

**В.Г. Вилков, А.Д. Деев, С.А. Шальнова**

**КЛАСТЕРЫ ФАКТОРОВ РИСКА ХРОНИЧЕСКИХ  
НЕИНФЕКЦИОННЫХ ЗАБОЛЕВАНИЙ  
(по данным проспективных исследований  
российской и американской когорт)**

**Москва**

**2010**

ББК 53

УДК 616-079.7

Рекомендовано к печати Ученым советом ФГУ «Государственный научно-исследовательский центр профилактической медицины Росмедтехнологий»

Рецензенты:

доктор медицинских наук, профессор, чл.-корр. РАМН *В.В. Кухарчук*

доктор медицинских наук, кандидат технических наук, профессор

*В.П. Невзоров*

**Вилков В.Г., Деев А.Д., Шальнова С.А.** Кластеры факторов риска хронических неинфекционных заболеваний (по данным проспективных исследований российской и американской когорт) – М.: ООО "Центр полиграфических услуг «Радуга»", 2010. – 120 с.

ISBN 978-5-94013-143-4

В книге посредством вторичного анализа данных популяционных проспективных исследований в России (по данным Государственного научно-исследовательского центра профилактической медицины) и США (the First National Health and Nutrition Examination Survey – NHANES I, NHANES I Epidemiologic Followup Study – NHEFS, the Second National Health and Nutrition Examination Survey – NHANES II, NHANES II Mortality Study – NH2MS) разработан оригинальный способ расчета риска фатальных событий в самоорганизующихся кластерах факторов риска. Построены таблицы 10-летних и 20-летних рисков смерти от сердечно-сосудистых заболеваний и всех причин с учетом возраста, пола, статуса курения, уровня образования, величин систолического артериального давления и частоты сердечных сокращений.

Использованы вычислительные эксперименты и многомерные статистические методы, включая искусственные нейронные сети.

Для врачей и научных сотрудников.

Издание содержит 40 рисунков, 15 таблиц, библиография – 90 наименований.

Электронная версия монографии представлена в Internet.

ББК 53

**Вилков Владимир Галикович, Деев Александр Дмитриевич,  
Шальнова Светлана Анатольевна**  
**КЛАСТЕРЫ ФАКТОРОВ РИСКА ХРОНИЧЕСКИХ  
НЕИНФЕКЦИОННЫХ ЗАБОЛЕВАНИЙ (по данным  
проспективных исследований российской и американской когорт)**

Печатается с представленного авторами оригинал-макета

Подписано в печать 25.06.2010 г.

Формат 60х90/16

Усл. печ.л. 7,5. Тираж 500. Заказ № 163-06/10

Отпечатано в ООО «Центр полиграфических услуг «Радуга»

Тел.: (495) 739-5680

<http://www.raduga-print.ru>, <http://www.radugaprint.ru>

ISBN 978-5-94013-143-4

© Вилков В.Г., Деев А.Д., Шальнова С.А., 2010

## ПРЕДИСЛОВИЕ

В монографии представлены результаты вторичного анализа данных популяционных проспективных исследований в России и США.

Обследование российской когорты выполнялось сотрудниками Института профилактической кардиологии Всесоюзного кардиологического научного центра АМН СССР, впоследствии Всесоюзного научно-исследовательского центра профилактической медицины и Государственного научно-исследовательского центра профилактической медицины, которым авторы выражают свою искреннюю благодарность.

Анализ связей между факторами риска (ФР) и смертностью в когорте США стал возможен благодаря свободному доступу к данным the First National Health and Nutrition Examination Survey (NHANES I), NHANES I Epidemiologic Followup Study (NHEFS), the Second National Health and Nutrition Examination Survey (NHANES II) и NHANES II Mortality Study (NH2MS), предоставленному the National Center for Health Statistics (NCHS) США [1].

Ответственность за результаты анализа, интерпретацию и выводы лежит на авторах, NCHS ответственен только за первичные данные.

## ВВЕДЕНИЕ

*С.А. Шальнова, В.Г. Вилков, А.Д. Деев*

Концепция ФР уже более полувека является основополагающей для ряда клинических и теоретических медицинских дисциплин и в значительной степени определяет профилактические, лечебные и диагностические стратегии в отношении хронических неинфекционных заболеваний и сердечно-сосудистых заболеваний (ССЗ) в частности. Поэтому изучение влияния ФР на заболеваемость и смертность продолжает оставаться актуальным.

У конкретного индивида нередко имеется сочетание нескольких ФР, многие ФР взаимосвязаны на уровне патогенеза высокозначимых для смертности хронических заболеваний, в частности атеросклероза. Следовательно, при изучении ФР необходимо использовать многомерные статистические методы. При этом возникает ряд проблем.

1. Для медико-биологических систем характерны сложность, большое число элементов и многообразие связей между ними.

Знания, отражающие состояние современной медицинской науки, как правило получены с применением методов статистического анализа, подразумевающих линейность зависимостей и нормальный закон распределения величин показателей. Однако уже при одномерном анализе связей ФР со смертностью в ряде случаев очевиден их нелинейный характер (например, U-образная зависимость для индекса массы тела). Складывается впечатление, что в данной предметной области нормальные распределения и линейные зависимости являются скорее исключением, чем правилом. Проведенные нами ранее исследования показали, что среди 23 показателей, характеризующих ФР ССЗ, не было ни одного, распределение величин которого строго подчинялось нормальному закону [2].

Вышеизложенное позволяет предположить, что использование параметрических линейных методов статистического анализа может привести к существенному искажению результатов.

2. «Проклятие размерности». При множественных сравнениях и использовании многомерных моделей с увеличением числа переменных быстро возрастает необходимое для надежных выводов число наблюдений, для обеспечения удовлетворительной точности требуются объемы выборок, трудно достижимые даже в крупномасштабных популяционных исследованиях.

3. Ряд используемых в настоящем исследовании переменных (пол, статус курения, уровень образования) являются номинативными, что накладывает дополнительные ограничения на применение традиционных параметрических методов статистики.

Одним из перспективных путей дальнейшего развития концепции ФР мы считаем вторичный анализ накопленных к настоящему времени данных популяционных исследований с использованием ориентированных на ненормальные распределения и нелинейные зависимости методов, в частности для выявления неоднородности многомерного пространства характеризующих риск факторов.

Использовавшиеся до сих пор способы прогнозирования риска основаны на многомерных регрессионных моделях, описывающих связи между ФР и исходами [3-7]. Однако эти модели учитывают линейные составляющие связей исходов с включенными в модель ФР, нелинейные составляющие утрачиваются, если они не включены в модель явным образом.

Основная особенность предлагаемой нами концепции состоит в том, чтобы использовать не моделирование с неизбежно присущими ему ограничениями, упрощениями и искажениями [8], а прямой расчет рисков в однородных в отношении ФР (с учетом их взаимосвязей) частях когорты. Преимущество такого подхода заключается в учете всех взаимодействий ФР без присущих моделированию упрощений и ограничений линейных статистических методов.

Очевидно, что качество прогнозирования течения заболевания и его исхода у индивида зависит от сходства между ним и

группой людей, на изучении которых основан прогноз [9]. Предлагаемое нами решение состоит в выделении из когорты кластеров\*, объединяющих лиц с близкими величинами показателей, характеризующих учтенные ФР. Для каждого кластера по реальным данным рассчитывается абсолютный риск, величина которого в дальнейшем соотносится с соответствующими ячейками таблицы риска. Более подробно технические детали представлены в разделе 2.

Цель настоящего исследования – разработка нового способа оценки суммарного риска развития фатального события.

Для решения поставленных задач использованы данные выполненных в России и США крупных проспективных популяционных исследований и базирующиеся на нейросетевых технологиях методы анализа.

Использовали данные проспективных российских исследований, выполненных в разные годы в Государственном научно-исследовательском центре профилактической медицины [10-11, 7] и нескольких исследований неорганизованной популяции США серии NHANES (the First National Health and Nutrition Examination Survey – NHANES I, NHANES I Epidemiologic Followup Study – NHEFS, the Second National Health and Nutrition Examination Survey – NHANES II, NHANES II Mortality Study – NH2MS). Характеристики когорт представлены в разделе 1 и таблицах 1-2.

Искусственные нейронные сети (ИНС) – это класс аналитических методов, построенных на принципах, сходных с гипотетическим функционированием мозга мыслящих существ и позволяющих прогнозировать значения некоторых переменных в новых наблюдениях по данным других наблюдений (для этих же или других переменных) после прохождения этапа так называемого обучения на имеющихся данных [12, 13]. ИНС воз-

---

\* Такие кластеры названы нами «самоорганизующиеся кластеры факторов риска» (СОК ФР), поскольку ключевая роль в использованной статистической технике принадлежит самоорганизующимся картам Кохонена (Self-Organizing Maps – SOM) – см. ниже раздел 2.

ники из исследований в области искусственного интеллекта, в частности из попыток воспроизвести способность биологических нервных систем обучаться, моделируя низкоуровневую структуру мозга [14, 12]. Для медицинских задач характерны большой объем входной информации, невозможность учета всех влияющих на ответ факторов, отсутствие точного алгоритма нахождения ответа, приблизительный характер последнего и наличие нескольких вариантов решения. Это соответствует основным принципам функционирования ИНС и способу выдачи ими результата. Поэтому задачи медицины и биологии являются перспективным полем для нейросетевых технологий и именно в этой области наблюдается наиболее яркий практический успех нейросетевых методов [15, 16]. Существует обширная литература по теории ИНС и их прикладному применению [17, 15, 13, и др.]. Используемые в настоящей работе нейросетевые методы более подробно описаны в разделах 1 и 2.

Как упоминалось выше, в данной монографии использован *вторичный анализ данных*, под которым понимали «любой последующий анализ существующего массива данных, который предоставляет интерпретации, выводы или знание, дополнительное или отличное от предоставленного в первом отчете по исследованию в целом и его основным результатам» [18]. Такой подход часто применяется в социологических исследованиях, особенно в США [19, 20]. В настоящее время вторичный анализ данных нередко используется и в медицинских исследованиях, в основном зарубежных.

Использующие первичный и вторичный анализ данных исследования различаются тем, что в первом случае производится сбор и анализ данных, а во втором аналитические методы применяются к уже существующим данным.

Вторичный анализ данных следует отличать от метаанализа. В первом случае проводится новый, расширенный и т.п. анализ первичных данных, во втором используется синтез результатов многих авторов и обобщаются сделанные ими выводы.

При крупных многоцелевых исследованиях границы между первичным и вторичным анализом отчасти стираются, поскольку их результаты содержат огромный объем информации, значительная часть которой не использована первоначальными исследователями. Участие в анализе других исследователей, решающих стоящие перед ними задачи, приводит к более полному использованию данных и в итоге к экономии средств. Разные исследователи даже для решения схожих вопросов могут применять различные подходы и методы анализа, что, возможно, расширит понимание предмета.

Таким образом, преимуществами вторичного анализа являются значительная экономия ресурсов, отсутствие связанных со сбором данных проблем и, как правило, более высокий уровень обобщения.

Ограничения вторичного анализа связаны с доступностью нужных данных и наследованием методических ошибок первичного исследования. Методические погрешности в крупномасштабных тщательно спланированных исследованиях типа NHANES минимальны, а первая проблема решается интенсивно развивающимися архивами данных – организациями, которые собирают, классифицируют данные и предоставляют их исследователям в удобной форме (примером в области социологии является International Data Library and Reference Service, Survey Research Center, University of California, Berkeley) [21, 20].

Определения использованных в книге терминов представлены в словаре.

Цветные таблицы 10- и 20-летних рисков фатальных событий на базе традиционного и предлагаемого нами подходов представлены в конце книги на цветных рисунках 1-32.

Большое количество таблиц и графиков обусловлено стремлением предоставить читателю максимум возможностей для самостоятельного ознакомления с фактами.



# 1. ИСТОЧНИКИ ДАННЫХ И МЕТОДЫ СТАТИСТИЧЕСКОГО АНАЛИЗА

*В.Г. Вилков, А.Д. Деев*

## 1.1. Данные

### *Российские исследования*

Для построения таблиц риска фатального события использовали данные проспективных российских исследований, протокол которых опубликован ранее [10, 11]. Первичное обследование простой пропорциональной случайной выборки проводилось в 1975-1982 гг., когорта включала 11136 мужчин и 5744 женщин в возрасте 18-72 лет, последняя информация о жизненном статусе обследуемых была получена в 2002 г.

В настоящей работе использовали данные проспективного наблюдения за включенными в упомянутую когорту 9165 мужчинами и 4511 женщинами в возрасте 35 лет и старше (табл. 1), у которых при первичном обследовании были зарегистрированы перечисленные в таблице 2 показатели. В указанной выборке наблюдалось 4673 смертей на 168 тысяч «человеко-лет» наблюдения за мужчинами и 1043 смертей на 71 тысячу «человеко-лет» наблюдения за женщинами. Анализировали смертность от ССЗ и общую смертность. Причины смерти кодировались по международной классификации болезней 8 пересмотра (МКБ-8). Причина смерти идентифицировалась только при наличии врачебного свидетельства о смерти. Если в качестве причины смерти указывалось сердечно-сосудистое или связанное с ним заболевание, на начальном этапе исследования собиралась дополнительная информация из больницы, от лечащего врача, от родственников и от лиц, непосредственных свидетелей смерти, изучались протоколы аутопсий. В последующем, после демонстрации хорошего совпадения с данными о причинах смерти из официального врачебного свидетельства о смерти, использовали последние.

Таблица 1

**Число наблюдений в изученных когортах российских и американских проспективных исследований**

| Когорта                 | Наблюдения              | Причина смерти |           |                                 |           |
|-------------------------|-------------------------|----------------|-----------|---------------------------------|-----------|
|                         |                         | все причины    |           | сердечно-сосудистые заболевания |           |
|                         |                         | муж.           | жен.      | муж.                            | жен.      |
| Российские исследования | все                     | 9165           | 4511      | 9165                            | 4511      |
|                         | цензурированные         | 4492           | 3468      | 6835                            | 4034      |
|                         | завершенные             | 4673           | 1043      | 2330                            | 477       |
|                         | <i>доля заверш. (%)</i> | <i>51</i>      | <i>23</i> | <i>25</i>                       | <i>11</i> |
| NHEFS                   | все                     | 2397           | 2725      | 2397                            | 2725      |
|                         | цензурированные         | 1349           | 1931      | 1874                            | 2337      |
|                         | завершенные             | 1048           | 794       | 523                             | 388       |
|                         | <i>доля заверш. (%)</i> | <i>44</i>      | <i>29</i> | <i>22</i>                       | <i>14</i> |
| NH2MS                   | все                     | 3793           | 4249      | 3793                            | 4249      |
|                         | цензурированные         | 2579           | 3392      | 3222                            | 3848      |
|                         | завершенные             | 1214           | 857       | 571                             | 401       |
|                         | <i>доля заверш. (%)</i> | <i>32</i>      | <i>20</i> | <i>15</i>                       | <i>9</i>  |

Примечание. Условные обозначения – см. Список сокращений.

*Американские исследования*

Использовали данные двух проспективных исследований – NHANES I Epidemiologic Followup Study (NHEFS) и NHANES II Mortality Study (NH2MS), состав когорт представлен в таблице 1.

Данные NHEFS получены при проспективном наблюдении за лицами, участвовавшими в одномоментном исследовании the First National Health and Nutrition Examination Survey (NHANES I). Детальное описание NHEFS представлено в [22], файлы данных свободно доступны через Internet [23].

NHANES I представляет собой выполненное в 1971-1975 гг. одномоментное исследование неорганизованной популяции гражданского населения США в возрасте 1-74 лет. Репрезентативная выборка состояла из 32 тыс. человек обоего пола. Методика исследования описана в [24, 25], файлы данных свободно доступны через Internet [26].

Когорта исследования NHEFS включала 14407 человек, обследованных при NHANES I в возрасте 25-74 лет, последняя информация о жизненном статусе обследуемых была получена в 1992 г.

В настоящей работе использованы данные обследования в NHANES I лиц в возрасте 35 лет и старше (только белых), показатели и их размерности приведены в таблице 2. Изученная нами когорта состояла из 2397 мужчин и 2725 женщин (табл. 1), у которых при NHANES I были зарегистрированы все перечисленные в таблице 2 показатели и имелись данные проспективного наблюдения. В указанной выборке наблюдалось 1048 смертей на 35 тысяч «человеко-лет» наблюдения за мужчинами и 794 смерти на 45 тысяч «человеко-лет» наблюдения за женщинами. В анализ включали смертность от ССЗ и общую смертность. Причины смерти кодировались по МКБ-9.

Данные NH2MS получены при проспективном наблюдении за лицами, участвовавшими в одномоментном исследовании the Second National Health and Nutrition Examination Survey (NHANES II). Детальное описание NH2MS представлено в [27], файлы данных свободно доступны через Internet [28].

NHANES II представляет собой выполненное в 1976-1980 гг. одномоментное исследование неорганизованной популяции гражданского населения США в возрасте 0,5-74 лет. Репрезентативная выборка состояла из 28 тыс. человек обоего пола. Методика исследования описана в [29], файлы данных свободно доступны через Internet [30].

Когорта исследования NH2MS включала 9250 человек, обследованных при NHANES II в возрасте 30-75 лет, последняя информация о жизненном статусе обследуемых была получена в 1992 г.

В настоящем разделе использованы данные обследования в NHANES II лиц в возрасте 35 лет и старше (только белых). Когорта состояла из 3793 мужчин и 4249 женщин (табл. 1), для которых имелись данные проспективного наблюдения и при NHANES II были зарегистрированы перечисленные в таблице 2 показатели. В указанной выборке наблюдалось 1214 смертей на 48 тысяч «человеко-лет» наблюдения за мужчинами и 857 смертей на 58 тысяч «человеко-лет» наблюдения за женщинами. В анализ включали смертность от ССЗ и общую смертность. Причина смерти закодирована по МКБ-9.

Когорты исследований NHEFS и NH2MS были объединены, так как их дизайн и время проведения достаточно близки. Объединенная когорта состояла из 6190 мужчин (2262 смерти на 83 тысячи «человеко-лет» наблюдения) и 6974 женщин (1651 смертей на 103 тысячи «человеко-лет» наблюдения).

## **1.2. Обоснование выбора характеризующих ФР показателей**

Известные к настоящему времени ФР исчисляются сотнями, однако основных «классических» ФР значительно меньше, порядка десятка.

Во всех использованных нами наборах данных доступны следующие характеризующие ФР показатели: возраст, пол, статус курения, уровень образования, концентрация общего холестерина в сыворотке крови, величины артериального давления (АД), частоты сердечных сокращений (ЧСС), индекс массы тела.

Некоторые из перечисленных показателей, зарегистрированные в российском и американских исследованиях, несколько различались, в связи с чем были произведены преобразования для получения унифицированных для всех исследований переменных.

*Уровень образования.* При анализе данных использовали номинативную переменную с двумя значениями: «образование среднее или ниже», «образование выше среднего».

В российских исследованиях вариантами ответа на данный

вопрос были: образование ниже среднего, среднее, выше среднего. Значение унифицированной переменной «образование среднее или ниже» получали, объединяя две первые градации.

В исследованиях серии NHANES ответ на вопрос об образовании более развернутый. Значение унифицированной переменной «образование среднее или ниже» присваивали лицам с образованием до 12-го класса нормальной школы включительно, значение «образование выше среднего» – начиная с 1-го года обучения в колледже.

*Статус курения.* При анализе данных использовали номинативную переменную с двумя значениями: «некурящие» и «курящие».

В российских исследованиях значение унифицированной переменной «некурящие» присваивали лицам никогда не курившим либо бросившим курить, а значение «курящие» – курящим в настоящее время вне зависимости от количества выкуриваемых сигарет или папирос.

Опросники исследований серии NHANES содержат большое количество вопросов относительно употребления табака. Информация была агрегирована следующим образом. К «некурящим» относили лиц, выкуривших за время жизни не более 100 сигарет (50 сигар, 3 пачек трубочного табака), а также лиц, отрицательно ответивших на вопрос об употреблении табака в момент обследования вне зависимости от количества выкуренных за время жизни сигарет (сигар, трубок). К «курящим» относили лиц, выкуривших за время жизни более 100 сигарет (50 сигар, 3 пачек трубочного табака) и положительно ответивших на вопрос об употреблении табака в момент обследования.

*АД, измеренное методом Н.С. Короткова в положении сидя.*

*ЧСС в положении сидя по данным пальпации пульса.*

*Концентрация общего холестерина в сыворотке крови.*

*Индекс массы тела, рассчитанный как отношение массы тела (кг) к росту (м) в квадрате.*

При анализе данных исследований серии NHANES были использованы такие переменные, как *уровень физической активности* и *потребление кофе и/или чая* (раздел 8).

*Характеризующие результаты проспективного наблюдения*

переменные включали код причины смерти (на базе МКБ-8 или МКБ-9), цензурирующие переменные для смерти от ССЗ и всех причин, продолжительность проспективного наблюдения.

Лиц, о которых не было полных сведений после проведения первичного медицинского обследования, из анализа исключали.

Перечень унифицированных переменных представлен в таблице 2.

За исключением показателей проспективного наблюдения все данные были зарегистрированы при первичном обследовании.

Методики исследований описаны в [24-25, 29, 10-11, 22, 27].

Для отбора наиболее информативных переменных (в контексте поставленных задач) проведен предварительный анализ связей перечисленных выше ФР с выживанием.

Раздельно для мужчин и женщин когорт России и США были построены модели пропорциональных интенсивностей Кокса [31-33], включающие в качестве зависимой переменной время дожития в отношении ССЗ или всех причин, а в качестве независимых переменных:

- возраст;
- систолическое и диастолическое АД;
- ЧСС;
- индекс массы тела;
- концентрацию общего холестерина в сыворотке крови;
- статус курения;
- уровень образования.

Обнаружено, что связь индекса массы тела с дожитием заметно слабее в сравнении с другими исследованными ФР, указанные связи статистически незначимы в большинстве когорт (исключая мужчин российской когорты).

Концентрация общего холестерина в крови является значимым предиктором для дожития преимущественно в отношении ССЗ у мужчин, в отношении смерти от всех причин эта переменная, как правило, менее информативна. Последнее согласуется с результатами [34], которые по данным 15-летнего проспективного наблюдения за когортой из 21766 мужчин устано-

вили, что показатели липидов крови не являются независимыми предикторами смертности вследствие всех причин.

При выборе между систолическим и диастолическим АД исходили из следующих соображений. Нами был проведен предварительный анализ с использованием моделей пропорциональных интенсивностей Кокса с различными комбинациями независимых переменных, включая систолическое и диастолическое АД, который показал, что в большинстве случаев связи между дожитием и систолическим АД сильнее в сравнении с диастолическим АД. Этот результат согласуется с данными литературы. Ряд авторов на основе анализа нескольких проспективных популяционных исследований, включая Фремингемское и Multiple Risk Factor Intervention Trial, сделали вывод о том, что значимость систолической артериальной гипертензии как ФР ССЗ сопоставима [35-38] или превышает [39-41] значимость диастолической артериальной гипертензии. На материалах российских популяционных исследований было показано, что систолическое АД сильнее влияет на смертность от ишемической болезни сердца (ИБС) и мозгового инсульта в сравнении с диастолическим АД [42].

При включении ЧСС в число учитываемых в наших таблицах ФР руководствовались результатами описанного выше предварительного анализа, а также ряда популяционных исследований, свидетельствующих о роли ЧСС как независимого ФР. Известно, что с увеличением ЧСС возрастает риск внезапной смерти [43], возникновения ССЗ и смерти от ССЗ и мозгового инсульта [44-49]. Относительно наличия значимой связи ЧСС с риском смерти от всех причин сведения менее однозначны: по данным исследований серии NHANES [46, 47] и ряда других, включая российские [Mensink G.B., 1997 258 /id}, 51], такая связь значима по крайней мере в части подгрупп; по данным других исследований [52] связь ЧСС с общей смертностью была слабой. Результаты метаанализа эпидемиологических исследований с общим числом наблюдений 116 тысяч человек в возрасте 18-80 лет при длительности наблюдения 5-36 лет свидетельствуют об увеличении риска смерти от всех причин и ССЗ у лиц с высокой ЧСС, причем ЧСС остается значи-

мым ФР после учета таких известных ФР ИБС, как возраст, пол, АД, индекс массы тела и наличие сахарного диабета [53, 54].

Механизмы положительного влияния относительно низкой ЧСС на продолжительность жизни не вполне ясны. Есть основания считать, что ЧСС отражает скорость метаболизма, и, как следствие, связана с ожидаемой продолжительностью жизни [55, 56]. Заслуживают внимания сведения, полученные при изучении ЧСС с позиций биологии и сравнительной физиологии. Складывается впечатление, что общее количество сердечных сокращений в течение жизни живой особи является относительно постоянной величиной, типичная продолжительность жизни составляет несколько месяцев у мыши при ЧСС 600 в минуту и 2-3 десятилетия у крупных животных (кита, слона) при ЧСС 20 в минуту [57-59].

В прикладном аспекте выглядит достаточно обоснованным предположение о желательности достижения у индивида относительно невысокого уровня ЧСС [54], особенно учитывая соответствие такого вывода приведенным выше данным клинических и эпидемиологических исследований о связи ЧСС со смертностью. Немаловажной представляется перспективность учета ЧСС для коррекции, в том числе медикаментозной, этого ФР. Помимо бета-блокаторов, в этом плане перспективен представитель нового класса лекарственных средств ивабрадин, замедляющий синусовый ритм за счет избирательного ингибирования ионных токов  $I_f$ , не оказывающий отрицательного инотропного эффекта и не влияющий на атриовентрикулярную проводимость [54, 60].

Полученные с использованием моделей Кокса выводы были перепроверены посредством прямого расчета абсолютного риска в однородных кластерах ФР, построенных описанным в разделе 2 способом.

В итоге с учетом всей совокупности данных было принято решение при дальнейшем анализе использовать следующие переменные:

- пол;
- возраст;



- систолическое АД (САД);
- ЧСС;
- статус курения;
- уровень образования.

Вышеизложенное не означает, что мы отрицаем роль такого ФР, как дислипидемия, в частности гиперхолестеринемия. Однако применительно к задачам данного исследования и с учетом ограниченного числа характеризующих ФР показателей, которые могут быть учтены в таблицах риска, мы сочли более перспективным использовать ЧСС.

### **1.3. Обоснование выбора статистических методов**

Очевидно, что некорректное использование статистических методов анализа может привести к ложным выводам даже при правильно спланированном и осуществленном исследовании [61, 62].

В настоящей работе анализировались данные трех типов:

- номинативные (пол, статус курения, уровень образования, цензурирующие переменные, код причины смерти);
- количественные (возраст, продолжительность проспективного наблюдения, величины АД, ЧСС, концентрация в крови холестерина, индекс массы тела);
- порядковые (выраженные в баллах уровни физической активности, потребление кофе и чая).

При анализе количественных данных в медицинских исследованиях в настоящее время, как правило, используются параметрические методы, потенциально более мощные, удобные с технической точки зрения и наиболее привычные для исследователей. Однако эти методы позволяют корректно оценить лишь линейные составляющие связей между переменными [61, 62]. Практически все имеющиеся к настоящему времени представления о роли ФР ассоциированных с атеросклерозом заболеваний основаны на изучении их связей с выживаемостью посредством методов линейного моделирования, результаты которых тем лучше, чем ближе распределение величин показателей в выборке к нормальному [6].

**Таблица 2**

**Использованные при анализе данных показатели**

| <b>Показатель</b>  | <b>Единицы измерения</b> | <b>Примечания</b>  |
|--|--------------------------|--|
| Возраст  | годы                     | При первичном обследовании   |
| Пол  |                          | 1 – мужчины; 2 – женщины   |
| САД  | мм рт.ст.                | Аускультативным методом по I тону Короткова  |
| ЧСС  | уд./мин                  | Частота сердечных сокращений по данным пальпации пульса  |
| Индекс массы тела (ИМТ)                                      | кг/м <sup>2</sup>        |  |
| Общий холестерин (ОХ)  | ммоль/л                  | В сыворотке крови  |
| Статус курения   |                          | 1 – некурящие (включая никогда не куривших и бросивших курить);<br>2 – курящие в настоящее время |
| Уровень образования  |                          | 1 – образование среднее или ниже;<br>2 – образование выше среднего                               |
| Физическая активность в обычных условиях (ФА <sub>1</sub> )* | балл                     | 1 – отсутствует;<br>2 – умеренная;<br>3 – очень высокая  |
| Физическая активность в период отпуска (ФА <sub>2</sub> )*   | балл                     | 1 – отсутствует;<br>2 – умеренная;<br>3 – очень высокая  |
| Потребление кофе и / или чая                                 | балл                     | 1 – менее 1 чашки в день;<br>2 – 1-5 чашек в день;<br>3 – более 5 чашек в день                   |
| <b><i>Показатели проспективного наблюдения</i></b>           |                          |  |
| Код причины смерти   |                          | По МКБ-8 (Россия) или МКБ-9 (США)  |
| Цензурирующая переменная для смерти от всех причин           |                          | Завершенными наблюдениями считали смерть от всех причин, цензурированными – все остальные случаи |
| Цензурирующая переменная для смерти от ССЗ                   |                          | Завершенными наблюдениями считали смерть от всех ССЗ, цензурированными – все остальные случаи    |
| Продолжительность проспективного наблюдения                  | годы                     |  |

\* Только в наборах данных исследований серии NHANES.

Таблица 3

**Значимость ( $p$ ) отличий распределений величин  
количественных показателей от нормального закона в  
зависимости от числа наблюдений**

| Показатель                | Выборка   | $n$  |      |      |       |        |        |        |        |
|---------------------------|-----------|------|------|------|-------|--------|--------|--------|--------|
|                           |           | 30   | 50   | 100  | 200   | 500    | 1000   | 2000   | 4000   |
| Возраст,<br>лет           | РФ, жен.  | 0,35 | 0,97 | 0,09 | 0,10  | 0,002  | 0,0001 | 0,0001 | 0,0001 |
|                           | США, жен. | 0,51 | 0,51 | 0,03 | 0,04  | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
|                           | РФ, муж.  | 0,57 | 0,91 | 0,42 | 0,14  | 0,002  | 0,0001 | 0,0001 | 0,0001 |
|                           | США, муж. | 0,67 | 0,83 | 0,02 | 0,02  | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
| САД,<br>мм рт.ст.         | РФ, жен.  | 0,51 | 0,07 | 0,06 | 0,27  | 0,001  | 0,0001 | 0,0001 | 0,0001 |
|                           | США, жен. | 0,21 | 0,69 | 0,07 | 0,03  | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
|                           | РФ, муж.  | 0,45 | 0,10 | 0,08 | 0,02  | 0,01   | 0,0001 | 0,0001 | 0,0001 |
|                           | США, муж. | 0,55 | 0,75 | 0,03 | 0,04  | 0,001  | 0,0001 | 0,0001 | 0,0001 |
| ЧСС,<br>уд./мин           | РФ, жен.  | 0,77 | 0,19 | 0,21 | 0,01  | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
|                           | США, жен. | 0,73 | 0,07 | 0,41 | 0,03  | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
|                           | РФ, муж.  | 0,42 | 0,58 | 0,18 | 0,001 | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
|                           | США, муж. | 0,96 | 0,62 | 0,08 | 0,05  | 0,0001 | 0,0001 | 0,0001 | 0,0001 |
| ИМТ,<br>кг/м <sup>2</sup> | РФ, жен.  | 0,67 | 0,88 | 0,55 | 0,58  | 0,02   | 0,001  | 0,001  | 0,0001 |
|                           | США, жен. | 0,81 | 0,29 | 0,13 | 0,006 | 0,001  | 0,0001 | 0,0001 | 0,0001 |
|                           | РФ, муж.  | 1,0  | 0,83 | 0,74 | 0,46  | 0,04   | 0,18   | 0,01   | 0,0001 |
|                           | США, муж. | 0,32 | 0,86 | 0,55 | 0,53  | 0,06   | 0,004  | 0,001  | 0,0001 |
| ОХ,<br>ммоль/л            | РФ, жен.  | 0,30 | 0,91 | 0,38 | 0,01  | 0,04   | 0,03   | 0,0001 | 0,0001 |
|                           | США, жен. | 0,95 | 0,86 | 0,87 | 0,30  | 0,13   | 0,007  | 0,003  | 0,0001 |
|                           | РФ, муж.  | 0,79 | 0,89 | 0,43 | 0,30  | 0,04   | 0,16   | 0,001  | 0,0001 |
|                           | США, муж. | 0,84 | 0,98 | 0,67 | 0,24  | 0,16   | 0,23   | 0,002  | 0,0001 |

## Примечания.

1.  $p$  – значимость различий эмпирического и теоретического нормального распределений по критерию Колмогорова-Смирнова для одной выборки.
2.  $n$  – число наблюдений в случайным образом перемешанной подвыборке.
3. Условные обозначения – см. Список сокращений.

Неоднократно высказывалось мнение, что требуемые для применения параметрических статистических методов условия редко соблюдаются в медико-биологических исследованиях, в связи с чем в данной области вообще рекомендуют отказаться от использования параметрических критериев [63].

Исследуем этот вопрос более подробно применительно к анализируемым в настоящей работе данным.

Из ограничений параметрических методов на практике важное значение имеет нормальность распределения. Ее проверили в специально проведенном вычислительном эксперименте, который заключался в выделении из всей совокупности случайно перемешанных наблюдений подвыборок объемом от 30 до 4000 наблюдений. В каждой такой подвыборке для всех включенных в дальнейший анализ количественных показателей проверяли нормальность распределения посредством сравнения эмпирического распределения с теоретическим по критерию Колмогорова-Смирнова для одной выборки [64]. Результаты представлены в таблице 3.

Из таблицы видно, что для всех изученных показателей вывод об отсутствии значимых различий эмпирического и теоретического нормального распределений и, соответственно, обоснованности применения параметрических статистических методов, в значительной степени зависит от объема выборки. При числе наблюдений в пределах 100 в большинстве случаев (17 из 20) распределение величин показателей значимо не отличается от нормального ( $p > 0,05$ ). При дальнейшем наращивании объема выборки увеличивается число случаев, когда распределения величин показателя значимо отличаются от нормального ( $p < 0,05$ ). При выборках, включающих 2000 наблюдений и более распределения всех изученных показателей с высокой достоверностью отличаются от нормального,  $p$  не превышает 0,01 (табл. 3).

Проведенный вычислительный эксперимент позволяет сделать следующие выводы:

1. Объем выборки, достаточный для доказательства отличия эмпирического распределения от нормального, варьирует

для разных показателей и выборок от десятков до тысяч наблюдений. При типичных для клинических исследований объемах выборок порядка 30-100 наблюдений распределение большинства включенных в настоящий анализ показателей формально не отличается от нормального, хотя при дальнейшем увеличении числа наблюдений отличия от нормального закона становятся высокодостоверными.

