Predicting Abuse and Reassault Among Batterer Program Participants

By D. Alex Heckert and Edward W. Gondolf

2004
NCJ 199730
D. Alex Heckert, Ph.D., is with the Department of Sociology, Indiana University of Pennsylvania. Edward W. Gondolf, Ed.D., MPH, is with the Mid-Atlantic Addiction Training Institute (MAATI), Indiana University of Pennsylvania.

The authors wish to thank the program directors at the four research sites who offered essential assistance in developing and implementing the data collection, and staff at MAATI who provided a variety of supports. We also want to thank our research assistants, especially Angie Beeman, Bob White, Jeff Rowles, and Danielle Payne. We extend particular thanks to our statistical consultant, Dr. Alison Snow Jones, for her excellent statistical advice.

Data coding, management, and analysis were supported by award 98–WT–VX–0014 from the National Institute of Justice, U.S. Department of Justice. Data collection and coding were supported through grants from the Centers for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services (Grant number R49/CCR310525–02). Findings and conclusions of the research reported here are those of the authors and do not reflect the official position or policies of the U.S. Department of Justice or CDC.
Previous Research

Many practitioners and researchers are intensifying efforts to predict reassault among men referred to batterer programs. These efforts include developing lists of predictive factors (Saunders, 1995; Dutton and Kropp, 2000), batterer profiles and types (Holtzworth-Munroe and Stuart, 1994), and risk assessment scales or indexes (Dutton and Kropp, 2000; Roehl and Guertin, 2000). Some risk markers have been identified, such as alcohol or drug abuse, previous criminality, severe personality disorders, and program dropout (Dutton et al., 1997; Gondolf, 1997; Saunders, 1995; Tolman and Bennett, 1990). However, overall prediction is weak, with limited ability to correctly classify reassault on a better than chance basis (Limandri and Sheridan, 1995; Saunders, 1995).

The emphasis in practice and research is shifting to identifying specific types of batterers (e.g., through batterer typologies), especially high-risk offenders, and developing specialized interventions to accommodate these different types of offenders (Healey, Smith, and O’Sullivan, 1998; Saunders, 1996). Current typologies are based either on psychological characteristics (Holtzworth-Munroe and Stuart, 1994) or criminal justice factors such as demographic information, criminal histories, and substance abuse data (Goldkamp, 1996; Gondolf, 1988). With the exception of Goldkamp’s criminal justice-based typology, which was predictive of rearrest for domestic violence, these typologies have not yet been confirmed as predictive.

A concurrent development in the domestic violence field is the use of risk assessment inventories, such as the Spousal Assault Risk Assessment (SARA) (Kropp et al., 1995) and the Kingston Screening Instrument for Domestic Violence Offenders (K–SID) (Gelles and Tolman, 1998). These instruments have expanded on the long-standing efforts of practitioners to develop lethality checklists (Hart, 1994). To date, risk assessment inventories offer, at best, marginal prediction (Dutton and Kropp, 2000; Roehl and Guertin, 2000). These inventories do offer an improvement over clinical judgment, but they do not appear to correctly classify men at a clinically acceptable level. More development of the inventories or more sophisticated prediction research may show otherwise.

In general, the prediction research has been limited by several factors. It has relied on simplistic dichotomous outcomes of “success” and “failure,” it has not considered mediating or “conditional” variables, and the databases used have been too small or limited to address these problems (Monahan, 1996; Mulvey and Lidz, 1993). This study reexamines the prediction of abuse and reassault among batterer program participants by addressing these methodological shortcomings. It attempts to improve prediction of reassault with multiple outcomes, conditional variables, and a comprehensive multisite database.

Objectives of Current Research

An extensive, multisite, longitudinal database of batterers and their female partners was used to test several possibilities for prediction: the utility of risk markers, conditional variables, risk instruments, and batterer types. The database also allowed for the exploration of the dynamics of reassault and other alternative conceptions of violence. The database included intake interviews.
with 840 batterers and their partners and followup interviews every 3 months over 15 months. The followup response rate was 70 percent.

