

Модуль «Статистическая эпидемиология».

1. Модуль «Статистическая эпидемиология».....	1
1.1. Проблемы сбора, хранения и статистического анализа данных по распространению туберкулеза в РФ.....	1
1.2. Структура подсистемы «Статистическая эпидемиология»	3
1.3. Возможности подсистемы «Статистическая эпидемиология».....	3
1.4. Статистические методы в подсистеме «Статистическая эпидемиология».....	3
1.5. Отображение результатов обработки.....	4
1.6. База данных.....	4
Полезные ссылки:.....	4

1. Модуль «Статистическая эпидемиология».

1.1. Проблемы сбора, хранения и статистического анализа данных по распространению туберкулеза в РФ

Первичными данными по распространению туберкулеза в РФ являются отчетные формы Министерства здравоохранения РФ за год. В них для каждого региона РФ отражены следующие показатели (Шилова, 2004, 2005):

Смертность на 100 тыс. населения	
Смертность детей до года на 100 тыс. населения	
Заболеваемость ф.8 на 100 тыс. населения	
Заболеваемость ф.33 на 100 тыс. населения	
Заболеваемость детей на 100 тыс. населения	
Бактериовыделение среди впервые выявленных больных туберкулезом детей, %	
Бактериовыделение среди впервые выявленных больных туберкулезом подростков,	
%	
Бактериовыделение среди впервые выявленных больных туберкулезом, %	
Первичное инфицирование у детей (вираж) на 100 тыс. детского населения	
Первичное инфицирование у подростков (вираж) на 100 тыс.	
Распространенность (болезненность) на 100 тыс. населения	
Бактериовыделителей на 100 тыс. населения	
Охват населения осмотрами, %	
Доля умерших в первый год после выявления среди всех умерших от туберкулеза, %	
Соотношение числа больных с МБТ+ т Кав.+ , %	
Заболеваемость туберкулезом лиц, проживающих совместно с	
бактериовыделителями, на 100 тыс. контактирующих	
Заболеваемость туберкулезом подростков, проживающих совместно с	
бактериовыделителями, на 100 тыс. контактирующих	
Заболеваемость туберкулезом детей, проживающих совместно с	
бактериовыделителями, на 100 тыс. контактирующих	
Число коек для взрослых	
Число коек для детей	
Число дней функционирования больничной койки для взрослых	
Число дней функционирования больничной койки для детей	
Число дней функционирования санаторной койки для взрослых	
Число дней функционирования санаторной койки для детей	
Госпитализация всех бациллярных больных в больницы и санатории, %	
Госпитализация впервые выявленных бациллярных больных с ТОД в стационары, %	
Госпитализация всех впервые выявленных больных в стационары, %	

Доля больных туберкулезом, умерших в стационарах, среди всех умерших от туберкулеза, %

Доля оперированных больных с ТОД среди состоящих на учете, %

Доля оперированных больных с ФКТ среди состоящих на учете, %

Закрытие полостей распада среди впервые выявленных, %

Прекращение бактериовыделения, %

Показатель клинического излечения больных ТОД, %

Показатель клинического излечения больных с внелегочными локализациями, %

Абациллирование, %

Число больных ФКТ легких, абс

МЛУ у впервые выявленных бациллярных больных (первичная МЛУ), %

МЛУ у бактериовыделителей, %

МЛУ на 100 тыс. населения

Летальность впервые выявленных больных туберкулезом, %

Летальность всех состоящих на учете больных туберкулезом, %

Летальность всех состоящих на учете больных туберкулезом от других причин, %

Отношение числа клинически излеченных от туберкулеза к числу умерших больных туберкулезом от всех причин

Рецидивы туберкулеза, абс

Частота рецидивов туберкулеза на 100 тыс. населения

Частота рецидивов туберкулеза у снятых с ДУ в связи с выздоровлением, на 100 тыс. населения

Частота рецидивов туберкулеза у пациентов III ГДУ на 100 тыс. III ГДУ

Частота рецидивов туберкулеза у снятых с ДУ в связи с выздоровлением, среди всех с рецидивами, %

