Empirically Examining the Risk of Intimate Partner Violence: The Revised Domestic Violence Screening Instrument (DVSI-R)

KIRK R. WILLIAMS, PhD^a

STEPHEN R. GRANT, MA, CAGS^b

SYNOPSIS

Objective. This study extends recent research on assessing the risk of intimate partner violence by determining the concurrent and predictive validity of a revised version of the Domestic Violence Screening Instrument (DVSI-R) and whether evidence of such validity is sustained independent of perpetrator demographic characteristics and forms of intimate violence. The analyses highlight violent incidents involving multiple victims as an indicator of “severe” violence. Previous research did not address these issues.

Methods. Data were analyzed on 14,970 assessments conducted in the State of Connecticut from September 1, 2004 through May 2, 2005. Hierarchical regression and receiver operating characteristic analyses were used to address the objectives of this research.

Results. The empirical findings support the concurrent and predictive validity of the DVSI-R and show that it is robust in its applicability. The findings further show that incidents involving multiple victims are highly associated with DVSI-R risk scores and recidivistic violence.

Conclusion. Validating and demonstrating the robustness of a risk assessment instrument is only a first step in preventing violence involving intimate partners or others in family or family-like relationships. The challenge is to train professionals responsible for addressing the problem of such violence to link valid risk assessments to well-crafted strategies of supervision and treatment so that the victimized or other potential victims are protected and perpetrators are held accountable for their actions.

^aDepartment of Sociology, Robert Presley Center for Crime and Justice Studies, University of California, Riverside, Riverside, CA

^bDeputy Director, Family Services, Court Support Services Division, Judicial Branch, Wethersfield, CT

Address correspondence to: Kirk R. Williams, PhD, Department of Sociology, Robert Presley Center for Crime and Justice Studies, University of California, Riverside, Riverside, CA 92521; tel. 951-827-4604; e-mail <Kirk.Williams@ucr.edu>.

©2006 Association of Schools of Public Health

Risk assessment has become a common practice among professionals having the responsibility of determining what should be done to prevent intimate partner violence. Indeed, whether required as an official administrative process or an unofficial “guess” about the danger posed by perpetrators and the safety of victims, public health and criminal justice practitioners, victim advocates, shelter workers, and emergency room health care providers among others must estimate the likelihood of recurrent or escalating violence involving people in intimate relationships. Some consider risk assessment as the fundamental cornerstone of professional efforts to prevent this violent behavior.¹

Violence prevention researchers have mobilized to support this common practice by developing instruments informed by empirical evidence to guide risk assessments.²⁻⁵ These instruments consolidate numerous risk factors identified through empirical research as increasing the risk of violence or escalating its seriousness into a single measure, thus providing an efficient and comprehensive method of calculating quantitative estimates of risk (i.e., risk scores). Some researchers contend that preexisting instruments produced through prediction research on violence in general are sufficient,⁶ but most have focused their research and development efforts on the initiation,⁷ outcomes (i.e., mortality),⁸⁻¹¹ or persistence of intimate partner violence.¹²⁻¹⁷

However, if risk assessment is the cornerstone of violence prevention efforts, it must be solid. Otherwise, the constructed strategies and activities will crumble; that is, scarce resources will be misguided and those efforts will be ineffective. Recognizing the pressing need to assess risk accurately, violence prevention researchers have validated recently developed instruments. Using prospective designs to determine their predictive validity, risk scores derived from the application of those instruments are statistically associated with reports of repeated violence during a follow-up period after the initial risk assessment, with empirical results being quite promising (i.e., having a high degree of predictive accuracy using state of the art statistical methods for prediction research).^{10,14-18}

Despite the important contribution of recent validation studies, they have been limited in scope. Virtually none of them determines whether perpetrator demographic characteristics or other aspects of the risk assessment process influence derived risk scores, yet all of them focus exclusively on intimate partner violence rather than comparing it to other forms of intimate violence, such as within parent and child, sibling, or other family or family-like relationships. Moreover, validation samples have consisted exclusively

of men or women, with no study having a combination of both. These issues must be addressed to determine whether risk assessment instruments are limited or robust (i.e., empirically valid across populations of perpetrators of intimate partner violence, including men and women’s involvement in such behavior, and other forms of intimate violence). Accordingly, the present study expands the scope of validation research on assessing the risk of intimate partner violence by addressing the following issues, using a revised version of the recently validated Domestic Violence Screening Instrument (DVSI-R):^{14,18}

- Estimate the effects of age, ethnicity, and gender on risk scores.
- Estimate the effects of intimate partner violence compared with other forms of intimate violence on risk scores.
- Determine the concurrent validity of the DVSI-R by estimating the effects of alternative behavioral measures on risk scores, emphasizing violent incidents having multiple victims as an indicator of more “severe” violence.
- Determine the predictive validity of the DVSI-R by estimating the effects of DVSI-R risk scores on repeat violence, independent of type of intimate violence, other perpetrator demographic characteristics, and behavioral measures.

