#15 Risk Assessment in Intimate Partner Violence:
A Systematic Review of Contemporary Approaches

Tonia L. Nicholls, Michelle M. Pritchard, Kim A. Reeves, and Ed Hilterman

Introduction. Intimate partner violence (IPV) has profound and widespread health and economic implications at an individual, familial, and societal level. Violence risk assessment measures offer an evidence-informed approach to ascertain the degree of threat an abuser poses. Risk assessments are an essential means of informing professionals and victims alike regarding the nature and intensity of services required to help keep a victim safe as well as a transparent and defensible indicator of the rationale for intervening with an abuser (e.g., sentencing, probation conditions, required treatment). Violence risk assessment measures have been in existence for several decades and there has been a proliferation of IPV specific measures developed. However, there is little agreement in the literature with regard to the most appropriate *approach* (actuarial, structured clinical judgment) nor which specific *measure* has the strongest empirical validation behind it, leaving clinicians and policy makers with little clear guidance.

Methods. This paper summarizes the state of knowledge regarding risk assessment for IPV through a systematic examination of all English publications from westernized nations from 1990 to 2011. Three search engines (PsychINFO, Science Direct, and Social Science Index) identified 3361 potentially relevant articles. After dropping duplicates, examining titles and abstracts and removing articles that did not explicitly examine risk assessments for IPV we were left with a total of 39 articles. Our specific objectives were to: (a) identify all IPV risk assessment measures and relevant approaches (e.g., pilot tools; women's appraisals of their partner's risk, the application of non-IPV specific measures to IPV populations/risk); (b) describe briefly the purpose, development and use of the various risk assessment approaches; and (c) report the state of the validation of the various measures and evaluate the psychometric properties of these diverse approaches (inter-rater reliability, convergent validity, predictive validity).

Results. The review yielded studies reporting on the validity and reliability of eight IPV specific actuarial instruments and three general actuarial risk assessment measures (see Table). The range of area under the curve (AUC) values reported for the validity of the Ontario Domestic Assault Risk Assessment (ODARA; Hilton, et al., 2004) predicting recidivism was good to excellent (0.64 – 0.77). The single study that reported on the Domestic Violence Risk Appraisal Guide (DVRAG; Hilton, Harris, Rice, Houghton, & Eke, 2008) reported an AUC = 0.70 ($p < .001$). The inter-rater reliability for both instruments was excellent. The Domestic Violence Screening Inventory (DVSI, Williams & Houghton, 2004) and Domestic Violence Screening Inventory – Revised (DVSI-R; Williams & Grant, 2006) were found to be good predictors of new family violence incidents and IPV recurrence (AUC range 0.61 – 0.71). Three studies examined the Psychopathy Checklist – Revised (PCL-R; Hare, 1991, 2003) and Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993; Quinsey, Harris, Rice, & Cormier, 2006), neither of which are IPV specific, reporting AUCs ranging from 0.66 – 0.71 and 0.67 - 0.75, respectively. The Level of Service Inventory – Revised (LSI-R; Andrews & Bonta, 1995, 2000, 2001) and Level of Service Inventory – Ontario Revision (LSI-OR; Andrews, Bonta, & Wormith, 1995) were discussed in four articles, reporting two AUC values of 0.50 and 0.73, both of which were predicting IPV recidivism.

Two structured professional judgment instruments were included in the review, the Spousal Assault Risk Assessment guide (SARA; Kropp, Hart, Webster, & Eaves, 1994, 1995, 1999, 2008) and the Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER; Kropp, Hart, & Belfrage, 2005). The SARA research reports nine AUCs ranging from 0.52-0.65. The interrater reliability (IRR) for the SARA was excellent for total scores, good for the summary risk ratings, and poor for the critical items. Although neither of the articles examining the B-SAFER reported the predictive validity of the instrument one did report the IRR based on 12 cases with a mean interclass coefficient (ICC) of 0.57.

The current literature for the Danger Assessment (DA; Campbell, 1986; Campbell, Webster, & Glass, 2009) and unstructured Victim Appraisals do not provide a clear picture of the validity and reliability of these two approaches to ascertaining an abuser's future risk of IPV. The DA has the largest body of literature behind it, but there are limitations in the research that inhibit a clear determination of the psychometric properties of the measure, thus far. Victim

appraisals of the risk of future IPV show some evidence of predictive accuracy, even exhibiting greater predictive validity than some risk assessment instruments (Heckert & Gondolf, 2004) and adding significantly to regression models (Weisz et al., 2000); however, further research is needed to determine the best means with which to collect the victim's reports and determining the conditions (e.g., stalking) and characteristics of victims that should be considered (e.g., PTSD, substance use).

Overall, the literature reveals moderate postdictive/predictive accuracy across measures with little evidence to support one as being highly superior to others, particularly given the heterogeneity of perpetrators and victims, study limitations, and the small body of empirical literature to date. Although lethal assault (which might reflect femicide, filicide, and/or familicide) is of greatest concern, the necessary evidentiary basis for recommending a measure to assess for risk of lethal IPV violence is highly limited (also see Bowen, 2011; Guo & Harstall, 2008; Hart & Watt, 2008).

Discussion. Several themes emerged when we examined the synthesized literature: (1) There is a relatively small body of empirical evidence evaluating IPV violence risk assessment measures. (2) The need for continued advancements in the methodological rigor of the research including prospective studies, research that compares multiple measures within single studies, and research that uses large samples and appropriate outcome indicators. Particularly challenging is studying the predictive validity of measures intended to predict femicide. There also is a need for increased consistency in adhering to measure guidelines in evaluation studies (i.e., coding the measures in the manner recommended), obtaining criterion data from multiple and reliable sources and using outcome data that matches the intended use of the risk assessment measures. (3) A need to extend the investigation of the validity and reliability beyond North American borders and expand the cross-validation research to diverse samples (e.g., Gay- Lesbian – Bisexual-Transgendered; male victims/female perpetrators). (4) A particularly exciting development in IPV risk assessment research is evidence that risk assessments can serve to reduce risk levels (Belfrage et al., 2011).

In terms of clinical implications, the review demonstrates the considerable promise of several IPV risk assessment measures but generally reveals modest postdictive/predictive accuracy for most measures. Limited evidence for the superiority of actuarial vs. SPJ measures

was evident. Similarly, IPVspecific risk assessment measures were not found to consistently outperform general violence risk assessment measures; however, we would recommend considerable caution in interpreting this finding given the small number of studies examining non-IPV measures and the fact that only one of those studies actually used IPV recidivism as the outcome criterion. We would assert this may largely be a reflection of poor study designs and procedures (e.g., not using the SARA in the preferred manner), particularly given meta-analyses in the risk assessment field have found context/outcome specific measure to have an advantage over non-population/offence specific measures (Singh, Grann, & Fazel, 2011). Given the challenges in comparing across studies and the heterogeneity of partner abusers it seems premature to recommend one preferred assessment measure/approach to clinicians. Victim appraisals, while the research has a considerable ways to go, were found to have clinical relevance. However, preliminary evidence suggests that clinicians may want to be particularly cautious when working with some sub-groups when taking into account victims' perceptions (e.g., PTSD symptoms, substance use, stalking and severe abuse experienced) and supplement the woman's input with an additional structured assessment.

When clinicians and administrators are faced with the challenge of determining which measure(s) to use to assess risk of IPV they should carefully consider the *purpose* of the assessment (Heilbrun, 2009). Assessors also should take into account the *context*, *setting*, and *resources* when evaluating which measure best suits their needs. For instance, some structured professional judgment measures (e.g., SARA and PCL-R) may be more resource intensive than most actuarial measures making them inappropriate for certain circumstances (e.g., police responders; also see Coid et al., 2009)[1]. In addition, many of the measures considered here require extensive professional training and *expertise* of the evaluator (e.g., PCL-R). Finally, consideration must be given to the characteristics of the *population* to be assessed (e.g., age, gender, ethnicity, socio-economic status) and the extent to which a measure has been cross-validated in similar samples is required (Heilbrun,2009).When validation research and empirical evidence is limited evaluators should be particularly cautious in the interpretation of results and should make the limitations clearly evident to stakeholders in their risk assessment documentation and communications. It is also important that measures are used as intended (e.g.

---

[1] However, note Miles et al. (in press) recently concluded an SPJ measure was not unduly resource intensive

administering measures and conducting interviews, reviewing criminal records and clinical files); therefore, if the setting and context does not lend itself to accessing the required information and sufficient time to complete each recommended component of an assessment the measure may not yield accurate information. In particular, assessors want to be clear about the outcome of concern (verbal abuse, physical abuse, severe violence, stalking, femicide?) and knowledgeable about relevant base rates (Heilbrun, 2009).

Limitations. This review is intended to provide both researchers and clinicians with a comprehensive review of the state of the IPV risk assessment field; as such we were inclusive of studies by not excluding published findings based on study quality but rather describing the limitations of available research. Reflecting that objective and the heterogeneous nature of the research available we have provided a narrative review and did not provide common effect sizes via a meta-analysis. We also included only studies published in English from Westernized nations, published in peer-reviewed journals thereby limiting the generalizability of our conclusions. Based on the available literature, we are also unable to provide guidance on the clinical relevance and utility of these instruments with female perpetrators, male victims, and in same-sex relationships due to the lack of studies using relevant populations. The field at present is limited by the small number of studies that have addressed each instrument and due to diverse methodological limitations. The extant literature lacks prospective, longitudinal studies, studies comparing multiple instruments, studies that reflect the intended outcome and/or that utilize multiple sources of data to code outcome criteria, and studies that code the measures in the manner intended/include all items, thereby making conclusions tentative. For instance many studies of the SARA relied on file reviews in the absence of an interview. Also, the criterion variables either did not match the intended use of the instrument (e.g. the Danger Assessment was used to measure re-abuse) or relied on criminal records or self report, but rarely both. In recognizing these limitations we hope to guide future research. For the same reason we did not perform a meta analysis of the heterogeneous literature.

Conclusion. There is considerable room for further IPV risk assessment research. In particular, studies examining the incremental validity of using IPV specific variables or measures once taking into account general risk predictors (incremental validity) , prospective studies, and

rigorous designs comparing multiple measures in single studies and using relevant criterion variables are required.

About the Authors

Dr. Tonia Nicholls obtained a Ph.D. with a specialization in Law and Forensic Psychology from Simon Fraser University in 2002. She is Associate Professor, Department of Psychiatry, University of British Columbia and Senior Research Fellow at the Forensic Psychiatric Services Commission, BC Mental Health & Addiction Services. She currently holds operating grants from the Canadian Institutes of Health Research and the Mental Health Commission of Canada. Her scholarly work earned her three "Brain Star" awards from the Institute of Neurosciences, Mental Health, and Addictions (Canadian Institutes of Health Research), the American Psychological Association Award for Distinguished Professional Contributions, and the Canadian Psychological Association President's New Researcher Award. Dr. Nicholls' research examines the intersections of law and mental health related to the provision of services to marginalized populations (homeless mentally ill, forensic patients, persons in conflict with the law). She is particularly interested in the assessment and treatment of violence and criminality and the development and implementation of evidence-based practices that bridge the research-clinical practice chasm. Her publications have covered diverse topics including inpatient aggression in psychiatric settings, women in conflict with the law, psychopathy, the potential contribution of protective factors to the violence risk assessment field, and intimate partner abuse. She has been particularly active in knowledge exchange, publishing several manuals to inform evidence-based practice and being actively engaged in large-scale implementations and evaluations including measures to inform: violence risk assessments (Short-Term Assessment of Risk and Treatability (START), Webster, Martin, Brink, Nicholls, & Middleton, 2004; Webster, Martin, Brink, Nicholls, & Desmarais, 2009; START: Adolescent Version abbreviated manual, Nicholls, Viljoen, Cruise, Desmarais, & Webster, 2010; START:AV full manual, Viljoen, Nicholls, Cruise, Desmarais, & Webster, in prep), mental health screening in correctional

settings (Jail Screening Assessment Tool, Nicholls, Roesch, Olley, Ogloff, & Hemphill, 2005) and assessments and safety planning with women coping with or escaping intimate partner abuse (Decision-making In Abusive Relationships Interview, Nicholls, Hilterman, Tengstrom, 2010). In 2007, she received a Michael Smith Foundation for Health Research Career Scholar award (2007-2013). She also presently holds a Canadian Institutes of Health Research New Investigator salary award (2011-2016).

Contact: tnichols@forensic.bc.ca

Michelle M. Pritchard received her BSc in psychology from Portland State University in Portland, Oregon. Ms. Pritchard is currently completing a Masters in Public Health degree at Simon Fraser University, British Columbia, Canada. Her final Masters project is examining Intimate Partner Violence risk assessment instruments from a public health perspective. She is also a project coordinator with the Forensic Psychiatric Services Commission, BC Mental Health and Addiction Services in Port Coquitlam, BC as well as in the Department of Psychiatry, Faculty of Medicine, University of British Columbia, Vancouver, Canada.

Contact: MPritchard@forensic.bc.ca

Kim A. Reeves is a doctoral student of Clinical Psychology with a specialization in Forensics at Simon Fraser University in Burnaby, British Columbia, Canada. Ms. Reeves received her BA in Psychology with a minor in Criminal Justice from The College of New Jersey in Ewing, NJ. She then earned her Masters of Arts in Forensic Psychology from John Jay College of Criminal Justice. Ms. Reeves worked as the Research Coordinator at the Bellevue/NYU Program for Survivors of Torture for one and a half years. During that time, she was part of the research team who investigated the Expedited Removal and Asylum Seeking process under the mandate of the United States Commission on International Religious Freedom. She decided to continue with her education and received her Masters of Arts in Clinical Psychology at SFU where she is now pursing her Doctoral degree.

Ms. Reeves has been involved in a variety of research pertaining to torture, PTSD, and policy both in the US and abroad. During her time at John Jay and now at SFU, Ms. Reeves has focused on the study of risk assessment especially in the context of intimate partner violence and stalking. She has received several grants to support her research including being one of the inaugural winners of the Social Sciences and Humanities Research Council of Canada Vanier Canada Graduate Scholarship.

While studying at SFU, Ms. Reeves completed a practicum at the Vancouver Police Department in the Domestic Violence and Criminal Harassment Unit. She consulted with the police and helped them implement risk assessments in their daily practice and assisted in developing management plans for offenders and safety plans for victims. This work has informed her research on risk assessment as she firmly believes in the scientist-practitioner model in psychology.

Ed Hilterman is a Sociologist, director of a small Consultancy specialized in applied research in the justice sector in Barcelona, Spain and Consultant at the Open University in Barcelona, Spain. Under the supervision of Prof. Dr. Chijs van Nieuwenhuizen, from Tilburg University, he is currently working on his PhD on risk assessment in juvenile offenders.

In both the Netherlands and in Spain he has worked in the area of risk assessment and risk management. In the Netherlands he has worked as researcher in the TBS system and in Spain in the juvenile justice system in Catalonia. Since 2008 he has had his own Consultancy in Barcelona.

He has specialized in risk assessment and has translated several risk assessment tools into Dutch and Spanish. In collaboration with Tonia Nicholls and Anders Tengström he also worked on the development of the Decision-making in Abusive Relationship Interview (DIARI), a structured needs assessment tool for victims of Intimate Partner Violence (IPV).

Contact: ehilterman@justamesura.com

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
Study (full reference)	*N*	**Sample Characteristics** (e.g., married or dating, gender, race, SES, religion)	**Method and Design** (Cross-sectional or longitudinal; where and how data collected; time period considered; self or other-report)	**Results** (including measures used)
Dutton, D.G. (1995). A scale for measuring propensity for abusiveness. *Journal of Family Violence, 10,* 203-221. **Objective.** Development and validation of a brief self-report scale (Propensity for Abusiveness Scale) to assess propensity for male abusiveness of a female partner in intimate relationships.	140 men in treatment for wife assault and 63 of their female partners. An additional 44 demographically matched men and 33 of their partners were assessed. The authors administered the 29-item PAS	*Assaultive sample.* **Age.** *M* = 35, range 17-65 years of age. **Ethnicity.** Not reported. **Education.** *M* = Grade 12. **Employment/Income.** Modal family income $34,000 (37% self-identified as white collar). **Relationship.** 56% were living with the assault victim. **IPV History.** The mean self-reported CTS scale score for physical assaultiveness against their partner in the	**Design.** Test development. **Measure.** Propensity for Abusiveness Scale. **Development.** The PAS was composed of components of other scales, which had the strongest correlations with female partner reported abusiveness. The authors then determined the item-whole correlations to the selected scales, and factor analyzed the new scale. The total score is composed of the sum of the unit weighted items. The scores for the first twelve items have a 1 to 5 range; items 13-21 have a 1 to 4 range; items 23 to 29 have a zero to 3 range. The resulting scale had a mean	*Assaultive Sample*. **Total Scores.** *M*=59.2 (*SD* = 17.1) significantly different *t* = 3.41, *df* = 2 158 p < .05). *Control sample.* **Total Scores**. *M*=44.7 (*SD*=11.7) **Inter-rater Reliability.** Not reported. **Internal Consistency.** Coefficient Alpha is .92. **Convergent Validity.** The total scores on the PAS correlated with other measures as follows: BPO (*r* = .64), MAI (*r* = .63); TSC (*r* = .56); EMBU Rejection by Father (*r* = .74), Rejection by Mother (*r* = .50), Warmth by Father (*r* =-.53) Warmth by Mother (*r* = -.36); RSQ Fearful Attachment (*r* = .48), Preoccupied Attachment (*r* = .35), Dismissing (*r* = .09), Secure (*r* = -.30). **Criterion Validity.** The PAS correlated *r*=.51 with the Dominance/Isolation scale of the PMWI and *r*=.47 with the Emotional Abuse scale of the PMWI. **Predictive Validity.** The PAS generated a

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
	to a group of 60 assaultive men, 36 of their female partners, 40 control men and 40 of their female partners for cross validation.	prior year was 5.87 (*SD*=6.22). *Control sample.* **Age.** *M* = 35 (range 19-45). **Ethnicity.** Not reported. **Education.** *M* = Grade 12. **Employment/Income.** Modal family income $35,000. **Relationship.** 65% were married. **IPV History.** The mean self-reported CTS scale score for physical assaultiveness against their partner in the prior year was 1.34 (*SD*=3.4). **Setting.** The Assaultive group consisted of both court-referred and self-referred males and their partners. **Cross Validation**	score of 49.3 (*SD* = 16.8) with a range of 12 to 95 in the population sampled. **Administration.** All assessments were completed during the first 4 weeks of a 16 week treatment program for wife assault. Women partners completed assessments privately and independently of their partner. **Criterion Scale.** Psychological Maltreatment of Women Inventory (PMWI). **Other measures.** Borderline Personality Organization (BPO); the Multidimensional Anger Inventory (MAI); Trauma Symptom Checklist (TSC-33); Recollections of Early Childrearing (EMBU); the Relationship Scales Questionnaire. **Limitations.** Discriminant validity has yet to be reported for this measure.	significant discriminant function that correctly classified 80.0% of men on the Dominance/Isolation factor and 84.4% on the Emotional Abuse factor. The PAS corrected for social desirability performed similarly to the uncorrected version both for convergent and criterion validity. **Cross-Validation.** Internal consistency was .88. The PAS correctly identified 82.2% of men on Dominance/Isolation and 81.3% of men on Emotional Abuse. The PAS correlated *r*=.52 with Dominance/Isolation and *r*=.45 with Emotional Abuse.

