Intimate Partner Violence and Risk Assessment: A Systematic Review

3 authors:

Irina Fanarraga
City University of New York - John Jay College of Criminal Justice
6 PUBLICATIONS   **7** CITATIONS

Jennifer Yang
City University of New York - John Jay College of Criminal Justice
7 PUBLICATIONS   **91** CITATIONS

Deborah Koetzle
City University of New York - John Jay College of Criminal Justice
52 PUBLICATIONS   **778** CITATIONS

Intimate Partner Violence and Risk Assessment: A Systematic Review

PREPARED BY

Irina Fanarraga
Research Associate

Jennifer Yang
Research Assistant

Deborah Koetzle
Principal Investigator
dkoetzle@jjay.cuny.edu

Department of Criminal Justice
John Jay College of Criminal Justice

Acknowledgements

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Intimate Partner Violence and Risk Assessment

A Systematic Review

EXECUTIVE SUMMARY

The current study is designed to assess the utility of recidivism risk assessments for individuals charged with, or convicted of, Intimate Partner Violence (IPV) in pretrial settings. Building on prior reviews, we completed a systematic review and meta-analysis of risk assessments among IPV aggressors to summarize their predictive validity in multiple types of recidivism, including IPV, other violence, and general recidivism.

The study includes previously completed validation studies of both IPV-specific and non-IPV-specific risk assessments. To be included, studies had to: 1) report on the validation of a risk assessment using an adult offender population with a sample (or subsample) being charged with an IPV-related offense, 2) report on an offender-focused risk assessment, 3) report on recidivism as an outcome, 4) have sufficient data to calculate an effect size, and 5) be in English or Spanish.

A total of 48 studies with 49 samples were eligible for inclusion in the current review. Each of the selected studies was coded along four categories: study characteristics, sample characteristics, methodological characteristics, and outcome characteristics.

Twenty-eight risk assessments were identified, including 16 designed to assess risk of IPV recidivism and 12 designed to assess risk of general recidivism. Results suggest that risk assessment instruments predict reoffending among IPV offender samples with moderate accuracy. The estimated summary effect across all samples (k = 49) and effect sizes (#ES = 246) was $z = 0.23$ (SE = 0.015), $t(245) = 15.69$, $p < .001$, which is equivalent to an AUC value of 0.63, considered a small to medium effect (Rice & Harris, 2005). Mean effects ranged from 0.54 for the Iowa Risk Assessment (k=1) to 0.71 for the Domestic Violence Risk Assessment Guide (k=5).

Moderator analyses largely failed to detect significant differences across study, sample, and methodological characteristics. Only 9 studies specifically identified using pretrial samples, though the majority (k=29) identified using community-based samples with a mean effect of 0.635.

Results suggest that the use of validated risk assessments can provide important information to jurisdictions seeking to assess the likelihood of future crime for individuals charged with IPV, including within pretrial settings. In selecting a risk assessment tool for implementation, jurisdictions should consider both the predictive validity of the instrument along with practical considerations including ease of use, cost, and fit with local characteristics and practices. Included in this report is a summary of each instrument, delineating the main features of each tool and considerations to be kept in mind when deciding on the adoption of an IPV risk assessment tool.

INTRODUCTION

It is well known that the United States has one of the highest incarceration rates in the world. Recent efforts towards decarceration have included a focus on pretrial detention, which accounts for nearly two-thirds of individuals being held in local jails (Zeng & Minton, 2021). Historically intended for individuals who were deemed a flight risk, it is estimated that 90% of those being held in local jails are there because they cannot afford the bail set by the court (Subramanian et al., 2015).

The growth in the pretrial detention population can be traced (at least in part) to the Bail Reform Act of 1984, which granted judges the authority to release or detain individuals based on a number of factors including the nature of charges and evidence, the individual's character and community ties, and the type of danger posed to the community (Van Brunt & Bowman, 2018).

Today, judicial decisions around detention entail a number of considerations including the likelihood of the individual fleeing, the type of crime one is accused of, the perceived dangerousness of the defendant, and the potential risk to the community the individual may pose (American Bar Association, 2019).

The consequences of pretrial detention are numerous. Pretrial detention can negatively impact important protective factors including family relationships, employment, and social support (Applegate, 2011). Individuals detained pretrial are less able to participate in their own defense, are more likely to be plead guilty, and are more likely to be convicted (Lee, 2019), often setting up a number of long-term collateral consequences associated with having a conviction (Edkins & Dervan, 2018).

Pretrial detention reforms have included expanding pretrial services, eliminating money bail, expanding the types of charges eligible for citations rather than arrest, and enhancing due process protections (Hopkins et al., 2018). However, concerns about dangerousness remain, and courts have continued to grapple with determining who is, or is not, at risk of causing further harm. This is an issue of particular importance for individuals charged with intimate partner violence (IPV).

Detention and release decisions of IPV aggressors[1] are complicated (Duane & Vasquez-Noriega, 2018). Justice system actors may be inclined to err on the side of caution because of concerns about subsequent reoffending and bad publicity (see Goddard, 2014). Yet, not all IPV aggressors will reoffend, and there is no evidence to suggest that pretrial detention reduces reoffending (George, 2012; Trevena & Poynton, 2016). And some survivors may prefer release over detention (Duane & Vasquez-Noriega, 2018).

Despite this, pretrial and criminal justice reform efforts do not universally include IPV-related charges. For example, there was opposition to the inclusion of domestic violence (DV) related charges to the bail reform laws in New York (Feldman, 2020). And mandatory minimums for IPV aggressors were increased, rather than decreased, in Iowa (HF 2399, 2016). Successfully reforming the criminal justice system must include a consideration of IPV aggressors, including the use of pretrial detention and release for this population.

_____

[1] Following Duane & Noriega-Vasquez (2018), we use the term aggressor to reflect an individual who has come into contact with the criminal justice system because they have potentially used violence.

BACKGROUND

IPV is defined as abuse or aggression that occurs in a close relationship, and includes physical violence, sexual violence, stalking and psychological aggression by a current or former intimate partner (Breiding et al., 2015). "Intimate partner" refers to both current and former spouses and dating partners (Centers for Disease Control and Prevention, 2020). While the term "domestic violence" (DV) is understood to encompass a broader kind of phenomena (it can include child or elder abuse, or abuse by any member of a household) (Barocas et al., 2016; World Health Organization, 2012), IPV and DV are routinely used interchangeably.[2]

According to the Centers for Disease Control and Prevention's (CDC) National Intimate Partner and Sexual Violence Survey (NISVS), approximately 25% or 30.0 million women and 11% or 12.1 million men in the United States have experienced contact sexual violence, physical violence, and/or stalking by an intimate partner during their lifetime (Smith et al., 2018).

Historically viewed as a private matter, IPV policies have evolved since the 1980s. Some policies, including mandatory arrests and "no-drop" prosecution (Fagan, 1996; Hanna, 1996; Messing, 2014; Murphy-Geiss et al., 2015), have led to an increase in the number of IPV cases making their way through the courts, sometimes overwhelming already overburdened jurisdictions. The use of standardized risk assessments offers the courts and police a method for identifying and prioritizing IPV cases in need of greater attention (Dutton & Kropp, 2000).

One of the first IPV-specific risk assessments was the Spousal Assault Risk Assessment (SARA), developed by Randall Kropp and colleagues in 1994 (Kropp et al., 1994). Designed to assess the likelihood of subsequent IPV behavior, the SARA contains a range of static and dynamic factors related to criminal history, psychosocial adjustment, spousal assault history, and the current offense (see Helmus & Bourgon, 2011). Since its initial development, there have been concerted efforts at designing, validating, and implementing standardized tools to be used specifically with IPV-related populations.