BACKGROUND

The analysis bearing on these issues is part of the Family Violence Risk Assessment Project, one of many developments that have occurred in the State of Connecticut since the enactment of the 1986 Family Violence Prevention and Response Act. This Act declared family and household violence a crime and formalized working relationships between the judiciary and non-profit, non-governmental community-based domestic shelter programs providing services to battered women, their children, and other victims of domestic violence on both the state and local levels. Although the Act specifically defines “family and household members,” the definition covers a wide range of relationships, which facilitates the comparison of intimate partner violence with other forms of intimate violence. Whenever police officers in Connecticut determine that such violence has occurred, the Act mandates that the person(s) suspected of its commission be arrested and charged with the appropriate crime. This mandatory arrest policy applies to misdemeanors and felonies, both of which are included in the present study.

The purpose of the Family Violence Risk Assessment

Project is to augment the clinical skills of the Family Relations Counselors (FRCs) with a validated risk assessment instrument suitable to the hectic, demanding, and time-constrained conditions under which the assessments are conducted in Connecticut. Besides handling criminal cases involving domestic violence, FRCs provide counseling and mediation services in civil cases involving other family matters. These experiences, combined with previous educational training (84% of FRCs have master's degrees or higher) yield a professional staff with significant clinical expertise. However, buttressing this expertise with a risk assessment instrument enhances the efficiency and consistency of the assessment process and standardizes the language about risk throughout the judicial system. The ultimate goal is to ensure that the recommendations made to the court and the services offered are intricately tailored to the risk of continued violence posed by the numerous perpetrators the FRCs must assess and case manage.

The risk assessment instrument selected for Connecticut by Family Services is the Domestic Violence Screening Inventory (DVSI), originally developed and validated in Colorado. Connecticut adopted the DVSI in 2002, given the suitability of this instrument for the conditions under which FRCs conduct assessments and the promising findings from the Colorado validation study.¹⁴ A pilot phase was implemented after initial training sessions on the administration of the DVSI. Initial analyses of approximately 1,500 cases assessed with the DVSI indicated that a number of items were scored as "unknown," with "unknowns" ranging from 2.5% to 12.0% of the sample across the 13 items of the initial version of the instrument. In addition, feedback from the field staff indicated the FRCs were uncomfortable with a purely statistical tool because it did not permit the infusion of their professional expertise. Given this input from the pilot phase, the coding rules for the DVSI were clarified to eliminate the "unknown" scoring. FRCs were instructed to review all sources of information available, and if no information about the presence and intensity of a risk factor was found, to score it as "no evidence." Moreover, similar to the Spouse Assault Risk Assessment Guide (SARA),¹² summary risk ratings of "imminent risk of violence toward victim" and "imminent risk of violence toward others" were added to provide an opportunity for FRCs to exercise their professional judgment in assessing risk, as they requested. With subsequent training, "unknown" scoring was virtually eliminated, and FRCs endorsed the summary risk ratings. An additional 1,800 cases were compiled from the statewide implementation and were analyzed for quality assurance.

Further modifications were made in 2003. First, some of the individual items on the DVSI were confusing, and others appeared redundant to the FRCs (e.g., history of protection orders versus history of violations of those orders). With their input, the rewording and consolidation of items was completed. These modifications were checked against the empirical results of the Colorado study to ensure that they did not undermine the predictive validity of the instrument.

METHODS

Sample selection and characteristics

Individuals arrested on family violence charges in Connecticut are required to be arraigned the next day court is in session. During the approximately 24-hour period between arrest and the initial court appearance, FRCs conduct a pre-arraignment assessment of all cases entering the system. Although not always used in practice, these assessments (in principle) are based on perpetrator interviews, a review of police reports, criminal history and protective order registry reviews, and victim interviews conducted by victim advocates. The assessments guide recommendations made by the FRCs to the court on protective orders, placement (criminal prosecution or pre-trial diversion), and treatment or services for cases remaining in the jurisdiction of Family Services for pre-trial management.

