This bulletin has been independently peer reviewed.

CRIME AND JUSTICE BULLETIN

NUMBER 244 | FEBRUARY 2022

Improving police risk assessment of intimate partner violence

Felix Leung and Lily Trimboli

AIM

To develop a tool with better predictive accuracy than the NSW's Domestic Violence Safety Assessment Tool.

METHOD

The sample comprised 234,454 incidents between intimate partners recorded from January 2016 to December 2018 in the Central Referral Point (CRP) database, provided by Victim Services NSW. Four predictive models of a victim's risk of intimate partner re-victimisation were evaluated. These were a) NRAP-all: a model with all 16 risk factors identified in the National Risk Assessment Principles (NRAP) developed by Australia's National Research Organisation for Women's Safety (ANROWS); b) NRAP-10: a model with only the 10 high-risk factors identified by ANROWS in the NRAP; c) SAFVR: the Static Assessment of Family Violence Recidivism (SAFVR), developed by the New Zealand Police; and d) DVSAT-8: a model with the eight items from the NSW DVSAT that are most predictive of repeat victimisation. The best performing model was further simplified, keeping only the best predictors. This simplified model was then used to predict re-victimisation.

RESULTS

Of the four models, only the two NRAP models had acceptable predictive performance (i.e., with an Area under the Curve (AUC) above 0.7). The model with the highest AUC was the NRAP model that combined all 16 risk factors identified by ANROWS (with an average AUC of 0.732). When tested on unseen data, its out-of-sample AUC was similar, at 0.738. The simpler NRAP-10 model performed only slightly worse with an average AUC of 0.728.

Further analysis found that the best performing NRAP model could be simplified, with almost no loss in predictive performance. Of the 16 risk factors in the NRAP model, a simplified model using only the best five predictors delivered an out-of-sample AUC of 0.734.

CONCLUSION

This study demonstrates that a risk assessment instrument with a small number of variables can identify victims who are most at risk of future intimate partner violence. With an appropriately selected risk threshold to match service capacity, the simplified model could help control the volume of referrals at a desired level.

KEYWORDS

intimate partner violence | re-victimisation | risk assessment | predictive accuracy

Suggested citation: Leung, F. & Trimboli, L. (2022). *Improving police risk assessment of intimate partner violence* (Crime and Justice Bulletin No. 244). Sydney: NSW Bureau of Crime Statistics and Research.

INTRODUCTION

Domestic and family violence (DFV) are significant issues, both nationally and internationally. In NSW, domestic and family violence are broadly defined as "any behaviour, in an intimate or family relationship, which is violent, threatening, coercive or controlling, causing a person to live in fear" (NSW Government, 2014a, p.28). It can include physical, sexual, emotional and financial abuse. According to the 2016 Personal Safety Survey, it was estimated that 17 per cent of women (1.6 million) and 6.1 per cent of men (547,600) have experienced physical and/or sexual violence[1] by a partner[2] since the age of 15 years. In addition, 23 per cent of women (2.2 million) and 16 per cent of men (1.4 million) report having experienced emotional abuse[3] by a current and/or previous partner since the age of 15 years (Australian Bureau of Statistics [ABS], 2017). The lifetime prevalence of financial abuse is estimated to be 15.7% among women and 7.1% among men (Kutin, Russell, & Reid, 2017). In NSW, 31,187 DFV related assaults were recorded by police in 2019 (NSW Bureau of Crime Statistics and Research [BOCSAR], 2019). This was one of only two major crime categories that recorded a significant upward trend in the 24 months to December 2019.[4] In this two-year period, reports of DFV related assault increased among female and male victims, adult and juvenile victims, and intimate partner and family violence victims.

There are serious adverse physical, psychological, social and economic costs of DFV both for the individual and the wider community (e.g. Dillon, Hussain, Loxton, & Rahman, 2013; Howard, Trevillion, & Agnew-Davies, 2010; Lagdon, Armour, & Stringer, 2014; Lum On, Ayre, Webster, & Moon, 2016; Stylianou, 2018; World Health Organisation, 2013). In 2015/2016 violence against women and their children in Australia was estimated to cost $22 billion (KPMG, 2016). The cost of physical and sexual violence was estimated at $12 billion, and emotional abuse and stalking was estimated at $10 billion (in 2009, this was estimated to be $5 billion). Pain, suffering and premature mortality accounted for the largest proportion of the estimated cost ($10.4 billion) and reflects the long-term impact of violence on health, including mental health, chronic illness and pain, reproductive health, and smoking and substance abuse.

A major challenge facing governments in managing DFV is allocating scarce resources in responding to incidents of DFV.[5] A key consideration for the government in doing this is victim safety and their risk of further harm. Not all victims of DFV face similar levels of risk; some victims are at much greater threat of serious harm and even death. Providing the most support to these victims who are at the greatest risk is critical to reduce the extent and severity of the harm caused by DFV.

Risk assessment tools have emerged as a key tool to assist authorities in responding to incidents of DFV. These tools typically incorporate a range of risk factors (such as perpetrator and victim characteristics, criminal history and details of most recent incidents) to produce a single predictive metric of the risk of subsequent harm. This metric can then be used as the basis for prioritising victims for intensive support using a set of rules or thresholds to classify cases into different categories of risk.[6]

Accuracy, the ability of a tool to correctly discriminate between those at more and less risk of DFV, is paramount for DFV risk assessment tools. A tool with poor accuracy would mean that some victims at high risk could be insufficiently supported, compromising their safety. It could also cause the misallocation of already-scarce DFV support resources where they are less needed. Accuracy has commonly been assessed using measures such as the Area Under the Receiver Operating Characteristic Curve (AUC),

1 The 2016 *Personal Safety Survey* (Australian eau of Statistics [ABS], 2017, Glossary) defines 'violence' as 'any incident involving the occurrence, attempt or threat of either sexual or physical assault'.
2 'Partner' in the 2016 *Personal Safety Survey* (ABS, 2017, Glossary) describes 'a person the respondent lives with, or lived with at some point in a married or de facto relationship. This may also be described as a co-habiting partner.'
3 In the 2016 *Personal Safety Survey* (ABS, 2017, Glossary), 'emotional abuse' was defined as occurring 'when a person is subjected to certain behaviours or actions that are aimed at preventing or controlling their behaviour causing them emotional harm or fear. Interested readers are directed to the ABS (2017) for more information on the specific behaviours included in this category of abuse.
4 The other crime category with a significant upward trend was 'steal from retail store' (NSW Bureau of Crime Statistics and Research, 2019).
5 While the NSW government has committed a large sum of money to DFV, it splits this over several classes of activities such as primary prevention, perpetrator rehabilitation and accountability and support to victims. Thus, only a fraction of total funding is spent on victim support and responses to incidents of DFV (NSW Government, 2016).
6 An alternative method for risk assessment is for DFV responders to determine risk based on their own professional judgement. However, many studies have demonstrated the superior ability of prediction tools and algorithms to professional judgments, for example Grove et al. (2000).

where a tool with AUC values of 0.7 is considered acceptable, 0.8 or more good, and 0.9 or more excellent (Hosmer & Lemeshow, 2004).

In pursuing greater accuracy, developers of DFV risk assessment tools face tension between parsimony and comprehensiveness. Including multiple factors which strongly correlate with further victimisation or perpetration will likely improve predictive ability compared to a tool with fewer such factors. However, variables irrelevant or redundant for predicting repeat victimisation introduce noise into the predictions made by such tools. This means that well-meaning attempts to incorporate as many plausible factors affecting DFV into these tools may reduce accuracy. Tools with more predictors also increase the burden of assessment and data collection on assessors such as police. Thus, to maximise the effectiveness of both prediction and implementation, DFV tools should include as many strong predictors of future victimisation as possible, while excluding unnecessary factors.

The NSW Domestic Violence Safety Assessment Tool (DVSAT)

In NSW, the Domestic Violence Safety Assessment Tool (DVSAT) has been used state-wide by the NSW Police Force since 1 July 2015 to determine whether the life, health or safety of a victim are classified 'at threat' or 'at serious threat' due to DFV (NSW Government, 2014b). It is designed to identify victims at greatest risk early on, so that they can be given priority and offered timely and appropriate support. It is designed for both female and male victims aged 16 years or more and for both victims in intimate partner relationships and victims in non-intimate relationships (for example, siblings, parents and adult children). It is a fundamental element of the NSW Government's *Safer Pathway* program which is a coordinated approach to responses for victims at threat of DFV, prioritising those at serious threat.[7] *Safer Pathway* and the DVSAT are currently operational in 48 sites across NSW (Women NSW, 2020).

The DVSAT has two parts, Part A and Part B, and consists of a total of 30 questions (NSW Government, 2014c). These 30 questions were selected following a review of existing risk assessment tools, as well as Australian and international research on DFV related homicide (NSW Government, 2014b). Part A is a risk identification checklist for victims of IPV in intimate partner relationships and contains 25 questions asking about the perpetrator's background and criminal history, the victim-perpetrator relationship and the involvement of children in the relationship. Part B of the DVSAT is completed for both victims in intimate and non-intimate relationships. It is based on the professional judgement[8] of the police officer and consists of five questions – the level of fear felt by the victim of their partner (from 'not afraid' to 'terrified'); the concerns expressed by the victim about what the partner might do and to whom; whether there are any children in the household at risk of harm and any action taken by the officer; and whether there are any additional factors or circumstances that lead the officer to believe that there is a threat or serious threat to the safety of the victim and/or the children.[9]

The DVSAT uses these responses to classify victims into one of two risk categories: 'at threat' or 'at serious threat'. A victim is assessed as 'at threat' if they answer 'yes' to any of the 25 questions in Part A, and 'at serious threat' if they answer 'yes' to 12[10] or more of the questions in Part A.[11] The officer can use their professional judgement (for victims in intimate and non-intimate relationships) to assess a person as 'at threat' or 'at serious threat' based on their responses in Part B.[12] This classification determines whether a victim receives more intensive support. Specifically, if a victim in a *Safer Pathway* Police Area Command (PAC) or Police District (PD) is assessed to be 'at serious threat' for future harm, they are referred to a

7 Safer Pathway is 'an integrated and coordinated government and non-government response for victims identified at threat of domestic and family violence, with a particular focus on victims at serious threat... victims are placed at the centre of the system and are provided with more effective and efficient responses, more integrated service support, to help them recover from domestic violence and to prevent the escalation of abuse' (NSW Department of Justice, 2014, p. 14).

8 Professional judgement is 'an assessment based on information gathered with a victim, and the knowledge, skills and experience of the service provider completing the DVSAT' (NSW Government, 2014b, p. 4).

9 In Part B, the officer is required to consider any factors that may increase the victim's vulnerability and threat to her/his safety. These include substance misuse, mental health issues, disability, social or geographical isolation, cultural or language barriers, immigration issues, unwillingness to engage with support services, the perpetrator's access to weapons, or involvement with the NSW Department of Communities and Justice.

10 The original threshold was answering 'yes' to 10 items or more. This threshold was increased to 12 items or more on 9 March 2015.

11 For victims in intimate partner relationships

12 This is also used to classify victims of violence in non-intimate relationships.

Safety Action Meeting (SAM).[13] These are resource-intensive meetings involving several agencies where a plan is made to refer victims to appropriate support services. Female victims assessed 'at serious threat' who do not reside in a *Safer Pathway* site are referred to the local Women's Domestic Violence Court Advocacy Service (WDVCAS) where staff members provide case co-ordination. All male victims are referred to Victims Services NSW which provides case co-ordination and referral to local support services, if available.