Dutton and Kropp (2000) argue there are some advantages to assessing for risk of IPV rather than general violence for this population. First, base rates for "spousal physical assault" are relatively high (25% - 50%) compared to general violence; using IPV assessments help to reduce the false-positive error rate. Second, IPV is unique in that there is a survivor who can act as a source of critical information about the history and the personality of the aggressor.

Finally, the authors argue that IPV risk assessment instruments manage to avoid the limitations of more general violence assessment instruments as identified by Monahan and Steadman (1994) −that is, that they rely on weak criterion variables that are too broadly defined or unreliably measured− because there are research-based risk factors that are uniquely related to IPV offending.

---

[2] Because our focus is on intimate partner violence, we use the term IPV throughout.

CURRENT STUDY

The current study builds on prior reviews by including both published and unpublished studies of both IPV-specific and non-IPV specific risk assessments. Specifically, the current study summarizes the predictive validity of risk assessment in assessing the risk of IPV, non-IPV violence, and general recidivism among samples of IPV aggressors.

RESEARCH METHODS

Search Strategy

Multiple databases and sources were searched between December 2019 and April 2020 to identify published and unpublished validation studies. In addition, we searched reference lists of previously conducted meta-analyses.

Search terms included ("domestic violence" or "intimate partner violence") and ("risk assessment" or "danger assessment"). In addition, searches were conducted for specific risk assessments. A complete list of these assessments is in Appendix A.

Inclusion Criteria

To be included in the analysis, studies had to meet the following criteria:

1. Sample consists of adults (18+) charged with IPV-related offenses
2. Assessment designed to predict risk of recidivism
3. Reported on results of predictive validation study
4. Recidivism as an outcome
5. Reported sufficient data to calculate an effect size for the IPV/DV sample
6. Published in English or Spanish

Academic Databases
PyschInfo, ScienceDirect Web of Science, ProQuest Dissertations & Theses, Sociological Abstracts, Social Service Abstracts, Criminal Justice Abstracts, National Criminal Justice Reference, Sage Criminal Justice Journals, CINAHL Complete, PAIS International

Websites
National Institute of Justice, OpenGrey, and Violence Against Women

When multiple studies reported on the same sample, the study with the longer follow-up period was selected.

Additional details regarding the methods and analysis can be found in the technical appendix at the back of this report.

Coding

Each study was coded on a number of items including study, sample, and methodological characteristics, along with outcomes as indicated below. Each study was coded by the first and third author of this report. Any disagreements were discussed, and a consensus was reached.

A complete code sheet is available from the authors upon request.

Data Analysis

Mean effects were estimated using a random-effects approach. All estimated effects were reported as Area Under the Curve (AUC; Rosenthal, 1994) to aid interpretation and are represented as effect sizes (ES). A multi-level approach was used, allowing for multiple effects from a single study to be used, and all analyses were conducted in the statistical software environment R (version 4.0.2).

One effect size was identified as an outlier. In order to examine the effect of this outlier, we compared the results with and without this effect size. As there were no discernible differences in the results, the outlier was included in the final analyses.

Publication Bias

There are several forms of bias that may impact the results of meta-analyses including publication bias. We tested for publication bias using a funnel plot based trim-and-fill analysis (Duval & Tweedie, 2000). Though there was some evidence of bias in the distribution of effects, the analysis failed to find evidence of significant bias in the results.

Study — Publication type, publication year, jurisdiction size, risk assessment, modified version

Sample — Sample size, gender, race & ethnicity distributions, relationship type, pretrial status, setting

Methods — Researcher involvement, fidelity, construction sample, multiple metrics, blind review

Outcomes — Type of recidivism, type of charge, charge level, length of follow-up

Area Under the Curve (AUC): A measure of predictive accuracy. An instrument with a AUC=.5 has no predictive ability and is the same as chance. An instrument with a AUC=1.0 is 100% accurate.

Outlier: A value for a given variable that is drastically higher or lower than the other values for that same variable within a dataset.

Funnel plot based trim-and-fill analysis: A test for publication bias, this analysis provides an estimate of the mean effect size if missing studies were included in the meta-analysis. Publication bias exists when the estimated effect and the calculated effect are significantly different from one another.

RESULTS

Over 2,000 results were identified through the search process, representing 260 unique studies. Some studies were excluded based on the review of the abstract and title. Two-hundred and twelve studies were fully reviewed; of these 48 met our inclusion criteria and were included in the analyses.

As detailed in the table, potential studies were excluded for a number of reasons including sample type, lack of outcomes, and a focus on survivor assessment. Multiple efforts were made to contact authors for studies that we could not locate or reported insufficient data.

Reason for exclusion	*n*
Sample	72
Survivor-oriented assessment	23
Duplicate sample	8
No risk assessment of offending	17
No recidivism outcome	35
Insufficient data	1
Unable to locate	8

Study Characteristics

The majority of included studies were conducted in North America and published in a peer reviewed outlet. Publication dates ranged from 1995 to 2020. Validation studies were generally collected on state or country level populations; only two studies were reported to have been conducted in non-urban settings.

A total of 28 risk assessments were identified across the 48 studies. These included a mix of IPV-specific and general recidivism risk assessments. Six studies reported on modified versions of existing instruments.

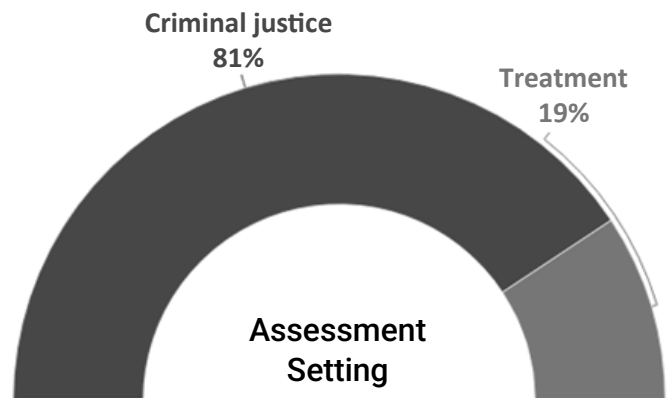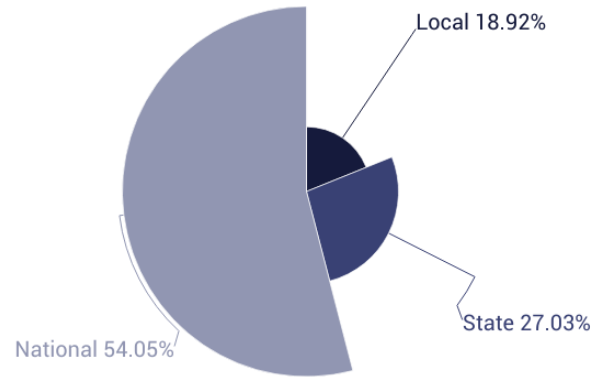| Study characteristics		%
Publication type	Peer reviewed	66.7
	All other	33.3
Publication Year	1995-2005	22.9
	2006-2013	39.6
	2014-2020	37.5
Country	North America	69.4
	All other	30.6
Jurisdiction	Urban	28.6
	Nonurban	5.7
	Mixed	65.7
Risk assessment version	Modified	13
	Original	87

Methodological Characteristics

Fifty-four percent of the studies included assessment results conducted by practitioners rather than researchers, and only 6 (12%) used construction samples. As a measure of quality, we considered whether studies reported on practices to monitor fidelity and whether multiple metrics were reported. Only 10 (20%) reported monitoring fidelity while 57% reported on multiple metrics, including reliability and construct validity.

