

## КОРРЕЛЯЦИОННЫЙ И ФАКТОРНЫЙ АНАЛИЗ ЭПИДЕМИЧЕСКОЙ СИТУАЦИИ ПО ТУБЕРКУЛЕЗУ В ОРЕНБУРГСКОЙ ОБЛАСТИ

**В статье рассматривается применение корреляционного анализа для определения парных обусловленностей и факторного анализа для выявления групповых обусловленностей эпидемиологических показателей по туберкулезу в Оренбургской области с целью оптимизации оценивания, анализа и контроля за эпидемической ситуацией.**

**Ключевые слова:** туберкулез, эпидемическая обстановка, заболеваемость туберкулезом.

Туберкулез является одним из самых значимых по своим трагическим последствиям для человечества инфекционным заболеванием [1], [2]. Эпидемическая ситуация по этой инфекции в Оренбургской области, несмотря на некоторую стабилизацию в последние годы остается напряженной [3], [4]. В регионе имеются дополнительные факторы, влияющие на неблагоприятное по туберкулезу: соседство с Республикой Казахстан, где пораженность населения туберкулезом в 2,5 раза выше, значительное влияние контингентов учреждений УФСИН, значительное количество населения с доходами ниже прожиточного минимума, очень высокий уровень ВИЧ-инфицированного населения, а так же нарастающая с каждым годом лекарственная устойчивость, включая множественную (DR) и экстремальную (MDR) [4]. Анализ эпидемической ситуации в регионе позволит более полно реализовать такие управленческие функции, как учет, контроль, анализ, оценивание и планирование противотуберкулезных мероприятий.

Целью работы явилась оптимизация оценивания, анализа и контроля над эпидемической ситуацией по туберкулезу в Оренбургской области. Задачей исследования является использование корреляционного анализа для определения парных обусловленностей и факторного анализа для определения групповых обусловленностей эпидемиологических параметров по туберкулезу.

### Материалы и методы

Для определения обусловленностей параметров исследования были выбраны основные эпидемиологические показатели за 12 лет: заболеваемость туберкулезом по 33 форме, за-

болеваемость туберкулезом по 8 форме, заболеваемость детей, заболеваемость подростков, распространенность туберкулеза, распространенность деструктивного туберкулеза, распространенность бациллярного туберкулеза, распространенность фиброзно-кавернозного туберкулеза и смертность от туберкулеза (таблица 1); показатели заболеваемости и распространенности ВИЧ-инфекции (таблица 2); уровень лекарственной устойчивости микобактерий туберкулеза (МБТ) (таблица 3) и регистрируемые по годам клинические формы туберкулеза.

### Результаты и обсуждение

Для определения парных обусловленностей был проведен корреляционный анализ. Для проведения корреляционного анализа была построена матрица исследования с параметрами-столбиками:

- (0. Годы)
- (1. Всего)
- (2. Очаговый)
- (3. Инфильтративный)
- (4. Казеозная пневмония)
- (5. ДТЛ)
- (6. Миллиарный)
- (7. Туберкулома)
- (8. Кавернозный)
- (9. ФКТ)
- (10. Цирротический)
- (11. Первичный туб. комплекс)
- (12. Туб. плеврит)
- (13. Туб. ВГЛУ)
- (14. Туб. трахеи, бронхов)
- (15. Ранняя туб. интоксикация)
- (16. Заболеваемость туберкулезом по 33 ф.)
- (17. Заболеваемость туберкулезом по 8 ф.)
- (18. Заболеваемость детей)

