DOI: 10.12737/1024-6177-2019-64-6-5-11

А.И. Горский[1], М.А Максютов[1], К.А. Туманов[1], О.К. Власов[1], Е.В. Кочергина[1], Н.С. Зеленская[1], С.Ю. Чекин[1], С.А. Иванов[1], А.Д. Каприн[2], В.К. Иванов[1]

АНАЛИЗ СТАТИСТИЧЕСКИХ СВЯЗЕЙ СМЕРТНОСТИ ОТ ЗЛОКАЧЕСТВЕННЫХ НОВООБРАЗОВАНИЙ С ДОЗОЙ ОБЛУЧЕНИЯ РАДИОНУКЛИДАМИ НАСЕЛЕНИЯ РЕГИОНОВ, ЗАГРЯЗНЕННЫХ ВСЛЕДСТВИЕ АВАРИИ НА ЧАЭС

1. Медицинский радиологический научный центр им. А.Ф. Цыба, Обнинск. E-mail: gorski@nrer.ru;
2. Национальный медицинский исследовательский центр радиологии Минздрава России, Москва

А.И. Горский – в.н.с., к.т.н.; М.А. Максютов – зав. отделом, к.т.н.; К.А. Туманов – зав. лабораторией, к.б.н.; О.К. Власов – зав. лабораторией, д.т.н.; Е.В. Кочергина – зав. лаб., к.м.н.; Н.С. Зеленская – н.с.; С.Ю. Чекин – зав. лаб.; С.А. Иванов – директор, д.м.н., проф. РАН; А.Д. Каприн – ген. директор, д.м.н., проф., акад. РАН; В.К. Иванов – зам. директора по научной работе, д.т.н., проф., чл.-корр. РАН

Реферат

Цель: Радиационно-эпидемиологическая классификация причин смерти от злокачественных новообразований (ЗНО) населения наиболее загрязненных радионуклидами вследствие аварии на ЧАЭС районов Брянской, Калужской, Тульской и Орловской областей, с использованием метода статистических связей.

Материал и методы: В анализе использованы данные по смертности от ЗНО за период наблюдения 1993–2017 гг., накопленные в Национальном радиационно-эпидемиологическом регистре (НРЭР). Число умерших мужчин за период наблюдения – 30771 человек (5407 смертей от ЗНО), женщин – 29033 человека (3472 смертей от ЗНО). Анализ статистических связей доз облучения и диагнозов причин смерти проведен методами Data Mining, свободными от априорных предположений о вероятностных распределениях доз и времени постановки диагнозов. Для анализа использованы таблицы сопряженности случаев смерти в двух дозовых категориях (группа 0 – до 0,014 Зв; группа 1 – 0,014+ Зв) и в трех возрастных группах (группа 0 – до 17 лет; группа 1 – 18–60 лет; группа 2 – 60+ лет). В анализе использована индивидуальная эффективная доза, накопленная до 1993 г. (около 90 % от дозы, накопленной к 2017 г.). Средняя эффективная доза равна 0,024 Зв.

Результаты: Для населения четырёх наиболее загрязненных вследствие аварии на ЧАЭС областей РФ не выявлено статистически значимых связей причин смерти от ЗНО с дозой облучения ни для всего класса ЗНО, ни для трёхзначных рубрик МКБ-10 внутри этого класса. Статистически значимые связи с дозой облучения наблюдались для отдельных диагнозов причин смерти: среди мужского населения – для ЗНО желудка неуточненной локализации (С16.9) и для ЗНО поджелудочной железы неуточнённое (С25.9); для женского населения – для ЗНО желудка неуточненной локализации (С16.9) и для ЗНО бронхов или легкого неуточненной локализации (С34.9). Эти диагнозы причин смерти в первую очередь должны быть предметом более чувствительного и специфичного радиационно-эпидемиологического анализа, учитывающего возможные мешающие и смещающие факторы.

Заключение: Представленная методика определяет первоочередные направления исследований для более точного радиационно-эпидемиологического анализа зависимостей доза–эффект среди населения, проживающего на территориях, загрязнённых радионуклидами вследствие аварии на ЧАЭС. Методика и результаты её использования могут быть полезны в работе экспертных советов по установлению причинной связи заболевания (смерти) с воздействием техногенных факторов.

Ключевые слова: *население, авария на ЧАЭС, загрязнение радионуклидами, смертность, злокачественные новообразования, доза облучения, радиационный риск, статистические связи, таблицы сопряженности, отношение шансов*

Поступила: 11.04.2019. Принята к публикации: 09.10.2019

Введение

После аварии на ЧАЭС большие группы населения России подверглись облучению радионуклидами – ликвидаторы последствий аварии (ликвидаторы) и население, проживающее на загрязненных территориях. Медико-дозиметрическая информация об облученных людях накапливается в Национальном радиационно-эпидемиологическом регистре (НРЭР) и служит базисом для исследования воздействия облучения на здоровье населения России. Аккумулированные данные этих исследований представлены в [1]. В основном, эти исследования касались состояния здоровья ликвидаторов и анализа связи рака щитовидной железы с дозой облучения среди населения регионов, загрязненных радионуклидами вследствие аварии на ЧАЭС.

Целью данной работы является радиационно-эпидемиологическая классификация причин смерти от злокачественных новообразований (ЗНО) среди населения наиболее загрязненных радионуклидами вследствие аварии на ЧАЭС районов Брянской, Калужской, Тульской и Орловской областей, с использованием непараметрического метода статистических связей, для выявления первоочередных направлений исследования зависимостей доза–эффект радиационно-эпидемиологическими методами, учитывающими возможные мешающие и смещающие факторы.