Following these normal operating procedures of processing intimate violence cases, a sample of 14,970 risk assessments was generated from September 1, 2004, through May 2, 2005. This sample covered all of the 23 judicial geographic areas in Connecticut, meaning the full population of perpetrators of intimate violence 16 years of age or older for the entire state during this period. The majority of the risk assessments involved men as perpetrators (71%), and the average age of those in the sample is 33, with about 65% of the sample being from age 22 to 44. Most of the risk assessments (66%) entailed some form of intimate partner violence, with 16% being parent and child violence and 18% involving other intimate relationships. The ethnicity of perpetrators had the following distribution: African American, 29%; Hispanic or Latino, 17%; Asian, Pacific Islander, and American Indian or Alaskan Native, less than 1%; and non-Hispanic white, 53%. The category of Asian, Pacific Islander, and American Indian or Alaskan Native was combined by the Connecticut Court Support Services Division in which Family Services is located because of the small number of cases for these ethnic categories ($n=115$), despite the resulting ethnic heterogeneity resulting from this combination.

The sample clearly was not generated through an

equal probability sampling method, thus raising questions about whether it is representative and whether generalizations are warranted. This limitation is outweighed by two important benefits from research on this agency-identified sample. First, a general population survey with an equal probability sampling design would require a huge (and costly) sample to generate enough cases of intimate violence to ensure the statistical power found with this agency-identified sample. Second, the purpose here is not to estimate the prevalence of different forms of intimate violence in the general population of Connecticut. The objective is to determine whether the DVSI-R is valid and robust in its predictions across a variety of intimate violence cases processed by Family Services and Connecticut Courts. Intimate violence cases coming to the attention of these agencies are, thus, strategic for the purposes of this analysis. The results have direct implications for addressing the specific needs of the population of intimate violence cases served by Family Services, an issue central to agency-identified samples in any jurisdiction across the United States.¹⁹

Risk assessment instrument: DVSI-R

The revised measure (DVSI-R) includes 11 items, seven of which primarily address the behavioral history of perpetrators (non-family assaults, arrests, or criminal convictions; family assaults, arrests, or criminal convictions; prior family violence intervention or treatment; violation of orders of protection or court supervision; prior or current verbal or emotional abuse; frequency of violence in past six months, and escalation of violence in past six months). The other four items pertain to substance abuse, objects used as weapons, children present during prior or current violent incidents, and employment status. The scoring categories range from zero (no evidence) to two or zero to three, depending on the item (contact the authors for more information about the DVSI-R form and coding instructions). FRCs scored each item after reviewing the available sources noted above and summed them to derive a total risk score. The potential range of risk scores is zero to 28, but the actual range in this sample is zero to 26, with a mean of 7.75 (standard deviation [SD]=5.57).

As noted above, the DVSI-R also includes two summary risk ratings for FRCs to provide their own professional assessment of the case. One of the items addresses the imminent risk of violence to the victim of the incident for which an arrest was made. The other item addresses the imminent risk to another person known to the perpetrator. Both items are scored as low, moderate, or high risk, with the following risk

distributions: victims: 35% low, 40% moderate, and 25% high; others: 53% low, 33% moderate, and 14% high. These summary risk ratings were used separately as criterion measures in testing the concurrent validity of the DVSI-R. Granted, they were scored in the context of administering this statistical instrument, yet they provide an alternative assessment of risk beyond a purely quantitative scoring procedure.

Behavioral measures

Two measures bearing on violent behavior involving intimates were used. First, the slightly more than eight-month time frame in which the 14,970 risk assessments were conducted included cases having a single assessment (75%) and cases having multiple assessments (25%). The multiple assessments represent perpetrators who were re-arrested and brought back to Family Services on intimate violence charges, with additional assessments being done. Although an admittedly crude measure, these cases were treated as repeat or recidivistic intimate violence cases and were used in the analysis of predictive validity of the DVSI-R. The 25% re-arrest and re-assessment rate is slightly lower but comparable to the overall recidivism rate (29%) in the original validation study of the DVSI conducted in Colorado.¹⁴

In principle, perpetrators entering the system earlier in the data collection time frame have a longer period in which to repeat their violence and thus be subjected to re-arrest and re-assessment. However, no evidence of this possibility was found in these data. Specifically, a month-by-month comparison of single assessment versus multiple assessment cases was conducted, and the percentages remained relatively constant (around 75% for single assessment and 25% for multiple assessment) across months. Hence, no statistically significant pattern of re-arrest and re-assessment was detected from September 1, 2004, to May 2, 2005.

