

Research article

Open Access

EURISWEB – Web-based epidemiological surveillance of antibiotic-resistant pneumococci in Day Care Centers

Sara Silva¹, Rodrigo Gouveia-Oliveira¹, António Marezek¹, João Carriço¹, Thorolfur Gudnason², Karl G Kristinsson², Karl Ekdahl³, António Brito-Avô⁴, Alexander Tomasz⁵, Ilda Santos Sanches^{1,6}, Hermínia de Lencastre^{1,5} and Jonas Almeida*^{1,7}

Address: ¹Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Av. República (EAN), PO Box 127, 2781-901 Oeiras, Portugal, ²Department of Pediatrics and Microbiology, Landspítali University Hospital, Reykjavik, Iceland, ³Swedish Institute for Infectious Diseases Control, Department of Epidemiology, Se-171 82 Solna, Sweden, ⁴Centro de Saúde de Oeiras, Av. Salvador Allende, 2780-163 Oeiras, Portugal, ⁵Laboratory of Microbiology, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA, ⁶Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Monte de Caparica, 2829-516 Caparica, Portugal and ⁷Dept Biometry & Epidemiology, Medical Univ South Carolina, 135 Cannon Street, Suite 303, PO Box 250835, Charleston SC 29425, USA

Email: Sara Silva - sara@itqb.unl.pt; Rodrigo Gouveia-Oliveira - rodrigo@itqb.unl.pt; António Marezek - antonio.marezek@microcortex.com; João Carriço - jcarriço@itqb.unl.pt; Thorolfur Gudnason - thorgud@landspitali.is; Karl G Kristinsson - karl@landspitali.is; Karl Ekdahl - karl.ekdahl@mtc.ki.se; António Brito-Avô - abritoavo@netcabo.pt; Alexander Tomasz - tomasz@mail.rockefeller.edu; Ilda Santos Sanches - isanches@itqb.unl.pt; Hermínia de Lencastre - lencash@mail.rockefeller.edu; Jonas Almeida* - almeidaj@musc.edu

* Corresponding author

Published: 08 July 2003

Received: 28 February 2003

BMC Medical Informatics and Decision Making 2003, 3:9

Accepted: 08 July 2003

This article is available from: <http://www.biomedcentral.com/1472-6947/3/9>

© 2003 Silva 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: EURIS (European Resistance Intervention Study) was launched as a multinational study in September of 2000 to identify the multitude of complex risk factors that contribute to the high carriage rate of drug resistant *Streptococcus pneumoniae* strains in children attending Day Care Centers in several European countries. Access to the very large number of data required the development of a web-based infrastructure – EURISWEB – that includes a relational online database, coupled with a query system for data retrieval, and allows integrative storage of demographic, clinical and molecular biology data generated in EURIS.

Methods: All components of the system were developed using open source programming tools: data storage management was supported by PostgreSQL, and the hypertext preprocessor to generate the web pages was implemented using PHP. The query system is based on a software agent running in the background specifically developed for EURIS.

Results: The website currently contains data related to 13,500 nasopharyngeal samples and over one million measures taken from 5,250 individual children, as well as over one thousand pre-made and user-made queries aggregated into several reports, approximately. It is presently in use by participating researchers from three countries (Iceland, Portugal and Sweden).

Conclusion: An operational model centered on a PHP engine builds the interface between the user and the database automatically, allowing an easy maintenance of the system. The query system is also sufficiently adaptable to allow the integration of several advanced data analysis procedures

far more demanding than simple queries, eventually including artificial intelligence predictive models.

Background

Social forces that produced Day Care Centers (DCCs) for preschool age children in many developed countries have – ironically – also created in these structures one of the major ecological reservoirs of drug resistant strains of *Streptococcus pneumoniae*, which spread globally and began to create serious complications in the chemotherapy of diseases caused by this dangerous pathogen [1–3]. Day Care Centers recruit in close physical proximity children of an age group that is characterized by high rate of carriage of *S. pneumoniae*, an immature immune system and frequent viral and bacterial respiratory tract infections leading to extensive use of antimicrobial agents which provide a powerful selective milieu for the emergence of resistant strains [4–7]. The best evidence that such strains can cause both pediatric and adult disease came from molecular epidemiological studies, which demonstrated that resistant clones of *S. pneumoniae* most frequently identified in disease [8,9] were also the ones frequently carried in the nasopharynx of healthy children in DCCs [10–12].

If DCCs are ecological reservoirs of resistant *S. pneumoniae* then reduction in the rate of carriage of such strains in DCCs should also impact on the frequency of infections caused by resistant pneumococci. Testing the efficacy of such a novel strategy was the purpose of the multinational initiative EURIS (European Resistance Intervention Study – Reducing Resistance in Respiratory Tract Pathogens in Children) [13] launched by the European Community in September of 2000 until 2003. Investigators from four countries (France, Iceland, Portugal and Sweden) supported by scientists from Germany and the USA joined forces to test the effect of a variety of different interventions methods (e.g. reduction in drug prescriptions; changing antibiotic dosing; improving hygienic conditions in DCCs etc.) on the frequency of nasopharyngeal colonization by resistant pneumococci – in carefully controlled studies.

The structure of EURIS is composed of four centers where strain collections and interventions are carried out: Portugal – Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa; Iceland–Landspítali University Hospital; Sweden – Swedish Institute for Infectious Diseases Control; France – Institut National de la Santé et de la Recherche Médicale. However EURISWEB represents data generated in only three of the four collection centers: Portugal, Iceland and Sweden, the three countries in which the timing, the age groups of the children and the

methods used were fully harmonized. The French initiative, while addressing the same issues, was not directly comparable, as it involved different age groups, different mode of sampling, and schools rather than Day Care Centers. Therefore the French data were deposited in a different database. Four additional collaborating units assist as reference centers for the harmonization of methods in clinical microbiology (Iceland: Landspítali University Hospital); molecular epidemiology of antibiotic resistant genes and clones (USA: Laboratory of Microbiology, The Rockefeller University; Germany: University of Kaiserslautern); data management and mathematical modeling of epidemiological aspects of EURIS (Portugal: Instituto de Biologia Experimental e Tecnológica).

The risk factors, *i.e.* the nature and number of the factors that influence the rate of carriage of drug resistant *S. pneumoniae* in preschool age children and their quantitative contribution to the degree of colonization, are not well understood. Furthermore, major risk factors for nasopharyngeal colonization may differ significantly from one setting to another [14–16], which makes analysis of data generated by a multinational study like EURIS, more complex. The evaluation and comparison of such massive amounts of surveillance data necessitated the construction of a computerized infrastructure organized in such a manner that it would assure not only data storage and retrieval but also an eventual bioinformatics analysis. The purpose of this communication is to describe such a web-based infrastructure specially designed to fit the purposes of EURIS – the EURISWEB.