Despite the importance of the DVSAT in allocating support to DFV victims, research indicates that it is a poor predictor of repeat intimate partner violence (IPV). Ringland (2018) evaluated the predictive ability of the DVSAT, examining data of 24,462 victims of IPV assessed on the DVSAT. There were several important findings of this study. First, responding 'yes' to 12 or more DVSAT Part A items did not predict repeat victimisation well, with an AUC of approximately 0.5. In other words, the DVSAT was found to be no better than chance at predicting re-victimisation. While the AUC increased for the 'at serious threat' risk classification (i.e., when including police professional judgment), it still did not reach an acceptable discrimination level of 0.7. Second, that some items on the DVSAT were found to be unrelated or even inversely associated with repeat victimisation. The inclusion of these factors in the risk assessment contributed to its low predictive accuracy.

Two process evaluations of the DVSAT have also suggested that the tool could be improved. The first study was based on interviews with 54 key stakeholders about aspects of the *Safer Pathway* program, including the DVSAT. Many stakeholders believed that there was scope to refine the DVSAT to better identify victims who are at serious threat of future injury or harm (BOCSAR, 2015; also see Wan, Thorburn, Poynton, & Trimboli, 2018). Stakeholders suggested two improvements. The first was to increase the weight DVSAT places on questions relating to more severe or serious events. Currently, each question is equally weighted. The second was to provide victims with a timeframe to which the questions refer, perhaps with a focus on the 'last 12 months'. Currently, the questions are not time-limited. Similar conclusions were reached by ARTD Consultants, who interviewed eight police officers while evaluating the *Safer Pathway* program (NSW Government, 2019).

The Static Assessment of Family Violence Recidivism (SAFVR) and the National Risk Assessment Principles

There are two risk assessment approaches that have so far been untested but could inform improvements to the DVSAT: a) the Static Assessment of Family Violence Recidivism (SAFVR),[14] from New Zealand, and; b) the ANROWS National Risk Assessment Principles (NRAP). The Static Assessment of Family Violence Recidivism (SAFVR), was developed by New Zealand Police (New Zealand Police, no date). It uses information from a number of police databases to predict the probability that a person will commit a family violence offence within the next two years. It considers characteristics of the offender, such as, gender and age at time of index offence/incident; past and current offending, such as, family violence index offence/incident codes, count of family violence occurrences in the past year, history of family harm, a count of previous offences (including overseas convictions), Ministry of Justice conviction record and sentences of more than 30 days. It is used as the first step in the New Zealand Police's 'Action Quality Family Harm Investigation' to determine further assessment and action.

Australia's National Research Organisation for Women's Safety (ANROWS) developed a set of National Risk Assessment Principles (NRAP) for DFV (Toivonen & Backhouse, 2018) as part of the Third Action Plan under the National Plan to Reduce Violence against Women and their Children 2010 – 2022 (Commonwealth Department of Social Services, 2016). The Principles are intended to provide an

13 SAMs are multi-agency meetings which are designed to provide a forum for effective information-sharing between senior representatives of relevant government and local non-government service providers working with domestic violence clients and perpetrators (NSW Government, 2017). One of their key functions is to develop targeted and time-specific safety plans to prevent or lessen a serious threat to the life, health or safety of domestic and family violence victims and their children.

14 Jolliffe Simpson, Joshi, & Polaschek (2021) tested the predictive ability of the SAFVR but within a considerably shorter follow-up period for measuring re-victimisation. The AUC was 0.64 for a 24-week follow-up window.

overarching national understanding of risk and managing risk in domestic and family violence. These Principles do not constitute a risk assessment tool. In fact, ANROWS notes that they do not replace existing state and territory frameworks or tools, but rather they "provide a guide for jurisdictions in developing, revising or evaluating risk assessment frameworks, tools and resources for various cohorts (adult survivors, perpetrators, children, other family members)" (Toivonen & Backhouse, 2018, p. 1).

The Principles were developed following an examination of Australian and international research, government reports and inquiries, and DFV death review reports. The literature examined by ANROWS dealt with the concepts of risk, safety, need, vulnerability, risk assessment and management approaches, intimate partner sexual violence and sexual assault, and multi-agency integrated service responses. The Principles also received substantial input from practitioners, policy-makers, government agencies, peak bodies and researchers. In addition, ANROWS considered the research and data available for people from diverse groups in the population and also from diverse life stages, such as Aboriginal and Torres Strait Islander women and families; migrants, refugees and people who are culturally and linguistically diverse; people with disabilities; lesbian, gay, bisexual, transgender, queer and intersex people; people with a mental illness; older women; women in pregnancy and early motherhood; people in regional, rural and remote areas and young women.

The ten lethality/high-risk factors identified by ANROWS as a guide for practitioners are:

1. a history of family and domestic violence;
2. separation (actual or pending);
3. intimate partner sexual violence;
4. non-lethal strangulation or choking;
5. stalking, including GPS tracking, persistent texting or telephoning;
6. perpetrators' threats to kill their partner, former partner, themselves or others, including children;
7. perpetrators' access to, or use of, weapons;
8. the escalation in frequency and/or severity of violence over time;
9. coercive or controlling behaviours, including verbal and financial abuse, psychologically controlling acts and social isolation; and
10. pregnancy and new birth.

In addition, ANROWS identified six other risk factors:

1. a victim's perception of their own risk of experiencing future violence;
2. a perpetrator's suicide threats or attempts;
3. court orders, such as intervention orders, and parenting proceedings;
4. alcohol and/or drug misuse and abuse;
5. isolation, including limiting interactions with family, friends, social supports and community support programs, and barriers to help-seeking; and
6. the abuse, cruelty or harm directed at pets and other animals.

The SAFVR and the risk factors identified by the NRAP each have particular features which could be promising in predicting re-victimisation. The SAFVR uses perpetrator characteristics from existing police databases to model risk, as opposed to information gathered from interviews with victims. An immediate advantage of this approach is that it would require less effort and resources to be collected and could potentially speed up response as police could have access to its risk classifications when responding to an incident. Meanwhile the NRAP risk factors were identified based on an extensive process of reviewing

prior research evidence and confirming these with experts. If these are empirically relevant for the NSW population of DFV victims, it could be used to predict re-victimisation with relatively high predictive accuracy.

Predictive accuracy of other DFV risk assessment tools

Before proceeding it is worth reviewing the findings of other evaluations of DFV risk assessment instruments. The most recent systematic review of DFV risk assessment tools finds that these instruments' ability to discriminate between recidivists and non-recidivists tend to be 'moderate' or 'acceptable', with a typical AUC in the range between 0.7 and 0.8 (Graham et al., 2021). This review, and others before it (Hanson, Helmus, & Bourgon, 2007; Messing & Thaller, 2013; Nicholls, Pritchard, Reeves, & Hilterman, 2013; Svalin & Levander, 2019; Wheller & Wire, 2014) have also noted that many tools may not be generalisable to different sub-groups of the population or to different settings (e.g. Singh, Grann, & Fazel, 2011). This is particularly relevant to jurisdictions such as NSW, where the definition of DFV is broader than male perpetrators and female victims in heterosexual relationships. Despite the focus on the AUC in the findings that follow , an acceptable AUC score should be considered a minimum requirement. Importantly, an acceptable AUC score does not imply that other metrics, such as the sensitivity and specificity, of the tools are also acceptable in practice.

Of the various tools that have been subject to empirical review, the Ontario Domestic Assault Risk Assessment (ODARA) has been shown to have strong predictive validity. Developed in Canada, the ODARA was designed specifically for police officers responding to incidents of IPV. It was intended to be a quick, brief assessment that is easy-to-use, easy to interpret and can be scored by officers immediately after completing a IPV investigation rather than during the situation itself. Much like the SAFVR, no victim interviews are undertaken. The tool was constructed by retrospectively coding potential risk factors from 589 cases of IPV recorded in the Ontario police database prior to 31 December 1996. The ODARA's AUC for the construction sample was .77 which is considered 'acceptable' and indicates a probability of .77 that a randomly selected recidivist would have a higher score than a randomly selected non-recidivist. In an initial cross-validation with a sample of 100 cases that were not used in the construction sample, the ODARA retained an acceptable AUC of .72 (Hilton et al., 2004). Since then, it has been validated on a number of sub-samples,[15] most notably an Australian sample. To do this, Lauria and colleagues (2017) used a sample of 200 family violence cases which involved male-to-female IPV with a history of assault and co-habitation occurring in Victoria over a five-month period. They found that ODARA accurately discriminated between perpetrators who re-offended with a further physical assault (i.e. attempted murder, homicide, abduction, assault, threat of assault and sexual assault) and those who did not (AUC = .68) and, similarly, between those who re-offended with non-physical abuse and those who did not (AUC = .72). Lauria et al. found that two items, prior non-domestic violence and prior violence toward another victim, were particularly strong predictors of further physical assault. However, the ODARA has limited practical use in the Australian context because of its restrictive inclusion criteria of male-perpetrated violence against a current or former co-habiting female partner. For example, in the study by Lauria et al., only 23.4 per cent of the 854 family violence cases that were reported to the police during the five-month data collection period met this eligibility criterion.

Studies of risk assessment tools employed in Australian jurisdictions have also found that these are only moderately predictive of future reports of domestic and family violence. Mason and Julian (2009) examined the Risk Assessment Screening Tool (RAST) which was used by Tasmania Police from September 2004 to assess victims' risk of experiencing future violence. The RAST contains 34 items related to the traits and behaviour of the offender, information on which is obtained from the victim by

15 These include: male offenders selected from police records (e.g. Hilton & Eke, 2016; Hilton & Harris, 2009; Ulmer, 2015) and correctional files (e.g. Gray, 2012; Hilton, Harris, Popham, & Lang, 2010), female offenders with a record of intimate partner violence against male partners following their release from a correctional treatment institution (e.g. Hilton, Popham, Lange, & Harris, 2014), and in sub-samples of offenders (e.g. men who killed their female intimate partners, Eke et al., 2011; and men who committed sexually motivated violent offences against their current or former intimate partners, Rettenberger & Eher, 2013).