The primary objective of the proposed research was to test a conditional prediction model of multiple outcomes of batterer intervention using multinomial (polytomous) logistic regression (see Heckert and Gondolf, 2002, for details). The main contribution of the model is that it considers multiple outcomes rather than simply a dichotomous success or failure. Partner self-reports about the batterer’s physical assault, verbal abuse, and controlling and threatening behaviors were used to identify five distinct outcomes for batterer program participants: nonabusive behaviors during followup (22 percent of the sample), controlling behavior/verbal abuse only (26 percent), threatening assault with no physical reassault (19 percent), one-time reassault (12 percent), and repeat reassault (21 percent). The authors hypothesized that this more sensitive measurement of the abuse outcome would improve prediction.

The second contribution of the model is that it considers the influence of conditional or situational variables that occur after program intake. Intervening factors assessed at 3 months following program intake were entered into the regression equations after program intake risk markers, batterer types, and program participation variables. These variables included batterer or victim employment, partner contact and new partners, batterer’s perceptions of sanctions for program dropout and reassault, alcohol and drug use, batterer alcohol or psychological treatment, the woman’s use of victim services and other help sources, and additional criminal justice intervention. The authors hypothesized that prediction of multiple outcomes would be improved by including these conditional factors in the multivariate models.

A third contribution is that it permits examining the predictive abilities of risk assessment inventories using the prediction model of multiple outcomes. Several popular risk assessment instruments were simulated using variables measured at intake. The authors hypothesized that risk assessment instruments would provide modest, but not substantial, prediction of the multiple outcomes.

A fourth contribution of the model is that it permits conducting case reviews of batterers to further clarify and substantiate the conditional prediction model. Personality profiles were elaborated for batterers using the Millon Multiaxial Clinical Inventory, Version III (MCMI–III; Millon, 1994) data. Batterer and victim narratives of reassault were also analyzed to describe the dynamics of the abuse and the extent and influence of various risk markers. The authors hypothesized that the men who repeatedly reassault are more likely to be psychopathic based on MCMI–III profiles; to commit excessive, unrelenting, escalating violent incidents; and to come from discussion-oriented as opposed to instructional programs.

**Methods**

**Database**

To address the research hypotheses, a multisite database of batterers and their female partners was used that included 840 men who were admitted to batterer programs in four cities—Pittsburgh, Dallas, Houston, and Denver. The database offered a large representative sample of batterers across four sites and diverse regions. The vast majority of the men (82 percent) were
mandated to the programs by the courts, while the rest (18 percent) voluntarily entered the
program. (See Gondolf, 1999, for a detailed description of the study design, sample recruitment,
and sample demographics.) Interviews were conducted at program intake with batterers and their
female partners, and with batterers, initial victims, and new female partners every 3 months for
15 months. The modalities of the four batterer programs conform to the parameters of the
prevailing State standards, which endorse cognitive-behavioral techniques taught in a group
setting. However, the selected programs represent a range of services and duration (see Gondolf,

At program intake, a background questionnaire was administered to the men that included
questions about the incident that led to referral to a batterer program. The men were asked a
series of open-ended questions, followed by the Conflict Tactics Scale (CTS) items for “physical
aggression” (Straus, 1979). The background questionnaire also asked about the men’s
demographic profile, living situation, mental health problems, alcohol use, prior treatment and
counseling, emotionally abusive behavior, previous arrests, partner’s response, and partner’s
help-seeking. An alcohol screening test and personality inventory were also administered.

Variables

The predictor variables were derived from background questionnaires administered to the men
and their partners at program intake. They included demographics, relationship status, past
behavior (including previous violence, arrests, substance use, and Michigan Alcohol Screening
Test results), and mental health (including MCMI–III results, psychiatric symptoms, and
psychological treatment). The men’s and women’s reports of past help-seeking and service
contact were also used. The women’s perceptions of their safety were obtained through
interviews conducted with them at the time of program intake. The women were asked, “How
safe do you feel?” and “How likely will your partner use violence again within the next 3
months?” using a Likert scale response. Conditional variables were identified from the 3-month
followup interval with the women. They included living arrangements, contact between partners,
employment status, substance use, further arrests, and use of additional services and treatment.
The multiple outcome variable was based on reports by the women regarding the men’s abusive
behavior. Men were classified in the following five categories based on their partners’ reports of
abuse during the 15-month followup:

♦ Repeat reassault: more than one incident that included one of the tactics on the physical
aggression subscale of the Conflict Tactics Scale (Straus, 1979).

♦ One-time reassault: only one incident of physical aggression.