Several potential conflicting attributes had to be accommodated in the design of such an infrastructure. On the one hand, it was to provide full integration of data from different countries in a common normalized repository, fully accessible to all EURIS participants. On the other hand, it was also supposed to exhibit the properties of a local database with full separation between the countries involved. Finally, it was anticipated that, eventually, EURISWEB would be made available for wider usage for research and public health management at a later stage, with steep requirements of stability, scalability, security, user-friendly access, low cost portability, and transparent implementation for subsequent independent development. In the design of EURISWEB we took into account the multiple goals of such a web-based infrastructure which now includes a relational online database coupled with data retrieval and analysis tools, where registered users can access data and tools by using a personal login

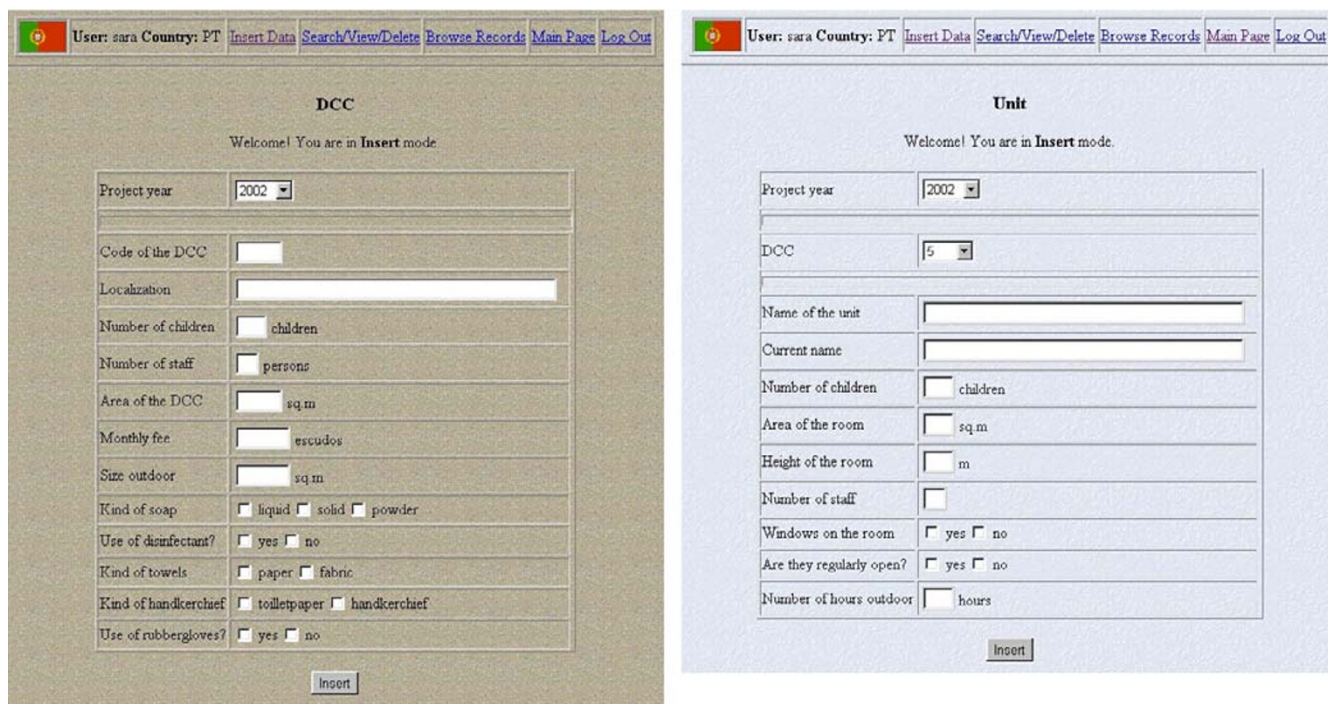


Figure 1
Data entry forms Data entry forms for the DCC and room (unit within the DCC) questionnaires. Upon pressing the Insert button, the validation procedure checks the data and either inserts it in the database or informs the user about errors in the data.

name and password through a standard web browser. Ultimately three of the four centers (Iceland, Portugal and Sweden), in which the nature of the pediatric population and mode of sampling were most comparable, chose to combine all data for deposition in EURISWEB, which now covers a large number of participant institutions: 16 DCCs in Portugal, 30 DCCs in Iceland and 25 DCCs in Sweden; a wide variety of sources of data, including demographic, socio-economic factors, clinical data, patterns and types of drug use and drug prescription; microbiological data on the antibiograms and serotypes as well as molecular types of the pneumococcal isolates and DNA fingerprints of resistant genes.

Availability

A demo version of EURISWEB is available to the general public [17], accessible with username euris and password welcome. For those who intend to receive the e-mails sent by the query agent (see User-Friendly Query System), please request a personal account to the authors. As any modifications applied to the current implementation of the infrastructure will automatically be reflected on the demo version, new features may be already apparent

when compared with illustrations and examples used in this manuscript.

Methods

Data and data acquisition

The diverse surveillance data (demographic, clinical, microbiological etc.) generated in project EURIS are used to fill five different types of Questionnaires which serve as the source of information to be introduced into the EURISWEB database. The relationship between the five questionnaires is described and illustrated later in this report (see Database structure and Database tables versus online forms). Typically each site will update surveillance information at least once per year. Questionnaires 1 and 2 are provided by the staff of each participating DCC.

- Questionnaire 1 contains information regarding physical features of the DCC (address, number of rooms and windows, area inside and outside the facility, number of children and staff, hygiene protocols and practice) – see Figure 1.

- **Questionnaires 2** provide the same type of information for each room (also referred to as "unit") in the particular DCC – see Figure 1.
- **Questionnaires 3** are filled at least once every year by the parents of the children. They contain demographic information on the household and environment where the child lives, including number and age of siblings, shared bedrooms, and specific conditions such as smoking in the house.
- **Questionnaires 4** are filled by the parents just prior to each strain collection. They provide information on antibiotic consumption prior to sampling (type of antibiotic, taken when and for how long). Also provided are data on illness and hospitalizations of the child.
- **Questionnaires 5** are filled by the participating microbiology and molecular biology laboratories. They contain characterization of each *S. pneumoniae* isolate for serotype; antibiotic type (susceptibility to oxacillin, chloramphenicol, erythromycin, clindamycin, tetracycline, sulfamethoxazole-trimethoprim, and levofloxacin); MIC values for penicillin and ceftriaxone; molecular type by PFGE (Pulsed-Field Gel Electrophoresis) and MLST (Multilocus Sequence Typing) (for selected isolates); DNA probes for antibiotic resistant factors; and RFLP (Restriction Fragments Length Polymorphisms) for *pbp* (penicillin binding protein) genes of selected penicillin resistant isolates. All data in questionnaire 5 are obtained by common harmonized methods.

