

Статистический анализ влияния сильного стресса на уровень заболеваемости населения раком

Е. А. Арутюнян, Г. К. Базилян, О. В. Григорян, П. А. Петросян

Республиканский противоопухолевый диспансер РА
Институт проблем информатики и автоматизации НАН РА и ЕрГУ
E-mail: evhar@ipia.sci.am

Аннотация

Длительные стрессовые нагрузки уменьшают сопротивляемость организма к различным заболеваниям в том числе и к онкологическим. Цель данной работы доказать значимость стрессовых нагрузок как одну из причин возникновения онкопатологических новообразований на примере севера Армении, подвергшегося в декабре 1988 года катастрофическому землетрясению.

Проведен статистический анализ заболеваемости раком в городе Гюмри (до 1992 г. Ленинакан) по сравнению с Ереваном до и после разрушительного землетрясения.

Показано, что заболеваемость раком, бывшая до землетрясения в Гюмри ниже, чем в столице, в годы после землетрясения оказалась значимо большей, чем в Ереване.

Фактор стресса оказал особенно сильное влияние на мужское население, тогда как заболеваемость женщин в Гюмри, по сравнению с Ереваном тоже стала больше, но незначимо.

1. Постановка задачи

Цель работы — исследование статистической значимости влияния фактора стресса на частоту онкологических заболеваний путем сравнения регионов Армении, пострадавших и не пострадавших от разрушительного землетрясения в декабре 1988 года.

Источником информации для исследования явилась компьютерная база данных, разработанная и внедренная в Республиканском противоопухолевом диспансере Республики Армения.

Разработка систем сбора и анализа данных заболеваемости злокачественными новообразованиями предусматривается различными противораковыми программами и предполагается как обязательный раздел деятельности онкологического диспансера. Последние годы для таких систем общепринято использовать термин "канцер-регистр". Возможности статистического анализа материалов канцер-регистра на примере рака желудка описаны, например, в [4].

Компьютерная система диспансерного учета — ОНКО, созданная в Республиканском онкодиспансере предназначена для решения задачи сбора, хранения и актуализации данных для их последующей обработки. Обобщение и обработка данных об онкозаболеваниях ведется по двум отчетным формам, рекомендованным ВОЗ, предполагающим

разбивку по половозрастным группам и по группам патологий. Объем базы — около 33 тысяч карточек. Программа управления базой данных предусматривает также получение таблиц и графиков, представляющих распределение заболеваний по районам Армении, гистологической структуре и некоторым другим факторам. Умение подбирать подходящие программные средства статистической обработки и умение интерпретировать результаты статистического анализа, как отмечается в [1], представляет определенную сложность для пользователей, не являющихся специалистами в математической статистике. Различные вопросы математической обработки данных медицинских исследований освещаются в работах [3-9]. Однако для повышения эффективности работы программа управления базой данных нуждается в автоматизированных статистических процедурах. Поэтому результаты настоящего исследования могут послужить также для выработки рекомендаций при последующей автоматизации тех или иных моментов анализа.

Для сравнительного анализа были выбраны Ереван, как город не претерпевший заметного ущерба от землетрясения, и Гюмри, как большой город, значительно пострадавший от землетрясения.

Данные о численности заболеваний среди различных половозрастных групп были извлечены из базы данных и сгруппированы в виде матрицы $\|Z_{ij}^k\|$, состоящей из трех столбцов: заболевания мужчин, заболевания женщин, суммарное число заболеваний обоих полов и четырех строк по возрастным группам: от 0 до 14 лет, от 15 до 34 лет, от 35 до 59 и свыше 60 лет, $k = 1$ обозначает Ереван, $k = 2$ — Гюмри. Элементы матрицы $\|N_{ij}^k\|$, $i = \overline{1,3}$, $j = \overline{1,4}$, представляют сведения о количестве населения по тем же половозрастным группам. Частота p_{ij}^k заболеваний в каждой половозрастной группе определяется как отношение:

$$p_{ij}^k = \frac{Z_{ij}^k}{N_{ij}^k}, \quad i = \overline{1,3}, \quad j = \overline{1,4}, \quad k = \overline{1,2}.$$

Частоты заболеваний за 1993 г. для городов Еревана и Гюмри приведены в таблице 1. Обозначим общую частоту заболеваний в Ереване p^1 в Гюмри p^2 . Как видно из таблицы в 1993 г. $p^2 > p^1$. Такое соотношение между частотами заболеваний наблюдалось и в 1994 и в 1995 годах. Цель работы — выяснить является ли это соотношение статистически значимым и, если да, то в какой именно половозрастной группе наиболее значимо проявляется это различие. Кроме того, если значимое различие объясняется стрессом, наблюдалось ли обратное неравенство, а именно $p^2 \leq p^1$ в годы предшествующие землетрясению.

2. Результаты обработки данных

Мы предполагаем, что существует некоторое неизвестное число p_k представляющее вероятность заболевания, постоянное для жителей всех половозрастных групп k -ого города, $k = \overline{1,2}$, в течении данного года. Кроме того, полагаем, что для Еревана эта неизвестная вероятность совпадает с наблюдаемой частотой, т. е. $p^1 = p_1$, а для Гюмри, мы по наблюдаемой частоте p^2 хотим проверить нулевую гипотезу $H_0: p_2 = p_1$.

Согласно [2], при выполнении условия

$$N_{ij}^2 p_{ij}^2 (1 - p_{ij}^2) > 9 \quad (1)$$

частота p^2 имеет асимптотически нормальное распределение с математическим ожиданием $Ep^2 = p_2$ и дисперсией $Dp^2 = \frac{p_2(1-p_2)}{N^2}$.

