MEDINFO 2010
C. Safran et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-588-4-939

Data mining techniques for analyzing stroke care processes

Silvia Panzarasa[a], Silvana Quaglini[b], Lucia Sacchi[b], Anna Cavallini[c], Giuseppe Micieli[c], Mario Stefanelli[b]

[a] CBIM, Pavia, Italy
[b] Dept. of Computer Engineering and Systems Science, University of Pavia, Italy
[c] IRCCS Foundation C. Mondino, Pavia, Italy

Abstract

Controlled randomized clinical trials and meta-analyses show that stroke patients benefit from access to specialized Stroke Units, in terms of mortality, disability and dependency. However, many issues relating to stroke diagnosis and therapy and to the organization of stroke care remain to be solved and little is known about what interventions make Stroke Units more effective. It is also agreed that compliance with clinical practice guidelines improves health outcomes for these patients, but little is known about the relative weight of the different guideline recommendations. Over the last decade, many hospital- or population-based stroke registers have been set up with the aim of identifying specific key indicators able to monitor the quality and adequacy of acute stroke care. Registers seem to be adequate tools for collecting the data needed to analyze care processes, providing data useful for both national healthcare planning and scientific research. In this paper we applied data mining techniques to data collected within the stroke register of the Lombardia region in Italy. From our analyses both expected and unexpected results have been found: not always compliance to recommendations is related to a good patients' outcome.

Keywords:

Data mining, Stroke care, Clinical practice guidelines.

Introduction

Stroke is a leading cause of disability and death in developed countries. Since many years, national scientific communities work to support specialists with the most appropriate approach to stroke care. In Italy, the Stroke Prevention and Educational Awareness Diffusion (SPREAD, www.spread.it) collaboration, active since 1998, involves multidisciplinary experts coming from 36 different organizations together with representatives of patients' associations. The goal of SPREAD is to make available clinical practice guidelines (GLs) for stroke management, from prevention to acute and post-acute phase. These GLs are a set of practical recommendations periodically revised (we are now currently at 5th edition) and based as much as possible on the best available scientific evidence: according to the strength of the supporting evidence, recommendations are labelled with decreasing degrees from A to D.

In previous works, we described how these GL recommendations have been translated in logical (IF-THEN) rules and integrated within an existing electronic patient record (EPR) through a careflow management system [1]. The result has been a real-time decision-support system based on clinical evidence, which has been implemented in our hospital. Moreover, a module has been implemented, called RoMA (Reasoning on Medical Actions) [2], able to detect non-compliances with GLs. In principle non-compliances could be weighted with the recommendations degree. RoMA is a flexible tool that can be integrated with different databases, such as different EPRs. It has been used on about 500 cases to evaluate adherence to SPREAD GL before and after the implementation of the decision support system [1], showing its effectiveness.

In order to obtain more significant statistics with respect to the analysis of a single EPR, RoMA has then been used to analyze the bigger amount of data collected in the Stroke Unit Network (SUN) register, still with the aim of verifying GL compliance and its correlation with patients' outcome. The results so far obtained show that good compliance improves the clinical outcome. However, this is an *overall* result, correlating the outcome with the total number of non-compliances, while no result exists about the influence of every single recommendation, or set of similar recommendations, on the outcomes.

Therefore, there are several open questions:

1. Which are the recommendations more related to a good/poor outcome? Is it possible to make some inference on the importance of the recommendations with low scientific evidence degree?
2. Which is the most effective organisational setting for a Stroke Unit? Which facilities, among those distinguishing a Stroke Unit from a regular ward, are more influential on the outcome?
3. Which are the major patterns of non-compliance, i.e. is it possible to individuate predictors of non-compliance for the different SPREAD recommendations?

4. Are these patterns more related to clinical or organisational factors?

As shown in [3], data mining techniques can be profitably used in this area. Thus we used decision trees and rule extraction methods on the SUN register data to try answering some of the above questions.

The paper organisation is the following: the next section illustrates the data available in the register; then the GL recommendations are illustrated together with data necessary for testing the compliance; then we describe the data mining techniques used; eventually, analyses and results are shown.

The Stroke Unit Network Register

The SUN register is part of a web-based network connecting 36 Stroke Units and Neurological Departments in the Lombardia region, Italy, with the aim at improving the quality of stroke care in the acute and post-acute phase. The SUN website facilitates storing homogeneous data, debating among specialists and sharing research projects and results.

The data collected within the SUN register can be split into different categories as shown in Table 1.

First, there are general information about patient's demographics, admission modalities, stroke type and severity scales.