2. Поэтому формальный подход к проверке нормальности распределения в малых выборках может провоцировать к использованию привычных и традиционно считающихся более мощными параметрических статистических критериев и линейных методов анализа.
3. Однако нормальное распределение нехарактерно для многих широко используемых в медицинских исследованиях количественных показателей, включая исследованные в данной работе основные ФР хронических неинфекционных заболеваний.
4. Хотя параметрические методы потенциально являются более мощными, при явном несоответствии распределения показателя нормальному закону их использование может повлечь существенное искажение результатов анализа.
5. Заслуживает внимания рекомендация отказаться от использования линейных методов и параметрических критериев в медико-биологических исследованиях [63]. Во всяком случае, целесообразно максимально применять нелинейные статистические методы и непараметрические критерии там, где это возможно.

Сформулированные выше выводы нельзя, конечно, безоговорочно распространять на все медико-биологические данные, для этого требуется проведение подобного вычислительного эксперимента для соответствующих показателей. Ранее нами были опубликованы [2] сходные результаты относительно нормальности распределения таких показателей, как:

- систолическое и диастолическое АД;

- разнообразные антропометрические показатели;
- типичные для опросников количественные показатели;
- лабораторные показатели, включая некоторые гормоны, показатели липидного и углеводного обменов.

Применительно к задачам настоящей работы представляется достаточно обоснованным вывод о необходимости использования нелинейных методов анализа.

На практике это было реализовано следующим образом.

Сравнение величин порядковых и количественных показателей в двух независимых выборках осуществляли с использованием *U*-критерия Манна-Уитни, в двух связанных выборках – критерия знаков и критерия Вилкоксона для разностей рангов.

Кластеризацию в многомерном пространстве ФР осуществляли посредством построения искусственной нейронной сети (самообучающейся карты Кохонена) с ее последующим обучением и тестированием [65-66, 12-13, 67-72], подробнее см. раздел 2.

При анализе выживаемости применяли метод множительных оценок Каплана-Мейера и процедуры сравнения выживаемости в нескольких выборках.

Модель пропорциональных интенсивностей Кокса [31-33] использовали на этапе предварительного анализа данных, а также для сравнения результатов нового и традиционного подходов к прогнозированию риска (в качестве прототипа).

Расчеты производили с использованием стандартных пакетов статистических программ для персональных компьютеров [71, 70, 73, 33, 61, 64].

## **2. СПОСОБ РАСЧЕТА РИСКА ФАТАЛЬНОГО СОБЫТИЯ В САМООРГАНИЗУЮЩИХСЯ КЛАСТЕРАХ ФАКТОРОВ РИСКА**

*В.Г. Вилков*

Предлагаемый нами подход основан на прямом расчете абсолютных рисков в кластерах, выделенных из когорты по принципу однородности в многомерном пространстве учтенных ФР.

Его преимущество в том, что влияние ФР учитывается во всей естественной сложности их взаимодействий, без присущего моделированию упрощения и связанных с использованием линейных статистических методов ограничений. Основная трудность состоит в обеспечении для каждого кластера, характеризуемого данным сочетанием ФР, не слишком малого числа наблюдений. Из-за этого простое решение, состоящее в подсчете абсолютных рисков для каждой ячейки таблицы типа приведенной на цветном рисунке 1, практически неосуществимо, т.к. при использовании представительной выборки из популяции в соответствующих некоторым ячейкам таблицы подгруппах число наблюдений оказывается явно недостаточным даже при очень больших, содержащих много тысяч наблюдений, когортах. Например, в ячейках таблицы на цветном рисунке 1, соответствующих курящим российским женщинам с САД 180 мм рт.ст., фактически имеются лишь единичные наблюдения (рис. 33-34).

Очевидно, что требуется каким-либо образом укрупнить подгруппы, в которых будет рассчитываться риск. Производить укрупнение по формальным признакам, например, объединяя несколько ячеек таблицы типа представленной на цветном рисунке 1, нецелесообразно, т.к. подобное формальное усреднение может нивелировать особенности отдельных категорий лиц, важные для последующей разработки дифференцированных профилактических и лечебных мероприятий.

Предлагаемое нами решение состоит в использовании кластерного анализа для выделения кластеров, объединяющих лиц

с близкими в многомерном пространстве данного набора ФР величинами этих факторов. Тем самым обеспечивается учет взаимосвязей исследуемых ФР. Затем для каждого такого кластера по реальным данным рассчитывается абсолютный риск, величина которого в дальнейшем соотносится с соответствующими данному кластеру ячейками таблицы риска.

Учитывая приведенные выше соображения о применении нелинейных методов, для кластеризации нами были использованы ИНС, а именно самоорганизующиеся карты Кохонена [65-66, 12-13, 67-72].

Принципы функционирования ИНС кратко могут быть охарактеризованы следующим образом. ИНС состоит из совокупности простых элементов, называемых искусственными нейронами, которые определенным образом соединены между собой и с внешней средой посредством связей, определяемых весовыми коэффициентами. Обычно в сети имеются входные, промежуточные (скрытые) и выходные нейроны. На первые подается закодированное входное воздействие, с последних снимается результат классификации или прогноза. В промежуточных нейронах осуществляется переработка информации. Нейроны могут объединяться в слои, содержащие совокупности нейронов с едиными входными сигналами. В целом с увеличением числа нейронов возрастают возможности сети к моделированию сложных зависимостей, однако при этом требуется большее число примеров для ее обучения и тестирования.

Самоорганизующиеся карты (сети) Кохонена (Self-Organizing Maps, SOM) представляют собой ИНС, основанные на воспроизведении топологических свойств человеческого мозга и предназначенные для решения задач кластеризации посредством неуправляемого обучения (классификации без учителя) [65, 13].

Способ реализации ИНС может быть различным, включая специализированные нейрокомпьютеры. В нашем случае среду нейрокомпьютера эмулировали с использованием программы-имитатора на обычном компьютере, функционирующем на принципах машины фон Неймана [13].



Применение ИНС обеспечивает следующие принципиальные преимущества:

- ИНС представляют собой исключительно мощный метод моделирования, позволяющий воспроизводить зависимости практически любой сложности;
- они нелинейны по своей природе, тогда как до сих пор рассматриваемая задача решалась средствами линейного моделирования;
- лучше по сравнению с традиционными методами работают при большом числе переменных;
- обучаются на примерах, не требуя априорных знаний о характере связей между входными данными и результатом, в нашем случае – о механизме влияния ФР на выживаемость.

Применение ИНС, в том числе карт Кохонена, сопряжено с рядом трудностей. В отличие от методов традиционной статистики, при использовании ИНС для конкретной практической задачи может быть построено много работоспособных и удовлетворяющих формальным критериям точности сетей, и, соответственно, найдено не одно, а множество решений, поэтому возникает проблема выбора лучшей сети. Применительно к кластеризации и сетям Кохонена указанный выбор усложняется отсутствием формальных критериев оценки успешности решения поставленной задачи данной конкретной сетью. Имеется также ряд технических вопросов (размер сети, ее конфигурация, выбор параметров пре- и постпроцессирования и др.), не имеющих на данный момент общепринятых способов решения, однако сильно влияющих на конечный результат.

Предпринятые в процессе настоящей работы попытки найти способ выбора единственной лучшей сети Кохонена не дали удовлетворившего нас результата. Полезной находкой, использованной в дальнейшем, оказался критерий оценки качества сети, основанный на стабильности результатов осуществляемой с ее помощью кластеризации в нескольких случайных подвыборках. Работа сети признавалась удовлетворительной, если величины характеризующих ФР и исходы показателей в дан-

ном выделенном с ее использованием кластере значимо не различались в нескольких случайных подвыборках. Однако описанная техника не позволяет однозначно выбрать единственную наилучшую сеть.

Предлагаемое нами решение заключается в следующем. Не пытаясь найти единственную наилучшую сеть, для каждой задачи обучали большое количество подходящих по размерам сетей Кохонена. Из их числа отбирали несколько (в настоящей работе порядка двух десятков) перспективных сетей, учитывая известные эмпирические правила построения ИНС [12] и описанный выше критерий, основанный на стабильности результатов кластеризации. Дальнейшие действия производили в два этапа.

На первом этапе работали с каждой отобранной сетью по отдельности. Вначале определяли номера кластеров, соответствующих ячейкам таблицы риска, образец которой представлен на цветном рисунке 1. Для этого через сеть прогоняли набор условных наблюдений, в которых величины учитываемых в таблице ФР соответствуют «центрам» ячеек таблицы. Например, для представленной на цветном рисунке 1 таблицы риска через сеть необходимо прогнать 96 условных записей. В первой записи значения ФР: пол – женский, САД 120 мм рт.ст, ЧСС 70 уд./мин, возраст 40 лет, статус курения – некурящие, образовательный уровень – среднее или ниже; во второй записи: пол – женский, САД 120 мм рт.ст, ЧСС 90 уд./мин, возраст 40 лет, статус курения – некурящие, образовательный уровень – среднее или ниже; и т.д. до последней записи: пол – женский, САД 180 мм рт.ст, ЧСС 90 уд./мин, возраст 60 лет, статус курения – курящие, образовательный уровень – выше среднего. В результате каждую условную запись сеть отнесет к одному из кластеров и всем ячейкам таблицы риска будут присвоены номера соответствующих кластеров.

Затем через сеть прогоняли весь набор реальных данных, каждое наблюдение сеть относил к одному из кластеров, после чего для кластера рассчитывали риск фатального события (к примеру, смерти от ССЗ) для нужного периода наблюдения

(например 10-летнего), этот риск соотносили с ячейками таблицы (типа представленной на цветном рисунке 1), соответствующих данному кластеру. В итоге каждой ячейке таблицы соответствует рассчитанный с использованием данной сети риск, выраженный как отношение числа умерших в течение указанного периода времени (числитель) к числу подверженных риску (знаменатель).

Описанную процедуру повторяли с каждой из отобранных сетей для сердечно-сосудистой и общей смертности и периодов наблюдения 10 и 20 лет.

На втором этапе производили усреднение результатов для каждой ячейки таблицы риска. Для этого суммировали для данной ячейки величины числителей и знаменателей, рассчитанные для использованных на первом этапе сетей, после чего посредством операции деления суммы величин числителей на сумму величин знаменателей вычисляли риск.

Такая техника имеет некоторое сходство с принципом «голосования» сетей, применение которого описано при решении задач прогноза или классификации искусственными сетями других типов, использующих обучение с учителем [15].

Описанный подход позволил обойти принципиальную трудность, связанную с выбором единственной лучшей сети, а также избежать чрезмерных колебаний величин риска в близких ячейках таблицы. Имеются и другие плюсы, связанные со спецификой нейросетевых методов, в частности становится возможным сочетать сети разных конфигураций (с разным числом нейронов и др. характеристиками), каждая из которых привносит в конечный результат свои преимущества. Увеличивая количество используемых на первом этапе сетей, можно добиться требуемого «сглаживания» величин риска при перемещении по таблице.

При построении каждой таблицы (цветные рисунки 1-16) использовали от 16 до 20 сетей, следовательно, выводы данной работы опираются на результаты кластеризации с использованием порядка 300 ИНС. С учетом предварительного анализа общее количество косвенно участвовавших в получении ко-

нечных результатов ИНС исчисляется тысячами.

Недостатком нашего способа является снижение надежности расчета риска для ячеек таблиц, которым соответствует малое число реальных наблюдений в когорте. Однако следует подчеркнуть, что эта проблема в неявном виде присутствует и при использовании основанного на моделировании традиционного подхода. Малое число лиц с данным сочетанием значений ФР в выборке при любом способе оценки риска обусловит ненадежный и нестабильный результат. В зависимости от техники расчета риска это может проявляться явно (при нашем способе) или маскироваться за счет усреднения, интер- и экстраполяции (при регрессионных методах).

В настоящей работе использовали следующее решение – в ячейках таблиц, которым соответствует менее 10 наблюдений в исследованной когорте, рассчитанные посредством кластерной техники величины риска были заменены на полученные с использованием соответствующей модели регрессии (см. ниже раздел 3). Величины риска в таких ячейках таблиц (цветные рис. 1-16) маркированы синим цветом.

Завершая посвященный статистической технике раздел, следует отметить, что при обучении ИНС такие показатели, как САД, ЧСС и возраст, были представлены непрерывными количественными переменными. Ступенчатые градации этих показателей использованы только в таблицах риска для лучшей визуализации. Например, возраст 40 лет в таблице риска соответствует значениям 35-44 в реальных наблюдениях, ЧСС 90 уд./мин в таблице риска соответствует значениям более 80 уд./мин, САД 180 мм рт.ст. в таблице риска соответствует значениям более 170 мм рт.ст. и т.д.

### 3. ТАБЛИЦЫ РИСКА НА ОСНОВЕ ТРАДИЦИОННЫХ РЕГРЕССИОННЫХ МОДЕЛЕЙ

*А.Д. Деев, Ф.А. Айду, В.Г.Вилков*

В известных способах прогнозирования риска используются многомерные регрессионные модели [3-4, 6-7, и др.]. Преимуществами такого подхода являются наличие хорошо разработанных статистических методов, относительная простота математического аппарата, наглядность, непротиворечивость и удобство интерпретации результатов. Поскольку модель базируется на всей совокупности наблюдений, число которых в популяционных исследованиях весьма велико, обеспечивается высокая статистическая достоверность. Однако это достигается за счет минимизации ошибок для наблюдений с наиболее часто встречающимися сочетаниями величин ФР, тогда как для относительно редко встречающихся наблюдений ошибки прогноза могут быть большими. Принципиальные недостатки данного подхода обусловлены налагаемыми линейным моделированием ограничениями, влекущими, в частности, утрату нелинейных составляющих связей между переменными.

Одним из прототипов предлагаемого нами способа является российская шкала 10-летнего риска смерти от ССЗ на базе регрессионной модели SCORE (Systematic Coronary Risk Evaluation), построенная с использованием той же когорты [7]. Поскольку в нашей работе набор ФР несколько отличается от таблиц SCORE – учтена ЧСС и отсутствует концентрация общего холестерина в крови, для сравнения нового и традиционного подходов с применением последнего были построены регрессионные модели и на их основе рассчитаны таблицы 10- и 20-летних сердечно-сосудистых и общих рисков для когорт России и США, которые считали ближайшим прототипом предлагаемого нами способа (цветные рис. 17-32).

Как и следовало ожидать, по данным регрессионных моделей связи риска с возрастом, полом, статусом курения, уровнем образования и величинами САД существенно не отличаются от известных [4, 7]:

- у мужчин риск выше по сравнению с женщинами;
- с возрастом риск повышается;
- у курящих риск выше по сравнению с некурящими;
- при более высоком уровне образования риск ниже;
- риск увеличивается по мере повышения САД.

Описанные закономерности наблюдаются для 10- и 20-летних сердечно-сосудистых и общих рисков.

Связь ЧСС с суммарным риском по данным регрессионных моделей однозначная – при тахикардии риск выше. В подгруппах с очень низким риском этот эффект может маскироваться вследствие округления риска до целых величин (цветные рис. 17-32).

Для российской и американской когорт регрессионные модели строились с использованием одинаковой статистической техники и единым набором переменных, это позволило сравнить величины рисков. Из цветных рисунков 17-32 видно, что при одинаковых уровнях учтенных ФР:

- у российских женщин суммарные риски равны или несколько выше в сравнении с американской когортой;
- у мужчин при той же направленности различия выражены сильнее – в российской когорте риски намного выше в сравнении с американской.

## 4. ТОЧНОСТЬ РАСЧЕТА РИСКА НОВЫМ СПОСОБОМ В СРАВНЕНИИ С ТРАДИЦИОННЫМ

*В.Г. Вилков*

При сопоставлении цветных рисунков 1-16 и 17-32 видно, что в части случаев рассчитанные предлагаемым нами способом риски сильно отличаются от полученных с использованием моделей регрессии. Возникает вопрос о точности расчета риска указанными способами.

Подобная проверка может быть осуществлена посредством прямого расчета риска для одной из ячеек используемых нами таблиц риска в случаях, когда число соответствующих данной ячейке наблюдений в когорте достаточно велико.

Для проверки требуется отобрать ячейки таблиц с большим числом наблюдений, желательно для разных сочетаний величин ФР.

На рисунках 33 и 34 представлены данные о числе наблюдений для каждой из ячеек таблиц риска в российской и американской когортах, штриховкой выделены ячейки, отобранные для проверки точности расчета риска.

В таблицах 4-15 представлены результаты расчета рисков фатальных событий тремя разными способами:

- новым, основанным на использовании СОК ФР;
- традиционным, использующим регрессионные модели;
- эталонным, когда риск рассчитывался напрямую по фактическим данным.

Такие расчеты проведены для каждой из 48 ячеек таблиц риска, выделенных серой штриховкой на рисунках 33-34 (общее число наблюдений 3517 человек). При выборе ячеек стремились, чтобы в анализе были представлены разные сочетания уровней ФР, обязательным условием было большое количество наблюдений в каждой ячейке.

Из таблицы 4 видно, что величины рисков, рассчитанные всеми способами, практически одинаковы. Этого можно было ожидать исходя из следующих соображений:

- В российской когорте на долю мужчин без явного повышения АД (САД 120-140 мм рт.ст.) и тахикардии (ЧСС 70 уд./мин) приходится большая часть всех имеющихся наблюдений (рис. 33).
- Многомерная регрессионная модель строится таким образом, чтобы суммарная ошибка по всей совокупности наблюдений была минимальной. Это достигается подгонкой модели, при которой наиболее часто встречающиеся наблюдения будут иметь минимальные ошибки.
- При кластерной технике велика вероятность, что относящиеся к такой ячейке наблюдения составят значительную долю всех наблюдений данного кластера.
- Поэтому рассчитанный кластерным и регрессионным способами риск должен быть близок к эталонному, что и наблюдается в действительности.

Иная картина наблюдается для подгруппы курящих 60-летних мужчин российской когорты с САД 160 мм рт.ст., без тахикардии и образованием средним или ниже (см. соответствующие ячейки таблиц на цветных рис. 2, 4, 6, 8, 18, 20, 22, 24). Из таблицы 5 видно, что рассчитанные нашим способом величины рисков в 3 из 4 случаев значительно ближе к эталонным в сравнении с регрессионным прототипом. Для 20-летнего риска смерти от всех причин отклонения результатов обоих способов от эталонной величины приблизительно одинаковые, в остальных 3 случаях регрессионные модели сильно завышают риск (для 10-летнего сердечно-сосудистого риска – почти в 2 раза). Рассматриваемой ячейке таблицы риска и соседним с нею соответствуют относительно небольшие числа наблюдений (рис. 33Б). В этих условиях наглядно проявляются преимущества основанного на СОК ФР способа в сравнении с традиционным.

В таблицах 6-15 приведены результаты еще нескольких подобных сравнений.



| А              |           |    |         |    |               |           |    |         |    |
|----------------|-----------|----|---------|----|---------------|-----------|----|---------|----|
| Образование    |           |    |         |    |               |           |    |         |    |
| среднее и ниже |           |    |         |    | выше среднего |           |    |         |    |
| САД            | некурящие |    | курящие |    | возраст       | некурящие |    | курящие |    |
| 180            | 336       | 76 | 7       | 1  | 60            | 97        | 14 | 7       | 1  |
| 160            | 263       | 60 | 19      | 2  |               | 81        | 14 | 10      | 0  |
| 140            | 224       | 50 | 22      | 2  |               | 89        | 20 | 9       | 1  |
| 120            | 127       | 9  | 16      | 2  |               | 71        | 9  | 7       | 0  |
| 180            | 103       | 26 | 8       | 3  | 50            | 35        | 11 | 6       | 0  |
| 160            | 132       | 21 | 12      | 4  |               | 41        | 10 | 4       | 2  |
| 140            | 200       | 32 | 14      | 2  |               | 120       | 20 | 11      | 2  |
| 120            | 236       | 21 | 32      | 2  |               | 191       | 26 | 23      | 5  |
| 180            | 22        | 7  | 1       | 1  | 40            | 7         | 3  | 2       | 2  |
| 160            | 46        | 15 | 7       | 1  |               | 28        | 13 | 6       | 0  |
| 140            | 136       | 36 | 32      | 5  |               | 89        | 17 | 23      | 8  |
| 120            | 323       | 51 | 80      | 9  |               | 371       | 63 | 93      | 13 |
|                | 70        | 90 | 70      | 90 | ЧСС           | 70        | 90 | 70      | 90 |

| Б              |           |    |         |     |               |           |    |         |    |
|----------------|-----------|----|---------|-----|---------------|-----------|----|---------|----|
| Образование    |           |    |         |     |               |           |    |         |    |
| среднее и ниже |           |    |         |     | выше среднего |           |    |         |    |
| САД            | некурящие |    | курящие |     | возраст       | некурящие |    | курящие |    |
| 180            | 68        | 12 | 41      | 21  | 60            | 36        | 9  | 16      | 1  |
| 160            | 104       | 14 | 85      | 18  |               | 51        | 11 | 24      | 5  |
| 140            | 124       | 19 | 132     | 28  |               | 125       | 11 | 57      | 9  |
| 120            | 104       | 9  | 117     | 20  |               | 135       | 15 | 69      | 6  |
| 180            | 86        | 29 | 118     | 45  | 50            | 45        | 16 | 33      | 8  |
| 160            | 142       | 30 | 255     | 89  |               | 100       | 21 | 55      | 20 |
| 140            | 350       | 47 | 492     | 112 |               | 316       | 38 | 164     | 27 |
| 120            | 357       | 36 | 564     | 85  |               | 500       | 31 | 309     | 28 |
| 180            | 18        | 10 | 30      | 17  | 40            | 12        | 1  | 10      | 2  |
| 160            | 55        | 24 | 106     | 33  |               | 31        | 8  | 27      | 10 |
| 140            | 148       | 23 | 298     | 62  |               | 182       | 29 | 130     | 28 |
| 120            | 293       | 24 | 538     | 71  |               | 467       | 37 | 357     | 40 |
|                | 70        | 90 | 70      | 90  | ЧСС           | 70        | 90 | 70      | 90 |

Рис. 33. Число наблюдений в ячейках таблиц риска в зависимости от возраста, уровней САД (мм рт.ст.) и ЧСС (уд./мин), статуса курения и уровня образования (Россия).

А – женщины, Б – мужчины.

Серой штриховкой маркированы ячейки таблицы, в которых проведен прямой расчет риска.

| А              |           |     |         |     |               |           |    |         |    |  |
|----------------|-----------|-----|---------|-----|---------------|-----------|----|---------|----|--|
| Образование    |           |     |         |     |               |           |    |         |    |  |
| среднее и ниже |           |     |         |     | выше среднего |           |    |         |    |  |
| САД            | некурящие |     | курящие |     | возраст       | некурящие |    | курящие |    |  |
| 180            | 195       | 181 | 41      | 33  | 60            | 24        | 26 | 2       | 4  |  |
| 160            | 263       | 249 | 61      | 54  |               | 60        | 52 | 10      | 8  |  |
| 140            | 452       | 333 | 104     | 109 |               | 116       | 74 | 19      | 22 |  |
| 120            | 489       | 260 | 147     | 111 |               | 131       | 76 | 35      | 29 |  |
| 180            | 25        | 16  | 11      | 18  | 50            | 2         | 2  | 0       | 3  |  |
| 160            | 62        | 62  | 23      | 39  |               | 9         | 14 | 3       | 2  |  |
| 140            | 145       | 94  | 41      | 57  |               | 34        | 26 | 9       | 9  |  |
| 120            | 279       | 138 | 160     | 116 |               | 89        | 35 | 44      | 35 |  |
| 180            | 3         | 14  | 5       | 6   | 40            | 2         | 0  | 0       | 1  |  |
| 160            | 13        | 16  | 11      | 9   |               | 8         | 0  | 1       | 5  |  |
| 140            | 57        | 78  | 34      | 56  |               | 16        | 24 | 8       | 10 |  |
| 120            | 276       | 221 | 201     | 189 |               | 148       | 90 | 50      | 50 |  |
|                | 70        | 90  | 70      | 90  | ЧСС           | 70        | 90 | 70      | 90 |  |

| Б              |           |     |         |     |               |           |    |         |    |  |
|----------------|-----------|-----|---------|-----|---------------|-----------|----|---------|----|--|
| Образование    |           |     |         |     |               |           |    |         |    |  |
| среднее и ниже |           |     |         |     | выше среднего |           |    |         |    |  |
| САД            | некурящие |     | курящие |     | возраст       | некурящие |    | курящие |    |  |
| 180            | 95        | 61  | 71      | 62  | 60            | 16        | 14 | 8       | 5  |  |
| 160            | 191       | 128 | 120     | 94  |               | 52        | 22 | 22      | 17 |  |
| 140            | 361       | 167 | 260     | 162 |               | 118       | 51 | 55      | 27 |  |
| 120            | 373       | 156 | 326     | 173 |               | 112       | 39 | 56      | 27 |  |
| 180            | 9         | 12  | 11      | 18  | 50            | 1         | 2  | 1       | 2  |  |
| 160            | 35        | 22  | 36      | 29  |               | 11        | 9  | 10      | 10 |  |
| 140            | 88        | 57  | 99      | 90  |               | 25        | 23 | 36      | 25 |  |
| 120            | 137       | 68  | 221     | 101 |               | 85        | 26 | 83      | 30 |  |
| 180            | 1         | 1   | 5       | 4   | 40            | 0         | 0  | 3       | 1  |  |
| 160            | 8         | 16  | 13      | 19  |               | 2         | 7  | 3       | 6  |  |
| 140            | 55        | 39  | 85      | 52  |               | 35        | 24 | 38      | 33 |  |
| 120            | 177       | 46  | 215     | 132 |               | 129       | 40 | 99      | 49 |  |
|                | 70        | 90  | 70      | 90  | ЧСС           | 70        | 90 | 70      | 90 |  |

Рис. 34. Число наблюдений в ячейках таблицы риска в зависимости от возраста, уровней САД (мм рт.ст.) и ЧСС (уд./мин), статуса курения и уровня образования (США).

А – женщины, Б – мужчины.

Серой штриховкой маркированы ячейки таблицы, в которых проведен прямой расчет риска.

Таблица 4

**Риск фатального события (%) для российских мужчин 40-летнего возраста, курящих, с образованием средним или ниже, САД 120 мм рт.ст., без тахикардии (n = 538)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 3                    | 3         | 2,8    |
| ССЗ, 20 лет                        | 10                   | 11        | 10,2   |
| Все причины, 10 лет                | 9                    | 9         | 10,4   |
| Все причины, 20 лет                | 31                   | 27        | 32,7   |

Таблица 5

**Риск фатального события (%) для российских мужчин 60-летнего возраста, курящих, с образованием средним или ниже, САД 160 мм рт.ст., без тахикардии (n = 85)**

| Причина смерти,<br>Срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 16                   | 28        | 15,3   |
| ССЗ, 20 лет                        | 39                   | 72        | 43,5   |
| Все причины, 10 лет                | 35                   | 46        | 37,7   |
| Все причины, 20 лет                | 89                   | 76        | 82,4   |

Таблица 6

**Риск фатального события (%) для американских мужчин 60-летнего возраста, некурящих, с образованием средним или ниже, САД 140 мм рт.ст., без тахикардии (n = 361)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 16                   | 9         | 15,8   |
| ССЗ, 20 лет                        | 30                   | 27        | 27,2   |
| Все причины, 10 лет                | 28                   | 16        | 29,6   |
| Все причины, 20 лет                | 55                   | 44        | 52,1   |

Таблица 7

**Риск фатального события (%) для российских женщин 60-летнего возраста, некурящих, с образованием средним или ниже, САД 180 мм рт.ст., без тахикардии (n = 336)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 20                   | 9         | 18,2   |
| ССЗ, 20 лет                        | 38                   | 31        | 34,2   |
| Все причины, 10 лет                | 32                   | 17        | 27,7   |
| Все причины, 20 лет                | 67                   | 54        | 60,1   |

Таблица 8

**Риск фатального события (%) для американских женщин 60-летнего возраста, некурящих, с образованием средним или ниже, САД 180 мм рт.ст., без тахикардии (n = 195)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 15                   | 7         | 13,9   |
| ССЗ, 20 лет                        | 29                   | 24        | 24,6   |
| Все причины, 10 лет                | 24                   | 12        | 22,1   |
| Все причины, 20 лет                | 49                   | 38        | 41,5   |

Таблица 9

**Риск фатального события (%) для американских женщин 60-летнего возраста, некурящих, с образованием средним или ниже, САД 180 мм рт.ст., с тахикардией (n = 181)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 19                   | 8         | 21,0   |
| ССЗ, 20 лет                        | 35                   | 27        | 34,8   |
| Все причины, 10 лет                | 32                   | 14        | 32,6   |
| Все причины, 20 лет                | 57                   | 44        | 55,3   |

Таблица 10

**Риск фатального события (%) для американских женщин 60-летнего возраста, некурящих, с образованием средним или ниже, САД 160 мм рт.ст., с тахикардией (n = 249)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 11                   | 6         | 11,7   |
| ССЗ, 20 лет                        | 25                   | 20        | 24,9   |
| Все причины, 10 лет                | 21                   | 12        | 22,1   |
| Все причины, 20 лет                | 45                   | 38        | 39,8   |

Таблица 11

**Риск фатального события (%) для американских женщин 60-летнего возраста, некурящих, с образованием средним или ниже, САД 160 мм рт.ст., без тахикардии (n = 263)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 9                    | 5         | 9,9    |
| ССЗ, 20 лет                        | 22                   | 18        | 20,9   |
| Все причины, 10 лет                | 16                   | 10        | 18,3   |
| Все причины, 20 лет                | 41                   | 33        | 37,3   |

Таблица 12

**Риск фатального события (%) для американских женщин 60-летнего возраста, некурящих, с образованием средним или ниже, САД 140 мм рт.ст., без тахикардии (n = 452)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 7                    | 4         | 6,0    |
| ССЗ, 20 лет                        | 17                   | 14        | 16,2   |
| Все причины, 10 лет                | 15                   | 8         | 12,4   |
| Все причины, 20 лет                | 36                   | 28        | 31,2   |

Таблица 13

**Риск фатального события (%) для американских женщин 60-летнего возраста, некурящих, с образованием средним или ниже, САД 140 мм рт.ст., с тахикардией (n = 333)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 9                    | 4         | 8,1    |
| ССЗ, 20 лет                        | 21                   | 15        | 15,6   |
| Все причины, 10 лет                | 19                   | 10        | 17,4   |
| Все причины, 20 лет                | 42                   | 33        | 34,2   |

Таблица 14

**Риск фатального события (%) для американских женщин 60-летнего возраста, некурящих, с образованием средним или ниже, САД 120 мм рт.ст., с тахикардией (n = 260)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 8                    | 3         | 6,9    |
| ССЗ, 20 лет                        | 19                   | 12        | 13,1   |
| Все причины, 10 лет                | 18                   | 8         | 15,4   |
| Все причины, 20 лет                | 39                   | 28        | 27,7   |

Таблица 15

**Риск фатального события (%) для американских женщин 60-летнего возраста, некурящих, с образованием средним или ниже, САД 120 мм рт.ст., без тахикардии (n = 489)**

| Причина смерти,<br>срок наблюдения | Способ расчета риска |           |        |
|------------------------------------|----------------------|-----------|--------|
|                                    | СОК ФР               | регрессия | эталон |
| ССЗ, 10 лет                        | 6                    | 3         | 5,3    |
| ССЗ, 20 лет                        | 15                   | 10        | 10,6   |
| Все причины, 10 лет                | 15                   | 7         | 14,7   |
| Все причины, 20 лет                | 34                   | 24        | 26,2   |

При сравнении точности нового и традиционного способов учитывали:

1) Разности рисков, рассчитанных соответствующим способом и эталонным. Из представленных на рисунке 35 гистограмм наглядно видны преимущества нового способа – отклонения от эталонных величин меньше, нет такой выраженной систематической ошибки, как при регрессионном способе.

2) Абсолютные величины рассмотренных в предыдущем пункте разностей. Для нашего способа средняя для всех 48 сравнений величина составила  $2,8 \pm 2,6\%$  ( $M \pm SD$ ) при максимальной величине 11,3%, для традиционного способа отклонения в 2 раза больше и равны  $6,0 \pm 5,3\%$  и 28,5% соответственно.

3) Из 48 сравнений в 30 случаях величины рисков, рассчитанных с использованием СОК ФР, меньше отличались от эталонных, чем рассчитанные на основе регрессионных моделей, в 8 случаях оба способа продемонстрировали приблизительно одинаковую точность, в 10 случаях основанный на СОК ФР способ уступал традиционному.

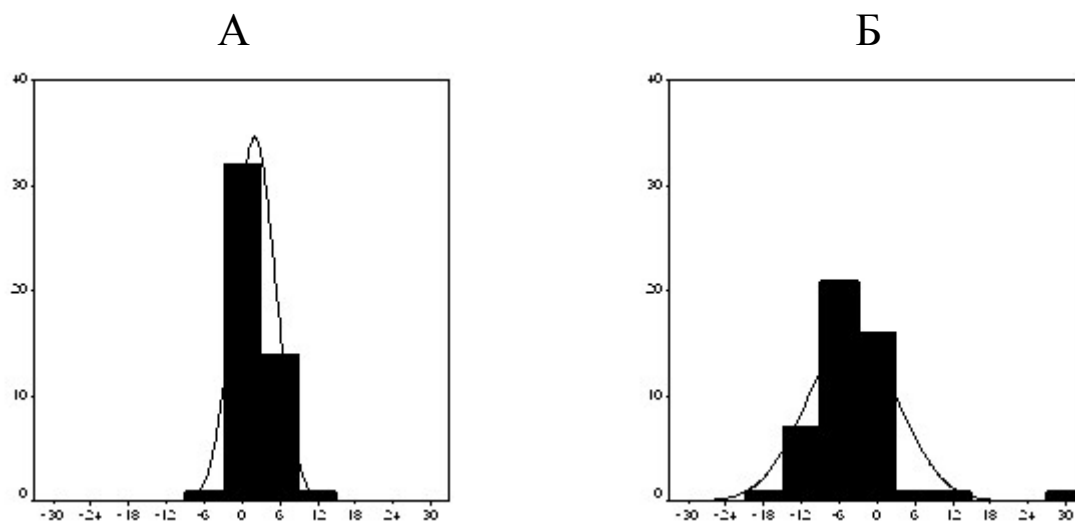


Рис. 35. Гистограммы ошибок рассчитанных новым и традиционным способами рисков.

По оси абсцисс – величины разности рассчитанных с использованием СОК ФР (А) или регрессии (Б) рисков и эталонных величин.

Приведенные данные свидетельствуют о том, что разработанный нами способ в большинстве случаев обеспечивает более высокую точность расчета рисков в сравнении с традиционным.

Следует отметить, что описанная выше проверка ставит сравниваемые способы в неравные условия – они более благоприятны для регрессионных методов, поскольку для них характерна подгонка модели с минимизацией ошибок для наиболее часто встречающихся наблюдений. Описанная проверка может применяться именно для таких подгрупп – с достаточно большим числом наблюдений. То, что новый способ продемонстрировал более высокую точность в таких условиях, является убедительным доводом в его пользу.