Second, 38% of the 14,970 cases had multiple victims involved in the violent incidents for which arrests were made, with those cases referred to Family Services for assessment. Multiple victims represent violence that is severe in nature, extending to the same individual victimized multiple times, different individuals victimized in different violent incidents, or different individuals victimized within the context of a single violent incident. Hence, multiple victim incidents were used in testing the concurrent validity of the DVSI-R, with the following question being central: Do incidents involving multiple victims score higher on this risk assessment instrument than those involving single victims?

Analysis plan

The analytical results presented below are the product of two procedures. First, hierarchical regression analysis was conducted, in which four functional sets of variables were entered sequentially to determine the independent and joint effects of these sets in accounting for variation in DVSI-R risk scores.²⁰ Demographic characteristics of perpetrators were entered first (Model 1 in Table 1), including gender, a continuous measure of age, and ethnic minority status, with three dichotomous dummy variables for African American; Latino; and Asian, Pacific Islander or American Indian, Alaskan Native. As noted above, the latter two categories were combined because of the small number of cases ($n=115$), despite the acknowledged ethnic heterogeneity produced by this procedure. Including these ethnic minority status dummy variables in the equations estimated essentially allowed a comparison between them and non-Hispanic whites in the sample.

The first stage of the analysis allowed for the determination of whether DVSI-R risk scores vary by these demographic perpetrator characteristics. The second functional set (Model 2 in Table 1) consisted of a single dichotomous dummy variable identifying whether the violent incident involved intimate partners, including unmarried couples, spouses, divorced or separated individuals, and those having a child in common. Inclusion of this dummy variable determined whether intimate partner risk scores were significantly different from risk scores for other forms of intimate violence. The third

and fourth stages of the analysis tested the concurrent validity of the DVSI-R. Specifically, the dummy variable identifying incidents involving multiple victims was entered to estimate the empirical relation between this alternative behavioral measure and the risk scores (Model 3 in Table 1). To the extent the DVSI-R is a valid assessment of behavioral risk, one would expect a strong association between these two measures, net of other demographic perpetrator characteristics and forms of intimate violence. The two summary risk ratings by the FRCs concerning their perception of the imminent risk of violence to the victim or to others were included in the last stage of the hierarchical regression (Model 4 in Table 1). Similar to the dummy variable identifying multiple victims, to the extent the quantitative DVSI-R scores represent gradations in behavioral risk, a separate judgment about that risk based on the professional expertise of the FRCs should also be strongly and independently associated with variation in these scores.

Next, a hierarchically structured logistic regression was conducted using the same four functional sets of variables described above, plus the addition of DVSI-R risk scores (stage three in the sequence) to predict the likelihood of repeat intimate violence (a dummy variable identifying those cases involving re-arrest and re-assessment by Family Services). Hence, Table 2 presents these results as five estimated Models, not four as in Table 1. The primary purpose of this analysis was to generate propensity scores used subsequently

Table 1. Results of the hierarchical regression analysis involving the DVSI-R risk scores

Variables	Estimated models			
	1	2	3	4
	Beta	Beta	Beta	Beta
Constant	5.83 ^a	5.78 ^a	4.07 ^a	2.34 ^a
Age	0.07 ^a	0.07 ^a	0.09 ^a	0.05 ^a
Men (1), Women (0)	0.13 ^a	0.13 ^a	0.09 ^a	0.01
African American (1), Other (0)	0.02 ^a	0.02 ^a	0.01	-0.04 ^a
Latino (1), Other (0)	-0.02	-0.02	-0.02	-0.04 ^a
Asian/Pacific Islander/American Indian/ Alaskan Native(1), Other (0)	-0.03 ^a	-0.03 ^a	-0.02	0.00
Intimate partner (1), Other (0)		0.01	0.04 ^a	0.01
Multiple victims (1), Single victim (0)			0.37 ^a	0.19 ^a
Imminent risk to victSim				0.56 ^a
Imminent risk to other				0.12 ^a
R ²	0.03	0.03	0.16	0.54

^a $p \leq 0.05$

Table 2. Results of the hierarchical logistic regression analysis involving repeat intimate partner violence