Recidivism Jurisdiction

Local 18.92%

State 27.03%

National 54.05%

Construction sample: Sample used to develop an instrument

Criminal justice 81%

Treatment 19%

Assessment Setting

Methodological Characteristics		%
Design	Prospective	53.2
	Retrospective	46.8
Assessment by	Researcher	45.8
	Other	54.2
Fidelity	Yes	20.4
	No	79.6
Construction Sample	Yes	12.2
Multiple Metrics	Yes	57.1
	No	42.9

Mean follow-up period in months
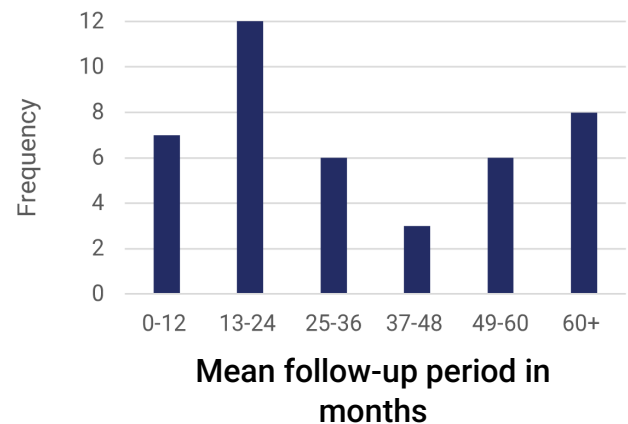
Sample Characteristics

The majority of studies utilized samples that were 100% male. Relatively few reported on the racial make-up of the sample; among those that did, the majority reported having diverse samples. Almost all of the samples included aggressors engaged in heterosexual relationships. Only 4 studies reported including some cases with same sex partners.

Gender

All male   Mixed
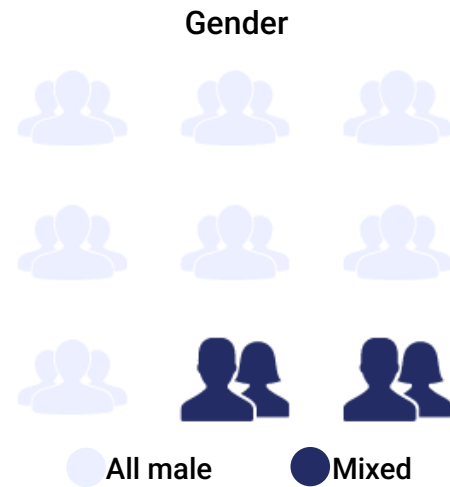
Criminal Justice Characteristics

Though the current study is primarily interested in considering the role of risk assessment for the management of IPV aggressors in pretrial settings, only 9 studies (18%) included reported the validity of risk assessments in pretrial settings. The remaining studies either were conducted in post-conviction settings or were unclear about the legal status of the sample.

However, the majority of studies were conducted with community-based samples, with assessments conducted in criminal justice settings (i.e., probation office, jail) rather than treatment settings. Thus, although few studies explicitly focused on pretrial cases, the majority of the studies represent settings and practices that are common to pretrial.

Sample Type

● Non-Pretrial Sample
● Pretrial Sample

18%

82%

Criminal Justice Setting

● Community ● Prison
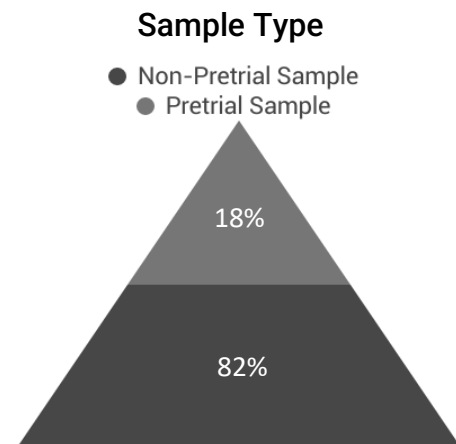
Overall Effect Size

Calculated effect sizes for each study are reported in the technical appendix. Across the 48 studies, there were a total of 246 effect sizes extracted. The mean effect size across the studies was 0.23 (SE=0.015, p<.001). This can be represented as an AUC value of 0.63.
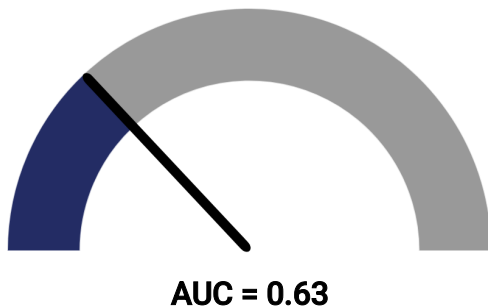
AUC = 0.63

An AUC = 0.63 is considered a small to medium effect (Rice & Harris, 2005) suggesting that the use of risk assessment is significantly more predictive than chance alone. In other words, overall, these tools significantly improve the odds of identifying individuals at risk of recidivism.

Overall, the results of the meta-analysis suggest small to moderate predictive validity of risk assessments among IPV aggressor samples. However, important differences exist in the validity and potential utility of the different tools.

We conducted additional analyses to explore whether the results varied across instruments or other factors. First, we calculated overall effect sizes for each risk assessment. Next, we examined whether the overall results varied by different approaches to risk assessment.

Summary of Instruments and Mean Effects

When selecting risk assessments for implementation, jurisdictions should be mindful of the validity of the tool, along with how well the tool fits with local context.

Table 1 reports the mean ES for each instrument. Following Table 1, we provide fact sheets for each that was included in more than one study and is available in the public domain.

Considerations for Implementation?

❑ **Prior validation**
Has the assessment been validated on populations similar to ours?

❑ **Cost**
What are the costs per use?
For training?

❑ **Staff qualifications**
Can front line staff be trained to complete the assessment?
Are special certifications required?

❑ **Time**
How long does it take to complete the assessment?

❑ **Training**
What type of training is required? Is it online or in-person?

❑ **Current assessments in use**
How does this assessment fit with our current process?

❑ **Format**
Is it available in a format that fits within our system?

Table 1. Mean Effect Size by Instrument

	k	# ES	Mean ES			k	# ES	Mean ES
Overall***	49	246			PAPS	1	2	0.64
B-SAFER 1	3	7			PCL-R***	5	9	0.63
B-SAFER 2**	2	10	0.56		PRA***	1	9	0.58
DRAOR*	1	4	0.63		PST-VC	1	2	0.67
DVRAG***	5	9	0.71		SAM*	1	8	0.57
DVSI***	5	24	0.61		SARA***	13	29	0.65
DVSI-R*	6	12			SIR-R1	1	3	0.59
DVSR*	2	3			SRA-PA***	1	9	0.58
FVIR***	2	10	0.63		SVRA-1	1	2	0.58
IRA	1	4			VP-SAFvR	1	2	0.62
LSI/R/CMI/OR***	5	18			VPR***	2	7	0.58
ODARA***	17	53			VRAG	3	3	0.70
OST/FROST***	1	7						

*<.05, **<.01, ***<.001

INSTRUMENT FACT SHEETS

When selecting risk assessments, jurisdictions should be mindful of both validity and how well the tool fits with local practice and resources. In the following section, we rate the mean ES for reach instrument and discuss important considerations for adopting and implementing these tools.

As detailed in the legend below, each fact sheet reports on a number of factors including:

- Assessment approach
- Type of recidivism
- Samples included

We also report on practical concerns, information sources, available formats, and provide links to more information.

Legend

APPROACH

SPJ Structured Professional Judgement	A Actuarial

RECIDIVISM

IPV Intimate Partner Violence	G General

EFFECT SIZE (ES)

S Small	M Moderate	NS Not Significant

AVAILABLE FORMATS

Pen and Paper	Computerized Version

Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER-1 and B-SAFER-2)

Approach	**SPJ**
Recidivism	**IPV**
ES = 0.59*	**S**

*V1 and V2 combined

Samples	Practical Concerns	Information Sources
✓ General	⏱ Varies	☑ Client Interview
✓ Pretrial	$ Per Use	☑ Survivor Input
✓ Women	# 10 - 15 Items	☑ Case File

Quick Facts

The **B-SAFER** is a briefer version of the Spousal Assault Risk Assessment Guide (SARA). It has 2 versions.