Study	N	Sample Characteristics	Method and Design	Results
		Sample. Demographics were similar to the original study samples.	In addition, it has only been correlated with partner's reports of abuse and not compared to future abuse.	
McFarlane, J., Parker, B., & Soeken, K. (1995). Abuse During Pregnancy - Frequency, Severity, Perpetrator, and Risk-Factors of Homicide. *Public Health Nursing, 12*(5), 284-289. **Objective.** To determine the frequency, severity, and perpetrator of abuse during pregnancy and the occurrence of risk factors of homicide.	1203 Pregnant women	**Age.** Most women (70%) were between the ages of 20 and 29 years, and 29.6% between 13 and 19 years. **Ethnicity.** Ethnically stratified cohort - 34.4% African-American, 34.2% Hispanic, and 31.3% Anglo- American. **Education.** Not mentioned **Employment/Income.** 94% below poverty **Setting.** Pregnant women recruited from public prenatal clinics in Houston, Texas and Baltimore, Maryland; **Relationship.** 36.4% married; **IPV History.** Of the	**Design.** Stratified, prospective cohort study. **Measure**. The DA, consisting of 14 items with yes/no response format. **Administration.** Women were assessed for abuse at first prenatal appointment and then two more times during pregnancy. If the woman reported abuse, the DA was administered. **Other measures.** All women were administered the Abuse Assessment Screen, Conflicts Tactics Scale (CTS), the Index of Spouse Abuse (ISA). **Limitations.** Women were recruited from public prenatal clinics limiting the generalizability of the study; In this article the DA is said to have 14 questions where as the DA	**Recidivism**. At baseline, 24.3% reported past year physical or sexual abuse. Of the women that had not experienced abuse at baseline, 5% reported the start of abuse at one of the follow-ups. **Total Scores**. Women abused during pregnancy had significantly higher scores on all instruments and more risk factors for homicide when compared with women abused prior to but not during pregnancy. **Inter-rater Reliability.** Not reported. **Internal Consistency.** α = 0.84. **Validity.** DA correlated with the ISA-P: *r* = .79. Scores on the DA were significantly different between the women who hadn't experienced abuse (0.65), women abused during past year but not while pregnant (3.32), and those abused while pregnant (3.82). The higher DA scores during pregnancy could suggest that there is an increased risk for homicide resulting from IPV during this time.

Study	N	Sample Characteristics	Method and Design	Results
		1203 pregnant women surveyed, 28.7% reported abuse at T1 or T2. 18% of the African American women, 13% of the Hispanic women, and 17% of the Caucasian women reported abuse during pregnancy.	has 15 items (Campbell, 1986). It is unknown which item was removed from DA and why it was removed.	
McFarlane, J., Soeken, K., Campbell, J., Parker, B., Reel, S., & Silva, C. (1998). Severity of abuse to pregnant women and associated gun access of the perpetrator, *Public Health Nursing, 15*(3): 201-206. **Objective.** To investigate the relationship between abuse during pregnancy and gun access of	199 pregnant, abused women	**Age.** Range = 14-42, *M* = 23.2; 29.6% were 19 or younger **Ethnicity.** Ethnically stratified sample: 35% African American women, 32% non-Hispanic Anglo-American women, 33% Hispanic women **Education.** Range: 2-26 years; *M* = 10.4 years. **Employment/Income.** All participants fell below the poverty level as determined by Women Infants Children (WIC)	**Design.** Prospective, ethnically stratified cohort design. **Measure.** Danger Assessment Scale (DA) **Administration.** The participants were interviewed in either Spanish or English and reported on the past year. **Other Measures.** Index of Spouse Abuse (ISA); The Severity of Violence Against Women Scale (SVAWS). **Limitations.** All information was collected as self-report; no control	**Recidivism.** Unknown. **Total Scores.** There was a significant difference in mean score of the DA between women whose partner had access to a gun (*M* = 7.0) and those that did not (*M* = 5.7; *t* = 3.10; *p* = .002). **Inter-rater Reliability.** Not reported. **Internal Consistency.** α = 0.72. **Validity.** DA correlated with the ISA - Physical items: *r* = 0.75. 41% of the participants' partners owned or had access to a hand gun. More specifically, 31.8% of Hispanic women, 44.4% of Anglo-American women, and 47.1% of African-American women reported that their partners had gun access. There was not a significant difference as to gun access based on age of the woman or her

Study	N	Sample Characteristics	Method and Design	Results
abuser.		standards. **Setting.** Public health clinic. **Relationship.** 47.7% were married or co-habitating with the abusive partner at time of first interview; 29.6% were separated from abusive partner; 19.1% were not living with abusive partner**.**	group; since it is a cross-sectional study temporality cannot be established.	partner.
Goodman, L. A., Dutton, M. A., & Bennett, L. (2000). Predicting repeat abuse among arrested batterers - Use of the danger assessment scale in the criminal justice system. *Journal of Interpersonal Violence,* 15(1), 63-74. **Objective.** To test the ability of the Danger Assessment Scale to predict short	49 female victims of IPV	**Age.** Participants were over 18 years. **Ethnicity.** 90% African American **Education.** Not reported. **Employment/Income.** 51% of participants were employed either full or part time. **Setting.** Women were recruited from the Domestic Violence Intake Center at Superior Court in Washington DC after they had been	**Design.** Prospective study (3 month follow-up). **Measure.** Danger Assessment Scale (DA) **Administration.** The participants completed the measures upon recruitment. Approximately 12 weeks after initial intake the research team made a follow up phone call in order to assess further threat and abuse. **Other measures.** Conflict Tactics Scale (CTS2) **Limitations.** Only 53% of the original 92 participants	**Recidivism.** 22% of participants reported at least one episode of abuse 3 months after the arrest of the perpetrator. **Total Scores**. Mean DA and CTS2 scores for women that experienced reabuse were 7.24 (SD=2.56) and 32.97 (SD=24.67), respectively. **Inter-rater reliability.** Not reported. **Internal Consistency.** Not reported. **Validity.** A score of one standard deviation higher on the DA was related to a fourfold increase (OR=4.18) in likelihood of re abuse within 3 months. A score one standard deviation higher on the CTS2 predicted a 2.77 increase in likelihood of reabuse. When scores from both DA and CTS2 were simultaneously taken into account only the DA was significantly associated with reabuse in the

Study	N	Sample Characteristics	Method and Design	Results
term repeat abuse by batterers in the CJS;		assaulted by an intimate partner. **Relationship.** 45% of batterers were ex-boyfriends and 40% were current boyfriends **IPV History.** All participants had at least one episode of IPV.	were available for follow-up (*N* = 49). Results may not be generalizeable to the general population. Reabuse, the outcome variable was re assessed with a single question that included both threats and reassault.	following three months.
Kropp, P. R. and Hart, S. D. (2000). The Spousal Assault Risk Assessment (SARA) Guide: Reliability and validity in adult male offenders. *Law and Human Behavior*, *24*(1), 101-118. **Objective**. To evaluate the reliability and validity of judgments concerning violence risk made	Sample 1: 1671 male probationer s; Sample 2 1010 male inmates.	**Age.** Median age = 32 years **Ethnicity.** 80% Caucasian, 15% Native Indian **Education.** Not reported. **Employment/Income.** Not reported. **Setting.** Sample 1: was recruited in British Columbia and consisted of probationers arrested for IPV. Sample 2 was made up of inmates serving more than 2 years in Canadian	**Design.** Postdictive. **Measure**. SARA. **Administration.** SARA scores were made by a range of corrections, mental health, and research professionals. The SARA was filled out based on an interview with the participant and a file review. **Outcome.** Severe IPV recidivism defined as: Threatened to kill, threatened with weapon, kicked, bit, or hit with fist, hit or tried to hit with something, forced sex,	**Recidivism.** Not applicable. **Total Scores**. Total SARA scores did not significantly differ between recidivists (17.69) and non recidivists (15.68) (p < 0.68). **Inter-rater reliability.** Individual items scored on a 3 point scale Median Intraclass Coefficient (ICC) = .65 with a range between .45 -.86. Total score ICC was reported as .84 (*p* , .05). Summary risk rating IRR was reported as .63 (*p* < .05). **Internal Consistency.** Item homogeneity was reported as a Mean Intraclass Coefficient (MIC) of 0.15. **Validity.** When comparing summary risk ratings to recidivism an AUC value of .70 (*SE* = .06) was reported.. The authors also reported on the validity of the

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
using the SARA in adult male offenders.		federal prisons that had **Relationship/Abuse.** Most of the participants with an IPV history had been married or lived common-law in the past.	choked/ strangled, beat up, used weapon. Other measures. Psychopath Checklist Short Version (PCL:SV), General Information on Recidivism Scale (GSIR), Violence Risk Appraisal Guide (VRAG). **Limitations.**	SARA total score compared to the PCL:SV ($r = .43$, $p < .001$), GSIR ($r = -.07$, ns) and VRAG. ($r = .29$, ns),
Weisz, A. N., Tolman, R. M. & Saunders, D.G. (2000). Assessing the risk of severe domestic violence: The importance of survivors' predictions. *Journal of Interpersonal Violence, 15*(1), 75-90. **Objective**: To determine if survivor's of IPV can more accurately predict severe IPV than a	177 female partners of batterers	**Age.** $M = 30.67$ **Ethnicity.** Caucasian (71.6 %), African American (27.7%). **Education.** Not reported. **Employment/Income.** 58.5% employed full-time; 13.6% employed part-time. **Setting.** Female partners of IPV perpetrators found guilty of misdemeanors. **Relationship/Abuse.** 61% living with abuser.	**Design.** Secondary analysis of a prospective study (4 month follow-up) (Harrell, 1991). **Measure.** A single question to assess victim report rated on a 10 point scale: "How likely would you say it is that your partner will become violent with you during a dispute in the next year?" Twelve items from the DA were also used. **Administration.** Researchers interviewed the participants twice. The first time was just after the case disposition,	**Recidivism.** Not reported. **Total Scores.** Not applicable for victim report; DA total scores not reported. **Inter-rater reliability**. Not applicable for victim report; DA IRR not given **Internal consistency**. Not reported. **Predictive Validity**. Bivariate analysis indicated a significant association between survivors' predictions and the recurrence of severe violence in the 4 month follow-up period [X^2 (9, N = 177) = 34.30, $p = .000$]. When survivors' predictions were added to a multivariate model of significant risk factors for recurrence of severe violence, the R^2 increased significantly from .15 to .25. A regression analysis was run that included indicators from the DA. When survivors' predictions were added to this model the R^2

Study	N	Sample Characteristics	Method and Design	Results
statistical approach, or if both methods in combination make a more accurate prediction.			approximately 17.5 weeks after the focal incident. The follow-up was 4 months after T1. **Outcome.** Victim report of severe violence between interviews. Severe violence defined as: threatened to kill, threatened with weapon, kicked, bit, or hit with fist, hit or tried to hit with something, forced sex, choke/strangle, beat up, used weapon. **Other Measures.** Expanded version of the CTS was used to assess outcome. **Limitations.** The victim prediction question only addresses violence in the context of a dispute; The victim's prediction question is asked at the end of an interview in which the victim has thoroughly reported on her IPV history; The victim report question asked about violence in the next year	increased significantly from .09 ($F = 1.39$, *ns*) to .22 ($F = 3.52$, $p < .001$).

Study	N	Sample Characteristics	Method and Design	Results
			however the follow-up occurred only 4 months later; The study did not have the data for all of the DA questions.	
Dutton, D.G., Landolt, M.A., Starzomski, A., & Bodnarchuk, M. (2001) Validation of the Propensity for Abusiveness Scale in diverse male populations. *Journal of Family Violence, 16*(1), 59-73. **Objective.** Cross-validation of the Propensity for Abuse Scale	Three groups of men (*N*= 363): non-violent clinical outpatient population (*N*=50), male college students (*N*=149), a sample of gay men in long-term relationships (N=104), and a new sample of known assaultive men (*N*=60). Only men with	***Clinical sample* Age.** *M*=34 (*SD*=5.1). **Ethnicity.** Not reported. **Education.** *M*=Grade 12. **Employment/Income**. Modal income $35,000 (*SD*=$9,600). **Setting.** Solicited from 14 outpatient clinics (collected from 12 clinics) located in the Vancouver area. ***Spousal Assault Treatment sample.*** See description in Dutton (1995). ***Gay sample.* Age.** *M*=34 (*SD*=5.8). **Ethnicity.** Not reported.. **Education.** *M* = 2	**Design.** Convergent and Criterion Validity. **Measure.** Propensity for Abusiveness Scale. **Administration.** Not reported.. **Outcome.** Responses on the criterion measures of abuse for all samples (see Other measures). **Other measures.** Psychological Maltreatment of Women Inventory (PMWI; Wives' Reports); Psychological Maltreatment Inventory (PMI); Conflict Tactics Scale (CTS); The Severity of Violence Against Women Scale (SVAWS), or the Severity of Violence Against Men Scale (SVAMS). **Limitations.** The	**Total Scores.** Assaultive sample *M*=66.1 (*SD*=33.7); Gay sample *M*=46.0 (*SD*=12.7); Clinical sample *M*=52.4 (*SD*=21.4); and College sample *M*=47.2 (SD=18.2). These scores are similar to the original self-referred assaultive sample (*M*=62.2, *SD*=17.1) and the original control group (*M*=44.7, *SD*=11.7). **Inter-rater Reliability.** Not reported.. **Internal Consistency.** Not reported. **Criterion Validity**. The Gay sample was significantly correlated with the Wives' Reports on the PMWI (*r* =.50 Dominance/Isolation and *r* =.36 Emotional Abuse), but not with the CTS. The Clinical sample was not correlated with the PMWI. The College sample was significantly correlated on the PMWI (*r* =.56 Dominance/Isolation, *r* =.40 Emotional Abuse) and *r* = .41 with the CTS. Lastly, the new assaultive group was significantly correlated with the PMWI (*r* = .64 Dominance/Isolation, *r* =.52 Emotional Abuse) and with the Marshall (*r* =.76 Threats, *r* =.30 Severe Violence). The PAS significantly predicted Emotional Abuse, use of Dominance Isolation, Physical Abuse, and the use of threats to kill or injure. The PAS was

Study	N	Sample Characteristics	Method and Design	Results
	intimate anger problems or alcoholism were rejected (not applicable to the known assaultive sample).	years University. **Employment/Income**. Modal income $60,000 (*SD*=$12,000). **Setting.** Solicited from advertisements in newspapers. *College sample.* **Age.** *M*=19 (*SD*=.96). **Ethnicity.** Not mentioned. **Education.** Comparable to the Gay sample in educational attainment. **Setting.** Solicited from college classes. *Total Sample.* **Ethnicity.** Predominantly Caucasian Euro Canadians; 20% non-Caucasian and from a variety of cultural backgrounds. **Relationship.** In an	administration of the measures was not detailed in the article so it was unclear what source of information was used to complete the measures. The internal consistency of the measure was not reported nor was there any information on reliability.	highly predictive of emotional abusiveness in all samples.

Study	N	Sample Characteristics	Method and Design	Results
		intimate relationship of at least 6 months duration. **IPV History.** Not reported..		
Hilton, Z., Harris, G.T., & Rice, M.E. (2001). Predicting violence by serious wife assaulters. *Journal of Interpersonal Violence*, 16(5), 408-423. DOI: 10.1177/08862600 1016005002. **Objectives**. The associations of a measure of psychopathy (PCL-R), other predictor variables with a hypothesized or defined relation with psychopathy, and violent recidivism are examined. Also are serious wife	508 male offenders, a sub sample of 88 offenders were wife assaulters.	**Age**. Wife assaulters were 37.23 (*SD* = 11.75) years old, while the other offenders were significantly (*t* = 8.46, *p* < .001) younger 25.25 (*SD*=9.35) **Ethnicity**. Not reported. **Gender**. All offenders were male. **Education**. Apart from elementary maladjustment (*M* = 1.40; *SD* = 0.89) there is no information on education level. **Setting**. All offenders had been admitted at Oak Ridge, a maximum-security psychiatric facility in Ontario, Canada.	**Design**. Retrospective study. **Measure**. PCL-R and VRAG. **Administration**. PCL-R and VRAG and other variables were scored from file information. Raters were blind for outcome. Data were obtained from the Coroner's Office, the Criminal Code Review Board, the Royal Canadian Mountain Police, the National Parole Service of Canada and the provincial correctional and parole systems. Data on recidivism was obtained up to January 31, 1993. The follow-up period for the wife assaulters was 82.5 months (SD = 56.0 months), the mean time at risk for the other offenders	**Total Scores.** PCL-R had mean score of 11.92 (SD=9.06) for the wife assaulters and 18.24 (SD=9.03) for the other offenders. The wife assaulters had a mean VRAG score of -9.27 (SD=9.76). **Inter-rater reliability**. Inter-rater correlations exceeded .80, kappas exceeded .70. **Internal consistency**. Not reported. **Validity**. The association between the PCL-R and violent recidivism for the wife assaulters was *r* = .39, *p* < .001, *n* = 74. The predictive accuracy of the VRAG with the subsample of wife assaulters was: AUC = .75 (SE = 0.07)

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
assaulters compared with other offenders on these variables and is the predictive accuracy of an existing actuarial instrument (Violence Risk Appraisal Guide; VRAG) in predicting violence among wife assaulters examined.		**Relationship**. Of the wife assaulters 2.3% were never married; of the other offenders 66.8% never had been married. **Abuse**. A subsample of 88 offenders had faced charges ranging from threatening to first-degree murder following acts of violence against their spouse.	was 69.3 months (SD = 59.5 months). Of the wife assaulters 24% recidivated in a violent offence. **Outcome**. Violent recidivism was defined as a new criminal charge for an offense against a person or readmission to a psychiatric facility for violent behavior. **Limitations**. The outcome was not specific intimate partner violence but general violent offences. The retrospective design of the research is a limitation in the measurement of repeated IPV.	