Database conception

Although there was an effort towards the normalization of data acquisition taking place in the different participant countries, the questionnaires delivered to the DCCs and to the parents contain various questions that reflect realities specific to the country involved. Accordingly, some questions only appear in the questionnaires of some countries. Also, the frequency with which updated information is collected differs between countries. Since discarding data was to be avoided at all costs in order not to confront local practices, the normalization process had to be extended to database conception itself. Instead of designing an optimal database structure for each country, an iterative consulting process was followed for nearly a year to produce a normalized database structure that fits the reality presented by all the countries involved. The final structure of the EURISWEB database accommodates both country specificity and common European health management practices.

Data retrieval

Besides providing comprehensive data storage, the web-based data management infrastructure must also allow

the easy querying and retrieval of the data it contains. Since some of the retrieval requests may generate large amounts of data, or may require intensive computation, the requests are processed as background processes managed by a software agent that send e-mails to the user with information on the execution state of each request and, finally, a link to the completed report. User-friendliness was the primary concern in building the interface available to make these requests, with current version reflecting extensive user feedback.

Software and Hardware

All the software used to implement this infrastructure is Open Source and is provided under public license. The scripts that generate the HTML (HyperText Markup Language) interfaces were written in PHP (PHP: Hypertext Preprocessor) [18] 4.x. The Database Management System is PostgreSQL [19] 7.1.x. The scripts for the agents that handle the data retrieval requests were written in shell (bash 2.04) and Perl [20] (5.x). The server is a PC, CPU 2x PIII (coppermine) @ 800 MHz, with 512 MB SDRAM, running Linux [21] (based on Slackware [22] 7.1, kernel 2.2.x), SSL (Secure Socket Layer) [23] enabled Apache Web-Server [24] (Apache/1.3.x Ben-SSL/1.x). Software versions described above were updated regularly throughout the course of the EURIS project (2000–2003) with no negative impact on its performance.

Results

Database structure

The basic internal structure of the database consists of 9 tables with an average of 14 fields per table. Figure 2 shows a simple model of this structure, where boxes represent tables and lines represent the relations between them. There is only one type of relation in this structure, which is "one-to-many", the "one" side being represented by the single line and the "many" side being represented by the forked line. A one-to-many relation between two tables means that one record from one table can be associated to several records from the other table. For example, one DCC can be associated to several rooms (units) in the same DCC; one unit can be associated with several children; and one child can be associated with several siblings.

The description illustrated in Figure 2 is country specific. A separate set of tables was defined for each of the three participant countries – Portugal, Iceland and Sweden, all inside the same database, but not formally connected to each other. Although the questionnaires for the different countries have significant differences, as some countries may lack many fields or even whole tables of this structure, the critical feature is that all the common fields can be found in exactly the same location in each country-specific structure. Equally critical, the key fields are

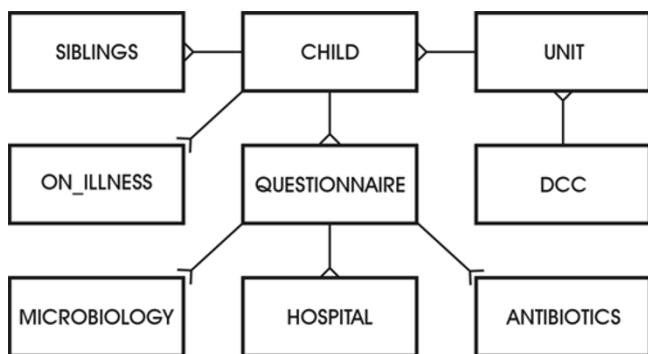


Figure 2
Database structure Boxes represent tables and lines represent the relations between them. There is a similar set of tables for each country. There is only one type of relation in this structure, which is "one-to-many". The single line represents the "one" side and the split line represents the "many" side. A one-to-many relation between two tables signifies that one record from one table can be associated to several records from the other table.

obligatorily shared by all countries, a feature that can only be easily achieved if a common host infrastructure is in place, which is the case in EURISWEB. Because the frequency with which updated information is collected differs between countries (see Data and data acquisition, and Database conception), many of the key fields are related to the specification of the sampling periods, playing an important role as temporal normalization features. Once the access restrictions are lowered, the conservation of ontology and structure enables intersection between country-specific structures to produce comprehensive data sets jointly describing epidemiological data, which are valid for all the participating countries. Furthermore, because all countries also share the same data retrieval system (see User-Friendly Query System), queries already built by different countries produce compatible results that can be promptly joined after removing the country-specific fields.

Online interface

The interface between the database and the users is made of standard HTML pages (no external applications, "plugins", needed on the client side). Data entering is performed through five online forms that mimic the original paper questionnaires, to facilitate the insertion task (see examples of two forms in Figure 1). All data entered in the forms is submitted to online validation procedures before entering the database, thus avoiding some of the most common user errors that may cause integrity or consistency violations in the database. Upon pressing the Insert button for submitting data, the user is promptly informed

of all its mistakes and given a chance of resolving them on the same page (example in Figure 3). Only after passing all the checks is the data effectively inserted in the database, and fitted into the respective internal data structure (see Database structure). Searching and visualizing data can be done on a record-by-record basis, using the same five forms format, or by browsing as a table that shows several records at the same time (example in Figure 4). Some simple statistics can also be requested online. For convenience, most of the tables presented can be directly viewed or saved in Excel format.

Data retrieval requests can also be made by filling a simple online form in which the amount of typing required is kept to a minimum (see User-Friendly Query System). The results can be viewed and downloaded in delimited text format, also readily importable into Excel.

Database tables versus online forms

The relationship between the internal database structure and the set of online forms available to the user is not a one-to-one association. Behind each form there can be more than one table, as shown in Figure 5. Although the mimicking of the original questionnaires by the online forms is meant to facilitate the user's adaptation to the data insertion and visualization, that is not the optimal data organization in a relational database. For example, the repeated set of questions about each antibiotic taken prior to sampling (see Data and data acquisition) should not result in a repeated set of fields in the same database table (table QUESTIONNAIRE, see Figure 2). Instead, each set of questions constitutes a row of fields in a different table (table ANTIBIOTICS, same figure).