Variables	Estimated models									
	1		2		3		4		5	
	β	OR	β	OR	β	OR	β	OR	β	OR
Age	-0.01 ^a	0.99	-0.01 ^a	0.99	-0.01 ^a	0.99	-0.01 ^a	0.99	-0.01 ^a	0.99
Woman (0), Men (1)	0.32 ^a	1.38	0.30 ^a	1.35	0.11 ^a	1.12	0.12 ^a	1.12	-0.05	0.96
African American (1), Other (0)	0.02	1.02	-0.03	0.97	-0.08	0.92	-0.10	0.91	-0.18 ^a	0.83
Latino (1), Other (0)	-0.33 ^a	0.72	-0.37 ^a	0.69	-0.42 ^a	0.66	-0.42 ^a	0.66	-0.51 ^a	0.60
Asian/Pacific Islander/American Indian/ Alaskan Native(1), Other (0)	-0.63 ^a	0.53	-0.58 ^a	0.56	-0.41	0.67	-0.37	0.69	0.07	1.07
Intimate partner (1), Other (0)			0.17 ^a	1.19	0.16 ^a	1.18	0.15 ^a	1.16	0.46 ^a	1.58
DVSI-R risk scores					0.14 ^a	1.14	0.13 ^a	1.14	0.08 ^a	1.08
Imminent risk to victim							0.07	1.10	0.06	1.06
Imminent risk to other							-0.07	0.93	-0.13 ^a	0.88
Multiple victims (1), Single victim (0)									2.48 ^a	11.91
Constant	-1.07 ^a		-1.18 ^a		-2.04 ^a		-2.03 ^a		-3.10 ^a	

^a $p < 0.05$ β = Beta

OR = odds ratio

in receiver operating characteristic analysis (ROC) to test the predictive validity of the DVSI-R, independent of perpetrator characteristics and forms of intimate violence. The additional contribution of the dummy variable identifying multiple victims and the summary risk ratings were also estimated in these analyses. Propensity scores are represented by the predicted probabilities of the logistic regression equations estimated at each stage in which the function sets of variables are sequentially included. Such scores have been used in previous ROC analyses of the predictive validity of risk assessment instruments.^{16,18}

Others have provided a thorough discussion of the history of ROC analysis and a description of the “statistical machinery” upon which it is based.²¹ This procedure has become the preferred analytical technique for prediction studies, especially those involving the prediction of violence in general^{22,23} or intimate violence in particular.^{16–18} It compares predicted outcomes and actual outcomes for various decision thresholds or cut-points on a prediction scale—in this case, values of the predictor variables or the derived propensity scores. “True positives” or sensitivity and “true negatives” or specificity represent agreement or correct classification between predicted and actual outcomes. “False positives” (violence was predicted but did not occur) and “false negatives” (violence was not predicted but did occur) represent disagreement or misclassification. ROC curves are constructed by plotting true positives

(sensitivity) against false positives (1–specificity) for decision thresholds varying from very stringent (no predictions of violence are rendered) to very lenient (all cases are predicted to be violent). The area under the curve (AUC) measures the accuracy of prediction. The AUC will equal 0.50 when the probability of true positives is virtually the same as false positives across all decision-making thresholds; thus, the risk assessment instrument does no better than chance in predicting behavioral outcomes. The AUC will equal 1.0 when predicted and actual outcomes are in complete agreement, meaning the accuracy of prediction is perfect.

RESULTS

The results of the hierarchical regression analysis are presented in Table 1. The relevant coefficients for each stage of the analysis are listed under Models 1 through 4, as described above. Observe that all of the perpetrator demographic characteristics (except the estimated effect for Latinos) are significantly associated with the DVSI-R risk scores. However, the effect sizes as expressed by the standardized coefficients are small, as is the variance explained ($R^2 = 0.02$). Men have slightly higher scores than women, older perpetrators have slightly higher scores than younger ones, and African American perpetrators have slightly higher scores than non-Hispanic whites, although Latinos as well as Asian, Pacific Islanders, American Indians, and Alaskan

Natives have slightly lower scores than non-Hispanic whites. The relevant coefficient in Model 2 shows that the estimated effect of the intimate partner dummy is positive, which suggests that perpetrators in such relationships tend to have slightly higher risk scores than those in other types of intimate relationships, although the effect size is again very small and statistically insignificant, with no increment to R^2 . When the severity of intimate violence is added to the analysis as indicated by multiple victim versus single victim incidents (Model 3), the explained variance ($R^2=0.16$) is significantly improved, and the estimated effect is moderately strong (Beta [β]=0.37). Adding the summary ratings of imminent risk of violence to the victim (IRV) or to others (IRO) in Model 4 also significantly improves the explained variance ($R^2=0.54$), and the effect sizes are moderate to strong ($\beta=0.12$ for IRO; $\beta=0.56$ for IRV).