Version 1 has two domains: Spousal Assault and Psychological Assessment.

Version 2 has three domains: Perpetrator Risk Factors, Psychosocial Adjustment and Victim Vulnerability Factors.

✖ The B-SAFER-1 was not significantly associated with recidivism.

⚠ The B-SAFER-2 has a significant but small effect size.

✓ The B-SAFER-2 is relatively quick to complete and can be used by frontline staff.

Available Formats: 🖥 📄

For More Information:
https://protect-international.com/

Domestic Violence Risk Appraisal Guide (DVRAG)

Approach	A
Recidivism	IPV
ES = 0.71	M

Samples	Practical Concerns	Information Sources
✓ General	🕐 20 to 30 minutes	✓ Clinical Data
✗ Pretrial	$ No Cost	✗ Survivor Input
✗ Women	# 14 Items	✓ Case File

Quick Facts

The **DVRAG** is a 14-item actuarial IPV risk assessment instrument.

It is a combination of the ODARA items and the total score from the PCL-R.

It was designed to be a more in-depth risk assessment tool than those used by frontline workers by including information available to clinicians and forensic professionals.

✓ The DVRAG is significantly associated with recidivism, with a moderate effect size.

⚠ Requires clinical background or access to clinical information.

✗ More resource intensive than other instruments, less useful for pretrial settings.

Available Formats:

For More Information:
Hilton et al. (2008)
https://doi.org/10.1007/s10979-007-9088-6

Domestic Violence Screening Instrument (DVSI) & Domestic Violence Screening Instrument Revised (DVSI-R)

Approach	**A/S**
Recidivism	**IPV**
ES = 0.60*	**S**

*DVSI/DVSI-R combined

Samples	Practical Concerns	Information Sources
⊘ General	🕐 Not specified	☒ Client Interview
⊘ Pretrial	$ Not specified	☒ Survivor Input
⊘ Women	# 11 - 36 Items	☑ Case File

Quick Facts

The **DVSI** is a 12-item actuarial IPV risk assessment instrument.

The **DVSI-R** is an 11-item IPV risk assessment instrument. While mostly actuarial, it includes two items that ask the evaluator to provide their own professional assessment regarding imminent risk of future violence to the victim and to some other persons known to the perpetrator and/or victim.

⚠ The DVSI is significantly associated with recidivism, with. a small effect size.

⚠ The DVSI-R is significantly associated with recidivism, with. a small effect size.

✓ The DVSI/DVSI-R can be quickly scored from criminal justice case file information.

Available Formats: Not specified

For More Information:
Williams & Grant (2006)
https://doi.org/10.1177/003335490612100408

Hare Psychopathy Checklist-Revised (PCL-R)

Approach	SPJ
Recidivism	G
ES = 0.63	S

Samples	Practical Concerns	Information Sources
✓ General	🕐 Up to 3 hours	✓ Client Interview
✗ Pretrial	$ Per Use	✗ Survivor Input
✗ Women	# 20 Items	✓ Case File

Quick Facts

The **PCL-R** is a 20-item psychological tool used to assess for psychopathy in forensic populations.

Ratings are based on interview and a review of collateral files and official records conducted by a highly trained clinician.

Demonstrates small to moderate predictive validity for general violence for men and women.

⚠️ The PCL-R is significantly associated with recidivism, with a small effect size.

⚠️ Requires highly trained clinical or forensic staff to administer it.

✗ Relatively long and involved assessment, might be impractical for use at the pretrial level.

Available Formats: 📄

For More Information:
https://www.pearsonassessments.com
http://www.hare.org/scales/pclr.html

Level of Service Inventory - Revised (LSI-R) Level of Service Inventory - Ontario Revision (LSI-OR), & Level of Service/Case Management Inventory (LS/CMI)

Approach   **A**

Recidivism   **G**

ES = 0.66*   **M**

LSI-R / LSI-OR / LS/CMI combined

Samples	Practical Concerns	Information Sources
✅ General	🕐 30 to 45 Minutes	☑ Client Interview
❌ Pretrial	$ Per Use	❌ Survivor Input
✅ Women	# 43 Items	☑ Case File

Quick Facts

The **LSI-OR** and **LS/CMI** have 43 items and are similar in format to the **LSI-R**, which has 54 items.

Results are based on a semi-structured interview and review of collateral information, including criminal records and brief interviews with family members or others.

The suite of assessments enjoy general and violent predictive validity across populations.

✅ The LS tools are significantly associated with recidivism, with a moderate effect size.

⚠️ A screening version is available but has yet to be validated with IPV aggressors.

✅ Frontline staff can be trained to properly conduct the assessment.

Available Formats: 🖥 📄

For More Information:
https://storefront.mhs.com/collections/lsi-r

Ontario Domestic Assault Risk Assessment (ODARA)

Approach	A
Recidivism	IPV
ES = 0.67	M

Validated Samples
- ✓ General
- ✓ Women
- ✓ Same-Sex

Practical Concerns
- 🕐 Varies
- $ No Cost
- # 13 Items

Information Sources
- ✗ Client Interview
- ✓ Survivor Input*
- ✓ Case File

* Optional

Quick Facts

The **ODARA** is a 13-item actuarial IPV risk assessment instrument.

Items cover domestic and non-domestic criminal history, threats and confinement during the index incident, children in the relationship, substance abuse, and barriers to victim support.

While created to be used with adult men, the ODARA has also been validated with female samples.

✓ The ODARA is significantly associated with recidivism, with a moderate effect size.

✓ Less time-consuming than other instruments, designed to be used by frontline staff.

⚠ Formal training is associated with improve scoring accuracy.

Available Formats: 📄

For More Information:
https://odara.waypointcentre.ca/

Spousal Assault Risk Assessment Guide (SARA)

Approach	SPJ
Recidivism	IPV
ES = 0.65	M

Samples	Practical Concerns	Information Sources
✓ General	🕐 60 to 90 Minutes	✓ Client Interview
✓ Women	$ Per Use	✓ Survivor Input
✓ Same-Sex	# 24 Items	✓ Case File

Quick Facts

The **SARA** is a set of structured professional guidelines for assessing risk of spousal violence. The latest version (SARA-V3) consists of 24-items divided into three scales: violent behavior, offenders' psychological adjustment factors, and victim vulnerability factors.

Access to clinical records and criminal justice case files are required. Interviews with the survivor and aggressor are also necessary.

✓ The SARA is significantly associated with recidivism, with a moderate effect size.

⚠ Requires a trained evaluator with experience working with perpetrators and victims of IPV.

✗ Relatively long and involved assessment, might be impractical for use at the pretrial level.

Available Formats: 📑

For More Information:
https://protect-international.com/

Violence Risk Appraisal Guide (VRAG)

Approach	A
Recidivism	G
ES = 0.70	NS

Samples

✓ **General**

✗ **Pretrial**

✗ **Women**

Practical Concerns

🕐 **60 to 240 Minutes**

$ **For Training**

\# **12 Items**

Information Sources

✗ **Client Interview**

✗ **Survivor Input**

✓ **Case File**

Quick Facts

The **VRAG** is a 12-item actuarial risk assessment instrument for the prediction of violent recidivism.

The items cover static demographic, childhood history, criminal history, and psychiatric assessment variables, including a PCL-R score.

The VRAG was designed to be scored based on clinical records rather than structured interviews.