In the second category there are patient's clinical data such as risk factors for cerebro-cardiovascular diseases and medical complications that arise during the hospital stay.

In the third category there are indicators of the care process. Some of them describe diagnostic patterns: whether or not ECG, neuroimaging and other basic tests have been performed in the emergency room, and all the instrumental tests performed during the hospital stay. Timing of each performed test is also recorded. Others are indicators of good medical assistance: again, both type and timing of every assistance procedure are stored. Finally there are indicators of therapeutic procedures, both for the acute phase and for the early clinical phase, treatment for comorbidities (diabetes, hypertension, etc.) and counselling for healthy lifestyle.

The fourth category is composed by outcome indicators, among which the Rankin scale, collected both at discharge and at three months follow-up.

Moreover, each clinical centre is provided with a description of the hospital to which it belongs (e.g. rural or not, university-related or not), of its organisation (e.g. human resources), and diagnostic facilities (e.g. presence and 24h availability of a service).

The register is provided with procedures for data quality control both in real time and in batch.

The register and its database have been designed and implemented during 2006, while the patients' recruitment started in January 2007. At January 2009, the register contained 5079 cases of ischemic strokes, which represent our study population.

Table 1 – Data collected within the Stroke Registry

<table>
<tr><th colspan="3">General information</th></tr>
<tr><td>Demographics</td><td colspan="2">Stroke onset</td></tr>
<tr><td>Birth date
Gender
Race</td><td colspan="2">Time
Rankin and NIHSS scales (higher values indicate higher severity)
Diagnosis (TOAST Classification)</td></tr>
<tr><th colspan="3">Clinical data</th></tr>
<tr><td>Risk factors</td><td colspan="2">Medical complications</td></tr>
<tr><td>Smoking; Previous stroke Atrial fibrillation; Hypertension; Dementia; Diabetes; …</td><td colspan="2">Hypertension; Hyperglycaemia; Myocardial Infarction; Hypertermia; Endocranial Hypertension; Seizures; …</td></tr>
<tr><th colspan="3">Process data</th></tr>
<tr><td>Diagnostic work-up</td><td>Medical assistance</td><td>Therapeutic patterns</td></tr>
<tr><td>ECG; Coagulation screening, carotid, vertebral, peripheral duplex, scan, transcranial Doppler; Echocardiography; Brain image; Cerebral angiography.</td><td>Neurological evaluation; Monitor of vital signs; Dysphagia screening; Early mobilization; Positioning of gastric probe and bladder catheter; Deep venous thrombosis prophylaxis.</td><td>Thrombolisis; Antiplatelets; Anticoagulants; Surgical procedures; Additional drugs for complications treatment; counselling.</td></tr>
<tr><th colspan="3">Outcome indicators</th></tr>
<tr><td colspan="3">Mortality, Recurrence; Disability (Rankin).</td></tr>
</table>

The guideline recommendations

As we are dealing with a disease register, we must point out that it contains less data, for an individual patient, with respect to a clinical chart, which reports much more detail. That means that we have fewer data fields available for testing GL compliance. On the other hand, a register allows a quick collection of a high number of patients, from a high number of clinical centers, thus allowing sounder statistics.

The information needed to assess the GL compliance has been matched with the data model of the SUN register in order to obtain IF-THEN rules as a formal representation of recommendations. The IF part of a rule is called *eligibility condition*, while the THEN part is the *action*. Three examples of this matching procedure are reported in Table 2. We remark that such a matching is not a trivial task, and it requires the collaboration of physicians in order to interpret the text, often ambiguous, of the GL, and to include some "hidden" knowledge that is not explicitly written in the GL, because part of the cultural background of any physician (e.g. which tests belong to a given diagnostic work-up, which diseases belong to a certain category, etc) or simply because obvious (e.g. the patient must be alive at a certain time for being evaluated).

Table 2 – Examples of matching the Register data with GL recommendations, in order to assess compliance

Recommendation	GL section	Register data used
5.5(D) - After a TIA or a stroke, transthoracic echocardiography (TTE) is recommended only when a heart disease is clinically suspected.	Diagnostic work-up	Alive at 3 days, TTE execution, myocardial infarction, atrial fibrillation, coronaropathy, heart failure, arrhythmias, prosthesis
9.7(D) - A non-contrast CT scan is recommended as soon as possible in the emergency care: to allow the differential diagnosis between ischaemic and haemorrhagic stroke, and with non-cerebrovascular lesions; to detect possible early signs of infarct.	Acute stroke:hospital admission (diagnostic procedures)	Time of stroke onset, time of hospital admission, time and setting of the neuroimaging execution
12.6a (B) - After stroke or TIA, lowering of high blood pressure is recommended, preferably using agents active on the renin-angiotensin system, calcium channel blockers and diuretics.	Secondary prevention	Alive at discharge, hypertension as risk factor or complication, antihypertensive therapy at discharge

