

Database

Open Access

Implementing a public web based GIS service for feedback of surveillance data on communicable diseases in Sweden

Per Rolfhamre¹, Katarzyna Grabowska¹ and Karl Ekdahl*^{1,2}

Address: ¹Department of Epidemiology, Swedish Institute for Infectious Disease Control, Nobels väg 18, SE-17182 Solna, Sweden and ²Department of Medical Epidemiology and Biostatistics, Karolinska Institute, SE-17177 Stockholm, Sweden

Email: Per Rolfhamre - per.rolfhamre@smi.ki.se; Katarzyna Grabowska - kasia.grabowska@smi.ki.se; Karl Ekdahl* - karl.ekdahl@smi.ki.se
* Corresponding author

Published: 10 June 2004

Received: 25 February 2004

BMC Infectious Diseases 2004, **4**:17

Accepted: 10 June 2004

This article is available from: <http://www.biomedcentral.com/1471-2334/4/17>

© 2004 Rolfhamre et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Surveillance data allow for analysis, providing public health officials and policy-makers with a basis for long-term priorities and timely information on possible outbreaks for rapid response (data for action). In this article we describe the considerations and technology behind a newly introduced public web tool in Sweden for easy retrieval of county and national surveillance data on communicable diseases.

Methods: The web service was designed to automatically present updated surveillance statistics of some 50 statutory notifiable diseases notified to the Swedish Institute for Infectious Disease Control (SMI). The surveillance data is based on clinical notifications from the physician having treated the patient and laboratory notifications, merged into cases using a unique personal identification number issued to all Swedish residents. The web service use notification data from 1997 onwards, stored in a relational database at the SMI.

Results: The web service presents surveillance data to the user in various ways; tabulated data containing yearly and monthly disease data per county, age and sex distribution, interactive maps illustrating the total number of cases and the incidence per county and time period, graphs showing the total number of cases per week and graphs illustrating trends in the disease data. The system design encompasses the database (storing the data), the web server (holding the web service) and an in-the-middle computer (to ensure good security standards).

Conclusions: The web service has provided the health community, the media, and the public with easy access to both timely and detailed surveillance data presented in various forms. Since it was introduced in May 2003, the system has been accessed more than 1,000,000 times, by more than 10,000 different viewers (over 12.600 unique IP-numbers).

Background

Surveillance of mandatory notifiable diseases has evolved as a corner stone of public health systems for control of infectious diseases in many countries [1]. Notification data allow for analysis of secular trends, providing public health officials and policy-makers with a basis for long-

term prioritizations and other decisions. An equally important use of the surveillance system is to give timely information on possible outbreaks – information used for rapid response.

The organization for collecting data differs between the countries, and so does the number of diseases being notifiable in single countries. In many European countries, the flow of information goes from clinicians and/or laboratories to local health officials and thereafter on regular intervals to the national surveillance institute, sometimes through a regional authority [2].

Regardless of the set-up of the surveillance system, it is of great importance that the data reach those in need of the information in a timely matter – data for action. Also, it is crucial, that the data can be easily interpreted and operated in various ways, without compromising confidentiality [3]. Modern computerized surveillance systems in combination with the general availability of the Internet could be useful tools in ensuring that relevant and updated surveillance information is always at hand where it is needed [4].

In this article we describe the considerations and technology behind a newly introduced public web tool in Sweden for easy retrieval of county and national surveillance data on communicable diseases, given as tables, maps, and trend curves. An increasing demand on detailed data and an ambition to present the data providers with more timely data for action were the driving forces behind the decision to use modern web technology and a geographical information system (GIS) software [5-7] to improve the feed-back of surveillance information.

Methods

The data

The statutory surveillance of communicable diseases is regulated in the Communicable disease act of 1989. The 54 statutory notifiable infectious diseases are divided, into "Diseases dangerous to society" (Groups 1.1, 1.2 and 1.3) and "Other notifiable diseases" (groups A.1 and A.2). All notifications are submitted both to the County medical officer (one in each of the 21 Swedish counties) and to the Swedish Institute for Infectious Disease Control (SMI). Diseases are notified both by the patient's physician (clinical notification) and the laboratory, having diagnosed the causative agent (laboratory notification). The notifications are required to be done within 24 hours of diagnosis (seven days for sexually transmitted infections – Group 1.3), and the clinical report shall contain information of epidemiological relevance, including country of infection, if known. The arrival date to the SMI of the first notification belonging to a case is used when presenting data.

A national computerized notification system (SmiNet), containing notification data from 1997 onwards, is used for feeding the national surveillance database (EpiArk). At the SMI, clinical and laboratory notifications from the same patient are merged, using the personal identification

number, issued to all residents in Sweden and used in all contacts with the health care sector. The personal identification number consists of the person's date of birth, and four additional digits, of which the third gives information on the sex of the patient; even figure for women and uneven for men (YYMMDD-XXXX). All diseases with the exception of those in Group 1.3 are notified with full patient identification. The data are also "cleaned" for duplicates and irregularities before being fed into the EpiArk database.

All data gathered are controlled for quality in several independent steps. 1) When data are inputted into the SmiNet, there are input field control and input field dependency validation. 2) When a notification arrives to the county medical officer and/or the SMI, the information is gone through by the officer responsible for the reported diagnosis. 3) At regular intervals the administrator of the EpiArk searches all database tables for obvious defects in the data, such as impossible date intervals, impossible birth dates or duplicates.

