Using Data Mining to Explore Complex Clinical Decisions: A Study of Hospitalization After a Suicide Attempt

Enrique Baca-García, M.D.; M. Mercedes Perez-Rodriguez, M.D.; Ignacio Basurte-Villamor, M.D.; Jeronimo Saiz-Ruiz, M.D; José M. Leiva-Murillo, M.Sc.; Mario de Prado-Cumplido, M.Sc.; Ricardo Santiago-Mozos, M.Sc.; Antonio Artés-Rodríguez, Ph.D.; and Jose de Leon, M.D.

Background: Medical education is moving toward developing guidelines using the evidence-based approach; however, controlled data are missing for answering complex treatment decisions such as those made during suicide attempts. A new set of statistical techniques called data mining (or machine learning) is being used by different industries to explore complex databases and can be used to explore large clinical databases.

Method: The study goal was to reanalyze, using data mining techniques, a published study of which variables predicted psychiatrists' decisions to hospitalize in 509 suicide attempters over the age of 18 years who were assessed in the emergency department. Patients were recruited for the study between 1996 and 1998. Traditional multivariate statistics were compared with data mining techniques to determine variables predicting hospitalization.

Results: Five analyses done by psychiatric researchers using traditional statistical techniques classified 72% to 88% of patients correctly. The model developed by researchers with no psychiatric knowledge and employing data mining techniques used 5 variables (drug consumption during the attempt, relief that the attempt was not effective, lack of family support, being a housewife, and family history of suicide attempts) and classified 99% of patients correctly (99% sensitivity and 100% specificity).

Conclusions: This reanalysis of a published study fundamentally tries to make the point that these new multivariate techniques, called data mining, can be used to study large clinical databases in psychiatry. Data mining techniques may be used to explore important treatment questions and outcomes in large clinical databases and to help develop guidelines for problems where controlled data are difficult to obtain. New opportunities for good clinical research may be developed by using data mining analyses.

(J Clin Psychiatry 2006;67:1124–1132)

Received Aug. 3, 2005; accepted Nov. 28, 2005. *From the Department of Psychiatry, Fundación Jiménez Diaz, Universidad Autonoma de Madrid (Drs. Baca-García and Basurte-Villamor); the Hospital Ramon y Cajal, Universidad de Alcalá (Drs. Perez-Rodriguez and Saiz-Ruiz); and the Department of Signal Theory and Communications, Universidad Carlos III (Messrs. Leiva-Murillo, de Prado-Cumplido, Santiago-Mozos and Dr. Artés-Rodríguez), Madrid, Spain; and the Mental Health Research Center, Eastern State Hospital, Lexington, Ky. (Dr. de Leon).*

This study was conducted at the Hospital Ramon y Cajal, Madrid, Spain, and recruitment was supported by a National Alliance for Research on Schizophrenia and Affective Disorders (NARSAD) Young Investigator Award to Dr. Baca-García and a grant from the Spanish Health Ministry (FIS 01/0077-01). The statistical analyses described in this article were conducted without external support.

The authors report no additional financial or other relationship relevant to the subject of this article.

The authors thank Lorraine Maw, M.A., for editorial assistance.

Corresponding author and reprints: Jose de Leon, M.D., U.K. Mental Health Research Center, Eastern State Hospital, 627 West Fourth St., Lexington, KY 40508 (e-mail: jdeleon@uky.edu).

Since the beginning of the Hippocratic tradition in Greece 2500 years ago, medical education is begun by working under the supervision of previously trained physicians who have accumulated medical experience, and then is furthered by acquiring new knowledge through personal observation, by using rational thinking to interpret the complex clinical reality, and more recently (fundamentally in the last 500 years) by progressively incorporating the scientific method to help interpret one's observations. The astronomical accumulation of medical knowledge at the end of the twentieth century has led to a paradigm change; medical education has changed from the traditional model to evidence-based learning. In this new learning paradigm, a group of experts reviews the literature to develop guidelines for diagnosing and treating medical problems. Information from the literature is classified according to the quality of the evidence. Meta-analyses of well-controlled studies following the traditional experimental paradigm are given more value (Evidence Ia) while observational naturalistic case reports are considered less valuable (Evidence III) and opinions of experts receive the lowest valuation (Evidence IV).[1] The progressive implementation of evidence-based medicine in the last 10 to 15 years has changed the way of measuring quality outcomes among physicians. In the traditional

model, experienced physicians were the ones who provided the best medical care since they had accumulated wisdom through years of making mistakes and learning from them. In the evidence-based paradigm, younger physicians appear to be more up-to-date and provide the best medicine, according to more recent medical guidelines.[2] In the traditional medical system, experience was of value; in the new medical system, the ability to update yourself is a fundamental asset, so you are not left behind.

GUIDELINES USING EVIDENCE-BASED APPROACH

Clinical Guidelines

The use of clinical guidelines is a major development that should be welcomed, but it is naive to think that the guidelines are free of problems. Three major problems are that (1) the results of well-controlled studies using the experimental method do not always translate well to the complex clinical world (efficacy vs. effectiveness); (2) there are no well-controlled clinical experiments that can be used to set guidelines for many of the most important clinical decisions; and (3) the experts developing the guidelines may have major philosophical and experiential differences with the busy clinicians trying to implement them.