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
Grann, M. and Wedin, I. (2002). Risk factors for recidivism among spousal assault and spousal homicide offenders. *Psychology, Crime & Law, 8*(1), 5-23. **Objective**: To test concurrent and predictive validity of the SARA while gaining cross-cultural experience.	88 Male adult offenders convicted of spousal assault or homicide	**Age**. *M* = 34.84 **Ethnicity**. 27% were not born in Sweden; 8% were second generation immigrants to Sweden. **Education**. 25% had some college; 17% did not complete elementary school. **Employment/Income**. At time of index offence 49% were unemployed and 52% relied on social assistance; 10% were homeless; 14% had never had a job longer than 8 months. **Setting**. This sample was taken from a group of Swedish, male violent offenders, convicted of spousal assault or homicide that underwent forensic psychiatric evaluations between 1988-1990 and were diagnosed	**Design**. Retrospective file review (7.8 years). **Measures**. Spousal Assault Risk Assessment (SARA); Violence Risk Appraisal Guide (VRAG); Psychopathy Checklist – Revised (PCL-R); **Administration**. The SARA and other measures were coded by a BA level psychology student from the forensic psychiatric evaluation files. The 20 items on the SARA were coded without the use of the critical items or risk judgement. If the file lacked enough information to rate an item then the item was omitted and the final score was pro-rated. The participants were followed from the index psychiatric evaluation, until December 31, 1995. Information was separated so that the rater of the SARA was blind to the participant's outcome.	**Recidivism**. Recidivism was defined as "any reconviction of a hands-off or hands- on violent behavior, such as homicide, assault, sexual crimes, threats, or violation of no-contact orders, involving a victim with whom the subject had an intimate, sexual relationship were counted as recidivism for the purposes of this study." During follow-up 28% of men were convicted of spousal assault **Total scores**. Mean SARA total score was 20.47 (SD=4.66); the mean SARA part 1 and 2 scores were 10.39 (SD= 3.21) and 10.08 (SD=3.03) respectively. **Inter-rater reliability**. IRR was tested using a random sub sample of n=18 cases rated by a Ph.D. psychologist. The ICC between the two rater's scores was ICC = .85 [F(2,18)= 15.70, *p* < .01] and somewhat lower for Part 2 [ICC=, .74; F(2,18)=7.01, *p* < .01] than for Part I [ICC, =.88; F(2,18)=21.13, *p* < .01]. For the individual items, Cohen's Kappa averaged *k* = .58, with a range from .30 to 1.0 **Internal consistency**. For individual items *k* = 0.58 (range: 0.30 – 1.0). **Concurrent/convergent validity**. Pearson's *r* for: total SARA score and PCL-R score is .59, total SARA and historical portion of HCR-20 is .46; and total SARA and VRAG

Study	N	Sample Characteristics	Method and Design	Results
		with a personality disorder. **Relationship**. Not reported. **IPV history**. All participants had been convicted of spousal assault or homicide. **Other**. Following the psychiatric evaluation, 31% were transferred to forensic psych hospitals on account of insanity, 59% were sent to prison, 10% put on probation. Prior to index, 50% had been convicted of a violent crime.	The PCL-R, VRAG, and H-10 had been completed as part of a previous study. **Other measures**. Historical part of the HCR-20 or H-10. **Limitations**. It is unknown how generalizeable the results from this specific sample would be to general spousal assaulters.	is .33. **Predictive validity**. Predictive validity was calculated based on 'opportunity time', or release from prison or psychiatric institution to recidivism or end of follow-up. The area under the curve (AUC), at 6 months, 1 year, 2 years and 5 years were: .52, .59, .63 and .65 respectively for total SARA scores. For this study, some of the general risk assessment tools more accurately predicted IPV recidivism at one year than the SARA. The AUC for the PCL-R, H-10, and VRAG are as follows: .71, .68, and .75, respectively. The OR for those that scored a 20 or above on the SARA was 2.70 (95% CI: .95-7.66).
McFarlane, J., Campbell, J.C., Sharps, P., & Watson, K. (2002). Abuse during pregnancy and femicide: Urgent implications for women's health. *Obstetrics &*	(*N* = 687) A group of female victims of completed femicide (*N* = 198) and another group of	**Age.** Attempted femicides (*M* = 33.2), completed femicides (*M* = 34.1) were an average of 3-4 years older than the abused controls (*M* = 30.6). **Ethnicity.** African Americans were	**Design.** Retrospective case control study **Measures.** The authors used 17 items from the DA (Campbell, 1995). The item regarding abuse during pregnancy was removed because it was the specific outcome of	**Recidivism.** Not reported. **Total Scores.** The scores on the DA for the attempted and completed femicide group that were abused during pregnancy and those not abused during pregnancy were 8.0 (*SD* = 3.3) and 5.7 (*SD* = 3.3), respectively (*p* < .01). The scores on the DA for the abused control group that were abused during pregnancy and those not abused during pregnancy were 5.8 (*SD* =

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
Gynecology, 100(1), 27-36. **Objective.** To describe the odds of femicide for women abused during pregnancy.	female victims of attempted femicide (*N* = 132) did not differ significantly so the authors put them into one group for some calculations (N = 330); Another group was made up of abused women that had been pregnant (*N* = 357).	overrepresented in the attempted femicide (53.4%) and completed femicide groups (38.1%) compared to the control group (22.1%). Caucasian women made up 48.2%, 22.1%, and 30.5% of the attempted femicide, completed femicide, and control groups respectively. Latina women made up between 22-25% of each group. **Education.** The majority of all women (62-82%) had at least a high school education. **Employment/Income.** Most women were employed (57-76%). **Setting.** Participants were recruited from 10 US cities, from 1994-2000, specifically shelters, medical	interest. **Administration.** 1-2 hour interviews focussed on the 12 months preceding femicide or attempted femicide or worst incident of abuse in the case of controls. *Completed Femicides*: Proxies of femicide victims were contacted from information found in police and medical examiner's records. A proxy was pre-screened to determine knowledge of victim, perpetrator and relationship and then took part in interviews. *Attempted femicides:* The victims of attempted femicide were contacted from police and district attorney files. Once consent was obtained they were interviewed using the same interview schedule as was used with the proxies in the completed femicide	3.7) and 2.6 (SD = 2.5), respectively (p<.01). **Inter-rater reliability.** Not reported. **Internal Consistency.** α = 0.74 for attempted femicides; α = 0.80 for femicides; α = 0.76 for abused controls. **Validity.** Women abused during pregnancy had 3.08 greater likelihood of becoming an attempted or completed femicide victim.

Study	N	Sample Characteristics	Method and Design	Results
		examiners, and district attorney's offices. For each attempted and completed femicide that had ever been pregnant, an abused control from the same city was matched. **Relationship.** Completed femicides had longer relationships with the perpetrators (7.4 years) compared to the attempted (5.6 years) and controls (4.7 years). Completed femicides were more likely perpetrated by former partners (32.3%) than both attempted femicide (29.0%) and control (21.4%) groups. **IPV History.** Only 7.8% of controls reported abuse during a pregnancy whereas 25.8% of the attempted and 22.7%	group. *Controls:* A woman was considered abused if physically assaulted, threatened with serious violence, or stalked by a current or former intimate partner during the past 2 years, as determined using a modified Conflict Tactics Scale with stalking items added. **Other**: 16 items from Violence against Women in America Survey (VAWA) and 10 items from HARASS. **Limitations.** The DA was completed retrospectively. The use of proxy respondents for completed femicide cases. Only women from a few large urban areas were participants. There was not an attempt made to verify self-reports.	

Study	N	Sample Characteristics	Method and Design	Results
		of the completed femicide groups reported abuse during a pregnancy.		
Cattaneo, L.B. & Goodman, L.A. (2003). Victim-reported risk factors for continued abusive behavior: Assessing the dangerousness of arrested batterers. *Journal of Community Psychology, 31*(4), 349-369. **Objective.** To investigate and determine if variables related to demographics, substance use, and past violence, have more predictive validity than a victim's assessment of the likelihood of re-	Baseline: 169 female victims of IPV; Follow-up: 96 female victims of IPV.	**Age.** Over the age of 18 **Ethnicity.** 91.1% African American; 5.9% missing /other; 1.8% Caucasian; 1.2% Latina. **Education.** 23.1% did not complete high school; 35.5% completed high school; 40.2% have at least some college. **Employment/Income.** 48.5% employed full-time, 8.3% employed part-time. 65.1% made less than $20,000 in past year, 24.3% made between $20,000 and $50,000 in past year. **Setting.** Participants recruited from the	**Design.** Prospective study (3-month follow-up). **Measure.** Victim report- A 10 point rating scale was used to rate the likelihood of the batterer hurting, threatening, or destroying the victim's property in next 3 months; a follow up interview ascertained the occurrence of these events. **Administration.** Initially participants filled out a questionnaire containing questions about demographic information, batterer's substance use habits, prior violence, and victim report; 3 month follow up was a telephone interview that collected information about violence, unwanted contacts from perpetrator, and feelings of safety in the past 3 months	**Recidivism.** Repeat abuse over 3 month period to include- unwanted contact, physical abuse, sexual abuse, threats to victim or her property; 11.5% of participants available for follow-up reported the occurrence of repeat abuse. **Total Scores.** NA. **Inter-rater reliability**. NA. **Internal consistency**. NA. **Validity**. Victim's predictions were found to be significantly related to re-abuse ($p < .01$), with a sensitivity of 74%, specificity of 58%, positive predictive value (PPV) of 40%, and the percentage of cases accurately classified was 66%.

Study	N	Sample Characteristics	Method and Design	Results
abuse.		Domestic Violence Intake Center at Superior Court in Washington DC after they had been assaulted by an intimate partner. **Relationship/Abuse.** 56.2% of victims living with batterer at time of assault; 65.6% were married, or dating batterer at time of assault.	**Other.** *Additional measures included:* General Violence Scale; Revised Conflict Tactics Scale (CTS2); Psychological Maltreatment of Women Inventory (PMWI)- short form. **Limitations.** Small sample size; relatively short follow-up; Low base rate of re-abuse; The participants are a specific subset of IPV victims, in that they had called police as a result of IPV and then also showed up in court limiting generalizability.	
Murphy, C.M., Morrel, T.M., Elliott, J.D., & Neavins, T.M. (2003). A prognostic indicator scale for the treatment of partner abuse perpetrators. *Journal of*	95 male participants .	**Age**. *M* = 34.6 years (*SD*=8.1; range 19 to 58) **Ethnicity**. Most participants were White (62%), Black (30%). Smaller proportions were Asian (3%), Native American (2%), Hispanic (1%), and	**Design**. Prospective study. **Measure**. Partner Abuse Prognostic Scale (PAPS), was constructed by summing 17 indicators. The indicators were divided into three subscales: Relationship violence problem severity (7 indicators); Substance use (4 indicators);	**Total Scores.** The range of the PAPS total score was 1-13 with a mean of 5.7 (Mdn=5; *SD*=2.9). **Interrater reliability**. Not reported. **Internal consistency**. Not reported. **Validity**. Correlations between self-report and victims report outcomes and subscales and total score PAPS: Subscale Relationship violence problem

Study	N	Sample Characteristics	Method and Design	Results
Interpersonal Violence, 18(9), 1087-1105. DOI: 10.1177/08862605 03254515. **Objectives**. To measure the effect of prognostic indicators (including, the Partner Abuse Prognostic Scale, psychological aggression, anger, self-esteem, motivational readiness to change, self-efficacy) on several outcome measures obtained through self-report, victim's report (both measured by the CTS2) and official criminal recidivism data		2% Other. **Gender**. Male. **Education**. Participants averaged 13.3 years ($SD = 2.2$) of formal education. **Income**. Participants had a median monthly net income of $1,300 ($M = \$1,820$; $SD = \$1,860$); 8% of participants were unemployed at intake. **Setting**. Adult male participants ($N = 95$) who presented for treatment of IPV at the Domestic Violence Center in, Maryland, USA. Of participants 68% had a court referral to counseling, 8% had a court case pending, but no court mandate to attend, and 23% reported no court involvement for IPV at the time of intake. **Relationship**. Not	Aggression history (4 indicators) and two 'other' indicators, Unemployment and Living together at time of intake. **Administration**. Participants and victims provided data during a 2-session intake assessment, at the end of the treatment program and at 6-months follow-up. **Outcome**. Three different outcome measures were used: 1. Self-report: Program completers provided post-intervention outcome data on the CTS2 (physical violence, severe violence and injuries) during the 15th session (of a 16-session program) and after a 6-month follow-up. 2. Victim's report: Information on outcome from the identified victim and/or most recent victim was obtained through the	severity with Any physical aggression ($r = .26$; $p < .05$), Any severe violence ($r = .27$; $p < .05$) and Injuries ($r = .19$; *ns*) at post treatment and Any physical aggression ($r = .37$; $p < .01$) at 6-months follow-up and Criminal recidivism ($r = .08$; $p > .05$) at 2-3 years. Subscale Substance use with Any physical aggression ($r = .29$; $p < .01$), Any severe violence ($r = .17$; *ns*) and Injuries ($r = .26$; $p < .05$) at post treatment and Any physical aggression ($r = -.20$; *ns*) at 6-months follow-up and Criminal recidivism ($r = .31$; $p < .01$) at 2-3 years. Subscale Aggression history with Any physical aggression $r = 17$; *ns*), Any severe violence ($r = .08$; *ns*) and Injuries ($r = .14$; *ns*) at post treatment and Any physical aggression ($r = -.10$; *ns*) at 6-months follow-up and Criminal recidivism ($r = .20$; *ns*) at 2-3 years. Total score PAPS with Any physical aggression ($r = .41$; $p < .01$), Any severe violence ($r = .35$; $p < .01$) and Injuries ($r = .31$; $p < .01$) at post treatment and Any physical aggression ($r = .24$; $p < .05$) at 6-months follow-up and Criminal recidivism ($r = .23$; $p < .05$) at 2-3 years.

Study	N	Sample Characteristics	Method and Design	Results
		reported.	CTS2 by telephone interview at the end of the program and at 6-month follow-up. 3. Criminal recidivism data was assessed through a review of criminal histories available at the electronic database for the State of Maryland. This data was obtained 22 – 36 months after the scheduled completion of treatment. Outcome data from one or both partners at post treatment was available for 76 participants, 22 (29%) were coded as recidivists for any physical assault, 14 (18%) for injuries, and 12 (16%) for severe violence. For 58 cases self-report and/or victim report was available at 6-month follow-up, 8 (14%) were coded as recidivists for any physical assault, 3 (5%) for injuries, and 3 (5%) for severe violence.	

Study	N	Sample Characteristics	Method and Design	Results
			Criminal recidivism through the database of the State of Maryland was obtained for 71% of the 95 participants, 12 (18%) were coded as recidivism for the following range of charges: assault, battery, violation of a protection order, malicious destruction of property, child abuse, telephone abuse, and assault with a deadly weapon. **Other measures**. CTS2 (Straus, Hamby, Boney-McCoy, & Sugarman, 1996); Alcohol Use Disorders Identification Test (Babor, de la Fuente, Saunders, & Grant, 1992). **Limitations**. The subscale Relationship violence problem severity (six months prior to intake) of the PAPS was constructed from subscales of the CTS2. For the post treatment and 6-month follow-up the CTS2 was	

Study	N	Sample Characteristics	Method and Design	Results
			also used. In this way there is an overlap between the predictor variables and the outcome.	
			It was not detailed who conducted the interviews with the participants and the victims.	
			Although it was described that the information for the CTS was gathered 6 months before intake it is part of the PAPS, this information was not described in detail. This makes it difficult to compare the group of offenders with other research on frequency and severity of previous IPV.	
Girard, L. & Wormith, J. S. (2004). The predictive validity of the Level of Service Inventory-Ontario Revision on general and violent recidivism among various	630	**Age.** 31.78 (*SD* = 9.69) years. **Ethnicity.** Most offenders were White (*n* = 551, 87.5%); Blacks (*n* = 30, 4.8%), Aboriginals (*n* = 24, 3.8%), Asians (*n* = 10, 1.6%), and a	**Design.** Prospective validation study. **Measure.** LSI-OR **Administration.** No incentive or reward was offered to offenders for participating in the study because a risk/need assessment is a standard	**Recidivism.** 54.4% (*n* = 343) of the adult male offenders had recidivated with at least one conviction. Overall, 24.1% (*n* = 152) of the sample recidivated with at least one violent conviction, with the institutional group again being more likely to recidivate violently than the community group (27.1% vs.16.5%, $X^2(1, n = 152) = 58.13, p < .001$). **Total Scores.** Inmates (*M = 22.90, SD = 7.63*)

Study	N	Sample Characteristics	Method and Design	Results
offender groups. *Criminal Justice and Behavior, 31*(2): 150-181. **Objective**. To conduct a longitudinal, predictive validation of the LSI-OR and to examine the contribution of its innovative components to the prediction of general and violent recidivism among samples of inmates and probationers (including IPV offenders).		group with unknown racial background ($n = 15$, .4%) made up the remainder. **Education**. Not mentioned **Employment/Income**. Not mentioned **Setting.** Adult male offenders consisted of 454 inmates and 176 probationers under community supervision. Five settings under the responsibility of the Ontario Ministry of Community Safety and Correctional Services (MCSCS) were identified for participation in the LSI-OR validation exercise. Settings included a small community jail, a correctional center, a treatment center, two large urban probation offices, and one small	part of the intake process within Ontario Ministry of Community Safety and Correctional Services (MCSCS). **Procedure.** The LSI-ORs were completed by institutional classification officers or community probation officers ($n = 16$) in the five locations for 6 months. All LSI-OR assessors had been previously trained in the use of the original LSI and participated in an additional 2-day training program on the LSI-OR prior to the current study. Offender recidivism was determined from two databases by a researcher who was blind to participants' LSI-OR scores. **Recidivism.** Follow-up took place over a 2 ½ year period. The average follow-up period was 932.73 days ($SD = 120.84$)	scored higher than probationers ($M = 13.04, SD = 7.65$) on the General Risk/Need section ($p < .001$). The probationers also had lower scores on all eight subscales (all $ps < .001$). They were consequently more likely to be placed in the high- and very high risk categories than probationers, $(4, n = 630) = 172.01, p < .001$. The institutional group also scored higher on the Specific Risk/Need section and its two subscales ($ps < .001$). Interestingly, the inmates also obtained significantly higher scores on the Other Client Issues subscale ($p < .01$). As expected, the institutional group scored significantly lower than the community group on Strengths ($p < .001$). **IRR.** Kappa coefficient of agreement between raters was .58. **Reliability.** The internal consistency of the 43 General Risk/Need items was quite high with an alpha of .91. Internal consistency was not as strong for the Specific Risk/Need section, with an alpha coefficient of .62. Alpha coefficients for the subscales varied considerably, from .32 for Family/ Marital (4 items) to .80 for Criminal History (8 items). The 1-month, test-retest reliability was .88 ($p < .001$) for the General Risk/Need section ($n = 18$) but only .12 ($p = ns$) for the Specific Risk/Need section.