Software and hardware

The application is written in Java™, using third party packages: The Big Faceless™ Graph library (version 1.0.13) [8], JavaMail™ API (version 1.2) [9], JavaBeans™ Activation Framework (version 1.0.1) [10], Apache Ant (version 1.4.1) [11], i-net MERLIA 2000 (version 1.03) [12], ArcIMS™ (version 4.0) [13] and Java™ 2 SE Platform (version 1.4.1_01) [14].

The EpiArk is stored at a Compaq Prolinea 800 with 256 Mb running Microsoft Windows 2000 and Microsoft SQL server 6.5.

A supporting computer (generating computer) was also used, namely a Compaq Deskpro SFF P3-1.0 GHz with 256 Mb ram running Microsoft Windows 2000.

The web server is a Compaq EVO D510CMT P4-2.0 GHz with 768 Mb ram running Microsoft Windows 2000 and Internet Information Server (IIS).

Results

Basic application design

To ensure a good security standard, the web server could not be directly connected to the EpiArk. The web server is open to public access, while the EpiArk must be protected inside. Therefore a supporting computer, in the middle (generating computer), was needed to filter the access to the database. Since we wanted to optimize the web server performance for incoming requests, it was suitable to let the generating computer, in the middle, do as much work as possible. Subsequently, the generating computer creates all the tables, graphs, maps and statistics later stored

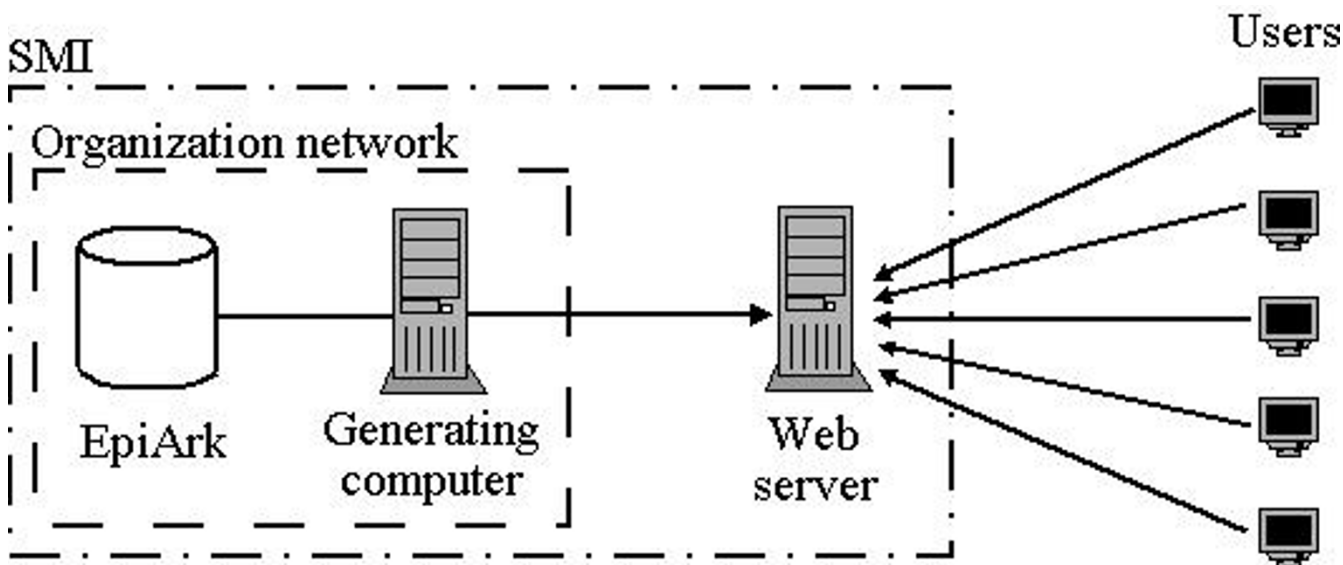


Figure 1
The system design. The generating computer starts the process by retrieving data from the EpiArk, generates the web service and copies it to the web server.

in the web server. At generation time, the application requests information from the EpiArk to retrieve and calculate the needed statistics. When the web pages are generated, they are copied (one-way connection) to the web server. The design is illustrated in Figure 1.

The work flow in eight steps

This procedure gives an overview of the application work scheme (GC – Generating computer, S – Web server).

1. (GC) Start up the environment. This practically means set up the error management, by creating and opening a log file, load the environment variables and setting all global variables for later usage.
2. (GC) Load the diseases to generate and their properties from a configuration file to a global storage structure.
3. (GC) Remove the old file system, to start from scratch. Copy all static files from a template directory to the file system, for each disease to generate. The file system is now complete apart from the files that need to be generated.
4. (GC) Open a connection to the EpiArk. Make queries for disease data and generate all non-static files into the file system, for each disease to generate. Close the connection to the EpiArk.

5. (S) Shut down all previous ArcIMS services, locking the active map files.
6. (S) Delete the old file system and copy the new from the generating computer.
7. (S) Start up the ArcIMS services, to be able to use the maps from the web.
8. (GC) Generate a summery report and mail to the administrator together with the log file. Close down the environment.

The downtime for the web server consists of step 6, except for the interactive maps whose downtime consists of step 5–7.

Database structure

The basic internal structures of EpiArk consist of 12 tables and 11 help tables. On February 19, 2004, the main table, which lists all notified cases, in Sweden, since 1997, contained about 340,000 cases. These cases are divided into 53 notifiable diseases. HIV and AIDS data are registered in a separate database; hence the system cannot generate these data automatically.