Guidelines for Hospitalization After a Suicide Attempt

Suicide is a fundamental area of clinical practice in psychiatry, but the available published information that can be reviewed for developing evidence-based guidelines is somewhat limited. Suicide's complex nature and the ethical issues involved make it difficult to conduct "clinical" experiments. Many guidelines have been designed for suicide assessment, but none are universally accepted.[3–23] One of the crucial clinical decisions after a suicide attempt is to determine whether the patient needs hospitalization or not. Most guidelines recommend that direct discharge from the emergency room should only be considered if a psychosocial assessment and aftercare plan can be arranged before discharge. Hospitalization is recommended when there is an imminent risk of suicide (Table 1).

It is not likely that, in the near future, well-controlled studies using the experimental method will illuminate the decision about hospitalization after a suicide attempt. There are 2 major reasons: ethical (it would be hard to justify experimental designs when the patient's life is in danger) and also methodological. The typical experimental approach works by designing an experiment that tests for a unique hypothesis by controlling any other confounding variables and using significance testing to rule out hypothesis results that may be explained by chance. As the risk of completed suicide is probably determined by many clinical variables, isolating only one clinical variable and testing it with an experimental design may not be terribly helpful. A 30-year-old article[24] that is probably the first attempt to develop guidelines for hospitalization after a suicide attempt illustrates the problem of developing guidance in this area. The authors consulted a committee of U.S. academic experts and asked them to develop guidelines for hospitalization, and then they tested the guidelines with 248 suicide attempters at a U.S. hospital. According to the expert guidelines, all 248 attempters should have been hospitalized, but 127 attempters (51%) were discharged by the psychiatry residents.[24] If one looks at these guidelines from the point of view of the clinician, one would think that the experts were the ones that made the mistake, erring too much on the side of safety.

Table 1. Review of Suicide Attempt Guidelines: Recommendations Regarding Hospitalization

Suicide Attempt Guideline
Summary by authors of all prior guidelines
Ominous signs that should always lead to hospitalization (1 is enough)[a]
Deep despair
Command hallucinations telling patient to kill him/herself
Definitive criteria for hospitalization (at least 1)[b]
Need for medical treatment after attempt
Need for psychiatric treatment of associated psychiatric disorder
Lack of social support
Recommended criteria for hospitalization (at least 1)[c]
Repeated suicide attempts in addition to a plan and a means to complete the plan
A psychiatric disorder and a recent crisis
A plan, access to a lethal means, recent social stressors, and some evidence of a psychiatric disorder
Insufficient evidence to recommend a specific clinical intervention[d]
American Psychiatric Association guidelines[e]
Admission is generally indicated after a suicide attempt or aborted suicide attempt if:
Patient is psychotic
Attempt was violent, near-lethal, or premeditated
Precautions were taken to avoid rescue or discovery
Persistent plan and/or intent is present
Distress is increased or patient regrets surviving
Patient is male and older than 45 years (especially with new onset of psychiatric illness or suicidal thinking)
Patient has limited family and/or social support (including lack of stable living situation)
Current impulsive behavior, severe agitation, poor judgment, or refusal of help is evident
Patient has change in mental status with a metabolic, toxic, infectious, or other etiology requiring further workup in a structured setting

[a]Based on Hirschfeld,[12,14] Hirschfeld and Russell,[13] Evidence Based Medicine Guidelines,[15] and Hamilton.[16]
[b]Based on Hider,[9] Hamilton,[16] Nicholas and Golden,[20] and Shaffer and Pfeffer.[21]
[c]Based on Hider,[9] Gliatto and Rai,[11] and World Health Organization.[17]
[d]Based on NHS Center for Reviews and Dissemination.[8]
[e]Published by American Psychiatric Association.[23]

Using Traditional Statistics to Explore Psychiatrists' Decisions to Hospitalize in the Real World

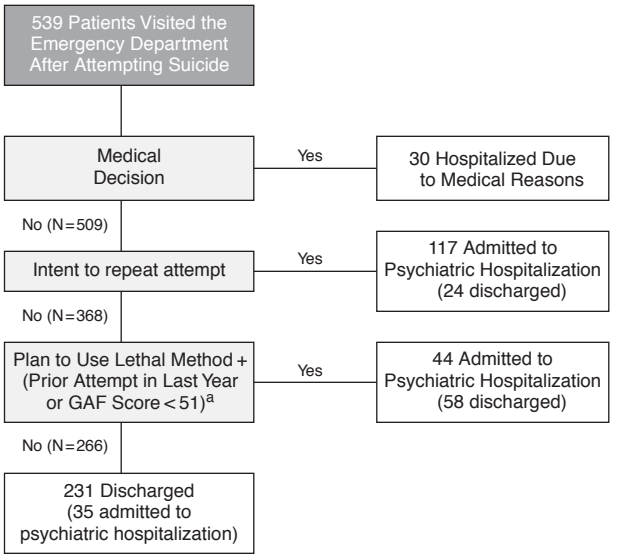
A clinician may defend another approach before developing any expert guidelines. One needs to know which variables are influencing clinician judgment in hospitaliz-

