

---

# Использование методов машинного обучения для идентификации типа вируса

---

A Preprint

Анастасия Алексеевна Елисова  
Факультет ВМК  
МГУ имени М.В.Ломоносова  
Москва, Россия  
s02210381@gse.cs.msu.ru

Олег Валентинович Сенько  
Факультет ВМК  
МГУ имени М.В.Ломоносова  
Москва, Россия  
senkoov@mail.ru

## Abstract

Данная статья посвящена выявлению зависимостей респираторных вирусных инфекций: гриппа, аденовируса, респираторно-синцитиального вируса и бокавируса от предоставленных данных (симптомов) на основе ансамблевых методов, а также тестов. Основным методом решения задач компьютерной диагностики является обучение алгоритма машинного обучения на основе выборки, которая состоит из описания отдельных пациентов. Однако такая информация часто ограничена рамками одного медицинского учреждения и имеет небольшой объем. В результате настоящей работы были проведены эксперименты и вычисления, позволяющие выявить статистическую и линейную зависимости между данными.

Keywords Машинное обучение · Статистический анализ · Вирусы

## 1 Introduction

Острые респираторные вирусные инфекции (ОРВИ) оказывают значительное влияние на системы здравоохранения по всему миру из-за высокой заразности и уровня госпитализаций, особенно среди детей, пожилых и людей с ослабленным иммунитетом. Эти инфекции являются одной из самых частых причин обращения за медицинской помощью, занимая лидирующие позиции по заболеваемости и смертности среди инфекционных заболеваний. Диагностика и дифференцировка этих инфекций по симптомам затруднены, особенно в условиях ограниченности данных, что ставит перед медицинскими специалистами задачу правильного и своевременного распознавания типа вируса. Ансамблевые методы машинного обучения и статистический анализ помогают повысить точность диагностики, а также выявить статистически значимые взаимосвязи между симптомами и вирусами.

Современная диагностика респираторных инфекций в основном базируется на клинических данных, таких как жалобы пациентов, анамнез, физикальные исследования, а также лабораторных тестах. Однако в условиях массовых обращений и недостатка лабораторных ресурсов точность диагностики может быть низкой. Традиционные методы, такие как ПЦР-диагностика или серологические тесты, зачастую требуют времени и значительных финансовых затрат. С этим связана необходимость внедрения более быстрых и доступных методов диагностики, например, с использованием алгоритмов машинного обучения, которые могут анализировать большое количество данных и предоставлять результаты с высокой точностью и в краткие сроки.

Вирус гриппа типа А — один из наиболее опасных вирусов, вызывающих как сезонные эпидемии, так и пандемии. Он имеет высокую изменчивость антигенов, что делает невозможным долгосрочное создание универсальной вакцины. Основные симптомы включают высокую температуру, озноб, слабость, боль в мышцах и головную боль, что отличает его от большинства других ОРВИ. Из-за высокой мутабельности

вируса, эпидемии происходят с периодической сменой штаммов, что требует регулярного обновления вакцин.

Аденовирус — вирус, вызывающий широкий спектр инфекций, включая респираторные, глазные и кишечные заболевания. Симптомы включают лихорадку, боль в горле, кашель, насморк и конъюнктивит. Аденовирусы могут поражать как верхние, так и нижние дыхательные пути и устойчивы к неблагоприятным условиям окружающей среды, что способствует вспышкам в закрытых детских учреждениях и медицинских организациях. Эти вирусы могут вызывать тяжёлые формы заболеваний у людей с ослабленным иммунитетом, таких как пациенты с онкологическими заболеваниями или трансплантированными органами.

Бокавирус — относительно недавно обнаруженный вирус, часто связанный с острыми респираторными инфекциями, особенно у детей. Он может вызывать симптомы, напоминающие грипп и РСВ, такие как кашель, лихорадка и затруднённое дыхание. Бокавирус часто сосуществует с другими вирусами, что затрудняет его диагностику и требует комплексного подхода для точного распознавания. Особенно важно учитывать его наличие при лечении пневмоний и бронхитов, когда обычные методы диагностики могут не дать точных результатов.

Риновирус — основная причина простуды и одна из наиболее частых причин заболеваний верхних дыхательных путей. Симптомы риновирусной инфекции включают насморк, боль в горле, головную боль и иногда кашель. В отличие от других ОРВИ, риновирусы редко сопровождаются высокой температурой, что делает их диагностику затруднённой, особенно в зимний период, когда происходит повышенная циркуляция других вирусов.

