Ubiquity: Technologies for Better Health in Aging Societies
A. Hasman et al. (Eds.)
IOS Press, 2006

Log files analysis to assess the use and workload of a dynamic web server dedicated to End-Stage Renal Disease

Mohamed BEN SAID, Loic LE MIGNOT, Jean Baptiste RICHARD, Christine LE BIHAN, Laurent TOUBIANA, Jean-Philippe JAIS, Paul LANDAIS [1]

Université Paris 5 R. Descartes – UPRES EA 222- AP-HP- Service de Biostatistique et d'Informatique Médicale, Hôpital Necker-Enfants Malades- Paris- France

Abstract

A Multi-Source Information System (MSIS), has been designed for the Renal Epidemiology and Information Network (REIN) dedicated to End-Stage Renal Disease (ESRD). MSIS aims at providing reliable follow-up data for ESRD patients. It is based on an n-tier architecture, made out of a universal client, a dynamic Web server connected to a production database and to a data warehouse. MSIS is operational since 2002 and progressively deployed in 9 regions in France. It includes 16,677 patients. We show that the analysis of MSIS web log files allows evaluating the use of the system and the workload in a public-health perspective.

Key words: Web log file; workload; Dynamic Web server; n-tier architecture; system use; Multi-Source Information System; Internet; End-Stage Renal Disease.

1. Introduction

A Multi-Source Information System (MSIS) [1] was set up, dedicated to collect continuous and completed records of all patients presenting with End-Stage Renal Disease (ESRD) and their clinical follow-up. MSIS collates in a standardized representation a minimal patient record elaborated by health professionals [2].
Our project was to provide the users a tool facilitating the access to useful information concerning ESRD demand and offer of care, allowing the national sharing of the results, and supporting public health decisions at regional and national levels for adapting the offer to the demand of ESRD care.
This paper describes a method to assess the use of the system and the workload through web log analysis.

2. Methods

2.1 MSIS organizational support

MSIS organization has been described elsewhere [3]. Briefly, MSIS is part of the Renal Epidemiology and Information Network (REIN). A national committee insures

[1] Corresponding author: Pr Paul LANDAIS – Service de Biostatistique et d'Informatique Médicale – Hôpital Necker Enfants Malades – 149, rue de Sèvres – 75743 Paris – FRANCE – email : landais@necker.fr

guidance and follow-up and involves several organizations: Société de Néphrologie, Société francophone de dialyse, Paris Descartes University and Grenoble J. Fourier University, Agence de la Biomédecine, Caisse Nationale d'Assurance Maladie des Travailleurs Salariés, INSERM, Direction de l'Hospitalisation et de l'Organisation des Soins, Institut de Veille Sanitaire and representatives of patients' associations.
In each French region, regional committees involve nephrologists, decision makers, public health insurers, epidemiologists and representatives of patients' associations. A nephrologist has been elected as program coordinator. A public health and epidemiology department provides resources and expertise for methodology and epidemiological studies. Once a year, a clinical research assistant performs the quality control for every patient record.
The agreement of the French "Commission Informatique et Libertés" (data protection act 78-17) was obtained, in accordance with the European Convention 108 and the directives n°95/46/CE and 2002/58/CE.

2.2 MSIS Design and Implementation

MSIS-REIN is based on a n-tier architecture [1]. Via a web browser, the client tier connects to a middle tier that is in relation with several databases: the identification database, the production database and the data warehouse. The middle tier supports client services through Web containers and business logic services through component containers. Business logic components in the middleware support transactions toward the databases.
At the client side, MSIS relies on existing local Internet networking facilities and on a widely spread computer configuration in medical settings. Maintenance and evolutions are made centrally which reduce deployment costs and delays.
The production database was structured according to users' profiles, administrative regions, care units and patients' information:
- Users' profiles concern end users, mainly physicians, regional coordinators, clinical research assistants and MSIS administrators (users roles are summarized on table 1).
Passwords are delivered individually to the users to access to patients' data of their own unit. The regional coordinator and the clinical research assistant have access to all units within their region in order to perform data quality control. MSIS access codes (Table 2) were provided to more than 430 nephrologists and 100 codes to their collaborators. Ten clinical research assistants and 9 physicians perform the quality control at the regional level. Eight university departments of medical informatics, epidemiology or public health, insure the patient data quality control using MSIS.
- Data of each administrative region is collected into a specific protected database subset.
- Patient records comprise two parts: patient identification and patient medical record. The patient medical record is composed of three parts: medical history, aetiology of ESRD and comorbidity at start of replacement therapy; recent medical record with information about access to the care facilities and to the national kidney-graft waiting list; an annual update of the actual renal dialysis method and context of treatment.
Admission, discharge and transfer event information's are documented and updated annually on the anniversary of first ESRD treatment. A decease record file, including standard medical codification of the decease is documented when necessary.

Table 1 – MSIS users' roles.