Table 2. Classification of Subjects Using Traditional Statistics and Data Mining

Method	Subjects Correctly Classified, %	Sensitivity: True Positive Rate, %	Specificity: True Negative Rate, %
Traditional statistics			
Univariate analysis[a]: 1 clinical variable (intent to repeat attempt)	88	63	92
Logistic regression[b]: 11 clinical variables (11 significant from a total of 46 variables)	83	81	91
Decision tree[c]: 3 clinical variables (3 selected from 11 significant variables)	77	82	27
Logistic regression[d]: 8 SIS items (8 significant from a total of 15 SIS items)	74	85	59
ROC analysis[e]: SIS total score	72	69	70
Data mining			
Forward selection: 5 first variables	99	99	100

[a]First, the association of the dependent variable, hospitalization (yes vs. no, or hospitalization vs. discharge), with any of the 47 clinical variables was initially explored by means of 2-way cross tabulations for univariate analyses. This provided 32 significant variables with p values < .05. Odds ratios (ORs) were used as a measure of the strength of the association and their 95% confidence intervals (CIs) were computed. The variable "intent to repeat attempt" had high specificity.
[b]Significant variables were then included as independent variables in a multivariate logistic regression that used hospitalization as the dichotomous dependent variable. The logistic regression provided adjusted ORs for the independent variables and reduced the significant variables to 11.
[c]The second model followed the idea of prior algorithms developed to assess imminent suicide risk[12–14]; a psychiatric decision tree (a simpler version of an algorithm) was developed using only the 3 most important variables of the logistic regression model (Figure 1).
[d]The third model did not provide better results. This model included the 15 items on the suicide scale in a logistic regression.
[e]The fourth model did not provide better results. This model included the total scale score in a receiver operating characteristic (ROC) analysis considering the best equilibrium between sensitivity and specificity.
Abbreviation: SIS = Beck's Suicidal Intent Scale.

Figure 1. Decision Tree to Explain Hospitalization After a Suicide Attempt



[a]Patients met 2 characteristics: plan to use lethal method and 1 of 2 variables (prior attempt during the last year or GAF score < 51).
Abbreviation: GAF = Global Assessment of Functioning.

ing patients after a suicide attempt and to consider these variables when developing new guidelines. We took that approach in one study that included a large sample of 509 patients assessed by psychiatrists, who used an interview including 47 clinical variables and a 15-item suicide scale.[25] Two research psychiatrists (E.B.G. and J.dL.), who work as clinicians, have research training, and are familiar with statistics, developed statistical models to explore which variables predicted hospitalization.

One variable was considered to be clinically relevant, and 4 statistical models were developed (Table 2 and Figure 1). First, the association of the dependent variable, hospitalization (yes vs. no, or hospitalization vs. discharge), with any of the 47 clinical variables was initially explored by means of 2-way cross tabulations for univariate analyses. Odds ratios (ORs) were used as a measure of the strength of the association and their 95% confidence intervals (CIs) were computed. Significant variables were then included as independent variables in a multivariate logistic regression[26] that used hospitalization as the dichotomous dependent variable (Table 2). The second model followed the idea of prior algorithms developed to assess imminent suicide risk[12–14]; it used a psychiatric decision tree (a simpler version of an algorithm) (Figure 1). The third model included the 15 items on the suicide scale in a logistic regression (Table 2). The fourth model included the total score of the suicide scale in a statistical analysis, considering the best equilibrium between sensitivity and specificity (Table 2).

A Review of Data Mining

Clinicians may not be familiar with data mining, a group of techniques that are not yet being used in psychiatric literature. Other equivalent or related terms are "machine learning" or "knowledge discovery in databases."[27] In spite of their lack of awareness, clinicians need to know that their personal and family data are subject to data mining since many industries rely on data mining. The list of industries includes marketing, manufacturing,

database providers, the travel industry, banking, telecommunications, engineering, and even the government.[28] The common theme of these industries is that they have massive amounts of information collected in different ways. In order to maximize the usefulness of the information, they use software to search for patterns and trends. The goal is to reduce complexity and extract or "mine" as much relevant and useful information as possible.

Thus the term *data mining* was adopted by computer scientists in the early 1990s and can be defined in a technical way such as "algorithmic and database-oriented methods that search for previously unsuspected structure and patterns in data"[29(p309)] or "the science of searching large bodies of data seeking interesting patterns and structures."[30(p305)] Thus data mining is "a child of the computer age,"[30] a direct consequence of the popularization of computers and the storage of large bodies of data.[30] Data mining is an eclectic discipline that uses ideas and tools from many disciplines. There are some reviews commenting on the potential of data mining for the biopharmaceutical industry[28] or nursing[27] but very limited information related to medicine. Two major applications for data mining in medicine are the mining of the medical literature and of patient databases.

Data mining techniques are based on solid statistical principles.[31,32] However, its approach to a particular problem is different from traditional statistics. While traditional statistics emphasizes inference and estimations, data mining emphasizes the fulfillment of a task such as classification, estimation, or knowledge discovery. Data mining is almost always done in a retrospective way, using data from naturalistic observations, and does not involve consideration of experimental design. As it happens in this study, the researchers conducting data mining may not be familiar with the meaning of the data. The fact that somebody with no medical knowledge looks for patterns in patient databases may be somewhat "unsettling" for clinicians, but the data mining experts would point out that their lack of knowledge minimizes the biases, while the clinician researcher might unconsciously have decided what results he/she expects to obtain. The machine learning multivariate methods are totally blind and unbiased.

