



Australian Government

Australian Institute of Criminology

Trends & issues in crime and criminal justice

No. 646 March 2022

Abstract | Risk assessment tools are used widely in the criminal justice response to serious offenders. Despite growing recognition that certain outlaw motorcycle gang (OMCG) members and their clubs are likely to be involved in crime, particularly serious crime, this is not an area where risk assessment tools have been developed and validated.

The nature of offending by OMCGs, and policing responses to OMCGs, requires a novel approach to risk assessment. This study uses machine learning methods to develop a risk assessment tool to predict recorded high-harm offending. Results are compared with those of a model predicting any recorded offending.

The model predicted high-harm offending with a high degree of accuracy. Importantly, the tool appeared able to accurately identify offenders prior to the point of escalation. This has important implications for informing law enforcement responses.

Predicting high-harm offending using machine learning: An application to outlaw motorcycle gangs

Timothy Cubitt and Anthony Morgan

Outlaw motorcycle gangs (OMCGs) are involved in both pervasive and distinctive criminal activity across Australia. While analogous to other types of gangs (Lauchs, Bain & Bell 2015), OMCGs appear to have become increasingly involved in organised criminal enterprise (Blokland et al. 2019; Dowling et al. 2021; Hughes, Chalmers & Bright 2020; Monterosso 2018; Morgan, Dowling & Voce 2020). There is increasing evidence that OMCGs engage in high-harm, organised and violent offending. This paper outlines the development of a risk assessment tool for high-harm offending among OMCG members in New South Wales.

Concentration of offending among outlaw motorcycle gangs

Recent research employing law enforcement data has presented a clearer picture of OMCG offending. Data show that the majority of OMCG members have a criminal record (Blokland et al. 2019; Klement 2016). Research with a large sample of Australian OMCG offenders found that, in the five years prior to the analysis, more than one in five members had been apprehended for a violence or intimidation offence, and one in eight for an organised-crime type offence (Morgan, Dowling & Voce 2020). OMCG members were much more likely than the general population to have a recorded criminal history by their mid-30s.



Serious & Organised Crime
Research Laboratory

Among criminal groups, criminal offending is concentrated within a relatively small subgroup (Ratcliffe & Kikushi 2019). This is also true of OMCGs, with serious offences, including violence and profit-motivated offences, heavily concentrated among a relatively small proportion of members and chapters (Morgan, Dowling & Voce 2020). The same is also true of geographically mobile OMCG members (Dowling & Morgan 2021). Voce, Dowling and Morgan (2021) showed an increasing trend towards violent crime, including weapons offences, among younger members, likely reflecting the changing culture of clubs (Dowling et al. 2021). While differences in offending between gangs has been observed in Australia and overseas (Blokland et al. 2019; Lauchs, Bain & Bell 2015; Morgan, Dowling & Voce 2020), there is evidence that certain individual and club-level factors are associated with a greater involvement in organised crime-type offences (Morgan, Dowling & Voce 2020).

Risk assessment and targeting of police resources

Risk assessment tools are used widely in the criminal justice response to serious offenders. Risk and threat assessments also feature in the response to organised crime (United Nations Office on Drugs & Crime 2010), though they rely much more heavily on the subjective assessment of law enforcement officers and analysts (Albanese 2008; Ratcliffe, Strang & Taylor 2014; Zoutendijk 2010).

Intelligence-led policing has been a cornerstone of the response to OMCGs, and most dedicated policing units are equipped with analyst capability to produce strategic and operational intelligence to inform law enforcement operations. This intelligence has been used to disrupt serious criminal activity by OMCGs (Bjørge 2019), as was reported recently in the Federal Bureau of Investigation and Australian Federal Police-led Operation Ironside, which disrupted several planned murders, some allegedly involving OMCG members (Australian Federal Police 2021).

Given the significant resources directed to policing OMCGs, the number of potential targets and the potential consequences of focusing on the wrong target, there is good reason to explore the potential of structured risk assessment. At the same time, the highly secretive nature and culture of OMCGs (Dowling et al. 2021), and the context in which law enforcement disruption occurs, means that the collection of additional information to inform actuarial assessments (as is done for domestic violence or correctional risk assessments) is prohibitively difficult. Novel methods that exploit existing criminal justice data holdings are therefore required.