Материал и методы

Общее описание когорты

В настоящее время в НРЭР зарегистрировано 357386 жителей Брянской, Калужской, Тульской и Орловской областей, подвергшихся радиационному воздействию вследствие аварии на Чернобыльской АЭС. На рис. 1 приведена карта загрязненности территорий радионуклидами и численности населения, проживающего на этих территориях.

В анализе использованы данные о случаях смерти за период наблюдения 1993–2017 гг., накопленные в НРЭР. Число умерших мужчин за период наблюдения 30771 человек (5407 смертей от ЗНО), женщин – 29033 человека (3472 смертей от ЗНО). Начало наблюдения 1993 г. было выбрано для учета минимального латентного периода индукции радиогенных раков.

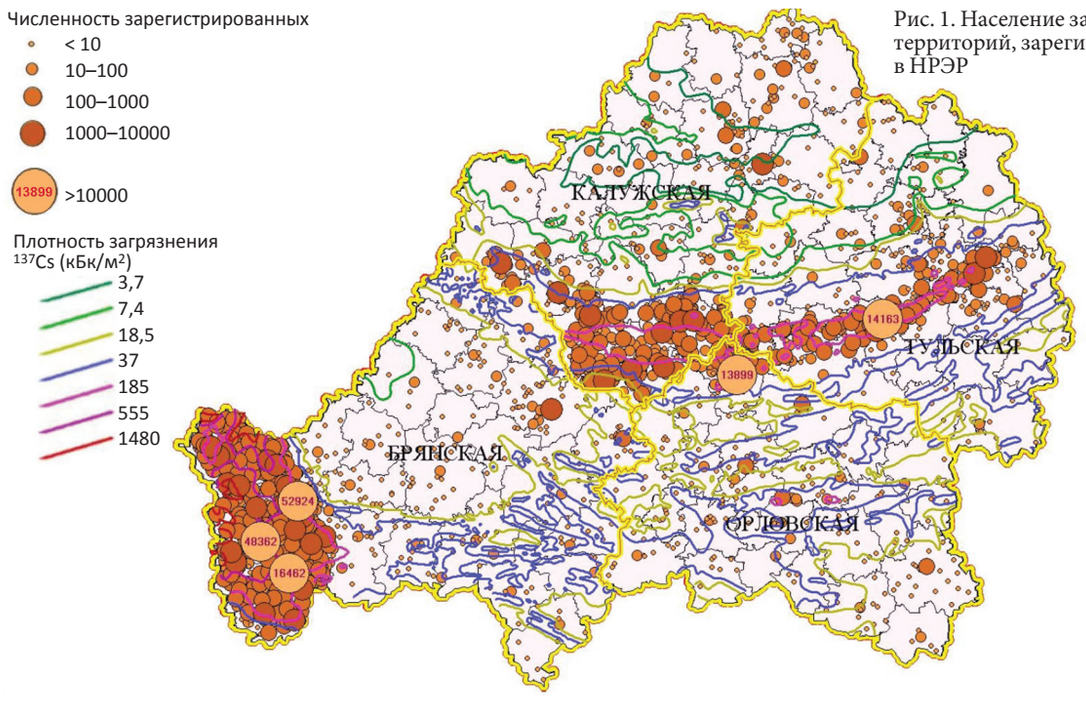

Рис. 1. Население загрязненных территорий, зарегистрированное в НРЭР

Все случаи смерти были разделены на две дозовые категории (группа 0 – до 0,014 Зв и группа 1 – 0,014+ Зв) и три возрастных категории (группа 0 – до 17 лет; группа 1 – 18–60 лет; группа 2 – 60+ лет).

В данной статье индивидуальная эффективная доза, рассчитанная в соответствии с МУ 2.6.1.579–96, условно принята в качестве меры радиационного воздействия для целей категоризации населения по дозе.

В анализе использованы значения эффективной дозы, накопленной до 1993 г. (около 90 % от дозы, накопленной к 2017 г.). Максимальная доза среди мужчин 0,232 Зв, максимальная доза среди женщин 0,424 Зв. Средняя эффективная доза равнялась 0,024 Зв.

Основные характеристики рассматриваемой когорты приведены на рис. 2, 3. На рис. 2 приведено возрастное распределение случаев смерти, на рис. 3 – распределение по возрасту случаев смерти от ЗНО.

Максимум смертей от всех причин наблюдается в возрасте 65–75 лет для мужчин и 75–85 лет для женщин. Для ЗНО максимумы наблюдаются в возрасте 60–70 лет для мужчин и 65–80 лет для женщин. Максимумы числа случаев смертей от всех причин в отличие от максимумов числа смертей от ЗНО смещены вправо, в сторону бóльших возрастов.

На рис. 4, 5 приведено распределение случаев смерти по дозе облучения.

Метод реконструкции эффективных доз на территориях указанных областей и их значения представлены в [2, 3].

Статистический метод анализа

Для анализа использовался программный модуль из пакета статистических программ «Статистика» [4] под названием Sequence, Association, and Link Analysis [4–7], который позволяет реализовать один из методов интеллектуального анализа данных – «Анализ последовательности, ассоциации и связи». Модуль определяет ассоциативные связи переменных, их временны́е последовательности и кластеры без априорных предположений о законах распределения переменных. Ниже приведены основные термины, используемые в этих методах анализа.