✗ The VRAG is not significantly associated with recidivism; this may be due to the small number of studies.

⚠ Training (of approximately 8 hours) and relevant professional background required.

✗ Relatively long and involved assessment, might be impractical for use at the pretrial level.

Available Formats: 📝

? **For More Information:**
http://www.vrag-r.org/

Moderator Analysis

In addition to estimating the overall effects we also explored whether the overall ES varied by a number of characteristics, including assessment characteristics, and study, sample, and methodological characteristics.

The full results are reported in the technical appendix, while key findings are detailed here.

Assessment Characteristics

- There were no significant differences in the predictive validity of assessments designed to predict general recidivism and those designed to predict IPV.

- Twenty-nine studies reported on structured professional judgement assessments and 39 reported on actuarial assessments. Both approaches performed equally well.

- Assessments performed slightly better predicting general recidivism compared to IPV recidivism. There were no significant differences when predicting non-IPV violence or general family violence.

- A total of 10 assessments included in the analyses require survivor input as part of the assessment process and 18 assessments did not require survivor input. Though those requiring survivor input had a slightly lower effect size, it was not significantly different

Sample Characteristics

- There were no significant differences in terms of race, gender, or relationship type.

Table 2. Moderator Analysis

	Moderator	AUC
Overall		0.63
Target	General instruments	0.64
	IPV instruments	0.63
Approach	Structured clinical judgement	0.62
	Actuarial	0.64
Charge Type	General	0.67
	Violence	0.62
	Domestic violence	0.63
	IPV	0.62
Gender	All Male	0.64
	Mixed	0.61
Race	>80% White	0.66
	<80% White	0.62
Relation-ship	Opposite Sex	0.63
	Mixed opposite & same sex	0.64

Note. All significant at the $p < .001$ level

Criminal Justice Characteristics

- Assessments demonstrated predictive validity in both **pretrial settings and other settings**.

- **The mean ES was lower for pretrial settings compared to other settings**. However, this may be a function of the number of studies (k=9) conducted with pre-trial populations.

- **Tools performed better with community-based samples** compared to prison-based samples.

- There were **no differences** between assessments conducted in criminal justice settings and those conducted in treatment settings.

- **More research** is needed on the validity of these tools with pre-trial populations.

Methodological Characteristics

- **Effects were stronger when assessments were completed by researchers** compared to practitioners. This is consistent with prior research showing effects tend to be stronger when researchers conduct the intervention under study.

- Monitoring **fidelity** was associated with **stronger effects**.

Table 2 (continued).

	Moderator	AUC
Pretrial	Yes	0.61
	No	0.64
Criminal Justice Setting	Community	0.64
	Prison	0.58
Assessment Setting	Criminal Justice	0.63
	Treatment	0.68
Assessment By	Researcher	0.67
	All others	0.61
Fidelity	Yes	0.68
	No	0.62
Construction Sample	Yes	0.68
	No	0.63
Multiple Metrics	Yes	0.64
	No	0.63

Note. All significant at the $p < .001$ level

DISCUSSION

Efforts to reform pretrial detention and money bail must consider the issue of individuals charged with violent offenses, including those charged with IPV. The current results demonstrate the ability of standardized risk assessment tools to predict the likelihood of subsequent criminal activity, including both violent and general recidivism.

Taken as a whole, however, the results are promising. We identified eight assessments that are available and have been subjected to more than one validation study. Of these, five were identified as having moderate size effects:

- Domestic Violence Risk Appraisal Guide
- Domestic Violence Screening Instrument-Revised
- Level of Service-Inventory
- Ontario Domestic Assault Risk Assessment
- Spousal Assault Risk Assessment

Three assessments were found to have evidence of small effect sizes:

- Brief Spousal Assault Form for the Evaluation of Risk, Version 2
- Domestic Violence Screening instrument
- Hare Psychopathy Checklist-Revised

The results indicate that these eight tools are better than 50/50 at identifying higher-risk individuals. These effects held across different study, sample, and methodological characteristics.

Jurisdictions seeking to implement risk assessments for IPV aggressors must consider a number of key considerations beyond the

validity of a given tool. These include issues like costs and resources, training and QA, and local context. An implementation checklist is provided in Appendix B to assist jurisdictions in this process.

Key Considerations

Though the findings support the validity of the majority of the tools, the use of standardized risk assessments is not without controversy, including concerns regarding racial bias. Though the current study failed to find differences across sample characteristics, only 20 studies reported details on race or ethnicity and additional research is needed on this front.

Similarly, the majority of studies reported on men who had been charged of convicted of aggression or violence against women. Few studies reported on same-sex partners and only one study reported results for female aggressors. Though the research on general risk assessments indicates that tools are valid across gender, more research is needed to assess the validity of these tools for individuals who identify as women or LGBTQ.

- Jurisdictions seeking to adopt risk assessments for IPV aggressors should include plans to validate and test for bias as part of the implementation plan.

Notably, assessments demonstrated greater predictive validity when they were conducted by researchers compared to practitioners or other actors. Though the reasons for this are unclear, it is consistent with research on program effectiveness, which finds improved outcomes when an evaluator is involved (Gendreau et al., 2006). This finding may reflect a difference in the quality of the

- When considering the use of assessments, it is critical that agencies plan for training staff on conducting assessments and interpreting the results.
- Jurisdictions should ensure that assessors are properly trained with oversite and quality assurance processes in place to ensure tools are used with fidelity.

The use of validated risk assessments can provide important and meaningful information to those tasked with making decisions regarding pretrial detention and community supervision. Jurisdictions seeking to improve their decision-making should adopt validated tools that can be easily implemented and used, in a manner consistent with local context and practice.

Finally, the current study was designed to assess the utility of risk assessments for IPV aggressors in a pretrial setting. Though the overall results are promising, it must be acknowledged that only 9 studies reported on pretrial populations. Although more research is needed in this area, the evidence is promising.

- There is promising evidence to suggest the tools identified above can be used in pretrial settings.

To summarize, more research is needed to assess the generalizability of these tools, across settings and populations. However, it is clear that risk assessments of IPV aggressors can provide meaningful information for identifying those likely to be arrested or charged for new offenses.

CONCLUSION

Keeping survivors of IPV safe should continue to be an important priority for criminal justice actors. However, this should not be taken to mean that IPV aggressors should be systematically detained on the basis of their charges alone. Doing so undermines efforts aimed at meaningful pretrial reform.

REFERENCES

American Bar Association. (2019, September 9). How courts work. https://www.americanbar.org/groups/public_education/resources/law_related_education_network/how_courts_work/bail/

Applegate, B. K. (2011). Jails and pretrial release. In M. Tonry (Ed.), The Oxford Handbook of Crime and Criminal Justice (pp. 795–824). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195395082.013.0025

Barocas, B., Emery, D., & Mills, L. G. (2016). Changing the domestic violence narrative: Aligning definitions and standards. Journal of Family Violence, 31(8), 941–947. https://doi.org/10.1007/s10896-016-9885-0

Breiding, M. J., Basile, K. C., Smith, S. G., Black, M. C., & Mahendra, R. (2015). Intimate partner violence surveillance uniform definitions and recommended data elements—Version 2.0. Centers for Disease Control and Prevention National Center for Injury Prevention and Control. https://nccpsafety.org/assets/files/library/Intimate_Partner_Violence.pdf

Centers for Disease Control and Prevention. (2020). Preventing intimate partner violence. https://www.cdc.gov/violenceprevention/pdf/ipv/IPV-factsheet_2020_508.pdf

HF 2399, 2399, Iowa General Assembly, 86 (2016).