Operational model

The operational model of the database interface is depicted in Figure 6, where the arrows represent flow of information between the various entities. The five online forms for record insertion and visualization, available to the user, are all built with the same general procedure (PHP engine). This program, written in PHP, reads files that contain all the information regarding the forms layout (layout files), designs the forms and manages all the interactions between the users and the database. Each layout file describes a form (for all countries) and consists of a few lines written in a subset of the PHP language, which indicate each field's properties, such as whether it is a numeric or Boolean field, a date or time field, and what are the range and type of values allowed. This program and the subset of PHP used to define the layout files are the core of the surveillance system reported here. Accordingly, to alter an existing form, or generate a new one, all the database manager has to do is update or build a layout file.


 User: sara Country: PT Insert Data Search/View/Delete Browse Records Main Page Log Out	
<h3>Child</h3> <p style="color: red; font-weight: bold;">Errors found!</p> <p>Please try again. You are in Insert mode.</p>	
Project year	2002
DCC and unit <i>Record unavailable for this year/semester!</i>	12a - Sala Alice
Code of the child	PT C 10
Date of birth <i>Invalid date!</i>	1995 02 29 yyyy mm dd
Unreturned ID form?	<input type="checkbox"/> (check if yes)
Gender <i>Only one option allowed!</i>	<input checked="" type="checkbox"/> male <input checked="" type="checkbox"/> female
Weight	<input type="checkbox"/> Kg

Figure 3
Validation checks Example of data entry form for the child records with error warnings to the user. The user must correct all the errors before being able to insert the record.

The layout files also include the description of the connection between the form fields and the actual database fields. This information must be in accordance to the internal database structure, which is managed by SQL (Structured Query Language) code also stored in files (structure files). Therefore, the database manager will need to keep them consistent with any changes in the structure files required by modifications in the online interface. These two simple tasks ensure both the automatic construction of personalized forms – together with online validation check procedures – and a smooth link between them and the database internal structure.

User-Friendly Query System

Although SQL is the standard way to access data stored in a database, using it requires some prior knowledge and experience from the user. The User-Friendly Query System, available to all the EURISWEB users, is an interface that facilitates query construction in order to make the

wide range of possibilities offered by SQL amenable to the untrained user. The users are presented with a series of selection boxes where they can select the fields they want to see, the restrictions they want to apply to the records returned, and how the returned records are to be grouped (Figure 7). The chosen options are then transformed into actual SQL formatted statements that are sent to the query management agent, through the PHP engine, as shown in Figure 8. The arrows in the figure represent flow of information between the various entities (see Figure 9 for the whole operational model).

The query agent manages all the requests and runs them exclusively in background, so that high usage rates and complex requests do not interfere with the normal usage of the database interface. The agent interacts with the database and informs the users, by e-mail, of when their requests start being processed and when they finish, including the information of whether the query was

User: sara Country: PT
[Insert Data](#)
[Search/View/Delete](#)
[Browse Records](#)
[Main Page](#)
[Log Out](#)

Browse Records

[Help](#) (new browser window)

Table name: Microbiology

Visible fields:	Shown records:	Ordering keys:
<div style="border: 1px solid gray; padding: 2px;"> sample_number isolate_code arrival_timestamp nurse dry_mucus </div>	<div style="border: 1px solid gray; padding: 2px; text-align: center;"> <input type="button" value="Previous 50 records"/> </div> <div style="border: 1px solid gray; padding: 2px; text-align: center;"> 110 to 113 </div> <div style="border: 1px solid gray; padding: 2px; text-align: center;"> <input type="button" value="Next 50 records"/> </div>	<div style="border: 1px solid gray; padding: 2px;"> First: sample_number </div> <div style="border: 1px solid gray; padding: 2px;"> Second: isolate_code </div>
<input type="button" value="Submit"/>		

Action	sample_number	isolate_code	result	benzylpenicillin	Type
<input type="button" value="View"/> <input type="button" value="Edit"/> <input type="button" value="Delete"/>	104	a	+	0.047 (S)	PSPN
<input type="button" value="View"/> <input type="button" value="Edit"/> <input type="button" value="Delete"/>	105	a	+	0.75 (I)	PRPN low level
<input type="button" value="View"/> <input type="button" value="Edit"/> <input type="button" value="Delete"/>	106	a	(+)	0.016 (S)	PSPN
<input type="button" value="View"/> <input type="button" value="Edit"/> <input type="button" value="Delete"/>	107	a	+	0.094 (I)	PRPN low level

Figure 4
Browsing records Browsing page for microbiology records. Here the user can browse through several records of the same table. Selection of visible fields and ordering by one or two fields is possible. Direct access to individual records is provided by the View, Edit, and Delete buttons.

successfully answered (the interface gives users enough freedom to request impossible things) or not, in which case the results presented are an empty text page. Due to security reasons, the results of queries are never sent by e-mail – they can only be downloaded from the server via an SSL connection.

Users can rerun, edit, or delete saved queries. They can also group queries into reports, so that a single request

will yield all the results from the several queries of that report. Furthermore, users can save restrictions used often, and apply them to other queries. To minimize the time and effort required of the users, we have provided several pre-made queries, already aggregated into several logical reports. This feature may prove particularly useful if standard reporting formats become a regulatory requirement.