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
		rural probation office. **Relationship/Abuse.** **Criminal History.** The most common convictions were for assault ($n = 105$, 16.7%).	for the entire sample. Because probationers were immediately eligible to recidivate, whereas inmates were not released for some time after their LSI-OR assessment, the follow-up period was significantly longer for the probationers than for inmates ($M = 1001.12$ days, $SD = 84.63$ vs. $M = 906.22$ days, $SD = 122.42$, $p < .001$, respectively; $t(457.70) = 11.05$, $p < .001$). Recidivism data were obtained from an automated Offender Management System (OMS) and from a national police database, the Canadian Police Information Centre (C-PIC). Three measures of recidivism were coded: any convictions, violent convictions, and offence severity. **Limitations.** Participants	

Study	N	Sample Characteristics	Method and Design	Results
			were not randomly selected; The inmates had a shorter amount of time with which they could recidivate than the probationers. A limitation of the article for our purposes is that the outcome variable is a conviction for any offense not just an IPV offense.	
Hanson, R. K. and S. Wallace-Capretta (2004). Predictors of criminal recidivism among male batterers. *Psychology Crime & Law, 10*(4), 413-427. **Objective.** To examine whether the same risk factors commonly associated with criminal recidivism among general offenders	320 Of the 356 men who completed the intake assessment, 153 failed to complete treatment, and 49 were still in treatment when the study ended. Missing data on the recidivism and/or post-	**Age**. $M = 35.2$ (SD = 8.8) years and 34.5 (8.8) for their female partners. **Ethnicity**. Not reported. **Education**. Men's average education was approximately 12 years (high school graduation), and the median household income was $45 000 (Canadian). **Employment/Income**. The median household income was $45 000 (Canadian). **Setting**. Male	**Design.** Prospective. **Measure.** A modified version of the Level of Service Inventory-Revised. (LSI-R. Andrews & Bonta, 1995). Predictor variables examined also included those commonly used with general offenders (e.g. prior criminal behaviour, lifestyle instability), factors previously cited as specific spousal assault recidivism risk factors (e.g. attitudes tolerant of wife assault; Kropp and Hart, 2000), and factors previously found to be correlated with male battering (e.g. marital	**Recidivism.** Of the 320 men with follow-up information, 55 (17.2%) recidivated with a violent offence and 82 (25.6%) recidivated with any offence. The follow-up period ranged from 39 to 73 months ($M = 58$, $SD= 7.7$). **Total Scores.** Not reported. **Inter-rater reliability.** Not reported. **Reliability.** Not reported. **Validity.** Correlation coefficient: Violent recidivism (IPV) = 0.32 ($p < 0.001$) Any recidivism = 0.40 ($p < 0.001$) AUC Violent recidivism (IPV) = 0.73 ($SD = 0.039$) Any recidivism = 0.76 ($SD = 0.031$) Noteworthy that the adaptations of the LSI-R

Study	N	Sample Characteristics	Method and Design	Results
are also associated with recidivism among male batterers.	treatment measures reduced the post-treatment sample sizes to a median of 98 men and 26 partners.	batterers attending 5 community treatment programs across Canada. **Relationship/Abuse.** 63.6% were married to victim upon index; Marital satisfaction was below the average (100) for both the men and their partners (Locke_ Wallace scores of 82.1 and 71.7, respectively). 45.3% of men had a prior conviction for IPV, however, criminal records were available for 166 of the 356 men.	distress). **Administration.** Although it is typically administered in an interview format, the current study used a self report version (Motiuk et al., 1992; Bonta et al., 2000), which, following recommendations from the scale's authors, was adapted to focus on the problems associated with spousal assault. **Procedure.** Men's characteristics were assessed using self-report questionnaires, which, for some variables, were supplemented with partners' reports (e.g. his substance abuse). Abusive men completed detailed questionnaires at intake and post treatment. Information was also collected from the partners of the abusive men. The number of partners providing information (*n*=116) was less than the	for abusive men were only partially successful. The original Criminal Associates scales showed greater predictive accuracy than the revised questions that specifically targeted association with abusive peers. Similarly, the partners' reports on the men's substance abuse did not improve upon the predictive accuracy provided by the men's own self-report of their substance abuse problems.

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
			number of men (*n*= 320) because many men were separated and participation was voluntary.	
			An important feature of the present design was the repeated assessments on some of the (potentially) changeable risk factors (e.g. attitudes, substance abuse).	
			Recidivism. The distribution of missing cases was similar across the sites. Overall, valid follow-up information was available for a combined sample of 320.	
			Recidivism information was based on both charges and convictions. Two outcome criteria were examined: (a) any violent recidivism and (b) any recidivism. Included in the violent category were assault offences, threats, criminal harassment and harassing telephone calls. Possession of a weapon	

Study	N	Sample Characteristics	Method and Design	Results
			was not considered a violent offence unless it was used in the commission of an offence (e.g. threatening, pointing a weapon). Also included in the violent category were one offender convicted of manslaughter and another offender convicted of robbery. **Limitations.** The LSI-R was not administered in the interview format. Several additional modifications were made to make the measure relevant to IPV. For instance, rather than simply addressing peer support for criminal behaviour, the revised version also addressed peer support for spousal assault (e.g. number of current friends who engage in abusive behaviour, friends with negative attitudes towards women). There was varying compliance with treatment and as the authors noted the men with	

Study	N	Sample Characteristics	Method and Design	Results
			the least engagement in treatment appeared the most likely to recidivate. It will be important to repeat the evaluation of the LSI-R (revised). This examined violent and general recidivism and did not report IPV recidivism specifically.	
Heckert, D. A. & Gondolf, E. W. (2004). Battered women's perceptions of risk versus risk factors and instruments in predicting repeat reassault. *Journal of Interpersonal Violence 19*(7) 778-800. Objective. To expand on previous research about women's perceptions of risk in comparison to and in conjunction with other risk	499 males admitted to batterer treatment and their female partners.	**Age.** Not reported*. **Ethnicity.** Not reported*. **Employment.** Not reported*. **Education.** Not reported*. *Sample characteristics are reported in Gondolf, 1999. **Criminal History.** 82% of the men in the treatment programs were mandated to treatment by the courts, the other men participated	**Design.** Prospective **Measures.** Victim Report, K-SID, SARA, DA **Administration.** At program intake each month, the first 20-25 men from each location were recruited into the sample. A background questionnaire was issued to the participants at intake collecting information about the index incident, CTS items for physical aggression, demographic information, alcohol use, prior treatment, arrest history, their partner's seeking help. Within 2 weeks, 82% the	**Recidivism.** Repeat re-assaulters – 23%, One-time re-assaulters – 12%; threatening reassault – 20%; controlling behavior or verbal abusers – 26%; no abuse- 19% **Total scores.** Not reported. **Inter-rater reliability.** Not reported. **Reliability.** *Victim report.* The two questions to assess the victim's perception of re-assault are: "How safe do you feel at this point?" and "How likely is it that your husband will become violent towards you during the next 3 months?" 26% of women reported that violence was unlikely and 42% reported that violence was very unlikely, 59% reported that they feel very safe. These two variables moderately correlated (Spearman's r = .54). **Concurrent validity.** Not reported. **Predictive validity.** The ability of the simulated K-SID total scores to accurately

Study	N	Sample Characteristics	Method and Design	Results
assessment instruments.		voluntarily. **Relationship.** For 14% of the batterers a new female partner was interviewed and for 8% of the batterers a new partner and the initial victim were both interviewed. **Setting.** Male batterers and their partners from 4 US cities, Dallas, Houston, Denver, and Pittsburgh were recruited from a database of men admitted to batterer treatment groups.	participant's partners were able to be reached and interviewed by phone. A questionnaire to verify the participant's answers was administered. Every 3 months during the15 month follow-up, the men and their partners were called separately and interviewed about drug and alcohol use, treatment, and relationship issues **Procedure.** Victim report was assessed by the following questions: 'How safe do you feel at this point?'; 'How likely is it that your husband will become violent towards you during the next 3 months?' Victims used a five point scale to rate their answer. The K-SID, SARA, and DA were simulated using the data collected in questionnaires. Based on the information collected the authors had 7 of 11	predict re-assault was fair (AUC =.57; sensitivity = 29%). The simulated SARA total scores were a better predictor of the outcomes (AUC = .64; Sensitivity = 43%). The DA was the best predictor in this study (AUC = .70). The women's assessments were moderate predictors (AUC for perception that violence is likely = .64; AUC for perception of safety = .63) When only the men's variables and reports were taken into account to assess re-assault, the AUC = .75 with a sensitivity of 55%. When women's characteristics but not their perceptions are taken into account the AUC = .79 with a sensitivity of 58%. Finally when women's perceptions of re-assault are in included in the model the AUC increases to .83 and a sensitivity of 70%.

Study	N	Sample Characteristics	Method and Design	Results
			items for the K-SID. For the 4 remaining items from the K-SID they had similar information.	
			For the DA, 13 of the 16 items were collected, with 8 items being the same as on the DA, 5 similar items and 3 items not available.	
			The simulated SARA consisted of 16 of the 20 items. 10 items are the same as on the instrument and 6 similar items.	
			Recidivism. No abuse, verbal abuse, controlling behavior, threats, one time reassault, and repeat reassault were the outcomes of interest. The outcome was assessed by the following: An open ended question 'How was the relationship going?', description of conflict and surrounding circumstances, and inventories such as the CTS, were administered for controlling behavior, verbal and physical abuse	

Study	N	Sample Characteristics	Method and Design	Results
			as well as threats. Follow-ups lasted 15 months in total; however, if a woman was followed to at least the 9 month point then her outcomes were included. **Limitations.** The risk instruments were simulated and not used as they would be in practice. In 79% of cases a female partner was interviewed. Only 67% of women were followed for the full 15 month period. However, the authors checked cross-classifications which resulted in a robust Kappa of .75, to justify the use of incomplete cases.	
Hilton, N.Z., Harris, G.T., Rice, M.E., Lang, C., Cormier, C.A., & Lines, K.J. (2004). A brief actuarial assessment for the prediction of wife assault recidivism: The Ontario Domestic Assault	589 male offenders. An additional 100 cases were selected for cross validation.	**Age.** *M*=38.2 years (*SD*=12.0) **Ethnicity.** Not reported. **Employment.** 20% unemployed **Education.** Not reported **Criminal History.**	**Design.** Retrospective follow-up. **Measure.** The authors used the ODARA, the DA, the SARA, and the DVSR. **Administration.** Measures were coded by two of the authors and senior graduate assistants based on file information only.	**Recidivism.** 175 men of 589 (30%) recidivated). The final constructed ODARA (via setwise and stepwise selection analyses) yielded an AUC of .77 (*SE* =.02 *CI* = +/- .04). Correlation between the ODARA and recidivism *r*=.434 (*p* <.001). Positive predictive power ranged from .297-.717 and negative predictive power ranged from .703-.957. **Total Scores.** ODARA mean score 2.89

Study	N	Sample Characteristics	Method and Design	Results
Risk Assessment. *Psychological Assessment, 16*, 267-275 DOI: 10.1037/1040-3590.16.3.267. **Objective:** The construction of an actuarial assessment (the ODARA) to predict male-to-female marital violence.		24% prior correctional sentence. **Relationship.** Average of .39 prior domestic violence incidents. In the index offence 9% of perpetrators used a weapon, 15% used threats, and 38% used severe violence against the victim as defined by the CTS. **Setting.** Consecutive IPV cases were drawn from the Ontario police database.	Probation and parole files were used to code the DA and SARA (55% of cases). **IPV.** Only cases involving a male who evidenced forceful physical contact against his current or former wife or common-law wife based on victim reports or police evidence. **Index offence**. The incident closest to, but not later than December 31, 1996. **Recidivism**. Any subsequent violent assault against an (ex) wife or (ex) common-law wife known to police regardless of whether charges were laid. The follow-up period was *M*=4.79 years after the index offense (*SD*=1.08). **Other measures.** CTS, Cormier-Lang Scale, information collected about substance use, injury to victim, prior criminal history, severity of IPV, victim barriers to support.	(*SD*=2.14). All measures significantly predicted the outcome but not as well as the ODARA as indicated by the ROC areas falling below the CI for the ROC area of the ODARA. **Inter-rater reliability.** Based on correlating ODARA scores for pre-index and index information with ODARA scores for post-index information. Independent raters coded the information. The ODARA ICC =.90; Recidivism ICC = .91.Two police officers who were not involved in the ODARA construction independently scored the ODARA for 10 cases. ICC of .95 (*p*<.001)**.** **Reliability.** Not reported. **Predictive Validity**. The DA AUC was .59, for the SARA the AUC was .64, for the DVSR the AUC was .67 and for the ODARA the AUC was .77 (+ or - .04), d=1.1. The ODARA was significantly correlated with recidivism (masked coding) *r*= .69. **Concurrent validity**. ODARA correlated with the DA (*r*=.43) with the SARA (*r*= .60) and the DVSR (*r*=.53; all p values < .01). **Cross Validation Sample.** In the cross-validation sample, the ODARA mean score was the same but the correlation with recidivism was smaller (base rate of recidivism 26%) and the AUCs of the DA, SARA, and DVSI did not significantly predict the outcome. The ODARA was significantly positively correlated with the

Study	N	Sample Characteristics	Method and Design	Results
			Limitations. It is postdictive instead of predictive and the information was only coded based on file information for both the scoring of the measures and the recidivism information.	sum of victim injury scores for subsequent IPV offences, the sum of the Cormier-Lang Scale scores for subsequent IPV offences, and the number of subsequent IPV incidents with acts of severe violence.
Williams, K.R. & Houghton, A.B. (2004). Assessing the risk of domestic violence reoffending: A validation study. *Law and Human Behavior, 28*(4), 437-455. DOI: 10.1023/B:LAHU.0000039334.59297.f0 Objective. Validation of the Domestic Violence Screening Instrument (DVSI) developed in the Colorado	1465 male offenders arrested for IPV against female partners. 125 women victims of IPV by a subsample of the male offenders to obtain data on self-reported reassault by male abuser.	**Age.** *M* = 32 **Ethnicity.** 49.5% Anglo, 5.6% African American, 43.7% Latino, and 1.2% other. **Gender.** Male. **Education.** Not mentioned **Setting.** Colorado Domestic Violence Risk Reduction Project was implemented in four of the 22 judicial districts of Colorado, USA. **Abuse**. All men included in the sample were charged for	**Design.** Prospective validation study **Measure.** Domestic Violence Screening Instrument (DVSI); SARA (Spousal Assault Risk Assessment). **Administration.** Family violence cases rated by Family Relation Counselors (FRCs) between July 1997 and March 1998 in four pilot judicial districts of Colorado, USA. Information was collected on each offender's history of domestic violence offending, DUI offending, restraining orders, and other criminal	**Recidivism.** Rearrests within the 18-month follow-up period for violations of DV retraining orders, 29% and for other types of criminal offending 53%. Reoffense within the 6-month follow-up based on self report by victims (*N*=125), 35% reported the use of some form of physical force, 80% threatening and/or verbally abusive behaviors, 65% engaged in controlling behaviors. **Total scores.** DVSI: *M*=6.5 (*SD*=4.8) with a range from 0 to 26 in this sample (DVSI has a range of 0-30). SARA: *M*=9.4 (*SD*=6.2) Range: 0-33. **Concurrent validity**. Correlation between DVSI and SARA, *r* = .539, and between DVSI total score and Structured Professional Judgment (SPJ) SARA: *r* = .567. **Interrater reliability**. The logistics of implementing the DVSI in the larger risk reduction project precluded collecting data on

Study	N	Sample Characteristics	Method and Design	Results
Domestic Violence Risk Reduction Project.		domestic violence, 35% had previous domestic violence convictions and/or restraining order violations.	involvement. **Outcome.** For a subsample (*N*=125) self-reported information from the female victims after a 6-month follow-up period was used. Interviews with victims were conducted by telephone using a questionnaire on three forms of control/threats: 1. Index of control; 2. Index of threats; 3. Index of severe threats. Similarly two indexes physically violent behaviors were divided into 4. Index of violence and 5. Index of very severe violence. Official re-arrest data: Outcome measure available for the whole sample: official re-arrest records after an 18-months follow-up period, This data was divided into: 1. Arrests for violations of domestic violence restraining orders and for domestic violence reoffending (base rate 29%), and 2. Arrests for	interrater reliability. **Internal consistency**. DVSI total score, α = .71. SARA: part 1 (general violence) α = .66, part 2 (spousal violence), α = .73. Victim self-report Threats: 1. Index of control (α = .85); 2. Index of threats (α = .79); 3. Index of severe threats (α = .63). Physical violence: Index of violence: α = .85; 2. Index of very severe violence =.77. **Predictive validity**. Self-reported victim data: DVSI on 1. Index of control: AUC = .58, *p* = .14; *r*= 13; 2. Index of threats: AUC= .56, *p* = .26; *r* = .09; 3. Index of severe threats: AUC=.68, *p*=001; *r*=.22, p<.05; 4. Index of violent behavior: AUC=.49; *p*=.92; *r*=.09; 5. Index of very severe violence: AUC = .65; *p* = .001; *r* = .22, *p* < .05. Re-arrest data: DVSI total score on dichotomy of re-arrest for DV: AUC = .61, p = .000, frequency of DV re-arrest; *r* = .18; *p* = .00. DVSI on dichotomy total offending: AUC = .65; *p* = .000; frequency of total re-arrest: *r* = .21; *p* < .00.