the responding police officer.[16] Mason and Julian drew a random sample of 1,406 domestic and family violence offenders from the Tasmania Police electronic database. Each offender had recorded at least one episode of family violence between January 2000 and July 2008. The researchers then examined the extent to which the RAST predicted further acts of violence against the victims. Of the total sample, 767 (54.5%) offenders had only one family violence incident during the study period while the remaining 639 (45.5%) had more than one incident. The analysis of the overall RAST score yielded an AUC of .602. However, if only the risk factors that were identified as significantly related to re-offending were included in the regression analyses, an improved AUC score of .726 was achieved. Significant factors included breaching protective orders, the offender's jealous or obsessive behaviour and whether they had made previous threats toward the victim. Two studies have examined Victorian approaches to DV risk assessment. McEwan, Bateson, and Strand (2017) compared the predictive accuracy of the Victoria Police Family Violence Risk Assessment and Management Report (or the VP Form L17), and two other approaches that had been developed for the jurisdiction. These were the Brief Spousal Assault Form for the Evaluation of Risk, B-SAFER (Kropp & Hart, 2004),[17] and a newly developed Priority Assessment Protocol (PAP). They found that both Form L17 and the PAP were unable to adequately discriminate between victims who experienced subsequent domestic and family violence incidents from victims who did not (AUC = .56 and .54, respectively). The B-SAFER case prioritisation tool achieved a slightly higher AUC of .63 but still demonstrated a low level of accuracy. This research led to the development of a new risk assessment tool, the Victoria Police Assessment for Family Violence Risk (VP-SAFvR), which was empirically developed using a randomly selected sample of 24,446 cases of domestic and family violence incidents recorded by Victoria Police in the 2013/2014 financial year and was cross-validated with two samples from the remaining 19,996 domestic and family violence incidents recorded in that year. Each case was followed for a period of 12 months to identify any subsequent domestic and family violence incidents involving the same two people as the index incident. McEwan, Shea, and Ogloff (2019) found that the VP-SAFvR achieved a moderate level of discrimination (AUC = .66) in predicting future reports of family violence incidents to Victoria Police. The new instrument correctly identified about three-quarters of cases with further reports of family violence in the 12-month follow-up period as high risk. The VP-SAFvR also correctly identified half of those cases that did not have further reports as low risk. The level of predictive validity of the VP-SAFvR was consistent across different types of relationships (i.e., intimate partner cases, non-intimate partner cases, cases involving only male perpetrators and cases involving only female perpetrators).[18] Finally, Dowling and Morgan (2019) recently examined the accuracy of the Family Violence Risk Assessment Tool (FVRAT) used in the Australian Capital Territory. The assessment comprises 37 items which are combined to produce an overall score that reflects the likelihood of repeat domestic violence, and is used to classify cases into low-, medium- or high-risk of repeat domestic violence.[19] Dowling and Morgan used a sample of 350 incidents of IPV reported between March and December

16 Of these factors, 18 are Group A or high-risk factors that may indicate that a sharp escalation in risk is imminent and the remaining 16 are Group B or 'other' risk factors for potential escalation (Winter, 2006). The 18 Group A risk factors are: the victim was assaulted in the past; the victim was assaulted in this incident; the level of violence escalated; bizarre, paranoid or delusional behaviour; separation after cohabitation; the victim is terrified; the offender breached restraint or Family Violence Order; suicide attempt; pregnancy or new birth; threats to kill the victim or children; the offender stalked the victim/others; possession or access to firearms; child was assaulted in the past; cultural or disability issues; child was assaulted in this incident; the offender killed family pet; sexual assault or rape arrest; and previous murder/manslaughter arrest. The 16 Group B risk factors are: the offender being jealous, obsessive, intimidating or bitter; the offender being unemployed; violence against any other person in the past; the victim being afraid; alcohol problem; the offender or the relationship being in financial difficulty; past threats to kill; drug use problem; current or history of depression; suicidal fantasy or threats; the victim or children being physically injured; diagnosed mental illness/disorder; homicidal fantasy or threats; not taking prescribed medication; injury or threat to pets; and strangling or suffocation (Winter, 2006).

17 The B-SAFER uses structured professional judgement, 'an approach that attempts to bridge the gap between actuarial and unstructured clinical approaches to risk assessment … [it provides] guidance in terms of which risk factors to consider, as well as operational definitions for the scoring of the factors … the final step of combining risk factors … is not done algorithmically' (Kropp & Hart, 2004, pp 4, 5). The B-SAFER is a checklist divided into two sections. The first section, entitled 'spousal violence', consists of five factors related to the perpetrator's history of intimate partner violence (actual or attempted physical and sexual assault; violent threats or ideation; increase in the frequency or severity of violence or of threats/ideation; violation of court orders for bail, probation, parole, restraining orders; and beliefs and values that encourage or excuse abusive, controlling and violent behaviour). The second section, entitled 'psychosocial adjustment', consists of five risk factors reflecting psychological and social functioning (other antisocial behaviour, intimate relationship problems, employment problems, substance use problems and mental health problems). Each factor is scored as being present or absent in two time periods – 'currently' (in the past four weeks) and 'in the past' (prior to the past four weeks). Information is collected from interviewing the offender, the victim and a review of police records (Kropp & Hart, 2004).

18 Intimate partner relationships: AUC = .65 (n = 26,246); other family relationships: AUC = .65 (n = 10,747); child-to-parent abuse: AUC = .63 (n = 6,827); male perpetrators: AUC = .66 (n = 33,987); female perpetrators: AUC = .65 (n = 10,258). However, for same-sex relationships, the effect size was small (AUC = .57), but the sample size was also small (n = 262) (McEwan et al., 2019, Table 5, p. 599).

19 While police officers can revise the risk ratings based on their professional judgement, Dowling and Morgan (2019) note that this did not occur during this wave of data collection.

2017 to assess whether the FVRAT's ratings accurately predicted repeat IPV involving the same victims and offenders within six months of the initial report. After controlling for police responses, Dowling and Morgan found that the FVRAT score was not a strong predictor of repeat IPV (AUC = .60). Dowling and Morgan found that 10 FVRAT items were positively associated with repeat IPV, and that restricting the tool to these 10 items and classifying cases into low-, medium- and high-risk groups produced an acceptable level of predictive validity (AUC = .73).[20]

In addition to the findings of Mason and Julian (2009) and Dowling and Morgan (2019), a study by Rahman (2018) indicates that repeat IPV can be reasonably predicted with only a small set of factors. In an initial attempt to improve the ability of police officers and others in NSW to identify those at high risk of repeat IPV, Rahman analysed a sub-set of 336 individuals who reported experiencing an incident of IPV by a current or former partner within the 12 months prior to data collection in the 2016 *Personal Safety Survey*. Rahman found five factors were significantly associated with the risk of being re-victimised with a physical or sexual assault within 12 months. These factors were experiencing emotional abuse in the form of a partner controlling or attempting to control contact with family, friends or the community; and various socio-demographic factors, namely living in an area within the most disadvantaged socio-economic quintile, living in an inner regional area, having low educational attainment and having a disability.[21] This model had an acceptable level of discrimination (AUC = .76) and correctly classified 69.3 per cent of cases. However, as Rahman acknowledged, the sample size used to estimate the model was very small and the study involved cross-sectional data, rather than repeated measures to measure further victimisation.

The current study

The research so far suggests that many Australian DFV risk assessment tools have only moderate levels of accuracy, including the NSW DVSAT. Meanwhile several Australian studies have demonstrated that it is possible to refine existing risk assessment tools (e.g. Ringland 2018, Dowling & Moran, 2019; Mason & Julian, 2009) and use a small set of strongly predictive factors (Rahman, 2018) to improve their accuracy.

The ultimate objective of the current study is to develop a better tool than the DVSAT that more accurately predicts future IPV re-victimisation in NSW. The refined tool should achieve an acceptable level of discrimination in predicting repeat victimisation (i.e., an AUC greater than .70). Ideally, the refined tool would be easy-to-use and quick to administer by NSW police officers, allowing them to rapidly assess the level of risk of domestic/family violence cases; be short, with a minimum number of items; and would help identify victims most at risk to be referred to the resource-intensive multi-agency Safety Action Meetings. To achieve these objectives, four of the risk assessment models described above (NRAP-all, NRAP-10, SAFVR, DVSAT) were examined for their predictive ability using recent data from the CRP database.

20 These items were: the offender has assaulted the victim during the most recent incident, the offender has injured or threatened to injure a family pet now or in the past, the victim was physically injured during the most recent incident, the victim and offender have recently separated after co-habiting, there is a pregnancy or new birth (less than 12 months) within the relationship, the offender has assaulted the victim in the past, the violence level (including property damage) in the relationship has escalated, the offender has committed offences of violence against any person in the past, the offender has breached a protection or any court order now or in the past, and the offender has money problems at the moment.

21 Of these factors, socio-economic disadvantage and emotional abuse by the partner were the two most strongly associated with repeat intimate partner violence assault within 12 months. The analyses showed that the odds of being a repeat victim of intimate partner violence within 12 months were 10 times higher for someone living in an area in the most disadvantaged socio-economic quintile relative to a person living in an area in the least disadvantaged socio-economic quintile. The odds of being re-victimised within 12 months were seven times higher if emotional abuse/controlling behaviour had been experienced by a victim in the last 12 months.

METHOD

Data source

We use a dataset of intimate partner incidents recorded between 1 January 2016 and 31 December 2019 in the CRP database provided by Victim Services, NSW Department of Communities and Justice. The CRP database contains all incidents of domestic and family violence attended by NSW Police, and includes both Safer Pathway and non-Safer Pathway referrals. As shown in Table 1, intimate partner incidents typically accounted for about 60 per cent of all recorded DFV incidents in recent years. In addition to the variables provided by Victim Services, we sourced the criminal history and other demographic variables of the persons of interest (POIs) from the NSW Bureau of Crime Statistics and Research's Reoffending Database (ROD). For incidents with multiple POIs, the POI with the most finalised court appearances as of the incident date was selected; if no records for any of the POIs were found in ROD, then the oldest POI as of the incident date was selected.[22] As Table 1 shows, the percentage of incidents involving intimate partners in which the POIs had records in ROD was similar in each year, between 66.9 and 67.1 per cent.

Table 1. Number of recorded incidents between 2016 and 2018

Year	All recorded DFV incidents	IPV incidents	POIs with records in ROD
2016	132,064	79,569 (60.3%)	53.380 (67.1%)
2017	129,044	75,946 (58.9%)	50.454 (66.4%)
2018	133,425	78,939 (59.2%)	52,828 (66.9%)

Note. Percentages are calculated with denominators in the column to the left.

The combined sample size across the three years is 234,454 incidents, with an average of about 78,151 incidents a year. We partitioned the records between 2016 and 2018 into two parts: incidents recorded in 2016 and 2017 as training data, and those recorded in 2018 as test data.[23] This train-test split allows for a fair evaluation of the models, with all model parameters estimated using the training data only and any predictive performance metrics measured using the test data. A temporal split also enables us to test whether the model's predictive ability persists over time.

Outcome variable

The key outcome variable, re-victimisation, is defined as any IPV recorded within 12 months of the index incident in the CRP database,[24] regardless of whether the POI in the subsequent incident is the same person in the index incident.[25] Classification of IPV is based on the recorded incident type. All incident types (e.g. common assault, ADVO breach, actual bodily harm) other than "No offence detected" and "Verbal argument" are classified as IPV.[26]

22 In the three years of incident data used in this report, only 2,825 (or 1.2%) out of the 234,454 incidents involved more than one POI. However, for these 1.2% of incidents we do not know the relationship between every POI and the victim, and can only infer that one of the POIs was an intimate partner of the victim. This is a limitation in the data provided to us. Incident data and POI data are in two separate tables, linked only through a unique incident number. However, the victim-perpetrator relationship is coded at the incident level only.

23 Records in 2019 provide the 12-month window needed to observe any repeat victimisations for incidents recorded in 2018.

24 We can distinguish between incidents that involved physical violence and those that did not, however we find that with these outcomes, the various models were generally less predictive (see Figure A1 in the Appendix).

25 About 88.3% of re-victimisations (i.e. 52,885 out of 59,884) involved the same perpetrator.

26 It may be argued that ADVO breaches are categorically different from other offences such as common assault that it should be treated separately, but given the inability to predict the seriousness of a repeat victimisation (see "Alternative outcome measures" in the Appendix) we chose to categorise ADVO breaches among other offences.

