XXIII ПРЕМИЯ
ГОРОДА МОСКВЫ
В ОБЛАСТИ
МЕДИЦИНЫ

2026

Автоматизированный алгоритм приоритизации вторичных находок для клинической интерпретации данных массового параллельного секвенирования

ФГБНУ «Медико-генетический научный центр имени академика Н.П. Бочкова»

Участники рабочей группы:
Рыжкова О.П. (руководитель), Басова А.М., Визеров Т.В., Боровиков А.О., Забудская К.Г., Бескоровайный Н.С., Малышева Е.В., Куцев С.И.

18+

Вторичные находки: что и зачем искать

Вторичные находки – часть случайных находок (клинически значимых вариантов, не связанных с причиной обследования), для которых существуют эффективные методы профилактики и лечения

84 гена
входят в список ACMG SF v3.3: находки в них рекомендуется сообщать пациенту

4,8 %
неродственных пациентов МГНЦ – носители вторичных находок (226 из 4 722)

≈ 4 млн
отличий от эталонной последовательности в геноме каждого человека

Наследственные онкологические заболевания
наследственный рак молочной железы и яичников (гены BRCA1 и BRCA2), синдром Линча; пожизненный риск рака у носителей – до 60–90 %

Сердечно-сосудистые заболевания
кардиомиопатии, аритмии, аневризмы аорты, семейная гиперхолестеринемия; первым проявлением может стать внезапная сердечная смерть

Врождённые нарушения обмена веществ и прочее
болезни Фабри, Помпе и Вильсона, наследственный гемохроматоз, злокачественная гипертермия

Эти болезни обычно проявляются во взрослом возрасте: если знать о риске заранее, их можно предупредить или вовремя начать лечение

Проблема: ручной анализ не масштабируется

Чтобы найти вторичные находки, врач должен оценить каждый вариант в генах списка, а их у человека сотни

≈ 4 млн
отличий от эталонного генома у каждого человека

≈ 500
вариантов в 84 генах списка на одного пациента

28 критериев
ACMG/AMP нужно применить к каждому варианту

Время врача
поиск вторичных находок отнимает время у основной задачи – поиска причины болезни пациента, а число геномных исследований быстро растёт

Разные заключения
при сравнении 9 лабораторий исходная согласованность классификации одних и тех же вариантов составила лишь 34 % (Amendola et al., 2016)

Нет готового решения
универсальные программы (InterVar, VarSome, Franklin, GeneBe) не учитывают правил ACMG SF: особых правил для генов, носительства, совпадения заболевания

Нужен автоматический, прозрачный и проверенный на клинических данных инструмент

Как работает алгоритм ACMG_SF_Classifier

Алгоритм получает варианты пациента после секвенирования, собирает о каждом сведения из баз данных, оценивает его по критериям ACMG/AMP и принимает одно из трёх решений

1. Данные

- варианты пациента из конвейера обработки NGS-data-Omics
- контроль качества прочтения

2. Базы данных

- частота у здоровых людей (gnomAD)
- описания в базах ClinVar и HGMD
- прогноз влияния на белок и сплайсинг (REVEL, AlphaMissense, SpliceAI, SQUIRLS)

3. Классификация

- 28 критериев ACMG/AMP
- баллы за каждое свидетельство
- особые правила для 84 генов
- предсказание распада мРНК

4. Решение

Автоматическое заключение
подтверждает врач

Проверка экспертом
список с указанием причины

Фильтрация
частые, доброкачественные, носительство

Балльная система (Tavtigian et al., 2018; ACGS, 2024)
Каждый критерий даёт от 1 до 8 баллов в зависимости от силы: свидетельства патогенности прибавляются, доброкачественности – вычитаются. 10 баллов и более – патогенный вариант, 6–9 – вероятно патогенный, 0–5 – неопределённого значения, от −1 до −5 – вероятно доброкачественный, −6 и менее – доброкачественный. Для патогенного класса нужно не менее двух независимых критериев, а каждое решение сопровождается объяснением.