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
			other types of criminal offending (base rate 53%).	
Clift, R.J.W., Thomas, L.A., & Dutton, D.G. (2005). Two-year reliability of the propensity for abusiveness scale. *Journal of Family Violence, 20*, 231-234. **Objective.** To establish the test re-test reliability of the Propensity for Abusiveness Scale	95 men and 67 women at Time 1 and 27 men and 37 women at Time 2 attending university during the 1998-1999 and 2000-2001 school years.	Of those who participated in Time 1, 55% of the women and 28% of the men agreed to return for Time 2. **Age.** *Females.* At Time 2 the mean age was 21.5 years (*SD*=1.56). *Males.* At Time 2 the mean age was 21.7 years (SD=1.18). **Ethnicity.** *Females.* At Time 2 27% Caucasian, 70% Asian, and 3% South Asian. *Men.* At Time 2 37% were Caucasian, 63% Asian. **Education.** Second and third year of university. **Employment/Income.** Not reported. **Relationship. An inclusion criterion**	**Design.** Test-retest reliability **Measure.** Propensity for Abusiveness Scale. **Administration.** Participants completed the PAS as part of a larger study on personality and relationships. Participants completed the PAS as part of a battery of tests and then were contacted two years later by phone or email. At Time 2, participants either picked up a package or were mailed a package. Completed packages were brought back by the participants to the Relationship Laboratory. **Limitations.** There was a high attrition rate even though there was no significant difference between those who did and did not participate in terms of their scores on the PAS.	*Females.* **Total Scores.** Time 1 mean score was 45.08 (*SD*=11.26) and at Time 2 the mean score was 44.76 (*SD*=10.99). *Men.* **Total Scores.** At Time 1, the mean score was 46.41 (*SD*=9.54) and at Time 2 the mean was 46.29 (*SD*=9.72). *Females.* **Test-retest Reliability.** The overall rest-retest reliability for females for the PAS total score was *r*=.851, *r*=.770 for the Affective Lability scale, *r*=.620 for the Trauma Symptoms scale, and *r*=.789 for the Recalled Negative Parental Treatment. *Males.* **Test-retest Reliability.** PAS total score *r*=.629; Affective Lability *r*=.553; Trauma Symptoms *r*=.526; Recalled Negative Parental Treatment *r*=.672. *Combined.* **Test-retest Reliability.** PAS total score *r*=.774; Affective Lability *r*=.680; Trauma Symptoms *r*=.567; Recalled Negative Parental Treatment *r*=.749. *Inter-rater Reliability.* Not reported. *Validity.* NA.

Study	N	Sample Characteristics	Method and Design	Results
		was a heterosexual relationship of at least 6 months duration. At Time 2, 41% of the women were still in a relationship with the same partner from Time 1, 59% were not and of those who were not, 45% reported a new relationship. For men, 70% were still in the same relationship, 30% were not and of those who were not in the same relationship, 38% reported a new partner. **IPV History.** Not reported **Setting.** Participants were recruited to "listen to some short audiotapes." Interested people were screened over the phone.	In addition, the sample is only university students and the PAS will likely be mostly used in the community and with known offenders.	
Berk, R. A., He, Y. & Sorenson, S.B. (2005). Developing a	Sample. 1,500 households	**Age**. Not reported. **Ethnicity**. Households with Anglo and Latin backgrounds are most	**Design.** Prospective development study. **Measure.** Development of a forecasting screener for	Recidivism. Of the households 109 (21.1%) had a subsequent call to the Sheriff's Department, and 407 (78.9%) did not, of the subsequent calls 29 (5.6%) were related to DV.

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
practical forecasting screener for domestic violence incidents. *Evaluation Review,* 29(4), 358-383. DOI: 10.1177/0193841 X05275333. **Objective**. The objective of the study was to find a set of approximately five questions to construct a forecast screening tool that deputies could administer and score very quickly at the scene and that could help them anticipate better whether future domestic violence was likely (p. 360).	that according to the authors were the original representative for the area of the Los Angeles Sheriff's Department. Complete data was obtained for 516 households. There was no specific information reported on the demographic characteristics of the sample.	common, but also African American households are frequent, as well as households from different countries in Asia. **Gender**. Not reported. **Education**. Not reported. **Setting**. The Los Angeles county in which the local Sheriff's Department has jurisdiction. **Relationship**. Heterogeneous composition of households. **Abuse**. Nearly three quarters of the households involved in the study had past experience (most recent occurrence within the past 6-months) with DV. Of these households 50% had called the police twice or more before,	Domestic Violence incidents, in the article the tool has not been given a specific name. **Administration.** Deputies of the Sheriff's Department employed an initial screener of 30 questions as part of their usual duties at the scene. Using data mining techniques (classification and regression trees) this was limited to four (all calls) and three (calls related to DV) questions. **Outcome.** The outcome of interest in this study was arepeat call to the same household during the 3-month follow-up period **Other measures.** No. **Limitations.** The outcome was repeated calls to the police, but repeated calls does not secure that DV has occurred. The base rate (5.6%) of DV related calls was very low and is more difficult to	**Total Scores.** Not applicable. **Interrater reliability.** Not reported. **Internal consistency**. Not reported. **Validity.** Using a cost ratio of 5:1 for false negatives (incorrectly forecasting no future calls) to false positives (incorrectly forecasting future calls) four of the original 30-items predicted future calls to the sheriff's office 60% of the time, and accurately forecasted the absence of domestic violence 50% of the time. With a cost ratio of 10:1 of false negatives to false positives, the accurate forecasting of domestic violence calls was 50% of the time, and accurately forecast the absence of domestic violence calls was 70% of the time.

Study	N	Sample Characteristics	Method and Design	Results
		approximately 25% had led to an arrest, 15% to a conviction. About 10% of the households had a restraining order in place.	predict.	
Ellis, D. & Stuckless, N. (2006). Separation, domestic violence, and divorce mediation. *Conflict Resolution Quarterly, 23*(4), 461-485. DOI: 10.1002/crq.150. The study was designed (p. 468) to (1) measure separation assaults as well as (2) conflict- and control- motivated assaults and emotional abuse, (3) to identify predictors of	The sample is a random, sequential sample of one hundred forty-seven male (n=67) and female (n=80) partners referred to divorce mediation by a family court judge or family lawyers, or who made the	**Age**. Female: M=34; Male M=36. **Ethnicity**. Not mentioned. **Gender**. Female (54,4%; n=80) and male (45,6%; n=67) partners. Relation with outcome variable after separation is reported for female partners only. **Education**. Approximately one quarter of male and female partners were university graduates. **Setting**. Divorce mediation in two urban family courts located in Ontario,	**Design**. Prospective study with one follow-up by telephone after 4-months. **Measure**. DOVE (Domestic Violence Evaluation), a risk assessment and management instrument with an integrated safety plan, is used to achieve these two objectives. **Administration**. DOVE was administered to couples by court connected/court based family mediators, prior to participation in divorce mediation. Questions were asked about the time when partners lived together and after they separated.	**Total Scores.** The authors mention (p. 469) that total score was calculated to categorize the individuals into categories of low, moderately high, high, and very high risk. But neither total scores nor the categorizations were detailed in the article. **Interrater reliability**. Not reported. **Internal consistency**. Not reported. **Validity**. Predicted validity was not available for the total score of the instrument, but it was on item level. Significant correlations ($p < .05$) of DOVE items with Serious physical harm after separation: Sexual assaults ($r=.259$), Being seriously hurt physically ($r=.943$), Called police ($r=.445$), Left home ($r=.250$), Partner's drinking ($r=.319$), Outbursts of anger ($r=.259$), Poor communication and social skills ($r=.264$), Extreme possessiveness, jealousy, emotional dependence ($r=.350$), and General control

Study	N	Sample Characteristics	Method and Design	Results
violence and abuse by male partners when they lived with their female partners and (4) by the same partners after they separated, and (5) to derive from the findings safety-promoting interventions for divorce mediators and other professional service providers whose mandate includes the prevention of domestic violence.	decision themselves. The two courts are located in Ontario, Canada.	Canada. **Relationship**. Male and female partners attended to divorce mediation in Ontario. **Abuse**. Abuse reported by the female partner before and after they separated from their partner. When living together: Physical assaults: 50.1%; Seriously hurt physically: 17.2%; Sexual assault: 9.8%; Emotional abuse: 84.4%; Seriously harmed emotionally: 71.3%. Definition and/or severity of abuse are not specified by authors. **Limitations**. The DOVE items that measure previous abuse were also used to measure future violence, by telephone interview after 4 months. The use of the	**Outcome**. Four months after divorce mediation was completed the same instrument (DOVE) as before mediation and additional questions (not specified by authors) was administrated by telephone. Post-separation violence against female partners: Physical assaults: 9.7%; Seriously hurt physically: 14.3%; Sexual assault: 9.8%; Emotional abuse: 69.9%; Seriously harmed emotionally: 47.8%. **Other measures**. No	(r=.297). Significant correlations of DOVE items with Physical assaults after separation: Physical assaults (r=.386), Complained about partner doing drugs (r=.332), Threats to kill self if she left (r=.274), and Threats to kill me if she left (r=.306). Significant correlations of DOVE items with Serious emotional harm after separation: Seriously hurt emotionally (r=.330), Called police (r=.330), General control, (r=.328) and Behavioral control (r=.340). Significant correlations of DOVE items with Emotional abuse after separation: Physical assaults (r=.274), Emotional abuse (r=.301), Left home (r=.334), Complained about partner doing drugs (r=.258), Blame (r=.416), Outbursts of anger (r=.268), Hard to get along with (r=.359), Extreme jealousy, possessiveness, emotional dependency (r=.308), General control (r=.412), and Behavioral control (r=.461).

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
		same items to measure future conjugal violence and abuse could enhance the predictive accuracy of the tool.		
Hendricks, B., Werner, T., Shipway, L., & Turinetti, G.J. (2006). Recidivism among spousal abusers: Predictions and program evaluation. *Journal of Interpersonal Violence, 21*, 703-716. **Objective.** To determine the relative effectiveness of two interventions for spousal abusers.	200 male offenders from rural Wisconsin who were charged with domestic abuse.	**Age.** Not reported **Ethnicity.** Not reported **Education.** Not reported **Employment/Income.** Not reported **Relationship.** Not reported **IPV History.** Not reported **Setting.** Male offenders from rural Wisconsin who were charged with domestic abuse and who were referred to Children's Service Society with intake dates between July 1999 and December 2001. One hundred twenty-five participants (62.5%)	**Design.** Prospective Cohort Study. **Measures.** LSI-R. **Administration.** At entry into the program, participants were administered the LSI-R. They then either were referred to an R&R program (cognitive skills training for social functioning) and then the SAFE program (IPV treatment) or just to the SAFE program. **Recidivism.** Offenders were followed-up at 6-month, 12-month and 19-month intervals after completing or withdrawal from their treatment. Outcome data (any offense for domestic violence toward an intimate partner)	**Recidivism.** Thirty-five participants (17.5%) were coded as treatment failures (more than half failed early in the process); 21 during and within 6 months of treatment; 8 between 6 and 12 months; and another 6 during the 12 to 18 month interval. Of participants that completed the 14 week program 10.6% recidivated Of participants that did not complete the program 38.8% recidivated **Inter-rater reliability.** Not reported. **Total scores.** LSI-R mean for the whole sample was 10.48 (*SD*=5.34). Those completing the SAFE program had significantly lower total LSI-R scores (*M* = 9.51, *SD* = 4.79) than those who did not (*M*=13.47, *SD*=5.87). Those referred to R&R had significantly higher total LSI-R scores. **Internal Consistency.** Not reported. **Predictive Validity.** When the authors entered the Risk and Need Scales independently into the logistic regression an overall classification

Study	N	Sample Characteristics	Method and Design	Results
		were convicted of abuse and were on probation; 75 (37.5%) participants entered the program under a deferred entry of judgment.	was obtained from a state criminal database. Violence before or during treatment was also treated as recidivism. **Limitations.** Limited generalizability given the specific sample. Also, because this sample was in treatment, this could impact recidivism and thus the predictive ability of the LSI-OR and the authors did not take this interaction into consideration.	rate of 64% results. When the LSI-R total score was used the overall classification rate was 66% with a specificity of 67% and a sensitivity of 60%. The authors determined the optimal cut-off score was 11.5 which was considerably below the value of 16 suggested by Andrews and Bonta (1995).
Williams, K.R. & Grant, S.R. (2006). Empirically examining the risk of intimate partner violence: The Revised Domestic Violence Screening Instrument (DVSI-R). *Public Health Reports, 121*, 400-408. Retrieved from: http://www.ncbi.nlm.nih.gov/pmc/arti	14,970 assessments	**Age**. M = 33 years, 65% of the sample between 22-44 years. **Ethnicity**. Non-Hispanic white, 53%; African American, 29%; Hispanic or Latino, 17%; Asian, Pacific Islander, and American Indian or Alaskan Native, less than 1% (*n* = 115). **Gender**. 71% male; 29% female **Education**. Not	**Design**. Field study. **Measure**. Domestic Violence Screening Instrument (DVSI) **Administration**. Family violence cases rated by FRC's between September 1, 2004 and May 2, 2005, in Connecticut, USA. In principle, assessments are based on perpetrator interviews, a review of police reports, criminal history and protective order registry reviews, as well as	**Total Scores.** DVSI-R mean score = 7.75; (SD=5.57) with a range from 0 to 26 in this sample (DVSI-R has a range of 0-28). **Interrater reliability**. Not reported. **Internal consistency**. Not reported. **Predictive validity**. DVSI-R: AUC = .71; Summary Risk Rating Imminent risk to victim: AUC = .64 (95% CI: .70-.72); Summary Risk Rating imminent risk to other: AUC = .61 (95% CI: .63-.65); Multiple victims (1), single victim (0), AUC = .79 (95% CI: .78-.80).

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
cles/PMC1525359/ pdf/phr121000400. pdf **Objectives**. 1. Estimate the effects of age, ethnicity, and gender on risk scores. 2. Estimate the effects of intimate partner violence compared with other forms of intimate violence on risk scores. 3. 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		available. **Setting**. Family Services and Connecticut Courts state wide in Connecticut. Family Relations Counselors (FRCs) of the Family Services rated the DVSI-R before court sessions in the 24-hour period between arrest and initial court appearance. **Relationship**. Multiple forms of family violence. **Abuse**. Most of the cases entailed a unspecified form of IPV (66%), 16% regarded violence between parent and child, and 18% of cases involved some other form of intimate relationships.	victim interviews conducted by victim advocates, but this information was not always available. **Outcome**. Multiple-assessments (25% of the 14,970 assessments) by the Family Service due to re-arrest, within the research period. **Other measures**. Not reported.	

Study	N	Sample Characteristics	Method and Design	Results
indicator of more "severe" violence. 4. 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.				
Cattaneo, L.B., Bell, M.E., Goodman, L.A. & Dutton, M.A. (2007). Intimate partner violence victims' accuracy in assessing their risk of re-abuse. *Journal of Family Violence, 22*(6), 429-440.	246 female IPV victims seeking help.	**Age.** *M* = 33 years. **Ethnicity.** 79% African American. **Education.** 74% completed high school. **Employment/Income.** 63% were employed full or part time; 92% made less than $30,000/year. **Setting.** Participants	**Design.** Longitudinal study **Measure.** Victims of IPV rated on a 5 point scale the likelihood that their partners would physically injure them and try to kill them in the next year. **Administration.** Phone interviews by research assistants every 3 months for first 12 months and then again at 18 months. During phone interviews	**Recidivism.** In the 18 month follow-up period 26% of participants experienced re-abuse. **Total Scores**. When victim reports were dichotomized into low or high risk for re-abuse 69% reported a low risk of physical re abuse while 31% reported being at high risk for physical re-abuse. **Inter-rater reliability**. NA **Internal consistency**. NA **Validity**. Approximately 66% of this sample assessed their risk accurately. Victims were equally able to predict future abuse as they were

Study	N	Sample Characteristics	Method and Design	Results
Objective: Examine victims' accuracy in predicting the likelihood of future physical IPV.		were recruited from shelters (14%), and criminal courts (27%) in a mid Atlantic city. **Relationship/Abuse.** Not reported.	the following information was collected: demographic information, PTSD symptomatology, substance use, recency of assault, interpersonal predictors, psychological abuse, relationship information, past IPV experience, and help seeking behaviors. **Outcome.** Physical IPV reassault. **Other Measures.** Yes/no version of CTS2; Psychological Maltreatment of Women Inventory- short form (PMWI); National Violence Against Women Survey Stalking items; Interpersonal Support Evaluation List (ISEL); PTSD checklist. **Limitations.** Study is not generalizeable beyond low income, African American women seeking assistance for IPV;; Exposure to risk not taken into account over	to predict no future abuse. Victims of abuse at follow-up were more likely to correctly predict re-abuse than not ($p=.01$).Women with PTSD were likely to overestimate the likelihood of re abuse while substance users were likely to underestimate their risk of re-abuse.

