-78 Система стандартов безопасности труда (ССБТ). Рабочее место при выполнении работ сидя. Общие эргономические требования // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200003913> (дата обращения: 16.05.2023).
- 13.СП 2.2.3670-20 "Санитарно-эпидемиологические требования к условиям труда" // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573230583> (дата обращения: 16.05.2023).
- 14.Трудовой кодекс (ТК РФ) «Рабочее время»// Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <http://base.garant.ru/12125268/> (дата обращения: 16.05.2023).
- 15.СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. 2021. – Режим

- доступа: <https://docs.cntd.ru/document/573500115> (дата обращения: 19.05.2023).
- 16.СП 52.13330.2016 «Естественное и искусственное освещение» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/456054197> (дата обращения: 16.05.2023).
- 17.СанПиН 1.2.3685-21 «Гигиенические нормативы и требования к обеспечению безопасности и (или) безвредности для человека факторов среды обитания» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573500115> (дата обращения: 19.05.2023).
- 18.ГОСТ Р 58698-2019 «Защита от поражения электрическим током» // Электронный фонд правовой и нормативно-технической документации [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200170001> (дата обращения: 16.05.2023).
- 19.Постановление Правительства Российской Федерации от 28.12.2020 № 2314 «Об утверждении Правил обращения с отходами производства и потребления в части осветительных устройств, электрических ламп, ненадлежащие сбор, накопление, использование, обезвреживание, транспортирование и размещение которых может повлечь причинение вреда жизни, здоровью граждан, вреда животным, растениям и окружающей среде» [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/573275660> (дата обращения: 19.05.2023).
- 20.СП 12.13130.2009 «Системы противопожарной защиты. Эвакуационные пути и выходы» [Электронный ресурс]. – Режим доступа: <https://docs.cntd.ru/document/1200071156> (дата обращения: 19.05.2023).
- 21.Постановлению от 31 декабря 2020 года №2398 «Об утверждении критериев отнесения объектов, оказывающих негативное воздействие на окружающую среду, к объектам I, II, III и IV категорий»

[Электронный ресурс]. – Режим доступа:
<https://docs.cntd.ru/document/573292854> (дата обращения: 19.05.2023).