Duplicate notifications and invalid notifications are removed before importing the cases to the main table

from the SMI, and all information, including changed and removed cases, must be traceable (logged), which makes the database structure complex.

The application uses information from three database tables: case information from the main database table, country of origin of infection data from the database table holding information of the clinical notifications and information from the population database table.

Online user interface

In order to enable international usage of the Swedish communicable disease statistics, all web pages are presented in both Swedish and English. The following statistics are published on the SMI web site (screenshots are presented in Figure 2 – Figure 8).

- **Table: Yearly and monthly statistics per county.** The data is presented in a crude manner. It consists of the number of reported cases (and incidence per 100,000 population) in a defined period (month or year from January 1997). This data is presented together with counts from previous years for comparison. The presented data sets can by one mouse click be changed between 1) all cases, 2) cases infected in Sweden, and 3) cases infected abroad (Figure 3).
- **Table: Age distribution.** Distribution of age per year. The data is based on all notified cases, regardless of country of infection.
- **Table: Sex distribution.** Distribution of sex per year. The data is based on all notified cases, regardless of country of infection.
- **Interactive maps.** Shows a map of Sweden, divided into the 21 Swedish counties, with county colors graded after the current measurement units. There are two types of maps; one for the total number of cases and one for the incidence (cases per 100,000 population and time period). The data set of the statistics can be changed between 1) all cases and 2) cases infected in Sweden. By clicking on a county, the user will retrieve detailed statistical information about the area (Figure 4).
- **Graph: Number of cases per week.** Graph showing the total number of cases per week during the last five years, in a crude manner.
- **Graph: Trend (moving average).** A moving average is the average number or reported cases per week at a specific point in time. The purpose of the moving average is to show the trend in a smoothed fashion. The graph displays two moving average curves: one for seasonal trends (based on six weeks) and one for yearly trends (based on

52 weeks). The algorithm for calculating the moving average is presented in algorithm 1 [see 1 'algorithm1.txt'], (Figure 5).

- **Country of infection facts.** Compilations for all diseases of the proportions of all infected that are infected in Sweden and infected abroad per year, as well as a top-15 list of the most frequent countries of infection for each year.
- **Manual web pages.** The SMI can create any kind of statistical web information and link from the disease page as 1) a table, 2) an interactive map, 3) a graph, or 4) country of infection facts.

All tabulated information can be downloaded for an alternative analysis; furthermore, the user can save all maps and graphs as images.

File generation

There are two directories crucial for file generation – the build directory and the template directory. These directories form the foundation of the file generation process.

The build directory is located on the generating computer as well as on the web server (mirrored). This directory is copied/updated to the web server after generation. It contains configuration files, data files for the map generation as well as the web pages, located in a directory mounted on the web server only (a file directory connected to the web server, making it available through the Internet).

The template files of the application are divided into two separate parts – one static for all diseases and one that must be generated once for each disease. Files like for instance help files, redirect files and some ArcIMS map files etc. are static for all diseases and can therefore be copied into the build file structure without any manipulation. The rest of the files can not be static, because that the disease name, statistical values or web URLs etc. needs to be inputted to a template file. The application implements methods for file generation as a process of exchanging a predefined delimiter with the actual value. These files need to be generated, one by one, into the build file structure, after interaction with the disease configuration file and the EpiArk.

Usage

The updating process takes about 90 minutes in total, which is divided into 60 minutes of generation (40% of which are response time of EpiArk) and 30 minutes of updating the web server. The updating consists of ArcIMS service shut down (2 minutes), deleting and copying the generated files (about 390 MB of data) (7 minutes) and restart of the ArcIMS services (20 minutes). This makes the downtime of the web server 7 minutes, except for the




Hepatitis B

 Notifiable Diseases

The information is based on cases notified both from the clinicians and the laboratories. An arbitrary case can be notified from either the clinician or the laboratory. It can also be notified from both instances, counted as one case.

All statistics are updated continually. Consequently early statistics may change with time.

Explanation and Help

-  Approach
-  Explanation of concepts
-  Example of Analysis

Statistics and Analysis











-  Table: Year and month statistics per county
-  Table: Age distribution
-  Table: Sex distribution
-  Map: All notified cases per county
Requires Internet Explorer version 4 or Netscape version 6 or later.
-  Map: All notified cases per 100.000 population per county and year
Requires Internet Explorer version 4 or Netscape version 6 or later.
-  Map: Domestic cases per county
Requires Internet Explorer version 4 or Netscape version 6 or later.
-  Map: Domestic cases per 100.000 population per county and year
Requires Internet Explorer version 4 or Netscape version 6 or later.
-  Graph: Number of cases per week
-  Graph: Trend (moving average)
-  Country of infection facts

Figure 2
 Screenshot of the disease main page. This page is presented for all notifiable diseases.

maps who have a 30-minute downtime. This is hardware dependent, as well as dependent of the implementation. About 34.000 files are generated and copied from the generating computer to the web server.

During first nine months of the system (May 2003 – January 2004), the web service page of the surveillance statistics recorded 1,030,000 hits in all, and was visited by over 12.600 unique IP-addresses (more than twice the

estimated overall number of persons working with communicable disease control in any capacity in Sweden).