Predictors

Table 2 describes the variables we selected from the CRP and ROD. They correspond to predictors in the four models:

NRAP-all: the 10 high-risk factors and six other risk factors in the NRAP developed by ANROWS;

NRAP-10: only the 10 high-risk factors in the NRAP;

SAFVR: the assessment tool developed by New Zealand Police; and

DVSAT-8: Eight question items from the DVSAT. These were the items most predictive of repeat victimisation for females and males, as identified by Ringland (2018).

We proxied 14 of the 16 NRAP risk factors by question items from the DVSAT that best describe those risk factors, as summarised in Table 2. For example, we measure the high-risk factor "separation (actual or pending)" using question 11 in Part A of the DVSAT, which asks the victim if there has been a recent separation in the 12 months before the incident or if one is imminent. The other two risk factors have no equivalent in the DVSAT. The "intimate partner sexual violence" risk factor was instead measured by the number of previous IPV incidents recorded in the CRP database in which a sexual offence was reported. The "history of family and domestic violence" risk factor was measured by the number of prior DV offences by the POI in the five years before the index incident, derived from the ROD, and the number of all prior incidents recorded in the CRP database. The latter measure of the "history of family and domestic violence" was included as previous research by Ringland (2018) found it to be a good predictor of re-victimisation. We derived the eight SAFVR predictors based exclusively on the POI's criminal record from the ROD to match those used by New Zealand Police as closely as possible. POIs with no matching records in ROD were assumed to have no prior finalised court appearances. The DVSAT-8 model used question items from the DVSAT that Ringland (2018) identified to be most predictive of repeat victimisation.

The descriptive statistics of the outcome variable and all the predictors are presented in Table 3, separately for the training data and test data. As Table 3 shows, 25.1 per cent of victims identified in IPV incidents in 2016 and 2017 recorded another IPV incident within 12 months. The figure was similar for IPV incidents in 2018, at 26.5 per cent. The other descriptive statistics also appear to be broadly similar in both samples. The standard deviations are only shown for numeric variables, and not for any binary variables. All binary variables were 0-1 coded. Unless otherwise indicated, a '1' means 'Yes'. Therefore, for these variables, the mean represents the proportion in the training or test data with a 'Yes' response.

Table 2. Predictors of each model

Variables	NRAP (high-risk factors in bold)	SAFVR	DVSAT-8
Number of prior incidents recorded in CRP	**History of family and domestic violence**		
Number of prior intimate partner incidents in which sexual offence reported	**Intimate partner sexual violence**		
POI Age		✓	
POI gender (Female, Male, Unknown)		✓	
Number of prior finalised court appearances with proven offences		✓	
Number of prior finalised court appearances with proven DV offences in previous year		✓	
Number of prior finalised court appearances with proven DV offences in last 5 years	**History of family and domestic violence**	✓	
Number of prior finalised court appearances with proven offences against justice procedures in last 5 years		✓	
If POI had prior conviction		✓	
If POI had been given full-time prison or control order in last 5 years		✓	
a01 - Has your partner ever threatened to harm or kill you?	**Threats to kill**		
a02 - Has your partner ever used physical violence against you?			✓
a03 - Has your partner ever choked, strangled, suffocated you or attempted to do any of these things?	**Non-lethal strangulation (or choking)**		
a05 - Has your partner ever harmed or killed a family pet or threatened to do so?	Abuse of pets and other animals		
a06 - Has your partner ever been charged with breaching an apprehended domestic violence order?			✓
a07 - Is your partner jealous towards or controlling of you?	**Coercive control**		
a08 - Is the violence or controlling behaviour becoming worse or more frequent?	**Escalation (frequency and/or severity)**		
a09 - Has your partner stalked or constantly harassed or texted/ e-mailed you?	**Stalking**		
a10 - Does your partner control your access to money?	Isolation and barriers to help-seeking		
a11 - Has there been a recent separation (in last 12 months) or is one imminent?	**Separation (actual or pending)**		
a13 - Is your partner unemployed?			✓
a14 - Does your partner have mental health problems (including undiagnosed conditions) and/or depression?			✓
a15 - Does your partner have a problem with substance abuse such as alcohol or other drugs?	Misuse of drugs or excessive alcohol consumption		✓
a16 - Has your partner ever threatened or attempted suicide?	Suicide threats and attempts		
a17 - Is/has your partner currently on bail, parole, served a time of imprisonment or has recently been released from custody in relation to offences of violence?			✓
a18 - Does your partner have access to firearms or prohibited weapons?	**Perpetrator's access to, or use of weapons**		
a19 - Are you pregnant and/or do you have children who are less than 12 months apart in age?	**Pregnancy and new birth**		✓
a22 - Is there any conflict between you and your partner regarding child contact or residency issues and/or current Family Court proceedings?	Court orders and parenting proceedings		
a23 - Are there children from a previous relationship present in the household?			✓
b01 - How fearful is the victim of the other party? (Not Afraid, Afraid, Unable to Answer, Terrified)	Victim's self-perception of risk		

Table 3. Descriptive statistics, training and test data

	Training data, 2016-2017 N=155,515			Test data, 2018 N=78,939		
Variables	Missing	Mean	SD	Missing	Mean	SD
Had another IPV incident within 12 months following the index incident	0	0.251	-	0	0.265	-
Number of prior incidents recorded in CRP	0	1.43	2.46	0	2.21	3.55
Number of prior intimate partner incidents in which sexual offence reported	0	0.000129	0.0113	0	0.000165	0.0128
POI Age	0	36.4	11.1	0	36.6	11.2
POI gender (1=Female)	51,681	0.158	-	26,111	0.167	-
Number of prior finalised court appearances with proven offences	51,681	5.08	6.11	26,111	5.23	6.34
Number of prior finalised court appearances with proven DV offences in previous year	51,681	0.141	0.414	26,111	0.149	0.427
Number of prior finalised court appearances with proven DV offences in last 5 years	51,681	0.475	0.951	26,111	0.525	0.996
Number of prior finalised court appearances with proven offences against justice procedures in last 5 years	51,681	0.694	1.27	26,111	0.744	1.35
If POI had prior conviction	51,681	0.765	-	26,111	0.762	-
If POI had been given full-time prison or control order in last 5 years	51,681	0.243	-	26,111	0.254	-
a01 - Has your partner ever threatened to harm or kill you?	0	0.227	-	0	0.222	-
a02 - Has your partner ever used physical violence against you?	0	0.320	-	0	0.326	-
a03 - Has your partner ever choked, strangled, suffocated you or attempted to do any of these things?	0	0.086	-	0	0.078	-
a05 - Has your partner ever harmed or killed a family pet or threatened to do so?	0	0.021	-	0	0.019	-
a06 - Has your partner ever been charged with breaching an apprehended domestic violence order?	0	0.113	-	0	0.126	-
a07 - Is your partner jealous towards or controlling of you?	0	0.281	-	0	0.279	-
a08 - Is the violence or controlling behaviour becoming worse or more frequent?	0	0.233	-	0	0.234	-
a09 - Has your partner stalked or constantly harassed or texted/ e-mailed you?	0	0.172	-	0	0.173	-
a10 - Does your partner control your access to money?	0	0.064	-	0	0.059	-
a11 - Has there been a recent separation (in last 12 months) or is one imminent?	0	0.328	-	0	0.352	-
a13 - Is your partner unemployed?	0	0.247	-	0	0.240	-
a14 - Does your partner have mental health problems (including undiagnosed conditions) and/or depression?	0	0.204	-	0	0.214	-
a15 - Does your partner have a problem with substance abuse such as alcohol or other drugs?	0	0.270	-	0	0.270	-
a16 - Has your partner ever threatened or attempted suicide?	0	0.113	-	0	0.110	-
a17 - Is/has your partner currently on bail, parole, served a time of imprisonment or has recently been released from custody in relation to offences of violence?	0	0.129	-	0	0.135	-
a18 - Does your partner have access to firearms or prohibited weapons?	0	0.022	-	0	0.021	-
a19 - Are you pregnant and/or do you have children who are less than 12 months apart in age?	0	0.046	-	0	0.045	-
a22 - Is there any conflict between you and your partner regarding child contact or residency issues and/or current Family Court proceedings?	0	0.076	-	0	0.079	-
a23 - Are there children from a previous relationship present in the household?	0	0.094	-	0	0.100	-
b01 - How fearful is the victim of the other party? (1=Afraid or Terrified)	0	0.331	-	0	0.327	-

Note. SD – standard deviation; POI – person of interest. For binary variables with a yes/no answer (i.e. for rows without a SD value), the mean is equivalent to the proportion. For variables with missing values, the mean or proportion has a sample size of (N – Missing).

As explained earlier, the refined DVSAT tool should:

1. have good predictive performance,
2. be parsimonious, with as few predictors as necessary to achieve good predictive performance, and;
3. accurately identify victims most at risk of re-victimisation.

To achieve (1), we evaluated the four candidate models discussed above on the training data, and selected the one with the best out-of-sample predictive performance. To achieve (2), we took the best-performing model and kept only the best predictors to produce an even more parsimonious model. To evaluate how accurate this parsimonious model could identify victims most at risk, we calculated how often the model predictions were correct (precision) and how often victims who were re-victimised within 12 months were correctly predicted (recall). These are explained in the following sub-sections.

Evaluating predictive performance

The four candidate models correspond to four different sets of predictors. Given a set of predictors as inputs, an important modelling decision is to select a classification method. Two classifiers were compared: logistic regression and gradient boosted trees. Logistic regression is commonly used for binary classification problems and serves as a benchmark. Another popular and competitive alternative is the gradient boosted trees classifier, which has achieved state-of-the-art results on many prediction problems (Chen & Guestrin, 2016). The two classification methods were applied to each of the four models, giving in total eight combinations. In the first part of the analysis, the goal was to identify the best combination, that is the set of predictors and classification method that achieved the highest cross-validated AUC.

The predictive performance of the four models was evaluated using the AUC. Two measures of the AUC statistic may be distinguished: the cross-validated AUC and the out-of-sample, or test-set, AUC. They are calculated the same way but on different sub-samples of the data and serve different purposes. The cross-validated AUC is calculated on the training data. It gives an approximate indication of how well a model would perform on new data and is used to compare the predictive performance of the four models. The model with the highest cross-validated AUC is then chosen. Its predictive performance is to be evaluated on the test data. The AUC that is calculated on the test data is referred to as the test-set AUC. In short, the cross-validated AUC is for model comparison and selection, and the test-set AUC is for final model evaluation.

Two other statistics are also reported. First is the precision of the predictions. This is the proportion of those above the selected risk threshold who were re-victimised within 12 months from the index incident. The second statistic is the recall, or the sensitivity, defined as the proportion of those re-victimised within 12 months from the index incident that the model correctly predicted would be re-victimised.