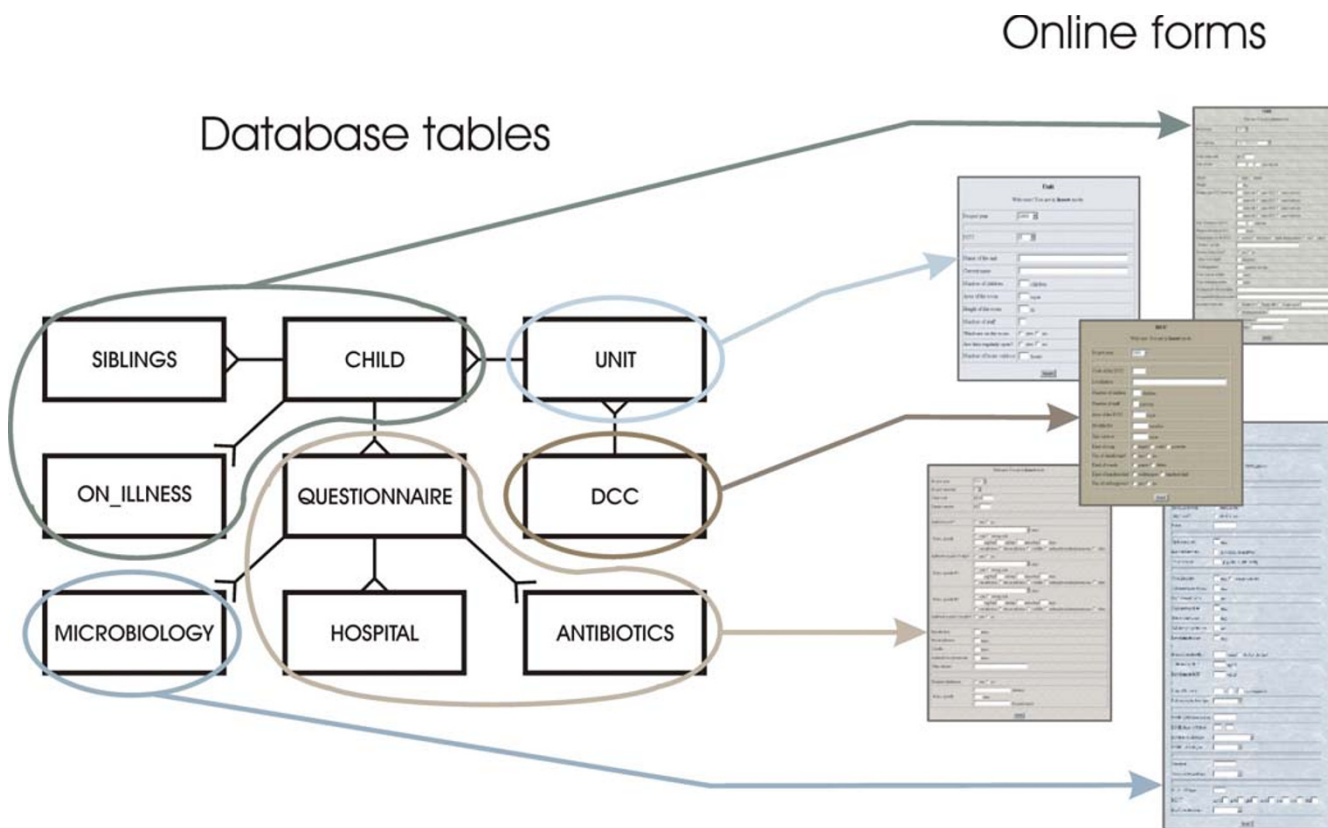


Figure 5
Tables versus forms Relationship between database tables and online forms. The optimal data organization in a relational database may not be agreeable to the user. There may be several database tables behind each online form.

Usage

The EURIS online database has been adopted as the data storage standard by three of the EURIS participant countries – Portugal, Iceland, and Sweden. Growing steadily since its birth, February 2001, it now has 24 registered users and contains a total of 213 DCC records, 720 unit records, 10991 children records, 13207 questionnaire records, and 13504 microbiology records, totaling more than 25 megabytes of data. The User-Friendly Query System, available since April 2002, now contains 400 pre-made and 786 user-made queries, aggregated into several reports.

Discussion and Conclusions

Privacy and security precautions

In EURISWEB, each user registry includes not only its login name and password, but also its country identification, which completely blocks access to data belonging to other countries. In fact, by using different table sets for each country, the central database can behave like several different local databases, and the user

is never aware that its access is restricted to only a subset of the complete system. In a near future, the fact that all countries use, after all, the same normalized structure, will allow simple queries and complex data analysis to be performed in the common data, as if we were dealing with a single country.

The user registration process also includes a level access number that defines restrictions for each type of user. Although the system was built anticipating this need, other precautions proved sufficient to monitor and to recover from possible destructive actions. All user inputs are scanned for invalid characters to prevent SQL code injection, and a record of all actions performed in the database is kept, including who did what, and when. Any accidentally deleted record can be promptly restored by the database manager; all the updates a record has undergone since its insertion can be tracked; and many database usage statistics can be easily performed.