Discussion

The surveillance data in Sweden, has traditionally been fed back monthly to the health community and public as aggregated number of notifications per county and month, through the SMI website and the monthly SMI bulletin, "Smittskydd". More in depth analysis and data broken down on age groups, gender, and country of infection has been published once a year only in the Annual report of the Department of epidemiology [15]. In between these annual reports, detailed data has only been available on special request, and each such request has generated substantial manual extra work. The system described in this article has in major ways improved and complemented these previous sources of information.

Design principles

When designing the application, it was crucial to decide some basic design principles to build the application upon:

1. Good security standards. The information in the EpiArk contains full patient identity, linked to a disease diagnosis, and the demands on security are therefore extremely high. If someone manages to hack into the web server, he/she must not be able to retrieve any more information about the data than from using the web service as intended.

2. Straightforward modification of the application and its behavior. There was a requirement of being able to manually change or add data between automatic updates as well as easily make adjustments in the database structure.

3. Intuitive presentation of the statistics. One major consideration in publishing a web statistics service like the application is that different target groups has widely different skills and abilities to understand and interpret the data correctly. This puts extra demand on the way statistics is presented, as well as on the user interface. The SMI uses a third party product to publish the interactive data, and it was seen as important that a user should not be able to distinguish the application as a different entity from the main web site.

Privacy and security considerations

As mentioned, confidentiality was the foremost priority when designing the application. This concept includes ensuring that no single patient identity could be deduced by cross-tabulating parameters (age, sex, county, time of diagnosis, country of infection). Therefore, we divided the notifiable diseases into three groups:

1. Diagnoses with no notified cases since 1997 when the EpiArk was launched (e.g. relapsing fever, yellow fever and polio) – no statistics is presented for these diseases.

2. Diagnoses with, in average, less than ten cases per year since 1997 (e.g. hepatitis D, viral hemorrhagic fevers and rabies) – only the year and month tables are presented for these diagnoses.

3. Diagnoses with in average ten cases or more per year since 1997 (e.g. salmonellosis, typhoid fever and syphilis) – the diseases in this group were considered to have enough cases to exclude the possibility of cross matching. Therefore, all parameters are presented.

A complete list of all diagnoses showed and which category they fall into is presented in table 1.

Table 1: Table showing the diseases covered. Low implies no notified cases since 1997 (no statistics is presented), Medium implies ten notified cases or less per year in average (only tabled statistics presented) and High implies more than ten notified cases or less per year in average (full statistical coverage).

Disease	Level of available information
Amoebiasis	High
Anthrax	Low
Atypical mycobacterial infection	High
Botulism	Medium
Campylobacteriosis	High
Cholera	Medium
Diphtheria	Medium
Enterohemorrhagic E. coli O157 infection (EHEC)	High
Genital Chlamydia infection	High
Giardiasis	High
Gonorrhoea	High
HIV infection/AIDS	Low
HTLV I infection	Medium
HTLV II infection	Medium
Haemophilus influenzae type b infection (invasive disease)	High

Table 1: Table showing the diseases covered. Low implies no notified cases since 1997 (no statistics is presented), Medium implies ten notified cases or less per year in average (only tabled statistics presented) and High implies more than ten notified cases or less per year in average (full statistical coverage). (Continued)

Hemorrhagic fever with renal syndrome	High
Hepatitis A	High
Hepatitis B	High
Hepatitis C	High
Hepatitis D	Medium
Hepatitis E	Medium
Hepatitis non A-E	Medium
Legionellosis	High
Listeriosis	High
Malaria	High
Measles	High
Meningococcal infection	High
Methicillin-resistant <i>S. aureus</i> infection (MRSA)	High
Mumps	High
Ornithosis	High
Paratyphoid fever	High
Penicillin-resistant pneumococcal infection (PRP)	High
Pertussis	High
Plague	Low
Polio	Low
Rabies	Medium
Relapsing fever	Low
Rubella	Medium
Salmonellosis	High
Shigellosis	High
Syphilis	High
Tetanus	Medium
Toxoplasmosis	High
Trichinosis	Medium
Tuberculosis	High
Tularemia	High
Typhoid fever	High
Typhus fever	Low
Ulcer Molle	Medium
Vancomycin-resistant enterococcal infection (VRE)	High
Viral hemorrhagic fevers	Medium
Yellow fever	Low
Yersiniosis	High

Performance and quality control of the notification system

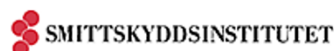
Before any notification is accepted into the EpiArk two epidemiologists (one at the county level, and one at the SMI) have scrutinized the content. If there are any doubts of the correctness of the data the notifying doctor is contacted.

The flow of information and timeliness has previously been studied in detail [16], and the median delay between diagnosis and registration of the report in EpiArk is 1–2 weeks. A new case will show on the web in the beginning of the month following the notification, but any corrections of the data from previous months will be shown after the weekly updates.

Currently a physician notifies diseases on a paper form. With a new communicable disease act in July 2004, and an upgraded version of the national electronic reporting system, notifications could from the second half-year of 2004 be notified electronically through the Internet. This will likely speed up the notification process considerably.

Scalability

The web statistics contain information from 1997 until today, with 1997 as a locked point in time. Each new year is accumulated to the statistics, hence providing a large amount of data to compare with, but also making the web statistic material larger and larger. At some point in time, the years presented will over swell the screen and the user will have trouble interpreting the information, because the information shown is too extensive. One way of



Statistics
 Notifiable Diseases
 Hepatitis B

Year and month statistics per county

Based on all cases, regardless of country of infection

↑ Hepatitis B

Options:

Change basis[i]

- All cases
 Domestic cases
 Cases infected abroad

Download statistic

- Current table (total number of cases)
 Current table (number of cases per 100.000 pop. and year).
 All tabled statistic for Hepatitis B

Click on a year for monthly statistics.