Duane, M., & Vasquez-Noriega, C. (2018). Pretrial strategy for handling intimate partner violence cases (p. 17). Urban Institute. https://www.urban.org/sites/default/files/publication/99167/pretrial_strategy_for_intimate_partner_violence.pdf

Dutton, D. G., & Kropp, P. R. (2000). A review of domestic violence risk instruments. Trauma, Violence, & Abuse, 1(2), 171–181.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics, 56(2), 455–463. https://doi.org/10.1111/j.0006-341x.2000.00455.x

Edkins, V. A., & Dervan, L. E. (2018). Freedom now or a future later: Pitting the lasting implications of collateral consequences against pretrial detention in decisions to plead guilty. Psychology, Public Policy, and Law, 24(2), 204–215. https://doi.org/10.1037/law0000159

Fagan, J. (1996). The criminalization of domestic violence: Promises and limits [National Institute of Justice Research Report]. National Institute of Justice.

Feldman, L. (2020, January 17). New York's bail reform ignores risks to victims of domestic violence. https://auburnpub.com/opinion/columnists/feldman-new-yorks-bail-reform-ignores-risks-to-victims-of-domestic-violence/article_6eb5d190-2c56-5c46-bf4e-cd8d57d9a4e5.html

Gendreau, P., Smith, P., & French, S. A. (2006). The theory of effective correctional intervention: Empirical status and future directions. In F. Cullen, J. Wright, & M. Coleman (Eds.), Taking stock: The status of criminological theory. (pp. 419–446). Transaction Press.

George, T. P. (2012). Domestic violence sentencing conditions and recidivism. Washington State Center for Court Research, Administrative Office of the Courts.

Goddard, R. G. (2014). When it's the first time every time: Eliminating the "clean slate" of pretrial diversions in domestic violence crimes. Valparaiso University Law Review, 49(1), 267–310.

Hanna, C. (1996). No right to choose: Mandated victim participation in domestic violence prosecutions. Harvard Law Review, 109(8), 1849–1910. JSTOR. https://doi.org/10.2307/1342079

Helmus, L., & Bourgon, G. (2011). Taking stock of 15 years of research on the Spousal Assault Risk Assessment guide (SARA): A critical review. International Journal of Forensic Mental Health, 10(1), 64–75. https://doi.org/10.1080/14999013.2010.551709

Hopkins, B., Bains, C., & Doyle, C. (2018). Principles of pretrial release: Reforming bail without repeating its harms. Journal of Criminal Law and Criminology, 108(4), 679–700.

Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1994). Manual for the Spousal Assault Risk Assessment Guide. British Columbia Institute on Family Violence.

Lee, J. G. (2019). To detain or not to detain? Using propensity scores to examine the relationship

between pretrial detention and conviction. Criminal Justice Policy Review, 30(1), 128–152. https://doi.org/10.1177/0887403416668016

Messing, J. T. (2014). Evidence-based prosecution of intimate partner violence in the post Crawford era: A single-city study of the factors leading to prosecution. Crime & Delinquency, 60(2), 238–260. https://doi.org/10.1177/0011128710362056

Monahan, J., & Steadman, H. J. (Eds.). (1994). Violence and mental disorder: Developments in risk assessment. University of Chicago Press.

Murphy-Geiss, G. E., Roberts, W. T., & Miles, D. J. (2015). One size does not fit all: A case study of an alternative intimate partner violence court. Feminist Criminology, 10(4), 348–367. https://doi.org/10.1177/1557085114554258

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. Psychological Bulletin, 86(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Rosenthal, R. (1994). Parametric measures of effect. In H. Cooper & L. V. Hedges (Eds.), The Handbook of Research Synthesis (pp. 231–244). Russell Sage Foundation.

Smith, S. G., Zhang, X., Basile, K. C., Merrick, M. T., Wang, J., Kresnow, M., & Chen, J. (2018). The National Intimate Partner and Sexual Violence Survey (NISVS): 2015 Data Brief – Updated Release. National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Subramanian, R., Delaney, R., Roberts, S., Fishman, N., & McGarry, P. (2015). Incarceration's front door: The misuse of jails in America. Vera Institute of Justice. https://www.vera.org/downloads/publications/incarcerations-front-door-report_02.pdf

Trevena, J., & Poynton, S. (2016). Does a prison sentence affect future domestic violence reoffending? Crime and Justice Bulletin, No. 190. https://www.bocsar.nsw.gov.au/Documents/CJB/Report-2016-Does-a-prison-sentence-affect-future-domestic-violence-reoffending-cjb190.pdf

Van Brunt, A., & Bowman, L. E. (2018). Toward a just model of pretrial release: A history of bail reform and a prescription for what's next. Journal of Criminal Law and Criminology, 108(4), 701–774. https://doi.org/0091-4169/19/10804-0701

World Health Organization. (2012). Understanding and addressing violence against women—Intimate Partner Violence (WHO/RHR/12.36). https://apps.who.int/iris/bitstream/handle/10665/77432/WHO_RHR_12.36_eng.pdf;jsessionid=B3A739C287A044C338CFBDB780F401CC?sequence=1

Zeng, Z., & Minton, T. D. (2021). Jail inmates in 2019 (NCJ 2556608). Bureau of Justice Statistics. https://bjs.ojp.gov/content/pub/pdf/ji19.pdf

APPENDIX A: LIST OF ASSESSMENTS

Brief Spousal Assault Form for the Evaluation of Risk B-SAFER

Chinese Risk Assessment Tool for Perpetrators CRAT-P

Classification of Violence Risk COVR

Correctional Offender Management Profiling for Alternative Sanctions COMPAS

Conflict Tactics Scale Revised / Revised Conflict Tactics Scales CTS2

Domestic Violence Inventory DVI

Domestic Violence Risk and Needs Assessment DVRNA

Domestic Violence Risk Appraisal Guide DVRAG

Domestic Violence Screening Instrument DVSI

Domestic Violence Screening Instrument–Revised DVSI-R

Domestic Violence Supplementary Report DVSR

Family Violence Risk Assessment Tool FVRAT

General Statistical Information on Recidivism GSIR

Hare Psychopathy Checklist–Revised PCL-R

Hare Psychopathy Checklist: Screening Version PCL-SV

Historical Clinical Risk Management-20 HCR-20

Historical part of the HCR-20 H-10

Kingston Screening Instrument for Domestic Violence K-SID

Jellinek Inventory for assessing Partner Violence J-IPV

Intimate Assault Screening Scale IASS

Intimate Femicide Screening Scale IFSS

Level of Service/Case Management Inventory LS/CMI

Level of Service Inventory LSI

Level of Service Inventory-Revised LSI-R

Offender Group Reconviction Scale OGRS

Ohio Risk Assessment System ORAS

Ontario Domestic Assault Risk Assessment ODARA

Partner Abuse Prognostic Scale PAPS

PErpetrator RaPid Scale PERPS

Police Screening Tool for Violent Crimes PST-VC

Primary Risk Assessment PRA

Propensity for Abusiveness Scale PAS

Public Safety Assessment PSA

Recidive Inschattingsschalen (RISc) [Recidivism Risk Assessment Scales in English]

Secondary Risk Assessment for Partner Abusers SRA-PA

Severe Intimate Violence Partner Risk Prediction Scale SIVIPAS

Spousal Assault Risk Assessment Guide SARA

Spouse Violence Risk Assessment Inventory SVRA-I

Two-Tiered Violence Risk Estimates Scale TTV

Valoración Policial del Riesgo VPR

Valoración Policial de la Evolución del Riesgo VPER

Victoria Police Screening Assessment for Family Violence Risk VP-SAFvR

VioGén

Violent Offender Risk Assessment Scale VORAS

Violence Risk Appraisal Guide VRAG

Violence Risk Scale VRS

Violence Risk Screening- Police Version V-RISK-POL

Violence Screening Checklist VSC

APPENDIX B: CHECKLIST FOR IMPLEMENTATION OF IPV-RELATED RISK ASSESSMENT

Step 1. Assessing for organizational readiness
- ❑ Stakeholders
- ❑ Administrators
- ❑ Staff

Step 2. Form a working group
- ❑ Include a maximum of 10 people
- ❑ Include a cross-section of staff, including those who will be responsible for conducting the assessment
- ❑ Arrive at a group decision on the following:
  - How will you educate stakeholders about the assessment?
  - How will the results be used?
    - ∗ Pretrial release decision
    - ∗ Diversion decision
    - ∗ Community supervision decision
- ❑ Identify external technical assistance if needed
- ❑ How will assessment information be shared with partners?
- ❑ Will reassessments be conducted? When?
- ❑ What resources are available to extract and analyze the assessment data?