Using machine learning to improve criminal justice decision making

The use of machine learning analytics in the criminal justice system is a relatively recent phenomenon. Machine learning analytics have been used in recent years to interrogate policing data with considerable accuracy (Berk et al. 2009; Berk, Sorenson & Barnes 2016; Cubitt, Wooden & Roberts 2020). This type of analysis has been used to forecast domestic violence (Berk 2019; Berk, Sorensen & Barnes 2016; Grogger et al. 2021), improve judicial decision making (Berk & Bleich 2014), forecast high-harm offence types (Berk et al. 2009), and improve the accuracy of criminal justice risk assessments (Berk 2019). Machine learning offers a viable alternative to more traditional analytical methods, allowing data to be interrogated in a more granular fashion, often with greater accuracy than may be expected from, for example, generalised linear models.

That said, there have been criticisms of this approach. Important ethical and jurisprudential concerns have been identified (Berk et al. 2018; Coglianese & Lehr 2017), though often these are overstated or based on misconceptions (Berk 2021). Although there are additional concerns relating to the quality and type of data used to develop prediction models, such as the need to be discerning in training models (Bennett Moses & Chan 2018), this analytical method is increasingly applied to policing data and commonly outperforms traditional analytics (Couronné, Probst & Boulesteix 2018; Grogger et al. 2021). Transparency is important in model development (McKay 2019), while the risks associated with machine learning approaches and applications, including misclassification of individuals, blind adherence to model predictions (Ridgeway 2013) and ineffective implementation strategies (Stevenson & Doleac 2021), need to be considered.

The primary aim of this study was to develop a risk assessment tool for OMCG members using law enforcement data, with a focus on recorded high-harm offending. Given the concentration of offending—especially serious offending—among a small proportion of gang members, the focus on the most harmful offences has obvious benefits for reducing harm to the community and maximising the use of police resources. A secondary aim of this study was to examine whether a risk assessment tool that identifies and focuses disruption efforts on recorded high-harm offending, rather than any recorded offending, could address criticisms of machine learning being applied to OMCG members. These criticisms include the potential for the method to be biased by proactive policing efforts targeting OMCGs, particularly for low-level regulatory offences, such as speeding, and the potential for the over-policing of gangs or gang members that are not responsible for causing harm to the wider community.

Methodology

Data

The current study involved linking three datasets. Extracts were taken from the NSW Police Force's Gangs database and Computerised Operational Policing System (COPS). The Gangs database comprises data relating to gang membership, while the COPS extract provided offence data for individuals affiliated with OMCGs. Data received from COPS consisted of 143,497 offences committed between January 1998 and February 2020 by 5,512 individuals identified as being affiliated with an OMCG. These datasets were linked with additional data on the custodial episodes of these individuals, sourced from the NSW Bureau of Crime Statistics and Research's Reoffending Database. Individuals were matched using the unique number attributed to each individual in the Central Names Index.

To establish a dataset comprising the recorded criminal history of individuals, their custodial history, demographics and gang affiliation, individuals with missing or incomplete data were removed, and the remaining data were then aggregated by Central Names Index identifier and date of birth to remove or resolve any duplicate identities. Finally, deceased individuals were excluded from the dataset. After matching, aggregation and cleaning of data, a total of 3,542 members of various ranks and 1,970 associates, who had collectively been proceeded against for a total of 99,793 offences, were available for analysis. The analysis was restricted to members who were still active ($n=2,246$), meaning former members and associates were excluded from the modelling process.

Outcome of interest

The primary outcome of interest was recorded high-harm offending by OMCG members in the five-year period between January 2015 and December 2019 (the reference period). High-harm offences were defined as offences that featured in the top 10 percent of harm as defined by a modified version of the Western Australian Crime Harm Index (WACHI). The WACHI was developed by House and Neyroud (2018) to assign each offence type in the Australian and New Zealand Standard Offence Classification a harm index based on court penalties. This has previously been used to analyse police data on offences by OMCG members and adapted for this purpose (Morgan, Dowling & Voce 2020). High-harm offences were defined as offences that featured in the top 10 percent of harm as defined by the WACHI for which there were recorded incidents among this group (with WACHI scores ranging from 52 and above). These included:

- murder;
- attempted murder;
- manslaughter;
- aggravated sexual assault;
- import illicit substances;
- aggravated robbery;
- non-aggravated robbery;
- property damage by fire or explosion;
- deal commercial quantities of illicit substances; and
- serious assault causing injury.