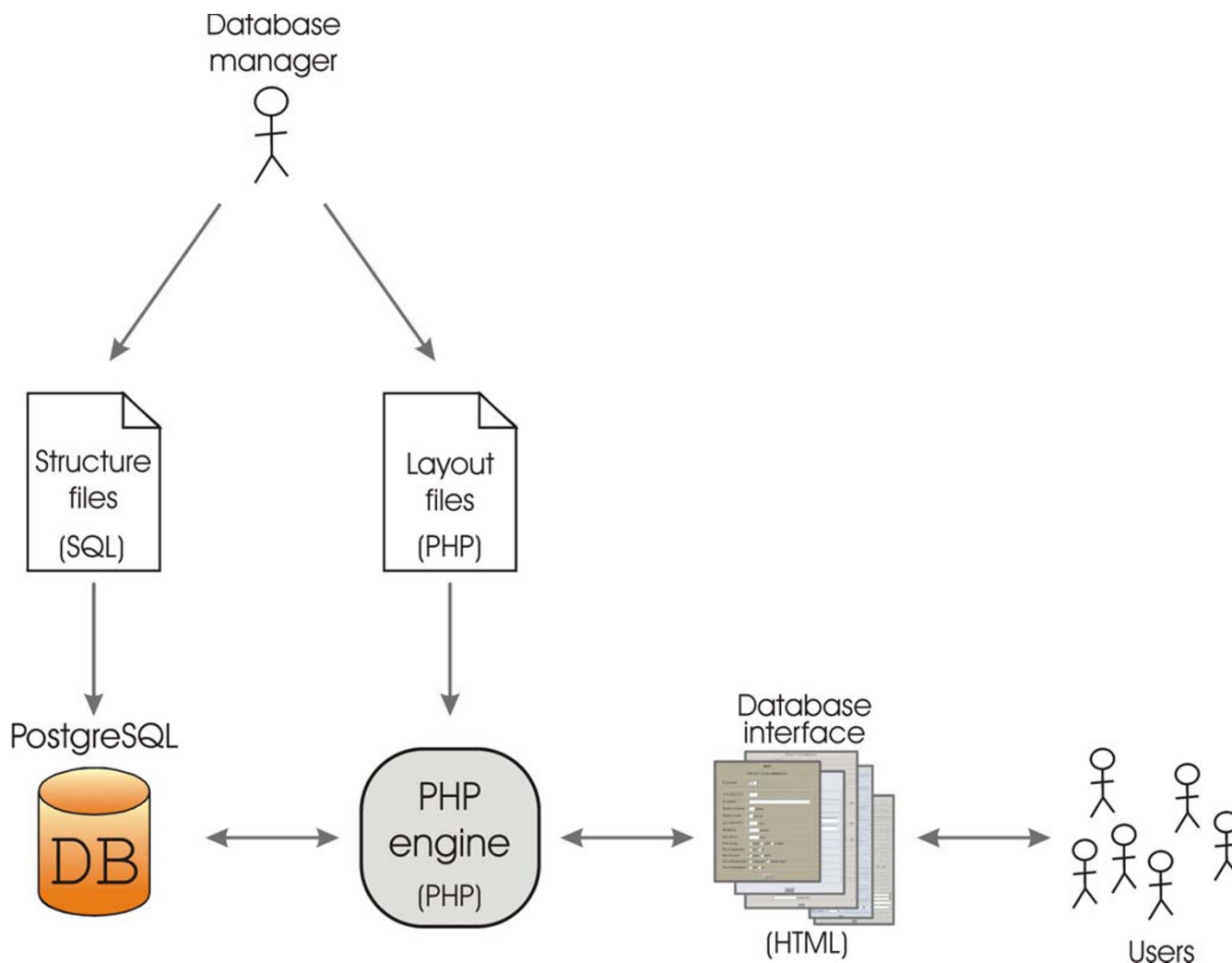


Figure 6
Operational model – database Operational model of the EURISWEB database. The arrows represent flow of information between the various entities. The database manager builds the layout files, used by the PHP engine to build the online forms, and the structure files, used to define the structure of tables and fields in the database. Layout and structure must be in accordance with each other. The PHP engine manages all the interactions between the database and the users.

Intrusion by unauthorized parties (hackers) is repelled by the need to log on with login name and password, and subsequent identification of the user with cookies protected by SSL, without which no page is ever shown and no query is run. A brute force attack is also limited by a delay introduced in the password checking cycle, and resources consumption at the server. Additionally, page accesses are monitored on a daily basis. Repeated login attempts would therefore be promptly detected before a sufficiently high number of probes take place. Furthermore, a firewall protects the server from being accessed on other ports apart from the HTTPS port, and the server soft-

ware (Apache, PHP, kernel, etc.) is promptly updated if any security breach is detected in the current versions.

In all cases, names of children and DCCs are not kept in the database, instead being replaced by codes and acronyms manually assigned prior to insertion.

Scalability

The operational model described in the Results section is the basis for easy improvements and extensions to the whole infrastructure of EURISWEB. As a consequence of the design described in that section, database management can be fully dealt with by manipulation of the

User: sara
Country: PT
[Instructions Page](#)
[Create New Query](#)
[Create Common Condition Fields](#)
[Queries List](#)
[Reports List](#)
[Main Page](#)
[Log Out](#)

User Friendly Query System

Create New Query

This page allows you to ask questions to the database in a more userfriendly environment than Structured Query Language. Your request is translated into SQL, submitted to the database and you are given the reply. Even though it is easy, one should read the [Instructions Page](#) before starting.

Specify - The contents of the fields selected here will be listed on the query result. The Distinct option is used to avoid repeated records in the result.

Specify multiple selection box:

- (none)
- Child-year
- Child-dcc_code
- Child-unit_name
- Child-child_code
- Child-birthdate_timestamp
- Child-gender_male
- Child-date_entrance_timestamp
- Child-daily_hours_dcc
- Child-weight
- Child-unreturned
- Child-smokers
- Child-total_cigarettes
- Child-occupation_father
- Child-occupation_mother

Distinct: No

Aggregate - Functions that calculate simple statistics on database fields. The Distinct option avoids repeated records in the calculations.

Function #1:

Field #1:

Distinct: No

Function #2:

Field #2:

Distinct: No

Function #3:

Field #3:

Distinct: No

Grouping - Group your results by one or more database fields.

Grouping field #1:

Grouping field #2:

Grouping field #3:

Grouping field #4:

Grouping field #5:

Conditions - Specify restrictions on the records the query should return.

Condition field #1:

Operator #1:

Value #1:

Condition field #2:

Operator #2:

Value #2:

Condition field #3:

Operator #3:

Value #3:

Condition field #4:

Operator #4:

Value #4:

Condition field #5:

Operator #5:

Value #5:

Condition field #6:

Operator #6:

Value #6:

Condition field #7:

Operator #7:

Value #7:

Special Conditions - Sets of Conditions previously built in [Create Common Condition Fields](#).

special cara-micro-year = 2001 AND child-gender_male = 1

aa-bb

empty special field-

Naming - Give a name and description to your query.

Please write a name for your query

If you wish, you can write down a description of this query

Figure 7
User-Friendly Query System interface Interface where the users build their queries, which are then translated into SQL. Saved queries can be rerun, edited, deleted, or grouped into reports. Commonly used sets of restrictions can also be saved and later applied to new queries.