Таблица 1. Эпидемиологические показатели по туберкулезу

Показатели	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Заболеваемость туберкулезом по 33 ф., на 100 тыс.	65,4	68,7	70,6	72,7	72,8	81,3	86,4	83,7	85,6	83	75,4	72,3
Заболеваемость туберкулезом по 8 ф., на 100 тыс.	83,3	75	96,8	87,1	91,8	98,8	110,9	108,8	113,3	105,7	96,1	94,2
Заболеваемость детей, на 100 тыс.	9,8	10,7	13,4	12,3	10,8	13,7	13,1	12,5	12,6	10,2	13,4	13,4
Заболеваемость подростков, на 100 тыс.	26	29,7	27,9	33,4	34,1	44,2	54,7	41,5	42,2	44,3	40,2	28,8
Распространенность туберкулеза, на 100 тыс.	234	237	252	255	201	204	207	214	218,7	217,9	213,8	212,3
Распространенность деструктивного туберкулеза, на 100 тыс.	72,9	75,7	81,9	82,5	84,9	91,7	96,6	90,4	93,7	89,6	❖	❖
Распространенность бациллярного туберкулеза, на 100 тыс.	77,9	82,3	84,5	85,6	85,0	91,5	94,5	92,3	89,2	88,1	83	84,5
Распространенность фиброзно-кавернозного туберкулеза, на 100 тыс.	11,8	11,8	11,4	11,6	12,1	12,7	11,9	11,8	11,8	12,3	11	12,7
Смертность от туберкулеза, на 100 тыс.	9,9	10,5	10,7	9,4	10,4	14	13,6	13,5	12,7	12,1	10,3	9,5

❖ – показатель не определялся

Таблица 2. Показатели заболеваемости и распространенности ВИЧ-инфекции

Показатели	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Число выявленных ВИЧ (абс.)	4127	4266	2618	1810	1681	1751	1887	2118	2082	2557	2688	1977
Заболеваемость ВИЧ-инфекцией, на 100 тыс.	168,2	178,8	109,1	75,5	71,5	75,8	79,9	90,2	88,3	93,9	97,7	71,5
Распространенность ВИЧ-инфекции, на 100 тыс.	❖	❖	❖	❖	674,9	634,1	758,4	948,9	894,3	820,1	977,7	832,6
Смертность от ВИЧ-инфекции, на 100 тыс.	0	0	0,2	0,4	1	1,3	4,2	8,9	7,5	9,8	7,2	10,8
Число выявленных случаев ТБ-ВИЧ (абс.)	0	15	29	38	119	178	197	213	248	360	429	480
Заболеваемость ТБ-ВИЧ, на 100 тыс.	❖	0,67	1,3	1,7	5,5	8,2	9,2	10	11,7	17	21,1	23,63
Распространенность ТБ-ВИЧ, на 100 тыс.	❖	❖	❖	❖	❖	❖	❖	❖	31,8	37,5	42,5	50,5

❖ – показатель не определялся

- (19. Заболеваемость подростков)
- (20. Распространенность туберкулеза)
- (21. Распространенность деструктивного туберкулеза)
- (22. Распространенность бациллярного туберкулеза)
- (23. Распространенность фиброзно-кавернозного туберкулеза)
- (24. Смертность от туберкулеза)
- (25. Число выявленных ВИЧ (абс.))
- (26. Заболеваемость ВИЧ-инфекцией, на 100 тыс.)
- (27. Распространенность ВИЧ-инфекции, на 100 тыс.)
- (28. Смертность от ВИЧ-инфекции, на 100 тыс.)
- (29. Число выявленных случаев ТБ-ВИЧ (абс.))
- (30. Заболеваемость ТБ-ВИЧ, на 100 тыс.)
- (31. Распространенность ТБ-ВИЧ-на 100 тыс.)
- (32. Первичная лекарственная устойчивость (ЛУ)-%)
- (33. Первичная множественная лекарственная устойчивость (МЛУ), %)

Строчками-наблюдениями в матрице исследования были значения параметров исследования по годам наблюдений.

**Результаты корреляционного анализа параметр – (3. Инфильтративный)**

-----  
 0.57| 0.94| 0.56| 1.00| 0.57| 0.29| 0.72|-0.27|-0.47|-0.21|  
 -----  
 -0.53| 0.07| 0.54| 0.12|-0.70|-0.56| 0.96| 0.93| 0.34| 0.89|  
 -----  
 -0.63| 0.82| 0.85| 0.15| 0.83|-0.63|-0.62| 0.65| 0.52| 0.45|  
 -----  
 0.42| 0.37|-0.62| 0.40|  
 -----