The second outcome of interest was any recorded offence by OMCG members in the five-year reference period.

Explanatory variables

There were 130 explanatory variables included in this dataset, primarily relating to the criminal history of each individual, including the number of each type of prior offence, and characteristics of prior offending. A weighted prior harm variable, which was the aggregate harm from all prior recorded offences, adjusted for free time (ie time not spent in custody) was also included. Additional information, such as member rank and gang membership, was derived from the NSW Police Force's Gangs database.

Analytic approach

Given the size and complexity of these datasets, a novel approach was employed for analysis. The ability of machine learning to identify complex structures among data such as recorded offence histories is an important emerging feature of crime analysis (Berk 2013). The random forest algorithm is particularly good at predicting offending behaviours. In this research, a random forest model was used to determine which characteristics of OMCG members were most associated with recorded high-harm offending, and then any recorded offending, in the five-year reference period.

Data were randomised and partitioned into a 70 percent training set and a 30 percent test set. The training set was used to train the algorithm, and the test set was used to test the algorithm (Hyndman & Anthanasopoulos 2014). Modelling was performed through application of pre-process design matrices. Analysis was undertaken using the statistical analysis software R and the 'randomForest', 'dplyr', 'pRoc', 'pdp', and 'ggplot2' packages. The random forest model was trained on instances of high-harm offending before being exposed to the test set. This process was repeated for any recorded offending.

A receiver operating characteristic (ROC) curve was used to identify the predictive accuracy of each model, through the area under the receiver operating characteristic (AUROC) curve. The ROC curve identifies the true positive rate of classification (y-axis), compared with the false positive rate (x-axis) at any threshold value. The AUROC score, which we refer to in simple terms as the predictive accuracy of the model, represents the probability that a randomly selected individual with a recorded high-harm offence will receive a higher risk rating than a randomly selected individual who did not commit a high-harm offence. The AUROC is calculated for the model when applied to the test dataset.

The results of this random forest model are interpreted through mean decrease Gini (MDG; Hong, Xiaoling & Hua 2016). The Gini coefficient is a measure of statistical dispersion, whereby MDG are interpreted as a proportion of the overall random forest model, relative to the AUROC produced by ROC curve. In simple terms, the AUROC identifies how accurate the model's predictions are, while each variable is attributed an MDG coefficient identifying its importance in making the prediction.

To supplement these analyses, a confusion matrix was produced for the test set of both models. The confusion matrix measures the performance of the trained model on the test set for each group, providing a measure of how often the model successfully or unsuccessfully made predictions (Barnes & Hyatt 2012). Several calculations can be made using the confusion matrix, but we focus on the false positive and false negative rates. False positives, which occur when members are incorrectly classified as high-harm offenders when they are not, have potential resource implications in that they can lead police to target individuals not at risk of high-harm offending. Conversely, false negatives—when members are classified as not being high-harm offenders when they do go on to commit a high-harm offence—may result in missed opportunities to disrupt crime and prevent harm to victims.

While we were primarily interested in predicting high-harm offending, as a means of comparison, a separate analysis was computed using the same analytical method but with a different outcome variable. We attempted to predict which OMCG members would commit any criminal offence in the five-year period, and compared the overall predictive accuracy, false positive and false negative rates and variable importance for the two models to understand whether there was a substantial difference in variables that were predictive.

Partial dependence plots (PDPs) were employed as a post-hoc analysis. These help to illustrate the nature of the relationship between predictors in the model and offending, and where the risk of high-harm offending or any offending is greatest. Functionally, PDPs indicate the contribution of the variable to the probability of classification to the dependent variable (ie high-harm offence) at different points within the range of that variable. This value is relative to the MDG produced by the random forest model, controlling for the influence of other variables, and is represented on the y-axis of figures below as a value between 0 and 1. Put simply, PDPs show the relationship between individual variables and high-harm offending at different points within the range of important variables.

