NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH Public Access
Author Manuscript
J Marriage Fam. Author manuscript; available in PMC 2006 October 25.

Published in final edited form as:
J Marriage Fam. 2001 May ; 63(2): 473–479.

The Hazards of Predicting Divorce Without Crossvalidation

Richard E. Heyman and **Amy M. Smith Slep**
State University of New York at Stony Brook

Abstract

Divorce prediction studies (e.g., Gottman, Coan, Carrere, & Swanson, 1998) suggest that couples' eventual divorce can be very accurately predicted from a number of different variables. Recent attention to these studies has failed to consider the need to crossvalidate prediction equations and to consider the prevalence of divorce in the population. We analyze archival data to demonstrate that accuracy and predictive value drops precipitously during crossvalidation. We conclude that results of studies without crossvalidation analyses should be interpreted with extreme caution, no matter how impressive the initial results appear to be.

Keywords

crossvalidation; divorce; overfitting; prediction studies

Can we really predict who will divorce from pre-marital data? The possibilities of such findings are staggering. The negative effects of marital decline and divorce are substantial and far-reaching on physical health (e.g., Burman & Margolin, 1992), psychological problems (e.g., Richards, Hardy, & Wadsworth, 1997), children's well-being (Grych & Fincham, 1990), and worker productivity (Forthofer, Markman, Cox, Stanley, & Kessler, 1996). Knowing who will eventually divorce would allow professionals, clergy, lay practitioners, and couples themselves to take steps to identify and ameliorate the factors that put them at risk before these factors take their toll. Accurate premarriage prediction of a couple's eventual divorce is truly the "Holy Grail" of marital and family preventionists.

Although a large number of studies have identified risk factors for divorce, only 15 published studies have predicted *who* will get divorced. These studies have used a variety of questionnaire, interview, and observational methods (Buehlman, Gottman, & Katz, 1992; Carrere, Buehlman, Gottman, Coan & Ruckstuhl, 2000; Crane, Soderquist, & Frank, 1995; Edwards, Johnson, & Booth, 1987; Fowers & Olson, 1986; Gottman, 1994; Gottman et al., 1998; Gottman & Levenson, 1999; Hill & Peplau, 1998; Jacobson, Gottman, Gortner, Berns & Shortt, 1996; Kurdek, 1993; Larsen & Olson, 1989; Lindahl, Clements & Markman, 1998; Matthews, Wickrama, & Conger, 1996; Rogge & Bradbury, 1999). The level of prediction accuracy is quite impressive, ranging from 67% to 95%. Most studies were prospective, following couples for 2–15 years. Sample sizes ranged from 54 to 286.

The six studies of Gottman and colleagues—Gottman et al. (1998) in particular—have received an extraordinary degree of professional and popular attention (e.g., Weiss, 2000), and thus a consonant degree of scrutiny is appropriate. We should note that Gottman et al. (1998) not only included estimates of prediction accuracy but also tested several theoretical models of

Department of Psychology, State University of New York at Stony Brook, Stony Brook, NY 11794-2500 (richard.heyman@sunysb.edu)..

Preparation of this article was supported Grant R01MH57779 from the National Institute of Mental Health. These analyses were derived from similar procedures conducted under contract support from a partnership of the U.S. Air Force Family Advocacy Program and the U.S. Department of Agriculture National Network on Family Resilience (contract CR-4953–545735).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

functional and dysfunctional couple behavior. Because a recent article in *Journal of Marriage and the Family* by Stanley, Bradbury, and Markman (2000) discussed problems with Gottman et al. (1998; many of which were disputed in a reply by Gottman, Carrere, Swanson, & Coan, 2000), the current article will be confined to an issue that Stanley et al. did not address: that the level of accuracy with which divorce can be predicted is being accepted and generalized without the necessary supporting statistical tests. That is, predictive equations must be crossvalidated for their true value to be known. No published study predicting divorce with general population couples has done this to date. (Crane et al. (1995) crossvalidated their predictive equation; however, they (a) used a marital therapy sample and (b) one of their two scales asked about cognitive and behavioral steps already taken toward divorce.) Two issues in particular—overfitting and the difficulty of predicting low prevalence events—plague the 15 prediction studies and make it imprudent to inform the public and clinicians that researchers can accurately predict who will divorce.