Новизна и оригинальность подхода

Алгоритм не только назначает критерии ACMG/AMP, но и учитывает правила ACMG SF и качество исходных данных

Оценка баз данных
сведения ClinVar и HGMD учитываются не по принципу «да или нет», а по уровню проверки записи, доле подтверждающих публикаций и совпадению заболевания

Распад мРНК (NMD)
алгоритм определяет, разрушится ли мРНК с преждевременным стоп-кодоном; от этого зависит, получит ли вариант очень сильный критерий потери функции PVS1

Правила ACMG SF
особые правила для всех 84 генов; одиночный вариант в гене рецессивной болезни – носительство, два варианта в одном гене – проверка экспертом

Частота без недооценки
берётся максимальная частота варианта по всем версиям и популяциям gnomAD; порог редкости зависит от типа наследования

Точечные варианты и CNV
делеция и патогенный вариант в одном гене рецессивной болезни направляются эксперту как возможное сочетание двух повреждений

Прозрачность
для каждого варианта записано, какие критерии назначены и почему принято решение; пороги вынесены в настройки, исходный код открыт

Модуль анализа CNV

CNV (вариации числа копий) – выпадения (делеции) и удвоения (дупликации) участков ДНК. Модуль находит их в генах списка ACMG SF и оценивает вместе с точечными вариантами того же пациента

1. Объединение
результаты двух программ поиска CNV (Canvas и Manta) объединяются в одно событие

2. Отсев артефактов
вызовы низкого качества, события, частые в запуске (> 1 %) и в базе геномов центра (> 0,05 %)

3. Зиготность
число копий, удалённые экзоны, затронуты одна или обе копии гена; учитывается Х-хромосома у мужчин

4. Значимость
баллы ACMG/ClinGen (AutoCNV и собственная оценка), сведения ClinVar и HGMD, решение с объяснением

Меньше рутины

артефакты и частые события отсеиваются автоматически: из 136 613 CNV клинической когорты эксперту направлено 180

Совместный анализ

делеция и патогенный точечный вариант в одном гене рецессивной болезни направляются эксперту как возможное сочетание двух повреждений

Без потерь

CNV, описанные как патогенные в ClinVar и HGMD, не отсеиваются по частоте в базе центра; события, которые не может оценить AutoCNV, оценивает сам алгоритм

Результаты на клинической когорте

Проверка на данных 7 827 пациентов МГНЦ: эталон – 425 вторичных находок, подтверждённых врачами

425 из 425
подтверждённых находок найдено (100 %)

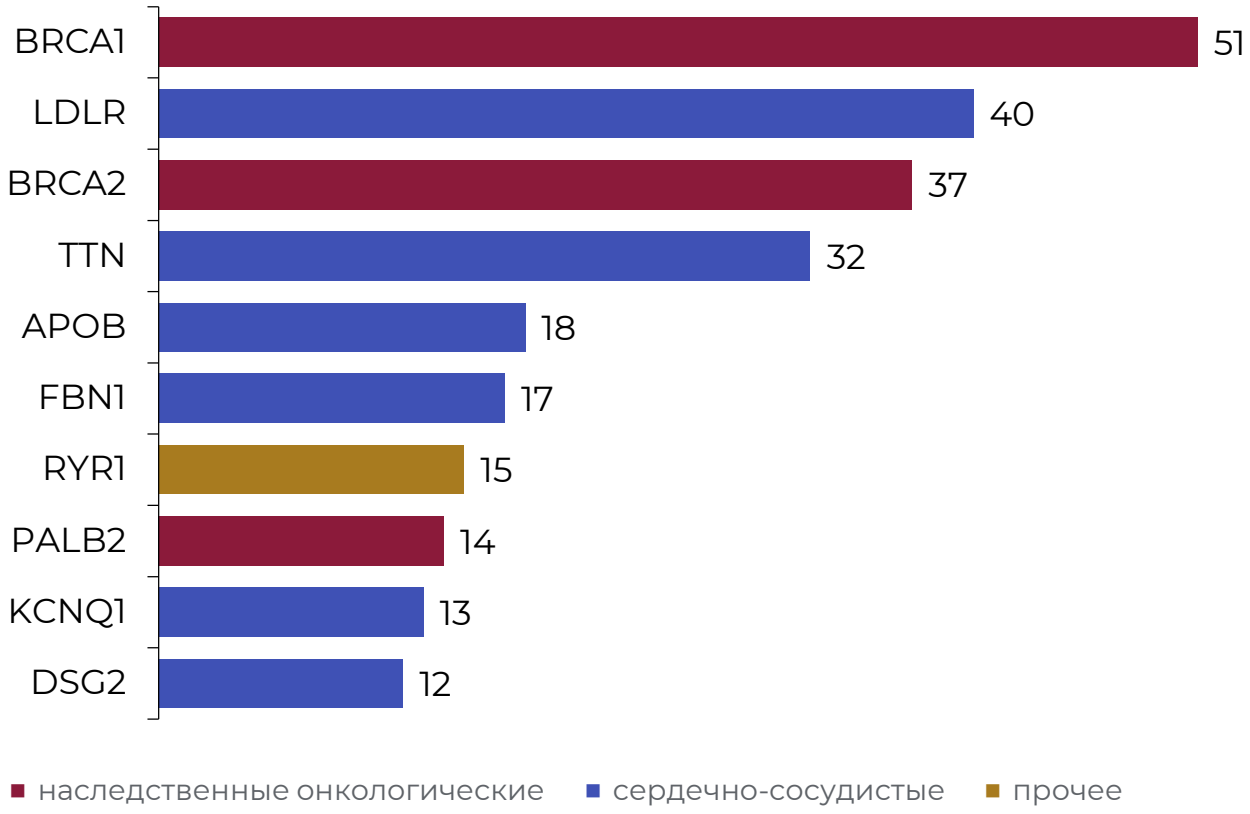
99,97 %
чувствительность и специфичность