Limitations

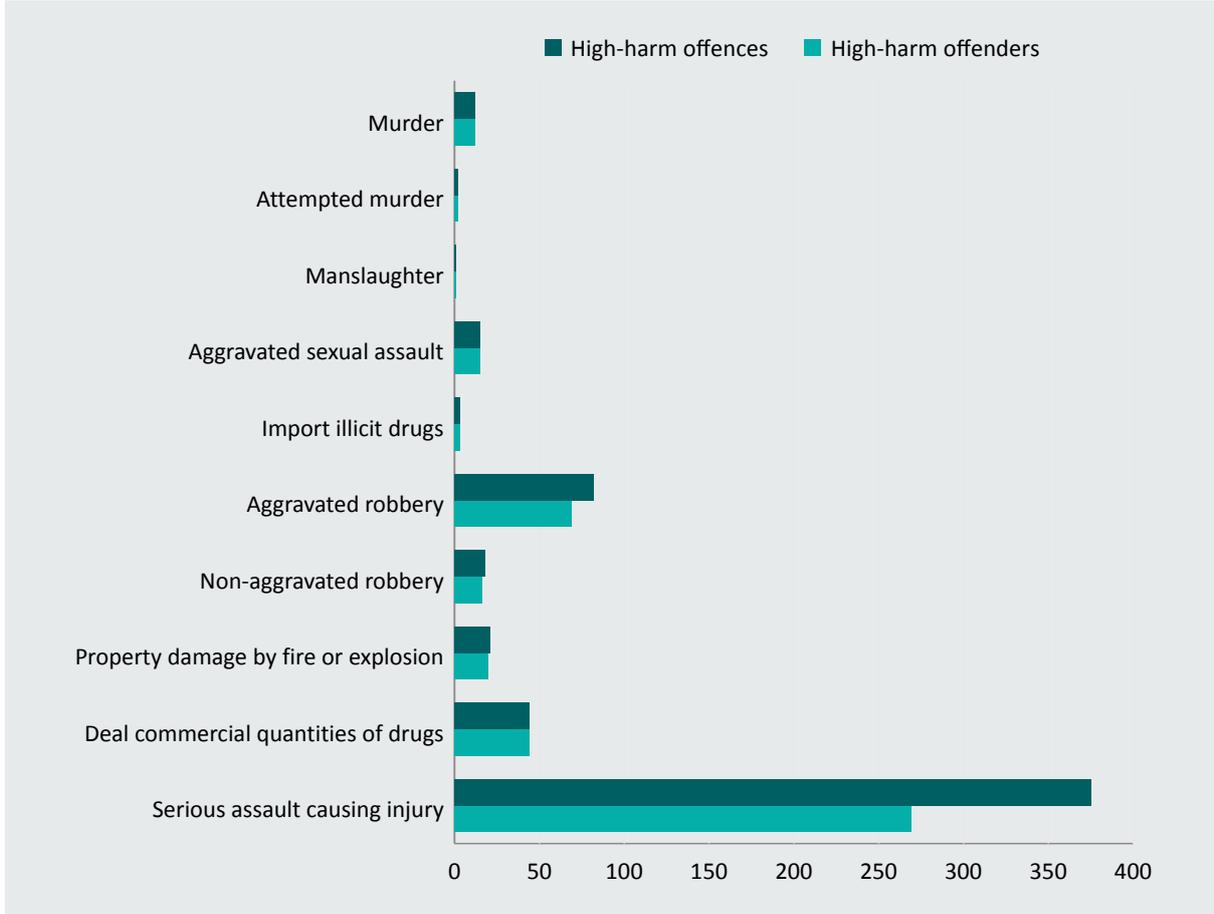
There are limitations to the model and the analysis and data that underpin it. The most obvious is the reliance on recorded offence data—namely, data on offences proceeded against by police. This model cannot account for offences that do not come to the attention of law enforcement (including prior offences or offences committed during the reference period). Conversely, there is a risk that the results are influenced by proactive enforcement and surveillance of OMCG members. However, the focus on high-harm offences likely mitigates this risk to an extent. This is reflected in differences between the models developed for high-harm offending and any offending, which suggest very different patterns of offending.

Further, though the random forest algorithm is uniquely placed to account for collinearity through bootstrap aggregation or bagging, the importance of specific variables can, on rare occasions, be impacted by collinearity (in which predictor variables are highly correlated). In addition, the model is derived from data that relate to patterns of offending among OMCG members at a particular point in time. A model which relies on historical offence data should ideally be regularly updated to reflect contemporary patterns of offending. Finally, the model was developed for OMCG members in New South Wales. The relevance of this model to other jurisdictions requires further examination. More importantly, the results presented in this paper are not applicable to other offending populations, especially those that do not have such frequent contact with the criminal justice system (see Morgan, Dowling & Voce 2020).

