J. Gouvernet

Service de l'Information Médicale,
Hôpital Timone Adultes,
Marseille, France

Synopsis

Decision Support Systems and Knowledge Processing

Four articles in this section deal with decision support in particular areas [1-4]: three focus on the crucial problem of prognosis, and one on dosage adjustment. The two other articles are in the field of knowledge processing and attempt to provide general solutions to institutional problems [5,6]: one apprach is the use of guidelines, and the other is based on the formalization of a patient problem list. We can roughly split these articles in those dealing with individual patient care management i.e. the first four, and those dealing with health-care management i.e. the latter two papers.

The paper by Smit et al. [1] presents two models of anticoagulation response to heparin regimen in dialyzed patients. The heparin dose must be sufficient to prevent blood clotting, while avoiding bleeding. The response to a standard dose varies from patient to patient, making adjustment to the individual patient a necessity. The classic pharmacodynamic approach provides good results but requires online measures of clotting time. The authors choose a population pharmacokinetics approach [7] based on routinely available data such as age, height, weight, gender, basal clotting time and dialyzer characteristics. They develop two models: the first by application of a non-linear mixed model with one compartment (NONMEM), the second using a neural network with hyperbolic tangent activation function and linear output function. The two models are achieved by forward selection of covariates using a training set and their performances are compared on a testing set. The two models handle roughly the same covariates. The results on the training set show a better prediction by the neural network essentially in the response to the bolus before the onset of dialysis; also there are more outliers with the NONMEM model. However, the NONMEM model provides estimations of clearance and volume of distribution which are explicit pharmacokinetic parameters and useful when the authors try to explain the decrease of clotting time during the first 15 minutes of dialysis. This paper offers a solution to an interesting and real problem.

In the second paper Hayden et al. [2] try to find a prediction of the presence or absence of cirrhosis in patients with chronic hepatitis C. The motivations are: prevalence of cirrhosis is about 20% after 20 years of infection, cirrhosis heavily changes prognosis and response to treatment, the only reliable procedure to assess presence of cirrhosis is liver biopsy which can miss cirrhosis and is associated with morbidity. The authors use artificial neural networks (ANN) to predict cirrhosis with 12 routine clinical and virological factors. They use a conventional ANN with 22 hidden neurons and an ANN with two types of transfer functions in a hidden layer of 24 neurons (Ward type ANN), in the two models there are 24 inputs. Their results are successful: sensibility and specificity are about 92% and 99%, respectively, with Ward type ANN. However, the size of the training set is 82, so the model can be overfitted. The authors compare the performances of ANN with a logistic regression and find that this latter method works poorly. However, logistic regression seems underspecified, in particular, the authors do not seek treshold values for continuous variables, do not enter interaction terms and retain only those covariates which reach significance, a criteria different from the one used with ANN. This may explain the absence in the logistic equation of age, albumin level and duration of infection, variables that appear to have heavy weights in ANN. Moreover, performances of logistic regression are probably evaluated with an a priori fixed cutpoint of 0.5.

The paper by Do et al. [3] explores to what extent computerization can help primary care physicians in the detection of severe angiographic coronary artery disease. The study is retrospective and includes 2385 hospitalized male patients all undergoing coronary angiography.

The authors show that computerized measurements of exercise tests have equivalent performances to visual interpretation by a cardiologist. A prediction by a logistic regression in-

cluding exercise test data and clinical data does not improve the classification obtained with the measurement of the ST-segment alone. The authors propose the use of computers by primary care physicians to both measure the exercise test ECG and calculate the prediction equation. Then they classify the patients in low, intermediate and high risk, and address for subsequent investigations the intermediate class. This strategy shows on their sample a sensitivity of 85% and a specificity of 86%. The authors carefully note: "The major limitations of this study are the lack of women, the retrospective design, and the failure to remove workup bias".

The paper by Marvin et al. [4] presents an application of a diffusion genetic algorithm to prognosis. They work on a sample of 237 gestational trophoblastic tumors , (a very large sample for this disease), to predict which women died two years after completing chemotherapy, and use eight predictors. To provide a basal performance they use a method based upon the adjusted odds ratio which reach as a correct classification in 95% of the survivors and 67% of the deaths. The genetic algorithm (GA) is applied to the entire set and the model with the better balanced performance in prediction of the survivors and of the deaths (using the lowest number of predictors) is chosen. The GA result is a linear combination of Boolean predictors and of an interaction term so that the model is readily interpretable by physicians. In this example, the pattern of the weight of the delays between pregnancy and diagnosis suggests two forms of the disease. On the entire set this model predicts 90% of the survivals and 87% of the deaths. Moreover, only 4 patients have a score between –100 and 100, thus the two subsets are well separated. The authors also undergo training testing, with a 10% testing set, the best selected model reaches a good performance. A questionable affirmation is "An important point to note is that the evolved model was trained on the entire sample. This is because the sample,..,is still relatively small", when the sample size is 237 and there are 8 predictors.