**с параметром – (0. Годы)**

- коэффициент корреляции = 0.575 с параметром – (1. Всего)
- коэффициент корреляции = 0.937 с параметром – (2. Очаговый)
- коэффициент корреляции = 0.557 с параметром – (3. Инфильтративный)
- коэффициент корреляции = 1.000 с параметром – (4. Казеозная пневмония)
- коэффициент корреляции = 0.568 с параметром – (6. Милиарный)
- коэффициент корреляции = 0.720 с параметром – (8. Кавернозный)
- коэффициент корреляции = -0.469 с параметром – (10. Цирротический)
- коэффициент корреляции = -0.529 с параметром – (12. Туб. плеврит)
- коэффициент корреляции = 0.545 с параметром – (14. Туб. трахеи, бронхов)
- коэффициент корреляции = -0.700 с параметром – (15. Ранняя туб. интоксикация)
- коэффициент корреляции = -0.560 с параметром – (16. Заболеваемость туберкулезом по 33 ф.)
- коэффициент корреляции = 0.961 с параметром – (17. Заболеваемость туберкулезом по 8 ф.)
- коэффициент корреляции = 0.927 с параметром – (18. Заболеваемость детей)
- коэффициент корреляции = 0.340 с параметром – (19. Заболеваемость подростков)
- коэффициент корреляции = 0.894 с параметром – (20. Распространенность туберкулеза)
- коэффициент корреляции = -0.629 с параметром – (21. Распространенность деструктивного туберкулеза)

Таблица 3. Уровень лекарственной устойчивости микобактерий туберкулеза

Показатели	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Первичная лекарственная устойчивость (ЛУ), %	68	59	45	51,6	42,1	49,4	40,8	57,4	51,2	54,1	52,9	56,6
Первичная множественная лекарственная устойчивость (МЛУ), %	7,3	10,3	8,6	6,7	8,5	12,7	11,7	14	16	23,8	26,5	22,6

коэффициент корреляции = 0.825  
с параметром – (22. Распространенность  
бациллярного туберкулеза)

коэффициент корреляции = 0.853  
с параметром – (24. Смертность от тубер-  
кулеза)

коэффициент корреляции = 0.829  
с параметром – (25. Число выявленных  
ВИЧ (абс.))

коэффициент корреляции = -0.627  
с параметром – (26. Заболеваемость ВИЧ-  
инфекцией, на 100 тыс.)

коэффициент корреляции = -0.620  
с параметром – (27. Распространенность  
ВИЧ-инфекции, на 100 тыс.)

коэффициент корреляции = 0.654  
с параметром – (28. Смертность от ВИЧ-  
инфекции, на 100 тыс.)

коэффициент корреляции = 0.524  
с параметром – (29. Число выявленных  
случаев ТБ-ВИЧ (абс.))

коэффициент корреляции = 0.454  
с параметром – (30. Заболеваемость ТБ-  
ВИЧ, на 100 тыс.)

коэффициент корреляции = 0.420  
с параметром – (31. Распространенность  
ТБ-ВИЧ-на 100 тыс.)

коэффициент корреляции = 0.366  
с параметром – (32. Первичная лекар-  
ственная устойчивость (ЛУ), %)

коэффициент корреляции = -0.615  
с параметром – (33. Первичная множе-  
ственная лекарственная устойчивость  
(МЛУ), %)

коэффициент корреляции = 0.401

Согласно корреляционному анализу, при значениях модуля коэффициента парной корреляции больше 0,75 мы имеем сильную линейную связь между двумя параметрами [5]. То есть один параметр можно определить через другой по линейному регрессионному уравнению  $y=ax+b$  с высокой точностью. При значениях модуля коэффициента парной корреляции больше 0,5 и меньше 0,75 мы имеем среднюю линейную связь между двумя параметрами и один параметр можно определить через другой по линейному регрессионному уравнению  $y=ax+b$  со средней точностью. И соответственно при значениях модуля коэффициента парной корреляции меньше 0,5 мы имеем слабую линей-

ную связь между двумя параметрами и один параметр можно определить через другой по линейному регрессионному уравнению  $y=ax+b$  с малой точностью.

Распространенность инфильтративного туберкулеза имеет сильную линейную связь с заболеваемостью туберкулезом по ф. 8 (0,927), заболеваемостью по ф. 33 (0,961), распространенностью бациллярного туберкулеза (0,853), распространенностью деструктивного туберкулеза (0,894) и смертностью от туберкулеза (0,829). Здесь мы имеем сильную линейную связь. Любой из параметров можно определить по линейному регрессионному уравнению через другой с высокой точностью.

