https://doi.org/10.47183/mes.2025-366

HIDDEN BURDEN OF CHRONIC NON-COMMUNICABLE DISEASES. PART 2: PREVALENCE OF GERMLINE VARIANTS ASSOCIATED WITH CANCERS AMONG RESIDENTS OF THE RUSSIAN FEDERATION

Vladimir S. Yudin[1], Aleksandra A. Mamchur[1], Ekaterina A. Nikolaeva[1], Mikhail V. Ivanov[1], Daria A. Kashtanova[1]✉, Elena A. Zelenova[1], Sergey I. Mitrofanov[1], Ekaterina A. Snigir[1], Lyudmila P. Kuzmina[2], Anton A. Keskinov[1]

[1] Centre for Strategic Planning and Management of Biomedical Health Risks of the Federal Medical and Biological Agency, Moscow, Russia
[2] Izmerov Research Institute of Occupational Health, Moscow, Russia

Introduction. Cancer is a leading cause of mortality worldwide. Hereditary (germline) mutations in cancer predisposition genes significantly contribute to oncogenesis (accounting for 3–12.6% of adult cases). Identifying carriers of pathogenic/likely pathogenic (P/LP) variants is crucial for cancer prevention and its early detection. Large-scale data on the prevalence of such variants in the general population of the Russian Federation, essential for healthcare resource planning, have been extremely limited until now.
Objective. To assess the prevalence and spectrum of pathogenic and likely pathogenic germline variants in genes associated with cancer risk in a Russian population cohort.
Materials and methods. An analysis of whole-genome sequencing data from 116,794 participants (a representative sample of the adult population from 86 federal subjects of the Russian Federation) was conducted. The search for P/LP variants in cancer-associated genes was performed based on the ClinVar database. Variant annotation was carried out considering the level of evidence. Statistical analysis was performed using Python (v3.9.12).
Results. P/LP variants were identified in 26 cancer-associated genes among 2643 participants (2.26%). The most frequent variant was rs36053993 (*MUTYH*, 0.28%), with its frequency in the Russian Federation being lower than in global databases (ALFA, gnomAD). Level A variants (only those with the "reviewed by expert panel" level of evidence) were carried by 0.8% of individuals collectively. Level B variants (those with the "reviewed by expert panel" and "criteria provided, multiple submitters, no conflicts" levels of evidence) were carried by 2.08%. The inclusion of less reliable evidence levels (C, D) increased this proportion only marginally, to 2.26%. A statistically significant decrease in the proportion of carriers of Level A variants with age was observed ($p = 0.007$), while the overall proportion of P/LP variant carriers decreased only slightly ($p = 0.17$). Substantial geographical variability in prevalence was identified, ranging from 865 (Bashkortostan) to 6250 (Chechnia) variants per 100,000 population. In Moscow, the prevalence was 2340 per 100,000.
Conclusions. The study revealed a significant "hidden burden" of P/LP oncogenic germline variant carriers (2.26%) in the Russian population. The primary burden is attributed to variants with a high level of evidence (Level B). The substantial variability in prevalence across regions of the Russian Federation necessitates a differentiated approach to healthcare resource planning. The data obtained substantiate the need for implementing genetic screening programs (especially for individuals with a family history) and enhanced oncological surveillance for carriers. This represents a strategic direction for reducing cancer incidence and mortality in high genetic risk groups within the Russian Federation.

Keywords: chronic non-communicable diseases; oncological diseases; germline genetic variants; whole-genome sequencing; hereditary predisposition

For citation: Yudin V.S., Mamchur A.A., Nikolaeva E.A., Ivanov M.V., Kashtanova D.A., Zelenova E.A., Mitrofanov S.I., Snigir E.A., Kuzmina L.P., Keskinov A.A. Hidden burden of chronic non-communicable diseases. Part 2: Prevalence of germline variants associated with cancers among residents of the Russian Federation. *Extreme Medicine*. 2025. https://doi.org/10.47183/mes.2025-366

Funding: the study was carried out without sponsorship.

Compliance with the ethical principles: the study was approved by the Local Ethical Committee of the Centre for Strategic Planning of the FMBA of Russia (Minutes No. 5 of 28.12.2020; Minutes No. 2 of 01.06.2021). For the creation of the Genetic Database (GDB) of the Centre for Strategic Planning of the FMBA of Russia, voluntary informed consent for participation in the study was obtained from all participants, and the study was conducted in accordance with the principles of the Declaration of Helsinki.

Potential conflict of interest: Anton A. Keskinov is the scientific editor of the Extreme Medicine journal. The remaining authors declare no conflict of interest.

✉ Daria A. Kashtanova DKashtanova@cspfmba.ru

Received: 4 Aug. 2025 **Revised:** 8 Oct. 2025 **Accepted:** 22 Oct. 2025 **Online first:** 28 Nov. 2025

УДК 631.523:616-006.04

СКРЫТОЕ БРЕМЯ ХРОНИЧЕСКИХ НЕИНФЕКЦИОННЫХ ЗАБОЛЕВАНИЙ. ЧАСТЬ 2: РАСПРОСТРАНЕННОСТЬ ГЕРМИНАЛЬНЫХ ВАРИАНТОВ, АССОЦИИРОВАННЫХ С ОНКОЛОГИЧЕСКИМИ ЗАБОЛЕВАНИЯМИ, СРЕДИ ЖИТЕЛЕЙ РОССИЙСКОЙ ФЕДЕРАЦИИ