Study	N	Sample Characteristics	Method and Design	Results
			follow up (i.e. batterer in jail).	
Belfrage, H., & Strand, S. (2008). Structured Spousal Violence Risk Assessment: Combining Risk Factors and Victim Vulnerability Factors. *International Journal of Forensic Mental Health, 7*(1), 39-46. DOI:10.1080/14999013.2008.9914402. **Objective**. 1. To investigate whether victim vulnerability factors contribute to the police officers risk assessment (correlate with the SPJ) and 2. To	540 alleged male aggressors assessed by police officers.	**Age**. Mean age 38 years (range13-76). **Index offenses**. The suspected index crimes were assault (58%, 312), illegal threat (18%, 95), violation of a woman's integrity (14%, 76), and other crimes (e.g., attempted murder, molesting, violation of no-contact orders, and invasion of privacy). **Ethnicity**. Not reported. **Gender**. 100% male. **Education**. Not reported. **Setting**. Investigating police officers that rate suspected spouse assaulters on B-SAFER. **Relationship**. All	**Design**. Prospective field study. **Measure**. B-SAFER (Kropp & Hart, 2000) with 5 additional victim vulnerability items (Inconsistent behavior/attitude; Extreme fear; Inadequate access to resources; Unsafe living situation; Personal problems) **Administration**. The investigating officers in two police counties in Sweden (Södertörn, a suburb of Stockholm and Kalmar, a medium sized city of approximately 100,000 inhabitants) rated the B- SAFER in cases of reported spousal violence during the period May 2005 – December 2006. **Outcome**. None (in terms of repeated abuse, see objective of the study). **Other measures**. Apart	**Total Scores.** Not reported, scores on individual item level were reported. **Interrater reliability**. Not reported. **Internal consistency**. Not reported. **Validity**. Not reported.

Study	N	Sample Characteristics	Method and Design	Results
study how common such factors are, and to determine if these factors can be coded in an acceptable way.		victims were spouses.	from the 5 additional victim vulnerability items no other measures were used in this study.	
Bell, M. E., Cattaneo, L. B., Goodman, L.A. & Dutton, M. A. (2008). "Assessing the risk of future psychological abuse: Predicting the accuracy of battered women's predictions." *Journal of Family Violence, 23*(2): 69-80. **Objective**: Examine victims' accuracy in predicting the likelihood of future psychological IPV.	244 female IPV victims seeking help.	**Age.** *M* = 33 years **Ethnicity.** 80% African American. **Education.** 74% completed high school. **Employment/Income.** 60% were employed full or part time; 92% made less than $30,000/year. **Setting.** Participants were recruited from shelters (13%), and criminal courts (28%) in a mid Atlantic city. **Relationship/Abuse.** Not reported.	**Design.** Longitudinal study **Measure.** Victims of IPV rated on a 5 point scale "the likelihood that their partners would control or dominate them or humiliate or degrade them in the next year. **Administration.** Phone interviews by research assistants every 3 months for 18 months. **Outcome.** Psychological IPV **Other Measures.** Yes/no version of CTS2; Psychological Maltreatment of Women Inventory- short form (PMWI); National Violence Against Women Survey Stalking items; Interpersonal Support	**Recidivism**. In the 18 month follow up 34% of participants said that the perpetrator had controlled or dominated them and 41% said that they had been humiliated or degraded by him. When these categories were collapsed 52% of the sample had been psychologically abused. **Total Scores**. When victim reports were dichotomized into low or high risk for re-abuse 55% of participants felt that they were at low risk for psychological re-abuse and 45% participants felt they were high risk for psychological re-abuse. **Inter-rater reliability**. NA **Internal consistency**. NA **Validity**. 62% of participants assessed their risk accurately. The multi-variate analysis revealed that the following variables significantly improved the fit of the model: a diagnosis of PTSD, recency of violence, psychological abuse and past stalking.

Study	N	Sample Characteristics	Method and Design	Results
			Evaluation List (ISEL); PTSD checklist. **Limitations.** Due to lack of previous research on the accuracy of victim predictions for psychological re- abuse, it is not known if the findings of this study are generalizeable.	
Glass, N., Laughon, K., Rutto, C., Bevacqua, J., & Campbell, J.C. (2008). Young adult intimate partner femicide - An exploratory study. *Homicide Studies, 12*(2), 177-187. **Objective.** To identify risk factors for intimate partner femicide in adolescents and young adults and compare whether they are same as	53 Abused control women; 23 Femicide cases; 5 Adolescent femicide cases	**Age.** Young adult femicide group between ages of 18 - 20 years old (*M* = 19.2); Abused control group was between the ages of 18 - 20 years old (*M* = 18.8); Adolescent femicide group aged 16 and 17 **Ethnicity.** *Abused control group -* African American 18%, White 31%, Latina 27%. *Femicide cases -* African American 57%, White 17%, Latina 24%.	**Design**. Retrospective; Secondary analysis of a case control study from Campbell et al., (2003). **Measures.** The Danger Assessment (DA); Conflict Tactics Scale (CTS) used to identify instances of abuse. **Administration.** *Abused control group* – Women between the ages of 18 and 50 that had been in a relationship in the past 2 years were called using stratified random-digit dialling. Those that were contacted that had been abused were included in	**Recidivism**. NA **Total Scores**. When the young adult (age18-20) femicide victim's overall DA scores were compared to older adult femicide victims (age 21 and up) (*N* = 310) there was not any significant difference. The mean DA score for the young adult femicide cases (5.4) was, higher than the mean score of the same age controls (*p*=.023). A total score for the adolescent femicide victims was not given. The femicide group also scored higher than the abused control group on every DA risk factor but the one assessing abuse during pregnancy. **Inter-rater reliability**. Not reported. **Internal consistency**. Not reported. **Validity.** Not reported.

Study	N	Sample Characteristics	Method and Design	Results
risk factors for older adults.		*Adolescents femicides* - African American *(n = 2)*, Latina *(n = 1)*, Asian American *(n = 1)*. **Education.** Information not given for the *Abused controls* or the *Femicide cases*. *Adolescent femicides-* All were high school students at the time of murder. **Employment/Income.** Information not given for the *Abused controls* or the *Femicide cases*. *Adolescent femicide cases-* 4 worked part-time. **Relationship.** In the *abused control group* 49% of perpetrators were ex-partners. In the *Femicide group* 61% of perpetrators	the parent study. Abuse was defined as threatened with a weapon or physically assaulted. In this secondary analysis only women from the parent study between the ages of 18-20 were included. *All femicide cases-* Police and medical examiner records from 1994-2000 in each of the study cities were examined for details between victim and perpetrator. All cases closed by the police in which the perpetrator was a current or ex-intimate partner were included. For each of these cases proxies were identified and screened, when possible. **Procedure**. Telephone or in-person interviews were used to collect information from the abused women and proxies of the femicide victims.	

Study	N	Sample Characteristics	Method and Design	Results
		were ex-partners.- *Adolescents femicides*- 2 current partners, 2 ex-parnters, 1 unknown **Setting.** All records of intimate partner female homicide, between 1994-2000, in 11 US cities were included in study if they were between 16-20 years, and a proxy could be located ($N =$ 23). The abused control group was made up of women age 18-20, that had been threatened with a weapon or physically assaulted by an intimate or ex-intimate partner in the past 2 years ($N = 53$). Five adolescent femicide cases were identified from the parent study and assigned controls of participants aged 18-20.	**Limitations.** Small sample size limits multivariate analysis as well as generalizability. No age matching for adolescent femicide victims because of consent laws. This secondary analysis could not address potential risk factors unique to adolescents and young adults. Proxy informants may lack important information.	

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
Glass, N., Perrin, N., Hanson, G., Bloom, T., Gardner, E., & Campbell, J.C. (2008). Risk for reassault in abusive female same-sex relationships. *American Journal of Public Health, 98*(6), 1021-1027. **Objective**. To revise the DA in order to apply it to female same-sex couples.	84 female victims of same-sex IPV	**Age.** *M* = 36.6 years **Ethnicity.** 65.9% white, 14% Latina, 13.6% African American **Education**. *M* = 12.89 years **Employment/income.** 64.5% of participants were employed; 16.5% of participants had a household income of less than $1000 per month and 24.1% had household incomes greater than $4001 per month. **Relationship/Abuse.** 73.3% reported that the perpetrator was an ex-partner. **Setting**. Participants were recruited from American cities with large Lesbian, Bisexual, and transgender communities (e.g. San Francisco, New York, Washington D.C.,	**Design.** Prospective (1 month follow-up). **Measures.** A revised version of the DA using 8 original items and 10 new items specific to female same-sex IPV risk assessment. This version of the instrument is referred to as the DA-R. **Administration.** In phase one of the project focus groups including victims and perpetrators of female same-sex IPV were assembled to review the risk factors in the 20 item DA and identify factors, not on the DA that would be important to consider in assessing risk in female same-sex IPV. The participants found all 20 items from the DA relevant to same-sex IPV, as well as came up with 59 other relevant items. Phase two of the project included a baseline and one month follow-up interview	**Recidivism.** One third of the sample reported being sexually or physically assaulted or threatened by an abusive partner or ex partner at follow-up. **Total Scores.** Statistical analysis led to 18 items being included in the DA-R. Eight of these items are from the original DA and 10 new items. The weighted score for participants that did not experience violence on follow up was 15.18 and the score for individuals that did experience violence on follow- up was 18.57. The DA-R was a significant predictor of threatened or actual, sexual or physical violence at 1 month. For the weighted DA-R, each additional point scored equated to a 1.21 greater likelihood that the individual would experience threatened or actual violence at one- month follow-up. **Inter-rater reliability**. Not reported. **Internal consistency**. Not reported. **Validity.** Not reported.

Study	N	Sample Characteristics	Method and Design	Results
		Seattle).	with women that had been abused in a female same-sex relationship. At baseline the participants reported their experiences in the last 6 months on the 20 DA items and 59 new risk factors for same-sex IPV. The 1 month follow-up interview included the same 79 items as at baseline but the women were reporting on their experiences in the past month. **Procedure**. Both the baseline and 1 month follow up interviews were conducted by telephone. **Limitations.** Small sample size.	
Hilton, N.Z., Harris, G.T., & Holder, N. (2008). Actuarial assessment of violence risk in hospital-based partner assault clinics, *Canadian*	111 women (71% larger community; 29% smaller community) who were victims of	**Age.** Range from 18 to 53 years (*M*=36.2; *SD*=10.0). **Ethnicity.** Not reported **Education.** Not reported **Employment/Income.**	**Design.** Correlational Design **Measures.** ODARA. The ODARA item related to victim concern about future violence was removed for analyses including the victim's perception of future violence.	**Recidivism.** No outcome data was collected. **Total scores.** The mean ODARA score based on the victim interview was 7.3. **Inter-rater reliability.** Not reported **Internal consistency.** Cronbach alpha = .65. **Concurrent Validity.** The correlation between the ODARA and CTS was r =.25, between the ODARA and sexual assault r =.22, between the

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
Journal of Nursing Research, 40(4) 56-70. **Objective.** To determine the utility of the ODARA to profile women attending partner assault clinics, especially with respect to level of risk and severity of injuries as assessed through a routine nursing assessment.	domestic violence. There was no formal record of those who declined, but staff noted about five clients declined due to time constraints at one of the sites.	Not reported **Relationship.** 39% of participants were cohabitating and 21% were married to the perpetrator. The average relationship length was 9.9 years (*SD*=7.0) and 60% of the couples had children under 18 years of age. **IPV History.** 79% were previously assaulted by the perpetrator. In the current assault, 66% of victims sustained an injury and included strangling (29%), kicking (20%) and sexual assault (11%). **Setting.** Participants either referred themselves or were referred by the ER after an admission for abuse and who were assessed between 2003 and 2006 at two	**Administration.** The ODARA was administered as part of the routine assessment (face-to-face) conducted by the nurse. **Recidivism.** No outcome data was collected. Instead, severity of the assault was coded from the participant's hospital file on a scale of 1 (no injuries) to 5 (wounds from weapons). In addition, participants rated the likelihood of experiencing an assault within the next year on a scale of 0 (no chance) to 10 (sure to happen). **Other Measures.** CTS-2. **Limitations**. The sample is not random and likely represents more severe cases of IPV since they are presenting to the ER. Thus, there was little variability in the ODARA scores. Also, it was not a predictive study so it is unclear whether victim perceptions are related to	ODARA and the five point injury scale was r=.25, between the ODARA and the presence of potentially lethal acts was r=.19, and between the ODARA, prior medical treatment for assault by the perpetrator was r=.26, and stalking behaviours r=.32 (all ps = or <.05). Victim perception of future violence was not significantly related to the ODARA. Participant ratings were most strongly associated with reports of increasing severity of violence (r=.29 p<.05) but not with perceived increase in frequency of assault.

Study	N	Sample Characteristics	Method and Design	Results
		Ontario hospital-based clinics that provide various medical, psychological, and legal services for women who have been recently sexually assaulted or experienced domestic violence (Sexual Assault/Domestic Violence Treatment Centres). Women were eligible if assaulted by a male partner.	accurate predictions based on the ODARA.	
Hilton, N.Z., Harris, G.T., Rice, M.E., Houghton, R.E., & Eke, A. W. (2008). An in-depth actuarial assessment for wife assault recidivism: The Domestic Violence Risk Appraisal Guide. *Law and Human Behavior, 32*, 150-163. **Objective.** To	**Sample 1:** 303 male offenders with a history of IPV (part of the original ODARA construction sample) **Sample 2:** 346 new male offenders	*SAMPLE 1:* **Age.** *M*=35.5 years (*SD*=10.1); **Ethnicity.** Not reported. **Employment.** 14% were unemployed. **Education.** Not reported. **Relationship.** 37% were married to the victim at the index offence and 33% separated from the	**Design.** Retrospective follow-up. **Measures.** ODARA, DA, SARA, PCL-R, and the DVRAG. **Administration.** All measures were coded by researchers and senior graduate students based on file information only. **Recidivism.** Any subsequent assaults against a current or former wife or common-law partner	**Recidivism.** The recidivism rate for Sample 1 was 49% and for Sample 2 it was 41%. **Total Scores**. *Sample 1*: ODARA *M* =4.05 (*SD*=2.15); SARA *M*=4.63 (*SD*=4.94); DA *M*=.73 (*SD*=1.20); DVSI *M*=2.68 (*SD*=2.38); PCL-R *M*=8.00 (*SD*=6.81); VRAG *M*=-2.17 (*SD*=6.73) *Sample 2:* ODARA *M*=3.54 (*SD*=2.00); SARA *M*=4.01 (*SD*=4.00); DA *M*=.75 (*SD*=1.23); DVSI *M*=2.25 (*SD*=2.09); PCL-R *M*=8.35 (*SD*=6.67); VRAG *M*=-3.66 (*SD*=6.56). In the combined sample (n=649) the mean DVRAG score was *M* =2.88 (*SD*=7.76). **Inter-rater Reliability.** IRR of DVRAG: two independent blind raters coded 10 randomly

Study	N	Sample Characteristics	Method and Design	Results
determine if the predictive accuracy of the ODARA could be improved with the addition of more detailed clinical information. This results in the construction of the DVRAG.	with a history of IPV.	victim prior to the index offence. The duration of the relationship was an average of 80.8 months (*SD*=93.9). **IPV History.** The mean number of prior DV incidents was.16 (*SD*=.56). *SAMPLE 2:* **Age.** *M*=35.3 years (*SD*=10.0). **Ethnicity.** Not reported. **Employment.** 14% were unemployed. **Education.** Not reported. **IPV History.** The mean number of prior IPV incidents was .27 (*SD*=.71). **Relationship.** 45% were married to the victim at the time of the index offence and 29% separated prior to	during the follow-up period. The follow-up was on average 5.01 years (*SD*=1.44). Outcome data was collected from federal and provincial police databases. **Other.** *DVRAG Construction.* The DVRAG is the ODARA plus the PCL-R. Weights were assigned to each variable based on the following calculation: for each 5% (rounded) deviation from the sample base rate in the recidivism rate among offenders with a given value range, a weight of plus or minus one was given to that value range. *IPV.* Only cases involving a male who evidenced forceful physical contact against his current or former wife or common-law wife based on victim reports or police evidence *Index Offence.* Physical	selected cases; *r*=.92 intra-class correlation was .90 (both *p*<.001). Forensic clinician compared to research assistant for 16 cases *r* =.83. **Internal Consistency.** Not reported. **Predictive Validity. *Sample 1*:** ODARA *AUC*=.67, 95% CI=.61-.73; All measures were significantly related to recidivism: VRAG *r* =.19, PCL-R *r* =.22; SARA *r* =.18, DA *r* =.12, DVSI *r* =.17. The addition of any of the other measures to the ODARA did not significantly increase the prediction of recidivism. Only the PCL-R added significantly to the prediction of the number of recidivism events, the number of severe violence incidents and for total recidivism injury. *Sample 2:* ODARA *AUC*=.65 (95% CI = .59-.71), *d*=.55, DVRAG *AUC*=.70, *d*=.75. The DVRAG represented an improvement in predictive validity over the ODARA score. *Combined sample:* DVRAG *AUC*=.70, *d*=.75, VRAG *AUC*=.67; PCL-R *AUC*=.66; SARA *AUC*=.59; DA *AUC*=.56; DVSI *AUC*=.61 (all significant).

Study	N	Sample Characteristics	Method and Design	Results
		the index offence. The duration of the relationship was an average of 99.8 months (SD=106). **Setting.** Participants were drawn from the Ontario Provincial Police database – a subsample of the original 689 used in the construction sample who had files compiled by the Ministry of Corrections (assessment reports) as a result of charges pertaining to the index or any other offense. For the second sample half of the participants were selected in the same manner as the ODARA construction sample and half were drawn from the records management systems of two urban police services in the Greater Toronto area.	assault (as defined above) committed between 1995-1997. **Limitations.** All measures were coded based on file information alone. Part of the sample was in the original construction sample.	