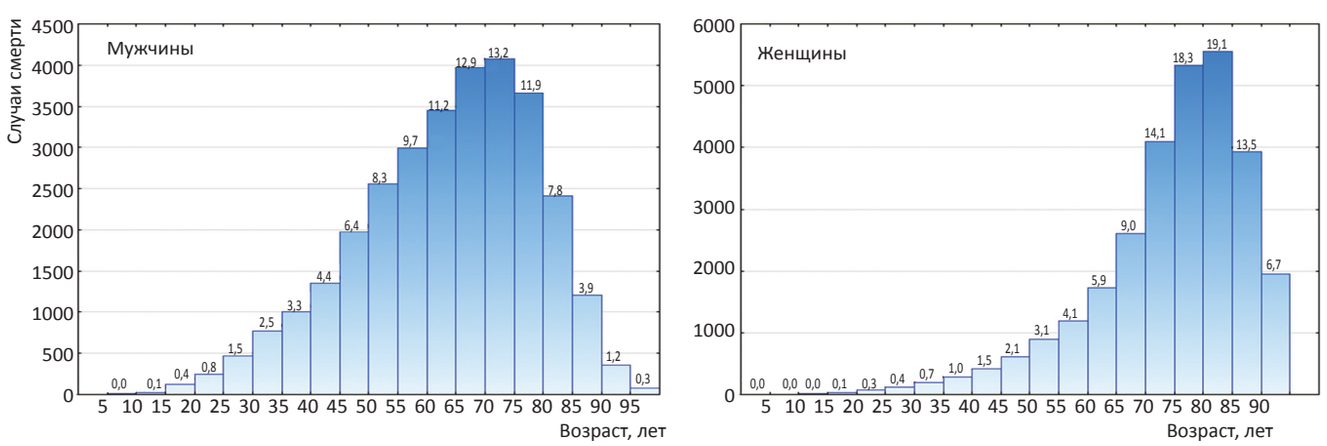

Рис. 2. Распределение случаев смерти (все причины) по возрасту на момент смерти, %