Users profiles	Roles
Regional Coordinator	- Coordinates regional activities - Delivers access authorizations to the Information System - Organizes data analysis - Represents the nephrologists in the meetings with the health authorities
Regional Clinical Research Assistant	Performs : - User training and accompaniment - Data completeness checking : systematic cross matching with lists of patients treated in each unit - Data quality control : in each unit, using sampling procedures, data assessment by going back to the patient medical record - Data consolidation to allow feeding the data warehouse - Preparing feed back reports based on the region data
Nephrologists and health professionals	- Provide medical expertise and fulfills patient records - Validate patient records

Consolidated patient data are loaded periodically to a data warehouse coupled with a Geographic Information System called SIGNe [4, 5]. SIGNe provides MSIS users with dynamic analysis views in the form of tables, charts and graphics, didactic maps as well as preformatted ready-to-print reports. Interactive queries allow representing regional or inter-regional profiles such as: incidence and prevalence of ESRD; Patient treatment trajectory, transportation modes, distances and duration to reach their care unit, care units catchments areas; comorbidities, typology, and impact on ESRD health care; follow up of ESRD treatment, control of anaemia and erythropoietin treatment; distribution of dialysis treatment methods: haemodialysis, peritoneal dialysis, or transplantation according to the typology of care units: centre, medical unit, self-dialysis, home dialysis.

Table 2 – MSIS – User profiles in eight administrative regions where MSIS is deployed.

User Profile / Year	2002	2003	2004	2005 semester 1	2005 semester 2
Nephrologists	14	75	199	261	434
Health care professionals in Nephrology setting		13	25	30	105
Clinical Research Assistants	1	3	5	7	10
Medical informatics and/or epidemiology and/or public health professionals	3	5	7	8	10
Total	18	96	236	306	559

The log files of the JSP/Servlet web server container (Tomcat 5) keep track of every single client request and server response in a multi line text record. We studied the Web log files over a period of 17th months from June 16th 2004 to Jan 15th 2006. They consist of large text files which size 17.8 gigabytes. We first used AWK programming language [6] to process them and extract information of interest. Then we used Mysql© to re-arrange the information for analysis. Analysis was developed according to several topics: exploring the frequency of requesting of specific functionalities, of an executable program, or downloading a given document. We then rapidly moved to focus on the sessions' key identifiers and how to organize the information. It made computations easier to retrieve specific information or the MSIS user's profile [7].

3. Results

As of January 30th 2006, 16,677 patient records, 1504 transplantations, 2300 transfers from a dialysis unit to another and 4012 ESRD deaths are documented in the production database. The active file includes 11,146 patients who undergo dialysis (detailed information are presented in table 3). According to the national survey of prevalent ESRD dialyzed patients in June 2003 [8], MSIS active file includes 35 % of the nation wide ESRD dialyzed patients. As longitudinal data accumulate, annual follow-up of the ESRD cohort are progressively and systematically taking place in the regions.

We identified 17,281 user's sessions during the 17th month period of observing the log files. They correspond to 348 identified users who connected several times. This observations also means that 2 out of 3 authorized users (348/559; 62%) have been connected effectively to MSIS.

We observed the duration of the session associated with each user, the number of sessions and the total duration a user had over a period of 17 months. Cross matching with MSIS tables allowed retrieving information sorted by user's profile and by region (table 4).

Epidemiologists and public health professionals at the Biostatistics and Medical Informatics Department at Necker Hospital in Paris - France, connected to MSIS mainly to assist online users and the clinical research assistants in the regions. During the study period, they created 3271 valid sessions representing a total duration of 1641 hours 49 minutes with an average of 29:37 per session.

Consolidated data by the clinical research assistant was progressively and periodically loaded into the data warehouse for analysis and data presentation for decision making.

4. Discussion

Nephrologists direct involvement in using MSIS to update patient information varies according to the size of the region or the number of units of care. In two small regions with small number of nephrologists, the Clinical Research Assistant updates patient information using MSIS. While in regions where more nephrologists are involved in the ESRD patient care, the nephrologists and the health care professionals directly use the MSIS to update patient information.

The web log files analysis showed an important involvement among the nephrologists and their teams to directly use the MSIS and document patients' information. They

Table 3 – MSIS deployment (Jan 30th,2006). The total population in the regions where MSIS is used represents 26,211,000 inhabitants (43% of the French population).

Region	Date of inclusion	Population	Number of cases in the active file
Limousin	Jan 22nd 2002	711,000	374
Languedoc-Roussillon	Jun 2nd 2002	2,296,000	1,537
Champagne-Ardenne	Jan 1st 2003	1,342,000	700
Centre	Jan 1st 2004	2,440,000	1,303
Provence-Alpes-Côte-d'Azur	Jan 1st 2004	4,506,000	3,120
Ile-de-France	Nov 1st 2004	10,952,000	1,878
Midi-Pyrénées	Feb 1st 2005	2,552,000	1,425
Basse Normandie	Feb 1st 2005	1,422,000	682
Paediatrics virtual region	Jan 1st 2004	N/A	127

Table 4 – Patterns of use of MSIS in the regions during the period of 17 months between June 16th 2004 and Jan 15th 2006.