В.С. Юдин[1], А.А. Мамчур[1], Е.А. Николаева[1], М.В. Иванов[1], Д.А. Каштанова[1]✉, Е.А. Зеленова[1], С.И. Митрофанов[1], Е.А. Снигирь[1], Л.П. Кузьмина[2], А.А. Кескинов[1]

[1] Центр стратегического планирования и управления медико-биологическими рисками здоровью Федерального медико-биологического агентства, Москва, Россия
[2] Научно-исследовательский институт медицины труда им. академика Н.Ф. Измерова, Москва, Россия

Введение. Онкологические заболевания — ведущая причина смертности в мире. Наследственные (герминальные) мутации в генах предрасположенности к раку вносят существенный вклад в онкогенез (3–12,6% случаев у взрослых). Выявление носителей патогенных / вероятно патогенных (P/LP) вариантов критически важно для профилактики и ранней диагностики рака. Крупномасштабные данные о распространенности таких вариантов в общей популяции Российской Федерации для планирования ресурсов здравоохранения ранее были крайне ограниченны.
Цель. Оценить распространенность и спектр патогенных и вероятно патогенных герминальных вариантов в генах, ассоциированных с риском развития онкологических заболеваний, в российской популяционной когорте.
Материалы и методы. Проведен анализ данных полногеномного секвенирования 116 794 участников (репрезентативная выборка взрослого населения из 86 субъектов РФ). Поиск P/LP-вариантов в онкоассоциированных генах выполнен на основе базы данных ClinVar. Аннотация вариантов проводилась с учетом уровня доказательности. Статистический анализ выполнен с использованием Python (v3.9.12).
Результаты. P/LP-варианты обнаружены в 26 онкоассоциированных генах у 2643 (2,26%) участников. Наиболее частый вариант — rs36053993 (*MUTYH*, 0,28%), частота в РФ ниже мировых баз (ALFA, gnomAD). Варианты уровня A (только варианты с уровнем доказательности «проверено экспертной комиссией») имели суммарно 0,8% людей, уровня B (варианты с уровнями доказательности «проверено экспертной комиссией» и «критерии предоставлены несколькими заявителями, конфликтов нет») — 2,08%, добавление менее надежных уровней (C, D) увеличивало показатель лишь до 2,26%. Наблюдали статистически значимое снижение доли носителей вариантов уровня A с возрастом (p = 0,007), общая доля носителей P/LP-вариантов снижалась незначительно (p = 0,17). Выявлена значительная географическая вариабельность распространенности: от 865 (Республика Башкортостан) до 6250 (Чеченская Республика) вариантов на 100 тыс. населения. В Москве — 2340 на 100 тыс.
Выводы. Исследование выявило значимое «скрытое бремя» носительства P/LP онкогерминальных вариантов (2,26%) в российской популяции. Основное бремя обусловлено вариантами с высоким уровнем доказательности (уровень B). Значительная вариабельность распространенности по регионам РФ требует дифференцированного подхода к планированию ресурсов здравоохранения. Полученные данные обосновывают необходимость внедрения программ генетического скрининга (особенно для лиц с семейным анамнезом) и усиленного онконаблюдения для носителей, что является стратегическим направлением для снижения онкозаболеваемости и смертности в группах высокого генетического риска в РФ.

Ключевые слова: хронические неинфекционные заболевания; онкологические заболевания; герминальные генетические варианты; полногеномное секвенирование; наследственная предрасположенность

Для цитирования: Юдин В.С., Мамчур А.А., Николаева Е.А., Иванов М.В., Каштанова Д.А., Зеленова Е.А., Митрофанов С.И., Снигирь Е.А., Кузьмина Л.П., Кескинов А.А. Скрытое бремя хронических неинфекционных заболеваний. Часть 2: распространенность герминальных вариантов, ассоциированных с онкологическими заболеваниями, среди жителей Российской Федерации. *Медицина экстремальных ситуаций*. 2025. https://doi.org/10.47183/mes.2025-366

Финансирование: исследование выполнено без спонсорской поддержки.

Соответствие принципам этики: исследование одобрено локальным этическим комитетом ФГБУ «ЦСП» ФМБА России (протокол № 5 от 28.12.2020, протокол № 2 от 01.06.2021). Для создания базы данных популяционных частот (GDB) ФГБУ «ЦСП» ФМБА России от всех участников получено добровольное информированное согласие на участие в исследовании, проведенном в соответствии с принципами Хельсинкской декларации.

Потенциальный конфликт интересов: А.А. Кескинов — научный редактор журнала «Медицина экстремальных ситуаций», остальные авторы заявляют об отсутствии конфликта интересов.

✉ Каштанова Дарья Андреевна DKashtanova@cspfmba.ru

Статья поступила: 06.08.2025 **После доработки:** 06.10.2025 **Принята к публикации:** 27.10.2025 **Online first:** 28.11.2025

INTRODUCTION

Cancer remains one of the most significant challenges in global public health and a leading cause of mortality worldwide [1, 2]. Combating oncological diseases requires a comprehensive approach, including understanding and accounting for their risk factors. Despite advancements in diagnostic methods and increased coverage of screening programs, malignant neoplasms are currently definitively identified once they have become clinically apparent. Therefore, developing technologies for the preventive detection and prediction of their development is critically important.

Hereditary (germline) mutations significantly contribute to oncogenesis, accounting for approximately 3–12.6% of adult cancer cases and 8.5–10% of childhood cancer cases [3, 4]. The influence of heredity on cancer risk varies from low-penetrance susceptibility (caused by common germline variants that increase relative risk by 1.5–2 times) through moderate penetrance (risk increases by 2–5 times) to high-penetrance predisposition associated with rare germline variants (risk increase more than fivefold) [5, 6]. Predispositions with moderate and high penetrance are typically inherited in an autosomal dominant pattern and are associated with mutations in Cancer Predisposition Genes (CPGs), of which over 100 have been identified [7]. Screening for such variants is particularly important in families with a significant family history of cancer [8–10].