Removing extraneous predictors

Next, we tried to remove any extraneous predictors which do not contribute to the accuracy of the risk predictions. By removing any irrelevant or weak predictors, it may be possible to obtain an even more parsimonious model which is as predictive as the best performing one. To do this, we first ranked the importance of the set of predictors in the best performing model using a variable importance measure developed by Fisher, Rudin, and Dominici (2019) . Their measure is a permutation-based measure that can be applied to any classification method. It is rooted in the idea that the correlation between a strong predictor and the outcome would be "removed" if the observed values of the predictor were randomly permuted, or shuffled, across observations. The variable importance measure we use is the reduction in the model's predictive performance after shuffling predictor values across observations. This is calculated as:

$$\text{AUC} - 1/20 \sum_{s=1}^{20} \text{AUCs}$$

where AUC is the model AUC, and $AUCs$ is the AUC of each shuffle. This reduction in performance would be large for an important predictor, whereas for an irrelevant predictor, permutation would have no effect on the model's accuracy. Therefore, by permuting the values of each predictor in turn and looking at how much the prediction accuracy deteriorates, the relative importance of all the predictors can be inferred. We then quantified any difference in predictive performance by comparing the test-set AUC of the simplified model to that of the best model.

Identifying victims most at risk of re-victimisation

Using the simplified model, we then predicted the risk of being re-victimised for each individual in the test dataset. This predicted risk quantifies the likelihood of re-victimisation and falls between 0 (least likely) and 1 (most likely). Those with a higher predicted risk would be, according to the model, more at risk of future harm than those with a lower predicted risk. The predicted risk could therefore be used to identify those most at risk and most in need of support.

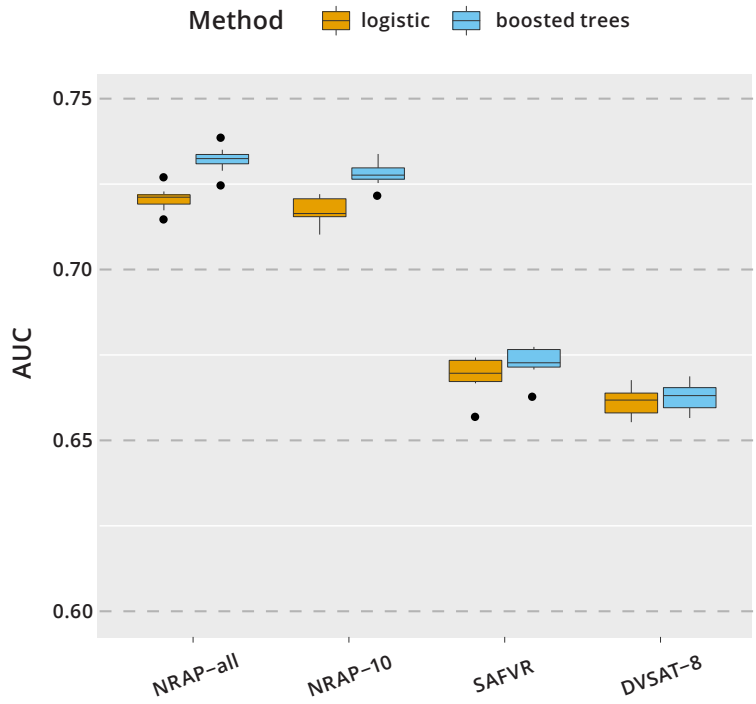
An important consideration is how a predictive model could aid planning of victim support services, such that those most at risk could be prioritised. This would require setting a threshold such that only those victims with a predicted risk above the threshold would be referred. The challenge is to select this threshold to match planned service capacity. To assess this, a risk threshold would first need to be determined using the training data. In the analysis, a 10 per cent service capacity level was assumed. That is, in a given year, it was assumed that existing services could provide support to 10 per cent of victims most at risk. This threshold can be calculated as the 90th percentile of the predicted risks in the training data. Given this threshold, we can calculate the predicted risks of victims in the test data and check what percentage of victims had a predicted risk above the threshold. By comparing this percentage to the hypothetical 10 per cent service capacity, we can measure how well the model could be used to prioritise victims to be offered support. To measure the quality of the predictions, we re-calculated the precision and statistics at the selected risk threshold.

RESULTS

Predictive performance

Figure 1 shows the 10-fold cross-validated AUC of the four models, estimated using logistic regression and boosted trees. The thickest line in each box is the median of the 10 cross-validated AUCs. The overall best model is the NRAP-all model with a gradient boosted trees classifier, with an AUC of .732. The worst performing model is the DVSAT-8 model, with an AUC of only .661. Both NRAP models, regardless of classifier used, have an AUC above 0.7. For each of the four models, the gradient boosted trees classifier achieved a slightly higher AUC than the logistic regression.

Figure 1. Predictive performance of the models



These cross-validated AUCs approximate the test-set AUCs. As Table 4 shows, for the NRAP-all model, the average cross-validated AUC (.732) and the test-set AUC (.738) are almost identical. The test-set precision is slightly lower than in the cross-validation precision, at .552. In other words, out of 100 predictions, only 55 were correct. The test-set recall is higher than the cross-validated recall, at .292. In other words, out of 100 IPV incidents that had another IPV incident within 12 months, 29 were predicted by the model.

Table 4. Predictive performance metrics of the four predictive models

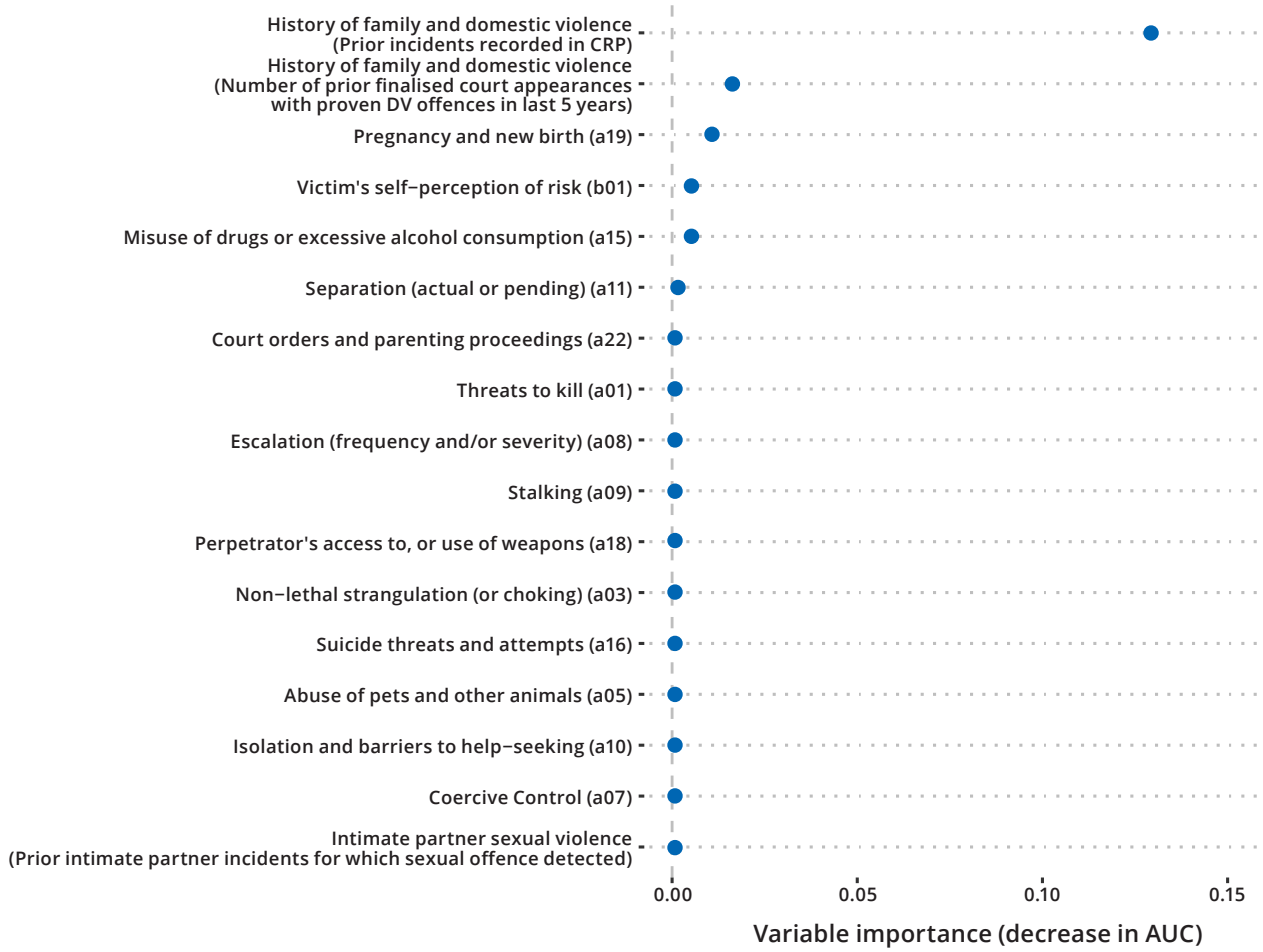
Model	Classifier	AUC	Precision	Recall
Test set				
NRAP-all	**Boosted trees**	**.738**	**.552**	**.292**
Avg. of 10 validation sets				
NRAP-all	**Boosted trees**	**.732**	**.564**	**.195**
NRAP-all	Logistic	.721	.576	.155
NRAP-10	Boosted trees	.728	.561	.190
NRAP-10	Logistic	.717	.573	.148
SAFVR	Boosted trees	.673	.547	.049
SAFVR	Logistic	.669	.520	.053
DVSAT-8	Boosted trees	.663	.492	.060
DVSAT-8	Logistic	.661	.496	.070

Note. AUC - Area under the Curve. Precision is also known as positive predictive value. Recall is also known as sensitivity. Precision and recall are calculated at the 0.5 threshold.

A simplified model

Figure 2 shows the importance of all 16 predictors in the NRAP-all model, ranked from the most important (at the top) to the least important (at the bottom), calculated using training data only. For each variable, the dot indicates how much the model's performance, as measured by the AUC, would decrease on average if the values of that variable were randomly permuted. The average was calculated over 20 repetitions. The larger the decrease is, the more important the predictor is. For example, if history of family and domestic violence (as measured by the number of previous incidents recorded in the CRP), the most important predictor of re-victimisation, is excluded from the model, the AUC decreases by about .126. As Figure 2 shows, the top five predictors are: the two measures of history of family and domestic violence; pregnancy and new birth; the victim's self-perception of risk; and misuse of drugs or excessive alcohol consumption. The remaining predictors contributed little to the model's predictive performance.

Figure 2. Variable importance of predictors in the NRAP-all model



It may come as a surprise that most of the risk factors contribute so little to the model's performance. This does not mean these risk factors are unimportant for other purposes, but rather that after having included the top 5 predictors in the model, the additional predictive performance from the remaining predictors is low.

If most of the predictors do not augment the predictive performance of the model, it follows that accuracy would not decrease much if these weaker predictors were excluded from the model. Keeping only the top five predictors, a simplified model was re-estimated on the training data. Table 5 compares the test-set AUC of the full NRAP-all model and that of the simplified model. The various metrics are almost identical, suggesting that the simplified model is as good as the NRAP-all model in predicting repeat victimisation.

Table 5. Predictive performance for NRAP-all and the simplified model on test data

Model	Classifier	AUC	Precision	Recall
NRAP-all	Boosted trees	.738	.552	.292
Simplified	Boosted trees	.734	.548	.291

Note. AUC - Area under the Curve. Precision is also known as positive predictive value. Recall is also known as sensitivity. Precision and recall are calculated at the 0.5 threshold.

Prioritising victims most at risk

So far we have shown that the NRAP-all model using a boosted trees classifier is the best of the four candidate models, and that keeping only the most important predictors, it can be made simpler with no decrease in predictive performance.