The numbers in the table below is presented as
 Total number of cases/Number of cases per 100.000 pop. and year.

County:[i]	1997	1998	1999	2000	2001	2002	2003	2004 (up to Jan)
Blekinge	19/ 12.5	15/ 9.9	21/ 13.9	35/ 23.2	37/ 24.6	29/ 19.3	39/ 25.9	2/ 1.3
Dalarna	29/ 10.1	23/ 8.1	17/ 6.0	23/ 8.2	14/ 5.0	39/ 14.0	26/ 9.3	0/ 0
Gotland	5/ 8.6	3/ 5.2	5/ 8.7	3/ 5.2	4/ 6.9	3/ 5.2	3/ 5.2	0/ 0
Gävleborg	33/ 11.5	31/ 10.9	29/ 10.3	37/ 13.2	30/ 10.7	63/ 22.6	69/ 24.8	8/ 2.8
Halland	39/ 14.3	37/ 13.5	46/ 16.8	43/ 15.6	46/ 16.6	33/ 11.8	42/ 15.1	2/ 0.7
Jämtland	6/ 4.5	5/ 3.7	10/ 7.6	3/ 2.3	11/ 8.5	31/ 24.1	13/ 10.1	2/ 1.5
Jönköping	41/ 12.4	42/ 12.8	36/ 11.0	35/ 10.6	50/ 15.2	36/ 10.9	30/ 9.1	4/ 1.2
Kalmar	18/ 7.4	37/ 15.5	27/ 11.4	52/ 22.0	49/ 20.8	51/ 21.7	35/ 14.9	14/ 5.9
Kronoberg	23/ 12.8	23/ 12.9	28/ 15.8	25/ 14.1	20/ 11.3	35/ 19.7	59/ 33.3	5/ 2.8
Norrbottn	22/ 8.3	13/ 4.9	11/ 4.2	29/ 11.3	32/ 12.5	42/ 16.5	65/ 25.5	6/ 2.3
Skåne	194/ 17.3	146/ 13.0	140/ 12.4	123/ 10.8	116/ 10.2	201/ 17.6	230/ 20.1	14/ 1.2
Stockholm	574/ 32.5	455/ 25.5	449/ 24.8	405/ 22.2	444/ 24.2	434/ 23.5	496/ 26.8	31/ 1.6
Södermanland	48/ 18.6	41/ 15.9	61/ 23.8	37/ 14.4	60/ 23.3	32/ 12.3	54/ 20.9	2/ 0.7
Uppsala	65/ 22.3	52/ 17.8	54/ 18.4	49/ 16.6	35/ 11.8	56/ 18.8	42/ 14.1	2/ 0.6
Värmland	12/ 4.2	24/ 8.6	33/ 11.9	43/ 15.6	49/ 17.8	59/ 21.5	75/ 27.3	0/ 0
Västerbotten	35/ 13.5	21/ 8.1	32/ 12.4	21/ 8.2	43/ 16.8	26/ 10.1	76/ 29.7	2/ 0.7
Västernorrland	48/ 18.8	32/ 12.7	38/ 15.2	44/ 17.8	43/ 17.4	49/ 20.0	57/ 23.2	2/ 0.8
Västmanland	37/ 14.3	49/ 19.0	33/ 12.8	30/ 11.6	29/ 11.2	43/ 16.6	40/ 15.4	4/ 1.5
Västra Götaland	322/ 21.6	197/ 13.2	202/ 13.5	608/ 40.6	328/ 21.8	295/ 19.6	272/ 18.0	18/ 1.1
Örebro	51/ 18.5	28/ 10.1	42/ 15.3	24/ 8.7	44/ 16.0	60/ 21.9	119/ 43.5	6/ 2.1
Östergötland	62/ 14.9	45/ 10.9	45/ 10.9	40/ 9.7	32/ 7.7	116/ 28.0	99/ 23.9	2/ 0.4
All of Sweden	1683/ 19.0	1319/ 14.8	1359/ 15.3	1709/ 19.2	1516/ 17.0	1733/ 19.4	1941/ 21.7	126/ 1.4

Figure 3
 Screenshot of the table presenting yearly and monthly statistics per county.