0,5
варианта на одного пациента проверяет врач

≈ 1 с
на один геном; вся когорта – за 2,5 часа

57 %
пациентов без вторичных находок

Гены с наибольшим числом вторичных находок (из 425)



Практическая значимость и внедрение

Внедрён в практику ФГБНУ «МГНЦ»: вторичные находки анализируются у каждого пациента, прошедшего секвенирование генома

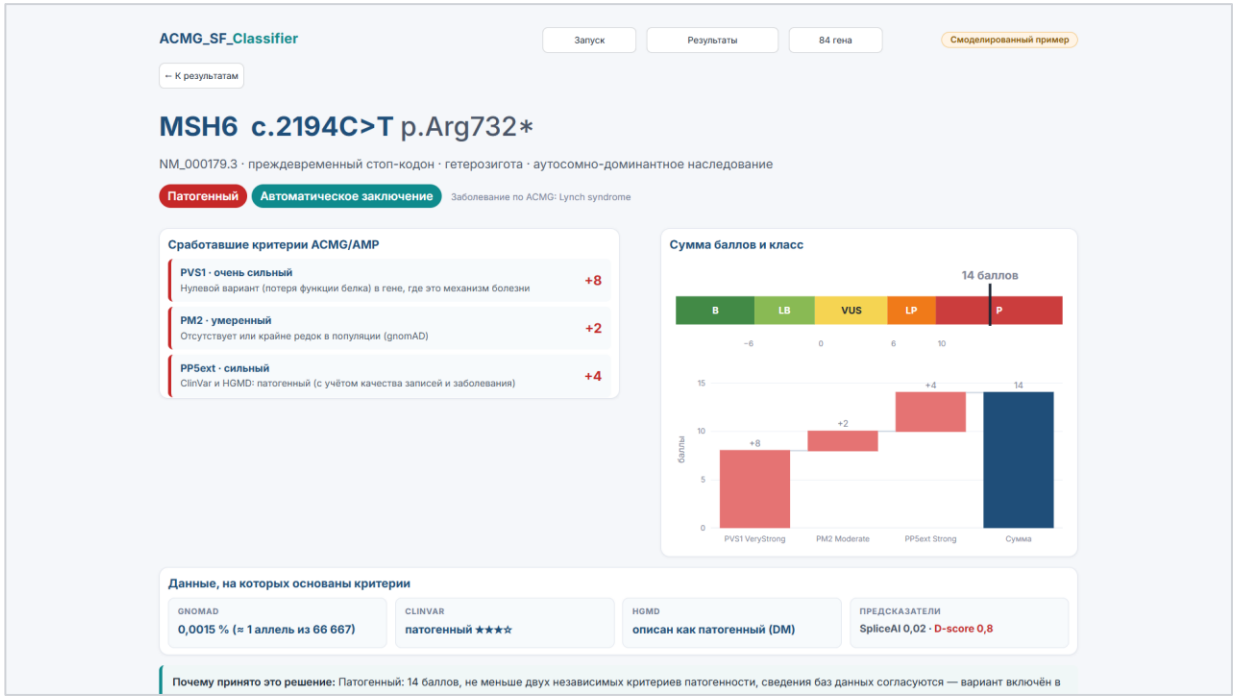
Собственный веб-портал: врачи сами запускают анализ и смотрят результаты с объяснением каждого решения; в 2027 году – модуль портала NGS-data

