

ЭЛЕКТРОННАЯ СИСТЕМА ХРАНЕНИЯ, ПРЕДСТАВЛЕНИЯ И АНАЛИЗА ЭПИДЕМИОЛОГИЧЕСКОЙ ИНФОРМАЦИИ В МАСШТАБЕ ФЕДЕРАЛЬНОГО ОКРУГА

DOI: 10.17691/stm2017.9.4.21

УДК 528.94:616–036.22

Поступила 3.05.2017 г.

**Л.А. Солнцев**, к.б.н., зав. лабораторией ГИС-технологий и биоинформатики¹;**Н.Н. Зайцева**, к.м.н., руководитель Приволжского окружного центра по профилактике и борьбе со СПИД¹;**Е.И. Ефимов**, д.м.н., профессор, директор¹; зав. кафедрой микробиологии и иммунологии²¹Нижегородский научно-исследовательский институт эпидемиологии и микробиологии

им. академика И.Н. Блохиной Роспотребнадзора, Н. Новгород, 603950, ул. Малая Ямская, 71;

²Нижегородская государственная медицинская академия, Н. Новгород, 603005, пл. Минина и Пожарского, 10/1

Цель исследования — разработка эпидемиологического атласа, функциональным ядром которого является база данных эпидемиологического процесса актуальных инфекционных заболеваний в Приволжском федеральном округе (ПФО).

Материалы и методы. Эпидемиологический атлас состоит из следующих компонентов: банка данных, топографической основы, инструмента прямого доступа к данным и визуализатора эпидемиологической обстановки. Топографическая основа и банк данных формируют базу данных. Визуализатор и средства прямого доступа — графическую оболочку для работы с базой данных. Проект реализован в виде веб-сайта. Материалом для создания базы данных послужили отчеты субъектов ПФО по статистической форме №2 «Сведения об инфекционных и паразитарных заболеваниях». База данных находится под управлением системы MySQL, что позволяет хранить информацию в виде связанных таблиц. С помощью языка PHP информация преобразуется в сводные таблицы, графики временной динамики и картосхемы.

Результаты. Для организации процесса сбора, хранения и поддержания в актуальном состоянии базы данных эпидемиологического атласа разработан и внедрен в практику протокол обмена данными, созданы необходимые программные средства для полуавтоматической и автоматической обработки отчетов и их ввода в базу. В качестве обменного формата был выбран формат электронных таблиц XLS (XLSX).

В эпидемиологическом атласе реализованы следующие функции: автоматическая генерация отчетов из базы данных; анализ внутригодовой динамики инцидентности; оценка текущей помесечной динамики уровня заболеваемости по сравнению с ретроспективной.

Все генерируемые материалы можно экспортировать в виде таблиц в формате XLS (CSV) или графиков в форматах PNG, JPEG, PDF.

На сегодняшний день атлас представляет собой постоянно функционирующую и регулярно обновляющуюся систему, аккумулирующую данные по инфекционной заболеваемости по всем 14 субъектам ПФО с пространственным разрешением до административного района субъекта и временным разрешением в 1 мес.

Заключение. Аналогов, подобных «Эпидемиологическому атласу ПФО», в Российской Федерации на сегодняшний день не существует. Наш проект представляет собой завершенную геоинформационную систему, где сочетаются возможности доступа к исходным данным, получения аналитико-прогностической информации и экспорта результатов.

Ключевые слова: геоинформационные системы; эпидемиологический атлас; база данных эпидемиологической информации; мониторинг эпидемиологического процесса.

Как цитировать: Solntsev L.A., Zaitseva N.N., Efimov E.I. Electronic system of storage, representation, and analysis of epidemiological information on a federal district scale. Modern Technologies in Medicine 2017; 9(4): 170–176, <https://doi.org/10.17691/stm2017.9.4.21>

English

Electronic System of Storage, Representation, and Analysis of Epidemiological Information on a Federal District Scale

L.A. Solntsev, PhD, Chief of the Laboratory of GIS-Technologies and Bioinformatics¹;**N.N. Zaitseva**, MD, PhD, Head of the Privolzhsky District Center on Prophylaxis and Struggle against AIDS¹;

Для контактов: Солнцев Леонид Аркадьевич, e-mail: solntsev.l.a@nniem.ru

E.I. Efimov, MD, DSc, Professor, Director¹; Head of the Department of Microbiology and Immunology²

¹Blokhina Scientific Research Institute of Epidemiology and Microbiology of Nizhny Novgorod, Russian Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing (Rosпотребнадзор), 71 Malaya Yamskaya St., Nizhny Novgorod, 603950, Russian Federation;

²Nizhny Novgorod State Medical Academy, 10/1 Minin and Pozharsky Square, Nizhny Novgorod, 603005, Russian Federation

The aim of the study was to work out an epidemiological atlas, the functional core of which is the database of the epidemiological process of actual infectious diseases in Privolzhsky Federal District (PFD).