Overfitting: The Need for Crossvalidation

Discriminant function analyses and logistic regression can be used to predict a categorical outcome (e.g., divorce) in any given sample. More accurately, however, we should say that the analyst asks the software to *reconstruct*, rather than predict, because the computer develops an equation to optimally reconstruct an already-known group status. This is not a trivial, semantic distinction. The equation makes some variables more important, and some less important, by assigning weights to each variable. Thus, the equation is tailor-made for that sample and may be overly influenced by idiosyncrasies in a particular data set, a situation labeled by statisticians as “overfitting” (Tabachnick & Fidell, 1996). Overfitting can cause extreme overinflation of predictive powers, especially when oversampled extreme groups and small samples are used, as was the case with Gottman et al. (1998; $n = 60$ couples for the prediction analyses) and nearly all of the other divorce prediction studies.

To establish the accuracy of a predictive equation, one must crossvalidate it in an independent sample. First, one must use the *exact* weights from the original (i.e., development) sample. Second, the precise cut-point used in the development sample (i.e., the probability value specified such that participants with odds greater than that cut-point are predicted to divorce, and those below are predicted to not divorce) must also be used in the crossvalidation. In other words, crossvalidation only provides evidence of the accuracy of a predictive equation when the original weights and cut-points are used; developing another “highly predictive” equation with new weights and cut-points is not sufficient to say that the original predictive equation has been validated for predicting individual couples’ outcomes.

Difficulty Predicting Low-Prevalence Events

A predictive test’s or equation’s

> validity has two components: sensitivity and specificity. *Sensitivity* is defined as the ability of a test to identify correctly those who have the disease. *Specificity* is defined as the ability of a test to identify correctly those who do *not* have the disease. (Mausner & Kramer, 1985, p. 217)

For our purposes, divorce substitutes for disease. Claims that one can predict whether an *individual* couple will get divorced are, in essence, referring to the positive predictive value (PV_{pos}) of the test (i.e., “the proportion of true positives among all of those who have positive test results” Mausner & Kramer, p. 220).

When assessing the accuracy of equations intended to be applied in general populations, we also must consider that prevalence and predictive value are intertwined. As Sedlak, (1988, p. 327) wrote:

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

> The predictive value of a measure . . . is a function of three factors: (a) its sensitivity; (b) its specificity; and (c) the true prevalence of [the problem]:
>
> $$PV_{pos} = (\text{Prevalence} \times \text{Sensitivity}) \div ((\text{Prevalence} \times \text{Sensitivity}) + (1 - \text{Prevalence})(1 - \text{Specificity})).$$

According to this formula, the predictive value of any measuring device will be low whenever prevalence itself is low, *even for high levels of specificity and sensitivity*.

Although Gottman et al. (1998) did not break down their 80% correct divorce prediction into sensitivity and specificity, let's assume 80% sensitivity and 80% specificity and approximately 16% prevalence rate of divorce among couples married 3–6 years (Clarke, 1995; as opposed to the artificially imposed prevalence of 33% by Gottman et al.'s procedures). The positive predictive value for Gottman et al.'s equation would be 43%. In other words, if one were to use an "80% correct" equation to inform couples at premarriage whether they would divorce, one would be wrong more than half of the time when one told couples "you are likely to get divorced." Of course, one would still be improving on chance, but not with the degree of precision that a casual observer might expect.

The purpose of this paper is to use an archival data set to explore these issues by developing and crossvalidating a predictive equation for divorce. The purpose is to highlight statistical issues in prediction, not to test a theoretical model. We hypothesized that, because of overfitting and a relatively low prevalence, an equation that was highly predictive in the development subsample would have far poorer predictive abilities in the crossvalidation subsample. The analyses are meant to be illustrative of the statistical issues relevant to determining the true accuracy of a prediction equation. We should note that the archival data set used was cross-sectional, unlike the prospective data of Gottman and colleagues. Although the data are appropriate for the illustrative purposes of this paper, the data collection strategies are not identical (i.e., cross-sectional vs. prospective, representative of the U.S. population vs. convenience sample).

Method

Survey Methodology

For our analyses, we used the 1985 National Family Violence Survey (1985 NFVS; Gelles & Straus, 1994), a nationally representative data set on family matters comprising participants in divorced and intact relationships. For the 1985 NFVS, Louis Harris and Associates, the national polling firm, conducted 30-minute telephone interviews covering issues related to family life and family violence. These data were collected using a methodology designed to obtain nationally representative data and to allow detailed subgroup analysis (see Straus & Gelles, 1990, p. 530 for a detailed description of the sampling strategy). Participants (n = 6,002) were men and women aged 18 years or older who met one of the following criteria: (a) presently married, (b) presently living as a male-female couple, (c) divorced or separated within the last 2 years, or (d) single parent with a child under the age of 18.