Clinical whole-exome (WES) and whole-genome (WGS) sequencing have become key tools for identifying hereditary cancer syndromes. WES effectively and economically detects pathogenic variants in the coding regions of known CPGs and Cancer Driver Genes (CDGs) [11], rendering it the primary method for clinical practice and screening.

Whole-genome sequencing offers a unique advantage through the ability to detect pathogenic variants

across the entire genome, including non-coding regulatory regions and structural variants potentially involved in hereditary predisposition [12]. However, its application is limited by higher costs and the complexity of interpreting findings in non-coding regions.

In 2013, the American College of Medical Genetics and Genomics (ACMG) issued recommendations directing laboratories to actively seek and report to patients the identification of pathogenic variants in a specific list of 56 genes associated with pathological but potentially manageable conditions [13]. This list, which includes several CPGs (e.g., *BRCA1*, *BRCA2*, *TP53*), was revised in 2017 (ACMG v2.0, 59 genes) and further in 2021 [14]. Although the original recommendations relied on a binary variant classification ("known pathogenic" / "expected pathogenic"), a transition to the modern, five-category ACMG pathogenicity system was strongly recommended.

This approach inevitably uncovers genetic variants unrelated to the primary indication for testing but possessing independent clinical significance. Such findings, known as secondary findings (SF), may indicate a risk for developing severe diseases (including some forms of cancer) for which effective preventive measures or early-stage treatments exist [15]. Identifying these germline mutations is critical for cancer prevention and early diagnosis in at-risk groups.

Determining the true frequency of pathogenic/likely pathogenic (P/LP) variant carriers in CPGs within the general population, not just among cancer patients, is of particular interest to healthcare systems. Knowledge of such a "genetic burden" allows for the proactive planning of medical interventions, such as estimating the need for genetic counseling, genetic testing, enhanced screening programs, and preventive measures (e.g., mammography + breasts MRI for *BRCA1/2* mutation carriers, colonoscopy for Lynch syndrome). Such planning enables the assessment of the burden on the healthcare system: forecasting the demand for expensive genetic tests, specialized diagnostics, and preventive treatments on a regional or national scale, ultimately optimizing cancer prevention strategies for individuals at high risk.

Conducting large-scale population studies to assess the frequency of germline mutations presents significant challenges. These include high costs, especially when utilizing WGS, and the necessity of examining large cohorts to obtain statistically significant data, particularly for rare genetic variants and in ethnically diverse populations. A particular ongoing difficulty lies in the interpretation of the results from such studies, due to the need for accurate variant classification according to stringent criteria (e.g., ACMG/AMP (Association for Molecular Pathology)) [16]. This demands substantial expert and computational resources, especially for automation (e.g., using InterVar) [17] and for the analysis of non-coding variants and structural variants. Furthermore, challenges entail ethical and organizational aspects: obtaining informed consent, ensuring data confidentiality, and organizing the collection of biomaterials.

Despite the challenges, research in this sphere is underway. Projects such as The Cancer Genome Atlas [17] and The Pancancer Analysis of Whole Genomes [4] have provided data on the frequency of germline variants among cancer patients (8% and 17%, respectively). Studies in the general population, primarily involving cohorts of European-American and African-American ancestry, estimate the prevalence of P/LP variants in genes from the American College of Medical Genetics and Genomics (ACMG) secondary findings list to range 0.8–7%. However, data derived from tumor tissues do not reflect the frequency in the healthy population.

As far as the Russian Federation is concerned, large-scale studies on the prevalence of germline P/LP variants in CPGs/CDGs in the population are extremely limited. Our previous study, which included data from nearly 77,000 individuals, was a significant step toward describing the genetic burden of hereditary cancer in the Russian population [19] but required expansion.

In the present study, we aim to assess the prevalence and spectrum of pathogenic and likely pathogenic germline variants in genes associated with the risk of developing oncological diseases within an expanded Russian population cohort. The results obtained are of significant importance for understanding the hidden genetic burden of hereditary cancer in the Russian Federation. This information is necessary for planning and optimizing resources within regional healthcare systems required for the prevention, early diagnosis, and management of patients from high genetic risk groups.

MATERIALS AND METHODS

Study participants and ethical considerations

This study is a continuation of a large-scale epidemiological project initiated in 2019 by the Centre for Strategic Planning of the FMBA of Russia to investigate the prevalence of genetic variants associated with the risk of chronic diseases, including oncological conditions, among the adult population across 86 federal subjects of the Russian Federation.

The study was conducted using a population-based sample from the Genetic Database (GDB) of the Centre for Strategic Planning of the FMBA of Russia.[1] The randomized representative sample comprised 116,794 individuals from 86 federal subjects of the Russian Federation. The demographic characteristics of the participants (age and sex distribution) are presented in Figure 1. The median age was 51 [22;73] years for men and 51 [23;75] years for women.

[1] Genetic Database (GDB). FMBA of Russia. Application v1.1.3 of 17.03.2025. Database v59.1 of 03.10.2024. URL: https://gdbpop.nir.cspfmba.ru (access date 02.08.2025) (In Russ.)