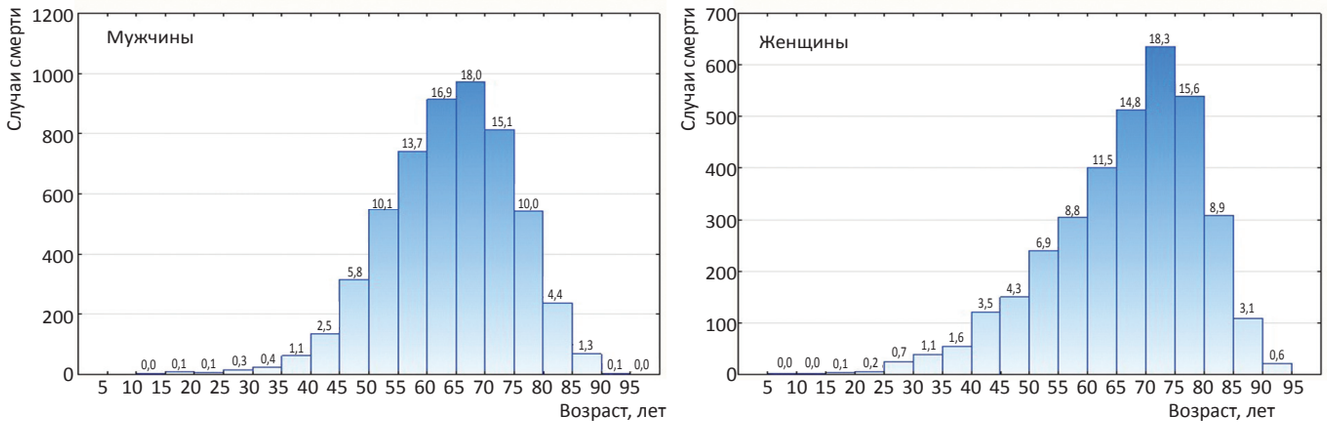

Рис. 3. Распределение случаев смерти от ЗНО по возрасту на момент смерти, %

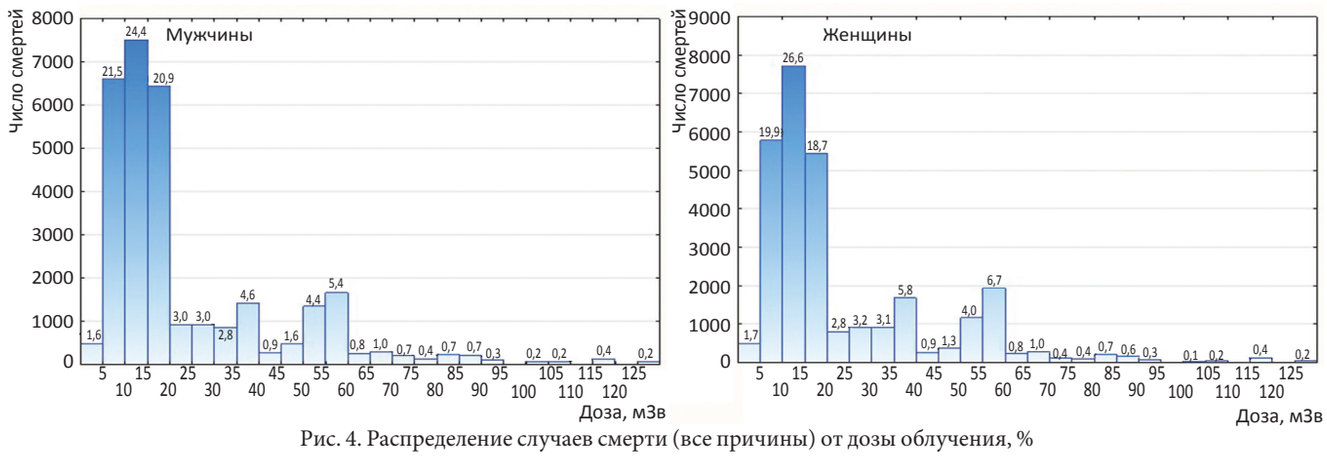

Рис. 4. Распределение случаев смерти (все причины) от дозы облучения, %

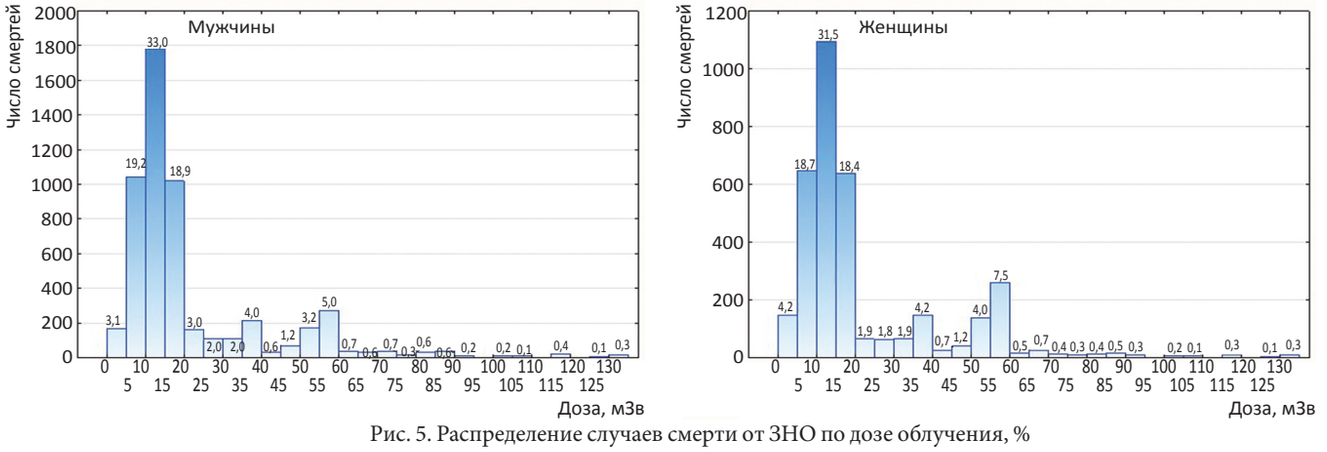

Рис. 5. Распределение случаев смерти от ЗНО по дозе облучения, %

Транзакция (*Transaction*) в контексте данной задачи – дозовая группа, возрастная группа, диагноз смерти и доза на одного умершего члена когорты.

Поддержка (*Support – sup*) – отношение числа транзакций (числа членов когорты), содержащих определенный набор данных (набор доз, возрастов и диагнозов смерти), к общему количеству транзакций (в данном случае умерших членов когорты). Например, Поддержка набора данных, состоящих всего из двух элементов данных, *A* и *C*, является отношением числа транзакций, содержащих элементы *A* и *C*, к общему числу транзакций, т.е. совместной вероятностью событий вхождения элементов *A* и *C*, *P*(*A*,*C*), в набор данных.

Достоверность (*Confidence – conf*) показывает вероятность осуществления правила, что из наличия в транзакции (у члена когорты) набора *A* следует наличие в ней набора *C*: *Confidence* (*if A→C*)=*Support*(*A*,*C*)/*Support*(*A*). «Достоверность» в данном контексте представляет собой условную вероятность *C* при условии *A*, *P*(*C*|*A*)=*P*(*A*,*C*)/*P*(*A*). Достоверность правила не является достаточной характеристикой для использования правила на практике.

Полезность правила (*lift*) определяется как *lift* (*if A→C*)=*Confidence* (*if A→C*)/*Support*(*C*)=*P*(*AC*)/(*P*(*A*)×*P*(*C*)).

Из этого определения следует, что если *lift* = 1, то события *A* и *C* независимы, и статистическая связь *A* и *C* отсутствует. Чем больше величина *lift*, тем больше мера статистической связи *A* и *C*, и тем более полезно правило связи. Критерием полезности правила считается условие *lift*>1.

Для определения статистической значимости правила используем вероятности Support (*sup*), Confidence (*conf*), меру связи *lift*, таблицы сопряженности 2×2 и статистику χ^2:

$$\chi^2 = \sum_{0\le i,j\le 1} \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}}, \qquad (1)$$

где $O_{i,j}$ и $E_{i,j}$ – наблюдаемые и ожидаемые частоты в ячейке (i, j) таблицы сопряженности 2×2 соответственно.

При расчете статистики χ^2 для пары переменных необходимо создать две таблицы (табл. 1, 2) сопряженности наблюдаемых (*O*) и ожидаемых значений (*E*).

Табл. 1 представляет наблюдаемые частоты, имеет четыре ячейки, соответствующие четырем возможным комбинациям событий *A* и *C*. Черта над переменной означает отсутствие события, или событие, дополнительное по отношению ко всему множеству рассматриваемых событий. Например, когда групповая переменная доза, как в данном исследовании, может принимать два значения – 0 и 1. Если событие *A* будет обозначать наличие члена когорты дозы в группе 1, то дополнительное к нему событие $\overline{A}$ – дозы в группе 0. Если событие *C* будет обозначать определенный диагноз смерти у члена когорты, то дополнительное к нему событие $\overline{C}$ – отсутствие данного диагноза.

В табл. 2 приведены ожидаемые (*E*) частоты для статистически независимых событий *A* и *C*.