Participants

We attempted to replicate Gottman et al.'s (1998) data set construction strategy, which took the total number of divorced couples (n = 20) and selected an equal number of those intact couples with the highest and lowest Marital Adjustment Test (MAT; Locke & Wallace, 1959) scores at the time of initial assessment. Although the 1985 NFVS is not a prospective study—and thus, unlike Gottman et al., we cannot analyze divorce over time—there were 176 recently divorced participants out of the 5,335 participants who were married, living together,

NIH-PA Author Manuscript

or divorced. To select contrast groups, we converted four MAT-like questions (about the frequency of disagreements about money, household chores, social activities, and sex/affection) into *z* scores and created a scale (α = .70). The 176 married or cohabiting participants with the highest disagreement scores and the 176 with the lowest disagreement scores were retained, along with the 176 divorced participants.

Procedures

Splitting retained sample into two subsamples—We wanted to (a) develop a predictive equation in one sample and (b) crossvalidate the equation in another sample. As discussed above, this is necessary because the development process uses knowledge of the values both of the predictors and of the dependent variable (i.e., divorce status) to create a prediction equation. It is critical, therefore, to test the accuracy of the equation in a completely separate sample. We therefore randomly split the full sample into two independent subsamples (i.e., development and crossvalidation subsamples).

We randomly split the 176 divorced participants and the 352 married or cohabiting participants into two subsamples, each with 88 divorced and 176 married or living together participants. Thus, the two subsamples were both randomly drawn from our full retained sample and were completely independent of each other (i.e., no cases appeared in both the development and crossvalidation subsamples). Demographics and characteristics of the two random subsamples were as follows: age of respondent (divorced: *M* = 43.60, *SD* = 16.52; married or cohabiting: *M* = 42.68, *SD* = 15.76), gender (divorced: 34% men, 66% women; married or cohabiting: 34.2% men, 65.8% women), fulltime employment status (divorced: 68% [men], 38% [women]; married or cohabiting: 68% [men], 43% [women]), median education (high school graduate for both groups and genders), race (divorced: White = 68%, African American = 15%, Hispanic = 15%; other = 3%; married or cohabiting: White = 56%, African American = 21%, Hispanic = 15%; other = 8%), and 1984 median family income (divorced: $20–30,000; intact: $20–30,000).

Development and crossvalidation of predictive equation—We examined the data set for theoretically viable predictor variables (listed below). We then determined the optimal combination of predictors and their weights through backward stepwise hierarchical logistic regression analyses. Logistic regression is a type of regression analysis that allows one to predict a dichotomous (i.e., yes-no) outcome from a set of predictor variables that may be continuous, categorical, or a mix of both. We chose to use logistic regression over the discriminant function analyses that Gottman et al. (1998) and others used because logistic regression is more robust: “Unlike discriminant function analysis, logistic regression has no assumptions about the distributions of predictor variables; in logistic regression, the predictors do not have to be normally distributed, linearly related, or of equal variance within each group” (Tabachnick & Fidell, 1996, p. 575). Hosmer and Lemeshow (1989) described the goal of stepwise logistic regression as finding the best fitting and most parsimonious model to describe the relationship between the dependent variable and a set of independent (i.e., predictor) variables. Thus, logistic regression allows us to evaluate the simultaneous contribution of a number of predictors, including interactions (when specified), to predict divorce. The regression analyses evaluate the contribution of each proposed independent variable to the prediction of the dependent variable, identifying predictors that do not add to the equation’s ability to predict divorce status.

The dependent variable was divorce status (i.e., divorced within the last 2 years vs. married or living together). The potential predictor variables were respondent’s age; number of children under 5; severity of drinking (abstinent, low, low moderate, high moderate, high, binge; Kantor & Straus, 1990); man’s and woman’s frequency of being high on drugs in the last year; man’s

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

and woman's education (none, 1–7 years, 8 years, some high school, high school graduate, some college, college graduate, some post graduate, advanced degree); a male-by-female education interaction; man's and woman's employment status (full time, part time, unemployed, retired, student, homemaker, disabled, and other); man's and woman's number of marriages; man's and woman's religion (none, Catholic, Protestant, Jewish, other); and male-to-female minor and severe physical aggression.