We now consider how well this simplified model would be used to identify victims most at risk to prioritise support. To do this, we determine whether the best model can be used to identify the 10 per cent of victims at the highest risk. First we calculated the 90th percentile of predicted risk in the training data. To minimise any temporal differences in the likelihood of re-victimisation, only 2017 data was used. This threshold is .505, and 10 per cent of victims in 2017 had a predicted risk above this threshold. The distribution of the predicted risk is shown in the top panel of Figure 3. We find that 13.8 per cent of the victims in the test data had a predicted risk above the threshold. The discrepancy of 3.8 percentage points is therefore a factor to be accounted for when using the model to prioritise victims. This also illustrates that even though the proportion of incidents with a subsequent IPV incident within 12 months is almost identical in the training and test data (see Table 3), the 90th percentile could be higher in the test data.

The precision and recall statistics depend on the risk threshold.[27] Reported in Table 6 are various performance metrics at this new risk threshold of 0.505. The metrics reported in Table 4 and Table 5 have an assumed threshold of 0.500. In general, raising the threshold would likely lead to fewer false positives (hence an increase in precision) and fewer true positives (hence a decrease in recall). This is the case. However, given the small change in the threshold, from 0.500 to 0.505, the changes in precision and recall are almost negligible. Same as before, 55 out of 100 victims predicted to be re-victimised had an IPV incident in 12 months, but out of every 100 victims re-victimised, the model only correctly predicted 29.

27 All reported performance metrics other than the AUC would change if the risk threshold changes.

Figure 3. Comparing the proportions of victims above the selected threshold of 0.505 in 2017 and 2018

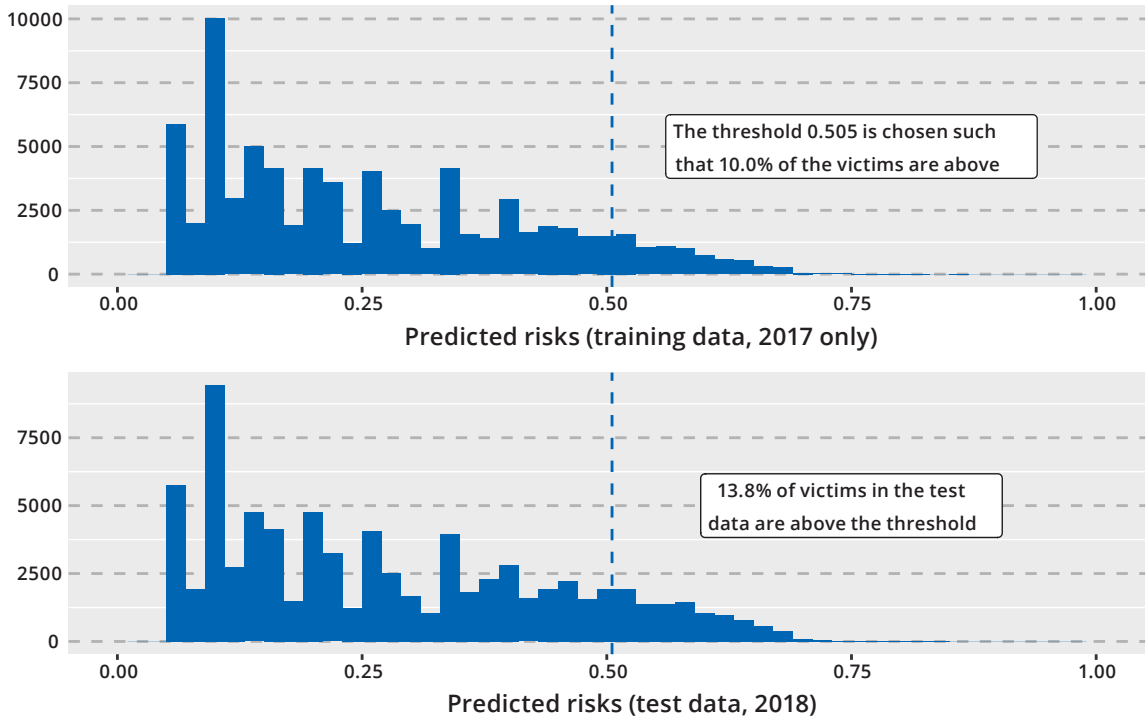


Table 6. Predictive performance for the simplified model at different thresholds

Model	Classifier	Threshold	AUC	Precision	Recall*
Simplified	Boosted trees	.500	.734	.548	.291
Simplified	Boosted trees	.505	.734	.549	.286

* Recall is also known as sensitivity.

DISCUSSION

The primary aim of this study was to improve the accuracy of the DVSAT in predicting repeat victimisation. Using NSW recorded victim data, we compared the predictive accuracy of four risk assessment models in identifying victims of IPV who are at risk of being re-victimised in the subsequent 12 months. Three of these models are Australian. Two models used risk factors that were identified by Australia's National Research Organisation for Women's Safety (ANROWS). In this study, these models are referred to as 'NRAP-10' (National Risk Assessment Principles; this model contains the ten 'high-risk' factors identified by ANROWS) and 'NRAP-all'; the latter comprises 16 factors (a combination of the ten 'high-risk' factors and the six 'other' factors identified by ANROWS). The third Australian model in this study, the Domestic Violence Safety Assessment Tool (DVSAT), was developed in NSW for use by police officers; in these analyses, a refined version of the original DVSAT was used, only including the eight items that Ringland (2018) found to be independently associated with intimate partner repeat victimisation within 12 months. The last model was the Static Assessment of Family Violence Recidivism (SAFVR), developed by New Zealand Police.

Of these four models, only the two NRAP models were able to adequately discriminate between victims who experienced subsequent incidents of IPV from victims who did not. The NRAP-all model was the better-performing of these two models, achieving an AUC of .738, which is regarded as an acceptable level of discrimination in predicting repeat victimisation. Its predictive performance was not compromised when the model was simplified by removing the weak, or less relevant, predictors and retaining only the strongest, or more important, predictors. This empirically simplified model achieved an AUC of .734, which is an improvement over the AUC of .659 of the DVSAT-8 model. This result is comparable to that achieved by the NRAP-all model and better than the 13-item ODARA scale which had the highest average weighted AUC (.666) out of five IPV risk assessment instruments reviewed by Messing and Thaller (2013). In the current study, the five variables that most strongly predicted repeat victimisation in the subsequent 12 months were: two measures of previous history of family and domestic violence; pregnancy and new birth; the victim's self-perception of risk of future violence; and misuse of drugs or excessive alcohol consumption. The analyses showed that the predictors that were least important, or the weakest, in predicting repeat victimisation were: a separation (either actual or pending); court orders, such as intervention orders, and parenting proceedings; the partner's threats to kill; an escalation in frequency and/or severity in the partner's violence over time; stalking by the partner; the partner's access to, or use of, firearms; non-lethal strangulation or choking by the partner; coercive control by the partner in the form of controlling the victim's access to money; the partner's threats or attempts to commit suicide; the partner's actual or threatened abuse, cruelty or harm of family pets and other animals; partner control of victim's access to money; coercive control in the form of jealousy or controlling behaviour; and known incidents of sexual violence. In other words, a risk assessment tool can discriminate between repeat and non-repeat victims of IPV even with the removal of a large number of questions about the perpetrator's past or current behaviour. It should be emphasised that these weaker predictors of re-victimisation could still be predictive of other outcomes, such as escalation in the frequency or severity of re-victimisation.

This study clearly demonstrates that it is possible to rely on a risk assessment tool with a small number of variables to predict IPV re-victimisation in NSW. The practical application of these results could have substantial and far-reaching implications for the victims of IPV. In short, providing effective interventions to those victims at the highest risk could improve their safety and quality of life, as well as prevent further, potentially fatal, incidents. In addition, there would be considerable associated benefits for the NSW Police Force, the various agencies involved in the resource-intensive Safety Action Meetings and the services providing support to the victims. Since it could be limited to including only the five strongest predictors of repeat victimisation, such a risk assessment tool would clearly be short, easy-to-use and quick to administer by NSW police officers, allowing them to rapidly assess the level of risk in IPV cases and reduce the impact on police resources. In fact, information regarding one of the five predictors – the history of family and domestic violence, at least that which has been reported to the police – is readily accessible and can be extracted from the police database. Police officers would require minimal, perhaps no, specific training in administering the tool. Being asked only a handful of questions means that very little information would be required from victims at the time of the incident, decreasing the stress levels for everyone involved. Not surprisingly, some victims become annoyed when asked questions that they perceive to be irrelevant, particularly if the questions are asked in a mechanical manner; these feelings would be exacerbated if police officers attend the same location frequently (Her Majesty's Inspectorate of Constabulary, 2015).[28]

A further advantage of the identified risk assessment approach is the ability to control the volume of referrals at a desired level. Some of the obvious corollaries of fewer referrals to the fortnightly Safety Action Meetings (SAMs) are less preparation time for the attendees and shorter meetings, allowing more time and resources to be focused on those victims who are genuinely at serious threat of future harm or injury. Victim safety could be improved by allocating scarce resources more effectively towards those cases genuinely at greatest risk. These issues have been of concern to meeting members. Each of the

28 Neither the number of victims interviewed nor the number giving these responses is reported.

54 key stakeholders[29] who were interviewed in 2015 during BOCSAR's process evaluation of the Safer Pathway program (unpublished) commented on the unexpectedly large number of referrals to SAMs; stakeholders noted that this meant that they spent substantial periods of time collating information regarding each victim, her/his children and alleged perpetrator in preparation for discussion at the SAMs. All stakeholders stated that meetings were too long, often lasting between three and four hours instead of the expected two hours.[30] Fewer referrals to the meetings would decrease these concerns.

The effectiveness of government-funded IPV programs could also increase if violence, or even death, is prevented by the more accurate prediction of intimate partner repeat victimisation and more focused targeting of resources to those victims at greatest risk of harm. Indeed, referrals to SAMs made because of inaccurate predictions may have been a reason Wan et al. (2018) obtained mixed findings in their evaluation of the *Safer Pathway* program. In other words, the program could have been ineffective in some locations because both high- and low-risk cases were being identified and referred to SAMs.

It is evident from this study that even a refined DVSAT is less accurate than our simplified model in predicting intimate partner repeat victimisation within 12 months (AUC = .66). This is perhaps not surprising since the DVSAT was not developed empirically, but instead items were selected following a review of existing risk assessment tools, as well as Australian and international research on DFV related homicide. Furthermore, it was not tested empirically prior to its state-wide implementation in 2015. The DVSAT's poor predictive accuracy is, however, concerning given the tool is currently being used across NSW to determine the type and extent of support and intervention that is offered to victims of domestic and family violence.

There is a degree of consistency between the results of this study and those of other Australian researchers who also found that a risk assessment tool with a small number of key variables can accurately predict future re-victimisation (e.g. Dowling & Morgan, 2019; Mason & Julian, 2009). However, there is limited overlap between the studies regarding which specific variables are the strongest predictors. For example, only two variables were found to be strong predictors of repeat victimisation by both Dowling and Morgan and this study: pregnancy and new birth, and a history of family and domestic violence (or 'the offender has committed offences of violence against any person in the past', in Dowling and Morgan's research). Conversely, some variables were found to be significant predictors of repeat victimisation in this study but not by Dowling and Morgan, for example, misuse of drugs or excessive alcohol consumption. One explanation could be that the two strong predictors identified by both our and Dowling and Morgan's study are easier for police to accurately assess.