## **5. ТАБЛИЦЫ РИСКА НА ОСНОВЕ САМООРГАНИЗУЮЩИХСЯ КЛАСТЕРОВ ФАКТОРОВ РИСКА ДЛЯ РОССИЙСКОЙ КОГОРТЫ**

*В.Г. Вилков, А.Д. Деев, С.А. Шальнова*

На цветных рисунках 1-8 представлены таблицы рисков обусловленных ССЗ и всеми причинами фатальных событий в течение 10 и 20 ближайших лет для российской когорты. На цветных рисунках 17-24 приведены прототипы – таблицы тех же рисков, построенные с использованием традиционных для анализа дожития процедур регрессии.

Относительно роли отдельных ФР на основании представленных на цветных рисунках 1-8 результатов можно отметить следующее.

### *Пол*

Как упоминалось в разделе 3, по данным способа-прототипа смертность от ССЗ и всех причин у мужчин во всех случаях выше в сравнении с женщинами. Наш способ демонстрирует аналогичную закономерность, но при той же направленности различий их выраженность в разных подгруппах варьирует заметно сильнее, чем по данным регрессионных моделей (цветные рис. 1-8 в сравнении с рис. 17-24).

В частности, в российской когорте у части лиц с высоким риском (60-летние некурящие с образованием средним или ниже) риск у мужчин отличается от такового у женщин заметно меньше, чем можно было предполагать по данным регрессионных методов.

### *Возраст*

В отношении влияния возраста полученные нашим способом результаты (цветные рис. 1-8) в принципе согласуются с прототипом (цветные рис. 17-24) – при равных величинах прочих ФР с возрастом риск увеличивается.

Однако по данным нашего способа у лиц с максимальными рисками (60-летние курящие мужчины со средним или ниже

образованием) прирост риска в сравнении с 50-летними заметно меньше, чем предполагалось на основании регрессионных моделей. Следует отметить, что в данном случае было проведено сравнение точности расчета риска нашим и традиционным способами посредством сопоставления с прямым расчетом по фактическим данным (раздел 4, табл. 5), которое продемонстрировало убедительное преимущество основанного на СОК ФР способа: рассчитанные им величины риска отклонялись от эталонных в пределах  $-5 \div +7\%$ , тогда как для регрессии отклонения составили  $-6 \div +28\%$ .

### *Курение*

По таблицам SCORE, построенным как по данным российских исследований [7], так и для европейских популяций с низким и высоким риском [4], курение оказывает одинаковое влияние на 10-летний фатальный сердечно-сосудистый риск – у курящих в сравнении с некурящими риск выше приблизительно в 2 раза.

Построенные традиционным способом для использованного нами набора переменных таблицы (цветные рис. 17-24) демонстрируют сходные результаты как для сердечно-сосудистого, так и общего рисков.

Из разработанных на основе СОК ФР таблиц (цветные рис. 1-8) видно, что в большинстве случаев риск у курящих выше, причем это касается как 10-летних, так и 20-летних не только сердечно-сосудистых, но и общих рисков. Однако при использовании нашего способа роль курения выглядит не такой однозначной, как в основанных на регрессионных моделях таблицах.

В построенных нами таблицах имеются отдельные области, в которых риски у курящих и некурящих в некоторых случаях близки по величине с небольшими колебаниями в ту или иную сторону:

- при очень низких суммарных абсолютных рисках (0-2%);
- у 60-летних женщин с образованием средним или ниже и нормальной ЧСС;

- 10-летний сердечно-сосудистый риск у женщин с САД 180 мм рт.ст. при образовании среднем или ниже;
- сердечно-сосудистый риск у 60-летних мужчин с образованием средним или ниже – в единичных случаях, преимущественно при тахикардии.

Рассмотрим данную ситуацию подробнее на примере подвыборки мужчин старше 54 лет, с САД более 150 мм рт.ст, ЧСС не более 80 уд./мин и образованием выше среднего. Для этой категории лиц сердечно-сосудистые риски у курящих и некурящих близки по величине, у последних иногда даже несколько выше (цветные рис. 2, 4). В подгруппах некурящих ( $n=87$ ) и курящих ( $n=40$ ) были построены кривые выживания Каплана-Мейера (рис. 36), из которых видно, что у курящих мужчин уровень дожития визуально несколько ниже при всех сроках наблюдения, однако эти различия статистически незначимы ( $p=0,3$ ). Таким образом, в рассмотренном случае речь идет о наличии ожидаемых различий дожития в соответствующих данным ячейкам таблицы подвыборках курящих и некурящих мужчин с перечисленными выше значениями остальных ФР. Различия величин риска по кривым Каплана-Мейера и таблицам невелики (единицы процента) и обусловлены техникой расчета – при кластеризации на приписанное данной ячейке таблицы значение риска могут отчасти влиять наблюдения, относящиеся не только к этой, но и соседним ячейкам.

В других случаях построенные нашим способом таблицы (цветные рис. 1-8) выявляют более выраженную негативную роль курения по сравнению с традиционным способом. Так, рассчитанный нашим способом риск у курящих значительно больше, чем в 2 раза, превышает риск у некурящих:

- у 40- и 50-летних женщин с образованием средним или ниже, нормальными или умеренно повышенными величинами САД и нормальной ЧСС;
- у 40-летних мужчин с образованием выше среднего, тахикардией и САД менее 150 мм рт.ст. (10-летний общий, 10- и 20-летний сердечно-сосудистый риски).

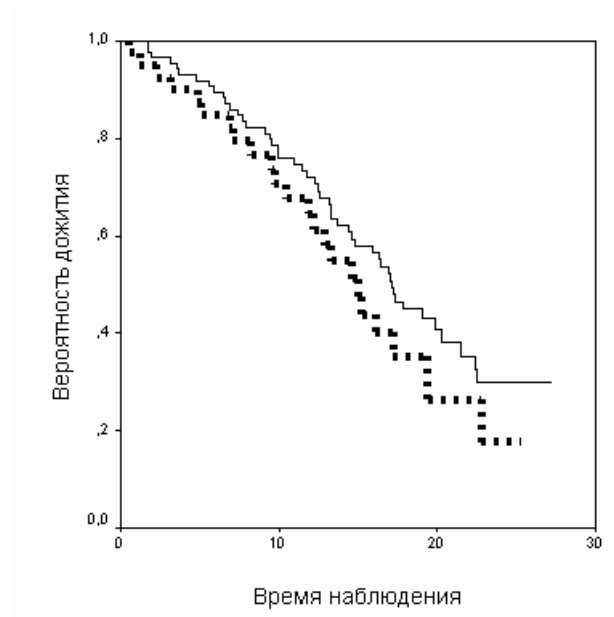


Рис. 36. Кривые выживания для смерти от ССЗ у некурящих (сплошная тонкая линия) и курящих (точечная линия с темными маркерами) мужчин старше 54 лет с образованием выше среднего, уровнем САД 150 и более мм рт.ст. и ЧСС не более 80 уд./мин.

По оси абсцисс – время от момента первичного обследования до завершения наблюдения (годы); по оси ординат – кумулятивная доля выживших (доли единицы).

Таким образом, в отношении курения результаты традиционного и нашего способов отчасти различаются.

По данным основанных на системе SCORE таблиц (как для российской, так и для европейских популяций с различным уровнем риска) 10-летний сердечно-сосудистый риск в 2 раза выше ( $\pm 1\%$ ) у курящих по сравнению с некурящими вне зависимости от пола, возраста, уровня образования, величин САД, концентрации общего холестерина в плазме крови и уровня суммарного риска [7, 4]. Построенные нами на основе регрессионных моделей прототипы (цветные рис. 17-24) демонстрируют сходные результаты как для сердечно-сосудистых, так и общих 10- и 20-летних рисков – курение ассоциируется с повышением смертности в 1,5-2 раза. Такое единообразие во всем диапазоне значений всех учтенных ФР выглядит маловероятным и вряд ли одинаково точно характеризует истинную роль курения у разных категорий лиц.

Вероятно, это обусловлено техническими причинами, связанными с использованием линейного моделирования. Более корректным будет утверждение, что для большинства включенных в когорту лиц курение при 20-летнем наблюдении в среднем в 1,5-2 раза повышает сердечно-сосудистый и общий риски. Этого в принципе достаточно для общего вывода о важной и безусловно негативной роли этого ФР, и соответствующих практических рекомендаций, ориентированных на популяцию в целом.

Наш способ позволяет дать более дифференцированную оценку.

Выявлены отдельные категории лиц, у которых значимость курения особенно велика. Это женщины 35-55 лет без высшего образования с нормальным или умеренно повышенным САД и нормальной ЧСС, а также мужчины 35-45 лет с высшим образованием, тахикардией и без явного повышения САД. В указанных подгруппах риск у курящих более чем в 2 (до 7) раз выше в сравнении с некурящими.

В ряде подгрупп полученные нашим способом результаты близки к прототипам – курению сопутствует увеличение риска в 1,5-2 раза, например у мужчин среднего возраста с нормальной ЧСС, без высшего образования и явного повышения САД.

Имеются также подгруппы, в которых по нашим данным статус курения не оказывает существенного влияния на риск, один из таких примеров рассмотрен выше (рис. 36). В качестве общей закономерности можно отметить более слабую связь суммарного риска с курением при высоком уровне САД и в возрасте старше 55 лет.

### *Образование*

По данным основанных на системе SCORE таблиц для российской когорты зависимость 10-летнего сердечно-сосудистого риска от уровня образования выглядит следующим образом. Если условно принять риск у лиц с образованием выше среднего за 100%, то при равных значениях прочих ФР риск у лиц со средним образованием выше на 10-20% у мужчин и в среднем

на 60% ( $\pm 10\%$ ) у женщин, а при образовании ниже среднего риск повышается еще приблизительно на 20% [7].

Построенные для использованного нами набора переменных регрессионные модели (цветные рис. 17-24) демонстрируют сходные результаты – у лиц с образованием средним или ниже риск выше в сравнении с аналогичными подгруппами с образованием выше среднего приблизительно на 30-100% у женщин, у мужчин при той же направленности различия выражены слабее, чаще в пределах 20-50%.

Такая однонаправленная и постоянная во всем диапазоне значений учтенных ФР зависимость напоминает рассмотренную выше ситуацию с ролью курения. Можно полагать, что и причина та же, обусловленная техническими особенностями использования линейного моделирования для расчета влияния номинативной переменной на суммарный риск. Все упомянутые при обсуждении связи с курением соображения о положительных и отрицательных сторонах данного подхода применимы и к уровню образования.

В построенных нашим способом таблицах (цветные рис. 1-8) влияние уровня образования на риск имеет ту же направленность (чем выше образовательный уровень, тем ниже риск фатальных событий), однако влияние данного фактора не так однородно, как при использовании основанных на регрессионных моделях таблиц. В части случаев риск мало зависит от образования (например, 10-летний сердечно-сосудистый риск у некурящих мужчин с нормальной ЧСС), в других при более низком уровне образования наблюдается значительное увеличение риска (например, у некурящих женщин 50-60 летнего возраста с тахикардией).

### *Систолическое АД*

По данным обоих способов – традиционного и нашего, этот ФР оказывает сильное и однозначное влияние на смертность – по мере повышения САД риск увеличивается. Как и следовало ожидать, зависимость вычисленных по регрессионным моделям рисков от САД практически линейная.

Результаты нашего способа в принципе схожи с прототипом, описанная зависимость наблюдается как для общей, так и сердечно-сосудистой смертности, у женщин и у мужчин (цветные рис. 1-8). Однако применение нелинейных методов позволило выявить ряд особенностей.

Влияние САД слабо выражено при низком суммарном риске (в пределах 2-3%, например 10-летний риск ССЗ у 40-летних некурящих женщин, цветной рис. 1).

С другой стороны, в ряде случаев обращает внимание более выраженный в сравнении с прототипом прирост риска при высоком САД (180, иногда 160-180 мм рт.ст.) по отношению к предыдущим градациям САД. Например, при изменении САД от 160 до 180 мм рт.ст. 10-летний сердечно-сосудистый риск у некурящих женщин со средним или ниже образованием увеличился с 9 до 20% в возрасте 60 лет и с 3 до 7% в возрасте 50 лет (цветной рис. 1). По таблице-прототипу в соответствующих категориях лиц прирост риска не превышает 1-2% (цветной рис. 17). Сходная картина наблюдается при использовании таблицы SCORE – например, у некурящих женщин с образованием ниже среднего и концентрацией общего холестерина в сыворотке крови 6 ммоль/л прирост риска составляет 1-3% [7].

Выраженный прирост риска при повышении САД до 160-180 мм рт.ст. обнаруживается преимущественно при умеренных величинах рисков у некурящих с нормальной ЧСС.

Различия прироста риска по мере повышения САД в зависимости от способа построения таблиц риска проиллюстрированы на примере 10-летних сердечно-сосудистых рисков у женщин (рис. 37). На рис. 37А представлена зависимость рассчитанного предложенным нами способом риска от уровня САД у некурящих женщин со средним или ниже образованием и нормальной ЧСС, а на рис. 37Б – аналогичная зависимость по данным построенной с использованием регрессионного прототипа таблицы (цветной рис. 17).

Из рис. 37 видно, что при увеличении САД от 160 до 180 мм рт.ст. и величинах суммарного риска более 2-3% рассчитанный нашим способом риск резко увеличивается в сравнении с предыдущими градациями САД. Для рассчитанных по регресси-

онной модели рисков такой эффект малозаметен, зависимость риска от САД близка к линейной.

### *Частота сердечных сокращений*

По данным основанного на СОК ФР способа (цветные рис. 1-8) между ЧСС и риском фатальных событий имеется связь (при тахикардии риск выше), но этот эффект не является настолько универсальным, как при использовании регрессионных моделей (цветные рис. 17-24) – он наблюдается приблизительно в половине случаев. Обратное соотношение рисков (т.е. более низкий риск при тахикардии) встречается редко, различия выражены слабо и, вероятно, не являются клинически значимыми.

Таким образом, для данного ФР характерны два варианта – приблизительно в половине случаев его роль незначима, в остальных подгруппах тахикардия сопряжена с повышенным риском. Первый вариант более характерен для женщин с образованием выше среднего, в особенности курящих, а также для всех категорий лиц с очень низким (в пределах 2-3%) суммарным риском. Второй вариант чаще наблюдается у мужчин и некурящих женщин с образованием средним или ниже.

Вышеизложенное в целом согласуется с данными ряда эпидемиологических исследований, в которых выявлены ассоциации смертности с ЧСС в покое, наблюдавшиеся, однако, не во всех группах обследованных лиц [44-47, 74, {Mensink G.B., 1997 258 /id}, 75]. Этот вопрос будет подробнее рассмотрен при описании таблиц риска для когорты США в сравнении с данными литературы для исследования NHEFS.

Ранее упоминалось, что одним из ограничений нашего способа является невысокая надежность оценки риска в случаях, когда в когорте мало лиц с данным сочетанием уровней ФР. Одним из способов проверки таких оценок может быть использование другой, независимой выборки. В следующем разделе представлены построенные по такому же принципу таблицы риска для неорганизованной популяции США.



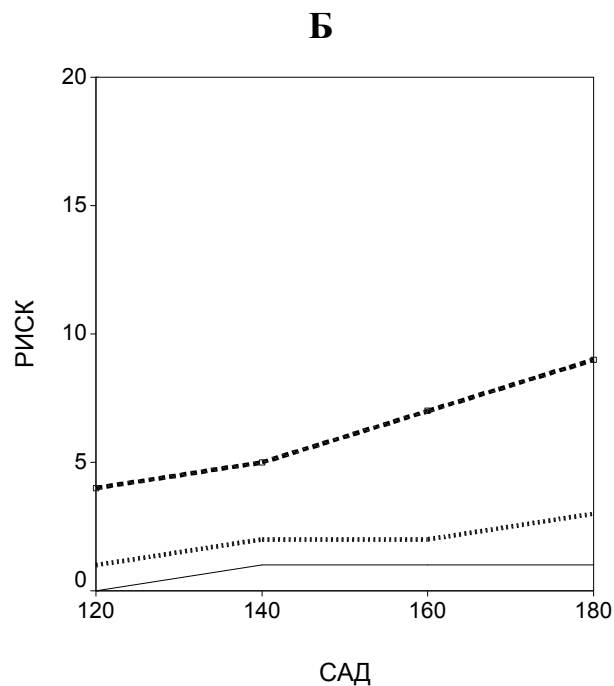
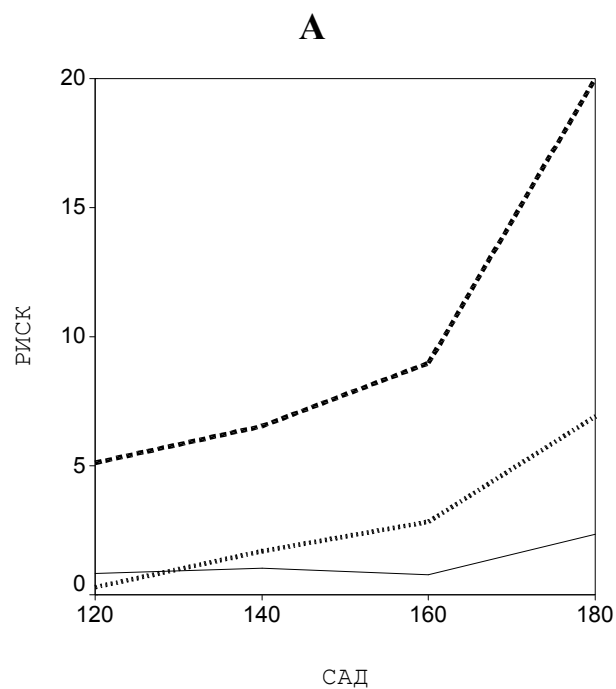


Рис. 37. Зависимость десятилетнего сердечно-сосудистого риска, рассчитанного в российской когорте предложенным нами способом СКОК ФР (А) и по регрессионной модели (Б) от уровня САД у некурящих женщин с нормальной ЧСС и образованием средним или ниже.

Сплошная, точечная и прерывистая линии соответствуют возрастным группам 40, 50 и 60 лет.

По оси абсцисс – САД (мм рт.ст.); по оси ординат – риск фатально-го сердечно-сосудистого события (%).

## **6. ТАБЛИЦЫ РИСКА НА ОСНОВЕ САМООРГАНИЗУЮЩИХСЯ КЛАСТЕРОВ ФАКТОРОВ РИСКА ДЛЯ КОГОРТЫ США**

*В.Г. Вилков*

На цветных рисунках 9-16 представлены таблицы рисков обусловленных ССЗ и всеми причинами фатальных событий в течение 10 и 20 ближайших лет для когорты США, на рисунке 34 – число наблюдений в исследованной выборке, соответствующих ячейкам указанных таблиц.

Из рисунка 34 видно, что у женщин когорты США в сравнении с российской реже встречаются ячейки таблицы риска, которым соответствует очень малое (до 10) число наблюдений, для мужчин ситуация обратная – таких ячеек меньше в российской когорте.

### **6.1. Связь между смертностью и отдельными ФР**

*Пол*

Как и в российской когорте, для мужчин характерны более высокие риски в сравнении с женщинами. Это наблюдается для сердечно-сосудистых и общих рисков при 10- и 20-летних сроках наблюдения.

В целом имеется соответствие результатов основанного на СОК ФР метода в когортах России и США.

*Возраст*

С увеличением возраста при равных уровнях прочих ФР риски повышаются, т.е. общая закономерность такая же, как в российской когорте.

В сравнении с российской когортой у мужчин когорты США обращают внимание значительно более низкие абсолютные риски в возрасте 40 и 50 лет. У 60-летних американских мужчин риски приближаются к таковым в российской когорте, а при среднем или ниже образовании нередко наблюдается более высокий риск в сравнении с соответствующими российскими

подгруппами. Следствием является более выраженный прирост риска у 60-летних мужчин в сравнении с 50-летними в когорте США в сравнении с российской.

### *Курение*

Из построенных нашим способом таблиц видно (цветные рис. 9-16), что в когорте США риск у курящих выше в сравнении с некурящими.

В отличие от российской когорты реже встречаются случаи с приблизительно одинаковыми рисками у курящих и некурящих. В когорте США это наблюдается, как правило, при низких суммарных рисках (в пределах 0-2%).

### *Образование*

Как и в российской когорте, в США риск смерти ниже при более высоком уровне образования.

Исключения встречаются у курящих лиц с нормальным или незначительно повышенным САД (120-140 мм рт.ст.). В указанных случаях величины риска различаются незначительно, эффект проявляется не во всех возрастных группах.

В целом можно заключить, что как в России, так и в США для большинства лиц имеется отчетливая связь уровня сердечно-сосудистого и общего рисков с образованием – при более высоком уровне образования риски ниже, в российской когорте указанная зависимость носит более универсальный характер.

### *Систолическое АД*

Складывается впечатление, что в когорте США уровень САД оказывает менее выраженное влияние на риск внутри каждой возрастной группы по сравнению с когортой России.

В когорте США полностью подтвердилась описанная для российской когорты общая закономерность увеличения смертности по мере роста уровня САД (цветные рис. 9-16).

Подтвердился также ряд особенностей связей величин риска и САД, описанных ранее для российской когорты:

- относительно слабое влияние САД при невысоком суммарном риске (в пределах 5%);

- более выраженный прирост риска при САД 180 (иногда 160-180) мм рт.ст. по отношению к предыдущим градациям САД в части подгрупп.

Рассмотрим последний феномен более подробно. Иллюстрирующий его пример для когорты США представлен на рисунке 38, демонстрирующем отчетливое сходство с российской когортой (рис. 37).

Обобщив данные для когорт России и США можно заключить, что выраженный прирост риска при повышении САД до 160-180 мм рт.ст. характерен для подгрупп с умеренными величинами суммарных рисков, абсолютные величины которых различны в зависимости от вида смертности и срока наблюдения. На фоне относительно высоких суммарных рисков данный эффект становится менее заметным. Можно предположить, что это обусловлено сочетанным воздействием высоких уровней других ФР.

При низких суммарных рисках наблюдается слабая связь между их величинами и уровнями отдельных ФР, включая САД. Это видно на рисунках 37-38 в подгруппах женщин 40-летнего возраста, у которых 10-летние сердечно-сосудистые риски не превышают 3% (цветные рис. 1 и 9).

Выявление описанных выше сходных особенностей связей риска с уровнем САД на материале двух больших популяционных исследований в разных странах представляется сильным аргументом в пользу работоспособности предлагаемого в данной работе способа оценки риска и корректности разработанных нами таблиц риска.

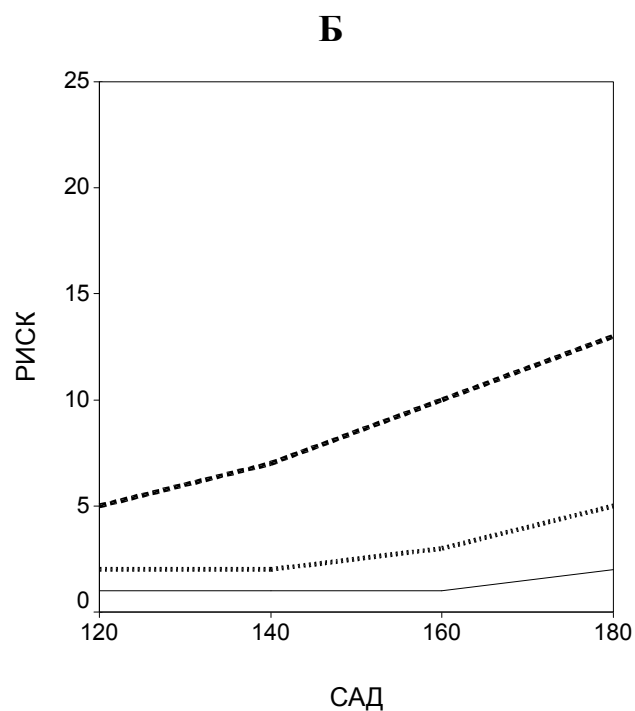
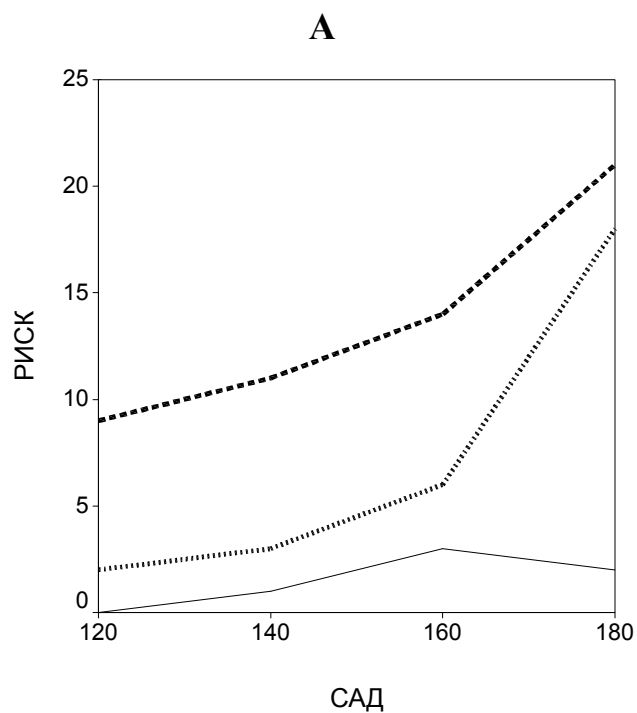


Рис. 38. Зависимость десятилетнего сердечно-сосудистого риска, рассчитанного в когорте США предложенным нами способом СОК ФР (А) и по регрессионной модели (Б) от уровня САД у курящих женщин с нормальной ЧСС и образованием средним или ниже.

Сплошная, точечная и прерывистая линии соответствуют возрастным группам 40, 50 и 60 лет.

По оси абсцисс – САД (мм рт.ст.); по оси ординат – риск фатально-го сердечно-сосудистого события (%).

### *Частота сердечных сокращений*

По данным нашего способа (цветные рис. 9-16) у белых мужчин и женщин США риск фатальных событий выше при тахикардии в сравнении с нормокардией примерно в 1/3 подгрупп, т.е. несколько реже, чем в российской когорте. В когорте США этот эффект чаще наблюдается применительно к общей смертности у лиц со средним или ниже образованием в возрасте 45 лет и более, чаще у некурящих женщин и курящих мужчин. В большинстве остальных подгрупп различия величин риска очень слабые, на единицы %, однако их направленность та же – риск выше при тахикардии. Обратное соотношение рисков встречается редко и различия при этом минимальны.

Описанные особенности отчасти совпадают с таковыми в российской когорте (повышенный риск при тахикардии более свойственен некурящим женщинам и лицам с образованием средним или ниже, а также отсутствие влияния ЧСС на риск при очень малых величинах последнего).

Сопоставление полученных результатов с данными литературы позволяет отметить следующее. Как уже упоминалось, опубликованные данные о ЧСС как независимом ФР неоднозначны. Наши данные свидетельствуют о значимости тахикардии как фактора повышенного риска (с учетом влияния пола, возраста, САД, статуса курения и образования), действие которого проявляется, однако, не во всех случаях, а приблизительно в 1/3 – 1/2 части когорты. Они также согласуются с мнением ряда исследователей о более сильной связи ЧСС с общей смертностью в сравнении со смертностью от ИБС и сердечно-сосудистой [76-77, 60]. Это не укладывается в представления, согласно которым тахикардия ассоциируется именно с повышенным риском смерти от ИБС, а связь с общей смертностью существует за счет вклада в нее смертности от ИБС [53].

Следует подчеркнуть, что уточнение роли данного ФР представляет большой практический интерес, поскольку создает предпосылки для его коррекции, в частности медикаментозной. В обзорных публикациях В.И. Маколкина и Ф.Н. Зябрева о роли ЧСС как ФР сердечно-сосудистых заболеваний [54, 60] от-

мечено, что недостатком популярных в настоящее время шкал для оценки риска, к примеру, SCORE, является отсутствие ЧСС среди учитываемых показателей.

Имеется теоретическая база в виде доказательств прямого влияния ЧСС на прогрессирование коронарного атеросклероза [39, 78]. Известны экспериментальные данные о том, что снижение ЧСС замедляет липид-индуцированный атеросклероз у приматов [79], в том числе при снижении ЧСС посредством лечения бета-блокаторами [80]. В целом, можно считать, что есть экспериментальные и клинические свидетельства о том, что ЧСС является вероятным детерминантом прогрессирования атеросклероза и его клинических манифестаций.

Выводы и рекомендации, опирающиеся на обобщенные данные по когортам обеих стран, сформулированы в заключении.

## **6.2. Роль сочетаний ФР**

В известных способах оценки риска, в частности системе SCORE, основанных на моделях регрессии, по существу используется суммирование эффектов учитываемых ФР. В результате по мере повышения АД риск увеличивается по линейному закону, курение удваивает риск и т.п. (цветные рис. 17-32, [7]). Это вытекает из ограничений использованных статистических методов, в первую очередь допущения о линейном характере связей между включенными в модель переменными.

В разработанном нами способе использованы нейросетевые методы, которые нелинейны по своей природе и в принципе не имеют ограничений, свойственных упомянутым выше прототипам. При анализе построенных в соответствии с предложенным нами подходом таблиц риска (цветные рис. 1-16) удалось выявить неочевидные ранее закономерности:

- для областей таблиц с близкими к нулю (приблизительно 0-2%) уровнями суммарного риска характерна монотонность – связь с изменениями величин отдельных ФР практически отсутствует;

- в областях таблиц с наиболее высокими величинами риска, обычно при выраженном повышении АД, связь суммарного риска с уровнями отдельных ФР (таких как курение, ЧСС) выражена слабее, чем при средних величинах риска;
- влияние отдельных ФР сильнее выражено при средних для данного вида риска и срока наблюдения уровнях суммарного риска.

Вышеизложенное позволяет сформулировать некоторые гипотезы:

1. у контингента лиц с высоким риском прежде всего целесообразно стремиться к коррекции наиболее значимого для прогноза ФР, как правило, это повышение САД более 170 (в части случаев более 150) мм рт.ст.;
2. после снижения риска до среднего уровня профилактические мероприятия могут быть дифференцированы в зависимости от особенностей подгруппы, к которой относится данный индивид.



## 7. ПРИМЕРЫ ОЦЕНКИ ПРЕДПОЛАГАЕМОГО ЭФФЕКТА КОРРЕКЦИИ ФАКТОРОВ РИСКА

*В.Г. Вилков*

Рассмотрим несколько гипотетических примеров коррекции риска в некоторых значимых для клинической практики ситуациях.

### 7.1. Мужчины с высоким риском

Максимальные величины суммарного риска наблюдаются при 20-летнем наблюдении у 60-летних мужчин со средним или ниже образованием (цветные рис. 8, 16).

В данной подгруппе российской когорты (цветной рис. 8) важная роль принадлежит ЧСС – на фоне тахикардии уменьшение риска по мере снижения САД выражено меньше, чем при нормальной ЧСС.

У некурящих российских мужчин 20-летний общий риск намного (как правило, на 15-20%) ниже в сравнении с курящими во всем диапазоне значений САД и ЧСС.

У 60-летних мужчин со средним или ниже образованием в когорте США (цветной рис. 16) связи между общим 20-летним риском, статусом курения, САД и ЧСС принципиально не отличаются от российской когорты.

С учетом вышеизложенного для 60-летних мужчин со средним или ниже образованием приоритетные рекомендации можно сформулировать следующим образом (если ориентироваться на снижение в первую очередь общего риска при максимальных сроках наблюдения):

1. Безусловно необходим отказ от курения.
2. У склонных к тахикардии лиц с артериальной гипертонией наиболее перспективны обладающие отрицательным хронотропным эффектом антигипертензивные лекарственные средства.
3. У склонных к тахикардии нормотоников при нормализации ЧСС есть основания ожидать снижения общего риска,

однако вопрос нуждается в дальнейшем изучении, особенно в части фармакологических воздействий (прежде всего в плане сопутствующих негативных эффектов). Физиологические способы немедикаментозного снижения ЧСС можно рекомендовать более уверенно (проверка некоторых гипотез приведена ниже в разделе 8).

У 50-летних мужчин российской когорты со средним или ниже образованием риски заметно меньше у лиц без явного повышения САД в сравнении с 60-летними, однако в целом закономерности изменений общего риска в зависимости от САД и ЧСС сходны с таковыми у 60-летних (цветной рис. 8). Соответственно приведенные выше рекомендации справедливы и для данной подгруппы.

У 50-летних мужчин в когорте США общие риски намного ниже в сравнении с 60-летними, данная подгруппа не может рассматриваться среди лиц с высоким риском.

## **7.2. Лица со средним уровнем риска**

У лиц со средними величинами 20-летних рисков смерти от всех причин (риски порядка 20-40%) нередко наблюдается сильная связь между суммарным риском и ЧСС (цветные рис. 7-8) – тахикардия ассоциируется со значительно более высоким риском в сравнении с аналогичным контингентом с нормальной ЧСС, особенно у некурящих лиц (например, в российской когорте у 40-летних некурящих мужчин со средним или ниже образованием или 50-летних некурящих женщин со средним или ниже образованием). Данный феномен малозаметен у лиц с выраженным повышением САД (более 170 мм рт.ст.). В ряде случаев он наблюдается и для 10-летних рисков (при величинах общих рисков порядка 10-30%), но его проявления менее демонстративны (цветные рис. 5-6), т.е. он проявляется при больших сроках наблюдения.

В когорте США описанный феномен также обнаруживается, но выражен менее отчетливо по сравнению с российской (цветные рис. 13-16).

Следовательно, для указанного контингента со средним уровнем общих рисков приоритетной задачей можно полагать нормализацию ЧСС (исключая больных артериальной гипертензией с САД более 170 мм рт.ст., у которых первоочередной задачей является его снижение). Так, например, в российской когорте у 40-летних некурящих мужчин со средним или ниже образованием (цветной рис. 8) снижение САД от 160 до 120 мм рт.ст. сопряжено с уменьшением общего 20-летнего риска на 7% (с 22 до 15%), а нормализация ЧСС при неизменном САД ассоциируется со снижением риска на 17-19% (с 32-39 до 15-22%).

## **8. ВЫЖИВАЕМОСТЬ И НЕКОТОРЫЕ ПОТЕНЦИАЛЬНО ЗНАЧИМЫЕ ДЛЯ СНИЖЕНИЯ ЧАСТОТЫ СЕРДЕЧНЫХ СОКРАЩЕНИЙ ПОВЕДЕНЧЕСКИЕ ФАКТОРЫ**

*В.Г. Вилков*

Представленные в предыдущих разделах книги результаты свидетельствуют о том, что у значительного числа людей (порядка  $1/3 - 1/2$  популяции) имеется связь между величиной ЧСС и суммарным риском смерти – при высокой ЧСС риск больше. Некоторые исследователи, изучавшие данный феномен, пришли к сходным выводам [43-51, 53-54].