♦ Threatening reassault: no physical tactics but any threats (i.e., to hit, attack, or harm; to kill;
to take away children or harm them, to kill or seriously harm other people; to kill or hurt
himself).

♦ Emotional abuse: no threats or physical tactics, but any controlling behaviors or verbal abuse
(i.e., kept from talking on phone; kept from friends; stopped from going some place;
followed partner; kept from using family income; took or stole money from partner; swore or
screamed; accused partner of being with another man; insulted or put down; threw, smashed, hit, or kicked something; destroyed property; or hurt a pet or pets).

♦ No abuse: no reports of physical assault, threats, or emotional abuse over 15 months.

A second multiple outcome variable was also constructed using interviews starting at the 6-month followup (that collected information from 3 to 6 months after intake) through the 15-month followup. This outcome excluded the first 3 months after intake and allowed testing of the conditional variables encountered from intake to 3 months.

To explore the last hypothesis, qualitative coding was used. The psychological characteristics of repeat reassaulters were investigated by interpreting the men’s MCMI–III profiles (Gondolf and White, 2001). The interpretations recommended in the instrument manuals were followed, with one revision. The authors identified psychopathic tendencies according to profile configurations recommended by experts on psychopathy and the MCMI (Blackburn, 1998; Millon and Davis, 1998). Any evidence of psychopathic tendencies was given priority over other possible interpretations to ensure the maximum inclusion of such tendencies. The broader and more liberal conception of psychopathy is likely to identify more men than narrower conceptions previously used in the field.

Qualitative Analysis

To assess the violent behavior of the men, the research team coded the women’s descriptions of the violent incidents using a sequential, situational conception of violence (Monahan, 1996; Mulvey and Lidz, 1993). First, research assistants coded the issues, circumstances, precipitants, alcohol use, man’s emotional state, couple interaction, pattern of tactics, and woman’s and man’s response to the violence. The codes for the various components were then cross-tabulated with the categories for reassault (no, once, repeat) to identify differences across the outcomes. The assistants also wrote their overall impressions and observations of the violence in each case, and other researchers summarized this information and used it to confirm and elaborate the cross-tabulations.

Results

Risk Markers

To address the study’s hypothesis about risk markers, logistic regression models were estimated using a dichotomous outcome of any reassault versus no reassault. The results confirmed previous research; significant predictors of reassault included younger age, race, living with partner, no children, heavy drinking, emotional abuse or threats, high likelihood of hitting, low help-seeking by the woman, and the woman’s shelter use. The dichotomous model was significant but had modest ability to predict reassault cases (overall accuracy = 75 percent; sensitivity = 44 percent).

The research team conducted a multinomial logistic regression analysis using only variables collected at program intake. The analysis was based on 499 cases for which the multiple outcome variables could be constructed and data on predictors were available. Two multinomial logistic
regression equations were estimated using the same predictors described above: an ordered multinomial logistic regression (cumulative log model or proportional odds model) and an unordered multinomial logistic regression for comparison. Based on a likelihood ratio test and tests of the assumptions of parallel lines, the results demonstrated that the unordered multinomial model was significantly better than the ordered model. Thus, the multiple outcomes variable should be treated as a nominal outcome variable, rather than an ordinal outcome variable. The multiple outcome categories do not necessarily represent a progression of least to most severe abuse.

The logistic equations were further examined to assess which outcome categories were best predicted or distinguished. First, the sets of predictors that distinguished repeat reassaulters from no abuse and repeat reassaulters from verbal abuse/controlling categories are very similar. Thus, “no abuse” and “verbally abusive/controlling” batterers are essentially indistinguishable based on variables available at program intake. Second, the best discrimination by variables available at program intake is between the repeat reassault and the no abuse categories. There are fewer variables that discriminate between the repeat reassault and use of threats categories. Third, there are few factors that discriminate between the repeat reassault and the one-time reassault categories (age, race, occupation, use of controlling behaviors within 3 months of intake, women’s perceptions of risk, and use of shelter prior to intake). However, the odds ratios suggest they are reasonably strong predictors. Fourth, the overall model does a reasonably good job of predicting repeat reassault (sensitivity for repeat reassault = 70 percent); however, it does have a high enough rate of false negatives (batterers who are predicted to not be repeat reassaulters who are = 30 percent) to cause concern about using risk markers for decisionmaking in the criminal justice system.

In sum, the first hypothesis was only partially supported. Prediction is improved with a multiple outcome but is still relatively weak.