Респираторно-синцитиальный вирус (РСВ) — вирус, вызывающий инфекции нижних дыхательных путей, особенно у младенцев и пожилых людей. Он является одной из ведущих причин бронхолита и пневмонии. Основные симптомы РСВ включают лихорадку, кашель, хрипы и затруднённое дыхание, что требует своевременной диагностики для назначения адекватного лечения. Существует также повышенный риск осложнений у людей с хроническими заболеваниями лёгких и сердечно-сосудистой системы.

Настоящее исследование сосредоточено на использовании ансамблевых методов машинного обучения для повышения точности диагностики респираторных вирусных инфекций. Целью работы является выявление ключевых симптомов, отличающих каждую из предложенных инфекций, и определение типа модели, которая на основе этих симптомов будет эффективно распознавать тип вируса. В ходе исследования используются статистические методы для оценки значимости симптомов, а также алгоритмы машинного обучения для того, чтобы улучшить диагностику и лечение ОРВИ в клинической практике. Модели машинного обучения, такие как случайные леса, градиентный бустинг и нейронные сети, способны эффективно обрабатывать большое количество признаков и выявлять сложные зависимости между симптомами и вирусами. На основе этих данных могут быть построены системы раннего оповещения и диагностики, что поможет снизить заболеваемость и смертность от ОРВИ.

## 2 Постановка задачи

В данной работе использована таблица с информацией о характеристиках гриппа, аденовируса, респираторно-синцитиального вируса и бокавируса у 1533 пациентов. Симптомы, рассматриваемые в данном исследовании, включают температуру, кашель, одышку, боли в горле, боль в груди, ринит, головную боль, диарею боли в животе, слабость и многие другие.

Для оценки качества используется метрика ROC-AUC, как рекомендуемая Министерством Здравоохранения для использования в доказательной медицине.

### 3 Анализ вирусов

	Количество нулей	Количество единиц
Адено	108	1425
Бока	62	1471
Рино	313	1220
ГриппА	157	1376
РС-вирус	211	1322

Таблица 1: Соотношение количества позитивных и негативных ответов для всех типов вирусов

Таким образом, можно сделать вывод о том, что данные не сбалансированы.

Симптом \ Тип вируса	Грипп	РС	Адено	Бока	Рино
Температура	216	218	209	234	207
Кашель	897	818	947	935	818
Одышка	99	81	97	89	66
Боли в горле	73	80	69	82	69
Боль в груди	42	47	42	47	42
Ринит	601	543	625	634	514
Головная боль	42	47	42	47	42
Диарея	51	43	55	54	50
Боли в животе	45	46	43	48	38

Таблица 2: Информация о количестве зафиксированных симптомов для каждого из заболеваний

### 4 Анализ симптомов

#### 4.1 Точный тест Фишера

Точный тест Фишера — это статистический тест, используемый для проверки гипотезы о независимости двух категориальных переменных, особенно при малых выборках. В данном случае, мы применяем этот тест для анализа зависимости между типом вируса и симптомом на основе таблицы сопряженности 2 на 2. В отличие от более традиционного критерия независимости Хи-квадрат, точный тест Фишера не делает предположений о размере выборки и распределении данных, что делает его более подходящим для анализа малых выборок, когда другие методы могут быть не применимы.

Тест Фишера используется для вычисления точного р-значения, которое помогает установить, существует ли статистически значимая зависимость между двумя переменными, или же они независимы друг от друга.

##### 4.1.1 Сравнение с критерием Хи-квадрат

Тест независимости Хи-квадрат является традиционным методом для анализа связи между категориальными переменными. Он используется для оценки того, насколько данные, полученные из наблюдений, отличаются от теоретически ожидаемых, если переменные независимы. Однако, распределение выборки по методу Хи-квадрат является приближением правильного распределения, и при малых значениях ячеек таблицы (меньше 5) результаты могут быть недостоверными.

В отличие от этого, точный тест Фишера не требует предположений о размере выборки, что делает его более подходящим для случаев с малыми выборками, где распределение по Хи-квадрат может быть ненадежным.

##### 4.1.2 Гипотезы

Для точного теста Фишера устанавливаются следующие гипотезы:

- $H_0$  (нулевая гипотеза): Две переменные независимы.
- $H_1$  (альтернативная гипотеза): Две переменные зависимы.

Таким образом, задача теста — проверить, есть ли статистически значимая зависимость между типом вируса и симптомом, или же эти переменные независимы.