Whole-genome sequencing

DNA was extracted from whole blood samples using the MagAttract HMW DNA Kit (Qiagen, Germany) on an automated robotic station. Whole-genome sequencing libraries were prepared using the Illumina DNA Prep kit (Illumina, USA) according to the manufacturer's instructions. Whole-genome sequencing was performed on the Illumina NovaSeq 6000 platform using the S4 Reagent Kit (300 cycles) (Illumina, USA) for paired-end reads of 2 × 150 bp.

During the demultiplexing stage, sequencing data in BCL format were converted to FASTQ format using Illumina bcl2fastq v2.20. The sequencing run quality was controlled using Illumina Sequencing Analysis Viewer v2.4.7, while the read quality (FASTQ.GZ format) was assessed using the FastQC v0.11.9 bioinformatic tool. Alignment to the reference genome (GRCh38) was performed on the Dragen Bio-IT platform (Illumina, USA). Quality control of alignments (BAM format) was conducted using DRAGEN, FastQC v0.11.9, samtools v1.13, and mosdepth v0.3.1. All samples passed quality control based on key metrics, including the percentage of duplicated reads and the number of unaligned reads. The final mean genome coverage per sample was at least 30x. Detection of single nucleotide polymorphisms (SNPs) and short insertions/deletions (up to 50 bp) was performed using Strelka2 [9]; only variants passing the PASS filter were selected.

Search for cardiac germline variants

A search for pathogenic and likely pathogenic genetic variants associated with the development of oncological diseases was conducted within the study sample, based on the ClinVar database. The following levels of evidence for the relationship between a genetic variant and the development of the corresponding phenotype, as provided by ClinVar (in descending order of strength), were used for annotating the obtained list of variants:

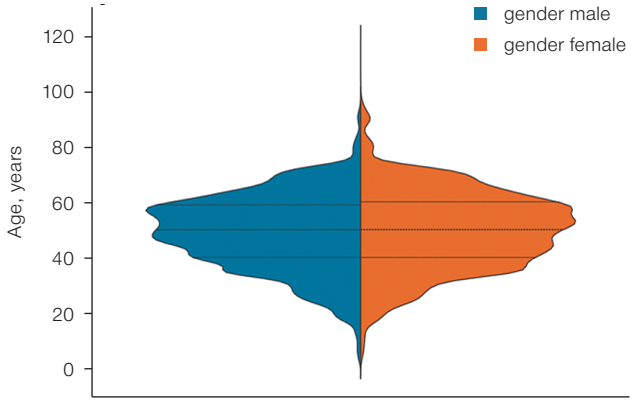

Figure prepared by the authors based on original data

Fig. 1. Age and sex distribution of the study participants

- "Practice guideline"
  Highest level. Interpretation is based on official clinical guidelines from authoritative professional societies. This information represents the most reliable data for clinical decision making.
- "Expert panel"
  Very high level. The variant has been thoroughly evaluated and approved by an independent expert group (e.g., within the ClinGen consortium). This also constitutes an extremely reliable basis for clinical action.
- "Multiple submitters"
  High reliability. Multiple laboratories have independently reached the same interpretation of the variant, and no conflicting data exist. This is a reliable indicator for use in practice.
- "Single submitter"
  Moderate reliability. The interpretation originates from only one laboratory or research group. Such data require caution and verification through scientific publications or other sources.
- "No assertion provided"
  Low reliability. The submitter provided information about the variant but did not offer any clinical interpretation. This information should not be used for clinical decision making.

Statistical Data Processing

Statistical data processing and result visualization were performed using the Python programming language (v3.9.12). Intergroup comparison of the number of carriers of genetic variants was conducted using the chi-square test.

RESULTS

Annotation results of variants associated with oncological diseases

In the analyzed sample, carriers of pathogenic and likely pathogenic variants were identified across 26 genes associated with various oncological diseases: *APC*, *BMPR1A*, *BRCA1*, *BRCA2*, *MEN1*, *MLH1*, *MSH2*, *MSH6*, *MUTYH*, *NF2*, *PALB2*, *PMS2*, *PTEN*, *RB1*, *RET*, *SDHAF2*, *SDHB*, SDHC, *SDHD*, *SMAD4*, *TMEM127*, *TP53*, *TSC1*, *TSC2*, *VHL*, *WT1*. At least one variant in these genes was detected in 2643 participants (2.26%). The number of unique variants identified was 467, with a cumulative frequency in the entire cohort of 1.15%, corresponding to 2307 variants per 100,000 population.

Variants in the *MUTYH* gene were the most frequently observed in the study sample. Specifically, the most common variant was rs36053993 (NC_000001.11:g.45331556C>T), with a frequency of 0.28%. According to the GDB, the frequency of this variant in the Russian population is statistically significantly lower than in other global projects. For

instance, its frequency in the ALFA project was 0.46% (p-value < 0.001), compared to 0.33% in gnomAD (p-value < 0.01).

The distribution of the identified variants by the level of evidence is presented in Figure 2A. Notably, none of the found variants had the "practice guideline" level of evidence. The largest number of variants (215 out of 467) had the "reviewed by expert panel" evidence level, followed by variants with the "criteria provided, multiple submitters, no conflicts" level (168 out of 467). The aforementioned most common variant, rs36053993, also belongs to this second group.

In total, Level A variants (only those with the "reviewed by expert panel" evidence level) were carried by 0.8% of individuals. Level B variants (those with "expert panel" and "multiple submitters, no conflicts" levels of evidence) were carried by 2.08%. Level C variants (those with "expert panel," "multiple submitters, no conflicts," and "single submitter" levels of evidence) were characteristic of 2.25% of individuals. Level D included all variants and was characteristic of 2.26% of individuals (the entire group of carriers of pathogenic and likely pathogenic variants).