Согласно определениям *sup, conf u lift*:

$$P(A\cap C)=sup;\ P(A)=\frac{sup}{conf};\ P(C)=\frac{conf}{lift};$$
$$P(A\cap\overline{C}) = P(A) - P(A\cap C) = \frac{sup}{conf} - sup, \qquad (2)$$
$$P(\overline{A}\cap C) = P(C) - P(A\cap C) = \frac{conf}{lift} - sup,$$
$$P(\overline{A}\cap\overline{C}) = 1 - P(A\cap C) - P(A\cap\overline{C}) - P(\overline{A}\cap C) =$$
$$= 1 - \frac{sup}{conf} - \frac{conf}{lift} + sup.$$

Подставив значения вероятностей в табл. 1, 2 и формулу (1), получим выражение для статистики χ^2 через вероятности *sup*, *conf* и меру связи *lift*.

$$\chi^2 = n\cdot(lift-1)^2\cdot\frac{sup\cdot conf}{(conf-sup)\cdot(lift-conf)}. \qquad (3)$$

Таблица 1

Таблица сопряженности для наблюдаемых частот

	C	$\overline{C}$
A	$n\cdot P(A\cap C)$	$n\cdot P(A\cap\overline{C})$
$\overline{A}$	$n\cdot P(\overline{A}\cap C)$	$n\cdot P(\overline{A}\cap\overline{C})$

Таблица 2

Таблица сопряженности для ожидаемых частот

	C	$\overline{C}$
A	$n\cdot P(A)\cdot P(C)$	$n\cdot P(A)\cdot(1-P(C)$
$\overline{A}$	$n\cdot(1-P(A)\cdot P(C)$	$n\cdot(1-P(A)\cdot(1-P(C)$

Табл. 1 сопряженности для наблюдаемых частот позволяет оценить отношение шансов (*OR*, odds ratio), которое в случае редких заболеваний представляет приближенную оценку относительного риска. Получим выражение для *OR* через вероятности *Support* (*sup*), *Confidence* (*conf*), меру связи *lift*, табл. 1 и соотношения (3).

Согласно табл. 1, отношение шансов события *C* при условии события *A* к шансам события *C* равно:

$$OR = \frac{P(A\cap C)\cdot P(\overline{A}\cap\overline{C})}{P(\overline{A}\cap C)\cdot P(A\cap\overline{C})}. \qquad (4)$$

После подстановки выражений для вероятностей (2) и упрощения получим:

$$OR = 1 + \frac{conf\cdot(1-lift)}{(conf-1)\cdot(conf-lift\cdot sup)}. \qquad (5)$$

Для расчетов приближенного 95 %-го доверительного интервала (ДИ) оценки *OR* используем приближение [8]:

$$OR^{(1\pm 1.96/\sqrt{\chi})}. \qquad (6)$$

Задание граничных значений для Поддержки и Достоверности имеет практическое значение. Задание больших значений для нижней границы Поддержки приведет к выявлению очевидных связей переменных, имеющих большие частоты, задание малых значений позволит выявить скрытые, неочевидные связи.

В данном исследовании Поддержка задана в пределах 1–100 %, Достоверность – в пределах 1–100 %.

В анализе ассоциаций рассматриваются правила, в которых из наличия одного набора элементов, который трактуется как причина (или условие, в терминах вероятностного анализа), следует наличие другого набора элементов, который трактуется как вероятностное следствие. Строго говоря, в анализе рассматривается не причинно-следственные связи, а статистические.

В данном исследовании в качестве причины выбиралась дозовая группа свыше 14 мЗв, а в качестве следствия – диагноз смерти (класс, рубрика или диагноз) в определенной возрастной категории. Правила связи между дозой облучения, возрастом и диагнозом смерти считались статистически значимыми, если нижняя граница 95 %-го ДИ оценки *OR* величина *low* превышала единицу.

Данный подход уже был использован для анализа смертности ликвидаторов последствий аварии на ЧАЭС [9].

Результаты

Для всего класса ЗНО (все диагнозы ЗНО в совокупности) статистически значимые связи смертности с дозой не выявлены ни для мужского, ни для женского населения четырёх исследованных областей РФ.

Мужчины

В табл. 3 приведены оценки *OR* для причин смерти мужского населения от ЗНО по трёхзначным рубрикам МКБ-10 [10] с Полезностью (*lift*) больше единицы в дозовой группе 1 (с большими дозами). Только в этом случае можно ожидать наличие дозовой зависимости причины смерти от дозы. Всего в заданных пределах Поддержки и Достоверности для первой дозовой груп-

пы выявлено 24 связи, из них Полезность больше единицы наблюдается только для 10 связей.

Прокомментируем обозначения в табл. 3. В первой колонке приведен фактор, связанный с причиной смерти, в данном случае дозовая группа 1, во второй – направление статистической связи, в третьей – в какой группе умерших появилась статистическая связь, в колонках 4–7 – значения параметров, необходимых для оценки *OR*, в колонках 8–9 – значение *OR* и нижний доверительный предел для *OR*. Напомним, что рассматриваются связи с Поддержкой (*support*) более 1 %.

В табл. 3 приведены следующие причины смерти мужского населения от ЗНО по трёхзначным рубрикам МКБ-10 с соответствующими локализациями: С64 – почка, С34 – бронхи, легкое, С25 – поджелудочная железа, С22 – печень, С18 – ободочная кишка, С16 – желудок.

Как следует из табл. 3, связи с дозой отдельных трёхзначных рубрик причин смерти от ЗНО статистически не значимы (нижняя 95 % доверительная граница *OR* меньше единицы). Тем не менее, следует отметить большое значение *OR* для ЗНО поджелудочной железы (рубрика С25), которое может стать статистически значимым при увеличении периода наблюдения за когортой (с увеличением статистической мощности исследования).