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
Campbell, J. C., Webster, D. W. & Glass, N. (2009). The Danger Assessment: Validation of a lethality risk assessment instrument for intimate partner femicide. *Journal of Interpersonal Violence*, *24*(4), 653-674. **Objective**. To assess and possibly revise the DA while identifying intimate partner homicide and attempted homicide risk factors.	310 femicide victims; 194 attempted femicide victims; 324 abused controls	**Age.** Not reported. **Ethnicity.** *Femicide victims.* 24.1% White; 47.3% African American; 24.1% Latina. *Attempted femicide. victims.* 21.6% White; 56.9% African American; 17.0% Latina. *Abuse victims.* 46.3% White; 20.6% African American; 24.2% Latina. **Education**. *Femicide victims.* 33.2% did not graduate high school; 27.6% high school graduate; 7.5% college graduate. *Attempted femicide victims.* 32.3% did not graduate high school; 24.5% high school graduate; 9.7% college graduate.	**Design.** Retrospective case-control. **Measures.** The Danger Assessment Scale (DA). **Administration.** Either telephone or in-person interviews were conducted with participants or proxies. **Procedure.** In addition to the DA, demographic and relationship information was collected, to include information about abuse during relationship. The proxies and victims of attempted femicide were asked to report on risk factors for the year prior to the femicide or attempted femicide. The abused controls reported on their worst incident of physical and sexual abuse in the past two years. **Other measures**. The CTS (with stalking items) was used to identify episodes of abuse **Limitations.** Bias is	**Recidivism**. NA. **Total Scores**. In order to identify risk and inform revision of the DA, mean scores were compared between femicide victims, attempted femicide victims, and an abused control group. The mean and median scores were similar for the femicide and attempted femicide groups (M=18.7; median= 18.0), both of which were greater than twice that of the abused control group (M=7.7; median= 7.0; $p < .001$). **Inter-rater reliability**. Not reported. **Internal consistency**. Not reported. **Validity.** When comparing the scores from the attempted femicide group to the control group the AUC was .916 ($p < .001$). Only 3.5% of the control group fell into the extreme danger category whereas 54.5% of the attempted femicides were categorized as at extreme risk. The sensitivity of the revised DA was anywhere between .55 -.99 depending if the high risk cut-off was a score above 18 or above 8.

Study	N	Sample Characteristics	Method and Design	Results
		Abuse victims. 17.9% did not graduate high school; 21.5% high school graduate; 28.5% college graduate. **Employment/income.** *Femicide victims.* 51.8% employed full-time; 34.1% unemployed. *Attempted femicide victims.* 42.9% employed full-time; 48% unemployed. *Abuse victims.* 52.2% employed full-time; 27.4% unemployed. **Relationship/Abuse.** *Femicide victims.* 39.9% spouse or common-law of abuser. *Attempted femicide victims.* 32.7% spouse or common-law of abuser. *Abuse victims.* 29.7%	introduced into the study because of the retrospective design; Due to study design the researchers could not calculate a meaningful positive predictive value (PPV).	

Study	N	Sample Characteristics	Method and Design	Results
		spouse or common-law of abuser. **Setting**. *Femicide victims*: Femicides between 1994 and 2000, in 11 cities were identified from police and ME records. The researchers tried to find proxy informants knowledgeable about the victim and her relationship, by reviewing these records. The proxy informants were then contacted and upon consent, interviewed either in person or by phone. *Attempted femicide victims*: Attempted femicide victims were identified and contacted from district attorneys, law enforcement offices, community advocates, and trauma centers.		

Study	N	Sample Characteristics	Method and Design	Results
		Abuse victims: A random digit dialling method was used to identify the women in the abused control group.		
Echeburua, E., Fernandez-Montalvo, J., de Corral, P., & Lopez-Goni, J.J. (2009). Assessing risk markers in intimate partner femicide and severe violence A new assessment instrument. *Journal of Interpersonal Violence, 24*(6), 925-939. **Objective.** To develop a prediction measure for lethal and severe IPV.	1081 male batterers: 269 severe batterers and 812 non severe batterers	**Age**. $M = 37.3$years for severe batterers and $M = 38.2$ for non-severe. **Ethnicity**. *Severe batterers*: 64.3% Spanish, 16.7% Latin American, 10.7% African. *Nonsevere batterers*: 74.1% Spanish, 13.8% Latin American, 6.3% African. **Education.** *Severe batterers*: 20.3% without primary education; 52.3% with primary education, 13.5% with secondary education, 12.6% with professional training. *Nonsevere batterers*:	**Design**. Case control. **Measure**. Pilot instrument. **Administration**. Batterers were interviewed by police when charges were pressed. **Outcome**. Homicide and severe violence. **Limitations.** Several different assessors were interviewing the participants and assigning participants to the severe and non severe groups; The instrument was not used to prospectively determine risk of future abuse.	**Recidivism.** Not reported. **Total Scores.** *Severe batterers*: $M = 9.2$, $SD = 3.6$; *Nonsevere batterers*: $M = 6.3$, $SD = 3.2$ ($t = 12.4$, $p < .001$). **Inter-rater reliability.** Not reported. **Internal Consistency.** Cronbach's alpha for the total sample, severe perpetrators and nonsevere perpetrators are .71, .69, and .66, respectively. **Validity.** Every item included on the instrument is valid at the $p < .05$ level. Using the authors' cut-off score of 10 for high risk the measure has a sensitivity of 47.96%, specificity of 81.40%, and a diagnostic efficacy of 73.1%.

Study	N	Sample Characteristics	Method and Design	Results
		16.7% without primary education; 47.5% with primary education, 18.2% with secondary education, 16.9% with professional training. **Setting**. Adult male batterers from Basque region of Spain charged between October 2005 and August 2006; The biggest difference between the severe and non severe two groups is nationality of perpetrator, with the experimental group having significantly more individuals from Latin America and Africa.		
Gondolf, E.W. & Wernik, H. (2009). Clinician Ratings of Batterer Treatment Behaviors in Predicting Reassault. *Journal*	A sample of 482 batterer program participants was drawn from a larger	**Age**. 67% of the men were under the age of 35 years. **Ethnicity**. 49% White. **Gender**. All participants were men. **Education**. Without	**Design**. Prospective study. **Measure**. The clinicians were requested to rate ten treatment behaviors. The 10 items fall into three constructs; 1. Treatment adherence factors	**Total Scores.** The mean of the 10 summed items was 35.52 (SD=9.66). **Interrater reliability**. Not applicable. **Internal consistency**. Cronbach's alpha = .97. **Validity**. Logistic regression analysis: Posttreatment 6-

Study	N	Sample Characteristics	Method and Design	Results
of Interpersonal Violence, 24(11), 1792-1815. DOI: 10.1177/08862605 08325493. **Objectives**. The objective of this study was to evaluate the utility of clinicians' ratings of treatment behaviors of batterer program participants in predicting reassault. The tested hypothesis was that negative clinician ratings would significantly predict reassault. The predictive strength of the individual items was also explored in the face of uncertainty regarding the	database of batterer program participants (N = 854), The sample included all men for whom complete clinician ratings were recorded at the time they exited the batterer treatment program.	high school education. 24%, versus 37% with more than high school education. Fully employed at intake was 68%, and 63% were blue-collar. Problems. 56% of the participants had alcoholic tendencies (e.g. scores >5 on the Michigan Alcohol Screening Test. Half of the men (49%) were previously arrested. **Setting**. Batterer programs in four sites (Pittsburg, Dallas, Houston and Denver). The programs had different durations depending on the site (Pittsburg & Dallas 3 months, Houston 5.5 months and Denver 9 months). **Relationship**. Not specified in the article. **Abuse**. 23% admitted to bruising or	(attendance, using techniques, help seeking, active engagement); 2. Problem behaviors (nonviolence, sobriety); 3. Changes in psychological mindfulness (acceptance, process consciousness, self-disclosure, use of sensitive language). The 10 items were scored on a 5-point Likert-type scale, ranging from 1 (*very little present*) to 5 (*extremely present*). **Administration**. The leaders of the batterer programs were requested to rate each participant on the ten treatment behavior items when the participant finished the program, dropped out on his own account, or was formally dismissed for irregular attendance or not paying fees. They rated the participants on each item based on what they observed in the group sessions and what the	months follow-up: The sum of the clinician's rating did not predict 'any reassault' during the posttreatment follow-ups significantly (Nagelkerke R^2 ranged from .0001 to .006). The rating sum was a significant predictor of 'severe reassault' only for completers of the programs (R^2 = .035; χ^2 = 4.14; $p < .05$; n = 337). Postintake 15-months follow-up: Any reassault (R^2 = .021; χ^2 = 5.46; $p < .05$; n = 382). Severe reassault (R^2 = .035; χ^2 = 7.92; $p < .01$; n = 382). AUC: Severe assault for 6-months posttreatment follow-up (only for completers of the programs) = .662. Postintake 15-months follow-up: Any reassault = .597, p < .01; Severe reassault = .640, p < .001. Small improvement if analysis were conducted with only program completers: Any reassault: .602, *p* < .01; Severe reassault: .646, *p* <.01; *n* = 302. The programs with a longer duration had higher effect sizes in predicting severe reassault at postintake 15-months follow-up: 9-month program: AUC = .744 (SE = .078, 95% CI = .592-.896, *n* = 134); 5.5-month program: AUC = .653 (SE = .055, 95% CI .545-761, *n* = 113). The clinical rating sum for the 3-months programs was not significant.

Study	N	Sample Characteristics	Method and Design	Results
influence of items related to group process as opposed to treatment adherence.		physically injuring their female partner sometime in the past; 85% were mandated to the programs by courts; 15% entered the program voluntarily. **Limitations**. The postintake 15-month follow-up included the period the program took place, information on the participant's behavior outside the program in this period could have influenced the clinical rating at the end of the program. This is especially relevant for the longer programs.	participants reported during the group. The programs during the study did not solicit women's reports of the men's behavior or other outside documentation to supplement the clinical observations. **Outcome**. Any reassault and severe reassaults reported by the particpant's female partners during follow-up interviews (at 6-months posttreatment and 15-months postintake follow-up). To gather this information the Physical Aggression scale of the CTS2 was used, results were dichotomized as 'Any reassault' and 'Severe reassualt' (severe reassault was defined as severe tactics on the Conflict Tactics Scale (i.e., hit with a fist, bit, kick; hit with something, attempted to hit with something; choked or burned; threatened with a knife or gun; used a knife	

Study	N	Sample Characteristics	Method and Design	Results
			or gun; forced sex against will). **Other measures**. CTS2 for follow up and the Michigan Alcohol Screening Test to detect alcohol problems.	
Hilton, N.Z., & Harris, G.T. (2009). How nonrecidivism affects predictive accuracy: Evidence from a cross-validation of the Ontario Domestic Assault Risk Assessment (ODARA). *Journal of Interpersonal Violence, 24*, 326-337. **Objective.** To cross-validate the ability of the ODARA to distinguish between subsequent	391 male intimate partner violence offenders. Of the original 737 cases drawn, all those with either a formal presentence, probation, or institutional assessment were removed for a separate study.	**Age.** *M*=38.1 years (*SD*=10.4). **Ethnicity**. Not reported **Education.** Not reported **Employment/Income.** 12% unemployed. **Setting.** Cases drawn from three police databases (greater metro of Toronto and Ontario). Drew a random sample of 105 unambigious nonrecidivists. Created 10 samples of 50 recidivists and 50 nonrecidivists by computerized random selection. **Relationship.** The	**Design.** Retrospective follow-up. **Measures**. The authors used the ODARA and the DVSR. **Administration.** Senior research assistants and graduate students coded all variables on the ODARA from the archival records. **IPV.** An assault against an adult female with whom he was or had been married or cohabitating. *Index offence*. Assault closest to, but before January 1998 (Toronto) or January 1997 (ON). **Recidivism.** Follow-up average of 60 months after the identified index offence. The outcome was	**Recidivism.** The base rate of recidivism was *n*=105 (27%). **Total Scores**. The mean ODARA score for the total sample of 391 cases was *M*= 2.42 (*SD*=1.83). The ODARA total score was significantly higher for recidivists (*M*=3.23; *SD*=1.96) than for all other cases (*M*=2.13, *SD*=1.68). The DVSR mean score was .89 (*SD*=1.07). **Inter-rater reliability.** Double blind IRR on 24 cases by two research assistants yielded IRR >.90 (previously reported in Hilton et al., 2004). **Internal Consistency.** Not reported. **Predictive validity.** The AUC for the ODARA based on the full sample was .67 (95% CI = .67 to .73), *d*=.6. The range of AUCs for the 10 samples of recidivists and nonrecidivists was .71 to .80 with a mean *AUC* of .74. For the DVSR, the *AUC* was .59 (95% CI = .52 to .68), *d*=.3 and the range across the 10 samples was .57 to .65 with a mean *AUC* of .61.

Study	N	Sample Characteristics	Method and Design	Results
recidivists and nonrecidivists.		average length of the relationship between perpetrator and victim was 98.9 months (*SD*=110.8). In the index assault 6% of perpetrators used a weapon and 48% of assaults resulted in victim injury. **IPV History.** The average number of IPV incidents in the relationship was .33 (*SD* = .80). **Other**. NA	wife assault recidvism in a police incident report or in an official criminal record in the national database. **Other.** NA **Limitations.** The ODARA was coded based on file information alone and recidivism data was based only on official records.	
Snider, C., Webster, D., O'Sullivan, C., and Campbell, J. (2009). Intimate Partner Violence: Development of a Brief Risk Assessment for the Emergency Department. *Academic Emergency Medicine*, *16*(11),	Baseline- *N* = 666 female IPV victims; Follow-up- *N* = 400 female IPV victims.	**Age.** *Baseline participants*: *M* = 30.2 years *Follow-up Participants*: *M* = 31.9 years **Education.** *Baseline participants*: Not a high school grad: 32.8%; High school/GED: 27.5%; Some college or more:	**Design.** Retrospective analysis of a prospective study. **Measures.** Pilot instrument based on Danger Assessment Scale (DA). **Administration**. Interviews were conducted at baseline and then approximately 9 months later. The baseline interview included the administration of the DA.	**Recidivism.** 14.9% of participants experienced a severe or potentially fatal IPV incident at follow-up. **Total Scores**. **Inter-rater reliability.** **Internal Consistency.** Not reported. **Predictive validity.** Using three of five 'yes' responses as a cut off for high-risk, sensitivity is 83%, specificity is 56%, and positive predictive value (PPV) is 25%. The five question screening instrument

Study	N	Sample Characteristics	Method and Design	Results
1208-1216. **Objective**. To develop a brief IPV risk assessment for use in acute care settings.		39.6%. *Follow-up Participants:* Not a high school grad: 33.3%; high school/GED: 32.3%; some college or more: 34.5%. **Employment/Income.** *Baseline participants*: Working full or part time: 39.4%; Other: 60.5%. *Follow-up Participants:* Working full or part time: 45.8%; Other: 54.3% **Setting.** IPV victims recruited from NYC and LA hospitals, family courts, shelters; **Relationship.** Approx 48% single, 41% married.	**Outcome**. The outcome at follow-up was whether the participant was the victim of severe or potentially fatal IPV. **Limitations.** Only 60% of participants were available for follow-up; Not a representative sample of the US; Women were not recruited from emergency departments where the instrument is meant to be used.	demonstrated a Hosmer-Lemeshow goodness of fit statistic of 0.12 when cross-validated. The pilot instrument was compared to report of self-perceived risk the AUC for each are 0.79 and 0.63 respectively.
Hilton, N.Z., Harris, G.T.,	150 male inmates	**Age.** *M*=30 years (*SD*=6.98) at the time	**Design**. Retrospective follow-up.	**Recidivism.** Base rate of IPV recidivism was 27%.

Study	N	Sample Characteristics	Method and Design	Results
Popham, S., & Lang, C. (2010). Risk assessment among incarcerated male domestic violence offenders. *Criminal Justice and Behavior, 37,* 815-832. **Objective**. To determine if the ODARA could predict recidivism in incarcerated male IPV offenders.	who were eligible for the domestic violence program. Of the original inmates, 37 cases had insufficient information and 9 cases were excluded because it was unclear if the person was released to the community during the follow-up period.	of the index offence. **Ethnicity.** 68% Caucasian; 30% Aboriginal. **Education.** Not reported **Employment/Income.** Not reported **Setting.** Participants were all male inmates from a Canadian correctional treatment institution admitted between 1995 to 2000. **Relationship.** 68% cohabitating; 7% dating; 7% separated at index offence. 54% of the couples had more than one child together. **IPV History.** 50% of the inmates had a prior domestic assault in official record. **Other**. Not reported	**Measures**. ODARA and the LSI-OR. **Administration.** The ODARA was coded based on institutional and medical files. The LSI-OR was completed as part of the regular intake procedure and recorded for the study. **Recidivism.** The average follow-up period was 7.98 years (*SD*=2.08). The time at risk was the time between the index and the next offence or until the date of retrieval of the records minus time in intervening custody. The average time at risk was 5.10 years *(SD*=2.98). The outcome was coded from federal criminal records and provincial offender tracking records and all postrelease criminal charges; RAs coded the information blind to ODARA and institutional variables. Charges where it was not clear who the	**Total Scores.** The mean LSI-OR score (*n*=140) was 31.04 (*SD*=6.08). ODARA recidivists (*M*=6.46, *SD*=1.69) and nonrecidivists (*M*=5.57; *SD*=2.15). Mean ODARA score for the total sample was 5.81 (*SD*=2.06). **Inter-rater Reliability.** The RAs and the first author coded 10 training cases (*ICC*=.95); IRR at follow-up on 10 cases *ICC*=.94 for time at risk and *ICC*=.76 for dichotomous recidivism. Discrepancies were resolved by consensus. **Internal Consistency.** Not reported. **Predictive Validity.** The ODARA *AUC* for recidivism was .638, 95% CI=.543-.732; the correlation between the ODARA and Cormier-Lang Score *r*=.34, *p* <.05; between the ODARA and the 5-point seriousness of postrelease DV incidents *r*=.21, *p*<.05; between the ODARA and the seriousness of criminal justice outcome *r*=.24, *p*<.01. The ODARA did not differ between individuals convicted and those charged and convicted of DV offences. Prior domestic assault and threat to harm at the index were significantly correlated with recidivism (*r*(121) =.22 and *r*(141) =.21 respectively). The prorated ODARA resulted in a small but not significant reduction in predictive accuracy. Adding an eighth ODARA category did not make a substantial difference to predictive accuracy. The ODARA was not significantly related to the elapsed time until recidivism.