While the results of this study are encouraging, further empirical research is necessary to convert them to a practical and effective tool. The next steps could involve constructing a risk assessment tool that includes the five predictors and testing the tool in genuine IPV incidents in two or three Police Area Commands (PACs)/Police Districts (PDs) in NSW to determine how accurately it predicts repeat victimisation within specific timeframes, for example, 3 months, 6 months, 12 months or longer. This would provide an indication of the model's performance in predicting more immediate re-victimisation as well. The performance of the new tool could then be compared to the performance of the existing DVSAT being used in other PACs/PDs which have similar characteristics (e.g. trends in DV re-victimisation) to the treatment PACs/PDs. Each tool could also be examined to determine how well it predicts repeat victimisation in different sub-samples of the population. For example, victims from specific ethnic groups, victims in same-sex relationships, victims with dependent children compared to those without, and victims who have experienced family violence. Clearly, such a test would require considerable co-operation from NSW Police Force at both the senior and the operational levels.

29 The vast majority of stakeholders who were interviewed were members of the Safety Action Meetings in the two initial Safer Pathway sites of Orange and Waverley launched in September 2014; the latter site comprised four NSW Police Local Area Commands (LACs) – Botany Bay LAC, Eastern Beaches LAC, Eastern Suburbs LAC and Rose Bay LAC.

30 According to the Safety Action Meeting Manual (NSW Government, 2020, p. 37), 'SAMs are generally two hours in duration'.

One of the strengths of this study is the large sample size of incidents (N = 234,454) used in the modelling. However, there are limitations. The second most important predictor in the simplified model, history of family and domestic violence, was derived from ROD. Given the time it takes to finalise court appearances, recent criminal charges immediately preceding an IPV incident may be awaiting a decision, a factor that needs to be accounted for when operationalising the model. Another limitation is that even the best model would still miss over 70 per cent (given a recall of 0.29) of victims who would be re-victimised. There is certainly room for improvement, for example if new predictors became available. Finally, a major weakness of this study is that it relied solely on reported victimisation episodes and may be vulnerable to under-reporting. A large proportion of DFV incidents are not reported to police. In fact, according to the 2016 *Personal Safety Survey* (Australian Bureau of Statistics, 2020), only 34 per cent of women reported their most recent incident of physical assault by a male partner to police. It is not clear to what extent the risk models developed here are relevant for prioritising services to victims who are not known to police. It could be that the models are more or less predictive when including unreported incidents of DFV, but that would require survey or other self-report data that is not vulnerable to under-reporting.

This study demonstrates that a risk assessment instrument with a small number of variables can both identify victims who are most at risk of future IPV and control the volume of referrals at a desired level. Such an instrument would allow government agencies not only to effectively target resources to improve the safety of those genuinely at greatest risk, but also to manage and potentially prevent repeat victimisation.

ACKNOWLEDGEMENTS

We would like to thank Feroz Sattar from Victims Services NSW for providing the data; Joanna Superina, Grace Beasant, Andrew Hurst, Sean McDermott, Trent Swinton from NSW Police Force, Alice Stiles and Fiona Quigley from Legal Aid NSW for the many helpful discussions; Suzanne Poynton, Sara Rahman and Clare Ringland from BOCSAR, and the two anonymous reviewers for their valuable feedback; and Florence Sin for desktop publishing the report.

REFERENCES

Australian Bureau of Statistics. (2017). *Personal Safety, Australia, 2016* (Cat. No. 4906.0). Retrieved 2 April 2020 from: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/4906.0~2016~Main%20Features~Key%20Findings~1

Australian Bureau of Statistics. (2020). *Partner Violence - In Focus: Crime and Justice Statistics.* Retrieved 2 April 2020 from: https://www.abs.gov.au/statistics/people/crime-and-justice/focus-crime-and-justice-statistics/latest-release

Commonwealth Department of Social Services. (2016). *National plan to reduce violence against women and their children: Third action plan 2016-2019 of the national plan to reduce violence against women and their children 2010-2022*. Retrieved 2 April 2020 from: https://www.dss.gov.au/sites/default/files/documents/10_2016/third_action_plan.pdf

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.

Dillon, G., Hussain, R., Loxton, D., & Rahman, S. (2013). Mental and physical health and intimate partner violence against women: A review of the literature. *International Journal of Family Medicine, 2013*. Retrieved 7 April 2020 from: http://dx.doi.org/10.1155/2013/313909

Dowling, C., & Morgan, A. (2019). *Predicting repeat domestic violence: Improving police risk assessment* (Trends & Issues in Crime and Criminal Justice, No. 581). Retrieved 2 April 2020 from Australian Institute of Criminology website: https://www.aic.gov.au/sites/default/files/2020-05/ti581_predicting_repeat_domestic_violence.pdf

Eke, A. W., Hilton, N. Z., Harris, G. T., Rice, M. E., & Houghton, R. E. (2011). Intimate partner homicide: Risk assessment and prospects for prediction. *Journal of Family Violence, 26*, 211-216.

Fisher, A., Rudin, C., & Dominici, F. (2019). All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. *Journal of Machine Learning Research, 20*(177), 1-81.

Graham, L. M., Sahay, K. M., Rizo, C. F., Messing, J. T., & Macy, R. J. (2021). The validity and reliability of available intimate partner homicide and reassault risk assessment tools: A systematic review. *Trauma, Violence and Abuse*, 1-23.

Gray, A. L. (2012). Assessing risk for intimate partner violence: A cross-validation of the ODARA and DVRAG within a sample of incarcerated offenders. Thesis for Master of Arts in Psychology, Carleton University, Canada. Retrieved 29 April 2020 from: https://curve.carleton.ca/system/files/etd/72e3aeb0-7550-4f38-9f47-d50f33431895/etd_pdf/6c6aacb2023c96d125559e194b77f097/gray-assessingriskforintimatepartnerviolenceacrossvalidation.pdf

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19–30.

Hanson, R. K., Helmus, L., & Bourgon, G. (2007).*The validity of risk assessments for intimate partner violence: A meta-analysis*. Public Safety Canada. Retrieved 29 April 2020 from: https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/ntmt-prtnr-vlnce/ntmt-prtnr-vlnce-eng.pdf

Her Majesty's Inspectorate of Constabulary. (2015). *Increasingly everyone's business: A progress report on the police response to domestic abuse.* London: HMIC. Retrieved 2 April 2020 from : https://www.justiceinspectorates.gov.uk/hmicfrs/wp-content/uploads/increasingly-everyones-business-domestic-abuse-progress-report.pdf

Hilton, N. Z., & Eke, A. W. (2016). Non-specialization of criminal careers among intimate partner violence offenders. *Criminal Justice and Behavior, 43*(10), 1347-1363.

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*(2), 326-337.

Hilton, N. Z., Harris, G. T., Popham, S., & Lang, C. (2010). Risk assessment among incarcerated male domestic violence offenders. *Criminal Justice and Behavior 37*(8), 815-832.

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 ODARA. *Psychological Assessment, 16*(3), 267-275.

Hilton, N. Z., Popham, S., Lange, C., & Harris, G. T. (2014). Preliminary validation of the ODARA for female intimate partner violence offenders. *Partner Abuse, 5*(2), 189-203.

Hosmer, D. W., & Lemeshow, S. (2004). *Applied Logistic Regression*. New Jersey: John Wiley & Sons.

Howard, L. M., Trevillion, K., & Agnew-Davies, R. (2010). Domestic violence and mental health. *International Review of Psychiatry, 22*(5), 525-534.

Jolliffe Simpson, A. D., Joshi, C., & Polaschek, D. L. (2021). Predictive Validity of the DYRA and SAFVR: New Zealand Police's Family Violence Risk Assessment Instruments. *Criminal Justice and Behavior, 48*(10), 1487-1508.

KPMG. (2016). *The cost of violence against women and their children in Australia: Final report.* Retrieved 2 April 2020 from: https://www.dss.gov.au/sites/default/files/documents/08_2016/the_cost_of_violence_against_women_and_their_children_in_australia_-_summary_report_may_2016.pdf

Kropp, R. P., & Hart, S. D. (2004). *The development of the Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER): A tool for criminal justice professionals*. Department of Justice Canada. Retrieved 2 April from: https://www.justice.gc.ca/eng/rp-pr/fl-lf/famil/rr05_fv1-rr05_vf1/rr05_fv1.pdf

Kutin, J., Russell, R., & Reid, M. (2017). Economic abuse between intimate partners in Australia: Prevalence, health status, disability and financial stress. *Australian and New Zealand Journal of Public Health, 41*(3), 269-274.

Lagdon, S., Armour, C. & Stringer, M. (2014). Adult experience of mental health outcomes as a result of intimate partner violence victimisation: A systematic review. *European Journal of Psychotraumatology, 5*(1). Retrieved 7 April 2020 from: https://www.tandfonline.com/doi/full/10.3402/ejpt.v5.24794

Lauria, I., McEwan, T. E., Luebbers, S., Simmons, M., & Ogloff, J. R. P. (2017). Evaluating the Ontario Domestic Assault Risk Assessment in an Australian frontline police setting. *Criminal Justice and Behavior, 44*(12), 1545-1558.

Lum On, M., Ayre, J., Webster, K., & Moon, L. (2016). *Examination of the health outcomes of intimate partner violence against women: State of knowledge paper* (Landscapes state of knowledge, Issue 3). Sydney, Australia's National Research Organisation for Women's Safety to reduce violence against women and children (ANROWS).

Mason, R., & Julian, R. (2009). *Analysis of the Tasmania Police Risk Assessment Screening Tool (RAST): Final report*. Tasmanian Institute of Law Enforcement Studies, University of Tasmania.

McEwan, T. E., Bateson, S., & Strand, S. (2017). Improving police risk assessment and management of family violence through a collaboration between law enforcement, forensic mental health and academia. *Journal of Criminological Research, Policy and Practice, 3*(2), 119-131.

McEwan, T. E., Shea, D. E., & Ogloff, J. R. P. (2019). The development of the VP-SAFvR: An actuarial instrument for police triage of Australian family violence reports. *Criminal Justice and Behavior, 46*(4), 590-607.

Messing, J. T., & Thaller, J. (2013). The average predictive validity of intimate partner violence risk assessment instruments. *Journal of Interpersonal Violence, 28*(7), 1537-1558.

New Zealand Police (no date). *Family harm new risk measures – handout: Police approach to family harm*. Retrieved 29 April 2020 from: https://www.police.govt.nz/sites/default/files/publications/police-approach-to-family-harm-infographic-web.pdf

Nicholls, T. L., Pritchard, M. M., Reeves, K. A., & Hilterman, E. (2013). Risk assessment in intimate partner violence: A systematic review of contemporary approaches. *Partner Abuse, 4*(1), 76-168. Retrieved 29 April 2020 from: https://www.domesticviolenceintervention.net/wp-content/uploads/2014/02/Nicholls.etal2013.Manuscript.pdf

NSW Bureau of Crime Statistics and Research. (2019). *NSW recorded crime statistics quarterly update, December 2019*. Retrieved 2 April 2020 from: https://www.bocsar.nsw.gov.au/Pages/bocsar_media_releases/2020/mr-NSW-Recorded-Crime-Statistics-Quarterly-Update-Dec-2019.aspx

NSW Bureau of Crime Statistics and Research. (2015). *Safer pathway program process evaluation*. Unpublished report.