Число инвалидов в связи с туберкулезом, абс

Доля инвалидов I группы, %

Доля инвалидов II группы, %

Доля инвалидов III группы, %

Доля инвалидов в связи с туберкулезом, среди всех состоящих на учете, %

Доля инвалидов в связи с туберкулезом на 100 тыс. населения, %

Впервые установленная инвалидность, абс

Вторичное заболевание туберкулезом больных ВИЧ-инфекцией, абс

Заболеваемость туберкулезом ВИЧ-инфицированных на 100 тыс. населения

Число заболевших туберкулезом, абс

Число заболевших туберкулезом, из них МБТ+, абс

Число умерших от туберкулеза, абс

Таблица 1. Первичные данные в формах Министерства здравоохранения РФ за год.

Эти данные доступны за несколько последних лет. Данные за прошлые годы характеризуются значительной неполнотой, обусловленной рядом объективных и субъективных обстоятельств, включая реорганизацию сети противотуберкулезных медицинских учреждений и государственных органов, смену отчетных форм и т.д. Основная форма хранения – архивы и опубликованные аналитические обзоры. Перевод на электронные носители фактически только начинается и разработка подсистемы “Статистическая эпидемиология” должна способствовать и этому процессу. Статистический анализ данных, в силу указанных причин, пока проводится на уровне отдельных показателей. Необходимость комплексного компьютерного анализа данных, включающего использование современных многомерных математических алгоритмов, очевидна.

1.2. Структура подсистемы «Статистическая эпидемиология»

Программа предназначена для информационного обеспечения мониторинга распространения туберкулеза, включая его множественные лекарственно-устойчивые формы, на территории Российской Федерации, моделирование и статистический анализ факторов, способствующих его распространению. Программа реализована в виде отдельной подсистемы «Статистическая эпидемиология». Основными компонентами подсистемы являются статистические данные, сценарии их обработки и средства визуализации получаемых результатов. Доступ к подсистеме осуществляется через Интернет.

Подсистема рассчитана на пользователей с различной компьютерной подготовкой. Начинающие пользователи могут обрабатывать имеющиеся данные, пользуясь заранее заготовленными сценариями, и просматривать полученные результаты в виде карт, графиков и таблиц. Пользователи с некоторым опытом общения с компьютером могут помещать свои таблицы в хранилище данных, редактировать имеющиеся сценарии (например, меняя имена файлов) и скачивать через Интернет результаты обработки в виде текстовых файлов и рисунков на свой локальный компьютер. Эксперты (продвинутые пользователи) могут создавать свои собственные сценарии обработки на основе имеющегося в подсистеме каталога команд и запускать их на выполнение.

1.3. Возможности подсистемы «Статистическая эпидемиология»

В подсистеме «Статистическая эпидемиология» предусмотрен набор возможных операций с таблицами - команд, например, объединение таблиц, вычисление относительных и интегральных характеристик, расчет корреляций между ними и т. д. Последовательность операций определяется сценарием обработки, записанном в виде xml-файла. Меняя сценарий, можно создавать гибкие варианты обработки таблиц, предназначенные для решения конкретных эпидемиологических задач.

1.4. Статистические методы в подсистеме «Статистическая эпидемиология»

Одномерный анализ статистических данных.

Вычисление основных статистических характеристик каждого показателя по выбранным регионам или годам: среднее, стандартное отклонение, коэффициент вариации, асимметрия, эксцесс.

Анализ относительных данных. Распространенным способом представления данных в медицинских исследованиях является отнесение их к 10000 или 100000 тысячам населения. Среднее значение относительного показателя по всем регионам в этом случае является естественной мерой «нормы», относительно которой можно судить о состоянии дел в конкретном регионе. Однако целесообразно вычислять его с учетом численности каждого региона, т.е., как средневзвешенное значение, иначе малочисленные регионы внесут в него непропорционально большой вклад, тем более, что данные по этим регионам, как правило, худшего качества и имеют гораздо большее рассеяние. Отклонение для каждого региона для сохранения масштаба целесообразно представлять в виде логарифма отношения значения для конкретного региона к средневзвешенному значению (чтобы разница в 2 раза или в $\frac{1}{2}$ раза отображалась одинаково). Кроме того, в этом случае знак логарифмического показателя сразу указывает на положение данного региона по отношению к средней по РФ: минус – ниже средней, плюс – выше средней.