Закономерно возникает мысль о возможности уменьшения риска смерти посредством изменения ЧСС. Очевидно, что связи между таким интегральным показателем, как ЧСС, другими ФР и выживаемостью имеют весьма сложный характер и вряд ли могут быть изучены в «чистом» виде. Более реальна сугубо прикладная задача – исследовать потенциальную возможность уменьшения риска при снижении ЧСС посредством изменения связанных с ЧСС модифицируемых факторов.

Снижение ЧСС с целью управления риском посредством лекарственных средств вряд ли оправдано без тщательного предварительного изучения в проспективных исследованиях с вмешательством и большими сроками наблюдения, прежде всего с точки зрения сопутствующих негативных эффектов. Это подразумевает большое количество регистрируемых показателей, включая периодическое разностороннее лабораторное обследование. Исключением могут быть ситуации, когда лекарственные средства данному индивиду в любом случае будут назначены (например, для лечения артериальной гипертензии) и речь идет лишь о дополнительном критерии выбора из числа хорошо изученных и примерно равноценных групп лекарственных средств в пользу обладающего отрицательным хронотропным эффектом препарата.

В отличие от применения фармакологических препаратов рекомендации по снижению ЧСС с использованием физиоло-

гичных воздействий типа повышения физической активности вызывают меньше сомнений.

Рассмотрим несколько гипотетических примеров связи выживаемости и немедикаментозных воздействий, одним из ожидаемых эффектов которых является снижение ЧСС.

Выбор модифицируемых поведенческих факторов диктовался имеющимися в наборах данных показателями. Используя данные исследований серии NHANES, мы изучили два фактора: уровень физической активности и потребление кофе и/или чая.

В анализ включили часть выборки, в которой можно ожидать наиболее сильной связи между суммарным риском и величиной ЧСС. Из этих соображений на основании таблиц 20-летнего общего риска (цветные рис. 15-16) была отобрана подвыборка, включавшая 1465 курящих мужчин из когорты США старше 34 лет с образованием средним или ниже и САД  $< 170$  мм рт.ст.

При анализе выживаемости источником систематических ошибок могут быть различия в исходных профилях ФР в сравниваемых когортах. Обычно роль отдельных ФР в отношении дожития учитывают посредством использования многомерного моделирования. Однако этот подход имеет свои недостатки, обусловленные техническими ограничениями выбранного класса моделей [9], в любом случае он позволяет учесть влияние только включенных в модель факторов.

В настоящем разделе мы использовали принципиально иной метод коррекции систематических ошибок, основанный на подборе для каждого исследуемого изучаемой группы соответствующего парного случая из числа лиц контрольной группы с близкими характеристиками значимых показателей, исключая изучаемый фактор [9, 2]. Парные случаи подбирали с учетом принадлежности к когорте (в данном случае США), пола, статуса курения, уровня образования, возраста, величин САД и концентрации общего холестерина в сыворотке крови при первичном обследовании. Для номинативных и порядковых показателей совпадение в парных случаях было полным, для количественных использовали некоторые диапазоны допустимых

отклонений от величины соответствующего показателя у данного члена исследуемой группы. Стремались использовать как можно более узкие диапазоны, обеспечивающие приемлемое число наблюдений в контрольных группах, при этом учитывали предполагаемую значимость данного показателя для выживаемости. Описанный подход позволяет минимизировать систематические ошибки, связанные с возможным влиянием различий в исходных уровнях учтенных факторов без использования моделирования с присущими ему ограничениями и искажениями, что обсуждалось выше.

### **8.1. Физическая активность, ЧСС и выживаемость**

Известно, что для физически тренированных людей характерна относительная брадикардия [81]. Как будет показано далее, этот эффект проявляется и на популяционном уровне – в части популяции с высокой физической активностью (по данным самооценки, полученным в процессе анкетирования) ЧСС ниже в сравнении с лицами, оценивающими свою физическую активность как более низкую.

В использованных нами наборах данных исследований серии NHANES имеется несколько зарегистрированных при первичном обследовании переменных, характеризующих уровень физической активности. Мы использовали для анализа две переменные, отражающие уровень физической активности в обычных условиях ( $\Phi A_1$ ) и в период отпуска ( $\Phi A_2$ ). Каждая переменная выражена в баллах и имеет 3 градации: активность отсутствует, активность умеренная и активность очень высокая (таблица 2). На базе этих двух переменных была создана еще одна переменная  $\Phi A_{\text{общ}}$ , позволяющая дать обобщенную оценку физической активности. Эта переменная имела 2 градации:

- $\Phi A_{\text{общ}}=1$  – низкая физическая активность ( $\Phi A_1=1$  и  $\Phi A_2=1$ , или  $\Phi A_1=1$  и  $\Phi A_2=2$ , или  $\Phi A_1=2$  и  $\Phi A_2=1$ , или  $\Phi A_1=2$  и  $\Phi A_2=2$ , или  $\Phi A_1=1$  и  $\Phi A_2=3$ );
- $\Phi A_{\text{общ}}=2$  – высокая физическая активность ( $\Phi A_1=3$  и  $\Phi A_2=3$ , или  $\Phi A_1=2$  и  $\Phi A_2=3$ , или  $\Phi A_1=3$  и  $\Phi A_2=2$ , или  $\Phi A_1=3$  и  $\Phi A_2=1$ )

Из описанного выше алгоритма расчета  $\Phi A_{\text{общ}}$  видно, что при оценке уровня физической активности в целом мы придавали несколько большее значение повседневной физической активности в сравнении с таковой в период отпуска. При этом руководствовались тем, что для долговременных эффектов регулярные физические нагрузки важнее более интенсивных, но ограниченных несколькими неделями отпуска в течение года.

Исследуемая группа включала 574 человека с высокой физической активностью ( $\Phi A_{\text{общ}}=2$ ), парная ей контрольная – 574 человека с низкой физической активностью ( $\Phi A_{\text{общ}}=1$ ). Корректность подбора пар подтверждена сравнением исследуемой и контрольной групп по  $U$ -критерию, не выявившем значимых различий по возрасту, САД, концентрации общего холестерина крови. Как уже упоминалось, значения таких переменных, как код популяции, раса, пол, статус курения и уровень образования в парных случаях полностью совпадали.

Средняя величина ЧСС в исследуемой группе с высокой физической активностью составила  $77,2 \pm 12,6$  уд./мин, а в контрольной группе с низкой физической активностью  $80,0 \pm 13,5$  уд./мин ( $M \pm SD$ ). Эти различия статистически достоверны как по параметрическому  $t$ -критерию Стьюдента ( $p=0,0001$ ), так и по непараметрическому  $U$ -критерию Манна-Уитни ( $p=0,001$ ). Таким образом, у лиц с высокой физической активностью ЧСС ниже при равных значениях таких факторов, как пол, возраст, САД, концентрация общего холестерина крови, статус курения и уровень образования.

На рис. 39 представлены кривые выживания у мужчин с высокой физической активностью в сравнении с парной контрольной группой с низкой физической активностью. При высокой физической активности уровень дожития достоверно выше как для сердечно-сосудистой ( $p=0,001$ ) так и общей смертности ( $p=0,0001$ ), суммарный абсолютный риск смерти от всех причин ниже на 13% в сравнении с низкой физической активностью (для сроков наблюдения 10-20 лет).

Таким образом, у курящих белых мужчин в когорте США со средним или ниже образованием при высокой физической активности уровень дожития достоверно выше при равных зна-

чениях таких ФР, как возраст, САД и концентрация общего холестерина в крови. Этот феномен сочетается с более низкой ЧСС. Следовательно, можно ожидать, что у данного контингента повышение физической активности с высокой вероятностью положительно повлияет на дожитие, целесообразно акцентировать внимание на этой рекомендации. При высокой физической активности в сравнении с низкой ожидаемая разница суммарного абсолютного риска смерти от всех причин для 10-20-летних сроков наблюдения составляет более 10%.

Возникает вопрос, насколько описанный эффект физической активности связан с ЧСС. Был проведен аналогичный описанному выше анализ с тем отличием, что при подборе парных случаев учитывали также величину ЧСС, при этом средняя величина ЧСС в исследуемой и контрольной группах не различалась ( $77,3 \pm 9,9$  уд./мин и  $77,4 \pm 10,0$  уд./мин соответственно,  $p > 0,7$ ).

На рис. 40 представлены кривые выживания в группе с высокой физической активностью в сравнении с парной контрольной группой с низкой физической активностью. При равных величинах ЧСС в группе с высокой физической активностью уровень дожития остается достоверно выше как для сердечно-сосудистой ( $p = 0,004$ ), так и для общей смертности ( $p = 0,001$ ). Однако величины суммарных абсолютных рисков в указанных группах различаются меньше в сравнении с предыдущим анализом, например для общих рисков при 10-летнем наблюдении на 5,5% против 13% на рис. 39.

Таким образом, если при подборе парных групп учесть ЧСС (величины ЧСС в парных группах равные), то различия в дожитии в связи с уровнем физической активности сохраняются, однако выраженность различий кривых Каплана-Мейера становится меньше (для общей смертности значимость различий 0,001 вместо 0,0001, различия суммарного риска при 10-летнем наблюдении 5,5% вместо 13%). Доля общих смертей при низкой физической активности осталась прежней (44%), при высокой физической активности она повысилась на 2% (37 вместо 35%).



Следовательно, проверка на парных наблюдениях показала, что при равенстве ЧСС связь уровня физической активности с дожитием сохраняется, но становится менее сильной.

Можно заключить, что в среднем при высокой физической активности ЧСС ниже. Связь между уровнем физической активности и дожитием не опосредована исключительно через ЧСС, однако сочетание высокой физической активности и более низкой ЧСС сопряжено с лучшей выживаемостью в сравнении с ситуацией, когда различается только уровень физической активности, а ЧСС одинакова.

## **8.2. Потребление кофе и чая, ЧСС и выживаемость**

Известно, что кофе и чай повышают АД, особенно у лиц страдающих артериальной гипертонией [82-84]. Ограничение потребления чая и кофе является традиционной рекомендацией и для коррекции тахикардии, поскольку кофеин относится к провоцирующим синусовую тахикардию факторам [85].

Мы попытались проверить значимость данного фактора на материале популяционных исследований серии NHANES. В использованных наборах данных имелся ряд зарегистрированных при первичном обследовании переменных, характеризующих уровень потребления кофе и чая. Для анализа использовали выраженную в баллах переменную, отражающую суммарное потребление кофе и/или чая и имеющую 3 градации: менее 1 чашки в день, 1-5 чашек в день и более 5 чашек в день (таблица 2). Как и в разделе 8.1, анализировали подвыборку из 1465 курящих мужчин из когорты США старше 34 лет со средним или ниже образованием и САД < 170 мм рт.ст. Применяли описанный выше подход, основанный на подборе парных случаев с учетом принадлежности к когорте, пола, статуса курения, уровня образования, возраста, величин САД и концентрации общего холестерина в сыворотке крови при первичном обследовании.

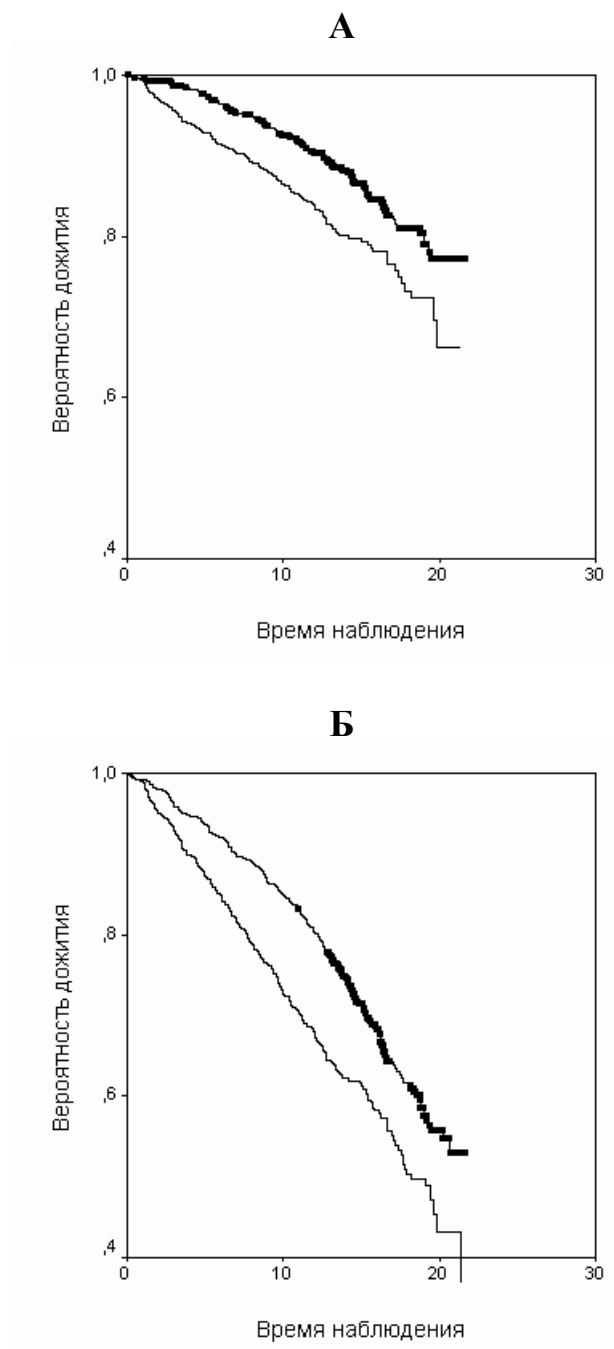


Рис. 39. Кривые выживания для смерти от ССЗ (А) и всех причин (Б) у курящих мужчин старше 34 лет с образованием средним или ниже и САД < 170 мм рт.ст. с высокой физической активностью (линия с темными маркерами) и в парной контрольной группе с низкой физической активностью (тонкая линия) в когорте США.

По оси абсцисс – время от момента первичного обследования до завершения наблюдения (годы); по оси ординат – кумулятивная доля выживших (доли единицы).

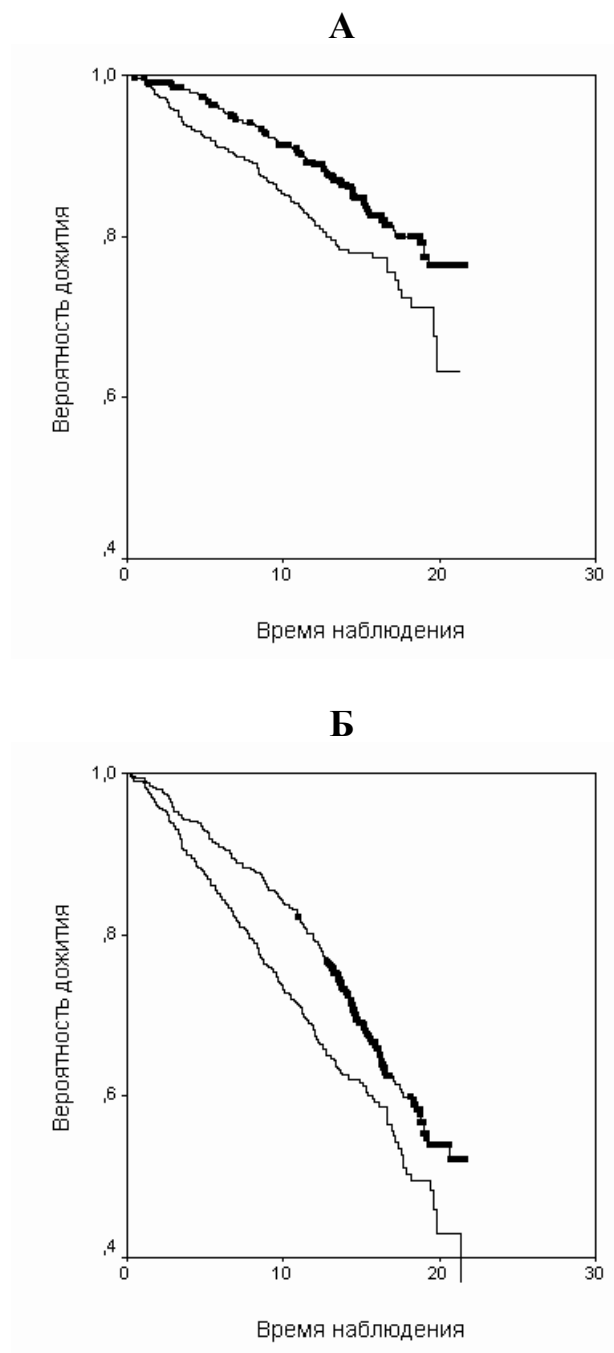


Рис. 40. Кривые выживания для смерти от ССЗ (А) и всех причин (Б) у курящих мужчин старше 34 лет с образованием средним или ниже и САД < 170 мм рт.ст. с высокой физической активностью (линия с темными маркерами) и в парной контрольной группе с низкой физической активностью (тонкая линия) в когорте США. В отличие от рис. 39 величины ЧСС в группах подобраны равными.

По оси абсцисс – время от момента первичного обследования до завершения наблюдения (годы); по оси ординат – кумулятивная доля выживших (доли единицы).

Исследуемая группа включала 99 мужчин, потреблявших менее 1 чашки кофе и/или чая в день, контрольная – 99 парных случаев из числа лиц, потреблявших более 5 чашек в день. Корректность подбора пар подтверждена сравнением исследуемой и контрольной групп по  $U$ -критерию, не выявившем значимых различий по возрасту, САД, концентрации общего холестерина крови. Значения таких переменных, как код популяции, раса, пол, статус курения и уровень образования в парных случаях полностью совпадали. Величины ЧСС в исследуемой и контрольной группах также не различались ( $p=0,3$  по непараметрическому  $U$ -критерию Манна-Уитни). Анализ дожития посредством построения кривых Каплана-Мейера не выявил различий между группами как для общей ( $p=0,7$ ), так и для сердечно-сосудистой смертности ( $p=0,6$ ).

Описанный выше анализ повторили для другой пары групп, одна из которых включала лиц, потреблявших более 5 чашек в день, другая была сформирована из числа лиц, потреблявших до 5 чашек в день. Результаты анализа не изменились – различий между парными группами по возрасту, САД, ЧСС, концентрации общего холестерина крови, статусу курения и уровню образования не было, уровни дожития не различались для общей ( $p=0,6$ ) и сердечно-сосудистой смертности ( $p=0,9$ ).

Таким образом, у курящих белых мужчин в когорте США со средним или ниже образованием с разным потреблением кофе и/или чая при равных значениях таких ФР, как возраст, САД, концентрации общего холестерина крови и уровень образования величины ЧСС при первичном обследовании и выживаемость в течение 20 лет не различаются.

Результаты анализа, представленные в данном разделе книги, позволяют предположить, что по крайней мере в части популяции при увеличении физической активности возможно повышение уровня дожития, это может сочетаться с некоторым снижением ЧСС покоя.

На популяционном материале не получено подтверждения гипотезы, что ограничение потребления кофе и чая способствует снижению ЧСС покоя и влияет на выживаемость.

## ЗАКЛЮЧЕНИЕ

*С.А. Шальнова, В.Г. Вилков*

Успешность профилактических мероприятий и эффективность медицинской помощи в целом в значительной степени зависят от точности оценки риска развития неблагоприятных событий. В течение последних десятилетий подобные исследования базируются на концепции ФР.

Известные к настоящему времени сведения о связи различных ФР со смертностью получены с использованием вариантов одного статистического метода – многомерного моделирования выживаемости [6], как правило модели пропорциональных интенсивностей Кокса [31, 32]. Следует отметить, что первое исследование риска смерти от всех причин на базе многомерных моделей опубликовано в 2005 году [34], до этого подобный подход использовали только для оценки риска смерти от ССЗ. Модели Кокса свойственны ограничения, присущие любому моделированию и связанные с упрощением, усреднением и т.п. Нужно также учитывать, что многомерные регрессионные модели наиболее успешно работают, когда первичные данные удовлетворяют характерным для параметрических методов условиям, включая соответствие распределений величин всех независимых переменных нормальному закону не только порознь, но и в многомерном пространстве входящих в модель признаков. На практике это означает, что будут учтены только линейные составляющие связей ФР с выживаемостью, за исключением частных случаев, когда нелинейные зависимости включены в модель в явном виде.

Положительными чертами подобных способов оценки риска являются удобство интерпретации, однозначность и непротиворечивость оценок и рекомендаций по коррекции ФР. К недостаткам, на преодоление которых нацелена настоящая работа, можно отнести нивелирование особенностей отдельных подгрупп, в связи с чем оценки и рекомендации могут оказаться недостаточно дифференцированными и, как следствие, в части случаев менее эффективными.

С вышеизложенным согласуются результаты анализа построенных нами в качестве прототипа таблиц риска, основанных на регрессионных моделях (цветные рис. 17-32), которые демонстрируют близкие к линейным монотонные зависимости между суммарным риском и уровнями отдельных ФР.

С другой стороны, линейные зависимости нехарактерны для медико-биологических данных [63, 2]. Применительно к использованным в настоящей работе показателям это подтверждено экспериментально (раздел 1.3).

Таким образом, нами обоснована перспективность нелинейных статистических методов для оценки суммарного риска по данным популяционных проспективных исследований.

С учетом приведенных выше соображений для решения поставленной задачи мы изначально ориентировались на использование искусственных нейронных сетей (ИНС) – нелинейных по своей природе методов, позволяющих отображать самые сложные зависимости. Второй принципиальной особенностью данной работы является применение не моделирования как такового (вне зависимости от использованной техники), а непосредственного расчета риска в части выборки. В рамках этого подхода разработан способ расчета риска в многомерных самоорганизующихся кластерах (СОК) ФР, основанный на статистической технике, обладающей рядом принципиальных отличий от применявшихся ранее способов. Особенностью нашего подхода является выделение в многомерном признаковом пространстве учтенных ФР самоорганизующихся кластеров, объединяющих лиц с близкими величинами характеризующих указанные факторы показателей, с последующим прямым расчетом фактически наблюдаемых в каждом кластере рисков. На завершающем этапе для каждой из ячеек таблиц риска (цветные рис. 1-16) осуществляется усреднение величин риска, рассчитанных с использованием нескольких ИНС.

При реализации нейросетевых методов использованы оригинальные решения некоторых важных для достижения успешного результата проблем (раздел 2). Технически кластеризацию осуществляли с использованием одного из нейросетевых

методов обучения без учителя – самообучающихся карт (сетей) Кохонена.

Прототипы, например таблицы риска на базе регрессионных моделей (цветные рис. 17-32) или системы SCORE [7, 4], дают однородную картину зависимости риска от уровней учтенных в модели ФР, это обусловлено особенностями использованной при их создании статистической техники, по существу сводящейся к линейному моделированию. Так, для российской популяции курение сопряжено с двукратным повышением 10-летнего сердечно-сосудистого риска, при более низком уровне образования указанный риск может повышаться на 30-80%, зависимость риска от уровня САД близка к линейной и т.п. [7].

Величины рассчитанных новым способом рисков смерти от ССЗ и всех причин в значительной части популяции заметно отличаются от таковых, полученных традиционным способом, использующим регрессионные модели дожития.

Сравнение точности расчета риска традиционным и новым способом продемонстрировало преимущества последнего (раздел 4).

Набор ФР, использованных в разработанных нами таблицах (цветные рис. 1-16), имеет некоторые особенности: наряду с такими классическими ФР, как пол, возраст, АД и статус курения, включены уровень образования и ЧСС, с другой стороны, не учтена концентрация холестерина в крови. Важная роль уровня образования для сердечно-сосудистого риска в российской популяции была показана ранее [7], в ходе проведенного в рамках настоящей работы анализа оказалось, что в популяции США данный фактор также является значимым. Предварительный анализ выявил, что в исследованных нами выборках при оценке сердечно-сосудистого и общего рисков роль ЧСС в целом выше по сравнению с концентрацией общего холестерина в крови, применительно к общей смертности это согласуется с результатами других проспективных исследований [34]. Исходя из этих соображений был сформирован окончательный перечень учтенных в таблицах ФР: пол, возраст, статус курения, уровень образования, величины САД и ЧСС.

Зависимости рассчитанного новым способом суммарного

риска от уровней классических ФР имеют ту же направленность, что и при использовании традиционного способа:

- риск выше у мужчин в сравнении с женщинами и у курящих в сравнении с некурящими;
- риск ниже при более высоком уровне образования;
- риск увеличивается по мере повышения САД и ЧСС.

Однако при анализе построенных нашим способом таблиц риска (цветные рис. 1-16) удалось выявить ряд особенностей.

1. Зависимость суммарного абсолютного риска от таких ФР, как пол, возраст, статус курения, уровень образования, величины САД и ЧСС не является столь однозначной, как считалось ранее, в частности исходя из системы SCORE.
2. Как при низком (до 2-3%), так и при очень высоком суммарном риске он относительно слабо связан с уровнями отдельных ФР.
3. При средних для данного вида смертности и срока наблюдения уровнях суммарного риска особенности влияния на смертность отдельных ФР проявляются сильнее. В частности, увеличение САД более 170 (в части случаев более 150) мм рт.ст. влечет более резкий прирост риска в сравнении с предыдущими градациями САД.
4. В когорте США в сравнении с российской смертность в целом ниже, различия наиболее выражены у мужчин в возрасте до 55 лет. У женщин и мужчин старше 55 лет различия смертности в указанных когортах выражены меньше.
5. При равных уровнях учтенных ФР суммарный риск ниже при высокой физической активности в сравнении с низкой, разница 20-летних рисков составляет более 10% (данные имеются только для когорты США).

Применив разработанный нами подход к данным более чем 20-летнего проспективного наблюдения за когортами из неорганизованных популяций России и США, первоначально обследованных в первой половине 70-х годов XX века, можно сформулировать некоторые дополнения к общепринятым [7, 4] рекомендациям.



1. У лиц с высоким суммарным риском фатальных событий целесообразно стремиться в первую очередь скорректировать наиболее значимые для прогноза ФР, а именно повышение САД более 170 (иногда более 150) мм рт.ст. и, в части подгрупп, курение.
2. После достижения среднего для данных возраста и пола уровня суммарного риска для максимальной эффективности профилактических мероприятий последние могут быть дифференцированы в зависимости от особенностей подгруппы с тем, чтобы сконцентрировать усилия на коррекции наиболее значимых для данного контингента ФР. В одних случаях более перспективно дальнейшее снижение уровня АД, в других – нормализация повышенной ЧСС, либо сочетание указанных мероприятий.
3. Лица, для которых характерна сильная связь суммарного риска с ЧСС, составляют значительную часть популяции. Для этого контингента снижение ЧСС может оказаться одним из приоритетных направлений профилактики. При наличии артериальной гипертензии у таких пациентов наиболее перспективны антигипертензивные лекарственные средства, обладающие отрицательным хронотропным эффектом. У нормотоников можно рекомендовать обратить особое внимание на повышение уровня физической активности.
4. Поддержание высокого уровня повседневной физической активности сопряжено со снижением суммарного риска до 10% (доказано только для когорты США, в российской когорте физическая активность не исследовалась).

Таким образом, как разработанный нами способ, так и основанные на моделировании дожития прототипы имеют свои плюсы и минусы. При разработке профилактических стратегий, ориентированных на популяцию в целом, представляется более удобным применение использующих регрессионные модели систем типа SCORE. В данном случае свойственные им эффекты усреднения не являются помехой, а простота и однозначность интерпретаций влияния ФР представляют собой

важное преимущество. Наш способ перспективен в случаях, когда требуется дать более дифференцированные рекомендации, учитывающие особенности конкретной подгруппы.

Важным сопутствующим результатом настоящей работы является подтверждение на большом популяционном материале ранее сделанного вывода [2] о желательности применения нелинейных статистических методов вместо линейных для изучения характеризующих ФР показателей. Об этом свидетельствуют приведенные в разделе 1.3 данные о том, что при достаточно большом числе наблюдений распределение величин всех проанализированных в настоящей работе количественных показателей (САД, ЧСС, индекса массы тела, концентрации в крови общего холестерина) значительно отличается от нормального.

До настоящего времени в публикуемых по теме данного исследования работах применялось многомерное моделирование, по существу использующее линейные связи между переменными. Полученные нами результаты свидетельствуют о том, что это не оправдано теоретически, а на практике в части случаев может повлечь существенное искажение результатов. Поэтому при использовании регрессии Кокса следует иметь в виду свойственные ей принципиальные ограничения в задачах данного типа, эффекты которых были обсуждены выше на конкретных примерах.

Еще одним, весьма важным, на наш взгляд, аспектом настоящей работы является демонстрация возможностей подхода, получившего название вторичного анализа данных, при котором для решения своих конкретных задач используются уже существующие в мировом информационном пространстве первичные данные. Можно ожидать, что это направление будет интенсивно развиваться, поскольку очевидны современные тенденции к расширению доступа индивидуумов к глобальным информационным ресурсам и стремительному увеличению объема последних. Это справедливо и в отношении части информационного пространства, содержащей данные медицинских исследований.

## ЭЛЕКТРОННАЯ ВЕРСИЯ КНИГИ

*В.Г. Вилков*

Электронная версия представлена в Internet по адресу [http://drvilkov.chat.ru/vilkov\\_book3.html](http://drvilkov.chat.ru/vilkov_book3.html). Она может быть также найдена по названию с помощью поисковых систем, например <http://www.rambler.ru/>, и в разделе «Медицина» каталога <http://list.ru/>.

## **ТАБЛИЦЫ РИСКА**

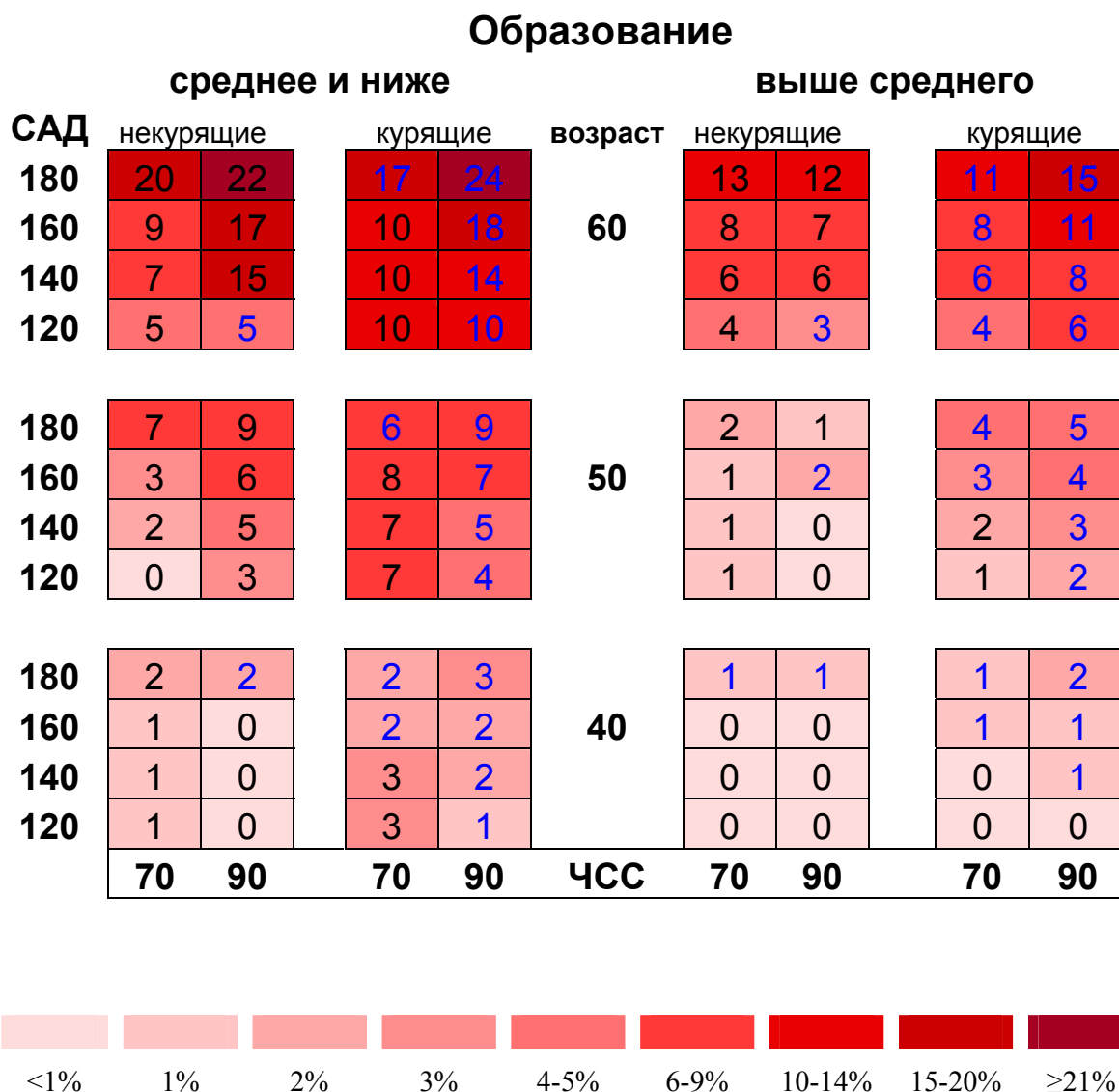


Рис. 1. Десятилетний риск смерти от ССЗ (%) для женщин (Россия).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 17).