Materials and Methods. The epidemiological atlas consists of the following components: a data bank, topographic base, direct data access tool, and visualizer of the epidemiological situation. The topographic base and data bank form a database, while the visualizer and direct access tools form a graphical shell for working with the database. The project was implemented as a website. Reports of PFD subjects filled-in in compliance with form No.2 "Information about infectious and parasitic diseases" served as the material for creating the database. The database is controlled by MySQL system, which provides the possibility to store the information in the form of linked tables. Using the PHP language the information is converted into pivot tables, diagrams of time dynamics, and schematic maps.

Results. To organize the process of acquisition, storage and maintaining the epidemiological atlas database in the running condition, a protocol of data exchange was designed and implemented into practice, the necessary software was created for semi-automatic and automatic report processing and inserting into the database. XLS (XLSX) format of the electronic tables was chosen as an exchange format.

The following functions are realized in the epidemiological atlas: automatic generation of reports from the database; the analysis of intra-annual incidence dynamics; evaluation of the current monthly dynamics of the morbidity rate compared to the retrospective information.

All generated materials can be exported in the form of XLS (CSV) tables or diagrams as PNG, JPEG, PDF files.

At present, the atlas represents a constantly functioning and regularly updated system accumulating data on infectious morbidity rate over all 14 PFD subjects with a spatial resolution to the administrative region subject and a time resolution of 1 month.

Conclusion. At present, there are no analogs to the "PFD Epidemiological Atlas" in the Russian Federation. Our project is a complete geoinformation system, where the possibilities of the initial data access, analytical and prognostic information acquisition, and export of the results are combined.

Key words: geoinformation systems; epidemiological atlas; epidemiological information database; epidemiological process monitoring.

Важнейшим критерием, характеризующим здоровье населения, является заболеваемость (инцидентность). Учет показателей инцидентности ведется на регулярной основе с высокой степенью детализации во всех медицинских организациях Российской Федерации, которые передают информацию в органы и учреждения Роспотребнадзора (приказ Росстата от 28 января 2014 г. №52, приказ Роспотребнадзора от 2 декабря 2014 г. №1203).

Собираемая информация отвечает требованиям, предъявляемым к пространственным данным [1]:

имеет пространственную привязку (населенный пункт, административный район, область, федеральный округ, где было зарегистрировано событие (случай заболевания)) и временную локализацию (день, месяц, год регистрируемого события);

имеет атрибутивную часть, которая содержит сведения о случаях регистрации той или иной инфекции в данном объекте и/или любую иную информацию (численность населения, качество водоподготовки, локализация в природно-географической зоне и т.д.);

имеет в качестве «субстрата» карту изучаемой (мониторируемой) территории, отдельные элементы которой (населенные пункты, административные рай-

оны, области, федеральные округа) выступают в роли объектов наблюдения.

Следовательно, отсутствуют теоретические и методические препятствия для использования геоинформационных систем и модели пространственных данных в качестве инструментов эпидемиологического анализа. Разработка таких систем открывает возможность проведения полуавтоматического или автоматического анализа собранной информации с помощью методов и алгоритмов, используемых в практике работы специалистов.

Цель исследования — разработка эпидемиологического атласа, функциональным ядром которого является база данных эпидемического процесса актуальных инфекционных заболеваний в Приволжском федеральном округе.

Для достижения поставленной цели необходимо было решить следующие задачи:

1. Разработать протокол получения исходных данных из центров гигиены и эпидемиологии (ЦГиЭ) и/или территориальных управлений Роспотребнадзора в субъектах Приволжского федерального округа (ПФО).

2. Спроектировать и реализовать структуру базы данных для хранения информации.

3. Создать графическую оболочку для работы с базой.