As noted above, the majority of identified variants belonged to the former two groups. Furthermore, these same groups accounted for the most prevalent variants. Consequently, Level B encompassed carriers of nearly all pathogenic and likely pathogenic variants. The inclusion of additional, less substantiated variants only marginally increased the overall carrier proportion.

It is important to note that the prevalence of pathogenic and likely pathogenic variants in a population typically exhibits an age-dependent dynamic. Carriage of these genetic variants is presumed to negatively impact survival and the likelihood of reaching old age. However, in our analysis (Fig. 2C), the proportions of P/LP carriers changed only marginally from one age group to another from a statistical standpoint (p-value = 0.17), although a trend suggesting a decrease in the carrier proportion with age was observed. This trend is more pronounced when considering only carriers of Level A variants (p-value = 0.007).

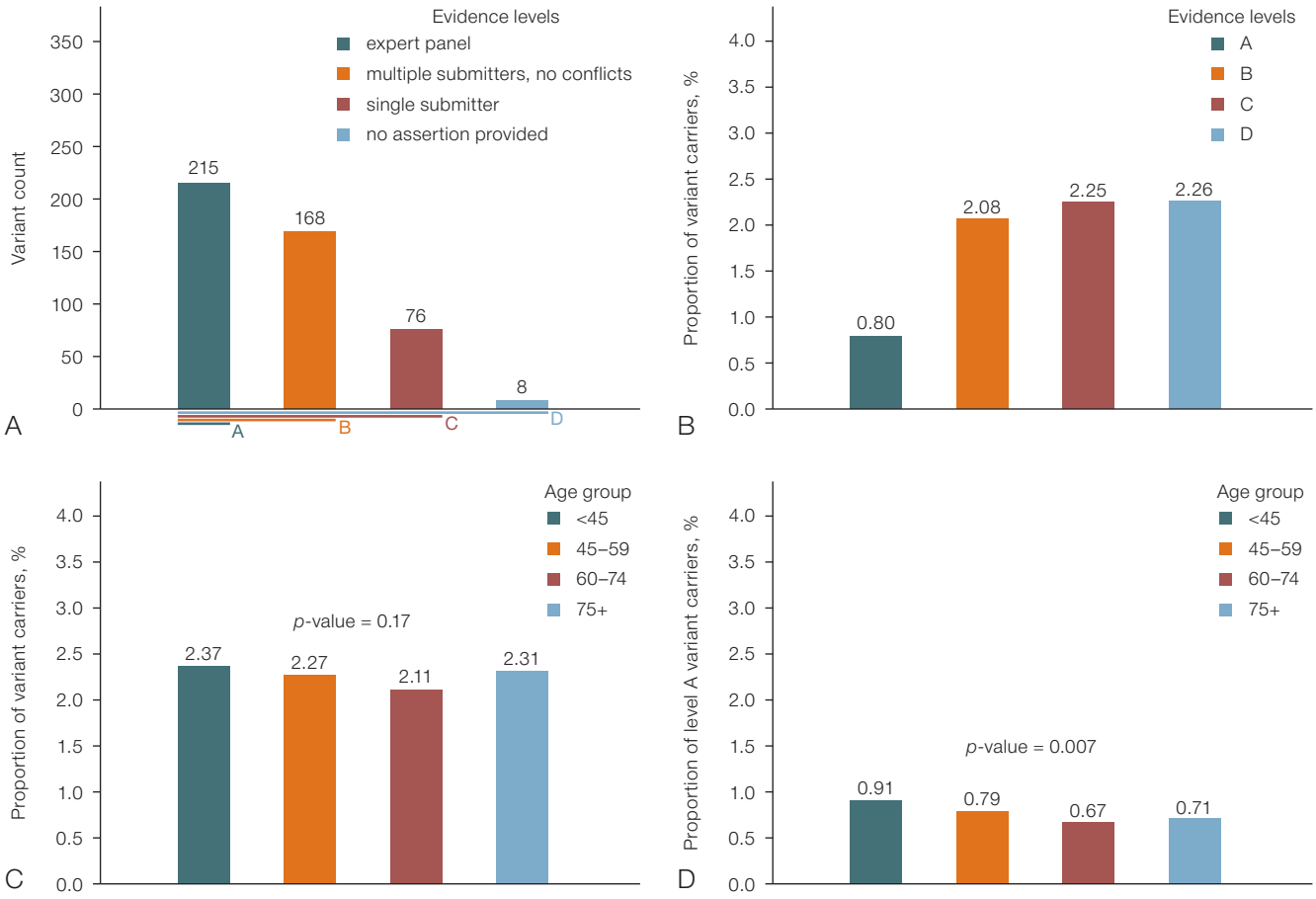

Figure prepared by the authors based on original data

Fig. 2. Annotation results of the identified genetic variants: A — frequency of variants with different evidence levels in the sample; B — proportion of carriers of variants with different evidence levels; C — proportion of carriers of all pathogenic and likely pathogenic variants associated with oncological diseases across different age groups; D — proportion of carriers of Level A variants across different age groups

Analysis of pathogenic and likely pathogenic variant distribution across Russian federal subjects

The conducted analysis also enabled the creation of a map illustrating the prevalence of pathogenic and likely pathogenic variants associated with oncological diseases across different regions of the Russian Federation (Fig. 3). The highest frequencies were observed in Chechnya (6250 variants per 100,000 population), the Republic of Komi (5484 per 100,000), Zabaykalsky Krai (4366 per 100,000), Kurgan Oblast (4219 per 100,000), and Lipetsk Oblast (3636 per 100,000). The lowest frequencies were found in the Republics of Sakha (Yakutia) (1179 per 100,000), Dagestan (1102 per 100,000), Chuvashia (1046 per 100,000), Mari El (967 per 100,000), and Bashkortostan (865 per 100,000). In Moscow, the prevalence of oncogenic germline variants was 2340 per 100,000 population, which aligns with the overall population value of 2307 mentioned earlier.