Results of the ROC analyses for each of the predictor variables are presented in Table 3, with AUC coefficients reported for those variables, along with the asymptotic 95% confidence interval (CI) around those coefficients. The AUC coefficients indicate the predictive accuracy of the perpetrator demographic characteristics, the dummy variable comparing intimate partners compared to other forms of intimate relationships, DVSI-R total risk scores, imminent risk of violence ratings, and multiple versus single victim incidents. Observe that the AUC coefficients for all perpetrator demographic characteristics hover around the “line of no information” (i.e., $AUC=0.50$), which is also the case for the intimate partner dummy variable. These findings suggest that such variables are virtually unrelated to repeat intimate violence during the time frame in which the data were collected. The predictor variables bearing on behavioral risk, however, perform quite well. Specifically, the DVSI-R has a strong empirical relation with repeat intimate violence ($AUC=0.71$), as do the two summary ratings of imminent risk to victims ($AUC=0.64$) and to others ($AUC=0.61$). However, notice the predictive accuracy is greater for the DVSI-R, suggesting that this “statistical” instrument of assessing risk performs better than the assessments based on “professional judgment,” an issue discussed in more detail below. The strongest predictor of repeat intimate violence is multiple versus single victim incidents ($AUC=0.79$).

The results reported in Table 3 are consolidated more succinctly through ROC analyses involving propensity scores. Recall that these analyses estimate the relation between predicted probabilities produced by hierarchical logistic regression and the likelihood of repeat violence, as indicated by re-arrest and re-

Table 3. ROC Analysis

Variables	AUC	Asymptotic 95% CI
Perpetrator demographic characteristics		
Age	0.50	0.49, 0.51
Men (1), Women (0)	0.53 ^a	0.52, 0.54
African American (1), Others (0)	0.51	0.49, 0.52
Latino (1), Others (0)	0.48 ^a	0.47, 0.49
Asian/Pacific Islander/American Indian/ Alaskan Native(1), Other (0)	0.50	0.49, 0.51
Intimate partner (1), Other (0)	0.52 ^a	0.50, 0.53
DVSI-R total score	0.71 ^a	0.70, 0.72
Multiple victims (1), Single victim (0)	0.79 ^a	0.78, 0.80
Imminent risk to victim	0.64 ^a	0.63, 0.65
Imminent risk to other	0.61 ^a	0.59, 0.62

^a $p \leq 0.05$

ROC = receiver operating characteristic

AUC = area under the curve

CI = confidence interval

assessment by Family Services. Results of the logistic regression to generate propensity scores are not discussed here but are shown in Table 2. Table 4 lists the AUC coefficients for each of the five estimated models generating the propensity scores and the asymptotic 95% confidence interval around those coefficients. Consistent with results reported above, propensity scores derived from perpetrator demographic characteristics have limited predictive accuracy, which marginally improves when incidents are classified as involving intimate partners. However, predictive accuracy increases substantially when DVSI-R risk scores are included in the analysis, with the AUC coefficients

Table 4. ROC Analysis: propensity scores predicting repeat intimate partner violence

Propensity scores	AUC	Asymptotic 95% CI
Perpetrator demographic characteristics (PC) only	0.55 ^a	0.54, 0.56
PC and intimate partners (IP)	0.56 ^a	0.55, 0.57
PC, IP, and DVSI-R total risk scores	0.71 ^a	0.70, 0.72
PC, IP, DVSI-R, and imminent risk to victim and other (IRV & IRO)	0.71 ^a	0.70, 0.72
PC, IP, DVSI-R, IRV, IRO and multiple victims	0.84 ^a	0.83, 0.85

^a $p \leq 0.05$

AUC = area under the curve

CI = confidence interval

increasing from 0.56 to 0.71. Adding the summary risk ratings to the propensity scores does not significantly increase the coefficient (AUC=0.71), but another substantial increase is achieved by including the dummy variable identifying the incidents involving multiple victims (AUC=0.84).