### Результаты факторного анализа

Для определения групповых обусловленностей был проведен факторный анализ [7] – [11]. Согласно факторному анализу, параметры, объединившиеся в одном факторе, имеют одну природу поведения, одну групповую обусловленность. То есть любой из параметров, объединившихся в факторе, можно определить по другим параметрам, попавшим в этот фактор, по соответствующим регрессионным моделям. В каждом факторе находился параметр с максимальной по модулю факторной нагрузкой – базовый параметр. По базовым параметрам можно определить все остальные из объединения в факторе по соответствующим регрессионным моделям.

Число базовых параметров меньше общего числа параметров исследования. Поэтому использование базовых параметров приводит к минимизации общего числа параметров исследования.

Это упрощает исследование без потери данных по небазовым параметрам, потому что значения небазовых параметров определяются по регрессионным уравнениям, параметрами-аргументами в которых будут базовые параметры. Например, в факторе 1 имеем:

Таблица 4  
Объединение по фактору 1

НОМЕР	НАЗВАНИЕ ПАРАМЕТРА	НАГРУЗКА
1	(0. Годы)	0.9348
7	(6. Миллиарный)	0.8553

11  (10. Цирротический)   -0.5607
14  (13. Туб. ВГЛУ)   0.8686
15  (14. Туб. трахеи, бронхов)   -0.5598
21  (20. Распространенность туберкулез    а)   -0.6348
22  (21. Распространенность деструктив    ного туберкулеза)   0.6049
28  (27. Распространенность ВИЧ-инфекц    ии, на 100 тыс.)   0.7945
29  (28. Смертность от ВИЧ-инфекции, н    а 100 тыс.)   0.8897
30  (29. Число выявленных случаев ТБ-В    ИЧ (абс.))   0.9805
31  (30. Заболеваемость ТБ-ВИЧ, на 100    тыс.)   0.9851
32  (31. Распространенность ТБ-ВИЧ-на    100 тыс.)   0.9577
34  (33. Первичня множественная лекарственная    устойчивость(МЛУ), %)   0.9709

В первом факторе базовый параметр – заболеваемость сочетанием туберкулеза и ВИЧ-инфекции (ТБ-ВИЧ) на 100 тыс. населения. По этому базовому параметру по соответствующим регрессионным уравнениям можно определить все параметры из прилагаемого объединения.

Таблица 5  
Объединение по фактору 6

НОМЕР	НАЗВАНИЕ ПАРАМЕТРА	НАГРУЗКА
5  (4. Казеозная пневмония)   -0.6252		
24  (23. Распространенность фиброзно-    кавернозного туберкулеза)   -0.9313		

В шестом факторе базовый параметр – распространенность фиброзно-кавернозного туберкулеза.

Для определения количественных обусловленностей параметров методом Брандона [6] строились регрессионные модели и по ним определялись вклады параметров-аргументов, которые являются оценкой количественной обусловленности зависимого параметра на

множестве параметров-аргументов. Чем больше вклад, тем больше обусловленность зависимого параметра от параметра-аргумента. Количественная обусловленность является самой приоритетной. Она выше парных и групповых обусловленностей, которые являются качественными обусловленностями. То есть парные и групповые обусловленности дают наличие связи, а количественные обусловленности дают по вкладам физическую оценку этой обусловленности.

Приведем вклады и модель для параметра – Заболеваемость туберкулезом по 8 форме.

Таблица 6

НОМЕР	НАЗВАНИЕ ПАРАМЕТРА	ВКЛАД В МОДЕЛЬ
3 (2. Очаговый)   0.0111		
4 (3. Инфильтративный)   0.9889		

$$y = + (111.6239411626) * (x_3)^0 (12) + (-0.5941178409) * (x_3)^1 + (0.0007802844) * (x_3)^2 + (-50.3421971262) * (x_4)^0 + (0.2289412254) * (x_4)^1 + (-0.0000561213) * (x_4)^2$$

где в (12) и далее ^ – возведение в степень.

Таблица 7

-Характеристики модели (12)

ХАРАКТЕРИСТИКИ МОДЕЛИ	ЗНАЧЕНИЯ
Коэффициент детерминации   0.9730	
Средняя абсолютная ошибка   3.70	
Средняя ошибка в процентах   3.80	

Согласно таблице 6, параметр «Заболеваемость туберкулезом по 8 ф.» более всего обусловлен клиническим параметром «инфильтративный», потому что его вклад 0.9889 наибольший.

Эта количественная обусловленность подтверждается результатами корреляционного и факторного анализов. Количественная обусловленность является приоритетной по сравнению с качественными парными и групповыми обусловленностями, которые определяются соответственно корреляционным и факторным анализами.