DISCUSSION

When discussing the genetics of oncological diseases, researchers and clinicians often refer to the accumulation of somatic mutations in specific cells, which stimulates the development of malignant neoplasms. However, hereditary predisposition to cancer, manifested as the carriage of pathogenic and likely pathogenic variants in cancer-associated genes, represents an equally critical risk factor. Moreover, while somatic mutations are spontaneous and often unpredictable, germline variants are detectable from birth. Identification of oncogenic germline variants is essential for estimating disease risks at both individual and population levels, including for a specific region or an entire country.

Our large-scale study, bridging epidemiology and genetics, has identified pathogenic or likely pathogenic variants of varying evidence levels in cancer-associated genes in 2.26% of participants from a representative sample of over 116,000 residents of the Russian Federation. This sample was not enriched with individuals with established oncological diagnoses, allowing us to reveal a “hidden” genetic burden within the ostensibly healthy population. Although all identified variants exhibit incomplete penetrance and disease manifestation depends on additional factors [21], data on the prevalence of oncogenic germline variants in the population can be valuable for developing preventive care protocols. While the identified variants are characterized by incomplete penetrance, their carriage significantly increases the lifetime risk of developing specific cancer syndromes [5–8]. The observed decrease in the proportion of P/LP variant carriers with age (particularly pronounced for Level A variants with the highest level of evidence, $p = 0.007$) aligns with the concept of increased mortality or morbidity among carriers who do not survive to older age groups. This indirectly confirms the clinical significance of the identified variants and the importance of pre-symptomatic detection.

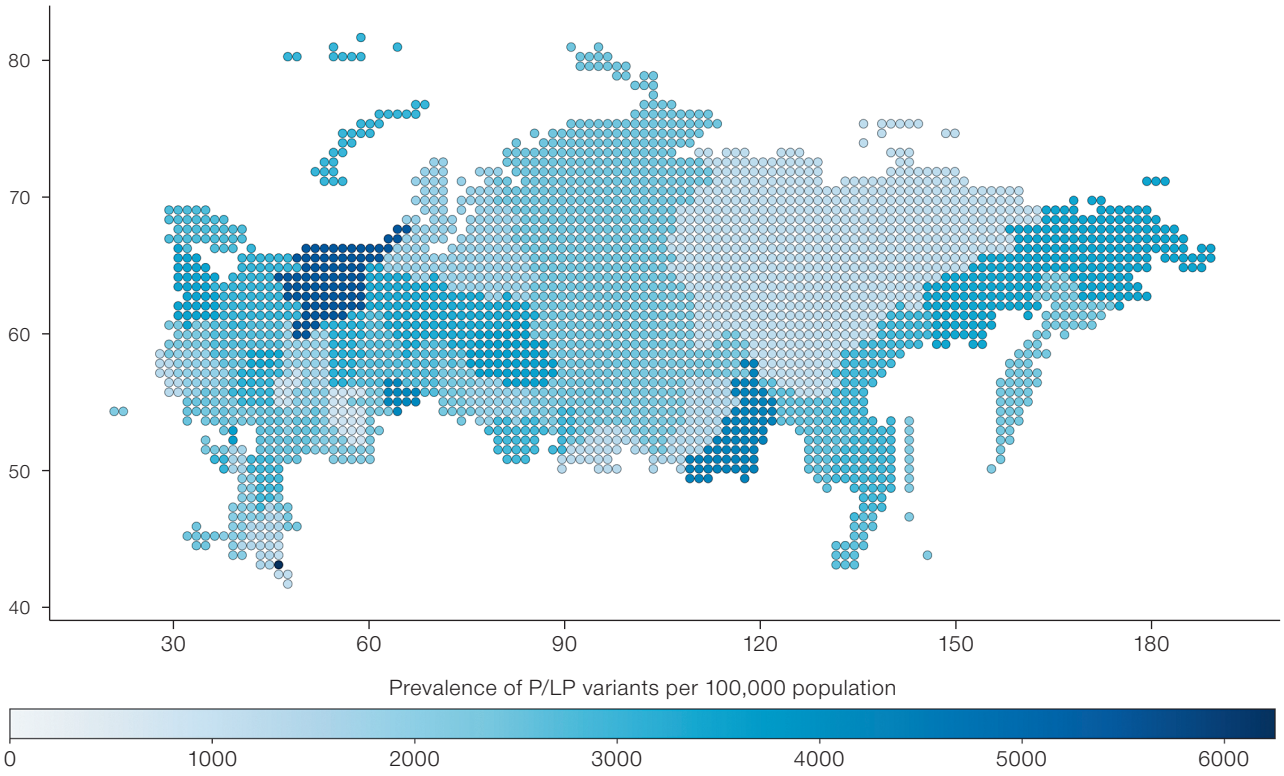

Figure prepared by the authors based on original data

Fig. 3. Prevalence map of oncogenic germline variants in the population of the Russian Federation

The use of ClinVar evidence levels enabled an assessment of the reliability of the identified associations. The fact that Level B (high and moderate evidence levels) accounts for 2.08% of carriers, while the addition of less reliable variants (Levels C and D) increases this figure only to 2.26%, holds significant practical importance. This indicates that the primary "genetic burden" is attributable to variants with a solid evidence base and also strengthens confidence in the results, justifying the prioritization of these variants when planning screening programs. The absence of variants at the "practice guideline" level underscores the need for further work on the validation and clinical interpretation of genetic findings in the Russian population.