Для выполнения точного теста Фишера необходимо построить таблицу сопряженности для двух категориальных переменных. Таблица сопряженности размером 2 на 2 выглядит следующим образом:

	Группа 1	Группа 2	Сумма по строкам
Категория 1	a	b	a + b
Категория 2	c	d	c + d
Сумма по столбцам	a + c	b + d	a + b + c + d = n

Таблица 3: Таблица сопряженности для точного теста Фишера

Здесь:

- $a$  — количество наблюдений, которые принадлежат обеим категориям (например, вирус 1 и симптом 1).
- $b$  — количество наблюдений, которые принадлежат категории 1, но не категории 2 (например, вирус 1 и симптом 2).
- $c$  — количество наблюдений, которые принадлежат категории 2, но не категории 1 (например, вирус 2 и симптом 1).
- $d$  — количество наблюдений, которые принадлежат обеим категориям (например, вирус 2 и симптом 2).
- $n$  — общий объем выборки.

#### 4.1.3 Вычисление p-значения

Точное p-значение для теста Фишера вычисляется по следующей формуле:

$$p\_value = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{a!b!c!d!n!} \quad (1)$$

Это выражение основано на перестановке всех возможных способов распределить наблюдения по ячейкам таблицы сопряженности, учитывая фиксированные суммы строк и столбцов. Чем меньше p-значение, тем более вероятно, что существует зависимость между переменными, и тем более убедительным является отклонение нулевой гипотезы.

После вычисления p-значения его сравнивают с заранее установленным уровнем значимости, обычно равным  $p = 0.05$ . Если p-значение меньше этого порога, то нулевая гипотеза отклоняется, и делается вывод о наличии статистически значимой зависимости между переменными. Если p-значение больше установленного уровня значимости, нулевая гипотеза не отклоняется, что указывает на отсутствие статистически значимой зависимости.

- Если  $p\_value > 0.05$ , то принимаем нулевую гипотезу и заключаем, что переменные независимы.
- Если  $p\_value \leq 0.05$ , то отклоняем нулевую гипотезу и заключаем, что существует зависимость между переменными.

(Для данного исследования в соответствии с рекомендациями Министерства Здравоохранения установлен критический уровень значимости  $p = 0.05$ .)

Симптом Тип вируса	Ринит	Боль в ухе	Боль в горле	Рвота	Нарушение сна	Субфебрилитет
Адено	0.086950	1.000000	0.001368	0.116175	0.037012	0.594665
Бока	0.601827	0.458161	0.252738	0.252072	1.000000	1.000000
Рино	0.084480	1.000000	0.484637	0.911745	0.476857	0.009667
ГриппА	0.349835	0.160121	0.575432	0.373212	0.616386	0.881994
РС-вирус	0.000025	1.000000	0.002916	0.601528	0.654996	0.895526

Симптом Тип вируса	Трудное дыхание	Боли в животе	Жидкий стул	Озноб	Чихание	Одышка
Адено	0.512465	0.380875	0.428021	0.285701	1.000000	0.041070
Бока	0.732766	0.259017	0.499241	1.000000	1.000000	0.005093
Рино	0.000028	1.000000	0.133674	1.000000	1.000000	0.001807
ГриппА	0.000091	0.471761	0.826338	1.000000	1.000000	0.000023
РС-вирус	0.436347	0.053472	0.028295	0.708457	0.447305	0.225688

Симптом Тип вируса	ЧБД	Аллергоанамнез	Антибиотики на догосп этапе	Температура 37-37,9
Адено	0.694798	0.034326	0.106807	0.058522
Бока	0.303352	0.541860	0.765073	0.078072
Рино	0.559318	0.000341	0.341007	0.053765
ГриппА	0.375182	0.028982	0.079896	0.859641
РС-вирус	0.170856	0.001241	0.347133	0.119859

Симптом Тип вируса	Температура 38-39	температура выше 39,1	Слабость	Головная боль
Адено	0.724977	0.000160	0.250843	0.375466
Бока	0.541186	0.378364	0.352762	0.257810
Рино	0.004546	0.008819	1.000000	0.099648
ГриппА	0.012971	0.908224	0.010300	0.809586
РС-вирус	0.793853	0.041301	0.001268	0.001787

Симптом Тип вируса	Мышечная боль	Кашель	ОРВИ	Бронхит	Пневмония	ЭБМ
Адено	1.000000	0.000000	0.000173	0.000138	0.022946	0.042460
Бока	0.000066	0.257810	0.770096	0.760127	0.331941	0.330939
Рино	1.000000	0.000053	0.000017	0.204727	0.000122	0.348508
ГриппА	0.013584	0.112970	0.057239	0.005271	0.019778	0.011198
РС-вирус	0.374557	0.000000	0.007488	0.000000	0.709289	0.357972