Следовательно, в случае справедливости гипотезы H_0 статистика

$$Z_0 = \frac{p^2 - p^1}{\left[\frac{p^1(1-p^1)}{N^2}\right]^{1/2}} \quad (2)$$

имеет стандартное нормальное распределение и для определения критических значений можно воспользоваться таблицами приведенными, например, в [2].

По частотам заболеваемости за 1993-1995 г. и по данным 1986 г. т. е. за два года до землетрясения получены следующие результаты.

Таб. 1. Частоты заболеваний и значения Z_0 -критерия для 1993 г.

Возраст (в годах)	Город	Частота			Z_0		
		мужчины	женщины	общая	мужчины	женщины	всего
0-14	Ереван	0.0000629	0.0000065	0.0000353	0.8293308	6.4940631	2.7475469
	Гюмри	0.0001010	0.0001047	0.0001028			
15-34	Ереван	0.0001460	0.0002331	0.0001901	-0.7781103	1.4884804	0.7120971
	Гюмри	0.0000935	0.0003571	0.0002284			
35-59	Ереван	0.0015718	0.0020949	0.0018515	4.3904339	-2.0084826	1.2408601
	Гюмри	0.0025948	0.0015801	0.0020723			
старше 60	Ереван	0.0092413	0.0067891	0.0078364	3.2952192	-2.4411328	0.3540770
	Гюмри	0.0126007	0.0050599	0.0080457			
все возрасты	Ереван	0.0014476	0.0015792	0.0015158	4.0016424	-2.8797861	0.6129977
	Гюмри	0.0019296	0.0012266	0.0015685			

Таб. 2. Частоты заболеваемости и значения статистики Z_0 .

Год	Ереван	Гюмри	Z_0
1986	0.00198	0.00115	-8.87
1993	0.00152	0.00157	0.613
1994	0.00154	0.00167	1.506
1995	0.00150	0.00188	4.435

Как видно из таблицы 2 в 1986 г. имеет место $p^2 < p^1$, а в годы после землетрясения (1993-1995) — обратное неравенство $p^2 > p^1$. Причем, поскольку при доверительной вероятности $\alpha = 0.05$, $Z_\alpha = 1.64$, то за 1993 г. и 1994 г. частота заболеваний в Гюмри незначимо превышает частоту заболеваний в Ереване, а в 1995 г. становится значимой.

Кажущееся уменьшение частоты заболеваемости в Ереване, видимо, связано с тем, что не учтен фактор миграции, имевшей место в последние годы, точные данные о которой, однако, отсутствуют.

Для того, чтобы выяснить в какой половозрастной группе влияние стресса наиболее значимо, статистика (2) была вычислена по данным таблицы 1, причем значение Z_0 сравнивалось с критическим значением Z_α только для тех строк, для которых выполнялось неравенство (1). Результаты приведены во второй части таблицы 1.

Таким образом, за 1993 г., влияние стресса оказывается значимым только среди мужского населения. Частота заболеваний женщин в Гюмри и после землетрясения среди всех возрастных групп кроме возрастной группы от 15 до 34 лет остается меньшей, чем в Ереване. Аналогичные результаты получены по данным 1994 и 1995 годов.

3. Выводы

- Показано, что одной из причин возникновения онконовообразований является фактор стресса.
- Стрессовые нагрузки оказывают статистически значимое воздействие на заболеваемость населения онконовообразованиями.
- Влияние стресса преимущественно сказывается на мужском населении.
- Среди женского населения воздействие стресса оказалось значимым в возрастной группе 15–34 лет.
- Эксплуатируемая в диспансере Республики база данных, совместимая с пакетом Microsoft Excel–97, даст возможность сравнительного статистического анализа данных об онкозаболеваниях.

Авторы выражают благодарность сотрудникам Республиканского противоопухолевого диспансера Республики Армения Алавердян А. С., Погосяну П. Б., Мурадян К. М., Алавердян А. А. за активное содействие.

Литература

- [1] С. А. Айвазян. Интеллектуализированные инструментальные системы в статистике и их роль в построении проблемно - ориентированных систем поддержки принятия решений. Обзорение прикладной и промышленной математики. Научное издательство "ТВП" Москва, 1997, с. 667–680.
- [2] А. Афифи, С. Эйзен. Статистический анализ. Подход с использованием ЭВМ. Пер. с англ. М. 1982.
- [3] Е. В. Гублер. Вычислительные методы анализа и распознавания патологических процессов. Медицина, Ленинград, 1978.
- [4] Л. Е. Денисов, Т. И. Ушаков, В. Д. Володин. Возможности обработки материалов канцер-регистра на персональном компьютере (на примере рака желудка). М. 1995.
- [5] А. М. Лисенков. Математические методы планирования многофакторных медико - биологических экспериментов. Медицина, Москва, 1979.
- [6] D. Clayton, M. Hills. Statistical Models in Epidemiology. Oxford University Press, 1993.
- [7] P. Elliott, Y. Cuzick, D. English, R. Stern. Geographical and Environmental Epidemiology. Oxford University Press, 1992.
- [8] J. Estewe, E. Benhamou, L. Raymond. Statistical Methods in Cances Research. Vol. Descriptive Epidemiology, Intern. Agency for Research in Cancer, Lyon 1994,
- [9] Understanding Statistics. Grahum Upton and Ian Coon, 1996, Oxford University Press.