The paper by Ohno-Machado et al. [5] must be read by anyone interested in guidelines. It includes an excellent discussion of the necessity of computer-based guidelines. The collaborative approach by four experimented teams results in comprehensive and sufficient specifications of guideline representation. The GLIF model is object oriented and contains:

a class guideline with attributes: name, intention, eligibility criteria, guideline steps, entry step,... ; a class of guideline steps with subclasses action step, conditional step, branch step and synchronization step; a class patient data; a class action; a class criterion and a class supplemental material. The section on the evaluation of encoding guidelines is of interest and we recommend the companion paper by Patel et al. [8] to anyone interested in this cognitive task. This first step in the development of GLIF is a very promising one.

Tuttle and colleagues [6] state that the problem-oriented record is preferable to the source-oriented record in medical information management, so that the problem list is a central point. Their aims are: 1) comparability of patient descriptions which is required for outcomes analysis, quality improvement, estimation of the efficacy of treatment, etc.;.2) scalability; 3) sustainability to insure comparability across changes in terminology; and 4) capturing novel problems. They present Metaphrase: a middleware component designed to help caregivers in formalizing problem lists. Metaphrase accepts casual input and returns authoritative input, aids conceptualization of the problem, typing, spelling and translation of terms. The initial repertoire of authoritative terms is the Unified Medical Language System metathesaurus augmented by problem dictionaries from the Mayo Clinic and Harvard Beth Israel Hospital. Metaphrase supports a "good enough updating" which permits new terms to be added to annual releases of the Metathesaurus. Metaphrase includes a top-down parser which returns the best matches first, and allows a source precedence (order on the terminologies) so that the desired authoritative term is often displayed at the top of the suggested list. Interesting is the display of concept definitions to caregivers and the observation: "..definition represents validation of an agreement between human and machine". The fact that formalization of a problem list improves the quality and utility of the problem list, even though not surprising, is interesting. That Metaphrase, designed to fulfil enterprise objectives, seems useful for caregivers is a very encouraging result.

The two articles on knowledge processing have a common aim: to achieve shareable and reusable models or artifacts, so it is not surprising that the two common traits are pragmatism and cooperative work. Two articles [1,2] use artificial neural networks and compare their performances with those of more classical approaches, but they do not compare the facility of interpretation of the models. This last point is prejudicial since it can be argued that the fact that parameters of the model are easy to interpret can prove very useful to the practitioner. Another point is the bias in comparisons arising from an underspecification or misspecification of the classical models [9].

References:

1. Smith BP, Ward RA, Brier ME. Prediction of anticoagulation during hemodialysis by population kinetics and an artificial neural

network.. Artif Organs 1998;22:731-9.
2. Haydon GH, Jalan R, Ala-Korpela M, Hiltunen Y, Hanley J, Jarvis LM, Ludlum CA, Hayes PC. Prediction of cirrhosis in patients with chronic hepatitis C infection by artificial neural network analysis of virus and clinical factors. J Viral Hepat 1998;5:255-64.
3 Do D, Marcus R, Froelicher V, Janosi A, West J, Atwood JE, Myers J, Chilton R, Froning J. Predicting severe angiographic coronary artery disease using computerization of clinical and exercise test data. Chest 1998;114:1437-45.
4. Marvin N, Bower M, Rowe JE. An evolutionary approach to constructing prognostic models. Artif Intell Med 1999;15:155-65.
5. Ohno-Machado L, Gennari JH, Murphy SN, Jain NL, Tu SW, Oliver DE, Pattison-Gordon E, Greenes RA, Shortliffe EH, Barnett GO. The guideline interchange format: a model for representing guidelines. J Am Med Inform Assoc 1998;5:357-72.
6. Tuttle MS, Olson NE, Keck KD, Cole WG, Erlbaum MS, Sherertz DD, Chute CG, Elkin PL, Atkin GE, Kaihoi BH, Safran C, Rind D, Law V. Metaphrase: an aid to the clinical conceptualization and formalization of patient problems in healthcare enterprises. Meth Inf Med 1998;37:373-83.
7. Sheiner LB, Rosenberg B, Marathe VV. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. J Pharmacokin Biopharm 1977;5:445-79.
8. Patel VL, Allen VG, Arocha JF, Shortliffe EH. Representing clinical guidelines in GLIF: individual and collaborative expertise. J Am Med Inform Assoc 1998;5:467-83.
9. Tu JV, Weinstein MC, McNeil BJ, Naylor CD. Predicting mortality after coronary artery bypass surgery: what do artificial neural networks learn? The Steering Committee of the Cardiac Care Network of Ontario. Med Decis Making 1998;18:229-35.

Address of the author:
Joanny Gouvernet
Service de l'Information Médicale,
Hôpital Timone Adultes,
264 rue Saint Pierre,
13385 Marseille Cedex 05, France
E-mail: jgouvern@ap-hm.fr