**Выводы**

Эпидемическая обстановка по туберкулезу несмотря на некоторое улучшение остается напряженной в Оренбургской области. Как показали наши исследования, на нее кроме социальных и географических факторов достоверно влияет нарастание лекарственной устойчивости возбудителя (МБТ) и увеличивающаяся распространенность ВИЧ-инфекции. Проведенный корреляционный и факторный анали-

зы показывают возможность минимизирования эпидемиологических показателей, необходимых для оценки эпидемической ситуации. Определение количественных обусловленностей эпидемиологических параметров позволяет определить зависимость между показателями относительно от их вкладов, что является приоритетной по сравнению с качественными парными и групповыми обусловленностями.

24.10.2012

**Список литературы:**

1. Ерохин, В. В. Туберкулез в России / В. В. Ерохин // Социально значимые болезни в Российской Федерации / Под ред. Л. А. Бокерия, И. Н. Ступакова. – М.: НЦССХ им. А. Н. Бакулева РАМН, 2006. – С. 263–266.
2. Перельман, М. И. Учение о туберкулезе / М. И. Перельман // Фтизиатрия: Национальное руководство. – Москва: ГЭОТАР, 2007. – С. 22–29.
3. Сазыкин, В. Л. Методологические аспекты комплексного оценивания деятельности фтизиатрической службы / В. Л. Сазыкин. – 2-е изд., дополненное. – М.: РИО ЦНИИОИЗ, 2010. – 224 с.
4. Сазыкина, И. Г. Взаимосвязь качества противотуберкулезной работы и эпидемиологического статуса территорий региона / И. Г. Сазыкина // Материалы XXVI областной научно-практической конференции фтизиатров / ред. В. Л. Сазыкин. – Оренбург: ООО Руссервис, 2010. – 56 с.
5. Михайловский, А. М. Особенности клинико-морфологических проявлений туберкулеза, сочетанного с ВИЧ-инфекцией в Оренбургской области: автореф. дисс. ... канд. мед. наук / А. М. Михайловский. – Москва, 2011. – 27 с.
6. Драйпер, Н. Прикладной регрессионный анализ / Н. Драйпер, Г. Смит. – М.: Статистика, 1973.
7. Brandon, D. V. Developing Mathematical Models for Computer Control / D. V. Brandon // USA Journal. – 1959. – V. 6, № 7.
8. Харман, Г. Современный факторный анализ / Г. Харман. – М.: Статистика, 1972.
9. Иберла, К. Факторный анализ / К. Иберла. – М.: Статистика, 1980.
10. Lawley, D. M. The estimation of factor loadings by the method of maximum likelihood / D. M. Lawley // Proc. roy. Soc. Edinb. Abo. 64-82(1940).
11. Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis / H. F. Kaiser // Psychometrika, 23, 187-200 (1958).

## Сведения об авторах:

**Михайловский Алексей Модестович**, заведующий патологоанатомическим отделением Оренбургского областного противотуберкулезного диспансера, кандидат медицинских наук 460041, г. Оренбург, Нежинское шоссе 6, e-mail: michailovsky2007@yandex.ru

**Чепасов Валерий Иванович**, заведующий кафедрой информационных систем и технологий Оренбургского государственного университета, доктор технических наук, профессор 460018, г. Оренбург, пр-т Победы, 13, e-mail: ist@unpk.osu.ru

UDC 616 – 002.5 – 036.22 – 084 (470.56)

Mihajlovskij A.M.<sup>1</sup>, Chepasov V. I.<sup>2</sup><sup>1</sup>Orenburg region tuberculosis clinic, e-mail: michailovsky2007@yandex.ru;<sup>2</sup>Orenburg state university, e-mail: ist@unpk.osu.ru**THE CORRELATION AND FACTORIAL ANALYSIS OF THE EPIDEMIC SITUATION ON THE TUBERCULOSIS IN THE ORENBURG REGION**

In article application of the correlation analysis for definition of paired conditioning and the factorial analysis to determine the parameters for study group conditioning epidemiological indicators on a tuberculosis in the Orenburg region for the purpose of optimisation of estimation, analysis and control over an epidemic situation is considered.

Key words: a tuberculosis, epidemic conditions, disease of a tuberculosis.