Таблица 4: Результаты вычисления точного критерия Фишера

#### 4.2 Тест Манна-Уитни

U-критерий Манна-Уитни, также известный как ранговый суммарный критерий Уилкоксона, является мощным непараметрическим тестом, предназначенным для сравнения двух независимых выборок. Он используется в случаях, когда данные не соответствуют нормальному распределению, и его основное преимущество заключается в том, что он не требует предположений о форме распределения данных, в отличие от традиционного t-критерия Стьюдента, который требует нормальности распределений.

Этот тест был предложен в 1947 году Манном и Уитни в контексте анализа рангов и стал основным инструментом для сравнения двух групп в статистике. U-критерий является предпочтительным выбором в случае, когда предположения о нормальности данных не могут быть выполнены, или когда данные представлены в порядковой шкале.

### 4.3 Принцип работы теста

U-критерий Манна-Уитни оценивает различия между двумя независимыми выборками на основе рангов их наблюдений. Процесс расчета статистики  $U$  включает несколько ключевых шагов:

1. Объединение выборок: Обе выборки объединяются в одну общую выборку.
2. Присваивание рангов: Каждому наблюдению присваивается ранг в пределах объединенной выборки. Для одинаковых значений наблюдений используются средние ранги.
3. Вычисление суммы рангов: Для каждой из выборок вычисляется сумма рангов.
4. Расчет статистики  $U$ : На основе сумм рангов рассчитывается статистика  $U$ , которая сравнивает их различия.

Если одна выборка склонна иметь более высокие или более низкие значения, статистика  $U$  будет соответствующим образом отражать это, и гипотеза о равенстве распределений будет отклонена.

### 4.4 Формула и интерпретация U-критерия Манна-Уитни

U-критерий Манна-Уитни вычисляется по формуле:

$$U = n_x n_y + \frac{n(n+1)}{2} - T, \quad (2)$$

где:

- $n_x$  и  $n_y$  — объемы выборок (количество элементов в выборках  $X$  и  $Y$  соответственно);
- $n$  — общий объем объединенной выборки;
- $T$  — сумма рангов для одной из выборок (например, для выборки  $X$  или  $Y$ ).

Если выборка  $X$  имеет меньшую сумму рангов, то соответствующее значение статистики  $U$  будет меньше, и наоборот. Кроме того, тест может быть использован для проверки гипотезы о том, что две выборки происходят из одинаковых распределений, или для проверки, что одна выборка в среднем больше (или меньше) другой.

### 4.5 Рассмотрение статистической значимости

Нулевая гипотеза ( $H_0$ ) для теста Манна-Уитни предполагает, что обе выборки принадлежат одинаковому распределению, то есть что они не имеют статистически значимых различий. Альтернативная гипотеза ( $H_1$ ) утверждает, что между выборками существуют различия.

- Если  $p$ -значение для статистики  $U$  меньше уровня значимости  $\alpha$  (обычно 0,05), то нулевая гипотеза отклоняется.
- Если  $p$ -значение больше уровня значимости, то нулевая гипотеза не отклоняется, и мы заключаем, что нет достаточных доказательств для утверждения о статистически значимых различиях.

Важно, что тест Манна-Уитни является двусторонним, то есть можно проверить гипотезу о том, что одна из выборок может быть статистически больше или меньше другой.

#### 4.5.1 Преимущества и недостатки

- Не требует нормальности данных: Тест Манна-Уитни не имеет строгих предположений о нормальности данных, что делает его идеальным для работы с негауссовыми распределениями.
- Устойчивость к выбросам: Из-за использования рангов, а не исходных значений данных, тест менее чувствителен к выбросам, что делает его надежным инструментом в реальных приложениях.
- Подходит для различных типов данных: Тест можно применять к порядковым, числовым и непрерывным данным.

- Простота в применении: В отличие от многих других статистических тестов, U-критерий Манна-Уитни легко реализуем и широко поддерживается в статистических пакетах.
- Низкая мощность для малых выборок: Как и многие непараметрические тесты, тест Манна-Уитни может иметь ограниченную мощность, особенно для малых выборок. Это может привести к неверному принятию нулевой гипотезы, если различия между выборками малы.
- Не предоставляет подробной информации о различиях: Тест только указывает на наличие или отсутствие статистически значимых различий, но не предоставляет информации о величине или характере этих различий.
- Зависимость от размера выборки: Когда объем выборки сильно различается между двумя группами, точность теста может снизиться, что влияет на его результаты.