The overall procedure result was that compliance can be verified for 12 recommendations of the acute phase and 4 of the secondary prevention phase, subdivided according to their scientific degree as follow: 2 with grade A, 4 with grade B, 1 with grade C, 8 with grade D and 1 only supported by experts consensus. While, in general, a guideline can contain recommendations that either suggest or advise against an action, all the 12 recommendations are of the first type and the non-compliance is defined as a case where the eligibility condition is satisfied and the action has not been performed.

The rules obtained have been used within the RoMA tool to assign every patient with a 1/0/null flag for any recommendation, representing non-compliance/compliance/no-eligibility respectively.

Data mining techniques

To analyze the impact of non-compliance to clinical GL on patients' outcome, we herein chose to exploit Data Mining (DM) techniques besides the classical statistical analysis that is the most frequent approach in these cases. The reason why DM can be useful is that it provides some tools to model the possible paths that lead to the outcome in a way that can be easily visualized and interpreted also by clinicians. These motivations led us to the choice of a specific DM algorithm, i.e. Classification Trees.

Classification Trees are a classification algorithm that works by recursively partitioning the examples space on the base of the most discriminant feature to predict the outcome (also known as class variable). Besides providing a method for class prediction, classification trees can be also a useful instrument for data exploration and understanding. Each branch of the extracted classification tree not only leads to the class prediction (in the leaves), but also identifies a clinical path that can be usefully considered for results interpretation. Each branch of the tree can be in fact seen as a classification rule where the body of the rule is represented by the conditions that characterize the patients belonging to that branch, while the head of the rule is the outcome. Such a representation is very suitable to identify the patterns that can lead patients to a particular outcome.

For Classification Trees construction we used the algorithm that is implemented in the Orange Data Mining Suite [4], choosing Gain Ratio as attribute selection criterion. Moreover, we set a minimum number of 100 patients in the leaves in order to avoid an excessive expansion of the tree obtaining too small and not significant patients sets. As we will see in the following paragraphs, we herein exploit classification trees mainly to analyze and discuss with our clinical partners about the clinical paths that lead to a certain outcome. For these reason the Orange widget "Classification Tree Graph" was very useful for results visualization and presentation.

Besides Orange, also S-Plus (Tibco Software Inc., Palo Alto, CA) was used for more classical statistical analysis such as logistic regression.

Analyses and Results

In order to produce results easily interpretable, different types of analyses have been performed using subsets of variables instead of using all the variables together, and continuous/ordinal variables have been discretized as follows:

Age: <=45years, 46-79, >=80

NIHSS scale: 0-7, 8-14, >14

Rankin pre-stroke scale: 0-1, 2-3, 4-5

Rankin scale at discharge/follow-up: good (0-2), bad (>2, including death). This binary value has been chosen as the main patients' outcome.

Delay between stroke onset and hospital admission: <=3 hours, 3-6, 6-24, >24.

The analysis has been performed using 5079 cases. The outcome at discharge was good for 2737 patients and bad for 2342, while the outcome at follow-up, where 2739 patients were available, was good for 1826 patients and bad for 913 (*good* and *bad* defined as illustrated above). On the average 3760 patients were eligible for each recommendation (range: 109 to 5079) and the average non-compliance rate was 36% (range: 0% to 80%).

The compliance flags, produced by the RoMA tool for the above mentioned 16 recommendations, have been used as explanatory variables for the health outcome, the discretized Rankin scale both at discharge and at 3-months follow-up. Age, gender, and severity at admission, that are known to be correlated with health outcomes, have been included in all the analyses as correction factors.

In the next paragraphs we show some results from different types of analyses. For sake of space, only the most significant results are shown.

Non-compliance vs outcome

In principle, GL compliance should be correlated to a good outcome, mainly when recommendations are based on scientific evidence deriving from clinical trials. This is not guaranteed in the real-world setting indeed, because the clinical routine can be very different from the *ideal* conditions in which trials are conducted. As an example, about short-term outcome, we obtained the classification tree in Figure 1. The tree has been derived considering the *Rankin scale at discharge* (good/bad) as outcome and the *non-compliance/compliance* to the set of recommendations of the acute phase as features. Figure 1 suggests that, after considering stroke severity represented by NIHSS and Rankin scales, that resulted the most discriminant variables, and age and gender (not appearing in the tree because not discriminant), the non-compliance to the recommendation 5.15 (namely *Transcranial Doppler is a complementary examination in patients with a recent TIA or stroke. It may provide additional information on patency of cerebral vessels, recanalization and collateral pathways; grade D*) is the most important to determine poor outcome in patients with less severe initial status. This is in agreement with the current physicians' feeling that this examination is very important to drive the subsequent diagnostic work-up, despite its low degree of scientific evidence.