**Conditional Prediction and Risk Assessment Instruments**

The second hypothesis about a conditional model of prediction was tested by entering conditional variables, measured at the 3-month followup, into the logistic regression equations, using the multiple outcome based on the 6- through 15-month followups. These logistic regressions did not improve prediction over the initial risk marker models (sensitivity for repeat reassault = 57 percent), although a number of conditional variables (e.g., relationship troubles and woman filed for a protection order) were significant predictors. The second hypothesis was not supported.

To address the third hypothesis about the risk inventories, three popular risk assessment instruments were simulated with the authors’ data—K–SID, SARA, and Campbell’s Danger Assessment Scale (DAS). The K–SID scores by themselves gave weak prediction of multiple outcomes (sensitivity for repeat reassault = 11 percent). The SARA total scores (sensitivity = 43 percent) and DAS (sensitivity = 66 percent) total scores worked substantially better than the K–SID scores but still offered modest prediction of multiple outcomes and high rates of false positives (predicting men to repeatedly reassault who do not do so; 27 percent for SARA and 33 percent for DAS). Interestingly, women’s perceptions (assessed at intake) of safety (sensitivity = 63 percent; false positives = 40 percent) and how likely the man is to use violence (sensitivity =
52 percent; false positives = 26 percent) were also modest predictors of multiple outcomes by themselves and were slightly better predictors than SARA. The best prediction was achieved by DAS (sensitivity rate = 66 percent).

Women’s perceptions of risk (at intake) had a higher rate of correct classification of repeat reassaulters than did two of the risk assessment instruments (K–SID and SARA). However, there were more false positives with the women’s perceptions as predictors. The combination of women’s perceptions and either the SARA total score or DAS were the best models in this set of analyses. Nonetheless, together they still offered only modest predictive ability and were not quite as predictive as the initial equations with individual risk factors, and they had a higher rate of false positives. Hypothesis three was, therefore, supported, but the prediction was still not at clinically acceptable levels.

**Additional Predictors**

A comparative analysis of the MCMI profiles of men in this study was conducted to explore for other differentiation that might help improve prediction (Gondolf and White, 2001). Previous batterer typology and personality research suggests that the men most likely to repeatedly reassault their partners tend to be antisocial and psychopathic. However, only about 11 percent of the 122 repeat reassaulters in the sample had personality profiles that suggested conventional or “primary” psychopathic disorder. The percentage of batterers who were categorized as having primary psychopathic disorder was similar across three groups of men: those who did not reassault their partner (8 percent; 33 of 394), those who reassaulted their partners once (9 percent; 6 of 68), and those who repeatedly reassaulted their partners (11 percent; 13 of 122) during a 15-month followup. The broadest possible conception of psychopathy, including “secondary” psychopathy and both psychopathic “disorder” and “style,” applies to 54 percent of the repeat reassaulters, 39 percent of nonassaulters, and 35 percent of one-time reassaulters ($p < .05$). Although a significantly greater portion of men in the repeat reassault category show some psychopathic tendencies, there were no significant differences across the three types of batterers with regard to personality dysfunction, psychopathic disorder, and personality type. In sum, a diversity of personality profiles seems to best characterize all three groups of men.

The analysis of the violence incidents also did not substantiate the researchers’ expectations (Gondolf and Beeman, 2003). A distinguishing mode of violence was not found, but a few circumstances did stand out. First, men in the repeat reassault category were more likely to be described as drunk, but alcohol use was not consistently indicated in the women’s descriptions. There were few differences in the other issues, precipitants, circumstances, or emotions. Second, men in the repeat reassault category were slightly more likely to use a chain of tactics, or multiple tactics, in their violent incidents. That is, their violence was more likely to be excessive and unrelenting. Third, the only substantial difference was in the women’s interaction during, and response after, the violence. The partners of men who repeatedly reassaulted were less likely to resist the violence during an attack and less likely to seek help in response to the violence. Fourth, when action was taken against the men who repeatedly reassaulted, they were less likely than the one-time reassaulters to be sanctioned or contained. Police did not arrest them, courts did not jail them, and social services did not refer them. In short, these men continued to get away with being violent.
Quantitative analysis showed that men in the repeat reassault category were not more likely to come from discussion-oriented programs rather than instructional programs. Overall, then, hypothesis four received minimal support.