The geographical distribution of the identified genetic variants obtained in this study is of particular value. The prevalence across different regions of Russia showed substantial variation, ranging from 865 variants per 100,000 population in Bashkortostan to 6250 variants per 100,000 in Chechnya. This disparity may be attributed to a multitude of factors. First, remote regions and isolated communities, which various ethnic groups may have constituted at different times, could be characterized by a "founder effect" or other evolutionary peculiarities.

Second, the demographic structure of a region can influence the frequency of P/LP variants. As our analysis demonstrated, the proportion of P/LP carriers decreases when moving from younger to older age groups. This is particularly significant for variants approved by an expert panel, i.e., those with the highest level of evidence. The observed trend is likely linked to the increased mortality of P/LP carriers before reaching old age, which reduces their numbers in the oldest age groups. Consequently, the prevalence of cancer-associated genetic variants will also vary depending on the demographic structure of a particular region and the predominance of younger versus older residents.

Third, the level of healthcare development in a given region is a significant factor. When cancer manifests in carriers of pathogenic/likely pathogenic variants in regions with advanced medical care, the chances of survival and achieving remission are significantly higher, thereby preserving the mutation in the population's gene pool. It is important to consider that for some regions with small sample sizes, the frequency estimate may be less precise.

Nevertheless, knowledge of regional characteristics is critically important for differentiated planning of healthcare system resources. The obtained data substantiate the necessity of ensuring access to confirmatory testing (especially WES/WGS), genetic screening programs (particularly for individuals with a family history), and specialized programs for enhanced cancer screening for carriers (e.g., "mammography + MRI" for *BRCA* carriers, colonoscopy for Lynch syndrome).

It is also known, for instance, that the presence of hereditary variants is significant for preventing adverse events associated with certain classical chemotherapy regimens [22], further underscoring the importance of such testing. Data on the prevalence of P/LP variants are crucial for assessing the cost-effectiveness of the widespread implementation of genetic testing and preventive programs at the population or regional level.

Thus, the early identification of carriers before the onset of disease opens the possibility for truly preventive interventions (enhanced screening, chemoprevention, prophylactic surgeries), which are capable of significantly reducing cancer incidence and mortality in this high-risk group.

CONCLUSION

The present study provides unique and large-scale data on the hidden burden of hereditary cancer predisposition in the Russian population. The identified significant carrier frequency of P/LP variants (2.26%), particularly its substantial regional variability, is of fundamental importance for understanding the genetic architecture of cancer risk in the country and serves as a powerful tool for planning and optimizing healthcare system resources.

The implementation of genetic screening programs and preventive management for carriers, based on the population data presented herein, represents a strategic direction for reducing the incidence and mortality from oncological diseases in high genetic risk groups.

Future research directions include conducting functional studies for pathogenic variants with an undetermined mechanism of action, as well as integrating genomic data with clinical and phenotypic manifestations to establish precise genotype–phenotype correlations and assess the contribution of modifying factors. Furthermore, expanding the spectrum of analyzed genes and incorporating the analysis of structural variants and copy number variations (CNVs) will provide a more comprehensive picture of hereditary predisposition.

References

1. GBD 2021 Forecasting Collaborators. Burden of disease scenarios for 204 countries and territories, 2022–2050: a forecasting analysis for the Global Burden of Disease Study 2021. *The Lancet*. 2024;403(10440):2204–56. https://doi.org/10.1016/S0140-6736(24)00685-8
2. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. 2024;74(3):229–63. https://doi.org/10.3322/caac.21834
3. Han SA, Kim SW. BRCA and Breast Cancer-Related High-