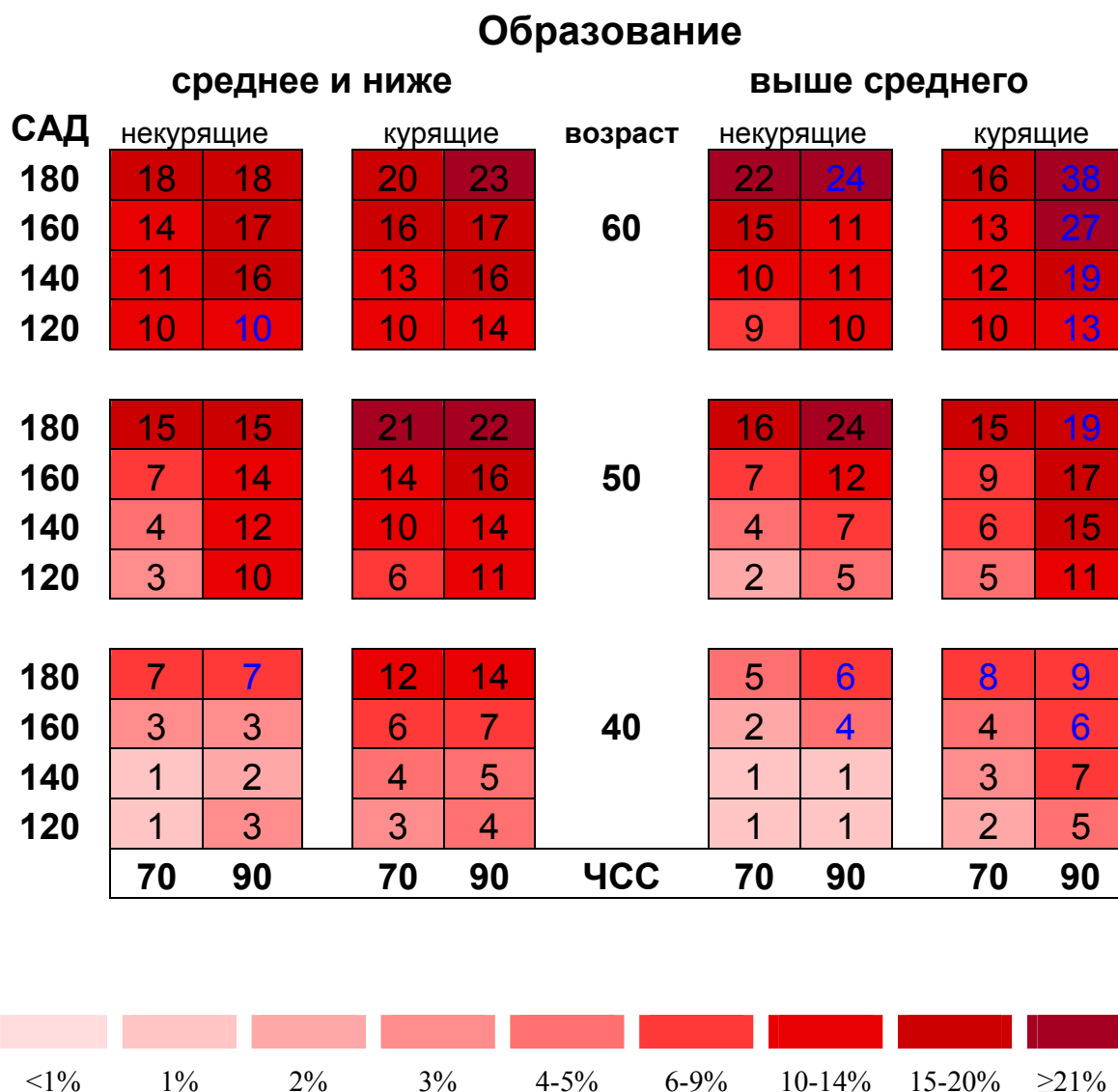


Рис. 2. Десятилетний риск смерти от ССЗ (%) для мужчин (Россия).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 18).

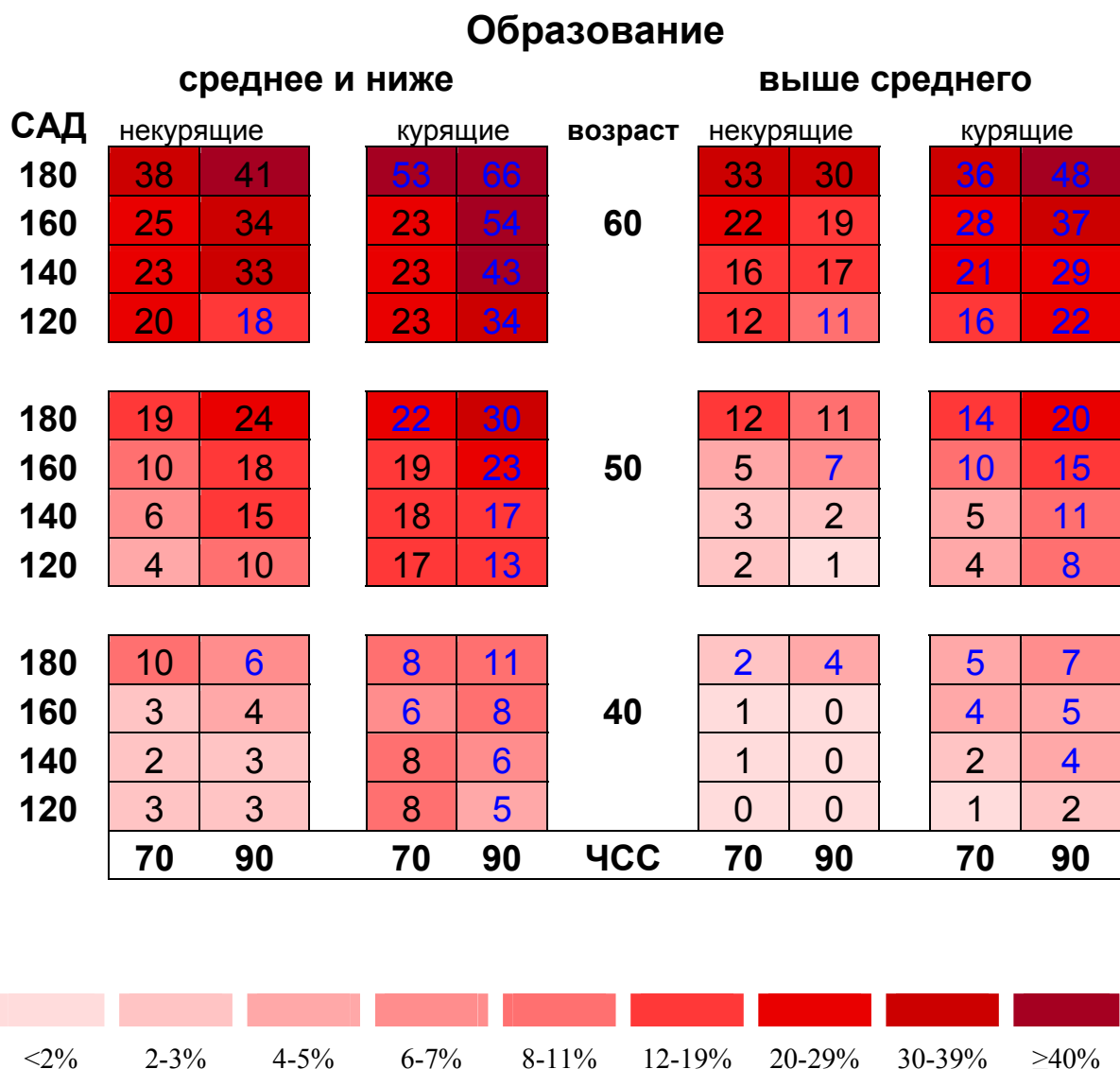


Рис. 3. Двадцатилетний риск смерти от ССЗ (%) для женщин (Россия).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 19).

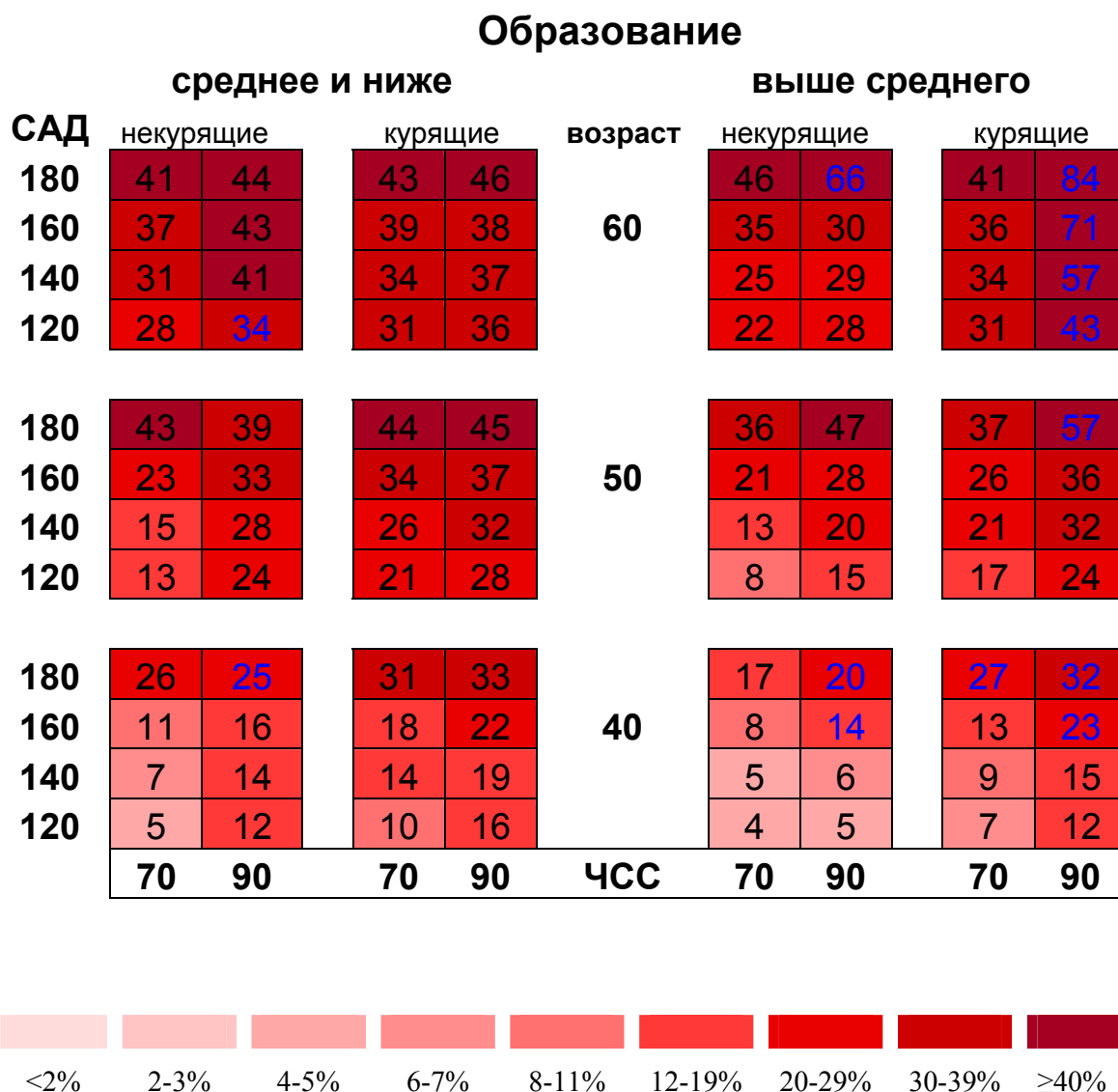


Рис. 4. Двадцатилетний риск смерти от ССЗ (%) для мужчин (Россия).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 20).



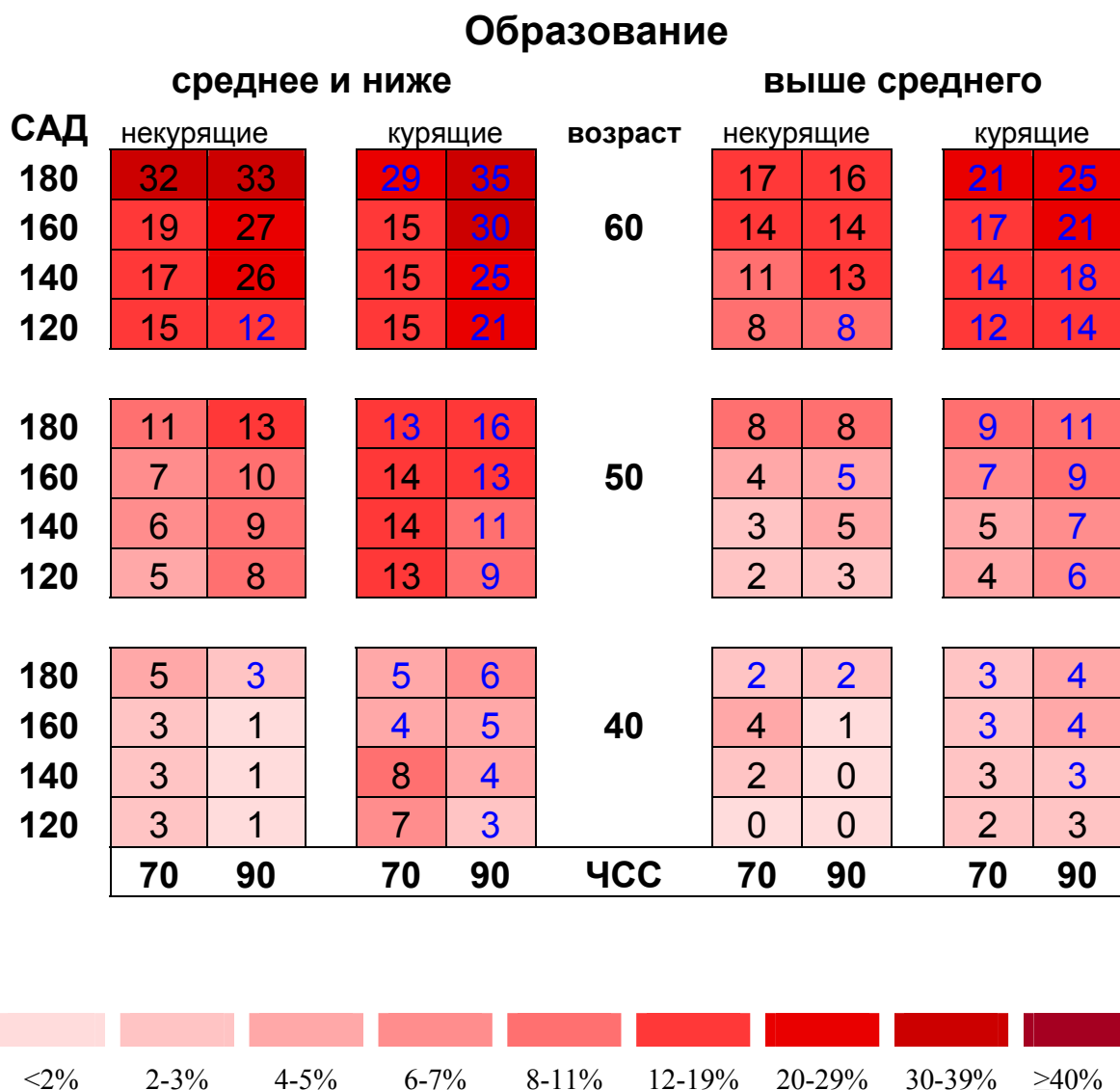


Рис. 5. Десятилетний риск смерти от всех причин (%) для женщин (Россия).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 21).

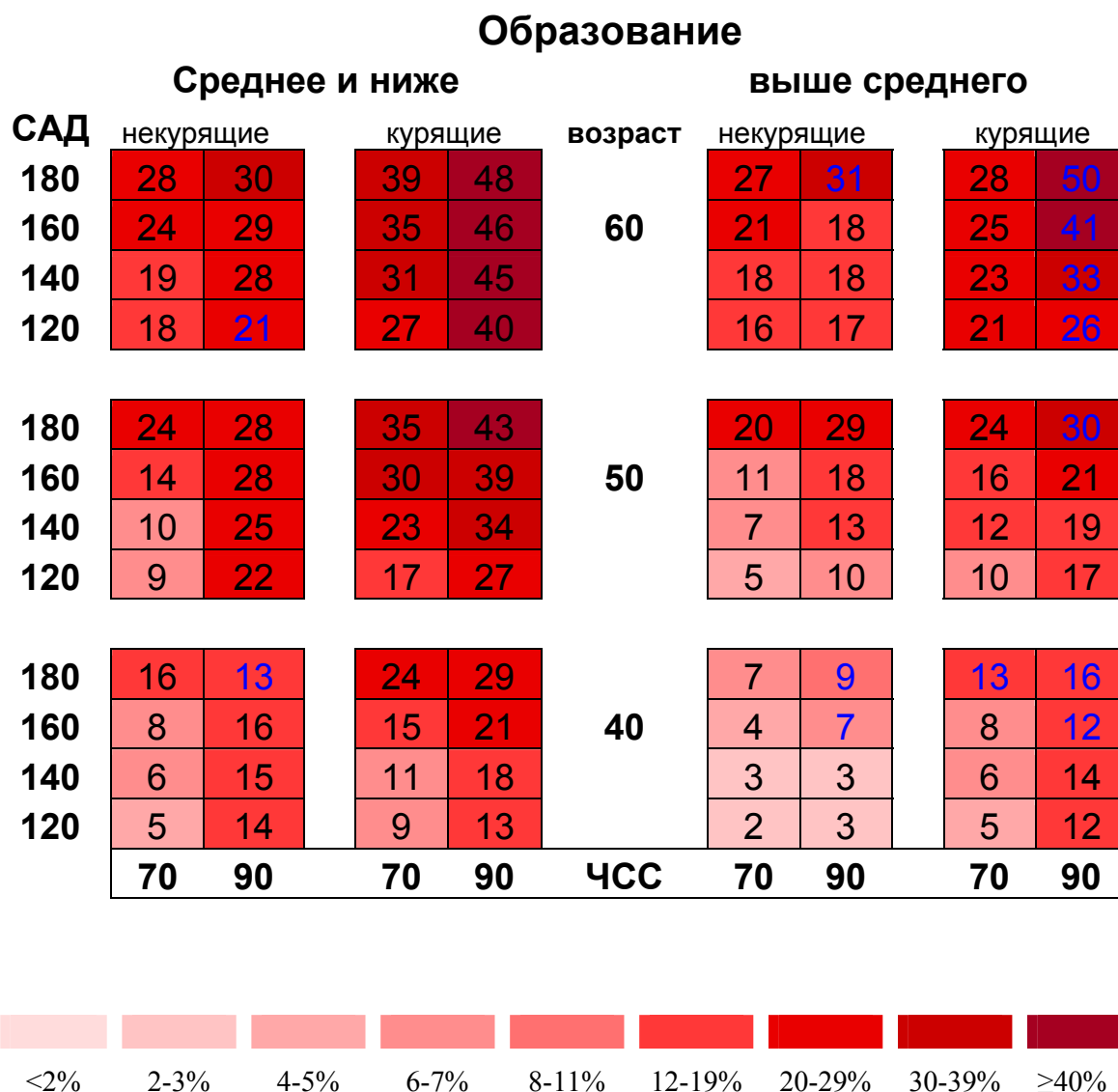


Рис. 6. Десятилетний риск смерти от всех причин (%) для мужчин (Россия).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 22).

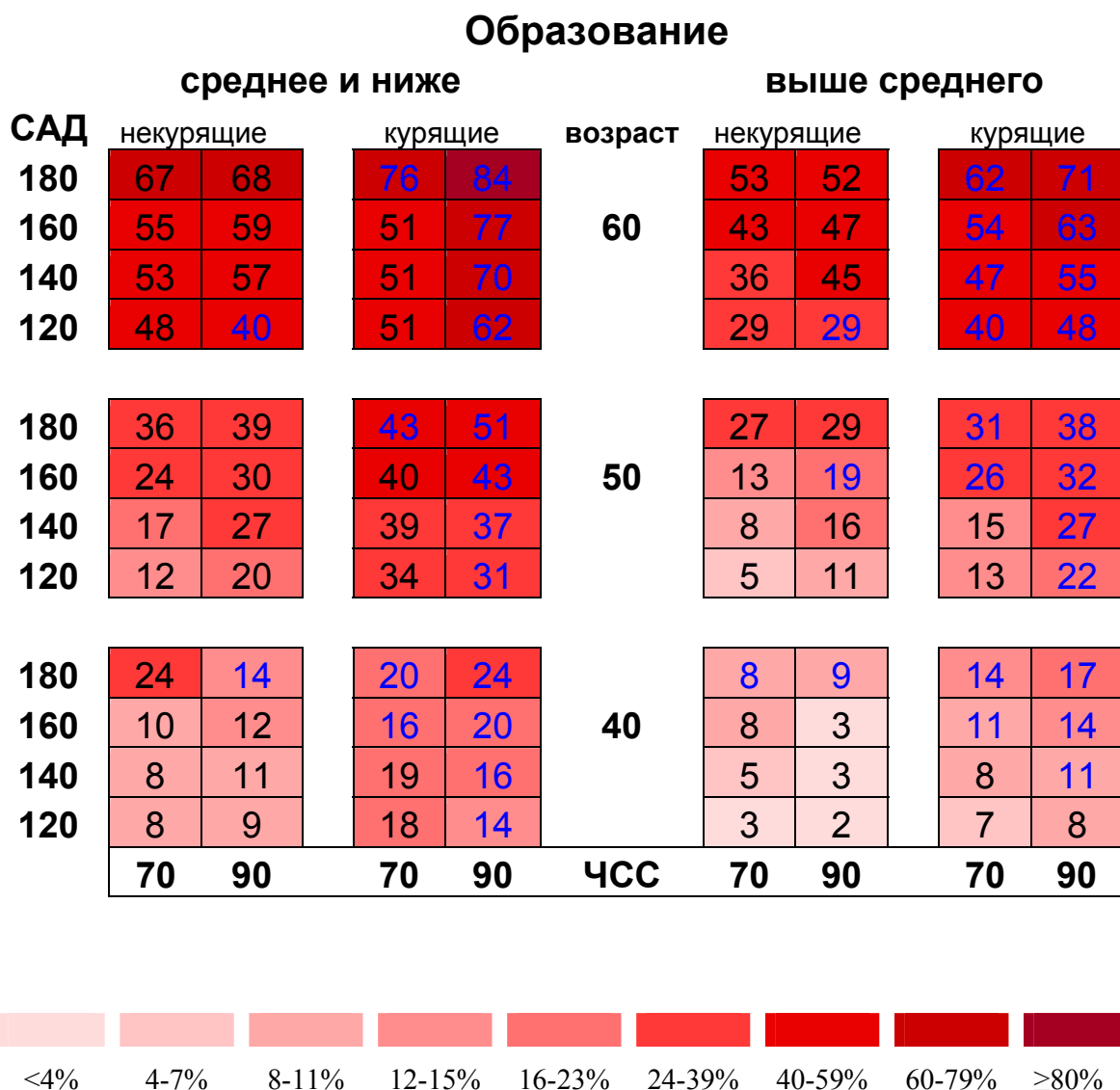


Рис. 7. Двадцатилетний риск смерти от всех причин (%) для женщин (Россия).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 23).

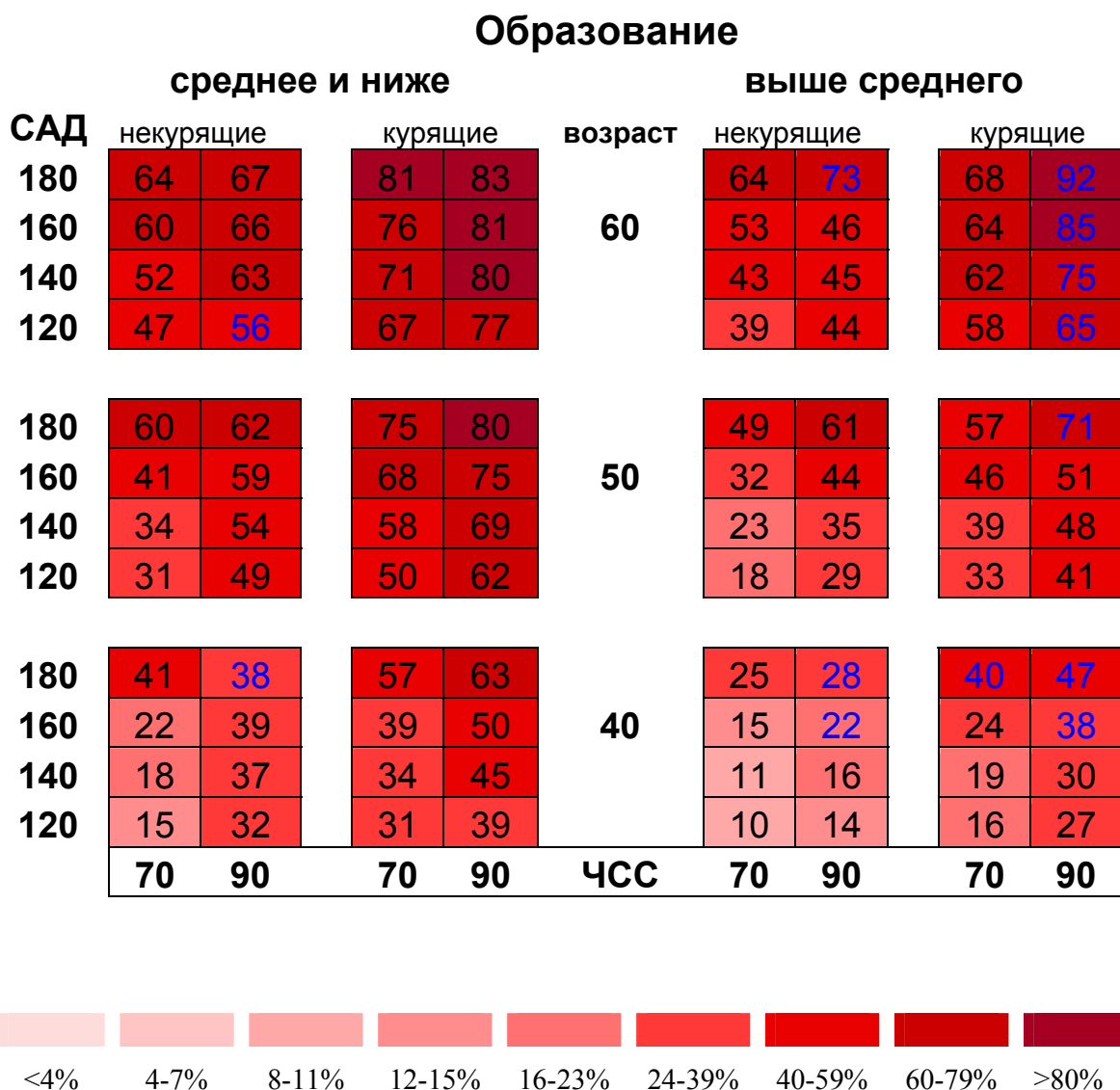


Рис. 8. Двадцатилетний риск смерти от всех причин (%) для мужчин (Россия).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 24).

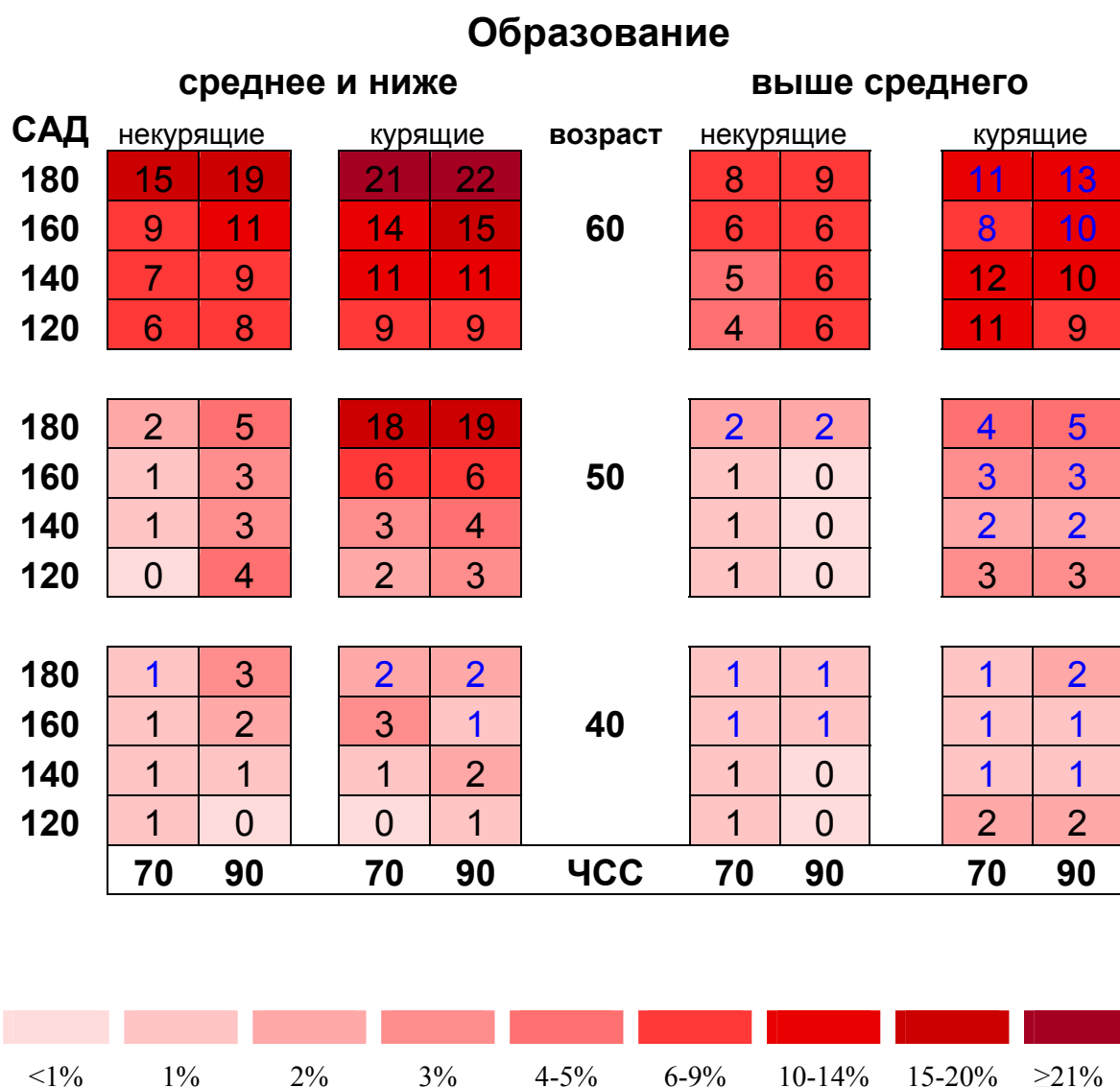


Рис. 9. Десятилетний риск смерти от ССЗ (%) для женщин (США).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 25).

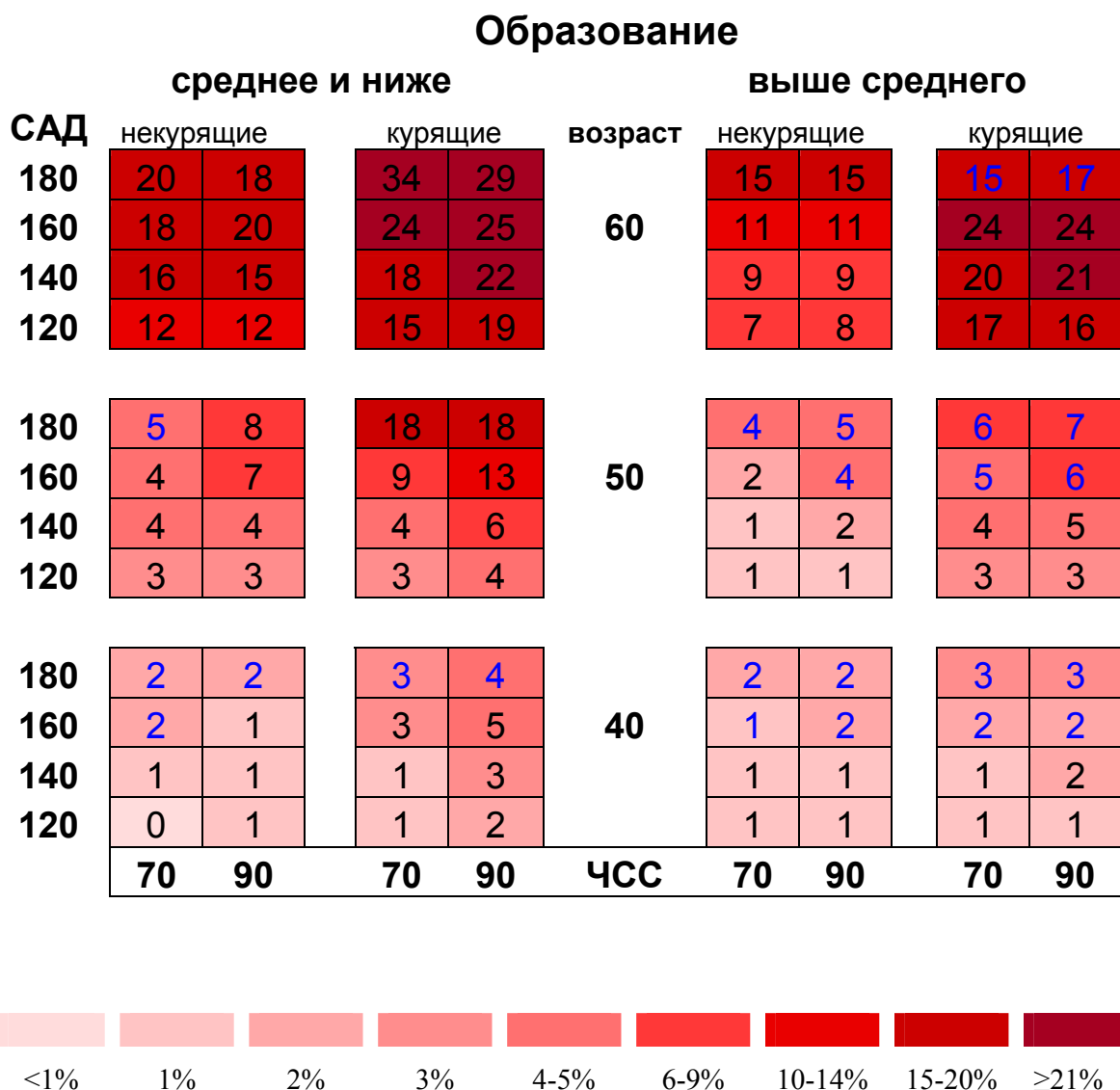


Рис. 10. Десятилетний риск смерти от ССЗ (%) для мужчин (США).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 26).