**Discussion**

**Prediction Improvements**

This study’s attempt to improve prediction of further abuse by batterer program participants produced some instructive findings. Using multiple outcomes does appear to improve prediction using intake risk markers, while the addition of conditional variables does not improve the prediction but identifies important predictors. The items from the risk assessment instruments also modestly predicted the outcomes, but only the DAS was more predictive than the women’s perceptions by themselves. The strongest prediction occurs by entering risk markers as individual items (and including women’s perceptions), rather than combining them into a composite index. As one might expect, the more sophisticated the prediction model, the better the prediction. There remains, however, a subjective decision about the utility of the improved prediction. The sophisticated models still only modestly predict the outcomes and do not appear to be sufficient for clinical decisions by themselves. The study’s qualitative exploration for other possible predictors or categorizations using the MCMI profiles and violent incidents did not produce other worthwhile considerations. Neither the profiles nor the incidents appeared to distinguish the outcome categories.

**Implications for Researchers**

This study demonstrates the importance of considering multiple outcomes in batterer research. Multiple outcomes not only modestly improve prediction, but they also expose different sets of predictors than do dichotomous outcomes. The findings suggest why causal research has produced inconsistent results (see Aldarondo and Sugarman, 1996). (“Causal research” refers to studies identifying factors that help to explain future reassault, as opposed to simply identifying who is most likely to reabuse.) Different predictors for repeat reassault compared with one-time reassault could cancel themselves out in an equation with a dichotomous outcome. Moreover, samples with fewer men in the repeat reassault category are likely to produce different predictors than samples with more men in the repeat reassault category. Excellent prediction can be derived with small samples, but such prediction is generally not replicable across samples because of variations in the influential subcategories of reabuse and reassault.

The findings raise some question for future research with multiple outcomes. It appears that additional variables modestly improve the prediction of multiple outcomes. However, it is uncertain how to substantially improve prediction or, indeed, whether it can be improved much further. Better measurement of the existing variables and the identification of additional influential variables (such as motivation) might improve prediction. The increased complexity, however, makes it more difficult to translate prediction into clinical practice. Further verification of risk instruments that use this approach is needed because the authors were able to simulate only instruments with limited items (either the same or similar items for 10 of 11 items for K–SID, 16 of 20 items for SARA, and 12 of 15 items for DAS).
Implications for Practitioners

The findings raise a few implications for clinical assessment of batterers, particularly the effort to identify and contain the most dangerous men. First, the results indicate the importance of distinguishing between one-time reassault and repeat reassault when attempting to identify high-risk batterers. The two groupings have different risk markers and may not be as readily identified if combined into one group. Second, the emphasis on personality traits and personality types failed to improve prediction of repeat reassault. Therefore, using psychological assessments to identify the extent of intervention or level of constraint may not be that useful.

Third, risk assessment instruments appear to offer only modest prediction in this study and should be used with caution by batterer programs and the criminal justice system, as previous research has recommended (see Roehl and Guertin, 2000). Results are improved somewhat by including additional items and women’s perceptions, reinforcing the importance of using instrument results in combination with a variety of other sources of information. Fourth, the predictive power of women’s perceptions suggests the importance of obtaining and heeding women’s appraisal of their situation, as advocates have long argued. Batterer program staff and the courts may have to work more closely with women’s advocates to obtain such information and incorporate it into their assessments.

The quantitative and qualitative findings, however, contradict overgeneralizations about high-risk batterers. These batterers are not readily or easily identifiable or “typed.” According to their personality profiles, many of the repeat reassaulters appear to be appropriate candidates for conventional batterer counseling. The findings also imply that conditional variables enhance prediction beyond personality factors. Shifting attention from intake assessment to ongoing risk management would likely improve identification and containment of the most dangerous men. Furthermore, this analysis, particularly of violent incidents, suggests that women’s characteristics (i.e., levels of assertiveness, help-seeking, satisfaction with services) warrant further consideration. Prevention efforts need to consider support and safety planning with the women, as much as containment and restraint of the men.

In sum, improvement of identification and containment of the most dangerous men requires not only further differentiation of batterers, but also consideration of a wide range of information, sources, and timeframes. Conventional intake assessment or risk instruments have limited predictive power, and even the more extensive and sophisticated predictions are not particularly strong.

References


Predicting Levels of Abuse and Reassault Among Batterer Program Participants