Traditional statistics distinguishes between testing hypotheses and generating hypotheses. Hypothesis-generating methods, or exploratory data analyses, have also been used by traditional statisticians.[33] Data mining can be seen as a hypothesis-generating method, but the introduction of statistical learning makes the approach different from traditional exploratory data analyses. Moreover, there is a completely different philosophical approach within the framework of interpreting scientific exploration.

The goal of the study was to determine which variables predicted psychiatrists' decisions to hospitalize patients in a large sample of suicide attempters who were assessed in the emergency department of a Spanish general hospital. The results, using traditional statistics performed by clinician researchers, have already been published.[25] This time, data mining techniques were applied to reanalyze the data, showing substantially different results.

METHOD

Brief Sample Description

The original study was approved by the hospital's Institutional Review Board and is thoroughly described above. It was conducted in a general hospital within the Spanish National Health System that provides medical coverage for all emergencies in a catchment area of 500,000 people in Madrid, Spain.[25] Suicide attempts were defined according to the definition given by the National Institute of Mental Health.[34] The study sample, recruited between 1996 and 1998, comprised 539 people over the age of 18 years (mean ± SD age was 35.9 ± 14.4 years) who had attempted suicide.[25] The structured patient assessment included clinical chart diagnoses; 47 clinical variables obtained from the literature; and a suicide scale, Beck's Suicidal Intent Scale (SIS).[35]

Thirty of the 539 attempters were initially hospitalized in medical units before psychiatrists became involved. The remaining 509 patients were included in the statistical analysis and were divided in a dichotomous way according to the psychiatrists' decisions: 196 (39%) were hospitalized in the psychiatric unit and 313 (61%) were discharged from the emergency room department.[25]

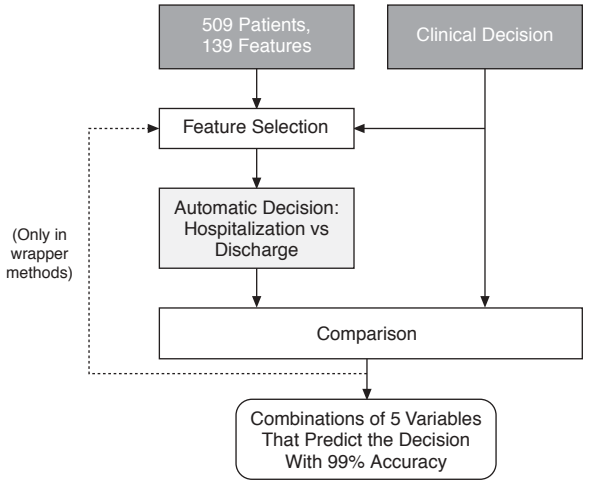
Data Mining

The method that followed in applying machine learning techniques to this study can be defined as building an artificial classification problem. The variable to be classified, or learned, is the decision of the clinician of whether or not patients who have attempted suicide were hospitalized. The method used in this task attempts to determine the most relevant variables in the classification task, i.e., those that predict classification with the greatest accuracy (Figure 2).

Feature selection methods[31,32,36,37] were used to rank the different variables according to their relevance to an artificial classification performance problem that emulates the psychiatrist's decision-making. This classification performance was conducted by an automatic decision system: in summary, a computerized system.

There are 2 types of feature selection methods: filter and wrapper methods.[31,32,36,37] The filter methods use a discrimination criterion to rank the variables according to some relevant criterion, taken individually for each of the features. These methodologies do not take into account the possible collinearity among variables. Thus, the ranking is carried out without using information about

Figure 2. Graphic Representation of the Data Mining Method Used in This Study[a]



[a]There were 139 variables (called "features" in data mining) from 509 patients. The feature selection methods, using the information from the clinical decision process, were used to select the individual variables that separated patients who were admitted versus hospitalized (the outcome [hospitalization versus discharge] is what is called "automatic decision" in data mining). The results of the feature selection methods were compared using the wrapper methods with the clinical decision results in a repeated and iterative way, providing the final combinations using 5 variables and predicting the clinical decision with 99% accuracy.

the classification performance of the variable. Four filter methods were used: Fisher's Discriminant, the Kolmogorov-Smirnov test, the Mutual Information Discriminant, and Recursive Feature Elimination (Table 3). While Fisher's Discriminant is a linear method, the other 3 methods try to capture nonlinear associations present in the data.

The wrapper methods use the classification performance of the variables as criteria to choose the most discriminative ones. Consequently, variables were selected by evaluating their incremental information relevant to the classification task, thus collinearity's disturbing effects are softened. Two wrapper methods were used: Random Forest and Forward Selection (see Table 3). Random Forest is based on the statistical processing of a large set of decision trees. The Forward Selection method is based on an incremental search that makes use of a support vector machine to carry out the classification.