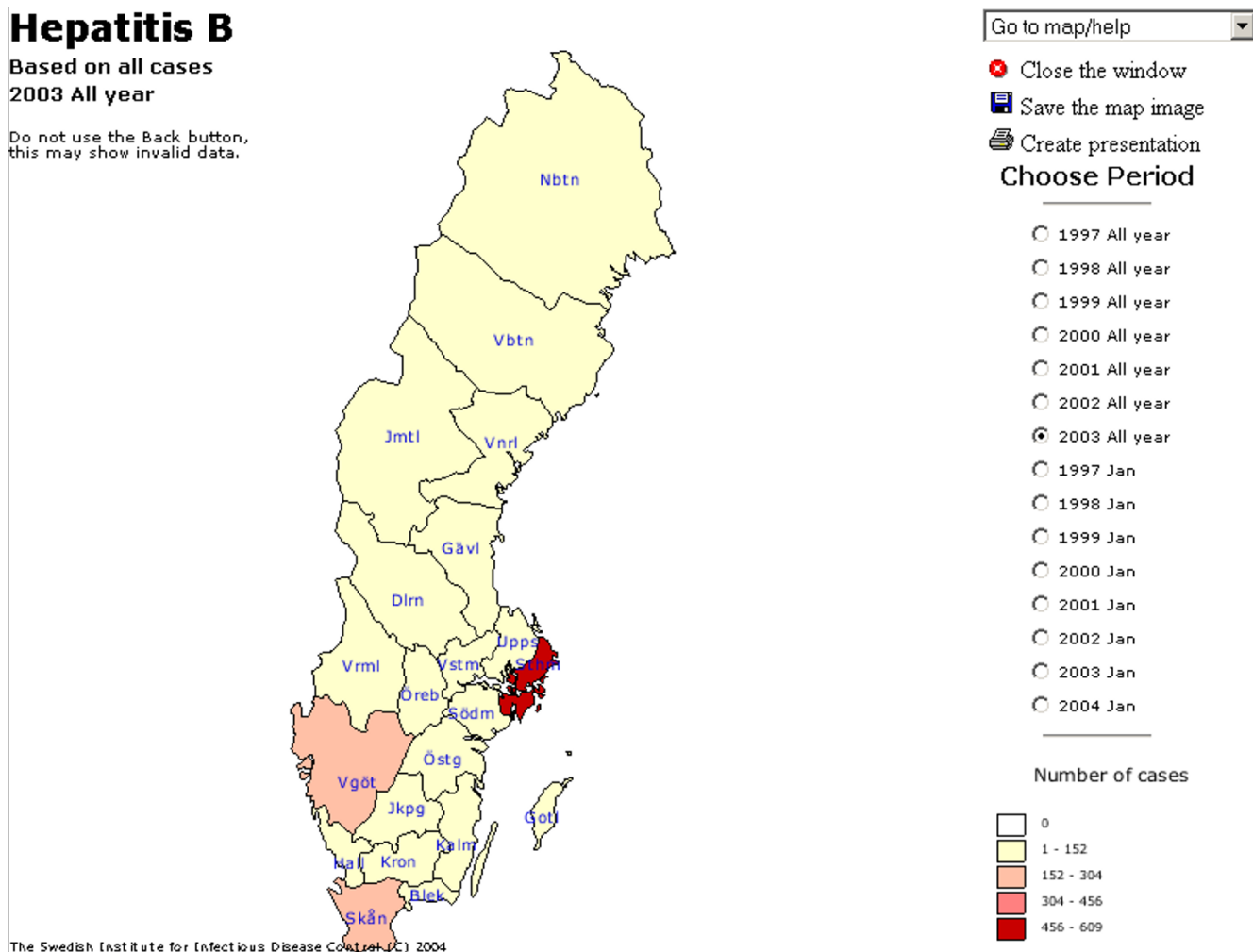


Figure 4
Screenshot of an interactive map, illustrating the geographical distribution in Sweden of where the cases were notified.

solving this problem is to have a start year constant, which may be changed, or choosing only to present information for i.e. the last four years.

Interpretation of the data



The web service contains a considerable amount of data, presented in various forms and allows the user a variety of interpretation. Users unfamiliar with the epidemiology of the single infectious diseases and the statistical methods used could have trouble interpreting the data. We have therefore included extensive help texts for the web service and also disease information and comments on the epidemiology of all diseases under surveillance (so far only in Swedish). Additional epidemiological information and trend analysis of each presented disease is offered, however not yet in English.

The application is based on all the notified cases with information derived from both the patient's physician and the laboratory, having diagnosed the causative agent. The notification should be submitted within 24 hours. However, in practice, this is not the case and the diseases may be reported days or weeks after the diagnosis. This may render the month of notification uncertain, but should not affect the total number of reported cases.

Since cases are reported both by the clinician and the laboratory, the overall sensitivity of the Swedish surveillance system is comparatively good. Using the capture-recapture technique [17] to estimate the proportion of cases being reported by either clinicians or laboratories, more than 98 percent of four different notifiable diseases (salmonello-

Statistics
Notifiable Diseases
Hepatitis B

Trend (moving average)