Материалы и методы

Исходные данные. На сегодняшний день территориальные органы и учреждения Роспотребнадзора ПФО для учета заболеваемости применяют разное программное обеспечение: программа «Анализ популяционной заболеваемости» (АПЗ); программа АС «СТАТИСТИКА РОСПОТРЕБНАДЗОР» (КРИСТА); программные решения, разработанные собственными силами учреждений. В каждом программном продукте данные хранятся в собственном формате, поэтому в ряде доступ к исходной информации возможен только в том программном обеспечении, в котором она была создана. Однако научная работа предполагает обращение именно к исходным данным, а также применение автоматизированных алгоритмов обработки информации, что предъявляет определенные требования к формату самих данных. Такой формат мы далее будем называть «машиночитаемый» (т.е. информация хранится в виде единообразно структурированных таблиц и т.п.). В настоящее время в системе учета заболеваемости Роспотребнадзора принцип хранения и предоставления данных в машиночитаемом виде редко реализуется на практике в полном объеме. Следует отметить, что экспортировать в эти программы информацию можно только в стандартизованных формах. Данное обстоятельство, учитывая характер программного обеспечения и ставящиеся перед ним задачи, не является недостатком, однако обуславливает создание неких промежуточных средств для конвертации данных из стандартизированных отчетов в машиночитаемый вид.

Исходным материалом для создания базы эпидемиологического атласа послужили отчеты по статистической форме №2 «Сведения об инфекционных и паразитарных заболеваниях», формирование которых

является одной из задач ЦГиЭ. Отчеты конвертировали в формат, пригодный для прямого экспорта в базу.

Архитектура системы. Эпидемиологический атлас состоит из следующих компонентов: банка данных, топографической основы, инструмента прямого доступа к информации визуализатора эпидемиологической обстановки. Топографическая основа и банк данных формируют базу данных. Визуализатор и средства прямого доступа — графическую оболочку для работы с базой. Общая архитектура атласа представлена на рис. 1.

Наполнение банка данных идет за счет информации, предоставляемой как ЦГиЭ и территориальными управлениями Роспотребнадзора, так и научными организациями, и референс-центрами (для специализированных разделов эпидемиологического атласа). Функционал визуализации эпидемиологической обстановки ориентирован на надзорные органы Роспотребнадзора. Средства прямого доступа к данным дают возможность получать информацию в машиночитаемом виде.

Функциональное ядро эпидемиологического атласа составляет база данных под управлением системы MySQL, что позволяет хранить информацию в виде связанных таблиц. С помощью языка PHP данные преобразуются в сводные таблицы, графики временной динамики и картосхемы. При этом пользователь не ограничен видами анализа и представления данных, заложенными разработчиками, он всегда имеет возможность получить доступ к исходной информации. Эпидемиологический атлас является законченной системой, т.е. помимо функции хранения и представления информации, имеет интерфейс для ввода новых данных, которые автоматически обрабатываются в системе и учитываются при формировании обобщающих материалов.

На рис. 2 приведена схема организации данных в базе применительно к отдельному субъекту ПФО.

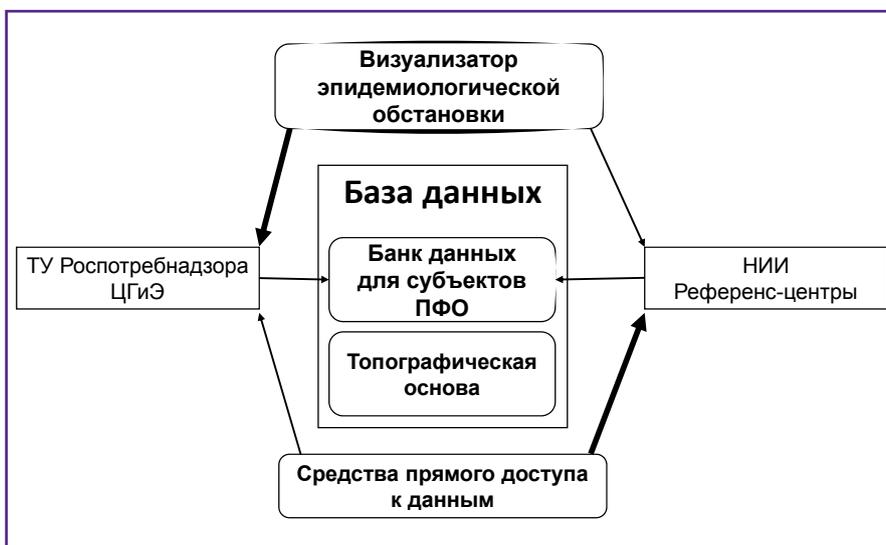


Рис. 1. Общая архитектура атласа
Стрелками обозначены пути движения информации в системе: жирные стрелки — пути, которые наиболее востребованы ТУ Роспотребнадзора и НИИ соответственно