Двумерный анализ статистических данных.

Вычисление по любым выбранным пользователем двум показателям коэффициента корреляции и линейной регрессии с отображением на графике.

Двумерный анализ относительных данных. Для относительных показателей – взвешивание на численность населения в регионе. При представлении относительных показателей в логарифмическом виде на двумерном графике диагональ графика дает

наглядное представление о степени превышения нормы одного относительного показателя по сравнению с другим.

Многомерный анализ статистических данных.

Статистические показатели часто коррелируют между собой. Поэтому можно предположить, что в основе их изменчивости лежит небольшое число общих факторов – интегральных показателей – и рассчитать эти факторы. В данном пакете для этой цели используется метод главных компонент. Как правило, имеет смысл рассматривать несколько первых интегральных показателей с наибольшим весом. Остальные характеризуются большим уровнем шума и их лучше исключать из анализа. (В этом пункте требуется очень тщательная предметная интерпретация. Сами по себе статистические соотношения ничего не говорят о породивших их причинах и возможных способах изменения ситуации).

Исследование динамики изменения статистических показателей.

Важной функцией программы является возможность исследования динамики роста либо спада с течением времени каких-либо значений в предоставляемой выборке данных, в простейшем случае – через линейную регрессию. С ее помощью можно получить представление о скорости роста ключевых для анализа параметров (например, уровня заболеваемости) и своевременно принять решение о необходимости вмешательства. Точно так же можно анализировать поведение обобщенных интегральных характеристик, полученных на предыдущем этапе.

1.5. Отображение результатов обработки.

Визуализация полученных результатов - карт, двумерных графиков и таблиц - осуществлена в виде растровых рисунков в png-формате. Это позволяет быстро перебрасывать их через Интернет и вставлять в любые отчетные формы, например, в документы Microsoft Office.

Основное предназначение описываемой программы заключается в наглядном и доступном для не имеющего специальную подготовку пользователя представлении исходных данных и получаемых в ходе работы результатов. Визуальный анализ данных, оформленных соответствующим образом (с включением диаграмм, показывающих значения параметров, отвечающих определенным географическим областям, а также с использованием схем цветовой градации), способен быстрее дать ответ на вопрос о необходимости и интенсивности вмешательства в развитие эпидемического процесса. При этом исследователю легче составить представление о ситуации в целом, чем при оперировании информацией, представленной в табличной форме. Для отображения информации по регионам используется карта, по годам – график с линейной линией тренда.

1.6. База данных

Статистические данные представлены в виде таблиц “регионы-характеристики” за ряд лет. Характеристики, как правило, берутся из отчетных форм Минздрава РФ, например, заболеваемость туберкулезом по регионам РФ за данный год. Однако подсистема может обрабатывать и любые другие характеристики, имеющие отношение к распространению туберкулеза, например, климатические или социально-экономические показатели. Для этого достаточно поместить соответствующие таблицы в хранилище данных.

Таблица состоит из блоков, например, блок 2003 года, блок 2004 года. Каждый блок состоит из строк и колонок (столбцов). Названия строк одинаковы для всех блоков.

Полезные ссылки:

Шилова М.В. Туберкулез в России в 2003 году. –М.: Фолиум, 2004. –106с.

Шилова М.В. Туберкулез в России в 2004 году. –М.: Фолиум, 2005. –108с.

Состояние противотуберкулезной помощи населению Сибирского и Дальневосточного федеральных округов по итогам работы в 2003 году. Аналитический обзор. –Новосибирск, НИИТ, 2004. –79с.