-  Hepatitis B
-  Save the graph

Hepatitis B - Trend

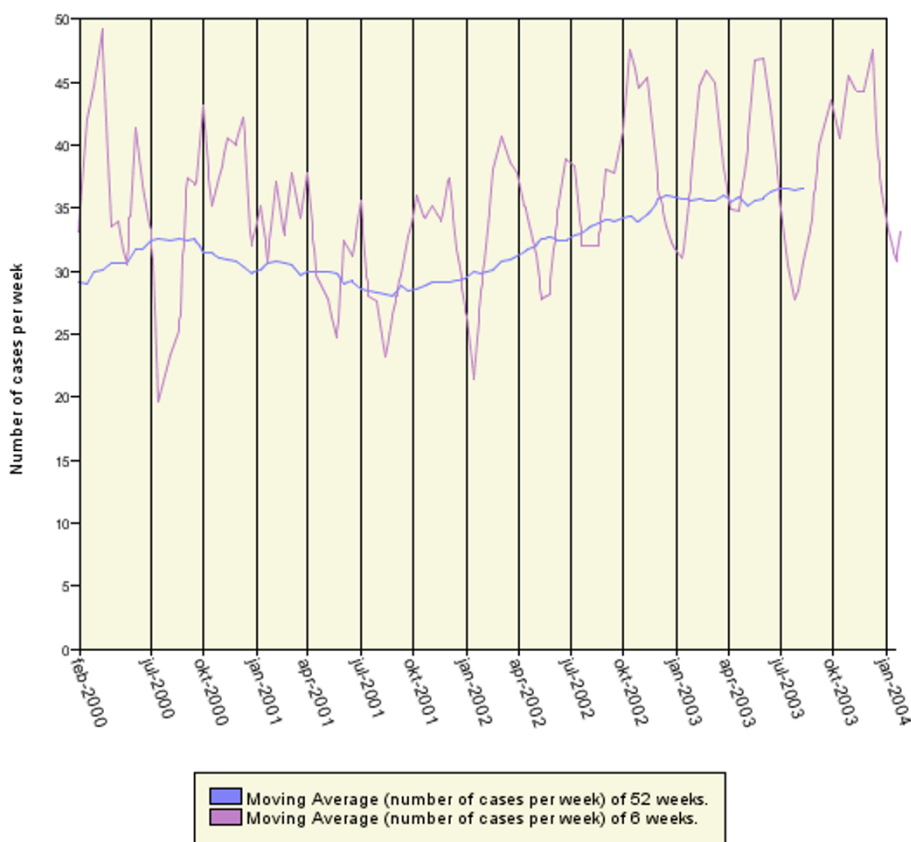


Figure 5
Screenshot of the graph illustrating trends. The red curve represents seasonal trends and the blue represents yearly trends.

sis, tularemia, penicillin-resistant pneumococci, and invasive meningococcal infection) were reported in 2002 [16].

In Sweden, as in other developed countries, the infectious disease panorama is affected by international travel. Surveillance statistics for travel-related diseases will give a

false picture of the epidemiology, unless data on domestic cases and cases with infection acquired abroad are presented separately. Therefore, all data are divided in cases infected in Sweden, cases infected abroad, and cases with no data on country of infection (mainly when clinical notification is lacking).

Without proper denominator data on travel, the data on country of infection will largely reflect the traveling pattern of the Swedes. However, since some years travel data from a commercial company are available, and could be used as a suitable denominator [18,19]. Presently, a project is initiated in which the surveillance data will annually be compared with the annually updated travel database, to have the system automatically generate risk estimates (per 1000 travelers) of being notified with each disease. This system will be built on the same GIS platform as the web service described in this paper.

Systems in other countries

A growing number of countries, e.g. the USA [20], Canada [21], England & Wales [22], Ireland [23], France [24,25] and Germany [26], put their national surveillance data in the public domain on the Internet. All update their data on a regular basis and presents the information in various ways; tables, graphs and maps. The difference between them is the availability. Some publish the information in a bulletin to download, while the rest publish the data on web pages. Canada provides an interactive map service for chronic diseases, which can be customized with several parameters (disease, age, notification year, sex and data type) [21]. To our knowledge, no other national surveillance systems publish as detailed and timely surveillance data as presently being published in Sweden.

The web system most similar to ours is the French Sentiweb [25,27]. This is based on the national general practitioner sentinel system and includes seven communicable diseases – influenza-like illness, diarrhea, measles, mumps, chickenpox, and male urethritis. The French system allows for syndromic reporting of influenza-like illness and diarrhea, which could not be reported in our web service, based on mandatory notifiable diseases. However, the Swedish system provides universally reported data for more than 50 diseases, giving far more diverse information. As a complement to our web service on mandatory notifiable diseases, we will soon put public on the Internet an electronic sentinel reporting and publication system for influenza and other syndromic sentinel reported diseases. This service will also include tables, maps and graphs, not dissimilar to the French Sentiweb.

Availability

The generated web service is available to the public, accessible at the SMI web site (<http://www.smittskyddsinsti.tutet.se> under Surveillance statistics). A technical report specifying the application, described in this paper, can also be downloaded from the SMI web site [28].

Future developments of the system

The application currently generates 34.000 files of HTML and java script, a huge amount that could be reduced, with a factor 500, if some servlet language like JSP, PHP or

ASP was used. This is a good way to optimize both the downtime of the web server and the flexibility of the web service further. This must be combined with a server side database though, and when using such, the security issues must be considered. If we were to rebuild the web service today, this is probably the way we would go.

The statutory notifiable diseases are not static, but may be changed. If a disease become statutory notifiable, the EpiArk nor the SMI often does not hold any information, before the date of the action. If this disease is to follow the "normal" flow of generation, the statistics will present all years before the action, as no cases were notified. One way to solve this problem is to introduce one start presenting statistics from year variable for each disease, making the statistic base only on real data.

In July 2004, there will be a new Communicable disease act in Sweden. Some of the present notifiable diseases will disappear and others be added to the surveillance system. In time for this a new web-based version of SmiNet will be implemented, greatly enhancing the speed and accuracy of the surveillance system. The presently described web service will need some minor adjustments to fit the SmiNet version 2.

Conclusion

The web service was developed to provide updated data on notifiable communicable diseases for a wide range of users. In order to meet visitor demands, the system was designed to provide an easy and weekly updated access to the data, without compromising the confidentiality aspects. The web service has provided the health community, the media, and the public with easy access to both timely and detailed surveillance data presented in various forms.

Competing interests

None declared.

Authors' contributions

Per Rolfhamre, as a computer scientist has designed, implemented and analyzed the web service. Katarzyna Grabowska has assisted with statistical issues and Karl Ekdahl has managed the project and assisted in issues related to surveillance and communicable diseases.