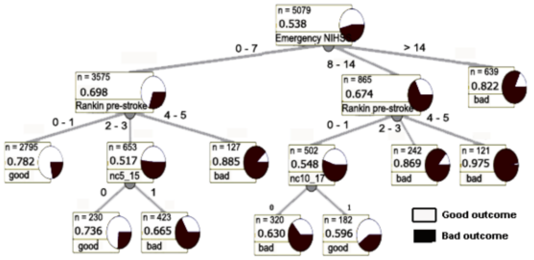

Figure 1 – The classification tree for the identification of the recommendations that are most related to the health outcome. "nc X" means "non-compliance with recommendation X". Bottom lines in the leaves indicate the majority class of the outcome (dichotomized Rankin scale), and the numbers represent the sample size (top) and the proportion of cases belonging to the majority class (middle).

On the contrary, in patients with a slightly more severe initial status (NIHSS 8-14), compliance to recommendation 10.17 (namely *Early mobilisation and hydration are always indicated to prevent Deep Venous Thrombosis. Graded compression stockings as well as intermittent pneumatic compression are recommended in combination with anticoagulation or as an alternative when anticoagulation is contra-indicate; grade D*) seems to be not as beneficial as expected (see Figure 1 again, bad outcome in 63% of compliance cases, and good outcome in about 60% of non-compliance cases) . The result is in agreement with a recent publication [5] that raised some doubts among such prevention procedures, which mainly consist in using graduated compression stockings. This is a very interesting exploitation of data mining, showing that a continuous monitoring of the collected data can anticipate and support some literature results, providing early indications for both GL developers and users.

The results shown in Figure 1, both for recommendation 5.15 and 10.17, are confirmed considering long-term outcome (Rankin scale at three-months follow-up). Among the 4 questions mentionned in the Introduction, the analyses just described are related to question #1.

Another type of analysis, still related to question #1, has been performed using every single recommendation, with the usual correction factors, as predictor of the health outcome. This analysis has been run through logistic regression, using age as a continuous variable. The rationale for this investigation is that some clinical trials enroll only a subset of the entire population (only males, only less than 80years, etc.) and the results are then extended to the entire population. Otherwise, a trial may involve unselected population, but lack of the so-called "subgroup analysis" may hide important distinctions about the treatment effect in different patients subsets. Therefore, we investigated whether some recommendations do not hold for the whole sample of our patients.

An interesting result has been obtained running logistic regression with recommendation 12.6a (hypertension treatment, see text in Table 2) as explanatory variable. In particular, the presence of a non-compliance to this recommendation resulted to have a negative correlation with the outcome (i.e. patients with non-compliance had better outcome). In order to gain a deeper insight into this result and since also the age variable turned out to be significantly correlated with the outcome in the logistic regression, we tried to stratify the patients sample on the base of their age. The main goal of this step was to see whether there exists a group of patients where this behavior is stronger. We created two age groups relying on the median of the age distribution (75 years) and we obtained a significant negative correlation between non-compliance and outcome with an odds ratio of 0.25 for the population of older patients with mild/severe impairment. Results on the younger patients group were instead not statistically significant.

After discussing this result with medical partners, they realised that it could be sensible, because the antihypertension treatment may induce a too high reduction of blood pressure, thus compromising the cerebral autoregulation mechanisms and causing secondary damage from cerebral hypoperfusion. This can be even more possible for elderly people living alone. For these people it is in fact difficult to monitor drug assumption and provide for prompt dosage correction if needed. The bad effect of lack of blood pressure response (and thus hypoperfusion) in stroke was already described in [6], but the study was limited to the acute phase. After finding the result described above, a literature search was performed and our finding has been strengthened [7,8].

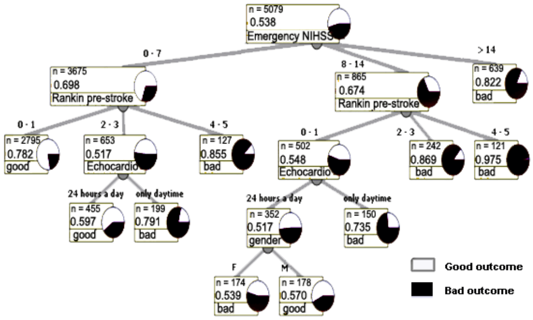

Figure 2 – The classification tree for the identification of the organizational variables that are most related to the health outcome.