Step 3. Select an Assessment
- ❑ Factors to consider include:
  - Prior validation
  - Cost
  - Staff qualifications
  - Staff resources
  - Time
  - Training
  - Current assessments in use
  - Format

Step 4. Staff Training
- ❑ Pretrial supervisors
- ❑ Case managers
- ❑ Other staff
- ❑ Plan for Training for Trainers

Step 5. Pilot the assessment
- ❑ Design pilot
  - Identify a start and end date
  - Identify the target population
  - Identify the procedures to be used
- ❑ Pilot assessment
  - Collect data on IPV aggressors
  - Collect data on assessment results
  - Track how many assessments were completed
  - Assess the length of time for assessment
  - Get feedback from stakeholders, staff, and clients, on the assessment process
- ❑ Data analysis (see step 6)
- ❑ Decide whether to
  - Implement the assessment as planned
  - Revise implementation plan
  - Calibrate the assessment to reflect local context
  - Try something different

Step 6. Analyze Data
- ❑ Assess distribution of risk scores
- ❑ Adjust cut-off scores as needed
- ❑ Consider stakeholder, staff, and client feedback

Step 7. Implement the Assessment
- ❑ Develop formal policy
- ❑ Policy may include:
  - Timing of assessment
  - Identify positions responsible for conducting assessments
  - Policies for overrides
  - Distribution of assessment results
  - Storage of assessment results
  - Procedures for reassessment
  - Quality assurance

Adapted from Koetzle, Mellow, Piñol, & Pugliese (2021). Practical Guide to Youth Risk and Need Assessments in Latin America and the Caribbean.

TECHNICAL APPENDIX

METHODS

Search Strategy

Multiple databases and sources were searched between December 2019 and April 2020 to identify published and unpublished validation studies. Databases included PyschInfo, ScienceDirect, Web of Science, ProQuest Dissertations & Theses, Sociological Abstracts, Social Service Abstracts, Criminal Justice Abstracts, National Criminal Justice Reference, Sage Criminal Justice Journals, CINAHL Complete, PAIS International, and PsycTests. In an effort to identify unpublished studies, we also searched National Institute of Justice, OpenGrey, and Violence Against Women websites. Finally, reference lists of previously conducted meta-analyses were searched. Search terms included ("domestic violence" or "intimate partner violence") and ("risk assessment" or "danger assessment"). In addition, searches were conducted for specific risk assessments as indicated in Appendix A of this report.

Inclusion Criteria

All of the studies collected through the search process were screened for eligibility. To be eligible, studies had to report on the validation of a risk assessment using an adult population (mean sample age 18+) with a sample or a subsample that was charged with IPV-related offenses and published in English or Spanish. Eligible studies reported on offender-focused risk assessment, reported on recidivism as an outcome, and reported sufficient data to calculate an effect size for the IPV/DV sample. When multiple studies reported on the same sample, the study with a longer follow-up period was selected. Studies reporting reliability statistics, but not validity, were excluded from the current analysis. See Figure 1 for the PRISMA flow chart.

Coding

Each study was coded along four categories. These included:

1. Study characteristics (e.g., publication type, publication year, author affiliation, jurisdiction size, instrument name, original versus modified instrument)

2. Sample characteristics (e.g., sample size, mean age, percent male, race and ethnicity distributions, relationship type, pretrial status, criminal justice setting, assessment setting, base recidivism)

3. Methodological characteristics (research design, data collection process, researcher-conducted assessment, fidelity, construction sample, statistical power, multiple metrics reported, blind review, reliability, recidivism jurisdiction type)

4. Outcome characteristics (type of recidivism, recidivism charge, charge level, length of follow-up)

Each study was double-coded by the first and third author of this report. Any disagreements were discussed, and a consensus was reached. The complete code sheet is available from the corresponding author.

Effect Sizes

Predictive validity statistics were coded for each study. Where multiple effects were reported, we extracted all of them and controlled for the dependence of effect sizes by using multi-level analysis (see data analysis). While this approach allows for a more precise estimation of the effect (Cheung, 2014), it limited our choices of the effect size to report on. For the current analysis, effect sizes were coded using Pearson's r. Where necessary, a calculator was used to convert reported outcome data or statistics to r (Rosenthal, 1994). Effects were then standardized using Fisher's z transformation. Following the estimation of the mean effect, we transformed the results to the Area Under the Curve (AUC; Rosenthal, 1994) to aid interpretation.

DATA ANALYSIS

Rather than a traditional two-level approach to meta-analysis, a three-level model was used, as this approach applies the logic of multilevel models to meta-analysis and allows for multiple effect sizes to be extracted from primary studies (Cheung, 2014; Konstantopoulos, 2011). Most studies reported on more than one effect size, and it can be expected that effect sizes extracted from the same study are more similar than effect sizes extracted from different studies. While conventional random-effects models take into account the within-effect-size variance (level 1), and the between-studies variance (level 3), three-level meta-analysis adds an additional level: the variance within study across different effect sizes (level 2). In other words, a three level meta-analytic approach was most appropriate as it can take into account effect size dependency, and power and accuracy are not artificially inflated (Hox et al., 2017).

Random effects approach

As each of the primary studies varied across design, samples, and other characteristics, a random-effects approach was deemed most appropriate for the analyses. Random-effects meta-analytic models regard studies to be a random sample out of a population of studies (Hedges & Olkin, 1985). Thus, the observed variation in effect sizes can be attributed to both sampling error as well as systematic variability across studies.

Following the syntax guide by Assink and Wibbelink (2016), the "rma.mv" function of the metafor package (Viechtbauer, 2010) was used in the statistical software environment R (version 4.0.2). First, the overall mean effect sizes were estimated. Then, in order to assess between-study and within-study heterogeneity, independent log-likelihood-ratio tests were performed in which the deviance of the full model was compared to the deviance with either the level 2 variance parameter or the level 3 variance parameter dropped from the model.

Outliers

One effect size was identified as an outlier with a standardized R value of -0.05. In order to examine the effect of this outlier, the three-level model was reproduced excluding the case and compared to the full overall model. As there were no discernible differences in the results, the outlier was retained in the final analyses.

Moderator analysis

When there was evidence for heterogeneity in effect sizes, moderator analyses were conducted. The three-level meta-analytic models were obtained for several subgroups based on the various study characteristics, sample characteristics, and methodological characteristics of the primary studies.

Publication bias

Despite best efforts to find as many relevant primary studies as possible, there are several forms of bias that may impact the results of meta-analyses such as coding bias or selection bias. For instance, publication bias refers to the tendency for studies reporting significant and large positive effects to be more likely to be submitted or accepted for publication compared to those reporting small, nonsignificant, or negative effects (Rosenthal, 1979).