Additional material

Additional File 1

Moving average algorithm. The algorithm used to calculate the moving average.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2334-4-17-S1.txt>]

References

1. Teutsch SM, Churchill RE, eds: **Principles and practice of public health surveillance**. 2nd edition. Oxford: Oxford University Press; 2000.
2. **Inventory of resources for infectious diseases in Europe** [<http://iride.cineca.org/>]. retrieved February 19, 2004
3. Friedman DJ, Anderka M, Krieger J, Land G, Solet D: **Accessing population health information through interactive systems: lessons learned and future directions**. *Public Health Rep* 2001, **116**:132-47.
4. Laporte RE, Barinas E, Barinas E, Chang Y, Libman I: **Global epidemiology and public health in the 21st century. Applications of new technology**. *Ann Epidemiol* 1996, **6**:162-7.
5. Rushton G: **Public Health, GIS, and Spatial Analytic Tools**. *Annu Rev Public Health* 2002, **24**:43-56.
6. Croner CM: **Public Health, GIS, and the Internet**. *Annu Rev Public Health* 2003, **24**:57-82.
7. Kistemann T, Dangendorf F, Schweikart J: **New perspectives on the use of geographical information systems (GIS) in environmental health sciences**. *Int J Hyg Environ Health* 2002, **205**:169-81.
8. **The Big Faceless Java Graph library, big.faceless.org** [<http://big.faceless.org/products/graph/>]. retrieved February 19, 2004
9. **JavaMail API, Sun Microsystems, Inc** [<http://java.sun.com/products/javamail/>]. retrieved February 19, 2004
10. **JavaBeans, Sun Microsystems, Inc** [<http://java.sun.com/products/javabeans/>]. retrieved February 19, 2004
11. **Apache Ant, The Apache Software Foundation** [<http://ant.apache.org/>]. retrieved February 19, 2004
12. **i-net MERLIA, i-net software** [<http://www.inetsoftware.de/English/Produkte/MERLIA/default.htm>]. retrieved February 19, 2004
13. **ArcIMS, ESRI** [<http://www.esri.com/software/arcims/index.html>]. retrieved February 19, 2004
14. **Java 2 Platform, Sun Microsystems, Inc** [<http://java.sun.com/j2se/>]. retrieved February 19, 2004
15. Ekdahl K, ed: **Communicable diseases in Sweden 2001. The annual report of the Department of epidemiology**. Solna: Swedish Institute for Infectious Disease Control (SMI) 2002 [<http://www.smittskyddsinstitutet.se/upload/Publikationer/Report2001.pdf>]. retrieved February 19, 2004
16. Jansson A: **Sensitivity and timeliness of case reporting in the Swedish statutory surveillance of communicable diseases 1998–2002**. Master thesis in Public health 2004 [<http://www.smittskyddsinstitutet.se/upload/Publikationer/MSc-surveillance.pdf>]. Karolinska Institute
17. Reintjes R, Termorshuizen F, Van De Laar MJW: **Assessing the Sensitivity of STD Surveillance in the Netherlands: an application of the capture-recapture method**. *Epidemiol Infect* 1999, **122**:97-102.
18. Lindbäck H, Lindbäck J, Tegnell A, Janzon R, Vene S, Ekdahl K: **Dengue fever in travelers to the Tropics, 1998 and 1999**. *Emerg Infect Dis* 2003, **9**:438-42.
19. **Swedish Travel and Tourist Data Base (TDB). Resurs AB, Sweden** [<http://www.resursab.se/>]. retrieved February 19, 2004
20. **NCID Office of surveillance, Centers for Disease Control and Prevention, USA** [http://www.cdc.gov/ncidod/osr/site/surv_resources/data_reports.htm]. retrieved February 19, 2004
21. **Disease surveillance on-line, Health Canada** [http://dsol-smed.hc-sc.gc.ca/dsol-smed/mcd-smcm/n_ind_e.html#maps]. retrieved February 19, 2004
22. **CDR Weekly, The Health Protection Agency, England** [<http://www.hpa.org.uk/cdr/default.htm>]. retrieved February 19, 2004
23. **The National Disease Surveillance Centre (NDSC), Ireland** [<http://www.ndsc.ie/IDStatistics/>]. retrieved February 19, 2004
24. **l'Institut de Veille Sanitaire, France** [<http://www.invs.sante.fr/surveillance/index.htm>]. retrieved February 19, 2004
25. **Le réseau Sentinelles, France** [<http://rhone.b3e.jussieu.fr/senti/en/php/navigation/accueil/>]. retrieved February 19, 2004
26. **Robert Koch Institut, Germany** [<http://www.rki.de/INFEKT/EPIBULL/EPI-HTML>]. retrieved February 19, 2004
27. Boussard E, Flahault A, Vibert JF, Valleron A-J: **Sentiweb: French communicable disease surveillance on the world wide web**. *BMJ* 1996, **313**:1381-2.
28. Rolfhamre P: **MAPAPP version 1.3.4: Documentation of application and generated service. SMI-report 1–2003**. Solna: Swedish Institute for Infectious Disease Control 2003 [[\[www.smittskyddsinstitutet.se/upload/Publikationer/SMI-rapport_1-2003.pdf\]\(http://www.smittskyddsinstitutet.se/upload/Publikationer/SMI-rapport_1-2003.pdf\)\]. retrieved: February 19, 2004](http://

</div>
<div data-bbox=)

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2334/4/17/prepub>

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