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
			victim was were coded as nonrecidivists. **IPV.** Police record of violence against a marital, cohabitating, dating or former partner or self-report of offender during intake. *Index offence.* Physically assaulting a current or former marital, cohabitating, or dating partner or making a credible threat of death with a weapon in hand in the presence of the victim. **Limitations.** Limitations included scoring the ODARA from institutional file reviews completed by research assistants rather than by institutional staff. For example, victim interviews were rarely documented in the files and it can be difficult to discern prior IPV offences when the criminal record simply records "assault" for instance. The authors also noted that missing information on ODARA	Offenders with higher ODARA scores (greater than 6) had a higher and faster rate of recidivism up to about 5 years of opportunity. Cox proportional hazards regression showed that the ODARA predicted both DV recidivism and any recidivism. The *AUCs* ranged from .638 to .664 with a mean of .649 using 1 year increments from 1 to 8 years; using exact year long follow-up periods the mean was $AUC = .651$. The LSI-OR AUCs are as follows: .502 for domestic violence, .477 for nondomestic violence; .623 (significant)for violence of unknown victim-offender relationship; .576 for any violent charges; .697 (significant) for any postrelease charges. The ODARA performed better than the LSI-OR in predicting DV recidivism, the LSI-OR performed better in predicting violence of unknown victim-offender relationship and both predicted the occurrence of any postrelease charges equally. Neither predicted overall violent recidivism. Correlations with other treatment variables is available in the article.

Study	N	Sample Characteristics	Method and Design	Results
			items and on the victim–offender relationships in subsequent violence. In some cases there was ambiguity regarding recidivism.	
Belfrage, H., Strand, S., Storey, J.E., Gibas, A.L., Kropp, P.R., & Hart, S.D. (2011). Assessment and management of risk for intimate partner violence by police officers using the spousal assault risk assessment guide. *Law and Human Behavior.* Advance online publication. DOI:10.1007/s10979-011-9278-0	429 male perpetrators of IPV.	**Age.** *M* = 39 years. **Ethnicity.** Most participants were born in Sweden and of Swedish descent but 27% were either first or second generation immigrants to Sweden. **Education.** Not reported. **Employment/Income.** Not reported. **Relationship/Abuse.** 51% of perpetrators were in a relationship with the victim at the time of arrest, while the relationships of the other 49% had broken up; **Setting.** Spousal assault cases in 3	**Design.** Prospective study (18 month follow-up). **Measures.** SARA. **Administration.** Trained police officers scored the SARA based on information collected during investigation. **Outcome**. Recidivism-actual, attempted, or threatened physical or sexual violence against a past or current intimate. **Other Measures.** None. **Limitations.** Recidivism was based on subsequent police reports; Researchers did not have access to the recommended risk management strategies only the ones that were implemented.	**Recidivism.** 21% of participants had further contact with police for IPV related incidents. **Total Scores.** *M* total score for recidivists = 11.45 (*SD* = 6.10); *M* total score for non-recidivists = 8.86 (*SD* = 5.84). The stability of total SARA scores between first and second contact was ICC=.76. **Inter-rater Reliability** Not reported. **Internal Consistency.** Not reported. **Predictive Validity.** After the first contact with police, 47% of perps were rated low risk, 39% were rated moderate risk, and 14% were rated high risk. Correlation between total score and recidivism was *r* = .18 (*p* < .001), with an AUC = .63 (*SE* = .03). Correlation between summary risk rating and recidivism was *r* = .092 (*p* = .056), with an AUC = .57 (*SE* = .03). Higher scores significantly predicted recidivism. Risk management recommendations were significantly associated with a decrease in

Study	N	Sample Characteristics	Method and Design	Results
		counties of Sweden during 18 month period beginning in 2000. **IPV History.** The most common offense was assault (66% of cases); Unlawful threat was the next most common offense (21% of cases).		recidivism amongst the high risk offenders but an increase in low-risk offenders.
Connor-Smith, J. K., Henning, K. Moore, S., & Holdford, R. (2011). Risk assessments by female victims of intimate partner violence: Predictors of risk perceptions and comparison to an actuarial measure. *Journal of Interpersonal Violence, 26*(12), 2517-2550. **Objective:** To investigate the	728 female IPV victims.	**Age.** *M* = 30.9 **Setting.** Records from a Domestic Violence Assessment Center in a large Southern city, from 2003 to 2006. **Relationship/Abuse.** 70.6% not married to batterer; 73.9% had children *Due to limited demographic information the following items reflect a sub-set of sample (193 women): **Ethnicity.** 75% African American,	**Design.** Cross sectional **Measure.** Victim's risk assessment was assessed with the following yes/no question: "Do you think he will be violent with you in the next year?" **Administration**. IPV victims were interviewed shortly after (median = 5 days) their male partner had been arrested for an IPV offense. The question regarding women's perceived risk of future aggression by their partner was asked at the beginning of the interview to avoid potential influence of	59% of participants thought it was unlikely that they would be re-assaulted. Risk factors associated with victims rating their risk as high included: criminal history variables, substance abuse, and IPV variables. One notable exception for IPV variables was the experience of being assaulted while pregnant, which was associated with lower expectations of future violence, but is often included in structured risk assessments. Counter to the IPV risk literature, women with younger partners, were not as likely to expect future abuse. Family constellation variables (e.g., presence of children at home, step-parenting role) were unrelated to women's risk assessments, but are often key risk indicators on risk assessment tools. **Recidivism. NA**

Study	N	Sample Characteristics	Method and Design	Results
factors IPV victims use in their assessment of risk for re-abuse and to distinguish the differences from structured risk assessment.		21% Caucasian. **Education.** 71.4% had high school education or less. **Employment/Income.** 40.4% employed full-time outside of home. **IPV History.** More than 70% of the women reported high levels of violence.	questions about IPV experiences and offender history on women's risk assessments. **Outcome**. NA. **Other Measures.** The ODARA was simulated and 10 of the 13 ODARA items were scored from a combination of the victim interview and offender's criminal records. **Limitations.** It is unknown if the victims that could not be reached for follow up differed from those that could, although the perpetrators in the two groups did not differ significantly.	**Total Scores.** NA **Interrater reliability**. NA **Internal consistency.** Not reported. **Validity**. Not reported. **Concurrent Validity.** Not reported. **Other.** In the article the authors refer to the explanation given by the victim for their rating and whether these explanations correspond to items in an IPV risk assessment instrument as IRR. Here it is being classified as concurrent validity. Of the explanations given by the victims that reported a high likelihood of re-abuse, an average of 96% were in agreement with items that would be present on an IPV risk assessment instrument. The kappas for this group averaged .78. The agreement from those that gave low-risk explanations was 96% and the average kappa was .80. Participants' risk assessments were compared with dichotomized ODARA assessments to examine agreement between the two assessments. There was a modest agreement in risk level (kappa = .34, 67% agreement): 38.7% agreed on low risk; 28.3% agreed at high risk; 12.6% women reported high risk but had ODARA low-risk assessments; 20.3% women

Study	N	Sample Characteristics	Method and Design	Results
				reported low-risk but had ODARA high risk assessment.
				To investigate risk factors that are different between those used by women and those used by the ODARA, two groups of participants were compared: 1) Agree Lower Risk group (women rated low risk and had low-risk ODARA score) and 2) Victims Perceive Higher Risk group (women rated high risk but had low risk ODARA score). These two groups had equivalent ODARA scores.
				Between group differences were noted for dynamic risk factors, with women in the Victim Perceive Higher Risk group reporting higher levels of factors such as escalating violence, violence after relationship ended, controlling behavior, emotional abuse, and forced sex. Women in the Victim Perceive Higher Risk group also reported higher levels of employment problems and friends with a criminal history. In contrast to expected findings, the Victim Perceive Higher risk group had partners with lower frequencies of prior violent offenses, failed conditional release and prior nonviolent offenses.
Eke, A.W., Hilton, N.Z., Harris, G.T., Rice, M.E., & Houghton, R.E. (2011). Intimate	146 males who murdered or attempted	*TOTAL SAMPLE:* **Age.** *M*=37.3 years (*SD*=12.2). **Ethnicity.** Not	**Design.** Postdictive. **Measures.** ODARA. **Administration.** The ODARA was scored only on the subsample of 30	**Recidivism.** Perpetrators were selected because they all murdered or attempted to murder their partner. **Total scores.** ODARA mean score was 8.9. **Inter-rater reliability.** Two raters

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
partner homicide: Risk assessment and prospects for prediction. *Journal of Family Violence, 26*, 211-216. **Objective.** To determine if assessing the risk for femicide is feasible.	to murder a female partner (ODARA was coded on 30 of these cases).	reported. **Education.** Not reported **Employment/Income.** 28% of participants were unemployed at the time of the index offence. **Relationship.** 26% were legally married to the victim. **IPV History.** 33% had a prior arrest for domestic violence. **Other.** 42% had prior criminal charges; 8% committed suicide; 15% had a prior psychiatric history. *SUBSAMPLE:* **Age.** *M*=38.1 years (*SD*=10.8). **Ethnicity.** Not reported. **Education.** Not reported. **Employment/Income.**	cases because they included more extensive information including mental health information (i.e., psychology reports). **Recidivism.** The death or severe injury to the victim in the index offence was the outcome. **Other.** *Index Offence.* The most recent documented incident of partner assault prior to the homicide (43% of 30 had a documented assault). **Limitations.** The study is postdictive in nature. The authors had to exclude those cases where there was not an assault as defined by the ODARA before the femicide or attempted murder. This could inflate the predictive ability of the ODARA. The ODARA was only completed on 30 of the cases which limits the generalizability.	independently coded 15 randomly selected cases, *r* =.98. **Internal Consistency.** Not reported. **Validity.** All but 1 of the perpetrators scored in the highest category of risk for recidivism.

Study	N	Sample Characteristics	Method and Design	Results
		37% were unemployed at the time of the index.		
		Setting. Cases were drawn from the Violent Crime Linkage Analysis System, a Canadian federal database of violent offenders. All 146 cases from 1996 until April 1998 in which an adult male killed an intimate partner (n=91; 62%) or in which death was likely (e.g., gunshots missed victim; n=55, 38%) were chosen. Extensive police case file information was extracted from three official criminal databases for 30 of the cases.		
		Relationship. 43% were legally married to the victim.		
		IPV History. 33% had a prior arrest for		

Study	N	Sample Characteristics	Method and Design	Results
		domestic violence. **Other.** 43% had prior criminal charges; 7% committed suicide; 10% had a prior psychiatric history.		
Thijssen, J. & de Ruiter, C. (2011). Identifying Subtypes of Spousal Assaulters Using the B-SAFER. *Journal of Interpersonal Violence, 26*(7), 1307-1321. DOI: 10.1177/08862605 10369129. **Objective**. 1. Classify subtypes of spousal assaulters along the three dimensions: severity of marital violence, generality of violence, and psychopathology/p	146 perpetrator s of IPV, 38% previous convictions for violence.	**Age**. *M* = 38.5 years (range: 20-62). **Ethnicity**. 32% Dutch, 19% Surinamese, 16% Turkish, 14% Moroccan, 10% Antillean, the remaining 9% other ethnic backgrounds. **Gender**. 94% male; 6% female. **Education**. Not reported. **Setting**. Community sentence probation offices in the Netherlands between 2004 and 2005. **Relationship**. Intimate partners, otherwise not specified. **Abuse**. Intimate	**Design**. Retrospective file study **Measure**. Four items of the B-SAFER (Kropp & Hart, 2000); Violent acts; General criminality; Substance use problems; Mental Health problems. **Administration**. Files of spousal assault cases retrieved from four regional probation offices. Files contained: (a) notes from the interview with the suspect by the probation officer, (b) an official victim statement, (c) an official statement of the suspect at the time of arrest, (d) an official criminal record until the moment of arrest, and (e) an advisory report of the probation service to the	**Total Scores.** Not reported. **Interrater reliability (IRR)**. *N* = 12; ICC's single measure between .21-74, mean ICC .57. IRR's: Violent acts (.65); General criminality (.74); Substance use problems (.69); Mental Health problems (.21). **Internal consistency**. Not reported. **Validity**. Not reported.

Study	N	Sample Characteristics	Method and Design	Results
ersonality disorders. 2. The evaluation of the relationship between the subtypes of spousal assaulters and recidivism rate.		partner violence.	court. Four coders rated these files after a 1-day training. Interrater reliability was obtained on the first 12 cases after the training. **Outcome**. Re-arrest for assault, threat or stalking intimate partner within the follow-up period of 16-39 months (M = 27 months). **Other measures**. NA.	
Williams, K.R. (2012). Family violence risk assessment: A predictive cross-validation study of the Domestic Violence Screening Instrument-Revised (DVSI-R). *Law and Human Behavior, 36*(2), 120-129. DOI 10.1007/s10979-011-9272-6. **Objective**. To	3,569 family violence perpetrator s older than 16 years. This includes: (a) Spouses, former spouses, (b) Parents and their children, (c) Persons eighteen years of	**Age**. *M* = 35 (range 18-81) **Ethnicity**. Non-Latino white 47.16% (*N* = 1,683); Black 26.67% (*N* = 952); Latino 19.25% (*N* = 687); other or unknown 6.92% (*N* = 247). **Gender**. Male: 70.55% (*N* = 2.518); Female: 29.45% (*N* = 1051). **Education**. No information provided. **Setting**. Family Relations Counselors	**Design**. Cross validation field study. **Measure**. Domestic Violence Screening Instrument-Revised (DVSI-R) and Summary Risk Ratings (SRR) of the FRCs on imminent risk-to-victim and imminent risk-to-others. **Administration**. Family violence cases rated by FRC's between September 1, 2004 and May 2, 2005 in Connecticut, USA. Cases coded on the basis of perpetrator interviews, a review of police reports,	**Total Scores.** DVSI-R total risk mean score = 8.28; range from 0 to 26 in this sample (DVSI-R has a range of 0-28). **Interrater reliability**. Inter-rater reliability was a part of the training process. Case files for three separate family violence cases were used in training all FRCs. An agreement rate of 80% was used to determine whether consistency among counselors was being achieved during the training sessions. No separate tests of inter-rater reliability post completion of the training sessions have yet been conducted. **Internal consistency**. α = .75; Average inter-item covariance .18, with a range from .24 (low; Children present during any prior or current family violence incident) to .72 (high; evidence of prior family assaults, threats or

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
examine whether the prediction of perceived imminent risk-to-victim or perceived imminent risk-to-others by the Family Relations Counselors corresponded with the prediction of risk by the DVSI-R total numeric risk scores.	age or older related by blood or marriage, (d) Persons sixteen years of age or older other than those persons in subparagraph (c) Presently residing together or who have resided together, (e) Persons who have a child in common regardless of whether they are or have been married or have lived together at any time,	(FRCs) of the Family Services in Connecticut rated the DVSI-R before court sessions in the 24-hour period between arrest and initial court appearance. **Relationship**. Household relationship types: Intimate partner violence (64.62%; N = 2,292); Parent child violence (13.64%; N = 484); Other types (21.74%; N = 771). **Abuse**. The majority of the sample accounted for an unspecified form of intimate partner violence (64.62%), 13.64% of violence between parent and child. A 21.74% of the sample engaged in unspecified other types of intimate violence.	criminal history and protective order registry reviews, and victim interviews conducted by victim advocates. **Outcome**. Re-arrest within 18-months follow-up for: 1. New family violence offenses only (33.80%; N = 872); 2. Violations of protective/ restraining orders (5.53%; N = 100); 3. New family violence offenses AND violations of court orders (19.09%; N = 403); 4. 1 to 3 together (44.60%; N = 1.375); 5. Total all re-arrests combined (52.14%; N = 1861). **Other measures**. A summary risk rating was added to the DVSI-R.	arrests). **Predictive validity**. DVSI-R total score: 1. New family violence offenses only: AUC = .62 (.60 - .64); 2. Violations of protective/restraining orders: AUC = .72 (.66 - .77); 3. New family violence offenses AND violations of court orders: AUC = .73 (.70 - .75); 4. 1 to 3 together: AUC = .66 (.64 - .68); 5. Total all re-arrests combined: AUC = .66 (.66 - .68). Summary Risk Rating risk-to-victim: 1. New family violence offenses only: AUC = .62 (.60 - .65); 2. Violations of protective/retraining orders: AUC = .73; (.68 - .78); 3. New family violence offenses AND violations of court orders: AUC = .73 (.70 - .76); 4. 1 to 3 together: AUC = .66 (.64 - .68); 5. Total all re-arrests combined: AUC = .66; (.64 - .68). Summary Risk Rating risk-to-others: 1. New family violence offenses only: AUC = .63 (.61 - .64); 2. Violations of protective/retraining orders: AUC = .73; (.68 - .78); 3. New family violence offenses AND violations of court orders: AUC = .73 (.70 - .76); 4. 1 to 3 together: AUC = .67 (.65 - .69); 5. Total all re-arrests combined: AUC = .66; (.65 - .68). Concerning the predictive validity there were

PASK#15 Online Table – Partner violence risk assessment instruments

Study	N	Sample Characteristics	Method and Design	Results
	and (f) Persons in, or have recently been in, a dating relationship.			no differences between male and female perpetrators, with the exception of the 3rd outcome measure (AUC = .79 females (.73 - .84); AUC = .70 males (.66 - .73).

Note. NA = Not applicable (i.e., there was no 'other' information considered essential to report or the information would not be relevant to the publication – for instance Internal consistency would not be reported for a study examining the predictive validity of victim's self-appraisal's of their partners' level of risk). Similarly, for studies that were not prospective recidivism is generally going to be coded NA. In comparison, Not mentioned is used to indicate that the information could foreseeably have been collected and reported or is considered potentially relevant (e.g., education or setting would generally be considered essential to report in any APA publication; APA, 2010). *ns* = not significant.