Organizational setting vs outcome

To answer question #2 of the Introduction, an analysis has been performed using organisational variables, again besides considering severity, age and gender. Among the resources that characterize the medical units in terms of diagnostic facilities, only the possibility of performing echocardiography over the 24 hours was found to be related to the health outcome, as shown in Figure 2. Note that the result holds only for a subset of patients with moderate impairment.

Conclusion

Our study demonstrates that data mining techniques are a useful tool for (a) continuous monitoring of GL recommendations when implemented in the real-world clinical routine, (b) promoting discussion among GL developers and users and (c) suggesting priorities for equipping stroke units with diagnostic and treatment facilities. In particular, the results obtained in this work stimulated useful discussion among clinical experts, rising doubts on the general applicability of GLs and providing a feedback to the SPREAD GL developers themselves. Interestingly, the debate fostered a deeper literature search, that eventually strengthened our findings. From another point of view, when data-mining techniques show unexpected results, these can suggest further clinical research. It must be stressed, in fact, that data mining cannot replace clinical trials, but only give suggestions for their planning.

About the support to organisational issues, we observed that transcranial doppler examination leads to a better outcome, both at short and long-term. This suggests this equipment, still rare, should be highly recommended: its initial cost will be very likely justified by long-term cost-saving. Similarly, enough personnel to ensure 24h echocardiography should be provided in every hospital that treats stroke patients.

In a near future further analyses will be performed in order to answer also questions #3 and #4 enumerated in the Introduction. In particular, individuating predictors of non-compliance and organisational pitfalls could allow to enact the opportune correction actions to modify the incorrect behaviours and to enhance clinical settings.

Acknowledgments

We are grateful to all the participants of the SUN project for providing data in the stroke register.

References

[1] Panzarasa S, Quaglini S, Micieli G, Marcheselli S, Pessina M, Pernice C, Cavallini A, Stefanelli M. Improving compliance to guidelines through workflow technology: implementation and results in a Stroke Unit. MEDINFO 2007 IOS Press, ed K. Kuhn et al; pp. 834 - 839 (2007)

[2] Panzarasa S, Quaglini S, Cavallini A, Marcheselli S, Stefanelli M, Micieli G. Computerised Guidelines Implementation: Obtaining Feedback for Revision of Guidelines, Clinical Data Model and Data Flow. AIME 2007, LNAI 4594 Proceedings, ed R. Bellazzi, A. Abu-Hanna, and J. Hunter (Eds.): pp. 461 - 466 (2007)

[3] Razavi AR, Gill H, Ahlfeldt H, Shahsavar N. Non-compliance with a postmastectomy radiotherapy guideline: decision tree and cause analysis. BMC Med Inform Decis Mak. 2008 Sep 21;8:41.

[4] Demsar J, Zupan B. Orange: From Experimental Machine Learning to Interactive Data Mining. White Paper (www.ailab.si/orange), Faculty of Computer and Information Science, University of Ljubljana. (2004).

[5] CLOTS Trials Collaboration, Dennis M, Sandercock PA, Reid J, Graham C, Murray G, Venables G, Rudd A, Bowler G Effectiveness of thigh-length graduated compression stockings to reduce the risk of deep vein thrombosis after stroke (CLOTS trial 1): a multicentre, randomised controlled trial. Lancet. 2009 Jun 6;373(9679):1958-65. Epub 2009 May 26.

[6] Marcheselli S, Cavallini A, Tosi P, Quaglini S, Micieli G. Impaired blood pressure increase in acute cardioembolic stroke. J Hypertens. 2006 Sep;24(9).

[7] Ungar A, Pepe G, Lambertucci L, Fedeli A, Monami M, Mannucci E, Gabbani L, Masotti G, Marchionni N, Di Bari M. Low diastolic ambulatory blood pressure is associated with greater all-cause mortality in older patients with hypertension. J Am Geriatr Soc. 2009 Feb;57(2):291-6.

[8] Acelajado MC, Oparil S. Hypertension in the elderly. Clin Geriatr Med. 2009 Aug;25(3):391-412.

Address for correspondence:

Silvia Panzarasa, silvia.panzarasa@unipv.it, via Ferrata, 1 27100 Pavia, Italy, tel +390382985981, fax +390382985373