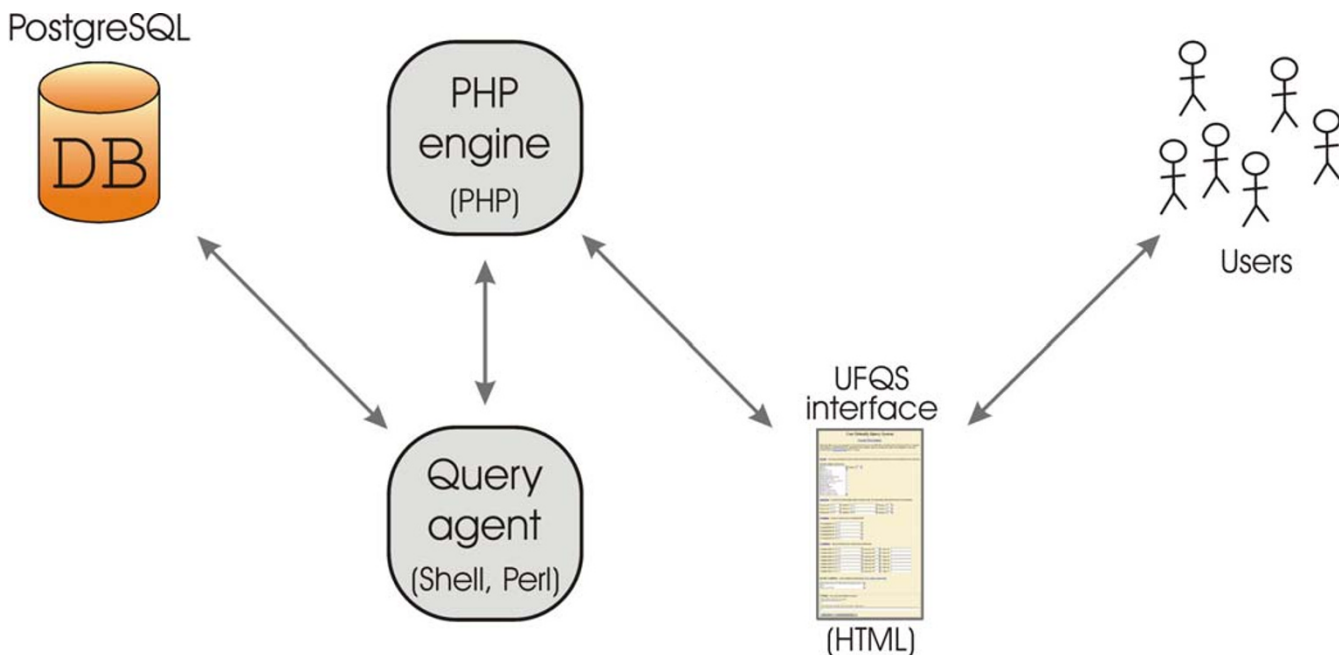


Figure 8
Operational model – query system Operational model of the EURISWEB query system. The arrows represent flow of information between the various entities. The user makes a request through the interface, which is translated into SQL by the PHP engine and sent to the query management agent. The query agent interacts with the database in background and informs the users, by e-mail, about the state of their requests.

layout and structure files. The core element of this operational functionality is supported by the PHP engine described in the Results section. As a result both maintenance and development scale well with increasing usage, particularly since availability of high performance hardware and Internet access have ceased to be an issue. It is noteworthy that the layout and structure files are particularly suited for extensions to the current model, including having new types of data integrated into the set already stored; having new countries and new country-specificities accommodated, while retaining previous accessibility and privacy. The demo version of the database shell and query system, made publicly available (see Availability for URL and login directions), was built by configuring a fictitious new country structure, which was achieved by performing minor modifications in the layout files.

This comes to illustrate that a possible useful extension to this system would be to allow selected users to manage their own tables and forms by providing a web-based interface to the layout and structure files. These users with management access permissions would not need to know the PHP or SQL languages, but simply interact with online forms with the same level of complexity as the regular database query forms. In our experience, most of the tasks

requested to the database manager are simple and pose no risk whatsoever to the data, like adding an item to a drop-down selection field, resizing a text edit field, or even adding a new field to an already existing table. Given the wide geographic distribution of the EURISWEB users, describing what needs to be done to the database manager is as time consuming as specifying it in such idealized management forms. The database manager could then be left with only the more complex and "dangerous" tasks. Figure 9 shows how the operational model implemented, including the query system, could be configured to greatly remove the need for low level data management.

Future directions

The extension of data management to include data analysis is particularly suited for web-based implementations – such as EURISWEB – since all computation takes place on the server side. This approach enables a bioinformatics approach to establish itself alongside data storage. As a consequence, advanced data mining tools such as multivariate statistical analysis or the identification of artificial intelligence predictive models using neural networks [25] and rule extraction by genetic algorithms can be made available alongside the data itself. This is mutually beneficial for usage and development and, on the other

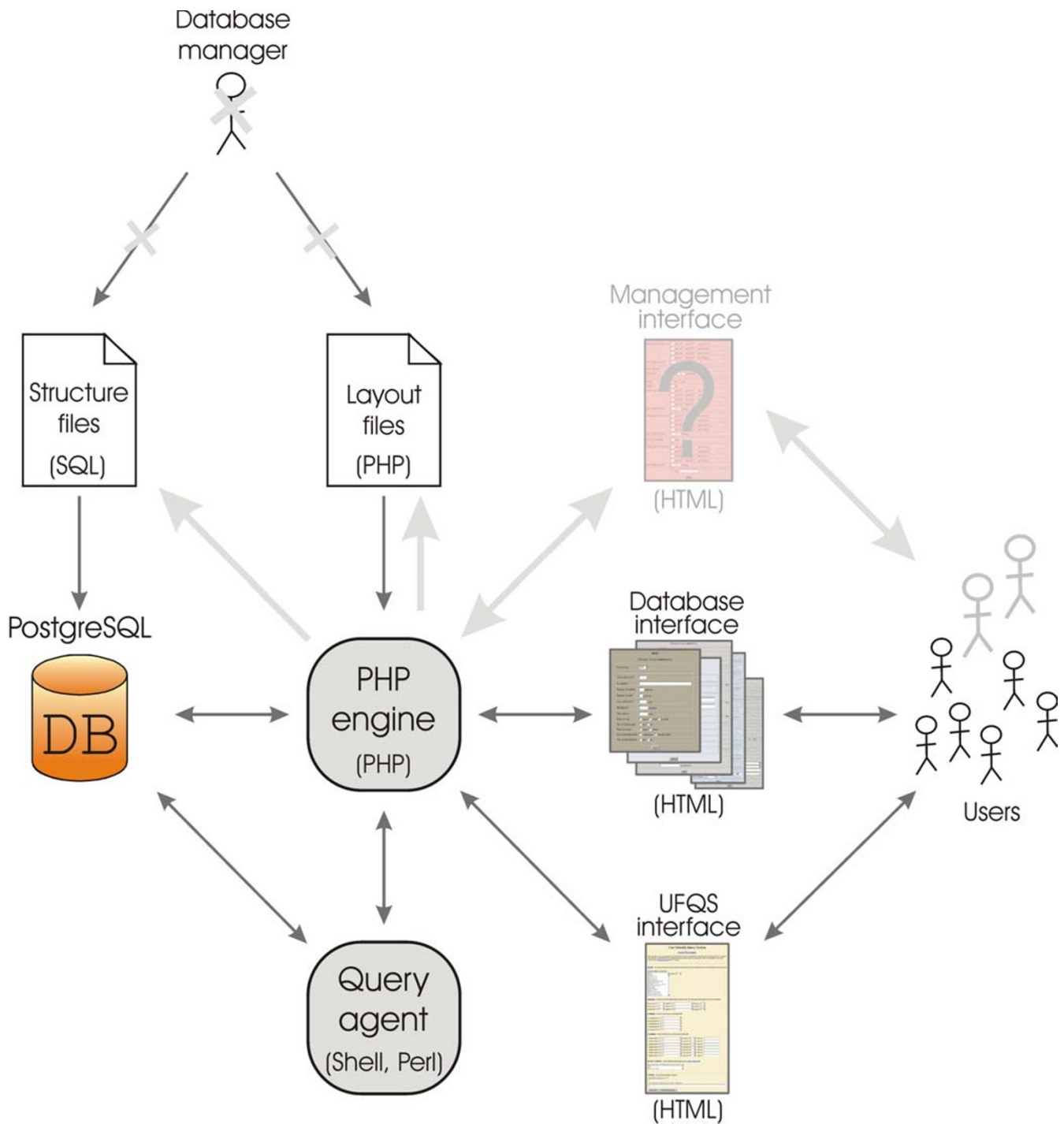


Figure 9
Operational model – the whole picture and hypothetical scenario Operational model of the whole EURISWEB infrastructure, by joining Figures 6 and 8, plus the hypothetical scenario where the database manager could be replaced by an extension to the PHP engine that would allow privileged users to manage both the layout and structure files through web pages.

hand, bioinformatics tool development has, in return, ready access to extensive datasets for validation as well as, even more important, facilitated interaction with domain experts that provide the context for its interpretation. This bi-directional integration enabled by a web-based development environment undoubtedly offers the best conditions for practical implementation of a full-fledged epidemiological information system.