NSW Department of Justice. (2014). *It Stops Here: Safer Pathway Overview*. Retrieved 19 March 2020 from: https://www.facs.nsw.gov.au/download?file=592949

NSW Government. (2014a). *It Stops Here: Standing together to end domestic and family violence in NSW, the NSW Government's domestic and family violence framework for reform*. Retrieved 2 April 2020 from: https://www.facs.nsw.gov.au/download?file=593053

NSW Government. (2014b). *Domestic Violence Safety Assessment Tool Guide*. Retrieved 2 April 2020 from: https://www.facs.nsw.gov.au/__data/assets/file/0009/593064/DVSAT_guide.pdf

NSW Government. (2014c). *Domestic Violence Safety Assessment Tool*. Retrieved 2 April 2020 from: https://www.facs.nsw.gov.au/download?file=592948

NSW Government. (2016). *NSW Domestic and Family Violence Blueprint for Reform 2016-2021: Safter Lives for Women, Men and Children.* Retrieved 10 August 2021 from: http://domesticviolence.nsw.gov.au/__data/assets/pdf_file/0004/379849/dfv-blueprint-for-reform.pdf

NSW Government. (2017). *Safety Action Meeting Manual*. Retrieved 2 April 2020 from: https://www.facs.nsw.gov.au/__data/assets/file/0004/592951/Safety_Action_Meeting_Manual_2017.pdf

NSW Government. (2019). *Safer Pathway evaluation: Final report*. Retrieved 2 April 2020 from: https://www.women.nsw.gov.au/download?file=650328

Rahman, S. (2018). *Assessing the risk of repeat intimate partner assault* (Crime and Justice Bulletin No. 220). Retrieved 2 April 2020 from NSW Bureau of Crime Statistics and Research website: https://www.bocsar.nsw.gov.au/Publications/CJB/2019-Report-Assessing-the-risk-of-repeat-intimate-partner-assault-CJB220.pdf

Rettenberger, M., & Eher, R. (2013). Actuarial risk assessment in sexually motivated intimate-partner violence. *Law and Human Behavior, 37*(2), 75-86.

Ringland, C. (2018). *The Domestic Violence Safety Assessment Tool (DVSAT) and intimate partner repeat victimisation* (Crime and Justice Bulletin No. 213). Retrieved 2 April 2020 from NSW Bureau of Crime Statistics and Research website: https://www.bocsar.nsw.gov.au/Publications/CJB/2018-Report-Domestic-Violence-Safety-Assessment-Tool-cjb213.pdf

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499-513.

Stylianou, A. M. (2018). Economic abuse experiences and depressive symptoms among victims of intimate partner violence. *Journal of Family Violence, 33*, 381–392. Retrieved 7 April 2020 from: https://doi.org/10.1007/s10896-018-9973-4

Svalin, K., & Levander, S. (2019). The predictive validity of intimate partner violence risk assessments conducted by practitioners in different settings – a review of the literature. *Journal of Police and Criminal Psychology*. Retrieved 29 April 2020 from: https://link.springer.com/article/10.1007%2Fs11896-019-09343-4

Toivonen, C., & Backhouse, C. (2018). *National Risk Assessment Principles for domestic and family violence* (ANROWS Insights 07/2018). Sydney, NSW: Australia's National Research Organisation for Women's Safety Ltd (ANROWS).

Ulmer, J. C. (2015). The Ontario Domestic Assault Risk Assessment (ODARA): A validation and comparison study for an Oregonian law enforcement agency. Doctor of Psychology Dissertation. Retrieved 29 April 2020 from: https://digitalcommons.georgefox.edu/cgi/viewcontent.cgi?article=1176&context=psyd

Victoria Police. (2019). *Code of practice for the investigation of family violence, 3rd edition.* Retrieved 29 April 2020 from: https://www.police.vic.gov.au/code-practice-investigation-family-violence

Wan, W. Y., Thorburn, H., Poynton, S., & Trimboli, L. (2018). *Assessing the impact of NSW's Safer Pathway Program on recorded crime outcomes – an aggregate-level analysis* (Crime and Justice Bulletin No. 210). Retrieved 2 April 2020 from NSW Bureau of Crime Statistics and Research website: https://www.bocsar.nsw.gov.au/Publications/CJB/2018-Report-Assessing-the-impact-of-NSWs-Safer-Pathway-Program-on-recorded-crime-outcomes-cjb210.pdf

Wheller, L., & Wire, J. (2014). *Domestic abuse risk factors and risk assessment: Summary of findings from a rapid evidence assessment*. College of Policing, UK Police. Retrieved 29 April 2020 from: https://whatworks.college.police.uk/Research/Documents/DA-ROR-Summary-December_14.pdf

Winter, R. (2006). *Researching family violence. Briefing paper No. 2*. Tasmanian Institute of Law Enforcement Studies. University of Tasmania. Retrieved 28 April 2020 from: https://www.utas.edu.au/__data/assets/pdf_file/0008/293696/Briefing_Paper_No_2.pdf

Women NSW (2020). *Safer Pathway*. Retrieved 28 May 2020 from: https://www.women.nsw.gov.au/programs/safer-pathway

World Health Organisation (2013). *Global and regional estimates of violence against women: Prevalence and health effects of intimate partner violence and non-partner sexual violence*. Retrieved 7 April 2020 from: https://apps.who.int/iris/bitstream/handle/10665/85239/9789241564625_eng.pdf?sequence=1

APPENDIX

DVSAT Questionnaire

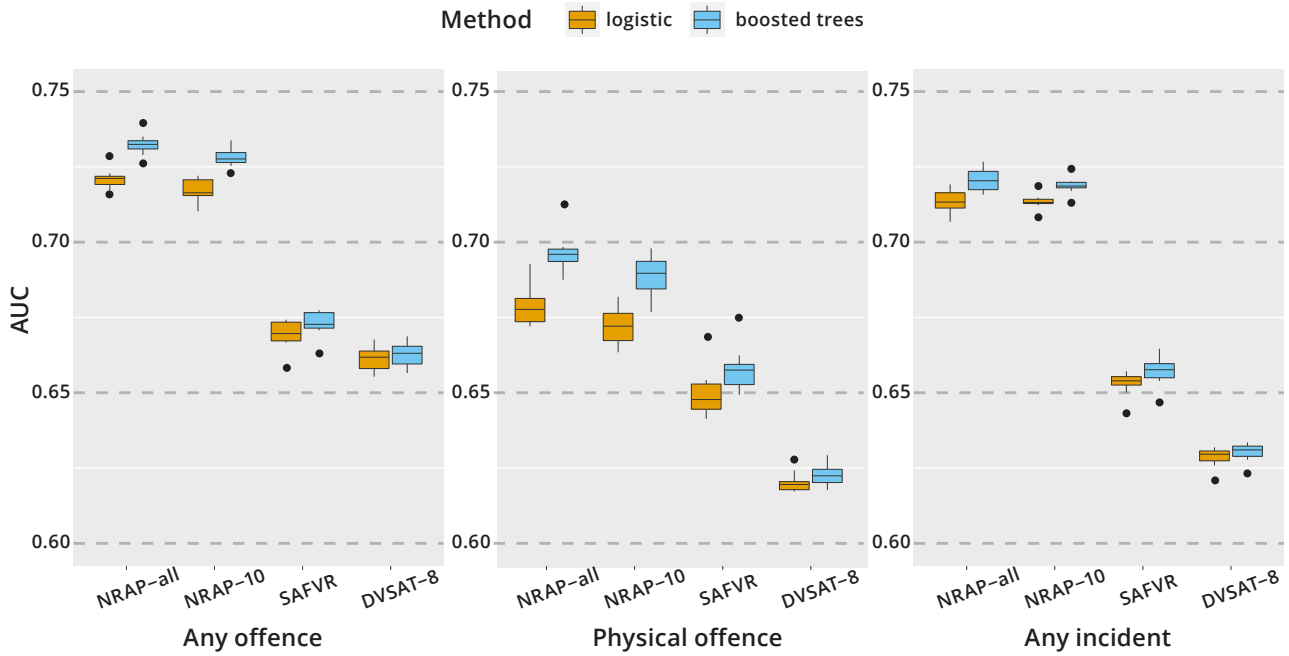
Table A1. DVSAT questions

No.	Questions
A1	Has your partner ever threatened to harm or kill you?
A2	Has your partner ever used physical violence against you?
A3	Has your partner ever choked, strangled, suffocated you or attempted to do any of these things?
A4	Has your partner ever threatened or assaulted you with any weapon (including knives and/or objects)?
A5	Has your partner ever harmed or killed a family pet or threatened to do so?
A6	Has your partner ever been charged with breaching an apprehended domestic violence order?
A7	Is your partner jealous towards or controlling of you?
A8	Is the violence or controlling behaviour becoming worse or more frequent?
A9	Has your partner stalked or constantly harassed or texted/ e-mailed you?
A10	Does your partner control your access to money?
A11	Has there been a recent separation (in last 12 months) or is one imminent?
A12	Does your partner or the relationship have financial difficulties?
A13	Is your partner unemployed?
A14	Does your partner have mental health problems (including undiagnosed conditions) and/or depression?
A15	Does your partner have a problem with substance abuse such as alcohol or other drugs?
A16	Has your partner ever threatened or attempted suicide?
A17	Is/has your partner currently on bail, parole, served a time of imprisonment or has recently been released from custody in relation to offences of violence?
A18	Does your partner have access to firearms or prohibited weapons?
A19	Are you pregnant and/or do you have children who are less than 12 months apart in age?
A20	Has your partner ever threatened or used physical violence towards you while you were pregnant?
A21	Has your partner ever harmed or threatened to harm your children?
A22	Is there any conflict between you and your partner regarding child contact or residency issues and/or current Family Court proceedings?
A23	Are there children from a previous relationship present in the household?
A24	Has your partner ever done things to you, of a sexual nature, that made you feel bad or physically hurt you?
A25	Has your partner ever been arrested for sexual assault?
B1	How fearful is the victim of the other party?
B2	Did the victim express any concerns?
B3	Were there children present or witness to the incident?
B4	If yes, did you sight them and check on their welfare?
B5	Are there any other additional factors, circumstances or details that you are aware of, which make you believe overall there is a serious threat to the safety of the victim or children?

Alternative outcome measures

Figure A1 presents the main results along with those relating to physical offences and any recorded incident. All outcomes relate to incidents involving intimate partners. "Any offence" is as defined in the main text. "Physical offence" includes any robbery, assault, sexual offences, and homicide. "Any incident" refers to any recorded incident in the CRP database, including when no offence was detected and verbal argument. We generally find that the models predicting any incident and physical offences performed worse than our main models. Specifically, these models are better at differentiating between any offence and no offence at all, but perform less well when differentiating the more serious incidents from the less serious ones.

Figure A1. Predictive performance of the models predicting three different outcome variables



NSW BUREAU OF CRIME STATISTICS AND RESEARCH - LEVEL 1, HENRY DEANE BUILDING, 20 LEE STREET, SYDNEY 2000
bcsr@justice.nsw.gov.au • www.bocsar.nsw.gov.au • Ph: (02) 8346 1100 • Fax: (02) 8346 1298
ISSN 2204-5538 (Online) • ISBN 978-1-922576-13-2