Время врача – основной диагностике: из 4 011 593 вариантов на проверку направлено 3 749 вариантов и 180 CNV, примерно в тысячу раз меньше

Секунды на геном: вся когорта из 7 827 пациентов – за 2,5 часа вместо более 14 суток ручного разбора

Единые правила классификации делают заключения воспроизводимыми и не зависящими от специалиста

Повторный анализ ранее исследованных геномов при обновлении баз ClinVar, HGMD и gnomAD



Карточка варианта в интерфейсе программы: назначенные критерии ACMG/AMP, баллы и причина решения. Смоделированный пример: показаны только варианты, описанные в ClinVar, номера образцов скрыты

Развитие алгоритма

Портал NGS-data – 2027

• сейчас врачи МГНЦ работают с алгоритмом через собственный веб-портал

• в 2027 году станет стандартным модулем общего портала анализа данных NGS-data

• анализ точечных вариантов и CNV в едином рабочем процессе врача

• повторный анализ ранее исследованных геномов при обновлении ClinVar, HGMD и gnomAD

Генные панели

та же логика для наборов генов, которые используют в диагностике отдельных групп заболеваний

Диагностика наследственных заболеваний

настройка списка генов и правил под конкретное заболевание

Регистрация и апробация

Свидетельство о государственной регистрации программы для ЭВМ

№ 2026613737

«ACMG_SF_Classifier», дата регистрации 09.02.2026

Правообладатель – ФГБНУ «МГНЦ»

Авторы: Басова А.М., Рыжкова О.П., Визеров Т.В., Боровиков А.О., Малышева Е.В., Бескоровайный Н.С., Забудская К.Г., Куцев С.И.

NGS 2026

XIV Всероссийская конференция центров геномных исследований мирового уровня, 16.04.2026, Москва; тезисы: сборник «Геномное секвенирование и редактирование», с. 10

«Научный телеграф»

мультидисциплинарная молодёжная конференция, 17.05.2026, Москва

ESHG-2026

конференция Европейского общества генетики человека, Гётеборг, 13–16.06.2026; электронный постер EP15.049

BGRS/SB 2026

международная конференция по биоинформатике, 6–11.07.2026, Новосибирск; тезисы: doi 10.18699/bgrs2026-717

Yandex Cloud

победа в «Конкурсе молодых учёных»

Правительство Москвы

МОСКОВСКАЯ МЕДИЦИНА

СОВЕТ ГЛАВНЫХ ВРАЧЕЙ ДЗМ

НИИ ОРГАНИЗАЦИИ ЗДРАВООХРАНЕНИЯ И МЕДИЦИНСКОГО МЕНЕДЖМЕНТА



СПАСИБО ЗА ВНИМАНИЕ!

ИНФОРМАЦИОННЫЕ ПАРТНЕРЫ:

МОСКОВСКАЯ МЕДИЦИНА cito

ЕЖЕКВАРТАЛЬНЫЙ НАУЧНЫЙ РЕЦЕНЗИРУЕМЫЙ ЖУРНАЛ ЗДОРОВЬЕ МЕГАПОЛИСА

ВРАЧАМ МОСКОВСКАЯ МЕДИЦИНА

18+