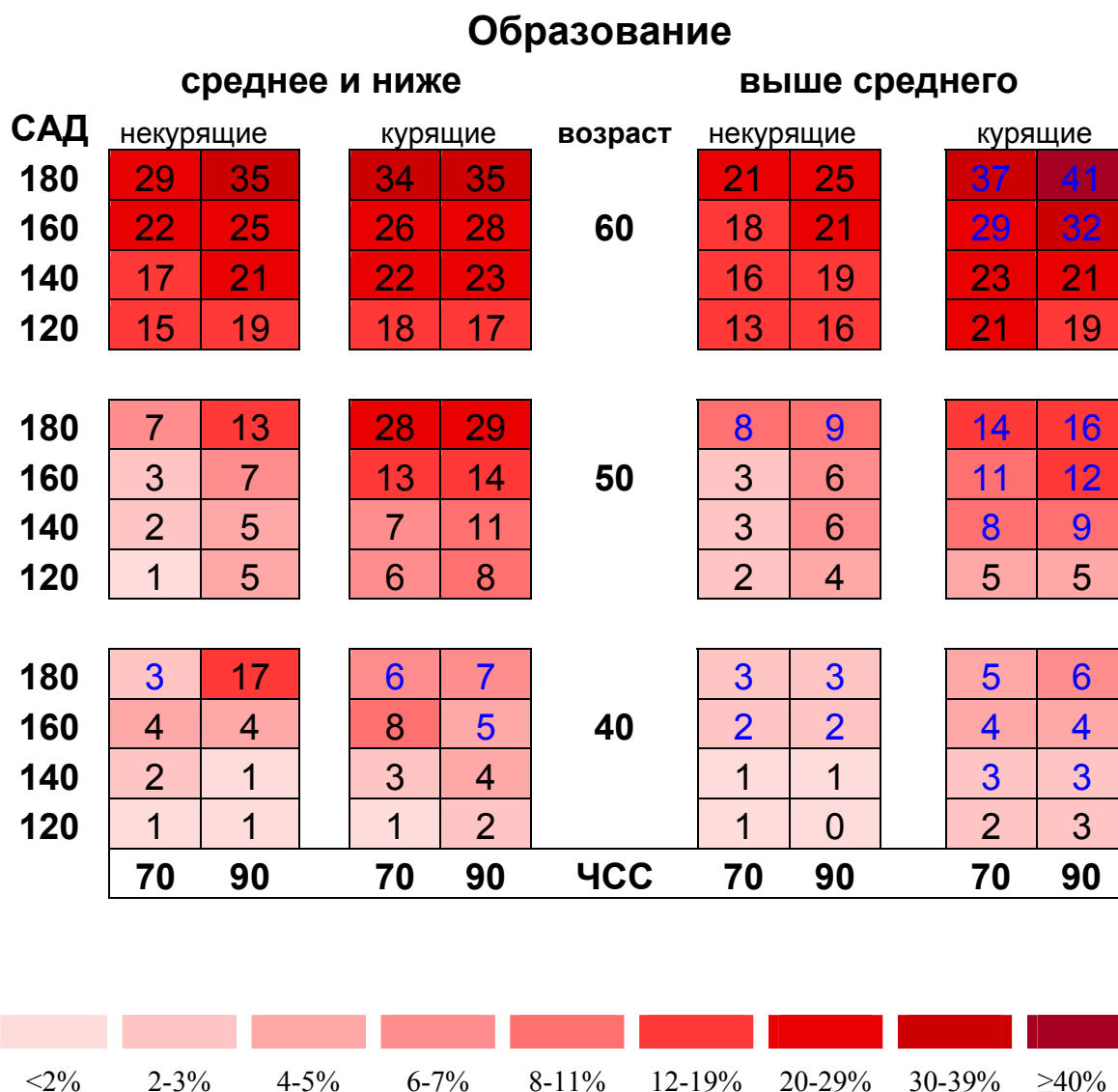


Рис. 11. Двадцатилетний риск смерти от ССЗ (%) для женщин (США).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 27).

| Образование    |           |    |         |    |               |           |    |         |    |
|----------------|-----------|----|---------|----|---------------|-----------|----|---------|----|
| среднее и ниже |           |    |         |    | выше среднего |           |    |         |    |
| САД            | некурящие |    | курящие |    | возраст       | некурящие |    | курящие |    |
| 180            | 39        | 37 | 46      | 44 | 60            | 30        | 31 | 45      | 49 |
| 160            | 32        | 34 | 37      | 38 |               | 23        | 24 | 34      | 34 |
| 140            | 30        | 29 | 30      | 34 |               | 19        | 21 | 30      | 30 |
| 120            | 23        | 22 | 27      | 31 |               | 16        | 19 | 27      | 25 |
| 180            | 18        | 27 | 32      | 32 | 50            | 14        | 16 | 21      | 23 |
| 160            | 11        | 21 | 20      | 24 |               | 6         | 13 | 17      | 19 |
| 140            | 9         | 11 | 13      | 15 |               | 6         | 7  | 8       | 11 |
| 120            | 6         | 8  | 10      | 11 |               | 5         | 6  | 7       | 7  |
| 180            | 7         | 8  | 11      | 13 | 40            | 6         | 6  | 9       | 10 |
| 160            | 6         | 5  | 9       | 12 |               | 5         | 5  | 7       | 8  |
| 140            | 3         | 4  | 6       | 7  |               | 2         | 2  | 4       | 4  |
| 120            | 2         | 4  | 4       | 6  |               | 2         | 2  | 4       | 3  |
|                | 70        | 90 | 70      | 90 | ЧСС           | 70        | 90 | 70      | 90 |

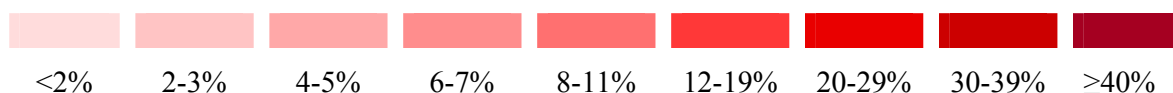


Рис. 12. Двадцатилетний риск смерти от ССЗ (%) для мужчин (США).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 28).



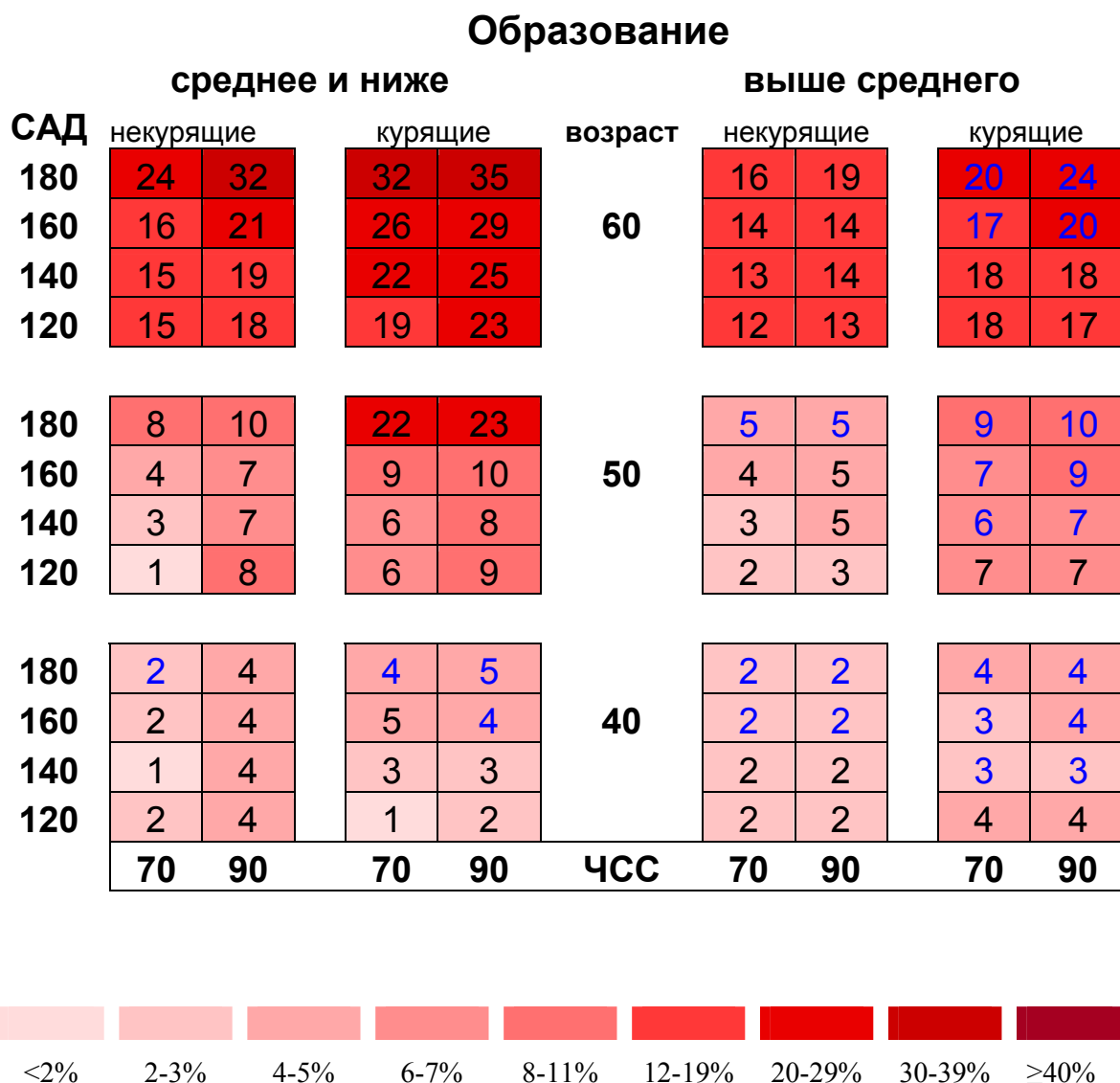


Рис. 13. Десятилетний риск смерти от всех причин (%) для женщин (США).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 29).

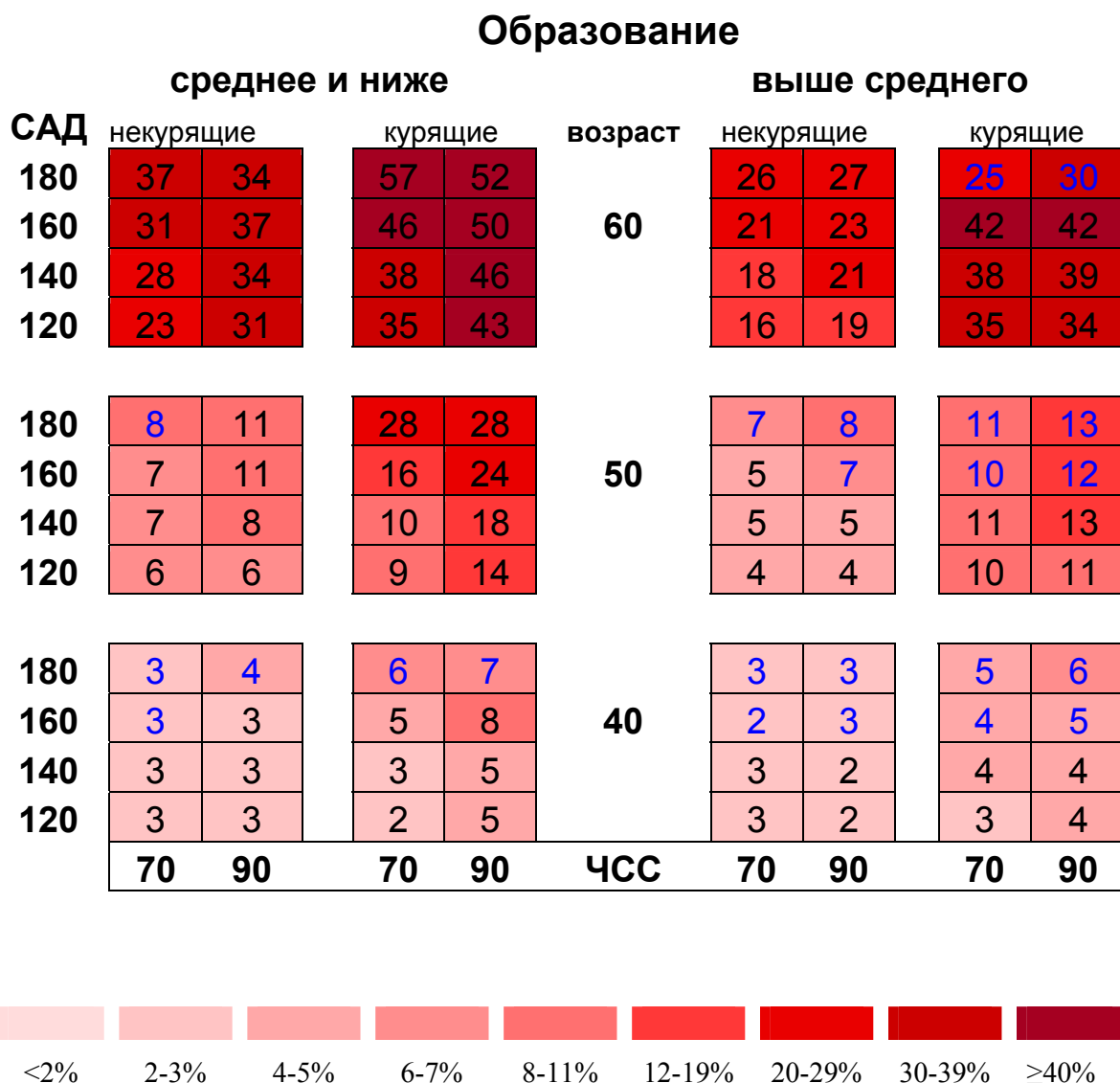


Рис. 14. Десятилетний риск смерти от всех причин (%) для мужчин (США).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 30).

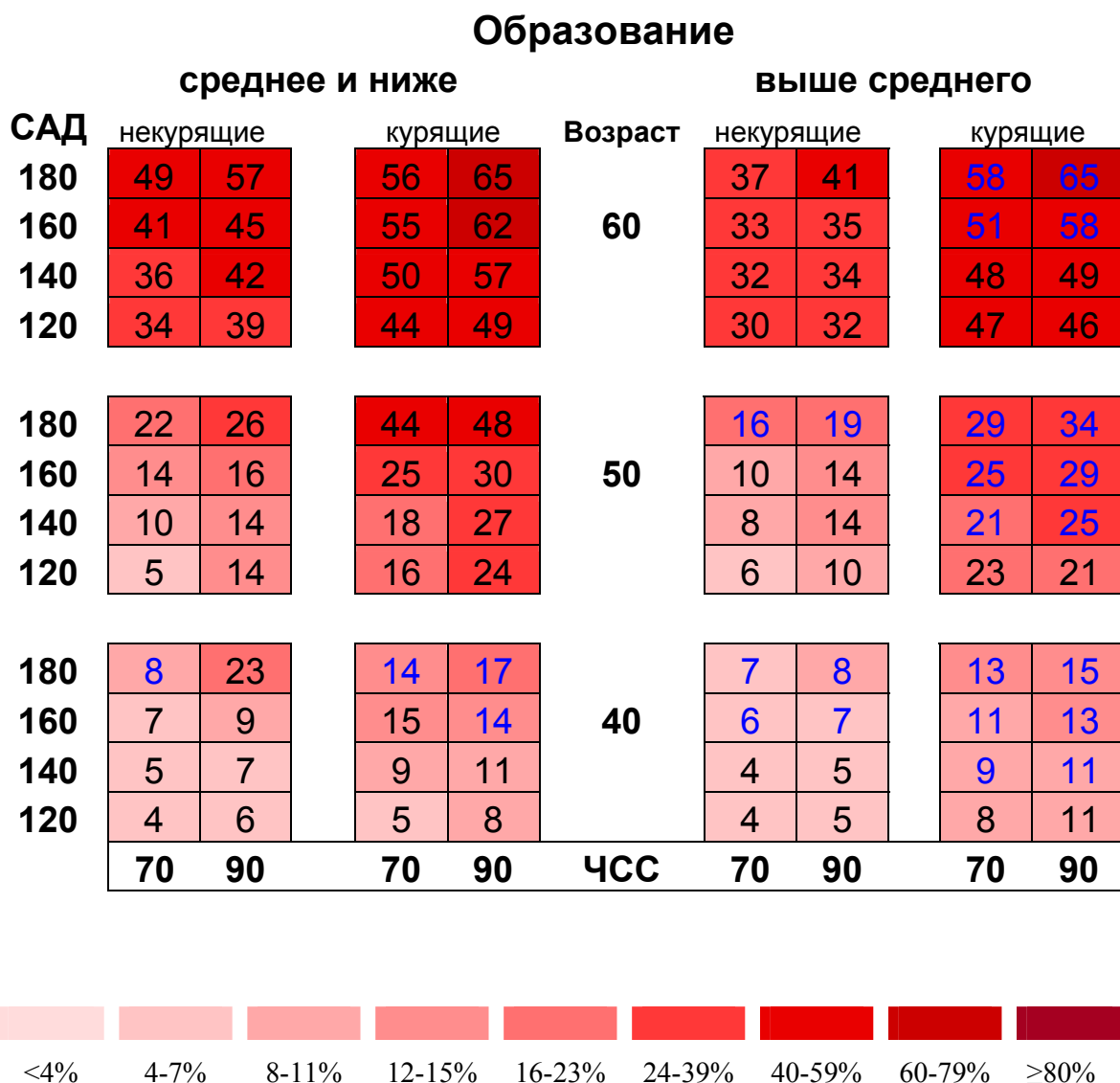


Рис. 15. Двадцатилетний риск смерти от всех причин (%) для женщин (США).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 31).

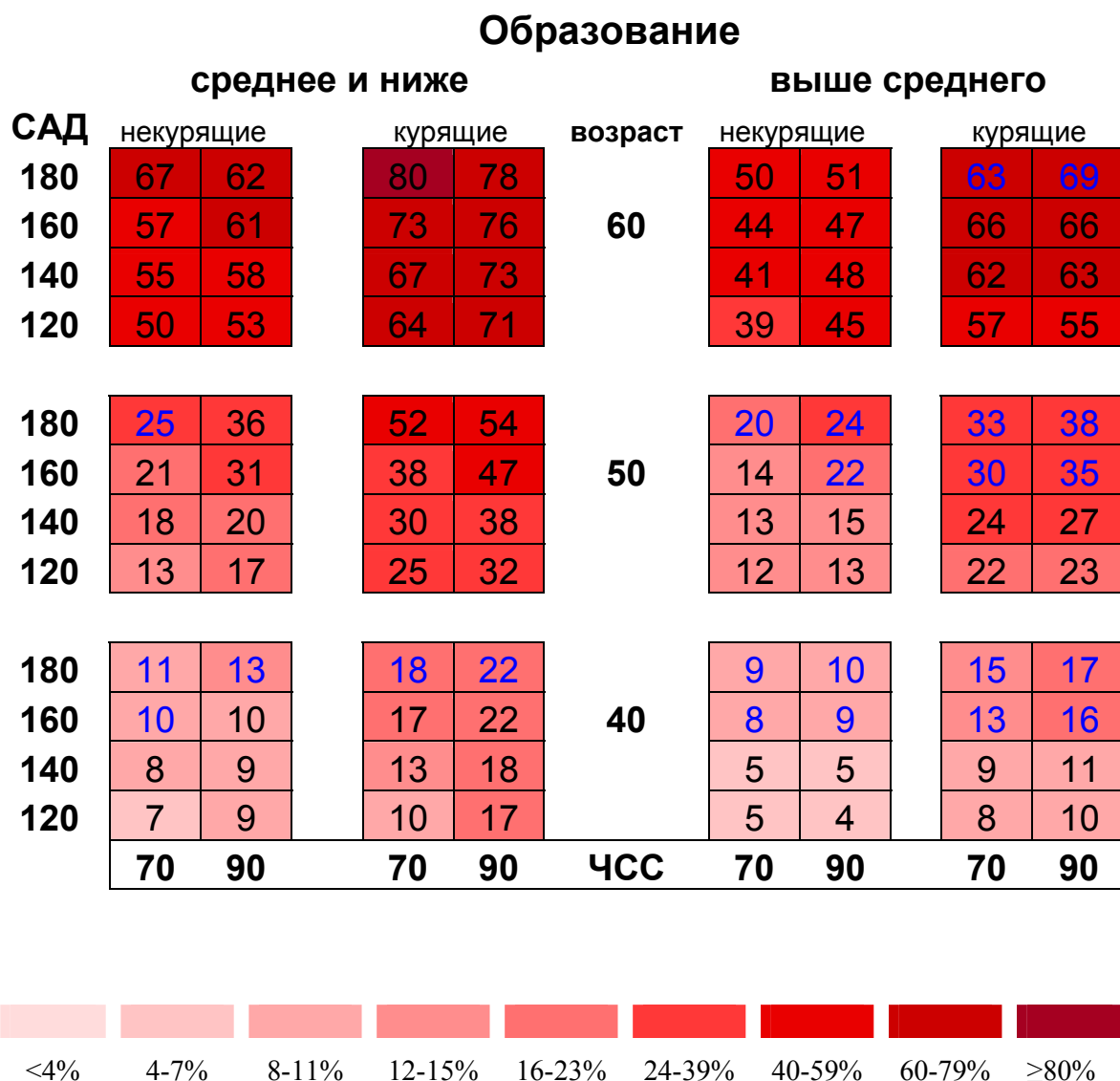


Рис. 16. Двадцатилетний риск смерти от всех причин (%) для мужчин (США).

Риск рассчитан для самоорганизующихся кластеров факторов риска, включающих возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения, уровень образования.

В ячейках таблицы, для которых число наблюдений в выборке не превышает 10, величины риска выделены шрифтом синего цвета, в этих случаях риск рассчитывали по регрессионной модели (см. рис. 32).

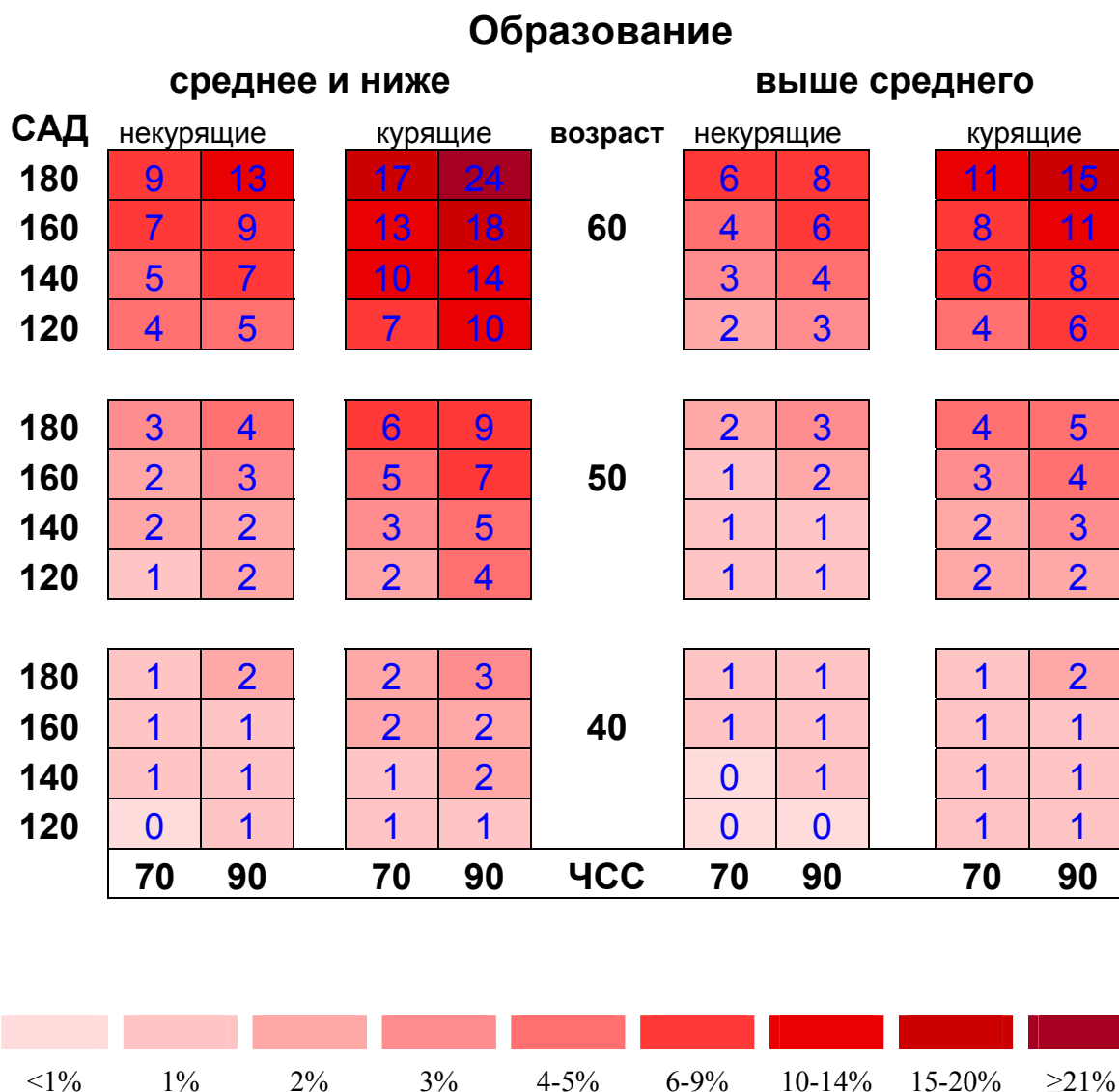


Рис. 17. Десятилетний риск смерти от ССЗ (%) для женщин (Россия).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

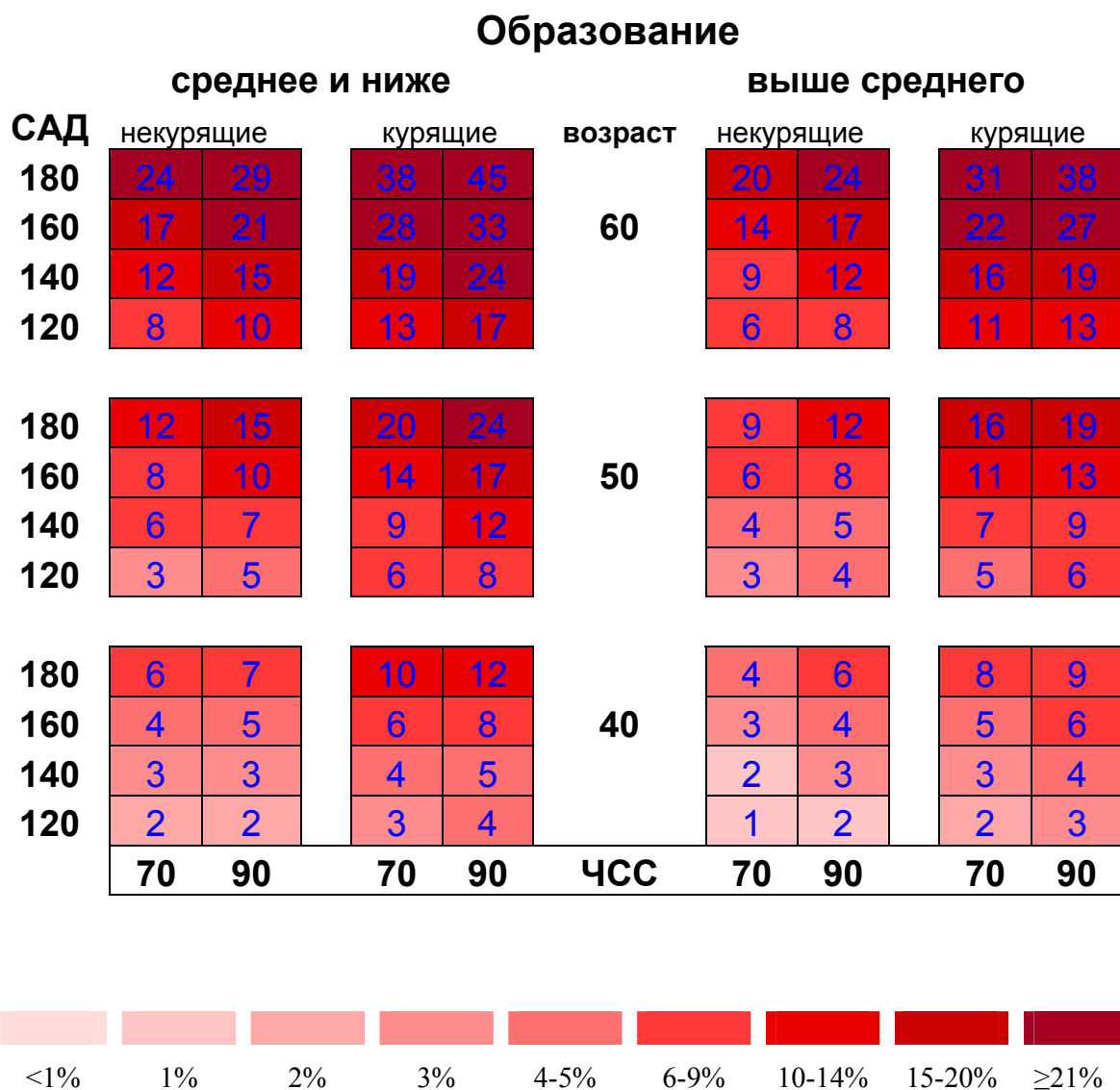


Рис. 18. Десятилетний риск смерти от ССЗ (%) для мужчин (Россия).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

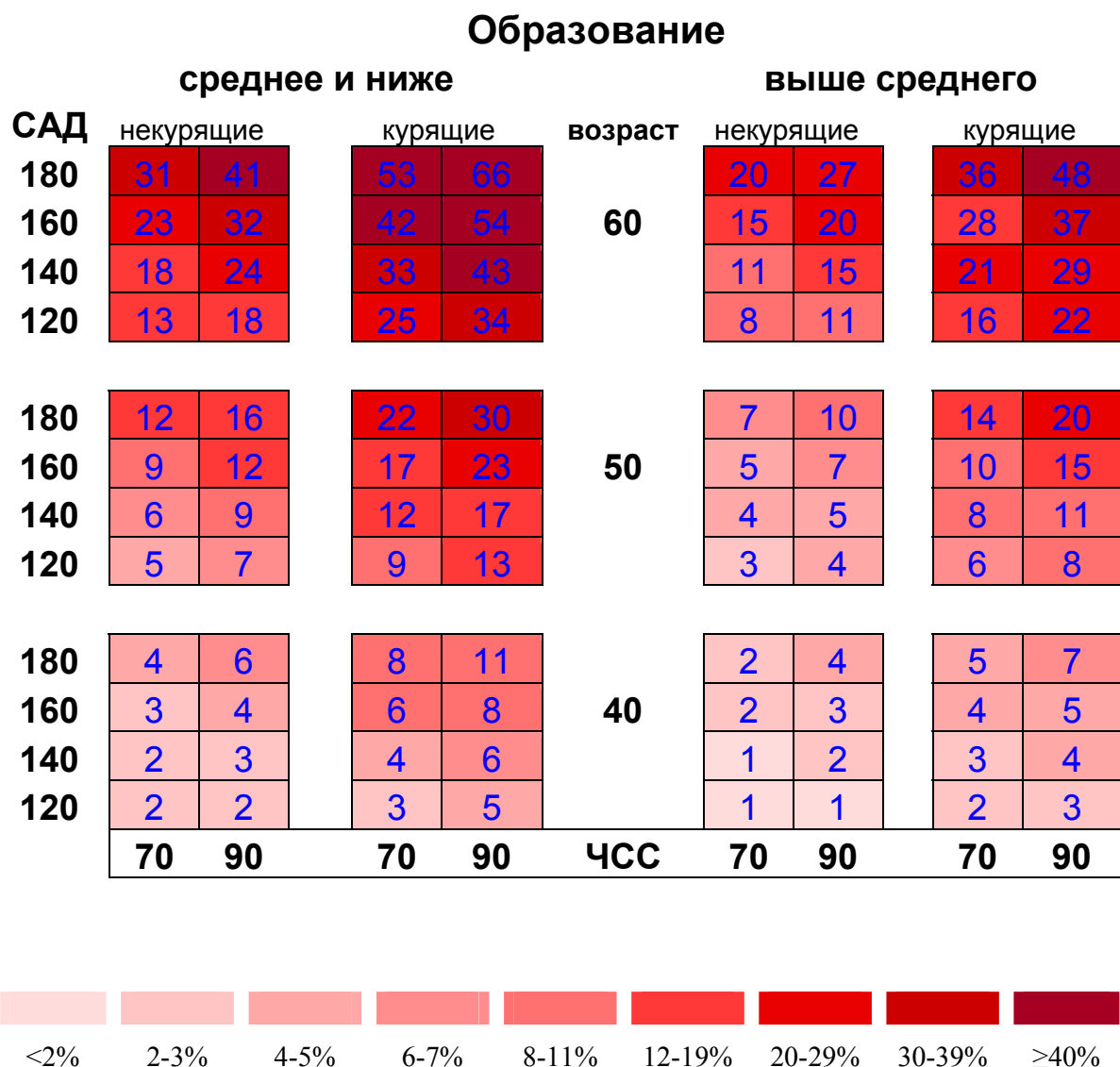


Рис. 19. Двадцатилетний риск смерти от ССЗ (%) для женщин (Россия).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

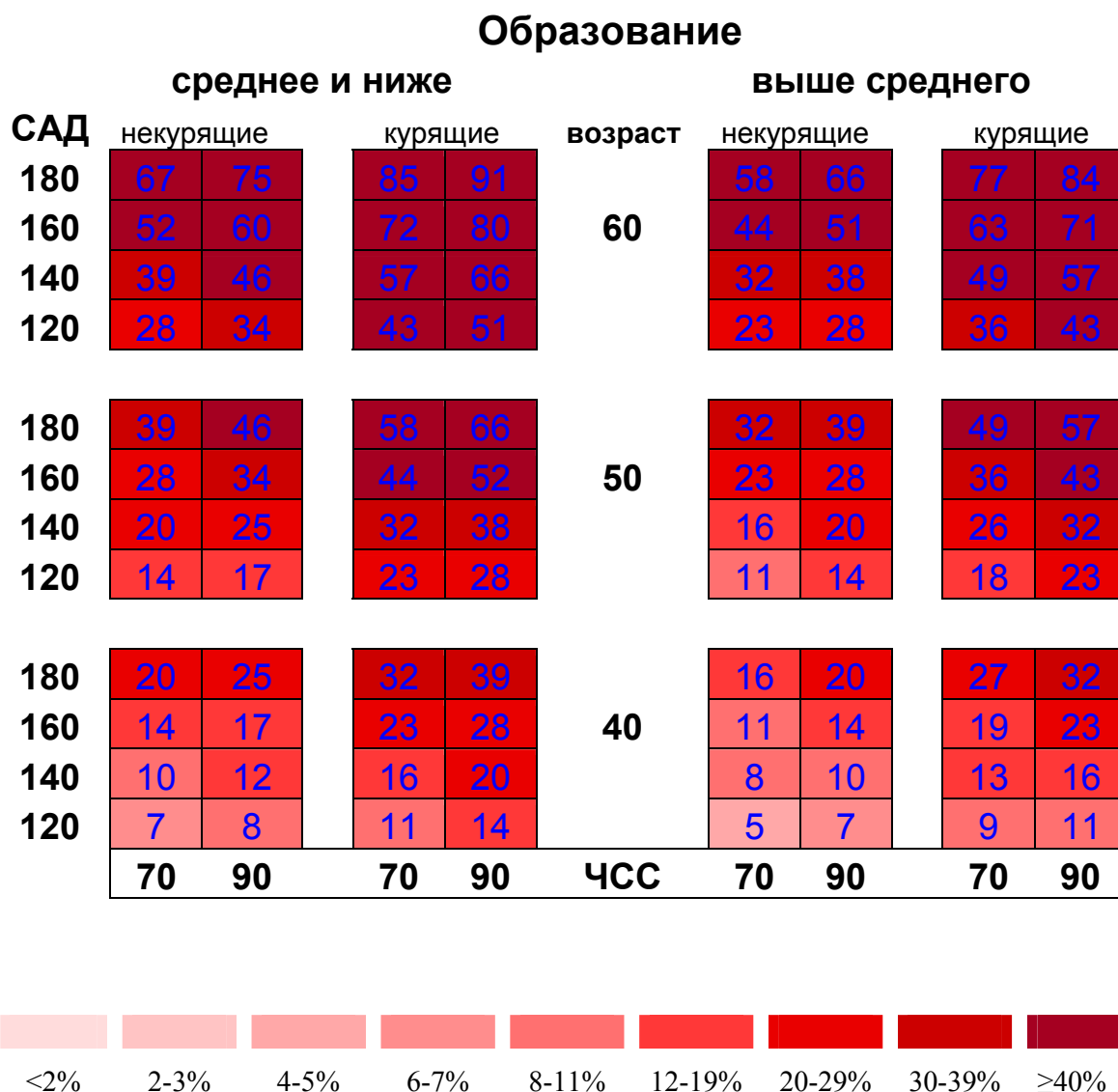


Рис. 20. Двадцатилетний риск смерти от ССЗ (%) для мужчин (Россия).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.