Penetrance Genes. *Translational Research in Breast Cancer.* 2021;1187:473–90. https://doi.org/10.1007/978-981-32-9620-6_25
4. Kolekar P, Balagopal V, Dong L, Liu Y, Foy S, Tran Q, et al. SJPedPanel: A Pan-Cancer Gene Panel for Childhood Malignancies to Enhance Cancer Monitoring and Early Detection. *Clinical Cancer Research*. 2024;30(18):4100–14. https://doi.org/10.1158/1078-0432.CCR-24-1063
5. Bader JS. The Panorama of Cancer Genetics. *Cancer Research*. 2021;81(10):2586–7. https://doi.org/10.1158/0008-5472.CAN-21-0885
6. Wu S, Zhu W, Thompson P, Hannun YA. Evaluating intrinsic and non-intrinsic cancer risk factors. *Nature Communications*. 2018;9(1):3490. https://doi.org/10.1038/s41467-018-05467-z
7. Hu C, Hart SN, Gnanaolivu R, Huang H, Lee KY, Na J, et al. A population-based study of genes previously implicated in breast cancer. *The New England Journal of Medicine*. 2021;384(5):440–51. https://doi.org/10.1056/NEJMoa2005936
8. Yang X, Song H, Leslie G, Engel C, Hahnen E, Auber B, et al. Ovarian and breast cancer risks associated with pathogenic variants in RAD51C and RAD51D. *Journal of National Cancer Institute*. 2020;112(12):1242–50. https://doi.org/10.1093/jnci/djaa030
9. Kastrinos F, Samadder NJ, Burt RW. Use of family history and genetic testing to determine risk of colorectal cancer. *Gastroenterology*. 2020;158(2):389–403. https://doi.org/10.1053/j.gastro.2019.11.029
10. Janssen B, Bellis S, Koller T, Tischkowitz M, Liau SS. A systematic review of predicted pathogenic *PALB2* variants: an analysis of mutational overlap between epithelial cancers. *Journal of Human Genetics*. 2020;65(2):199–205. https://doi.org/10.1038/s10038-019-0680-7
11. Wang Y, Zhou B, Ru J, Meng X, Wang Y, Liu W, et al. Advances in computational methods for identifying cancer driver genes. *Mathematical Biosciences and Engineering*. 2023;20(12):21643–69. https://doi.org/10.3934/mbe.2023958
12. Shickh S, Gutierrez Salazar M, Zakoor KR, Lazaro C, Gu J, Goltz J, et al. Exome and genome sequencing in adults with undiagnosed disease: a prospective cohort study. *Journal of Medical Genetics*. 2021;58(4):275–83. https://doi.org/10.1136/jmedgenet-2020-106936
13. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genetics in Medicine*. 2017;19(2):249–55. https://doi.org/10.1038/gim.2016.190
14. Miller DT, Lee K, Chung WK, Gordon AS, Herman GE, Klein TE, et al. ACMG SF v3.0 list for reporting of secondary findings in clinical exome and genome sequencing: a policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genetics in Medicine*. 2021;23(8):1381–90. https://doi.org/10.1038/s41436-021-01172-3
15. Bowling KM, Thompson ML, Finnila CR, Hiatt SM, Latner DR, Amaral MD, et al. Genome sequencing as a first-line diagnostic test for hospitalized infants. *Genetics in Medicine*. 2022;24(4):851–61. https://doi.org/10.1016/j.gim.2021.11.020
16. Harrison SM, Biesecker LG, Rehm HL. Overview of Specifications to the ACMG/AMP Variant Interpretation Guidelines. *Current Protocols in Human Genetics*. 2019;103(1):e93. https://doi.org/10.1002/cphg.93
17. Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *ASHG*. 2017;100(2):267–80. https://doi.org/10.1016/j.ajhg.2017.01.004
18. Cancer Genome Atlas Research Network, Weinstein JN, Collison EA, Mills GB, Mills Shaw KR, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics*. 2013;45(10):1113–20. https://doi.org/10.1038/ng.2764
19. Gusakova M, Dzhumaniiazova I, Zelenova E, Kashtanova D, Ivanov M, Mamchur A, et al. Prevalence of the cancer-associated germline variants in Russian adults and long-living individuals: using the ACMG recommendations and computational interpreters for pathogenicity assessment. *Frontiers in Oncology*. 2024;14:1420176. https://doi.org/10.3389/fonc.2024.1420176
20. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Kallberg M, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nature Methods*. 2018;15:591–4. https://doi.org/10.1038/s41592-018-0051-x
21. Virolainen SJ, VonHandorf A, Viel KCMF, Weirauch MT, Kottyan LC. Gene-environment interactions and their impact on human health. *Genes and Immunology*. 2023;24(1):1–11. https://doi.org/10.1038/s41435-022-00192-6
22. Kaehler M, Cascorbi I. Germline variants in cancer therapy. *Cancer Drug Resistance*. 2019;2(1):18–30. https://doi.org/10.20517/cdr.2019.05

Authors' contributions. All authors confirm that their contributions meet the ICMJE authorship criteria. The contributions are distributed as follows: Vladimir S. Yudin — development of the research methodology and conceptual framework, formulation of scientific objectives, writing, editing, and approval of the final manuscript; Aleksandra A. Mamchur — formal data analysis, synthesis and systematization of results, data visualization, manuscript editing; Ekaterina A. Nikolaeva — synthesis of results, manuscript drafting and formatting; Mikhail V. Ivanov — software development, data administration, data visualization; Daria A. Kashtanova — scientific supervision, critical revision of the manuscript, result verification; Elena A. Zelenova — formal data analysis, conducting the study; Sergey I. Mitrofanov — software development, conducting the study; Ekaterina A. Snigir — development of methodology, conducting the study; Lyudmila P. Kuzmina — research leadership and coordination, provision of resources, result verification; Anton A. Keskinov — project administration, provision of resources.

AUTHORS

Vladimir S. Yudin, Cand. Sci. (Biol.)
https://orcid.org/0000-0002-9199-6258
VYudin@cspfmba.ru

Aleksandra A. Mamchur
https://orcid.org/0000-0002-6025-7663
AMamchur@cspfmba.ru

Ekaterina A. Nikolaeva
https://orcid.org/0000-0002-7954-2560
Enikolaeva@cspfmba.ru

Mikhail V. Ivanov
https://orcid.org/0009-0004-7070-5636
Mivanov@cspfmba.ru

Daria A. Kashtanova, Cand. Sci. (Med.)
https://orcid.org/0000-0001-8977-4384
DKashtanova@cspfmba.ru

Elena A. Zelenova
https://orcid.org/0000-0003-4604-7261
Ezelenova@cspfmba.ru

Sergey I. Mitrofanov
https://orcid.org/0000-0003-0358-0568
Mitrofanov@cspfmba.ru

Ekaterina A. Snigir, Cand. Sci. (Biol.)
https://orcid.org/0000-0003-1245-7764
Esnigir@cspfmba.ru

Lyudmila P. Kuzmina, Dr. Sci. (Biol.)
https://orcid.org/0000-0003-3186-8024
lpkuzmina@mail.ru

Anton A. Keskinov, Cand. Sci. (Med.)
https://orcid.org/0000-0001-7378-983X
Keskinov@cspfmba.ru