Results

There were 2,246 OMCG members in the final dataset, 451 of whom committed a high-harm offence during the five-year reference period. Serious assault causing injury and aggravated robbery were the two most frequent high-harm offence types (Figure 1). This suggests that, among OMCG members, high-harm offending typically constituted violent offences. However, the next most common high-harm offence was dealing commercial quantities of drugs, reflecting OMCG involvement in drug supply.

Figure 1: Number of high-harm offences committed by OMCG members in the five-year reference period (n=2,246)

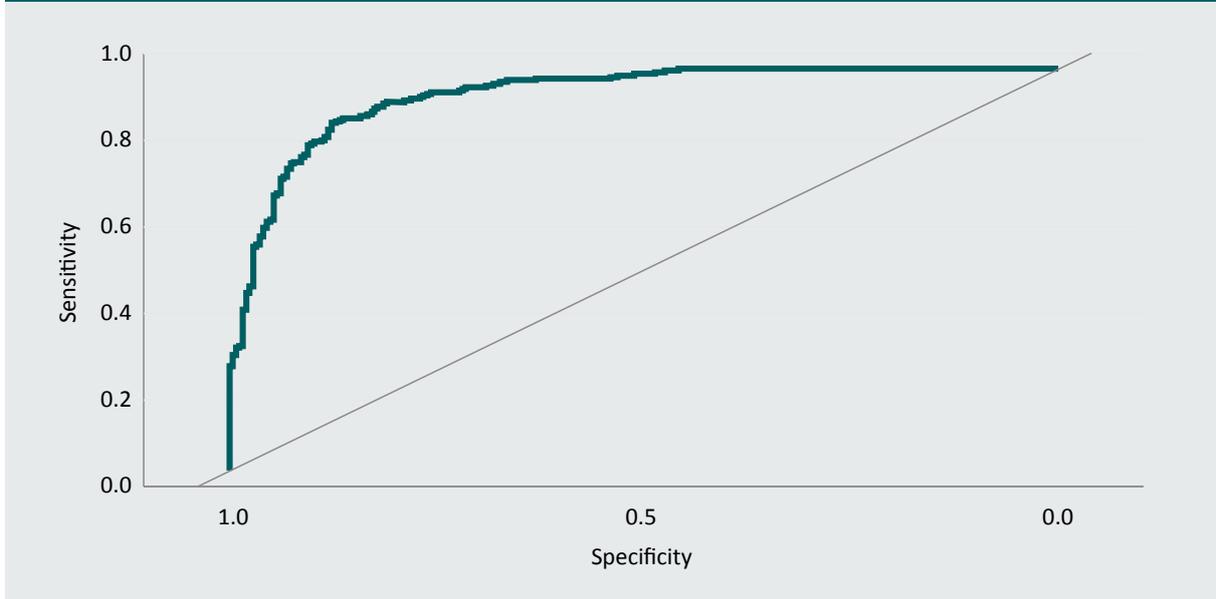


Source: NSW OMCG offending database [computer file]

Model predicting high-harm offending

The first model predicts high-harm offending among OMCG members in the five-year period between January 2015 and December 2019. This model featured a prediction accuracy of 91.4 percent, as demonstrated in Figure 2 (AUROC=0.914).