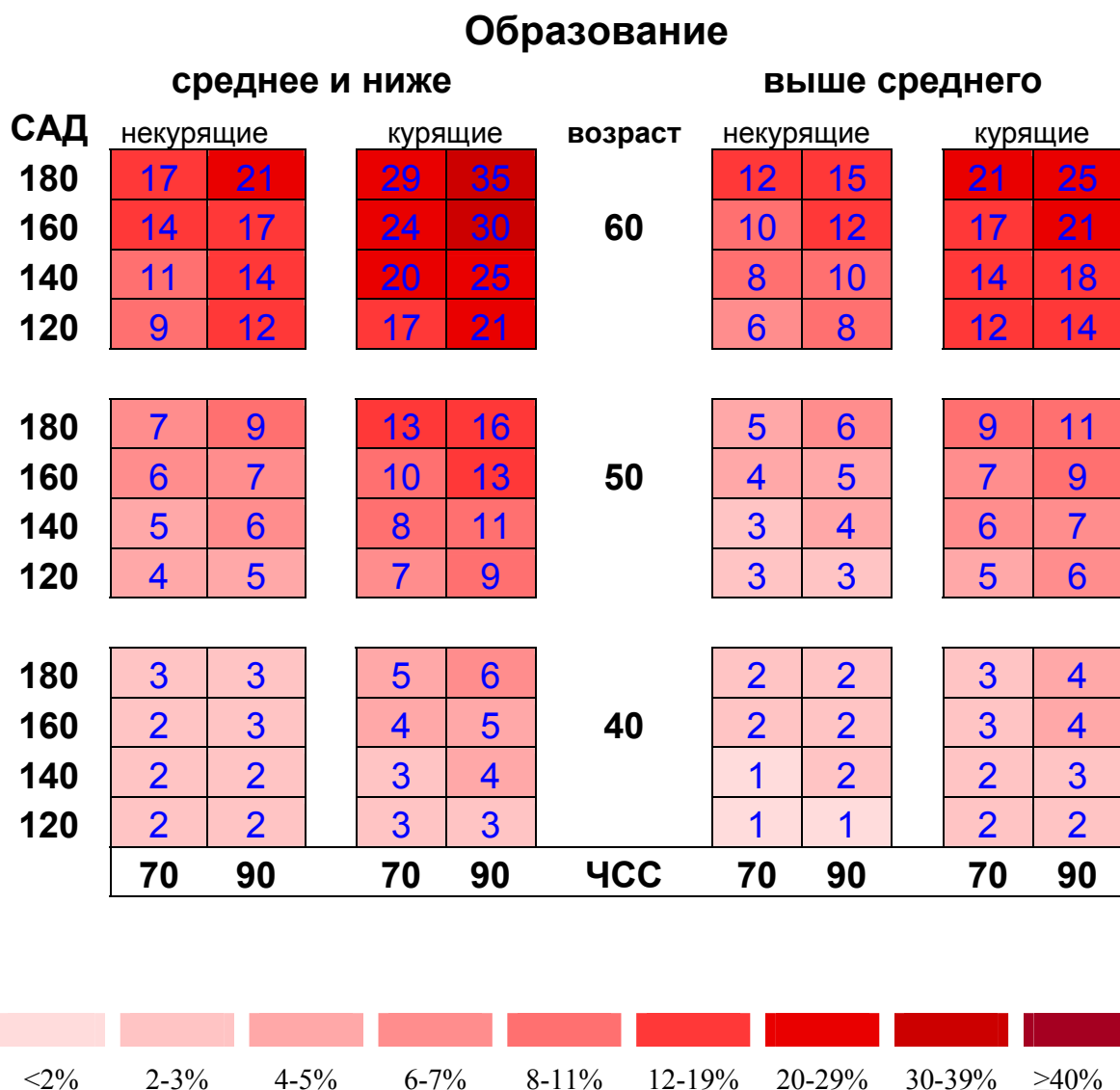


Рис. 21. Десятилетний риск смерти от всех причин (%) для женщин (Россия).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

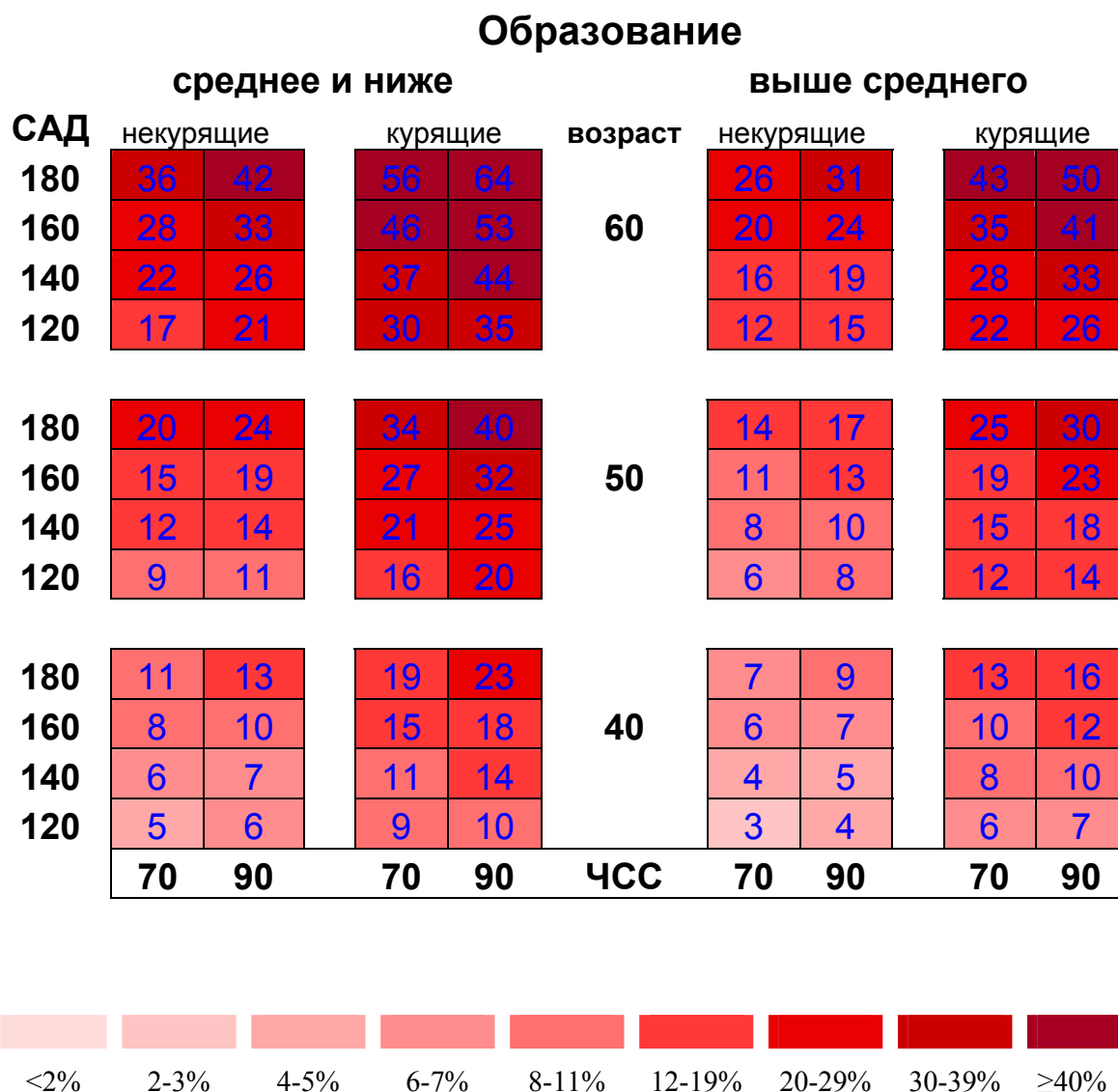


Рис. 22. Десятилетний риск смерти от всех причин (%) для мужчин (Россия).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

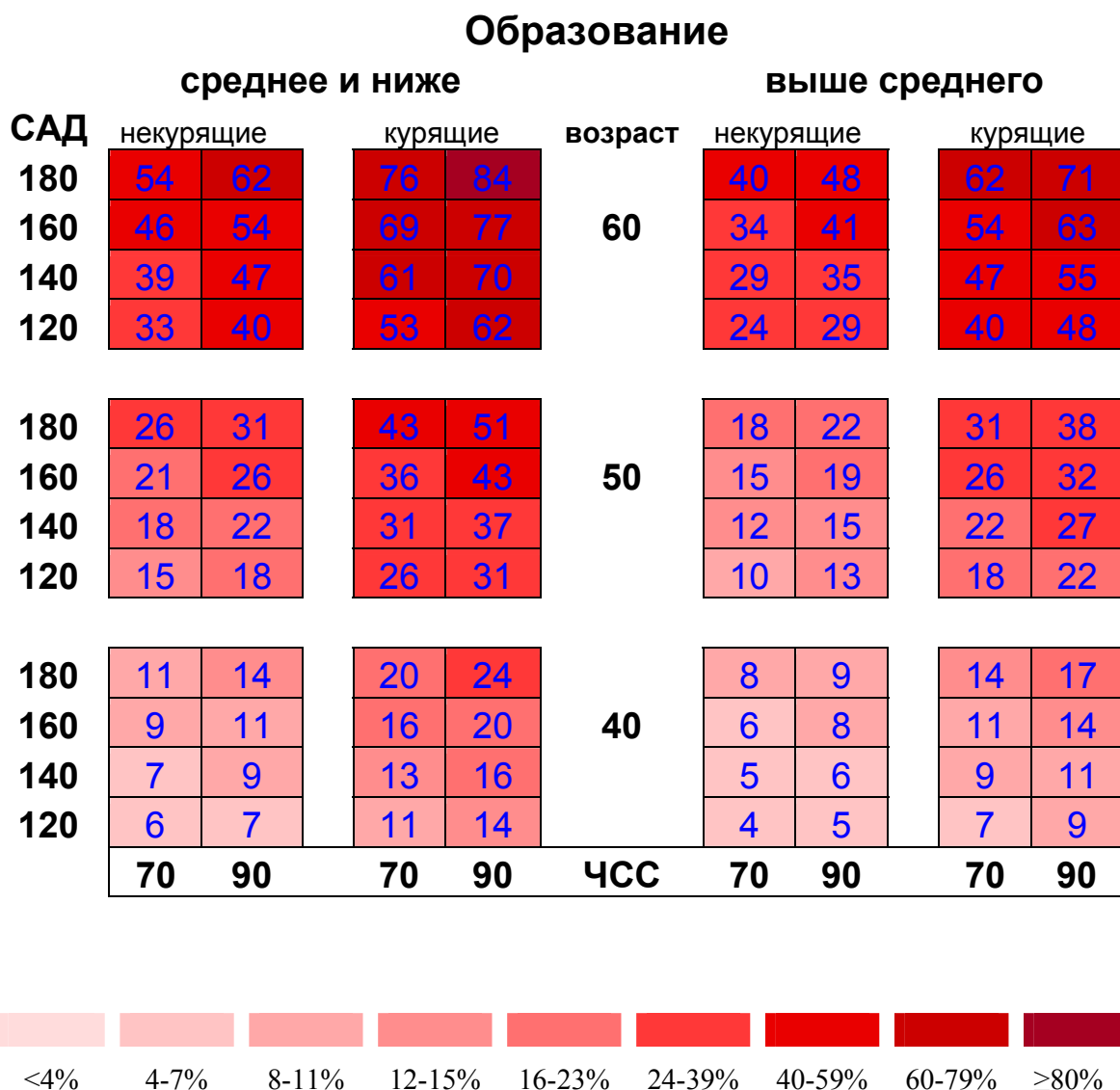


Рис. 23. Двадцатилетний риск смерти от всех причин (%) для женщин (Россия).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

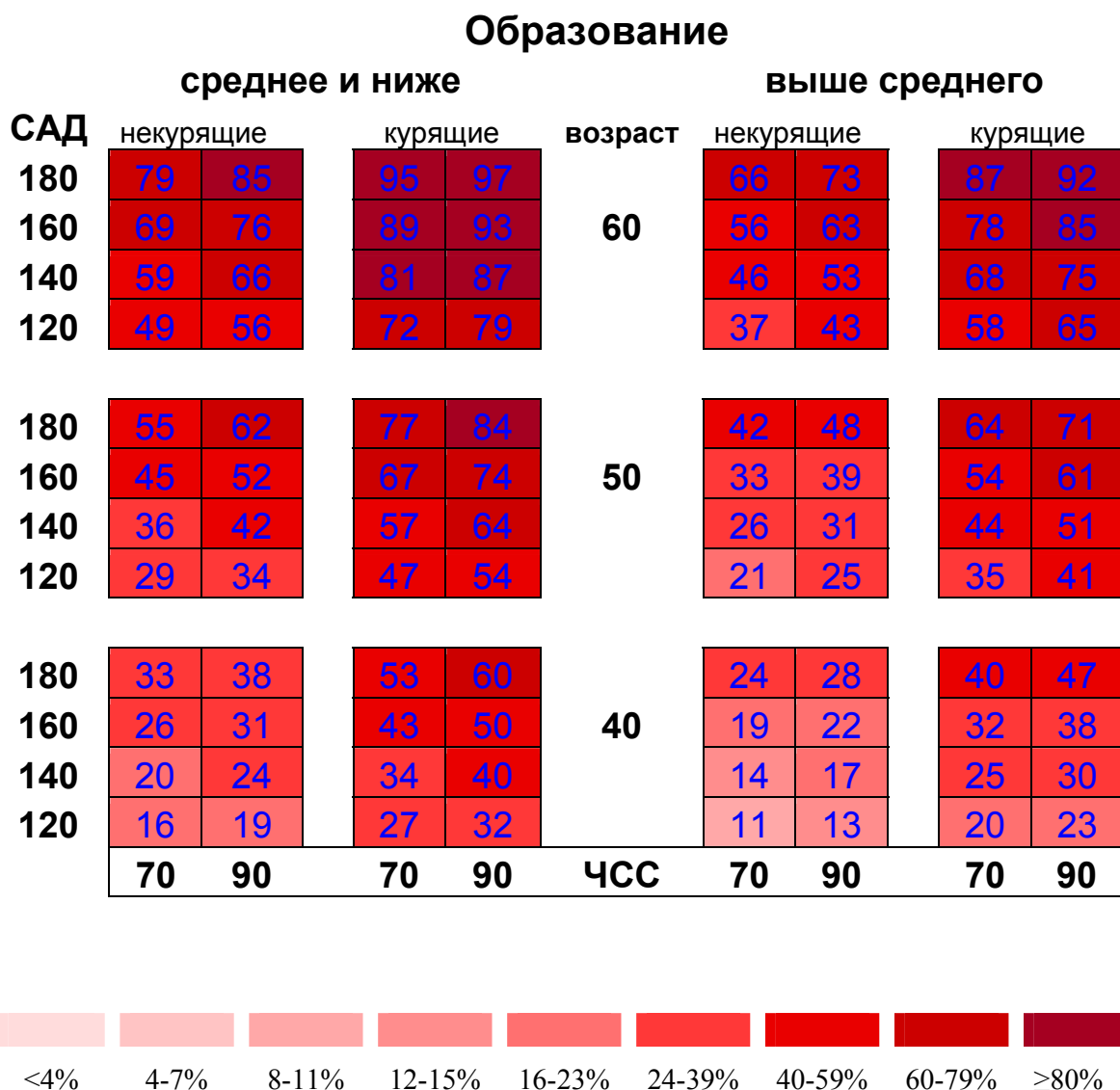


Рис. 24. Двадцатилетний риск смерти от всех причин (%) для мужчин (Россия).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

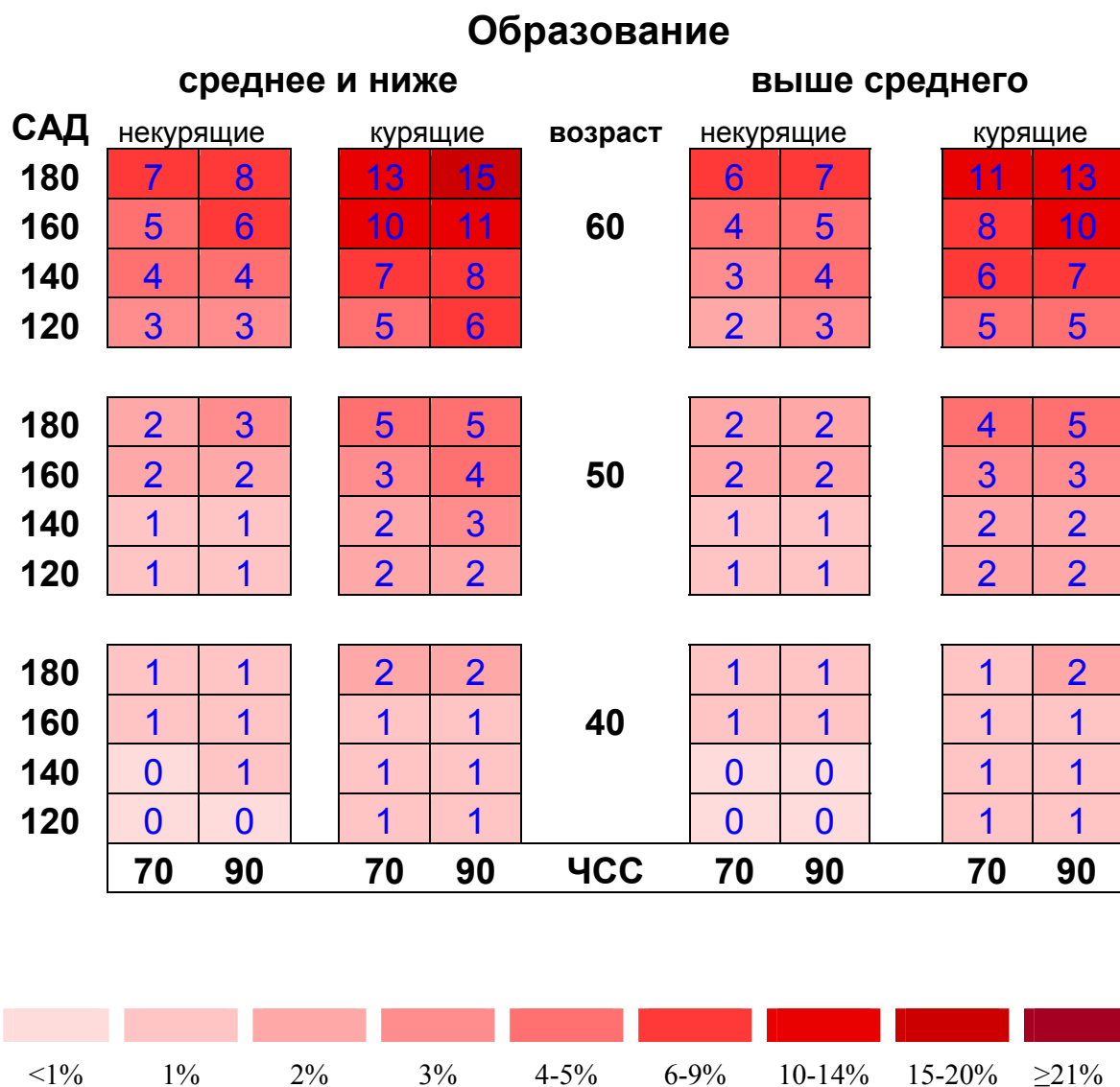


Рис. 25. Десятилетний риск смерти от ССЗ (%) для женщин (США).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

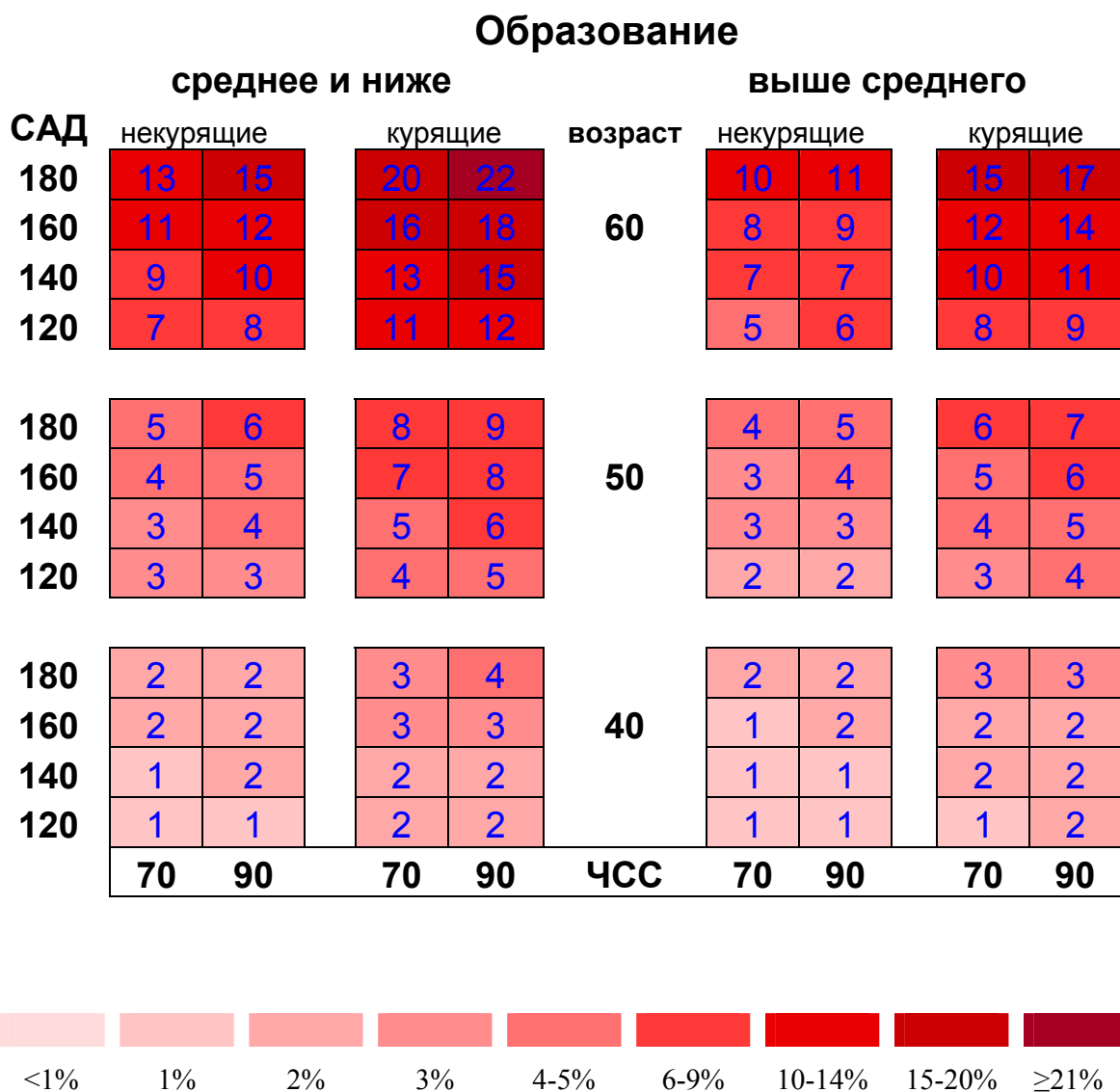


Рис. 26. Десятилетний риск смерти от ССЗ (%) для мужчин (США).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

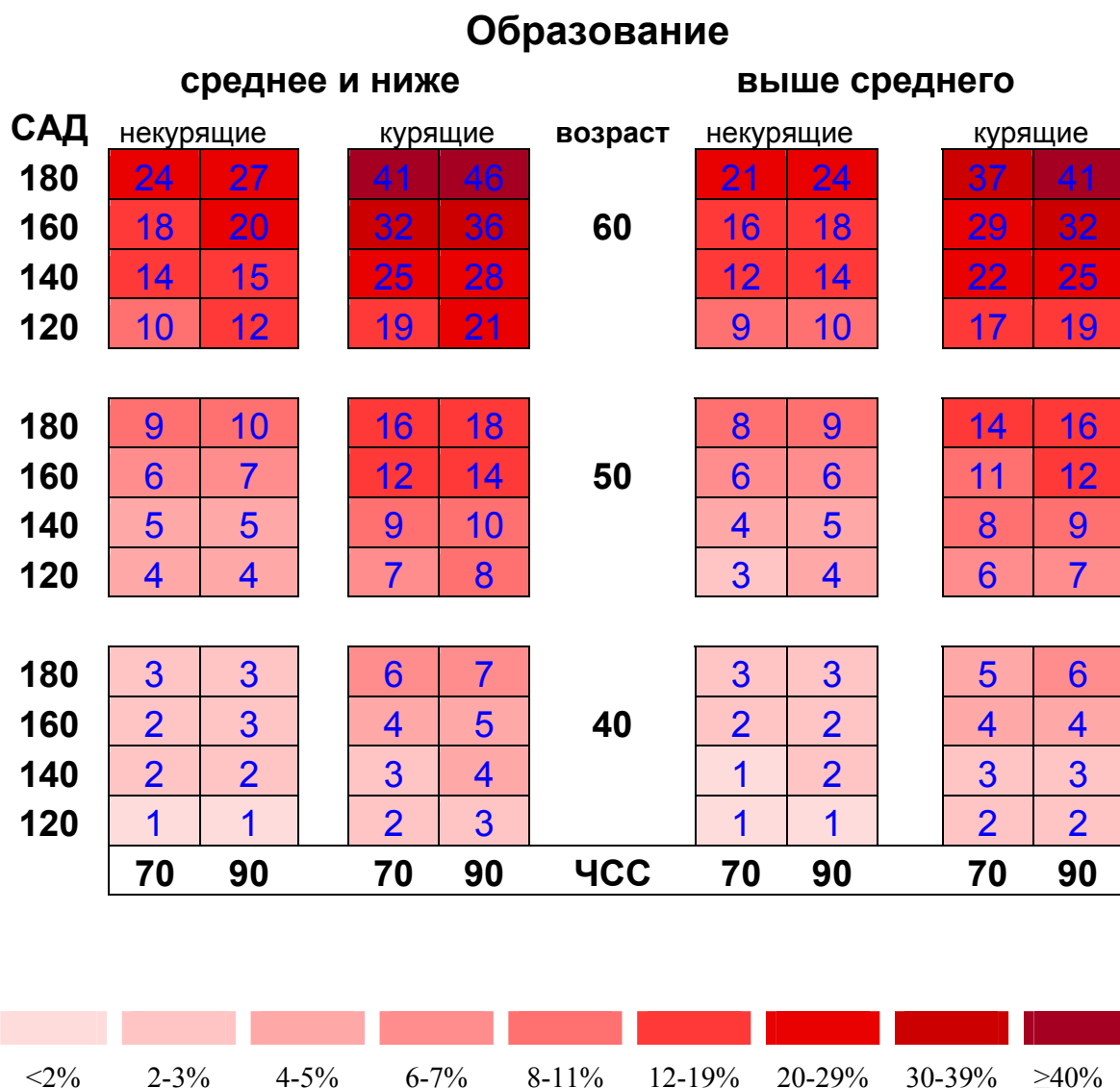


Рис. 27. Двадцатилетний риск смерти от ССЗ (%) для женщин (США).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

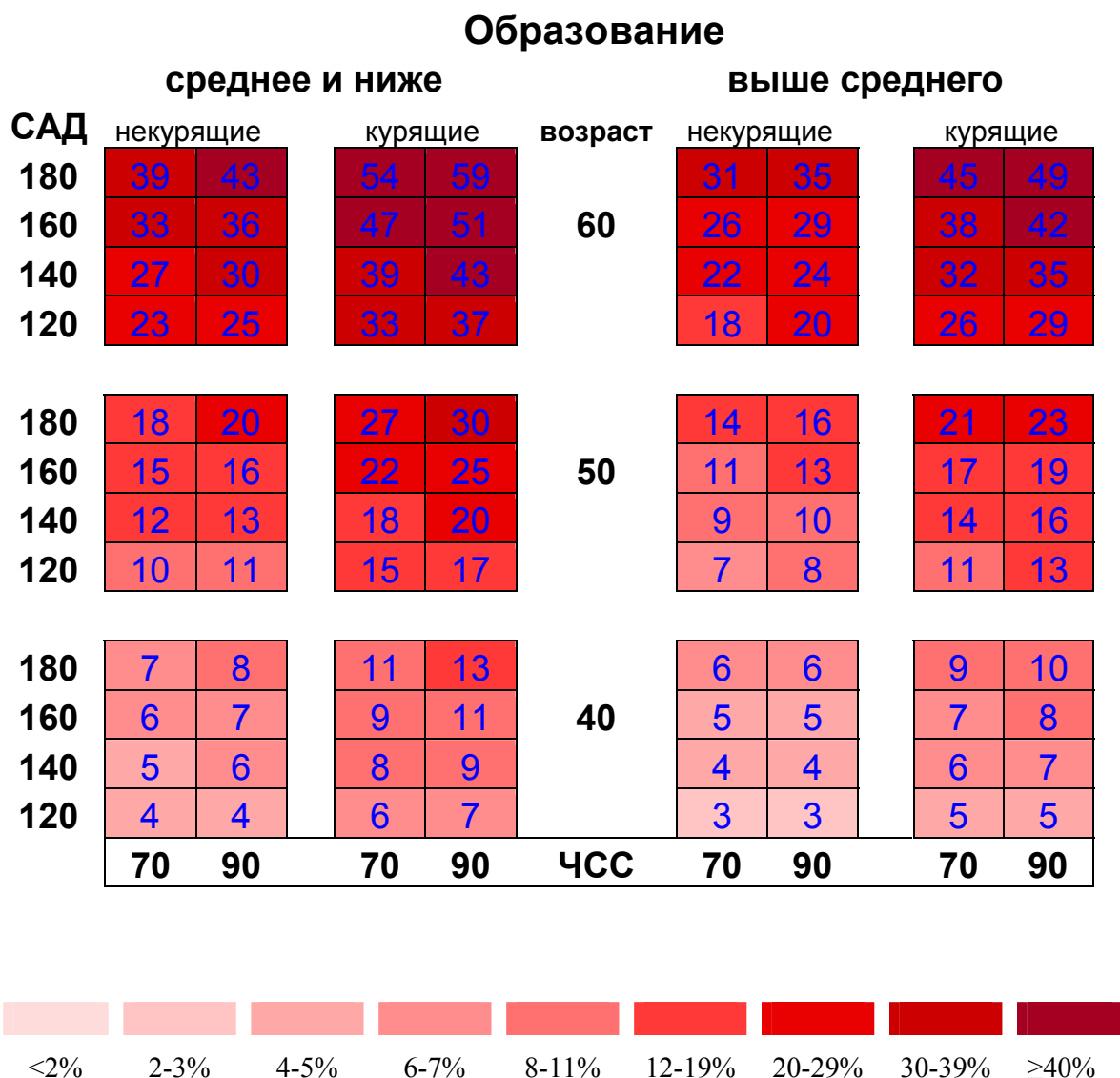


Рис. 28. Двадцатилетний риск смерти от ССЗ (%) для мужчин (США).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.