В табл. 4 представлены статистические связи причин смерти мужского населения с дозой для отдельных диагнозов (полезность больше единицы в дозовой группе 1): С64.0 – ЗНО почки (кроме почечной лоханки), С34.0 – ЗНО главных бронхов, С25.9 – ЗНО поджелудочной железы неуточнённое, С22.9 – ЗНО печени неуточненное, С16.0 – ЗНО кардии, С16.9 – ЗНО желудка неуточнённой локализации, С32.9 – ЗНО гортани неуточненное, С15.9 – ЗНО пищевода неуточнённое. Всего выявлены 18 связей, из них 11 имеют Полезность больше единицы.

Согласно данным табл. 4, статистически значимые связи причин смерти мужского населения с дозой выявлены для двух диагнозов: С16.9 – ЗНО желудка неуточнённой локализации и С25.9 – ЗНО поджелудочной железы неуточнённое.

Женщины

В табл. 5 приведены оценки *OR* для причин смерти женского населения от ЗНО по трёхзначным рубрикам

Таблица 3

Результаты оценки связи рубрик причин смерти от ЗНО с дозой облучения (мужчины)

1	2	3	4	5	6	7	8	9
Body	==>	Head	Support	Confidence	Lift	xisq	OR	low
dos_gr==1	==>	C64	0,012	0,026	1,118	1,549	1,251	0,879
dos_gr==1	==>	C34	0,126	0,272	1,015	0,358	1,038	0,92
dos_gr==1	==>	age_gr==1, C34	0,042	0,091	1,006	0,019	1,013	0,841
dos_gr==1	==>	age_gr==2, C34	0,084	0,181	1,019	0,348	1,043	0,907
dos_gr==1	==>	C25	0,021	0,045	1,141	3,712	1,31	0,995
dos_gr==1	==>	age_gr==2, C25	0,013	0,028	1,158	2,887	1,35	0,955
dos_gr==1	==>	C22	0,011	0,023	1,164	2,547	1,363	0,932
dos_gr==1	==>	C18	0,017	0,038	1,009	0,014	1,017	0,767
dos_gr==1	==>	age_gr==2, C18	0,014	0,029	1,034	0,16	1,068	0,774
dos_gr==1	==>	age_gr==2, C16	0,059	0,128	1,04	1,038	1,088	0,925

Таблица 4

Результаты оценки связи диагнозов причин смерти от ЗНО с дозой облучения (мужчины)

Body	==>	Head	Support	Confidence	Lift	xisq	OR	low
dos_gr==1	==>	C64.0	0,012	0,026	1,118	1,549	1,251	0,879
dos_gr==1	==>	C34.0	0,126	0,272	1,015	0,358	1,038	0,920
dos_gr==1	==>	age_gr==1, C34.0	0,042	0,091	1,006	0,019	1,013	0,841
dos_gr==1	==>	age_gr==2, C34.0	0,084	0,181	1,019	0,348	1,043	0,907
dos_gr==1	==>	C32.9	0,014	0,030	1,049	0,318	1,097	0,795
dos_gr==1	==>	C25.9	0,013	0,027	1,404	14,966	2,179	1,468
dos_gr==1	==>	C16.9	0,056	0,122	1,070	2,925	1,157	0,979
dos_gr==1	==>	age_gr==2, C16.9	0,041	0,089	1,127	6,418	1,291	1,059
dos_gr==1	==>	C16.0	0,013	0,029	1,077	0,747	1,157	0,832
dos_gr==1	==>	C15.9	0,012	0,027	1,084	0,830	1,173	0,832

Таблица 5

Результаты оценки связи рубрик смерти от ЗНО с дозой облучения (женщины)

Body	==>	Head	Support	Confidence	Lift	xisq	OR	low
dose_gr==1	==>	age_gr==2, C56	0,014	0,031	1,005	0,003	1,01	0,687
dose_gr==1	==>	age_gr==2, C34	0,022	0,046	1,054	0,413	1,112	0,804
dose_gr==1	==>	age_gr==2, C20	0,023	0,050	1,089	1,155	1,19	0,867
dose_gr==1	==>	age_gr==2, C18	0,028	0,060	1,065	0,753	1,136	0,852
dose_gr==1	==>	age_gr==2, C16	0,065	0,141	1,039	0,717	1,088	0,895
dose_gr==1	==>	age_gr==1, C50	0,031	0,066	1,037	0,279	1,076	0,82
dose_gr==1	==>	C34	0,029	0,061	1,06	0,657	1,125	0,846
dose_gr==1	==>	C20	0,030	0,064	1,124	2,83	1,278	0,96
dose_gr==1	==>	C18	0,033	0,071	1,053	0,598	1,111	0,851
dose_gr==1	==>	C16	0,079	0,169	1,000	0,000	1,001	0,838

Таблица 6

Результаты оценки связи диагнозов причин смерти от ЗНО с дозой облучения (женщины)

Body	==>	Head	Support	Confidence	Lift	xisq	OR	low
dose_gr==1	==>	age_gr==2, C34.9	0,016	0,035	1,193	3,357	1,448	0,974
dose_gr==1	==>	age_gr==2, C25.9	0,010	0,022	1,139	1,162	1,302	0,806
dose_gr==1	==>	age_gr==2, C20.0	0,022	0,048	1,083	0,960	1,175	0,851
dose_gr==1	==>	age_gr==2, C16.9	0,047	0,101	1,138	5,628	1,327	1,050
dose_gr==1	==>	age_gr==2, C16.0	0,010	0,022	1,139	1,162	1,302	0,806
dose_gr==1	==>	C53.9	0,015	0,032	1,131	1,488	1,285	0,859
dose_gr==1	==>	C50.9	0,056	0,120	1,048	0,911	1,107	0,898
dose_gr==1	==>	age_gr==1, C50.9	0,022	0,046	1,136	2,385	1,303	0,931
dose_gr==1	==>	C50.0	0,016	0,033	1,037	0,139	1,074	0,737
dose_gr==1	==>	C34.9	0,020	0,043	1,186	3,948	1,432	1,005
dose_gr==1	==>	C25.9	0,013	0,028	1,126	1,209	1,271	0,829
dose_gr==1	==>	C20.0	0,029	0,062	1,126	2,799	1,283	0,958
dose_gr==1	==>	C18.7	0,010	0,021	1,108	0,684	1,228	0,755
dose_gr==1	==>	C16.9	0,056	0,120	1,095	3,380	1,221	0,987
dose_gr==1	==>	C16.0	0,013	0,029	1,099	0,794	1,209	0,796