The classification performance, using a cross-validation procedure, followed several steps. First, the dataset is divided into a set of 3 folds. Then 3 classification experiments are made; the accuracy in each fold is evaluated by training with the other 2. Finally, the average performance of the 3 experiments is considered the expected performance. In this way, it is possible to assure the generalization ability of the classifier, avoiding overfitting to the data sample under consideration.[32]

First, each of the 4 filter methods was used to establish the top 20 variables that predicted hospitalization versus discharge. Table 4 describes only the top 10 of the 20 variables selected by each filter method. Second, the 2 wrapper methods were used to rank the 20 variables selected by the filter methods according to their ability to classify patients as hospitalized. The Forward Selection method provided an accumulative measurement of accuracy of classification using the database (the accuracy with the first variable was 96.5%; with each additional variable added, the accuracy got closer to 100%). Random Forest, based in resampling techniques, provided the variable importance as well as the identification of the outliers and the correlation between the variables (data not presented). Table 5 describes the top 10 variables selected by the 2 wrapper methods, compared with the top 10 variables (out of a total of 11) selected by the logistic regression model that was previously published.[25]

RESULTS

Filter Methods

The 4 filter methods showed similar most important variables associated with hospitalization (Table 4). Consuming drugs during the attempt was the first variable for 3 methods and fourth for the other. Lack of family support was first in 1, second in another, and fourth in 2 others. Consuming alcohol during the attempt was the second variable for 2 methods, fourth for another, and sixth for another. Intent to repeat the attempt was third for 2 methods and fifth for the other 2 methods.

Wrapper Methods

The 2 wrapper methods showed a common top variable: consuming drugs during the attempt (Table 5). To select the final model, sensitivity and specificity were reviewed. Different combinations of the top 5 variables of the Forward Selection method (drug consumption during the attempt, relief that the attempt was not effective, lack of family support, being a housewife, and family history of suicide attempts) reached an almost perfect discriminative power with 99% of the subjects correctly classified. These results are remarkably better than the attempts made by research psychiatrists using traditional statistics (Table 2).

DISCUSSION

According to the data mining methods, the main variables associated with the clinician's decision to hospitalize a patient after a suicide attempt are related to drug or alcohol consumption during the attempt, lack of family support, and attitude toward the attempt (criticism and intention to repeat). When using traditional statistical methods, drug or alcohol consumption during the attempt were

Table 3. Six Feature Selection Methods Used for Data Mining in This Study

Method	Description
Filter methods	These methods try to find the most relevant variables for separating the samples according to another variable or class. In this case, the class will be hospitalization or not.
Fisher's Discriminant	A ranking mechanism based on linear separation criteria. Looks for the maximum of the ratio between the distance between the means μ_0^i and μ_1^i of the classes (0 and 1) and the sum of the variances (σ_0^i and σ_1^i).[a] For each i feature this ratio is computed and then used to set the ranking.[b]
Kolmogorov-Smirnov Test	Based on the greatest distance between the probability distribution functions of 1 feature including both classes and the same feature of the positive class.[c] The probability distribution is obtained by means of the empirical distribution function, e.g., the one that assigns the same probability for each sample.[b]
Mutual Information Discriminant	Ranks the variables depending on their relevance by measuring Mutual Information between each of the variables and the class to which they belong. According to Shannon's Information Theory, Mutual Information ($I(X_i,Y)$) is a natural way of measuring the degree of knowledge of a variable (in this case the class X_i and Y) that can be obtained from another variable.[d]
Recursive Feature Elimination	Consists in the iterative elimination of the variable that conveys the smallest variation in the margin obtained by a support vector machine. The features are then sorted according to their importance, but no objective ranking is obtained. This method cannot be considered as a pure filter method since a support vector machine classifier needs to be trained.[e] Recursive Feature Elimination is able to find nonlinear associations in the data when an appropriate kernel function is used in the support vector machine.
Wrapper methods	These methods use the classification performance of the variables as a criterion to choose the most discriminative variables.
Random Forest	Builds an ensemble of weak classifiers using resampling techniques. Normally, these classifiers are simple decision trees, although any classifier can be used. Random Forest selects random sets of variables to train the classifiers using a new sample set selected randomly from the original training set. In this way, Random Forest assigns an importance to each feature based on its usefulness for separating the classes in the whole set of classifiers.[f]
Forward Selection	Consists in the iterative addition of features as long as they improve the classification performance. At step 1, the variable that provides the best cross-validation accuracy of the support vector machine classifier is chosen. At step 2, any new variable that, together with already chosen variables, provides the best accuracy is selected. This procedure is repeated until the desired number of variables is reached. Thus, the set of variables is built by a growing procedure and ranked according to their order of appearance.[g]

[a]Formula: $\rho(i) = \dfrac{(\mu_1^i - \mu_0^i)^2}{(\sigma_1^i)^2 + (\sigma_0^i)^2}$

[b]Based on Fukunaga.[37]

[c]Formula: $KS(i) = \sqrt{L}\sup(\hat{P}\{X_i \le x\} - \hat{P}\{X_i \le x, Y = 1\})$

[d]Formula: $I(X_i,Y) = H(X_i) - H(X_i|Y) = H(Y) - H(Y|X_i)$

$I(X_i,Y) = h(X_i) - h(X_i|Y) = h(Y) - h(Y|X_i)$

where $H(\cdot)$ is the entropia and $h(\cdot)$ is the differential entropia. $H(X) = -\sum_{x\in X} p_X(x)\log p_X(x)$

$$h(X) = -\int p_X(x)\log p_X(x)\, dx$$

The conditional entropia $H(X_i|Y)$ in the first formula is: $H(X_i|Y) = \sum_{k=1}^{N_c} p_Y(Y = c_k)H(X_i|Y = c_k)$

$$= -\sum_{k=1}^{N_c}\sum_{x\in X_i} p_{X_i,Y}(x, Y = c_k)\log p_{X_i|Y}(x|Y = c_k)$$

where N_c is the number of classes.