ROC analyses were also conducted to determine whether the predictive accuracy of DVSI-R risk scores is consistent across forms of intimate violence. The coefficient estimating the relation between these risk scores (not propensity scores) and repeat intimate partner violence (AUC=0.70) was quite similar to the coefficients for parent and child violence (AUC=0.72) and other intimate violence (AUC=0.73).

Finally, consistency in predicting repeat intimate violence was determined by estimating the relation between the DVSI-R risk scores and ten randomly selected sub-samples of approximately 1,500 cases from this population of intimate violence perpetrators. The mean of the AUC coefficients for the ten sub-samples was 0.70, with a range of 0.68 to 0.73, suggesting that the DVSI-R consistently predicts repeat intimate violence.

DISCUSSION

The results of the analyses can be summarized around the issues posed at the outset. First, the weak estimated effects of perpetrator demographic characteristics on both the DVSI-R risk scores and the indicator of repeat intimate violence suggest those scores differ little from one age to another, between men and women, or across ethnicity groups. Thus, this risk assessment instrument is generally applicable across categories of perpetrator populations. The findings suggest the DVSI-R risk scores are not biased by the age, gender, or ethnicity of the perpetrator.

Second, concerning violence in different forms of intimate relationships, the estimated effect of the intimate partner dummy variable on risk scores is very weak, and predictive accuracy differs very little across types of intimate relationships. However, one finding about intimate partner violence should be noted: Repeat violence is significantly more likely among intimate partners compared to others in family or family-like relationships (see Table 2). No doubt, violence prevention efforts should address all forms of intimate violence, but this finding provides an empirical justification for an intensified focus on intimate partner violence, especially since it is so prevalent in this sample.

Third, results concerning the involvement of multiple victims and the summary risk ratings are supportive

of the concurrent validity of the DVSI-R. The estimated effects of these two variables on the risk scores are moderate to strong in magnitude, and the inclusion of each adds significant increments to the explained variance of their respective Models. These results support the concurrent validity of the DVSI-R since these two alternative measures of behavioral risk have such strong associations with the risk scores derived from administering the DVSI-R. Regardless, an important and unanticipated finding of this research is the significant and substantial association between multiple versus single victim incidents and repeat intimate violence. The multiple victim dummy variable had a stronger empirical relation with this behavioral outcome than any other variable incorporated in the prediction analyses. Clearly, this factor should be explored further in future developments in devising instruments to assess risk of repeat intimate violence.

Fourth, the AUC coefficients representing the predictive accuracy of the DVSI-R scores themselves or the propensity scores including the DVSI-R are significant and substantial in magnitude, independent of perpetrator characteristics and type of intimate relationship. These results support the predictive validity of this risk assessment instrument. Moreover, the DVSI-R risk scores were more strongly associated with repeat intimate violence than the summary ratings of imminent risk to the victim or to others. Such findings are not definitive evidence bearing on the debate about "statistical" versus "clinical" assessments of risk, nor do they address the important issue of how both types of information may be invaluable in rendering placement and treatment decisions about perpetrators of intimate violence.^{14,22,24} Nonetheless, the findings suggest that at least in this sample of perpetrators and using these assessment procedures, the more subjective assessment of risk based on professional judgment did not significantly improve the prediction of repeat intimate violence based on the more statistical assessment of risk (i.e., DVSI-R).

Additional results corroborating predictive validity of the DVSI-R are the consistency with which its risk scores predicted different forms of repeat intimate violence, including intimate partner violence, as well as recidivism across randomly selected sub-samples of this population of perpetrators. The AUC coefficients reported here are respectable in magnitude and well within the range of those previously reported in intimate violence risk assessment research.^{14,16-18} In fact, these coefficients are greater in magnitude than those reported in validation research involving the original DVSI (AUC=0.61).

This article began by noting how the present research

would extend previous studies of intimate violence risk assessment by determining whether risk assessment instruments are limited or robust in their prediction of behavioral risk. The focus here has been on a specific risk assessment instrument, the DVSI-R. The results show that this risk assessment is robust, not limited, with evidence of concurrent and predictive validity found independent of characteristics of perpetrators and forms of intimate violence. However, further validation is necessary using more comprehensive measures of recidivism than multiple assessments alone. Validating and demonstrating the robustness of a risk assessment instrument is only a first step in preventing violence involving intimate partners or others in family or family-like relationships. The challenge is to train professionals responsible for addressing the problem of such violence to link valid risk assessments to well-crafted strategies of supervision and treatment so that the victimized or other potential victims are protected and perpetrators are held accountable for their actions.