Competing interests

None declared.

Authors' contributions

SS designed and manages the database, and prepared the manuscript draft. AM designed the query system and does the software and hardware maintenance. SS, RGO, AM and JC implemented the database and query system. TG, KGK, KE, AT, ISS, HL and JSA conceived of the EURIS Project, and participated in its design and coordination. ABA participated in the design of the study in Portugal. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by grant QLK2-CT-2000-01020 (EURIS) from the European Commission. JC was supported by grant SFRH/BD/3123/2000, and AM by grant POCTI/1999/BSE/34794 (SAPIENS), both by Fundação para a Ciência e a Tecnologia, Ministério da Ciência e do Ensino Superior, Portugal. We also thankfully acknowledge the EURIS Portuguese team (C. Simas, R. Mato, S. Nunes, N. Sousa, N. Frazão) and the Icelandic Team for early advice that was essential for the EURISWEB prototype design, as well as J. Saldanha (Portugal) for help in designing the questionnaires.

References

- Gray BM, Turner ME and Dillon HC Jr: **Epidemiologic studies of Streptococcus pneumoniae in infants. The effects of season and age on pneumococcal acquisition and carriage in the first 24 months of life** *Am J Epidemiol* 1982, **116**:692-703.
- Yagupsky P, Porat N, Fraser D, Prajgrod F, Merires M, McGee L, Klugman KP and Dagan R: **Acquisition, carriage, and transmission of pneumococci with decreased antibiotic susceptibility in young children attending a day care facility in southern Israel** *J Infect Dis* 1998, **177**:1003-1012.
- Butler JC, Dowell SF and Breiman RF: **Epidemiology of emerging pneumococcal drug resistance: implications for treatment and prevention** *Vaccine* 1998, **16**:1693-1697.
- Reichler MR, Allphin AA, Breiman RF, Schreiber JR, Arnold JE, McDougal LK, Facklam RR, Boxerbaum B, May D, Walton RO and Jacobs MR: **The spread of multiply resistant Streptococcus pneumoniae at a day care center in Ohio** *J Infect Dis* 1992, **166**:1346-1353.
- Arason VA, Kristinsson KG, Sigurdsson JA, Stefansdottir G, Molstad S and Gudmundsson S: **Do antimicrobials increase the carriage rate of penicillin resistant pneumococci in children? Cross sectional prevalence study** *BMJ* 1996, **313**:387-391.
- Forsell G, Hakansson A and Mansson NO: **Risk factors for respiratory tract infections in children aged 2-5 years** *Scand J Prim Health Care* 2001, **19**:122-125.
- Holmes SJ, Morrow AL and Pickering LK: **Child-care practices: effects of social change on the epidemiology of infectious diseases and antibiotic resistance** *Epidemiol Rev* 1996, **18**:10-28.
- McGee L, Klugman KP and Tomasz A: **Serotypes and clones of antibiotic-resistant pneumococci** In *Streptococcus pneumoniae. Molecular Biology & Mechanisms of Disease* Edited by: Tomasz A. New York: *Mary Ann Liebert*; 1996:375-379.
- McGee L, McDougal L, Zhou J, Spratt BG, Tenover FC, George R, Hakenbeck R, Hryniewicz W, Lefevre JC, Tomasz A and Klugman KP: **Nomenclature of major antimicrobial-resistant clones of Streptococcus pneumoniae defined by the pneumococcal molecular epidemiology network** *J Clin Microbiol* 2001, **39**:2565-2571.
- Sá-Leão R, Tomasz A, Sanches IS, Brito-Avô A, Vilhelmsson SE, Kristinsson KG and de Lencastre H: **Carriage of internationally spread clones of Streptococcus pneumoniae with unusual drug resistance patterns in children attending day care centers in Lisbon, Portugal** *J Infect Dis* 2000, **182**:1153-1160.
- Sá-Leão R, Tomasz A, Sanches IS, Nunes S, Alves CR, Brito-Avô A, Saldanha J, Kristinsson KG and de Lencastre H: **Genetic diversity and clonal patterns among antibiotic-susceptible and -resistant Streptococcus pneumoniae colonizing children: day care centers as autonomous epidemiological units** *J Clin Microbiol* 2000, **38**:4137-4144.
- de Lencastre H and Tomasz A: **From ecological reservoir to disease: the nasopharynx, day care centres and drug resistant clones of Streptococcus pneumoniae** *J Antimicrob Chemother* 2002, **50**(Suppl C):75-81.
- The EURIS project [<http://www.itqb.unl.pt:1111/euris/>]
- Kristinsson KG: **Effect of antimicrobial use and other risk factors on antimicrobial resistance in pneumococci** *Microb Drug Resist* 1997, **3**:117-123.
- Boken DJ, Chartrand SA, Moland ES and Goering RV: **Colonization with penicillin-nonsusceptible Streptococcus pneumoniae in urban and rural child-care centers** *Pediatr Infect Dis J* 1996, **15**:667-672.
- Muhlemann K, Matter HC, Tauber MG and Bodmer T: **Nationwide surveillance of nasopharyngeal Streptococcus pneumoniae isolates from children with respiratory infection, Switzerland 1998-1999** *J Infect Dis* 2003, **187**:589-596.
- EURIS Login [<https://www.itqb.unl.pt:1122/>]
- PHP: Hypertext Preprocessor [<http://www.php.net/>].
- PostgreSQL [<http://www.postgresql.org/>]
- Perl.com: The Source for Perl - perl development, perl conferences [<http://www.perl.com/>].
- The Linux Home Page at Linux Online [<http://www.linux.org/>]
- The Slackware Linux Project [<http://www.slackware.com/>]
- Apache-SSL [<http://www.apache-ssl.org/>]
- Welcome! - The Apache Software Foundation [<http://www.apache.org/>]
- Almeida JS: **Predictive non-linear modeling of complex data by artificial neural networks** *Curr Opin Biotechnol* 2002, **13**:72-76.

Pre-publication history

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

<http://www.biomedcentral.com/1472-6947/3/9/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