In order to determine whether a form of bias was present in the effect sizes examined in the current meta-analysis, a funnel plot based trim-and-fill analysis was conducted (Duval & Tweedie, 2000) using the trimfill function of the metafor package (Viechtbauer, 2010) in R. A funnel plot charts the standardized effect sizes against the precision of the effect size (1 divided by the standard error). Effect sizes from studies with smaller sample sizes are likely to be scattered widely at the bottom of the plot, and effect sizes from larger sample studies would be concentrated at the top of the plot, thus forming the shape of a funnel. This method is predicated on the assumption that when unaffected by bias, effect sizes are symmetrically distributed in the shape of a funnel around the "true" effect size. If the plot is asymmetric, it is remedied by imputing effect sizes that are estimated on the basis of existing effect sizes in the dataset. An adjusted overall effect can be estimated using a version of the dataset that includes the imputed effect sizes that were produced by the trim-and-fill algorithm.

RESULTS

Over 2,000 studies were identified through the search process, representing 260 unique records. After screening, we assessed 212 records for inclusion. More than half were excluded. The most common reason for exclusion was that the sample was not comprised of IPV aggressors, or did not report effects specifically for IPV aggressors. Twenty-three studies were excluded because they were evaluating victim-focused assessments, which was outside the scope of the current study. A number of studies could not be located or did not include sufficient data for extracting effect sizes. After contacting authors multiple times, we excluded 9 studies for these reasons. Following the screening and exclusion process, a total of 48 primary studies were included in the current analysis, with one study reporting on two distinct samples.

Overall effect size

The estimated summary effect across all studies (k = 49) and effect sizes (#ES = 246) was of z = 0.233 (SE = 0.015), t(245) = 15.694, p < 0.001, which is equivalent to an AUC value of 0.633. This indicates a 63% chance that a randomly selected recidivism will have a higher risk score than a randomly selected non-recidivism. An AUC = 0.633 is considered a small to medium effect (Rice & Harris, 2005) suggesting that the use of risk assessment is better than chance alone.

This overall effect should be interpreted with caution however, as the one-sided likelihood-ratio tests showed significant variance both on the second level $\chi^2(1) = 158.474$ $p < 0.0001$ and the third level χ^2 (1) 64.113 $p < 0.0001$. Of the total variance, 28.7% could be attributed to within-study differences in effect sizes (level 2), and 58.7% could be attributed to between-study differences.

Analysis of potential bias

A visual inspection of the funnel plot (see Figure 2) indicated asymmetry in the distribution of effect sizes, suggesting that there may be bias present in the dataset. Specifically, effect sizes were missing to the left of the estimated mean effect, meaning that small and negative effect sizes were missing in the dataset. Based on the trim-and-fill analysis (see Table 3), 33 effect sizes were imputed to the left of the estimated mean effect and added to the dataset to restore the symmetry of the funnel, and to estimate a "corrected" overall effect. Re-estimating the overall effect produced an effect size of $z = 0.146$ (SE = 0.018), $t(277) = 8.362$, $p < 0.001$, which equals an AUC value of 0.583. The difference between the two effects ($0.233 - 0.146 = 0.087$) is modest, indicating that the results were not seriously influenced by publication bias. Subsequently, the imputed values were dropped from the remaining analyses.

Moderator analyses

Given the significant level of variation between studies, we explored a number of potential moderators. Each potential moderator of interest was examined in a bivariate model. Potential moderators were classified into four categories: study characteristics, sample characteristics, criminal justice characteristics, and methodological characteristics. Detailed results are available from the authors.

Table 3. Overall Effects Before and After Trim-and Fill Analyses

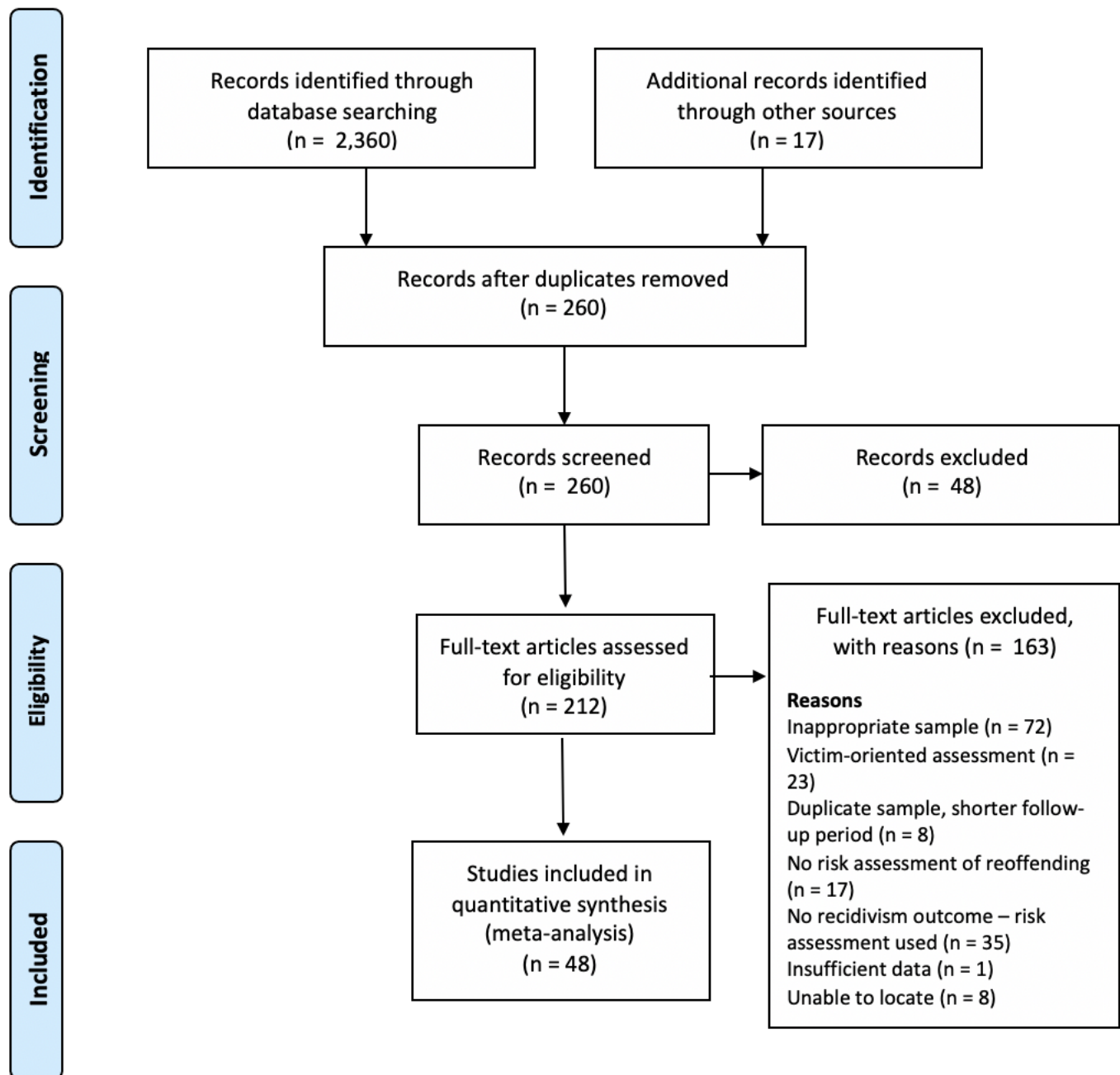|	Mean ES (SE)	95%CI	% var. at level 1	Level 2 variance	% var. at level 2	Level 3 variance	% var. at level 3
Overall effect before trim-	0.233*** (0.015)	0.204, 0.263	12.6	0.004***	28.7	0.008***	58.7
Overall effect after trim-and-	0.146*** (0.018)	0.112, 0.181	7.4	0.004***	15.3	0.018***	77.2

*<.05, **<.01, ***<.001

Figure 1. PRISMA Flow Chart

Figure 2. Funnel Plot