МКБ-10 с соответствующими локализациями: С56 – яичник, С34 – бронхи, легкое, С20 – прямая кишка, С18 – ободочная кишка, С50 – молочная железа, С16 – желудок. Всего выявлено 46 связей, из них 10 имеют Полезность больше единицы.

Для трёхзначных рубрик связи причин смерти женского населения с дозой облучения статистически не значимы.

В табл. 6 представлены статистические связи причин смерти женского населения с дозой для отдельных диагнозов (всего выявлено 47 связей, из них 15 имеют Полезность больше единицы): С34.9 – ЗНО бронхов или легкого неуточненной локализации, С25.9 – ЗНО поджелудочной железы неуточнённое, С20.0 – ЗНО прямой кишки С16.9 – ЗНО желудка неуточненной локализации, С16.0 – ЗНО кардии, С 53.9 – ЗНО шейки матки неуточненной части, С50.9 – молочной железы неуточненной части, С50.0 – соска и ареолы, С18.7 – сигмовидной кишки.

Согласно данным табл. 6, статистически значимые связи причин смерти женского населения с дозой выявлены для двух диагнозов: С16.9 – ЗНО желудка неуточненной локализации и С34.9 – ЗНО бронхов или легкого неуточненной локализации.

Обсуждение

Разработанная методика определения связи техногенного фактора (дозы) с его воздействием на здоровье (смертность) полезна как фильтр, определяющий набор гипотез о связи причины смерти с дозой, которые в первую очередь требуют более точного радиационно-эпидемиологического анализа с учётом возможных мешающих и смещающих факторов.

На первом этапе представленного подхода определяются связи для классов заболеваний, тем самым определяются большие группы людей, для которых эта связь проявляется, затем происходит уточнение на уровне рубрик и отдельных диагнозов при уменьшении численности соответствующих групп.

Основным преимуществом данной методики является анализ статистических связей не только на уровне классов и рубрик причин смерти, но и с учетом всей совокупности отдельных диагнозов, что трудно реализуемо при обычном радиационно-эпидемиологическом анализе.

Численность групп, для которых выявлены значимые связи дозы с причиной смерти, невелики по численности (200–300 человек) и отдельные связи могут быть фиктивными, особенно для малых групп людей вследствие неучтенных факторов гетерогенности популяции.

Полученные результаты свидетельствуют о том, что рассматриваемая популяция гетерогенна по величинам и направлению связи облучения с конкретными причинами смерти. Направление действия связей в обратном направлении (больше доза – меньше эффект) может вызвать смещение (bias) оценки изучаемого эффекта.

Данная методика представляется полезной для количественного определения связи техногенного фактора (в данном случае облучения) со смертностью и может быть полезна в работе экспертных советов по установлению причинной связи заболевания (смерти) с воздействием техногенного фактора.

Выводы

1. Для населения четырёх загрязненных после аварии на ЧАЭС областей РФ не выявлено статистически значимых связей причин смерти от ЗНО с дозой облучения ни для всего класса ЗНО, ни для трёхзначных рубрик МКБ-10 внутри этого класса.

2. Статистически значимые связи с дозой облучения наблюдались для отдельных диагнозов причин смерти: среди мужского населения – ЗНО желудка неуточненной локализации (С16.9) и ЗНО поджелудочной железы неуточнённое (С25.9); для женского населения – ЗНО желудка неуточненной локализации (С16.9) и ЗНО бронхов или легкого неуточненной локализации (С34.9). Эти диагнозы причин смерти в первую очередь должны быть предметом более чувствительного и специфичного радиационно-эпидемиологического анализа, учитывающего возможные факторы гетерогенности, мешающие и смещающие факторы.

3. Представленная методика определяет первоочередные направления исследований для более точного

радиационно-эпидемиологического анализа зависимостей доза–эффект среди населения, проживающего на территориях, загрязнённых радионуклидами вследствие аварии на ЧАЭС.

Для цитирования: Горский А.И., Максютов М.А, Туманов К.А., Власов О.К., Кочергина Е.В., Зеленская Н.С., Чекин С.Ю., Иванов С.А., Каприн А.Д., Иванов В.К. Анализ статистических связей смертности от злокачественных новообразований с дозой облучения радионуклидами населения регионов, загрязненных вследствие аварии на ЧАЭС // Медицинская радиология и радиационная безопасность. 2019. Т. 64. № 6. С. 5–11.

