Building a better prediction model

William A. Grobman, MD, MBA

In this issue of the Journal, Hamilton et al[1] address an important question: how can the results of observational data be presented in such a way that they are useful to clinicians at the bedside? The outcome on which the authors have chosen to focus, third- and fourth-degree lacerations, is certainly worthy of their attention because it has been associated with both short- and long-term morbidity. To accomplish their objective, the investigators have turned to classification and regression tree (CART) analysis, a methodology that has not been widely used in obstetrics but that has many characteristics that make it a potentially useful prediction tool.[2,3]

Often in observational studies, the independent association between factors and outcomes is estimated through techniques such as multivariable regression that yield a list of odds ratios. Although these techniques may identify several associations, these associations may not be particularly able to accurately predict outcomes; even if they are, it is not possible for a clinician at the bedside to easily combine multiple variables with different odds ratios into a patient-specific probability of outcome.

In contrast, in CART analysis, there is progressive partitioning of the population into subgroups as determined by the most predictive independent variables. The variables that are chosen, the discriminatory values of the variables, and the order in which the splitting occurs (ie, in what order the variables are used to split the population) are all produced by the analysis to maximize predictive accuracy. To use a tree, the clinician simply follows the paths of the tree that describe the characteristics of the patient being evaluated and arrives at the probability of the outcome of interest for that particular patient. Thus, patients are placed into the discrete groups that yield optimal overall prediction and that can be more easily utilized by clinicians. Additionally, the tree has an intuitive feel for clinicians, who are used to making decisions based on serially arising information.

Indeed, their work suggests how a CART approach can be elucidating for clinicians in the context of third- and fourth-degree lacerations. First, their approach illustrates how a single factor, in this case episiotomy, can be identified as being the single most useful characteristic in distinguishing high- and low-risk groups from one another. Their approach also reveals how a relatively small number of factors can be used to identify groups of women who have markedly different probabilities of a major perineal trauma.

Nevertheless, before such a model is adopted clinically, several concepts related to prediction models in general bear noting. Much to their credit, Hamilton et al[1] developed their model from a large dataset that included information from several institutions, thereby enhancing its generalizability. Yet a prediction model should be validated in additional populations.[4] First, this ensures that the discriminatory capacity of the original model does not, because of chance, appear to be better than it actually is. Second, this demonstrates that populations other than the one in which the model was developed will still have their outcomes accurately predicted by the model.

Several specific aspects of the model presented by Hamilton et al[1] also are important to consider. First, the very fact that episiotomy is of such primary importance in the model also suggests a potential weakness. In many institutions and for many providers, episiotomy is infrequent to the point of rarity. Thus, in these circumstances, the ability of the tree to provide further clinically usable information is extremely limited for the vast majority of the population. If a woman does not have an episiotomy, the only other discriminatory factor is the length of the second stage, and this can be used to distinguish women with point-estimate probability of 1.7% from those with a point-estimate probability of 8.9%.

The authors also cite the very powerful combination of episiotomy, instrumental delivery, and birthweight, noting, for example, that every one of the women who had episiotomy, vacuum, and a newborn with a birthweight 4312 g or more ended up with a third- or fourth-degree laceration. Although true, this is based on a sample size of just 7 women, meaning that the confidence interval for this probability is quite wide. More important, though, is the use of birthweight in the model. If the purpose of the CART is to contribute to an understanding of the factors, and the combination of the factors, most contributory to a given outcome, then birth weight is a completely acceptable factor to utilize. However, if the purpose of the CART is to help in prediction, then a factor such as birthweight is likely to be unhelpful.

Birthweight is a piece of data that is not known with certainty until after delivery, by which time the outcome of interest has occurred. Estimates of birthweight can be made, but these estimates are well known to be relatively inaccurate and cannot be simply substituted for the actual birthweight used in the model because the predicted probabilities may no longer be valid. A clinician faced with prospective decision making, and considering a vacuum or forceps delivery, will not know which

From the Department of Obstetrics and Gynecology, Feinberg School of Medicine, Northwestern University, Chicago, IL.

Received Jan. 4, 2011; accepted Jan. 6, 2011.

Reprints: William A. Grobman, MD, MBA, 250 East Superior St., Suite 05-2175, Chicago, IL 60611. w-grobman@northwestern.edu.

0002-9378/free

doi: 10.1016/j.ajog.2011.01.008

See related article, page 309

arm of the tree to follow or how to counsel accordingly. In general, a model designed to provide the prospective prediction of an outcome should utilize only variables that can be known prior to the outcome's occurrence.[5]

Regardless of these points, however, the fact remains that the work of Hamilton et al[1] should serve to pique the interest of the obstetric community with respect to CART and its potential uses. Prediction of meaningful outcomes via a method that can be incorporated into patient care remains an important clinical goal. Although CART has been applied by some, its use has been limited, and this manuscript should stimulate other researchers to consider whether the CART methodology is well suited to a problem at hand. ■

REFERENCES

1. Hamilton EF, Smith S, Yang L, Warrick P, Ciampi A. Third- and fourth-degree perineal lacerations: defining high-risk clinical clusters. Am J Obstet Gynecol 2011;204:309.e1-6.

2. Barnhart KT, Cassanova B, Sammel MD, Timbers K, Chung K, Kupl KL. Prediction of location of a symptomatic early gestation based solely on clinical presentation. Obstet Gynecol 2008;112:1319-26.

3. Grobman WA, Terkildsen MF, Solysik RC, Yarnold PR. Predicting outcome after cerclage using classification tree analysis. Am J Perinatol 2008;25:443-8.

4. Ransohoff DF. How to improve reliability and efficiency of research about molecular markers: roles of phases, guidelines, and study design. J Clin Epidemiol 2007;60:1205-19.

5. Grobman WA. Rates and prediction of successful vaginal birth after cesarean. Sem Perinatol 2010;34:244-8.