[e]Based on Vapnik.[32]

[f]Based on Breiman.[31]

[g]Based on Guyon and Elisseeff.[36]

not considered since they were not significant in the univariate analyses. The OR explored for the combination of alcohol or drug consumption during the attempt was 1.3 (95% CI = 0.84 to 1.1, p = .26). The individual OR (OR for alcohol consumption during the attempt and for drug consumption during the attempt) was not explored in the first analysis using traditional methods but provided no better results regarding significance.

Once the database was cleaned by a psychiatric researcher, researchers with no psychiatric knowledge used data mining techniques to provide a better way of classifying subjects (in this case, hospitalized suicide attempters). The percentage of subjects correctly classified by traditional statistical methods was between 72% and 88%, with a best result of 88% using only 1 variable. Data mining methods combining the top 5 variables had a 99% classificatory accuracy. This marked improvement in classification suggests that psychiatric researchers and clinicians need to pay attention to these new multivariate techniques called data mining.

Another important message is that large complex clinical databases can be explored using data mining

Table 4. Description of the Top 10 Variables Taken From the Top 20 Obtained by 4 Filter Methods

Rank	Fisher's Discriminant	Kolmogorov-Smirnov Test	Mutual Information Discriminant	Recursive Feature Elimination
1	Consuming drugs during attempt	Consuming drugs during attempt	Consuming drugs during attempt	Lack of family support
2	Lack of family support	Consuming alcohol during attempt	Consuming alcohol during attempt	Estimated patient reliability[a]
3	Critique[b] of attempt[c]	Intent to repeat attempt	Intent to repeat attempt	Expect to be found and helped[c]
4	Consuming alcohol during attempt	Lack of family support	Lack of family support	Consuming drugs during attempt
5	Intent to repeat attempt	Realistic perspectives on the future[c]	Attitude toward living/dying[d]	Intent to repeat attempt
6	Realistic perspectives on the future[c]	Attitude toward living/dying[d]	Critique[b] of attempt[c]	Consuming alcohol during attempt
7	Attitude toward living/dying[d]	Critique[b] of attempt[c]	Reaction to attempt[a]	GAF score < 51 before attempt
8	Relief attempt was not effective[c]	Conception of medical rescuability[d]	Conception of medical rescuability[d]	Ever being employed[c]
9	Seriousness of attempt[d]	Relief attempt was not effective[c]	Seriousness of attempt[d]	Using firearms for attempt
10	Clinician's confidence[a,e]	Reaction to attempt[a]	GAF divided in 3 levels	Childhood psychiatric disorders[f]

[a]Additional item included in the Suicide Intent Scale (SIS) but not used for total score.
[b]Critique of performing attempt (wrong solution) or of the motive (wrong motive).
[c]This variable was associated with discharge. The absence of this variable was associated with hospitalization.
[d]One of 15 items of the SIS.
[e]Clinician's confidence in his or her inference about objective SIS questions.
[f]History of any psychiatric disorder usually first diagnosed in infancy, childhood, or adolescence.
Abbreviation: GAF = Global Assessment of Functioning.

Table 5. Description of the Top 10 Variables Taken From the 20 Obtained by 2 Wrapper Methods and Comparison With the Top 10 Variables Identified by Logistic Regression

	Forward Selection		Random Forest		Published Logistic Regression[c]	
Rank	Variable	Accumulative Accuracy, %[a]	Variable	Importance[b]	Variable	OR[d]
1	Consuming drugs during attempt	96.5	Consuming drugs during attempt	72.8	Intent to repeat attempt	14.7
2	Relief that attempt was not effective[e]	97.8	Consuming alcohol during attempt	47.1	Plan to use a lethal method	8.3
3	Lack of family support	98.7	Lack of family support	34.7	Previous psychiatric hospitalization	3.5
4	Being a housewife[e]	99.1	Treatment for medical problems	18.7	GAF score < 51 before attempt	3.5
5	Family history of suicide attempts	99.3	Estimated patient reliability[e,g]	17.1	Suicide attempt during past year	3.4
6	Overt communication of intent[e,f]	99.1	Intent to repeat attempt	16.3	Realistic perspectives on the future[e]	0.30
7	History of depression	98.9	Assessed by a psychiatry resident	15.4	Relief attempt was not effective[e]	0.33
8	GAF score < 51 before attempt	98.7	Critique[h] of attempt[e]	14.2	Availability of a method to kill oneself[e]	0.36
9	Precaution against discovery/intervention[f]	99.1	Visualization of death[g]	14.2	Planning that nobody will try to save them	2.6
10	Being a student[e]	98.9	Using a lethal method	12.2	Belief that attempt will influence others[e]	0.45