DOI: 10.12737/1024-6177-2019-64-6-5-11

DOI: 10.12737/1024-6177-2019-64-6-5-11

Analysis of Statistical Correlation between Radiation Dose and Cancer Mortality among the Population Residing in Areas Contaminated with Radionuclides after the Chernobyl Nuclear Power Station

A.I. Gorski[1], M.A. Maksiutov[1], K.A. Tumanov[1], O.K. Vlasov[1], E.V. Kochergina[1], N.S. Zelenskaya[1], S.Yu. Chekin[1], S.A. Ivanov[1], A.D. Kaprin[2], V.K. Ivanov[1]

1. A.F. Tsyb Medical Radiological Research Center, Obninsk, Russia. E-mail: gorski@nrer.ru;
2. National Medical Research Radiological Center, Moscow, Russia

A.I. Gorski – Leading Researcher, PhD Tech.; M.A. Maksiutov – Head of Dep., PhD Tech.; K.A. Tumanov – Head of Lab., PhD Biol.; O.K. Vlasov – Head of Lab., Dr. Sci. Tech.; E.V. Kochergina – Head of Lab., PhD Med.; N.S. Zelenskaya – Researcher; S.Yu. Chekin – Head of Lab.; S.A. Ivanov – Director, Dr. Sci. Med., Prof.; A.D. Kaprin – General Director, Academician of RAS, Dr. Sci. Med., Prof.; V.K. Ivanov – Deputy Director for Science, Corr. Member of RAS, Dr. Sci. Tech., Prof.

Abstract

Purpose: To evaluate the evidence for the correlation between radiation exposure of the population following the Chernobyl accident and cancer mortality using statistical correlation.

Material and methods: Data on deaths and causes of death for the period between 1993 and 2017 were taken from the National Radiation Epidemiological Registry (NRER). The number of deaths in males is 30771 persons including 5407 cancer death, the number of deaths in females is 29033 persons including 3472 cancer death. For analysis of statistical correlation between radiation exposure and causes of death data mining algorithms free of a priori statements on probabilistic distributions of doses and diagnoses were used. Contingency tables of deaths in two dose groups: the group 0 – <0.014 Sv and the group 1 – ≥0.014 Sv; and in three age groups: group 0 – <17 years old, group 1 – 18–60 years old, group 2 – 60+ years old were used for analysis. About 90 % of individual effective doses were accumulated from 1986 till 1993. Mean effective dose is 0.024 Sv.

Results: For the population of four areas in the Russian Federation which are most contaminate after to accident on the CNPP significant association of all causes of death from cancer and for three-digit headings ICD-10 with an exposure dose was not found.

Statistically significant association between dose and pancreatic cancer (C25.9 ICD-10) and stomach cancer (C16.9) in males, stomach cancer (C16.9) and malignant neoplasms of bronchus and lungs (C34.9) in females was found. The association between dose and cancer death was demonstrated in 150–230 cases.

These diagnoses of causes of death first of all have to be a subject of more sensing and specific radiation and epidemiological analysis considering possible the confounding factors.

Conclusion: The presented method is exploratory in nature and can be used to find a direction of research necessary for making more accurate evaluation of epidemiological evidence for the correlation between dose and effect. The method may be useful for evaluation of the association between radiation exposure and pathological response (death).

Key words: *population, Chernobyl accident, contamination by radionuclides, mortality, malignant neoplasms, radiation dose, radiation risk, statistical correlations, contingency tables, odds ratio*

Article received: 11.04.2019. Accepted for publication: 09.10.2019

REFERENCES

1. Ivanov VK, Kaprin AD, eds. Health effects of Chernobyl: prediction and actual data 30 years after the accident. Moscow: GEOS; 2015. 450 p. (in Russian).
2. Brook GY, Bazjukin AB, Bratiliva AA, et al. Average effective exposure doses (accumulated for 1986–2016) of residents of the settlements of the Russian Federation carried to radiocontamination zones under the resolution of the Government of the Russian Federation of 08.10.2015 No. 1074 "About the approval of the List of the settlements which are in borders of zones of a radiocontamination owing to accident on the Chernobyl NPP". (in Russian).
3. Vlasov OK, Brook GY, Schukina NV. Development and verification of technology of reconstruction of effective exposure doses of the population of Russia after accident on the CNPP. Radiation and Risk. 2017;26(3):28-45. (in Russian).
4. Statsoft. Available from: http://www.statsoft.ru (cited 16.10.2018).
5. Piatetsky-Shapiro G. Discovery, analysis and presentation of strong rules. In: Knowledge Discovery in Databases. Piatetsky-Shapiro G, Frawley WJ, eds. Cambridge, MA: AAAI/MIT Press: 1991: 229-48.
6. Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases. Proc. 1993 ACM SIGMOD Int. Conf. Management of data (SIGMOD'93). New York, 1993: 207-16.
7. Hahsler M. A Probabilistic Comparison of Commonly Used Interest Measures for Association Rules, 2015. Available from: http://michael.hahsler.net/research/association_rules/measures.html (cited: 16.10.2018).
8. Mietenen O.S. Confounding and effect modification. Amer J Epidemiol. 1974;100:350-53.
9. Gorski AI, Maksioutov MA, Tumanov KA, Kochergina EV, Korelo AM. Association rules for discovery relationship between mortality among Chernobyl liquidators and radiation dose. Radiation and Risk. 2018;27(1):22-32. (in Russian).
10. International Statistical Classification of Diseases and Related Health, 10th revision (ICD-10). Vol. 1 (Part 1). Geneva: WHO; 1995. 696 p. (in Russian).

For citation: Gorski AI, Maksiutov MA, Tumanov KA, Vlasov OK, Kochergina EV, Zelenskaya NS, Chekin SYu, Ivanov SA, Kaprin AD, Ivanov VK. Analysis of Statistical Correlation between Radiation Dose and Cancer Mortality among the Population Residing in Areas Contaminated with Radionuclides after the Chernobyl Nuclear Power Station. Medical Radiology and Radiation Safety. 2019;64(6): 5-11. (Russian).

DOI: 10.12737/1024-6177-2019-64-6-5-11