The development of the predictive equation comprised three steps. First, a backward stepwise hierarchical regression analysis was conducted to determine the optimal combination of predictors and their weights. Main effects for mens' and womens' education were entered first, with the education interaction and the other predictors entered next with the backward stepwise procedure. The predictor variables retained in the equation were man's and woman's education, man's education × woman's education, woman's present employment status, man's and woman's frequency of getting high on drugs in the last year, respondent's drinking index score, and the number of children under 5.

Next, to determine the optimal cut-point, the probabilities generated by the logistic regression equation were compared to the actual group status via a receiver operating characteristic (ROC) curve (a plot that displays the cut-point that maximizes both sensitivity and specificity). ROC curves have been used in setting cut-points and in evaluating prediction in a variety of areas (Swets, 1988; e.g., engineering, medicine, weather forecasting, educational testing). We used the SPSS (SPSS, Inc., 1999) ROC graphical procedure, with the "coordinate points of the ROC curve" output option selected. This option provides a table of the sensitivity and specificity for every possible cut-point. We then calculated the difference between sensitivity and specificity for each possible cut-point. The cut-point that simultaneously maximized both sensitivity and specificity was 0.222157493.

Third, the predicted group status variable was created. Participants with probabilities above the cut-point were predicted to be divorced, whereas those below the cut-point were predicted to be in intact relationships. Fourth and finally, the cross-tab table (comparing predicted and actual divorce status) was generated and sensitivity, specificity, and predictive value statistics calculated. The formulas were as follows: Sensitivity = True Positives / (True Positives + False Negatives); Specificity = True Negatives / (True Negatives + False Positives); Positive Predictive Value = True Positives / All Positives; Negative Predictive Value = True Negatives / All Negatives (Mausner & Kramer, 1985).

To crossvalidate the regression equation and cut-point (i.e., to test if the accuracy, as measured by the predictive value, remains high), the equation was converted to a series of SPSS compute statements, and these computations were applied to the independent crossvalidation subsample. These computations calculate the natural logarithm of the predicted odds of divorce, known as the logit. Next, the logit was converted to the probability by taking the exponent of the logit divided by 1 plus the exponent of the logit. Finally, by applying the cut-point, the predicted status of each couple was obtained. The crosstab table of predicted and actual divorce status was generated and sensitivity, specificity, and predictive value statistics calculated.

Results and Discussion

As shown in Table 1, the prediction equation correctly classified 90% of couples, with a sensitivity of 92% and a specificity of 89%. When predicting a couple as divorced, the equation was correct 65% of the time (positive predictive value), whereas when predicting a couple as not divorced, the equation was correct 98% of the time (negative predicted value). Thus, this equation performed in the upper range of the various prediction studies of divorce. This is the point at which most papers stop. We would be accurate to conclude at this point that the results

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

of the predictive equation are quite encouraging. Usually, we count ourselves fortunate if we have a large enough sample to have sufficient power to test our hypotheses; sufficient data to test our hypotheses and then crossvalidate our results is quite rare in marital and family studies. The following analyses make clear how restrained our optimism should be, however.

As hypothesized and as shown in Table 2, the accuracy, or predictive value, of the predictions dissipates substantially when the prediction equation is applied to an independent crossvalidation sample. The equation's sensitivity dropped 45% and its sensitivity dropped 15%. Predicting a couple would divorce is now accurate only 29% of the time. Thus, although the prediction equation is clearly capturing real relations among the variables, the positive predictive value of 29% no longer affords us a level of prediction that so dramatically improves on chance. Certainly, one would not want to suggest to couples that their relationships were at risk solely on the basis of this equation because most of the participants in this crossvalidation sample that would have been classified as divorced would, in fact, be in relationships that are intact. Thus, if a paper stopped at the equation development stage, encouragement would certainly still be warranted, but caution regarding overgeneralizing would certainly need to be emphasized. These analyses clarify the impact of overfitting but do not yet address the impact of prevalence on predictive accuracy.

To examine the impact of prevalence, we controlled for the effects of the artificially inflated prevalence of divorce from our Gottman-parallel participant selection procedures (see Table 3). Thus, we took all married or living together participants who were either (a) included in the crossvalidation sample or (b) originally not included in either the development or crossvalidation sample. We weighted the married/living together group so that the final prevalence (i.e., after participants with missing data were dropped) was equivalent to that of couples who divorce within the first 3–6 years of marriage (i.e., 16%; Clarke, 1995) and applied our predictive equation. As noted earlier, predictive value is affected by prevalence (Sedlak, 1988). Readjusting the prevalence down to 16% caused a further drop in the proportion of couples predicted to be divorced who actually were divorced (i.e., PV_{pos}: 20.99%). This increment in prediction above the prevalence (from 16 to 21%) may still be important and have useful applications. Nonetheless, it clearly suggests that this equation does not hold the key to answering the question "Who will get divorced?"