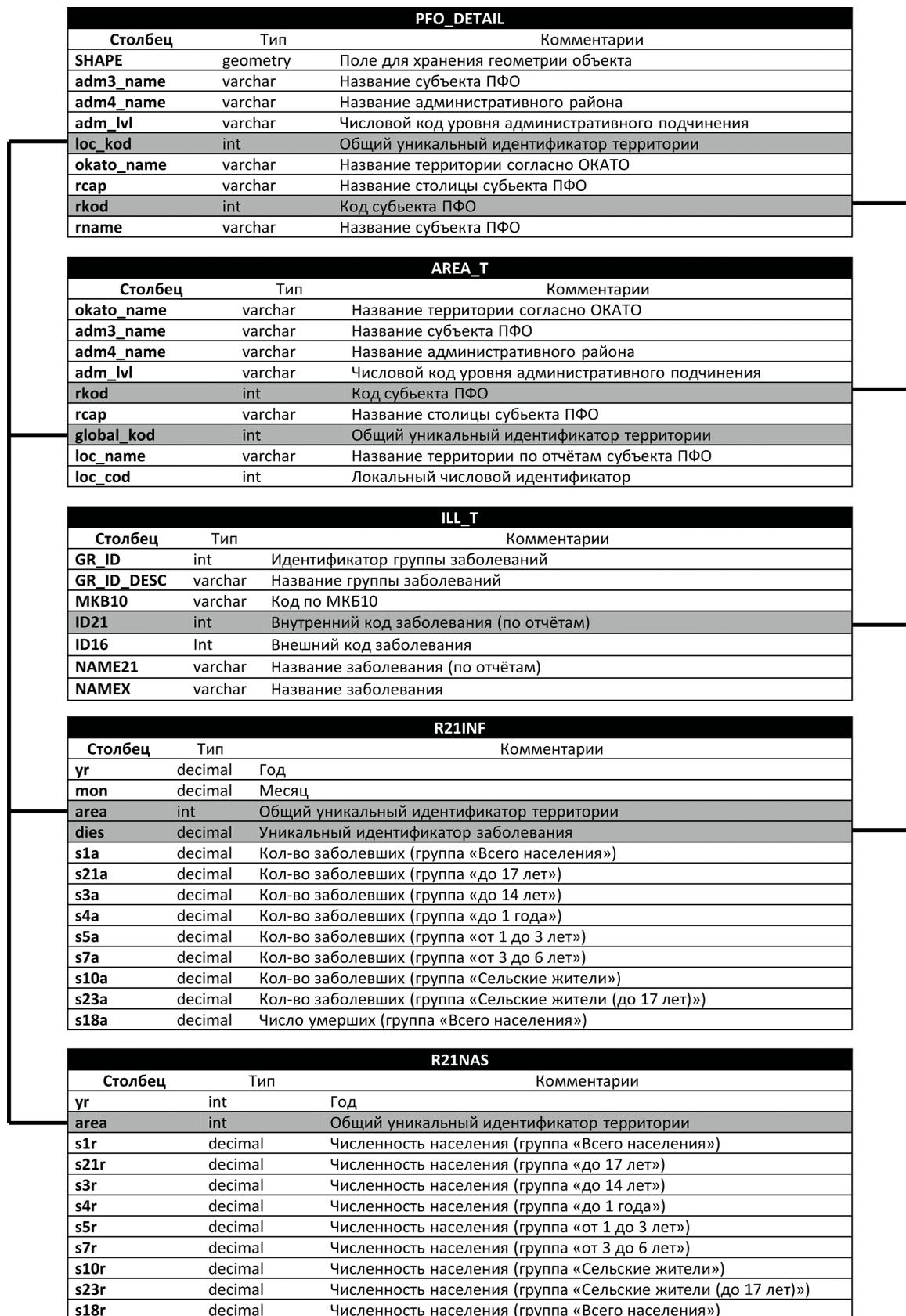


Рис. 2. Схема взаимосвязей таблиц данных в модуле отдельного субъекта ПФО (на примере Чувашской республики) PFO_DETAIL — пространственные данные по границам административных единиц субъекта; AREA_T — справочник административных единиц субъекта ПФО; ILL_T — справочник инфекций; R21INF — данные по числу случаев; R21NAS — данные по численности населения