The authors express special appreciation to William Carbone, Executive Director of CSSD, for the support of this research.

This research was supported and conducted in collaboration with Judicial Branch, Court Support Services Division, in the State of Connecticut.

REFERENCES

- Kropp PR. Some questions regarding spousal assault risk assessment. *Violence Against Women* 2004;10:1-22.
- Campbell JC, O'Sullivan C, Roehl J, Webster D. Intimate partner violence risk assessment validation study: the RAVE study. Final report to the National Institute of Justice, Washington; 2005.
- Dutton DG, Kropp PR. A review of domestic violence risk instruments. *Trauma Violence Abuse* 2000;1:171-81.
- Gondolf EW. Batterer intervention systems: issues, outcomes, and recommendations. Thousand Oaks (CA): Sage Publications; 2002.
- Roehl J, Guertin K. Current use of dangerousness assessments in sentencing domestic violence offenders. Pacific Grove (CA): State Justice Institute; 1998.
- Hilton NZ, Harris GT, Rice ME. Predicting violence by serious wife assaulters. *J Interpers Violence* 2001;16:408-23.
- Dutton DG. A scale for measuring the propensity for abusiveness. *J Fam Violence* 1995;10:203-21.
- Campbell JC. Assessing dangerousness: violence by sexual offenders, batterers, and child abusers. Thousand Oaks (CA): Sage Publications; 1995.
- Campbell JC, Webster D, Koziol-McLain J, Block C, Campbell D, Curry MA, et al. Risk factors for femicide in abusive relationships: results from a multisite case control study. *Am J Public Health* 2003;93:1089-97.
- Goodman LA, Dutton MA, Bennett L. Predicting repeat abuse among arrested batterers: use of the danger assessment scale in the criminal justice system. *J Interpers Violence* 2000;15:63-74.
- De Becker G. *The gift of fear*. Boston: Little, Brown, and Co.; 1997.
- Kropp PR, Hart SD. The Spousal Assault Risk Assessment (SARA) guide: reliability and validity in adult male offenders. *Law Hum Behav* 2000;24:101-18.
- Gelles R. Lethality and risk assessment for family violence cases. Presented at the 4th International Conference on Children Exposed to Family Violence; 1998; San Diego, CA.
- Williams KR, Houghton A. Assessing the risk of domestic violence re-offending: a validation study. *Law Hum Behav* 2004;28:437-55.
- Weisz AN, Tolman RM, Saunders DG. Assessing the risk of severe domestic violence: the importance of survivors' predictions. *J Interpers Violence* 2000;15:75-90.
- Heckert DA, Gondolf EW. Battered women's perceptions of risk versus risk factors and instruments in predicting repeat reassault. *J Interpers Violence* 2004;19:778-800.
- Hilton NZ, Harris GT, Rice ME, Lang C, Cormier CA, Lines KJ. A brief actuarial assessment for the prediction of wife assault recidivism: the Ontario Domestic Assault Risk Assessment. *Psychol Assess* 2004;16:267-75.
- Campbell JC, O'Sullivan C, Roehl J, Webster D. Intimate Partner Violence Risk Assessment Validation Study: The RAVE Study. Final Report, National Institute of Justice; 2005.
- Dodge KA, Pettit GS, Bates JE. How the experience of early physical abuse leads children to become chronically aggressive. In: Cicchetti D, Toth SL, editors. Rochester symposium on developmental psychopathology. Vol. 8: Developmental perspectives on trauma: theory, research, and intervention. Rochester (NY): University of Rochester Press; 1997. p. 263-88.
- Cohen J, Cohen P. *Applied multiple regression/correlation analysis for the behavioral sciences*, 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates, Publishers; 1983.
- Swets JA, Dawes RM, Monahan J. Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest* 2000;1:1-26.
- Monahan J, Steadman HJ, Silver E, Appelbaum PS, Clark Robbins P, Mulvey EP, et al. Rethinking risk assessment: the MacArthur study of mental disorder and violence. New York: Oxford University Press; 2001.
- Rice ME, Harris GT. Violent recidivism: assessing predictive validity. *J Consult Clin Psychol* 1995;63:737-48.
- Litwack TR, Schlesinger LB. Dangerousness risk assessments: research, legal, and clinical considerations. In: Hess A, Weiner I, editors. *The handbook of forensic psychology*. New York: John Wiley & Sons; 1999. p. 171-217.