Clearly, there are innumerable differences between the analyses presented here and Gottman et al. (1998) and the other divorce prediction studies summarized earlier. Yet, the message conveyed in the analyses is relevant. Our prediction equation, as evaluated in the development subsample, was extremely accurate. Without performing the crossvalidation analyses, there is no way to know how robust the equation would be. Given the nature of discriminant function and logistic regression analyses, however, one can be nearly certain that the equation will not perform as accurately when applied to any sample independent of the development sample. The analyses presented here included many predictor variables that can be measured with little error (e.g., education) and both our development and crossvalidation samples were randomly drawn from a single, representative data set. Both of these characteristics should enhance the robustness of the prediction equation. Yet, the accuracy of the equation still dropped markedly when applied to an independent subsample (and even further when the actual population baserate of early divorce was taken into account). An equation with an initial overall accuracy of 90% ended up with a positive predictive value of 21%. These analyses suggest a tempered reaction to initial, uncrossvalidated results. Large sample sizes would help mitigate some of the problems with overfitting, but because of the high cost and effort of collecting observational data such as that of Gottman et al. (1998), it may be unrealistic to expect the collection of 1,000 or so participants.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

In conclusion, published studies that find extraordinary initial predictive results may aid us in improving models of risk by identifying important risk factors. Nonetheless, dissemination of "predictive power" results in the popular media *must* await supportive data on sensitivity, specificity, and predictive value when the predictive equation is applied to independent samples. By recognizing both the value and limitations of predictive studies, professionals and the public alike will be served best.

References

Buehlman KT, Gottman JM, Katz LF. How a couple views their past predicts their future: Predicting divorce from an oral history interview. Journal of Family Psychology 1992;5:295–318.

Burman B, Margolin G. Analysis of the association between marital relationships and health problems: An interactional perspective. Psychological Bulletin 1992;112:39–63. [PubMed: 1529039]

Carrere S, Buehlman KT, Gottman JM, Coan JA, Ruckstuhl L. Predicting marital stability and divorce in newlywed couples. Journal of Family Psychology 2000;14:42–58. [PubMed: 10740681]

Clarke, S. C. (1995). Advance report of final divorce statistics, 1989 and 1990. *Monthly Vital Statistics Report, 43* (Suppl. 9). Hyattsville, MD: National Center for Health Statistics.

Crane DR, Soderquist JN, Frank RL. Predicting divorce at marital therapy intake: A preliminary model. American Journal of Family Therapy 1995;23:227–236.

Edwards JN, Johnson DR, Booth A. Coming apart: A prognostic instrument of marital breakup. Family Relations 1987;36:168–170.

Forthofer MS, Markman HJ, Cox M, Stanley S, Kessler RC. Associations between marital distress and work loss in a national sample. Journal of Marriage and the Family 1996;58:597–605.

Fowers BJ, Olson DH. Predicting marital success with PREPARE: A predictive validity study. Journal of Marital and Family Therapy 1986;12:403–413.

Gelles, R. J., & Straus, M. A. (1994). *Physical violence in American families, 1985* (3rd release) [computer file]. Durham, NH: University of New Hampshire Family Research laboratory [producer], 1988. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor].

Gottman J, Carrere S, Swanson C, Coan JA. Reply to "From basic research to interventions. Journal of Marriage and the Family 2000;62:265–273.

Gottman, J. M. (1994). *What predicts divorce?* Hillsdale, NJ: Erlbaum.

Gottman JM, Coan J, Carrere S, Swanson C. Predicting marital happiness and stability from newlywed interactions. Journal of Marriage and the Family 1998;60:5–22.

Gottman JM, Levenson RW. Rebound from marital conflict and divorce prediction. Family Process 1999;38:287–292. [PubMed: 10526766]

Grych JH, Fincham FD. Interventions for children of divorce: Toward greater integration of research and action. Psychological Bulletin 1990;111:434–454. [PubMed: 1594720]

Hill, C. T., & Peplau, L. A. (1998). Premarital predictors of relationship outcomes: A 15-year follow-up of the Boston Couples Study. In T. N. Bradbury (Ed.), *The developmental course of marital dysfunction* (pp. 237—278). New York: Cambridge University Press.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression.* New York: Wiley.