Результаты

Ввод и хранение данных. Для организации процесса сбора, хранения и поддержания в актуальном состоянии базы эпидемиологического атласа разработан и внедрен в практику протокол обмена данными между территориальными органами и учреждениями Роспотребнадзора и Нижегородским научно-исследовательским институтом эпидемиологии и микробиологии им. академика И.Н. Блохиной Роспотребнадзора, созданы необходимые программные средства для полуавтоматической и автоматической обработки отчетов и их ввода в базу данных. В качестве обменного формата выбран формат электронных таблиц XLS (XLSX). Данный формат доступен как средство экспорта во всех программных продуктах, используемых территориальными управлениями ПФО, и удобен для автоматической обработки и преобразования данных. По состоянию на апрель 2017 г. получены и обработаны архивные массивы данных (глубина до 2010 г.) и ежемесячные отчеты, которые в дальнейшем стали основной формой обмена информацией. В настоящее время процесс получения и обработки данных уже является рутинной процедурой, не требующей дополнительной корректировки. При этом осуществляется двойная система контроля корректности ввода: проверяется формальное соответствие структуры отчета исходно заданным критериям (расположение и название полей и т.д.) и проводится выборочный обратный контроль — полученные после конвертации отчетов значения сравниваются с исходными. Кроме того, разработаны инструментарий для точечной коррекции уже введенных данных, а также инструмент для преобразования полученных таблиц в SQL-транзакции.

Тематическое картографирование. Основным способом представления данных является тематическое картографирование и автоматическое районирование территории по атрибутивному признаку (уровню заболеваемости и т.д.). При этом чрезвычайно важен выбор масштаба районирования (для более точного выделения проблемных территорий нужно перейти от уровня субъектов ПФО до уровня административных единиц этих субъектов) и метода, по которому территориальные единицы будут объединяться в группы. При использовании для районирования одного признака оптимальным является применение одного из простых методов классификации. В эпидемиологическом атласе пользователь имеет возможность выбрать один из наиболее распространенных методов: метод равных интервалов, метод квантилей, метод геометрических интервалов, метод натуральных интервалов. В случаях, когда требуется провести районирование по набору признаков (как качественных, так и количественных), необходимы более сложные методы, такие как дискриминантный анализ, методы машинного обучения и т.д. Возможна комбинация нескольких параметров на одной карте-схеме. Например, можно отметить предполагаемые границы очага (согласно данным, предоставленным ЦГиЭ),

отобразить сопутствующую информацию и провести группировку районов по уровню инцидентности/распространенности того или иного природно-очагового заболевания.

Анализ данных. На сегодняшний день в эпидемиологическом атласе реализованы следующие возможности:

автоматическая генерация отчетов из базы данных (например, сравнительный анализ по текущему и прошлому году и т.д.) с функцией модификации полученного результата по той или иной утвержденной форме отчетности;

анализ внутригодовой динамики инцидентности. Например, данные о вкладе отдельных месяцев в общее число заболевших за год можно представить в виде «тепловой карты» и в виде линейного графика. Еще один вариант — это анализ внутригодовой структуры заболеваемости [2] и выделение отдельных функциональных компонент ряда динамики заболеваемости [3, 4]. Последний метод, насколько нам известно, в отечественной эпидемиологической практике применяется редко;

оценка текущей помесечной динамики уровня заболеваемости по сравнению с ретроспективой. Это позволяет прогнозировать уровень развития заболеваемости текущего года, ориентируясь на разброс значений данного месяца за предыдущие годы.

Все генерируемые материалы можно экспортировать в виде таблиц в формате XLS(CSV) или графиков в формате PNG, JPEG, PDF. Графического материала достаточно для использования в отчетах и докладах.