[a]The forward selection method provides an accumulative measurement of accuracy of classification using the database (the accuracy with the first variable was 96.5%, and with each additional variable added the accuracy got closer to 100%).
[b]Random Forest, based on resampling techniques, provided the variable's importance (maximum 100 and minimum 0). The importance is a normalized (in respect to its typical deviation) value.
[c]Data from Baca-García et al.[25]
[d]Odds ratios (ORs) are in order according to importance. An OR < 1 indicates decrease of risk. The absence of that variable is associated with inverse OR. An OR = 0.30 for realistic perspective on the future after the attempt corresponds to an OR = 3.3 for lacking realistic perspective on the future after the attempt. The inverse of an OR = 0.33 is 3.0, of an OR = 0.36 is 2.8, and of an OR = 0.45 is 2.2.
[e]This variable was associated with discharge. The absence of this variable was associated with hospitalization.
[f]One of 15 items of the Suicide Intent Scale.
[g]Additional item included in the Suicide Intent Scale but not used for total score.
[h]Critique of performing attempt (wrong solution) or of the motive (wrong motive).
Abbreviation: GAF = Global Assessment of Functioning.

techniques to answer important clinical questions that evidence-based guidelines cannot answer properly since there are limited controlled data. Clinicians need to begin to consider clinician-friendly ways of recording data using simple and reliable computerized assessment of their patients so that large databases can be built. These large databases can be studied using data mining techniques to explore effectiveness of treatment. Similarly, clinical trial databases can be studied with data mining to provide other clinical answers besides the original question that led to the clinical trial. These may be ways of approximating efficacy, effectiveness, and clinical practice and improving evidence-based guidelines. Complex clinical problems such as suicide, where evidence-based data are difficult to obtain, are the ones that will benefit more from studying good naturalistic clinical databases.

Limitations

Individual clinical decisions for each patient were not studied, but statistical analyses (first, traditional multivariate methods and, in this second article, data mining methods) were used to indirectly explore which clinical variables appear to better explain the clinical decision; it would be the equivalent of using statistical models that mimic the "ideal" clinician behind all individual physicians involved in the treatment of those patients.

Some may consider that reanalyzing previously published data is a limitation. However, in our opinion this is the only way to test and compare 2 different ways of using statistical methods to study a complex clinical problem. Moreover, the main goal of this article is not to demonstrate that the second set of analyses (data mining) is more valid than the first set of analyses. To prove that the second set of analyses are valid, one needs to replicate them, but different hospitals and physicians may produce different results. This is how clinical decisions work in the real world; clinical decisions do not follow the approach of experimental studies. The accumulation of information concerning how hospitalization decisions are currently made in the clinical world may help to shape guidelines in an area where more controlled data are missing and are unlikely to be obtained in the future. Additionally, the main goal of this article is not to defend the validity of the results but to use them as a way to present to clinicians a new exploratory data technique.

Similarly, one can argue that analyzing data by using data mining techniques without psychiatric knowledge is a limitation. However, this may also be considered an advantage, since it minimizes the biases that the researcher might unconsciously cause because he/she has previously decided what results he/she expects to obtain. The multivariate methods described above are totally blind and unbiased. More importantly, data mining techniques provided a better way of understanding the statistical nature of the clinical decision than those based in traditional statistics.

In summary, this reanalysis of a study of hospitalization after a suicide attempt fundamentally tries to make the point that a new set of multivariate techniques, called data mining, can be used to study large clinical databases in psychiatry. Data mining techniques may be used to explore important treatment questions and outcomes in large clinical databases and to help develop guidelines for problems where controlled data are difficult to obtain. The potential of data mining is highlighted by the better classification of subjects obtained by these methods versus that obtained by traditional statistics performed by psychiatric researchers familiar with traditional multivariate methods. Psychiatric researchers and clinicians may need to get familiar with data mining and consider collaborating with experts familiar with that set of techniques. Large detailed clinical databases may have new value as potential material for machine learning researchers. New opportunities for good clinical research may be presented by using data mining analyses' computer methods to emulate clinicians as they deal with complex problems using complex sets of variables, as clinicians usually do in their daily practices.