Jacobson NS, Gottman JM, Gortner E, Berns S, Shortt JW. Psychological factors in the longitudinal course of battering: When do the couples split up? When does the abuse decrease? Violence and Victims 1996;11:371–392. [PubMed: 9210278]

Kantor, G. K., & Straus, M. A. (1990). The "drunken bum" theory of female beating. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 203–224). New Brunswick, NJ: Transaction Books.

Kurdek LA. Predicting marital dissolution: A 5-year prospective longitudinal study of newlywed couples. Journal of Personality and Social Psychology 1993;64:221–242.

Larsen AS, Olson DH. Predicting marital satisfaction using PREPARE: A replication study. Journal of Marital and Family Therapy 1989;15:311–322.

Lindahl, K., Clements, M., & Markman, H. (1998). The development of marriage: A nine-year perspective. In T. N. Bradbury (Ed.), *The developmental course of marital dysfunction* (pp. 205–236). New York: Cambridge University Press.

Locke HJ, Wallace KM. Short marital adjustment and prediction tests: Their reliability and validity. Marriage and Family Living 1959;21:251–255.

Matthews LS, Wickrama KAS, Conger RD. Predicting marital instability from spouse and observer reports of marital interaction. Journal of Marriage and the Family 1996;58:641–655.

Mausner, J. S., & Kramer, S. (1985). *Epidemiology: An introductory text.* Philadelphia: Saunders.

Richards M, Hardy R, Wadsworth M. The effects of divorce and separation on mental health in a national UK birth cohort. Psychological Medicine 1997;27:1121–1128. [PubMed: 9300516]

Rogge RD, Bradbury TN. Til violence does us part: The differing roles of communication and aggression in predicting adverse marital outcomes. Journal of Consulting and Clinical Psychology 1999;67:340–351. [PubMed: 10369054]

Sedlak, A. J. (1988). Prevention of female abuse. In V. B. Van Hasselt, R. L. Morrison, A. S. Bellack, & M. Hersen (Eds.), *Handbook of family violence* (pp. 319–358). New York: Plenum Press.

SPSS, Inc. (1999). Statistical Package for the Social Sciences (SPSS 9.0) user's guide. Englewood Cliffs, NJ: Prentice-Hall.

Stanley SM, Bradbury TN, Markman HJ. Structural flaws in the bridge from basic research on marriage to interventions for couples: Illustrations from Gottman, Coan, Carrere, and Swanson (1998). Journal of Marriage and the Family 2000;62:256–264.

Straus, M. A., & Gelles, R. J. (1990). *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families.* New Brunswick, NJ: Transaction.

Swets JA. Measuring the accuracy of diagnostic systems. Science 1988;240:1285–1293. [PubMed: 3287615]

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics.* New York: Harper Collins.

Weiss P. Is this marriage on the rocks? An author and his spouse have their heads and hearts examined at the Love Lab. New York Times Magazine 2000 May 7;:61–64.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Classification Table: Development Subsample

Predicted Status	Actual Status: Divorced	Actual Status: Married or Living Together
Divorced	34 (true positive)	18 (false positive)
Married or living together	3 (false negative)	149 (true negative)

Note: Percent correct: 89.71%; sensitivity: 91.89%; specificity: 89.22%; PV_{pos}: 65.38%; PV_{neg}: 98.03%. Sixty subjects were dropped because of missing data.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 2

Classification Table: Crossvalidation Subsample

	Actual Status	
Predicted Status	**Divorced**	**Married or Living Together**
Divorced	17 (true positive)	42 (false positive)
Married or living together	20 (false negative)	123 (true negative)

Note: Percent correct: 69.31%; sensitivity: 45.95%; specificity: 74.55%; PV_{pos}: 28.81%; PV_{neg}: 86.01%. Sixty-two subjects were dropped because of missing data.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Classification Table: Full Crossvalidation Subsample, Weighted to Produce Divorce Rate of 16%

	Actual Status	
Predicted Status	**Divorced**	**Married or Living Together**
Divorced	17 (true positive)	64 (false positive)
Married or living together	20 (false negative)	130 (true negative)

Note: Percent correct: 89.71%; sensitivity: 45.95%; specificity: 67.01%; PV_{pos}: 20.99%; PV_{neg}: 86.67%.