Регистрация и внедрение базы данных в практику. База данных эпидемиологического атласа прошла процедуру государственной регистрации, заявка удовлетворена отделом регистрации программ для ЭВМ, баз данных и топологий интегральных микросхем Федерального института промышленной собственности Федеральной службы по интеллектуальной собственности (Москва), получено Свидетельство о государственной регистрации базы данных «Электронный эпидемиологический атлас Приволжского федерального округа» от 24.07.2015 г. №2015621143. Использование эпидемиологического атласа в практической работе одобрено приказом Роспотребнадзора №283 от 13.04.2015 г. «О реализации решения коллегии от 06.03.2015 г. №2 «О деятельности ФБУН «Нижегородский научно-исследовательский институт эпидемиологии и микробиологии им. академика И.Н. Блохиной» Роспотребнадзора за 2012–2014 гг.»; приказом Роспотребнадзора №1129 от 23.10.2015 г. «О проведении регионального совещания «Изучение положительного опыта использования электронного эпидемиологического атласа ПФО в организации и проведении эпиднадзора за актуальными инфекциями»; приказом Роспотребнадзора №1209 от 16.11.2015 г. «О создании регионального (окружного) центра по мониторингу за инфекционными болезнями ПФО на базе ФБУН «Нижегородский НИИЭМ им.

акад. И.Н. Блохиной Роспотребнадзора»; протоколом совещания «О дополнительных мерах по реализации поручений Правительственной комиссии по охране здоровья граждан по вопросу предупреждения распространения ВИЧ-инфекции в Российской Федерации» (8–10 декабря 2015 г., Суздаль).

Внедрение эпидемиологического атласа в практическую работу подтверждено актами управлений Роспотребнадзора ЦГиЭ в субъектах ПФО (Башкортостан, Мордовия, Татарстан, Марий Эл, Удмуртская и Чувашская Республики, Нижегородская, Пензенская, Саратовская, Самарская, Ульяновская, Пензенская, Кировская области, Пермский край).

Обсуждение. В Российской Федерации есть системы, которые внедрены в повседневную практику и условно доступны специалистам

Разработка Ставропольского научно-исследовательского противочумного института Роспотребнадзора [5] — картографическая программа ZikaMap для сбора и анализа информации о мероприятиях по мониторингу и регуляции численности комаров-переносчиков вирусных заболеваний, в том числе лихорадки Зика (в настоящее время используется на территории Краснодарского края). Эта программа в большей степени ориентирована на оперативный ввод данных энтомологического мониторинга и его визуализацию, а не на анализ. Кроме того, она является узкоспециализированной системой.

Разработка Ростовского научно-исследовательского противочумного института Роспотребнадзора [6] — геоинформационный портал по инфекционным болезням (холера, сибирская язва, клещевая геморрагическая лихорадка). Он представляет собой просто визуализацию на карте случаев регистрации инфекций из списка.

Разработка Российского научно-исследовательского противочумного института «Микроб» Роспотребнадзора [7] — портал «Оздоровительные мероприятия в Горно-Алтайском высокогорном природном очаге чумы» — наиболее проработанная система. Карта портала оснащена легендой, в которой содержится информация о задействованных слоях. На электронной карте имеется информация: ФИО, возраст, статус, дата заболевания, дата контакта, место предполагаемого заражения, форма заболевания, лабораторное подтверждение, метод лабораторного подтверждения, дата лабораторного подтверждения. Помимо этого, портал содержит информацию о 189 медицинских учреждениях Республики Алтай, включая сведения об адресе учреждения, количестве медицинского персонала (в том числе специалистов по особо опасным инфекциям), наличии боксированных палат.

Среди зарубежных систем можно отметить порталы WHO [8] и CDC [9]. Данные порталы предназначены для представления глобальной информации (до уровня страны в [8] или сугубо для административных единиц США [9]). Зарубежные системы ориентированы на небольшую территорию, их инструментарий

ввода, хранения и визуализации данных, авторам не известен. Скорее всего, подобные проекты ведутся специализированными организациями и представляются собой закрытые отраслевые решения.