REFERENCES

1. Oosterhuis WP, Bruns DE, Watine J, et al. Evidence-based guidelines in laboratory medicine: principles and methods. Clin Chem 2004;50: 806–818
2. Choudry NT, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. Ann Intern Med 2005;142:260–273
3. The General Hospital Management of Adult Deliberate Self-Harm: A Consensus Statement on Standards for Service Provision. London, England: Royal College of Psychiatrists; 1992
4. Gunnell D, Frankel S. Prevention of suicide: aspirations and evidence. BMJ 1994;308:1227–1233
5. Kachur SP, DiGuiseppi C. Screening for Suicide Risk. In: Guide to Clinical Preventive Services, Second Edition, Mental Disorders and Substance Abuse. Washington, DC: US Dept Health Human Services, Office of Disease Prevention and Health Promotion; 1996. Available at: http://cpmcnet.columbia.edu/texts/gcps/gcps0060.html. Accessibility verified April 4, 2006
6. McNamee JE, Offord DR. Prevention of suicide. In: Canadian Task Force on the Periodic Health Examination. Canadian Guide to Clinical Preventive Health Care. Ottawa, Canada: Health Canada; 1994:456–467. Available at: http://www.ctfphc.org/Tables/Ch40tab.htm. Accessibility verified April 4, 2006
7. Suicide Risk Advisory Committee of the Risk Management Foundation of the Harvard Medical Institutions. Guidelines for Identification, Assessment, and Treatment Planning for Suicidality; 1996. Available at: http://www.rmf.harvard.edu/files/documents/suicideAs.pdf. Accessibility verified April 4, 2006
8. NHS Center for Reviews and Dissemination, University of York. Deliberate self-harm decision makers. Effective Health Care 1998;4:1–12
9. Hider P. Youth Suicide Prevention by Primary Healthcare Professionals: A Critical Appraisal of the Literature. NZHTA REPORT 4. 1998. Available at: http://nzhta.chmeds.ac.nz/publications/nzhta4.pdf. Accessibility verified April 4, 2006
10. The Royal New Zealand College of General Practitioners. Guidelines for Primary Care Providers: Detection and Management of Young People at Risk of Suicide. 1998. Available at: http://www.nzgg.org.nz/guidelines/dsp_guideline_popup.cfm?guidelineCatID=32&guidelineID=29. Accessibility verified April 4, 2006
11. Gliatto MF, Rai AK. Evaluation and treatment of patients with suicidal ideation. Am Fam Physician 1999;59:1500–1506
12. Hirschfeld RMA. Algorithm for the evaluation and treatment of suicidal

patients. Primary Psychiatry 1996;3:26–29
13. Hirschfeld RMA, Russell JM. Assessment and treatment of suicidal patients. N Engl J Med 1997;337:910–915
14. Hirschfeld RMA. The suicidal patient. Hosp Pract (Minneap) 1998;33: 119–123,127–128, 131–133
15. Evidence Based Medicine Guidelines: Suicide. 2000. Available at: http://www.terveysportti.fi/pls/ebmg. Accessibility verified May 26, 2006
16. Hamilton NG. Suicide prevention in primary care: careful questioning, prompt treatment can save lives. Postgrad Med 2000;108:81–87
17. Mental and Behavioural Disorders. Department of Mental Health. World Health Organization. Preventing Suicide: A Resource for General Physicians. Geneva, Switzerland: World Health Organization; 2000
18. Magellan Behavioral Health Clinical Practice Guideline. Clinical Practice Guideline for Assessing and Managing the Suicidal Patient; 2000. Available at: https://www.magellanprovider.com/providing_care/clinical_guidelines/clin_prac_guidelines/2005suicide.pdf. Accessibility verified April 4, 2006
19. Isacsson G, Rich CL. Management of patients who deliberately harm themselves. BMJ 2001;322:213–215
20. Nicholas LM, Golden RN. Managing the suicidal patient. Clin Cornerstone 2001;3:47–57
21. Shaffer D, Pfeffer CR. Practice parameters for the assessment and treatment of children and adolescents with suicidal behavior. J Am Acad Child Adolesc Psychiatry 2001;40:S24–S51
22. Rihmer Z, Belsö N, Kiss K. Strategies for suicide prevention. Curr Opin Psychiatry 2002;15:83–87
23. American Psychiatric Association. Practice guideline for the assessment and treatment of patients with suicidal behaviors. Am J Psychiatry 2003; 160(suppl 11):1–60
24. Kirstein L, Prusoff B, Weissman M, et al. Utilization review of treatment of suicide attempters. Am J Psychiatry 1975;132:22–27
25. Baca-García E, Diaz-Sastre C, García Resa E, et al. Variables associated with hospitalization decisions by emergency psychiatrists after a patient's suicide attempt. Psychiatr Serv 2004;55:792–797
26. Hosmer DW, Lemeshow S. Applied Logistic Regression. 2nd ed. New York, NY: John Wiley & Sons Inc; 2000
27. Goodwin L, VanDyne M, Lin S, et al. Data mining issues and opportunities for building nursing knowledge. J Biomed Inform 2003;36: 379–388
28. Data mining. Nat Biotechnol 2000;18(suppl):IT35–IT36
29. Smyth P. Data mining: data analysis on a grand scale? Stat Methods Med Res 2000;9:309–327
30. Hand DJ. Mining medical data. Stat Methods Med Res 2000;9:305–307
31. Breiman L. Random forest. Machine Learning 2001;45:5–32
32. Vapnik VN. Statistical Learning Theory. New York, NY: John Wiley & Sons; 1998
33. Tukey J. Exploratory Data Analysis. Reading, Mass: Addison-Wesley; 1977
34. O'Carroll PW, Berman AL, Maris RW, et al. Beyond the Tower of Babel: a nomenclature for suicidology. Suicide Life Threat Behavior 1996;26:237–252
35. Beck AT, Schuyler D, Herman I. Development of suicidal intent scales. In: Beck AT, ed. Prediction of Suicide. Philadelphia, Pa: Charles Press; 1974:45–56
36. Guyon I, Elisseeff A. An introduction to variable and feature selection. J Mach Learn Res 2003;3:1157–1182
37. Fukunaga F. Introduction to Statistical Pattern Recognition. New York, NY: Academic Press; 1990