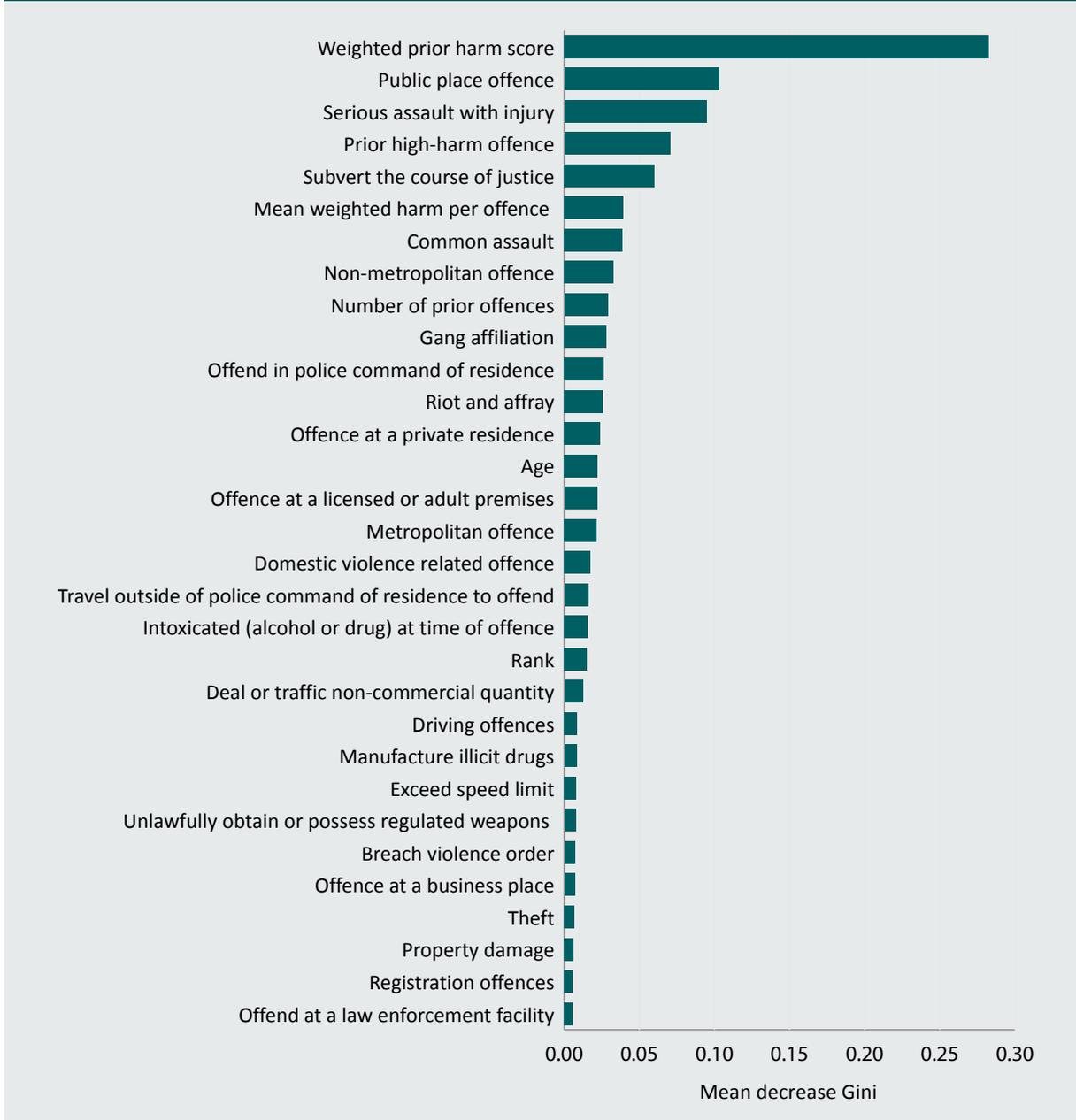
Figure 2: ROC curve for random forest model trained on high-harm offending by OMCG members



Source: NSW OMCG offending database [computer file]

Predictions about high-harm offending among OMCG members were based on the 130 variables included in this dataset, primarily relating to the criminal history of each individual. Thirty-four of these variables had no statistical interaction with the likelihood of a high-harm offence, leaving 96 variables with varying degrees of association with high-harm offending. Of these 96 variables, 31 accounted for 90.3 percent of the predictive power of the model, and the remaining 65 variables had little predictive power. Figure 3 identifies the 31 most predictive variables in this model and their importance in predicting high-harm offending.

Figure 3: Variable importance associated with high-harm offending among OMCG members