Созданный на базе Нижегородского научно-исследовательского института эпидемиологии и микробиологии им. академика И.Н. Блохиной Роспотребнадзора геоинформационный проект «Эпидемиологический атлас Приволжского федерального округа» (<http://epid-atlas.nniiem.ru>) является уникальной общедоступной разработкой подобного класса на территории РФ. На сегодняшний день атлас представляет собой постоянно функционирующую и регулярно обновляющуюся систему, аккумулирующую данные по инфекционной заболеваемости по всем 14 субъектам ПФО с пространственным разрешением до административного района субъекта и временным разрешением в 1 мес. Глубина архива — с января 2010 г. по настоящее время. Функционал эпидемиологического атласа позволяет реализовать все этапы эпидемиологического анализа — от получения разрозненных данных с субъектов округа до представления сводной аналитико-прогностической информации.

На основе эпидемиологического атласа возможно создание целого класса более узкоспециализированных систем (надзор за природно-очаговыми заболеваниями, мониторинг противоэпидемических мероприятий и т.д.). Так, на сегодняшний день функционирует специализированный раздел по мониторингу ВИЧ-инфекции (http://epid-atlas.nniiem.ru/a03_data_main_aid.html).

Внедрение эпидемиологического атласа в практику работы как научно-исследовательских, так и надзорных учреждений Роспотребнадзора позволяет:

- снизить количество запросов в территориальные органы;
- уменьшить время получения данных заинтересованными организациями;
- автоматизировать процесс подготовки данных к анализу;
- использовать готовые решения при типовом анализе.

Разработанный геоинформационный проект отвечает требованиям государственной политики по внедрению новых технологий в практику работы учреждений Роспотребнадзора и органов здравоохранения.

Заключение. Аналогов, подобных «Эпидемиологическому атласу Приволжского федерального округа», в Российской Федерации на сегодняшний день не существует. Наш проект представляет собой завершённую геоинформационную систему, где сочетаются возможности доступа к исходным данным, получения аналитико-прогностической информации и экспорта результатов.

Благодарности. Авторы выражают благодарность управлениям Роспотребнадзора по субъектам ПФО и центрам гигиены и эпидемиологии в субъектах ПФО за помощь в подготовке первичных материалов для базы данных.

Финансирование исследования и конфликт интересов. Исследование не финансировалось какими-либо источниками, и конфликты интересов, связанные с данным исследованием, отсутствуют.

Литература/References

1. Геоинформатика. Под ред. Тикунова В.С. М: Издательский центр «Академия», 2010. *Geoinformatika* [Geoinformatics]. Pod red. Tikunova V.S. [Tikunov V.S. (editor)]. Moscow: Izdatel'skiy tsentr "Akademiya", 2010.
2. Основы ретроспективного анализа инфекционной заболеваемости. Под ред. Шкарина В.В., Рахманова Р.С. Н. Новгород: Изд-во НижГМА; 2009. *Osnovy retrospektivnogo analiza infektsionnoy zbolevaemosti* [Fundamentals of retrospective analysis of infectious diseases]. Pod red. Shkarina V.V., Rakhmanova R.S. [Shkarin V.V., Rakhmanov R.S. (editors)]. Nizhny Novgorod: Izd-vo NizhGMA; 2009.
3. Cleveland R.B., Cleveland W.S., McRae J.E., Terpenning I. STL: a seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics* 1990; 6(1): 3–73.
4. Филатова Е.Н., Солнцев Л.А., Уткин О.В. Влияние сезонных факторов на динамику уровня заболеваемости инфекционным мононуклеозом в разных возрастных группах (на примере Нижнего Новгорода). *Эпидемиология и инфекционные болезни* 2017; 22(2): 79–85. Filatova E.N., Solntsev L.A., Utkin O.V. The impact of seasonal factors on the dynamics of incidence rate of infectious mononucleosis in different age groups (on the example of Nizhny Novgorod). *Epidemiologiya i infektsionnye bolezni* 2017; 22(2): 79–85.
5. *ZikaMap*. URL: <http://snipchi.ru/page.php?326>.
6. Геоинформационный портал. *Geoinformation portal*. URL: <http://gis.antiplague.ru/index.php>.
7. Оздоровительные мероприятия в Горно-Алтайском высокогорном природном очаге чумы. *Sanitary measures in Gorno-Altaysk natural focus of plague*. URL: <http://www.microbe.ru:6655/altaymap/>.
8. WHO. URL: <http://www.euro.who.int/en/health-topics/disease-prevention/vaccines-and-immunization/publications/surveillance-and-data/who-epidata>.
9. CDC. URL: <https://www.cdc.gov/surveillancepractice/data.html>.