Тест используется в любых ситуациях, где важно оценить различия между двумя группами, особенно если данные не подчиняются нормальному распределению.

Симптом Тип вируса	Возраст	Год	День болезни	Температура за несколько дней до поступлен.
Адено	0.000003	0.000025	0.011808	0.000029
Бока	0.000001	0.000003	0.000225	0.000533
Рино	0.002273	0.046081	0.028801	0.049020
ГриппА	0.000023	0.000039	0.005980	0.000341
Рс-вирус	0.000093	0.001092	0.012898	0.000647

Симптом Тип вируса	Эозинофилы	Моноциты	Мочевина	Креатинин	п/ядерные
Адено	0.000035	0.000102	0.000027	0.000021	0.000163
Бока	0.000009	0.000061	0.000000	0.000000	0.000012
Рино	0.071484	0.015527	0.000000	0.000006	0.002841
ГриппА	0.000001	0.000850	0.000000	0.000077	0.003311
Рс-вирус	0.000131	0.000559	0.000001	0.000006	0.003355

Таблица 5: Результаты вычисления U-критерия Манна-Уитни

#### 4.6 Перестановочный тест

Тесты на перестановку, или тесты рандомизации, представляют собой мощный инструмент для проверки гипотез и оценки статистической значимости. Эти тесты не делают строгих предположений о распределении данных и широко применяются для анализа различных типов данных. В отличие от традиционных параметрических тестов, таких как t-критерий, которые зависят от конкретных предположений о данных (например, о нормальности распределения), перестановочные тесты используют данные напрямую, переставляя метки наблюдений, чтобы оценить распределение статистики при нулевой гипотезе.

##### 4.6.1 Принцип работы перестановочного теста

Перестановочный тест основан на вычислении всех возможных значений тестовой статистики при перестановке меток в наблюдаемых данных. Процесс включает в себя следующие этапы:

1. Вычисление статистики теста для исходных данных: Для данных вычисляется значение тестовой статистики, например, разница между средними двух выборок.
2. Перестановка меток: При каждой перестановке метки (например, метки группы в эксперименте или лечение) случайным образом переносятся между наблюдениями, создавая новые выборки с перемешанными метками.
3. Вычисление статистики для каждой перестановки: Для каждой перестановки вычисляется новая тестовая статистика.

4. Оценка уровня значимости: После множества перестановок (обычно несколько тысяч) распределение тестовой статистики при нулевой гипотезе строится, и определяется, насколько редким является значение статистики, полученное на реальных данных. Уровень значимости  $p$ -значение вычисляется как доля перестановок, для которых статистика теста была больше или равна исходному значению.

Перестановочный тест может быть использован для проверки различных гипотез, включая гипотезы о разнице между средними, медианами или других статистиках. Стандартная формулировка гипотез в контексте перестановочного теста выглядит следующим образом:

- $H_0$  (нулевая гипотеза): Нет различий между группами, т.е. наблюдаемые различия могут быть объяснены случайными вариациями.
- $H_1$  (альтернативная гипотеза): Есть статистически значимые различия между группами, которые не могут быть объяснены случайными вариациями.

Таким образом, основной задачей перестановочного теста является проверка, можно ли объяснить наблюдаемое различие случайным образом, исходя из предположения о том, что группы или условия эксперимента не влияют друг на друга.

#### 4.6.2 Преимущества и недостатки перестановочного теста

- Не требует предположений о распределении: Перестановочные тесты не зависят от специфических предположений о нормальности данных или о форме распределения, что делает их универсальными и подходящими для широкого круга задач.
- Гибкость: Тесты можно применять для различных типов данных и статистик. Например, можно сравнивать средние значения, медианы, ранги, корреляции и другие статистики.
- Точность: Перестановочные тесты дают точные уровни значимости, так как основаны на эмпирическом распределении статистики теста.
- Простота интерпретации: Результаты перестановочного теста легко интерпретируются, так как  $p$ -значение указывает на вероятность получения наблюдаемого эффекта случайным образом.
- Вычислительная нагрузка: Перестановочные тесты могут быть вычислительно затратными, особенно если выборки большие и требуется большое количество перестановок для получения стабильных результатов.
- Ограничение на тип данных: Хотя перестановочные тесты очень гибки, они лучше всего работают с категориальными или числовыми данными, где перестановка меток имеет смысл.