Region	Nephrologists & Health care professionals			Clinical Research Assistants & Referral Nephrologists		
	Nbr sessions	Session duration h:mn:sec	Average session mn:sec	Nbr sessions	Duration h:mn:sec	Average session mn:sec
Limousin	2	00:12:20	06:10	915	195:16:23	13:10
Languedoc-Roussillon	719	484:14:59	40:24	1095	984:21:24	53:56
Champagne-Ardenne	578	268:10:25	27:50	454	139:45:05	18:28
Centre	361	184:55:55	30:44	1537	1121:14:35	43:46
Provence-Alpes-Côte-d'Azur	1145	638:26:36	33:27	1111	825:38:47	44:35
Ile-de-France	1885	831:47:55	26:28	1013	706:21:48	41:50
Midi-Pyrénées	998	348:03:33	20:55	572	261:56:40	27:28
Basse Normandie	5	20:32	04:06	416	50:07:21	07:13
"Paediatrics"	200	123:25:45	37:01	75	04:22:29	03:29

created 5893 sessions and connected during 2879 hours in the study period between June 16th 2004 and January 15th 2006.
The clinical research assistants insure the data quality control and the completeness of the patient information. They heavily rely on the MSIS to document patient files, to control the data quality and to prepare statistical analyses in the region for supporting health care decision making. They connected 7188 times during 4289 hours with an average of 35 minutes during the observation period of 17 months.
The department of Biostatistics and Medical informatics at Necker Hospital in Paris accompanied the nephrologists, clinical research assistants and all the MSIS users. This technical accompaniment proved necessary not only at beginning of the loading process, but also during its growth and evolution. Professionals from the laboratory created 3271 sessions which lasted 1614 hours.

5. Conclusion

Four years of on-line use of MSIS showed that the involvement of the professionals was effective as they bring the expertise and the accuracy to the patient information record and improve epidemiology and support to decision making. The technical research assistants brought an essential contribution to obtain the completeness of the data and permitted an efficient quality control assessed by web log analysis.

Acknowledgments

The nephrologists in charge of the ESRD units in Limousin, Languedoc-Roussillon, Champagne Ardenne, Centre, Provence-Alpes-Côte d'Azur, Basse-Normandie, Midi-Pyrénées, Ile-de-France, Paediatrics are acknowledged for their fruitful cooperation and comments. JP Necker and X Ferreira are acknowledged for their skilful help. This work was supported by a grant of the University of Paris 5 and by the Agence de la Biomédecine.

[1] Landais P, Simonet A, Guillon D, Jacquelinet C, Ben Saïd M, Mugnier C, Simonet M. SIMS@REIN : Un Système d'Information Multi-Sources pour l'Insuffisance Rénale Terminale. *CR Biol* 2002;325:515-528.

[2] Ben Said M, Simonet A, Guillon D, Jacquelinet C, Gaspoz F, Dufour E, Mugnier C, Jais JP, Simonet M, Landais P. A dynamic Web application within an n-tier architecture: a Multi-Source Information System for end-stage renal disease. Stud Health Technol Inform. 2003;95 ::95-100.

[3] Ben Said M, A Multi-Source Information System via the Internet for End-Stage Renal Disease: Scalability and Data Quality. Stud Health Technol Inform 2005 ;116 :994-999.

[4] Richard JB, Toubiana L, Le Mignot L, Ben Said M, Mugnier C, Le Bihan Benjamin C, Jais JP, Landais P. SIGNe: A Geographic Information System on the Web for End Stage Renal Disease. Stud Health Technol Inform 2005 ;116 :713-718.

[5] Richard JB, Toubiana L, Le Mignot L, Ben Said M, Mugnier C, Le Bihan Benjamin C, Jais JP, Landais P. A Web GIS for Health Care Decision Support. AMIA 2005 Symposium Proceedings, Page 365-369

[6] Aho AV, Kernighan BW, Weinberger PJ. The AWK Programming Language — Eds: Addison Wesley Publishing Company – Reading Massachusetts 01867 – 1988 – ISBN 0-201-07981-X

[7] Hoffer EP, Feldman MJ, Kim RJ, Famiglietti KT, Barnett GO. "DXplain: Patterns of Use of a Mature Expert System" in AMIA 2005 Symposium Proceedings – Page 321-324 – Washington DC October 22-26, 2005 – Charles Friedman Editor.

[8] Macron-Nogues F, Vernay M, Ekong E, Thiard B, Salanave B, Fender P, Allemand H. The Prevalence of ESRD Treated With Renal Dialysis in France in 2003. Am J Kidney Dis. 2005 Aug;46(2):309-15