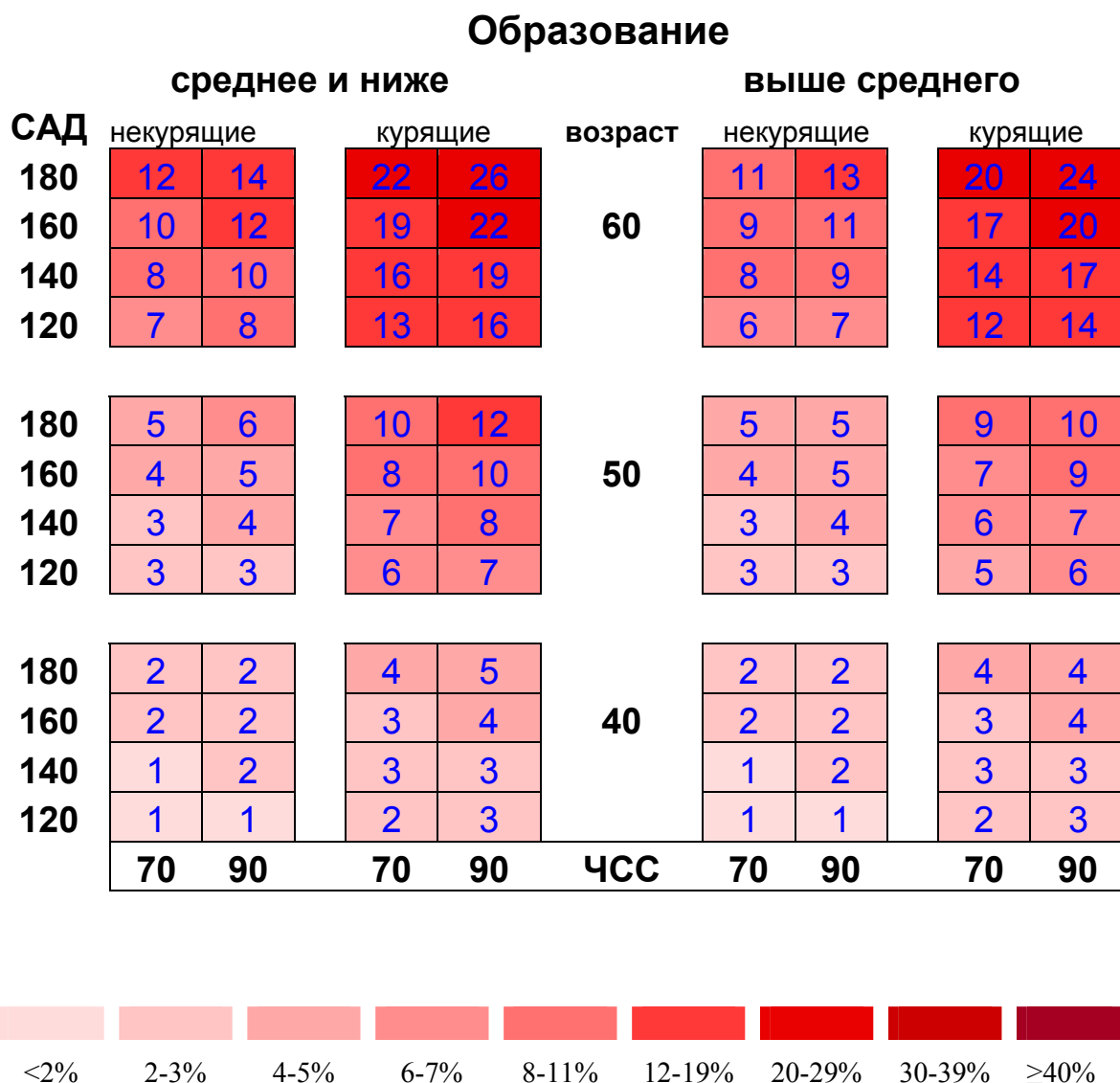


Рис. 29. Десятилетний риск смерти от всех причин (%) для женщин (США).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

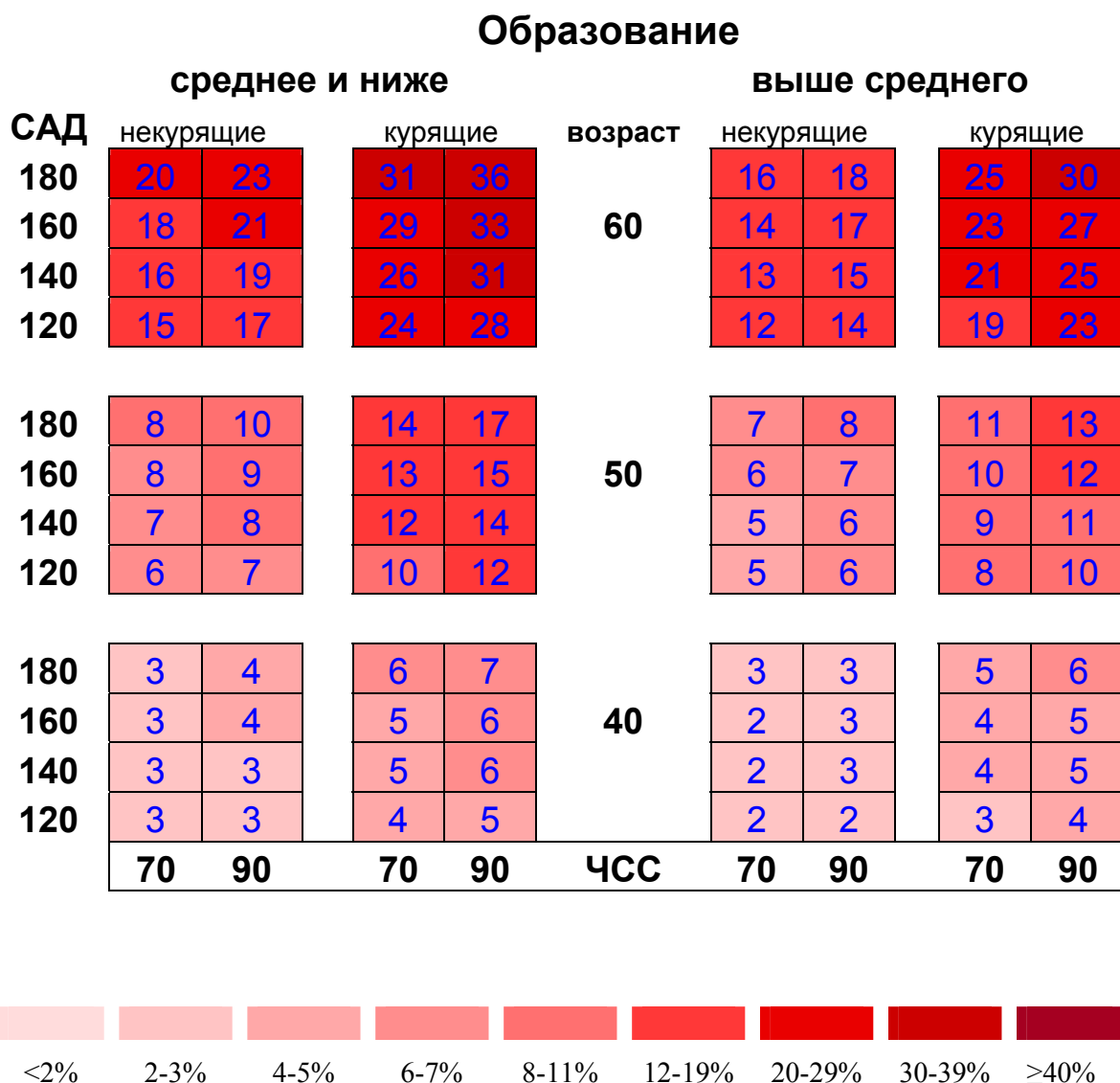


Рис. 30. Десятилетний риск смерти от всех причин (%) для мужчин (США).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

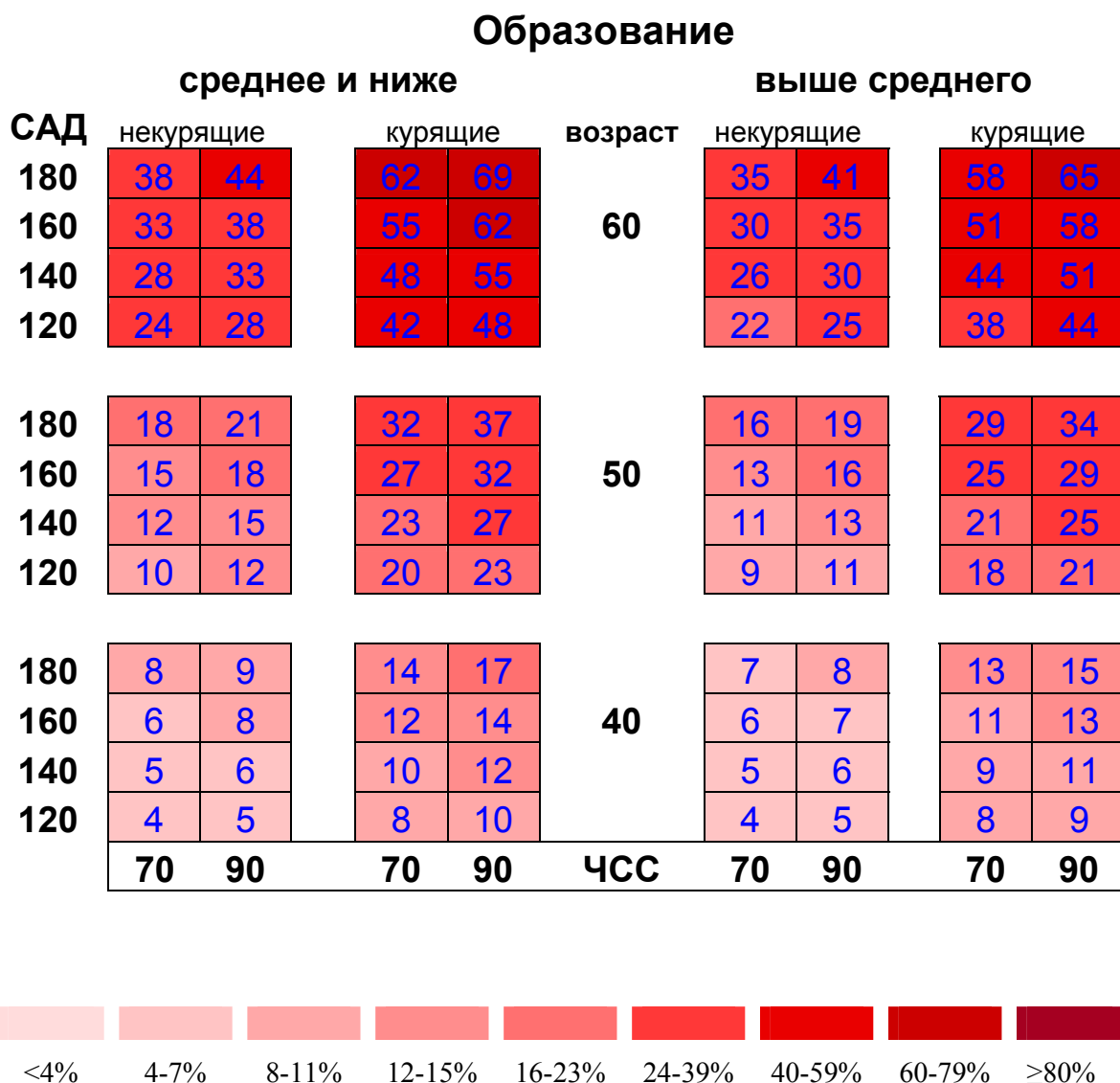


Рис. 31. Двадцатилетний риск смерти от всех причин (%) для женщин (США).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

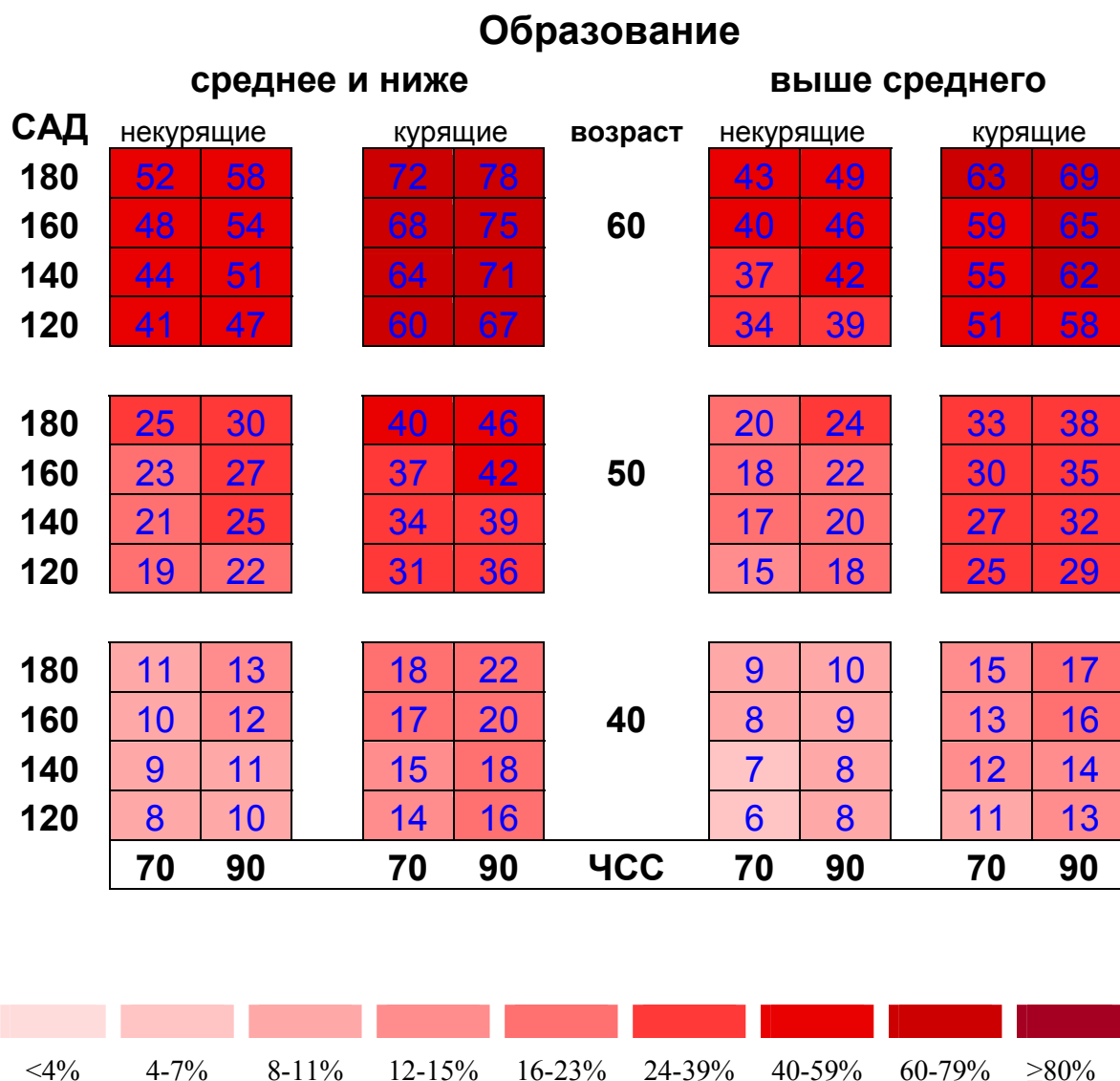


Рис. 32. Двадцатилетний риск смерти от всех причин (%) для мужчин (США).

Риск рассчитан по регрессионной модели, включающей независимые переменные: возраст, величины САД (мм рт.ст.) и ЧСС (уд./мин), статус курения и уровень образования.

## СПИСОК СОКРАЩЕНИЙ

|                   |  |
|-------------------|--|
| АД                | артериальное давление  |
| ИБС               | ишемическая болезнь сердца   |
| ИМТ               | индекс массы тела (Кетле)  |
| ИНС               | искусственная(ые) нейронная(ые) сеть(и)  |
| МКБ-8             | Международная классификация болезней 8 пересмотра  |
| МКБ-9             | Международная классификация болезней 9 пересмотра  |
| ОХ                | общий холестерин   |
| САД               | систолическое артериальное давление  |
| СОК ФР            | самоорганизующийся(еся) кластер(ы) факторов риска  |
| ССЗ               | сердечно-сосудистые заболевания  |
| ФА <sub>1</sub>   | уровень физической активности в обычных условиях   |
| ФА <sub>2</sub>   | уровень физической активности в период отпуска   |
| ФА <sub>общ</sub> | обобщенная оценка физической активности  |
| ФР                | фактор(ы) риска  |
| ЧСС               | частота сердечных сокращений   |
| <i>M</i>          | средняя арифметическая величина  |
| <i>n</i>          | число наблюдений в группе (подгруппе)  |
| NCHS              | the National Center for Health Statistics  |
| NH2MS             | NHANES II Mortality Study  |
| NHANES I          | the First National Health and Nutrition Examination Survey   |
| NHANES II         | the Second National Health and Nutrition Examination Survey  |
| NHEFS             | NHANES I Epidemiologic Followup Study  |
| <i>p</i>          | уровень значимости, вероятность $\alpha$ -ошибки (I рода), при $p < 0,0001$ в таблицах приведена величина 0,0001 |
| SCORE             | Systematic Coronary Risk Evaluation  |
| <i>SD</i>         | стандартное отклонение   |
| SOM               | Self-Organizing Maps   |
| <i>t</i>          | критерий Стьюдента   |
| <i>U</i>          | критерий Манна-Уитни   |

## СЛОВАРЬ ТЕРМИНОВ

**АГРЕГИРОВАНИЕ ИНФОРМАЦИИ** – представление информации в сжатом, укрупненном виде. Агрегирование осуществляется с помощью группировок, а также методов многомерного статистического анализа (кластерного анализа, факторного анализа и т. п.) и позволяет перейти к новому языку описания, например, от описания объектов к описанию типов (классов) объектов, от описания признаков объектов – к факторам и т. п. [86].

**ИСХОД** – клинически значимое явление, изучаемое в данном исследовании. В настоящей работе это смерть от ССЗ или всех причин [87].

**КЛАСТЕР** (от английского "cluster" – гроздь, скопление) – скопление однотипных объектов [88].

**КРИВАЯ ВЫЖИВАНИЯ (SURVIVAL CURVE)** – кривая, показывающая темпы убывания размера наблюдаемой группы населения начиная с начальной величины, принятой за 100% (или за 1,0), через последовательные промежутки времени до тех пор, пока информация о статусе (жив-умер) членов группы еще доступна [89, с изменениями]. Например, кумулятивная доля выживших (по Каплану-Мейеру) [33].

**РИСК (RISK)** – вероятность возникновения неблагоприятного исхода в течение конкретного промежутка времени. В настоящей работе это вероятность смерти от ССЗ или всех причин за 10 или 20 лет [89].

**САМООРГАНИЗУЮЩИЕСЯ КАРТЫ КОХОНЕНА (SELF-ORGANIZING MAPS, SOM)** – разновидность искусственных нейронных сетей, реализующая неуправляемое обучение (без учителя) [65].

**САМООРГАНИЗУЮЩИЙСЯ КЛАСТЕР ФАКТОРОВ РИСКА, СОК ФР (SELF-ORGANIZING CLUSTER of RISK FACTORS, SOC RF)** – подгруппа, объединяющая лиц с близкими величинами показателей, характеризующих учтенные ФР. Способ расчета риска в таких кластерах является предметом настоящей работы.

**СВЯЗЬ (ASSOCIATION)** – статистические взаимоотношения между двумя и более событиями, факторами или переменными [89].

**СМЕРТЬ ОТ ВСЕХ ПРИЧИН** – все известные случаи смерти.

**СМЕРТЬ ОТ СЕРДЕЧНО-СОСУДИСТЫХ ЗАБОЛЕВАНИЙ** включает случаи смерти с кодами 390-448 по МКБ-9 [90].

**СТАНДАРТНОЕ (РАБОЧЕЕ) ОПРЕДЕЛЕНИЕ СЛУЧАЯ (CASE DEFINITION)** – набор стандартных критериев, позволяющих отнести конкретное лицо к определенной группе. В настоящей работе это смерть от всех причин либо смерть от ССЗ [89, с изменениями].

**ФАКТОР РИСКА (RISK FACTOR)** – особенности поведения и уклада жизни, генетическая предрасположенность, и/или внешние воздействия, а также их взаимодействия, которые статистически связаны с возникновением неблагоприятного исхода<sup>1</sup> [9, 89, 61, 7].

---

<sup>1</sup> В настоящей работе - смерти от всех причин или смерти от ССЗ.

## СПИСОК ЛИТЕРАТУРЫ

1. <http://www.cdc.gov/nchs/Default.htm>
2. Вилков В.Г. Сердечно-сосудистый риск и дисфункция щитовидной железы (по данным популяционных исследований). М.: Издатель "Гайнуллин", 2004.
3. Kannel W.B., McGee D., Gordon T. A general cardiovascular risk profile: the Framingham Study. Am J Cardiol 1976; 38 (1): 46-51.
4. Европейские рекомендации по профилактике сердечно-сосудистых заболеваний в клинической практике. Кардиоваскулярная терапия и профилактика 2004; 3 (4): 99-112.
5. Pocock S.J., McCormack V., Gueyffier F., et al. A score for predicting risk of death from cardiovascular disease in adults with raised blood pressure, based on individual patient data from randomised controlled trials. Br Med J 2001; 323: 75-81.
6. Elousua R., Marrugat J. Южно-европейский средиземноморский парадокс. Кардиоваскулярная терапия и профилактика 2002; 1 (1): 84-85.
7. Шальнова С.А., Оганов Р.Г., Деев А.Д. Оценка и управление суммарным риском сердечно-сосудистых заболеваний у населения России. Кардиоваскулярная терапия и профилактика 2004; 3 (4): 4-11.
8. <http://www.megaputer.ru/download/book.zip>
9. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология: Основы доказательной медицины. М.: Медиа Сфера, 1998.
10. Константинов В.В., Жуковский Г.С., Тимофеева Т.Н. Ишемическая болезнь сердца, факторы риска и смертность среди мужского населения в связи с уровнем образования. Кардиология 1996; (1): 37-41.
11. Шальнова С.А., Деев А.Д., Шестов Д.Б. Прогностическая оценка эпидемиологических характеристик ишемической болезни сердца. Кардиология 1997; (9): 49-54.
12. Нейронные сети: STATISTICA Neural Networks. М.: Горячая линия - Телеком, 2001.



13. Круглов В.В., Борисов В.В. Искусственные нейронные сети: Теория и практика. М.: Горячая линия - Телеком, 2001.
14. Patterson D. Artificial neural networks. Singapore: Prentice Hall, 1996.
15. Горбань А.Н., Дунин-Барковский В.Л., Кирдин А.Н., и др. Нейроинформатика. Новосибирск: Наука, 1997.
16. Ежов А., Чечеткин В. Нейронные сети в медицине. Открытые системы 1997; (4): 34-37.
17. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика. М.: Мир, 1992.
18. Hakim C. Secondary analysis in social research: A quid to data sources and methods with examples. L.: George Allen & Unwin, 1982.
19. Steward D.W. Secondary Research. Newbury Park, Calif.: Sage, 1984.
20. <http://www.sociology.ru/forum/01-1-4khazeva.html>
21. Мангейм Дж.Б., Рич Р.К. Политология: методы исследования. Издательство "Весь Мир", 1997.
22. Cox C.S., Mussolino M.E., Rothwell S.T., et al. Plan and Operation of the NHANES I Epidemiologic Followup Study 1992. Vital Health Stat 1997; 1 (35).
23. <http://www.cdc.gov/nchs/about/major/nhefs/nhefspuf.htm>
24. Miller H.W. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971-1973. National Center for Health Statistics. Vital Health Stat 1973; 1 (10a,b).
25. Engel A., Murphy R.S., Maurer K., Collins E. Plan and operation of the HANES I Augmentation Survey of Adults 25-74 Years, United States, 1974-75. Vital Health Stat 1978; 1 (14).
26. <http://www.cdc.gov/nchs/about/major/nhanes/nhanesi.htm>
27. Loria C.M., Sempos C.T., Vuong C. Plan and operation of the NHANES II Mortality Study, 1992. National Center for Health Statistics. Vital Health Stat 1999; 1 (38).

28. [http://www.cdc.gov/nchs/r&d/nchs\\_datalinkage/nhanesii\\_data\\_linkage\\_activities.htm](http://www.cdc.gov/nchs/r&d/nchs_datalinkage/nhanesii_data_linkage_activities.htm)
29. National Center for Health Statistics. Plan and operation of the Second National Health and Nutrition Examination Survey, 1976-80. *Vital Health Stat* 1981; (15).
30. <http://www.cdc.gov/nchs/about/major/nhanes/nhanesii.htm>
31. Cox D.R. Regression models and life tables. *J Royal Statistical Society* 1972; 34: 187-220.
32. Кокс Д.Р., Оукс Д. Анализ данных типа времени жизни. М.: Финансы и статистика, 1988.
33. Боровиков В. STATISTICA: искусство анализа данных на компьютере. Для профессионалов. СПб.: Питер, 2001.
34. Janssen I., Katzmarzyk P., Church T.S., et al. The Cooper Clinic Mortality Risk Index: Clinical Score Sheet for Men. *Am J Prev Med* 2005; 29: 194-203.
35. Stoks J., Kannel W.B., Wolf P.A., et al. Blood pressure as a risk factor for cardiovascular disease: The Framingham study - 30 years of follow-up. *Hypertension* 1989; 13 (suppl. I): i13-i18.
36. MacMahon S., Peto R., Cutler J., Collins R. Blood pressure, stroke and coronary heart disease. Part 1. Prolong differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet* 1990; 335 (8692): 765-774.
37. Stamler J., Stamler R., Neaton J.D. Blood pressure, systolic and diastolic, and cardiovascular risks: US population data. *Arch Intern Med* 1993; 153: 598-615.
38. Prospective Studies Collaboration: Cholesterol, diastolic blood pressure and strokes in 450000 people in 45 prospective cohorts. *Lancet* 1995; 346: 1647-1653.
39. Keys A. Seven countries a multivariate analysis of death and coronary heart disease. Cambridge: Harvard University Press, 1980.
40. Lichtenstein M.J., Shipley M.J., Rose I.I. Systolic and diastolic blood pressure as a predictors of coronary heart disease mortality in the Whitehall study. *Br Med J* 1985; 291: 243-245.

41. Anderson W.T. Re-examination of some of the Framingham blood pressure data. *Lancet* 1994; (1139).
42. Шальнова С.А., Деев А.Д., Оганов Р.Г., и др. Роль систолического и диастолического артериального давления для прогноза смертности от сердечно-сосудистых заболеваний. *Кардиоваскулярная терапия и профилактика* 2002; (1): 10-15.
43. Castelli W.P., Levy D., Wilson P.W.F., et al. Sudden death: The view from Framingham. *The Prevention of Sudden Death*. New York: E Wiley-Liss, 1990.
44. Dyer A.R., Persky V., Stamler J. Heart rate as a prognostic factor for coronary heart disease and mortality: findings of three Chicago epidemiologic studies. *Am J Epidemiol* 1980; 112: 736-749.
45. Kannel W.B., Kannel C., Paffenbarger R.S.Jr., Cupples L.A. Heart rate and cardiovascular mortality: the Framingham Study. *Am Heart J* 1987; 113 (6): 1489-1494.
46. Gillum R.F. The epidemiology of resting heart rate in a national sample of men and women: associations with hypertension, coronary heart disease, blood pressure and other cardiovascular risk factors. *Am Heart J* 1988; 116: 163-174.
47. Gillum R.F., Makuc D.M., Feldman J.J. Pulse rate, coronary heart disease and death: the NHANES 1 epidemiologic follow-up study. *Am Heart J* 1991; 121 (1 - Pt 1): 172-177.
48. Shaper A.G., Wannamethee G., Macfarlane P.W., et al. Heart rate, ischemic heart disease and sudden cardiac death in middle-aged British men. *Br Heart J* 1993; 70: 49-55.
49. Kristal-Boneh E., Silber H., Harari G., et al. The association of resting heart rate with cardiovascular, cancer and all-cause mortality: Eight-year follow-up of 3527 male Israeli employees (the CORDIS Study). *Eur Heart J* 2000; 21: 116-124.
50. Mensink G.B., Hoffmeister H. The relationship between resting heart rate and all-cause, cardiovascular and cancer mortality. *Eur Heart J* 1997; 18 (9): 1404-1410.
51. Шальнова С.А., Деев А.Д., Оганов Р.Г., и др. Частота пульса и смертность от сердечно-сосудистых заболеваний у российских

- мужчин и женщин. Результаты эпидемиологического исследования. Кардиология 2005; (10): 45-50.
52. Wilhelmsen I., Berglund G., Elmfeldt D., et al. The multi factor primary prevention trial in Goteborg, Sweden. Eur Heart J 1986; 7: 279-288.
  53. Habib G.B. Is heart rate a risk factor in the general population? Dialog in Cardiovasc Med 2001; 6: 25-31.
  54. Маколкин В.И., Зябрев Ф.Н. Может ли частота сердечных сокращений рассматриваться в качестве фактора риска сердечно-сосудистых заболеваний. Кардиоваскулярная терапия и профилактика 2006; 5 (6): 5-9.
  55. Schmidt-Nelsen K. Animal Phisiology: Adaptation and Environment. New York: Cambridge University Press, 1975.
  56. Boraso A. Why is reduced heart rate beneficial? Dialog in Cardiovasc Med 2001; 6: 19-24.
  57. White P.D., Jenkins J.L., Benedict F.G. The electrocardiogram of the elephant. Am Heart J 1938; 16: 744.
  58. Spector W.S. Handbook of Biological Data. Philadelphia: WB Saunders, 1956.
  59. Levine H.J. Rest heart rate and life expectancy. J Am Coll Cardiol 1997; 30 (4): 1104-1106.
  60. Маколкин В.И., Зябрев Ф.Н. Частота сердечных сокращений как прогностический фактор у больных сердечно-сосудистыми заболеваниями. Consilium Medicum 2006; 8 (5).
  61. Власов В.В. Введение в доказательную медицину. М.: Медиа Сфера, 2001.
  62. <http://www.smolensk.ru/user/sgma/MMORPH/N-7-html/LEONOV/paradigma.htm>
  63. Hall J.C. Use of the t test in the British Journal of Surgery. Br J Surg 1982; 69 (1): 55-56.
  64. Бююль А., Цёфель П. SPSS: искусство обработки информации: Анализ статистических данных и восстановление скрытых закономерностей. СПб.: ООО «ДиаСофтЮП», 2002.

65. Kohonen T. Self-Organizing Maps. Heidelberg: Springer-Verlag, 1995.
66. Mueller J.-A., Lemke F. Self-Organising Data Mining. Hamburg: Libri, 2000.
67. Медведев В.С., Потемкин В.Г. Нейронные сети: MATLAB 6. М.: ДИАЛОГ-МИФИ, 2002.
68. <http://www.cis.hut.fi/research/som-bibl/>
69. <http://www.cis.hut.fi/research/som-research/nnrc-programs.shtml>
70. [http://www.cis.hut.fi/research/som\\_pak/](http://www.cis.hut.fi/research/som_pak/)
71. <http://www.cis.hut.fi/research/som-research/nnrc-programs.shtml>
72. <http://www.orc.ru/%7estasson/n3.zip>
73. SAS/STAT user's guide, version 6.04th ed. Cary: SAS Institute Inc., 1993.
74. Persky V., Dyer A.R., Leonas J. Heart Rate: a risk factor for cancer? *Am J Epidemiol* 1981; 114: 177-187.
75. Pytlak A., Piotrowski W., Broda G. Relationship between heart rate and general mortality and that caused by cardiovascular diseases. *Przegl Lek* 1997; 54 (1): 25-30.
76. Gillman M.W., Kannel W.B., Belanger A. Influence of heart rate on mortality among person with hypertension: the Framingham Study. *Am Heart J* 1993; 125: 1148-1154.
77. Shalnova S., Plavinskaya S., Williams D. The association of heart rate and mortality in Russian and US men aged 40-59. *Cardiovasc Risk Factors* 1997; 7 (1): 36-42.
78. Kaplan P.A., Manuck S.B., Clarkson T.B. The influence of heart rate on coronary artery atherosclerosis. *J Cardiovasc Pharmacol* 1987; 10 (suppl 2): S100-S102.
79. Beere P.A., Glagov S., Karins C.K. Retardings effects of lowered heart rate on coronary atherosclerosis. *Science* 1984; 226: 180-182.
80. Kaplan P.A., Manuck S.B., Adams M.R. Propranolol inhibits coronary atherosclerosis in behaviorally predisposed monkeys fed an atherogenic diet. *Circulation* 1987; 761: 1364-1372.

81. Дембо А.Г., Земцовский Э.В. Спортивная кардиология. Л.: Медицина, 1989.
82. Hodgson J.M., Puddey I.B., Burke V., et al. Effects on blood pressure of drinking green and black tea. J Hypertens 1999; 17 (4): 457-463.
83. Corti R., Binggeli C., Sudano I., et al. Coffee acutely increases sympathetic nerve activity and blood pressure independently of caffeine content: role of habitual versus nonhabitual drinking. Circulation 2002; 106 (3): 2935-2940.
84. Алешин С. Гипертония: ответный удар. Немедикаментозная программа, основанная на доказательствах. М.: «Орто.ру», 2004.
85. [http://rql.net.ua/cardio\\_j/2005/4D/sychov3.htm](http://rql.net.ua/cardio_j/2005/4D/sychov3.htm)
86. <http://www.uo01.mnogosmenka.ru/uo010040/uo010047.htm>
87. Краткий словарь терминов. Международный журнал медицинской практики 2005; (2).
88. <http://building.bk.kz/archiv/280305/8.shtml>
89. <http://www.doktor.ru/doctor/epid/diker/slovar.pdf>
90. <http://www.cdc.gov/nchs/data/dvs/2amanual.pdf>

## ОГЛАВЛЕНИЕ

|   |    |
|---|----|
| ПРЕДИСЛОВИЕ .....   | 3  |
| ВВЕДЕНИЕ. <i>С.А. Шальнова, В.Г. Вилков, А.Д. Деев</i> .....  | 4  |
| 1. ИСТОЧНИКИ ДАННЫХ И МЕТОДЫ<br>СТАТИСТИЧЕСКОГО АНАЛИЗА. <i>В.Г. Вилков, А.Д. Деев</i> .....  | 9  |
| 1.1. Данные .....   | 9  |
| 1.2. Обоснование выбора характеризующих ФР показателей..  | 12 |
| 1.3. Обоснование выбора статистических методов .....  | 17 |
| 2. СПОСОБ РАСЧЕТА РИСКА ФАТАЛЬНОГО СОБЫТИЯ В<br>САМООРГАНИЗУЮЩИХСЯ КЛАСТЕРАХ ФАКТОРОВ<br>РИСКА. <i>В.Г. Вилков</i> .....                                    | 23 |
| 3. ТАБЛИЦЫ РИСКА НА ОСНОВЕ ТРАДИЦИОННЫХ<br>РЕГРЕССИОННЫХ МОДЕЛЕЙ. <i>А.Д. Деев, Ф.А. Айду,<br/>В.Г.Вилков</i> .....   | 29 |
| 4. ТОЧНОСТЬ РАСЧЕТА РИСКА НОВЫМ СПОСОБОМ В<br>СРАВНЕНИИ С ТРАДИЦИОННЫМ. <i>В.Г. Вилков</i> .....  | 31 |
| 5. ТАБЛИЦЫ РИСКА НА ОСНОВЕ<br>САМООРГАНИЗУЮЩИХСЯ КЛАСТЕРОВ ФАКТОРОВ<br>РИСКА ДЛЯ РОССИЙСКОЙ КОГОРТЫ. <i>В.Г. Вилков, А.Д.<br/>Деев, С.А. Шальнова</i> ..... | 41 |
| 6. ТАБЛИЦЫ РИСКА НА ОСНОВЕ<br>САМООРГАНИЗУЮЩИХСЯ КЛАСТЕРОВ ФАКТОРОВ<br>РИСКА ДЛЯ КОГОРТЫ США. <i>В.Г. Вилков</i> .....                                      | 50 |
| 6.1. Связь между смертностью и отдельными ФР .....  | 50 |
| 6.2. Роль сочетаний ФР .....  | 55 |
| 7. ПРИМЕРЫ ОЦЕНКИ ПРЕДПОЛАГАЕМОГО ЭФФЕКТА<br>КОРРЕКЦИИ ФАКТОРОВ РИСКА. <i>В.Г. Вилков</i> .....   | 57 |
| 7.1. Мужчины с высоким риском .....   | 57 |
| 7.2. Лица со средним уровнем риска .....  | 58 |

## 8. ВЫЖИВАЕМОСТЬ И НЕКОТОРЫЕ ПОТЕНЦИАЛЬНО ЗНАЧИМЫЕ ДЛЯ СНИЖЕНИЯ ЧАСТОТЫ СЕРДЕЧНЫХ СОКРАЩЕНИЙ ПОВЕДЕНЧЕСКИЕ ФАКТОРЫ.

|   |     |
|---|-----|
| <i>В.Г. Вилков</i> .....                              | 60  |
| 8.1. Физическая активность, ЧСС и выживаемость .....  | 62  |
| 8.2. Потребление кофе и чая, ЧСС и выживаемость ..... | 65  |
| ЗАКЛЮЧЕНИЕ. <i>С.А. Шальнова, В.Г. Вилков</i> .....   | 69  |
| Электронная версия книги. <i>В.Г. Вилков</i> .....    | 75  |
| ТАБЛИЦЫ РИСКА .....                                   | 76  |
| Список сокращений.....                                | 109 |
| Словарь терминов.....                                 | 110 |
| Список литературы.....                                | 112 |