Симптом Тип вируса	Прокалицитонин при выписке	Базофилы	Эозинофилы %	П/ядерные %
Адено	0.000003	0.000132	0.000029	0.000163
Бока	0.000088	0.000718	0.000533	0.000012
Рино	0.000088	0.005729	0.049020	0.015154
ГриппА	0.000003	0.005729	0.000341	0.001000
Рс-вирус	0.000003	0.005729	0.000647	0.000225

Симптом Тип вируса	Моноциты	Общий белок	Билирубин	Мочевина	Креатинин
Адено	0.000102	0.000008	0.000024	0.000027	0.000021
Бока	0.000061	0.000008	0.000097	0.000000	0.000000
Рино	0.015527	0.000239	0.000647	0.000000	0.000006
ГриппА	0.000850	0.034390	0.031824	0.000000	0.000077
Рс-вирус	0.000559	0.001091	0.000647	0.000001	0.000006

Таблица 6: Результаты вычисления перестановочного теста



## 5 Выводы

В ходе проведённого исследования были сделаны следующие ключевые выводы:

- Зависимости между симптомами и вирусами. Анализ данных показал статистические зависимости между симптомами и типами респираторных вирусных инфекций, таких как грипп, аденовирус, респираторно-синцитиальный вирус (РСВ) и бокавирус. Это позволило глубже понять, какие симптомы наиболее характерны для разных вирусов, что способствует точной диагностике в условиях ограниченных данных.
- Эффективность ансамблевых методов машинного обучения. Ансамблевые методы, включая случайные леса, градиентный бустинг и нейронные сети, продемонстрировали высокую эффективность в классификации и анализе симптомов, что может значительно повысить точность диагностики респираторных инфекций. Эти методы могут быть использованы в клинической практике для разработки систем раннего оповещения.
- Ограниченность данных и её влияние на результаты. Одним из значимых аспектов исследования является ограниченность доступных данных, что требует применения методов, способных работать с малым объёмом выборок и компенсировать их недостатки. Это подчеркивает важность использования статистических методов, таких как точный тест Фишера, для корректного анализа при малых выборках.
- Проблемы с несбалансированностью данных. Несбалансированность классов в данных (например, неравномерное количество случаев для разных вирусов) представляет собой серьёзную проблему. Для преодоления этой проблемы необходимо использовать методы, которые способны корректировать несоответствия классов, чтобы избежать искажений в результатах.
- Применение статистических методов для улучшения диагностики. Применение статистических тестов, таких как точный тест Фишера и критерий Манна-Уитни, позволило выявить важные зависимости между симптомами и типами вирусов, а также оценить значимость тех или иных симптомов для диагностики. Это может послужить основой для разработки более точных и быстрых диагностических инструментов.
- Перспективы применения машинного обучения в медицинской диагностике. Результаты работы подтверждают, что использование машинного обучения для диагностики респираторных вирусных инфекций может значительно улучшить результаты клинической практики, повысить скорость диагностики и снизить затраты на лабораторные исследования. В будущем внедрение таких технологий в повседневную медицинскую практику может существенно повысить эффективность диагностики и лечения респираторных заболеваний.

## Список литературы

- [1] Материалы ММКЦ "Коммунарка"
- [2] Изучение влияния клинико-генетических факторов на течение дисциркуляторной энцефалопатии с использованием методов распознавания: Кузнецова А.В., Костомарова И.В., Водолагина Н.Н., Малыгина Н.А., Сенько О.В.
- [3] Сологуб Т.В., Осиновец О.Ю. Иммуномодуляторы в комплексной терапии ОРВИ: возможности применения препарата галавит. Русский медицинский журнал. 2013;3.
- [4] Инфекционные болезни: национальное руководство. Под общ. ред. Ющука Н.Д., Венгерова Ю.Я. М.: Гэотар-Мед.; 2009.
- [5] Empirical testing of institutional matrices theory by data mining I. L. Kirilyuk, A. I. Volynsky, M. S. Kruglova, A. V. Kuznetsova, A. A. Rubinstein, O. V. Senko.
- [6] Борисов, Л. Б. Медицинская микробиология, вирусология, иммунология / Л. Б. Борисов. М. : Медицинское информационное агентство
- [7] Р 50.1.033-2001. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть I. Критерии типа хи-квадрат. – М.: Изд-во стандартов. 2002. – 87 с.