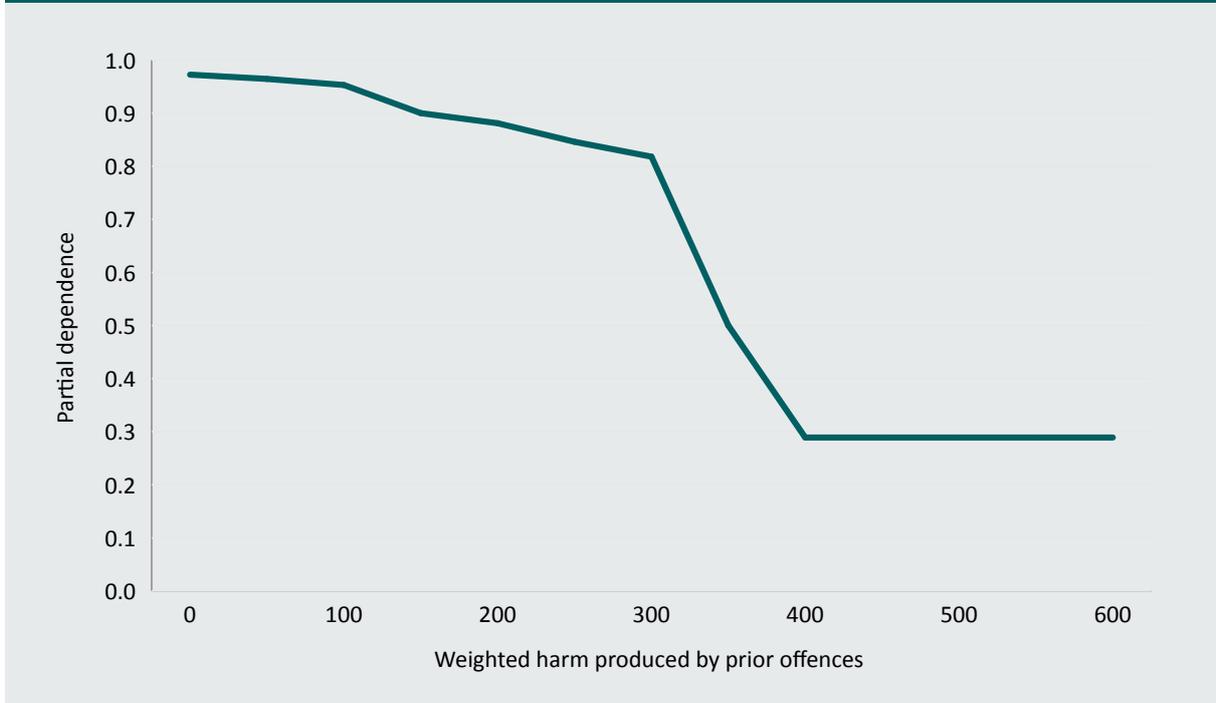
Note: The larger the mean decrease Gini coefficient, the stronger the association of that variable with high-harm offending

Source: NSW OMCG offending database [computer file]

Unsurprisingly, prior offending was an important predictor of high-harm offending. The weighted prior harm score was a particularly strong predictor of high-harm offences by OMCG members. This variable represents the total weighted harm of prior offences committed by an OMCG member, accounting for time spent in custody. The mean weighted harm per offence also emerged as important, as did prior violent offences.

Given its importance in the model, we examined the relationship between the weighted harm of prior offences and high-harm offending by inspecting the PDP (Figure 4). Risk of high-harm offending was highest where the aggregate weighted harm of prior offences was lower. In fact, though the relationship was not linear, the higher the total harm of prior offences, the lower the risk of further high-harm offending. The risk of high-harm offending was relatively stable up to an aggregate prior harm score of 300 (based on the WACHI), before a sharp decline. The relationship between the average weighted harm of prior offences and risk of high-harm offending (not shown) followed a similar pattern, as did the number of prior offences, suggesting that a history of relatively infrequent and most likely low- to medium-harm offences was more strongly associated with high-harm offending. The exception to this was serious assault resulting in injury, which—along with common assault—was an important predictor of high-harm offending. Given serious assault resulting in injury was classified as a high-harm offence, and the most frequently occurring high-harm offence, the results suggest that, as expected, prior violence was also an important predictor of future high-harm (violent) offending.

Figure 4: Risk of high-harm offence by aggregated weighted harm of prior offences



Note: Y-axis is the probability of classification to the dependent variable (ie high-harm offence) relative to the mean decrease Gini
Source: NSW OMCG offending database [computer file]

Table 1 shows where the model successfully predicted, and where it failed to identify, high-harm offending. This model was particularly good at identifying OMCG members who did not commit high-harm offences, which was important given they accounted for the majority of individuals in the test sample ($n=538$). Among OMCG members who did not commit high-harm offences during this time, the model incorrectly predicted high-harm offending in only 1.1 percent of cases. That is, the false positive rate was 1.1 percent (or a specificity of 98.9%). Conversely, the model had a false negative rate of 11.8 percent (or a sensitivity of 88.2%), meaning it incorrectly classified only 11.8 percent of members who did in fact go on to commit a high-harm offence. Overall, the misclassification rate for this model was just 3.3 percent.

Table 1: Confusion matrix for predicting high-harm offences among OMCG members

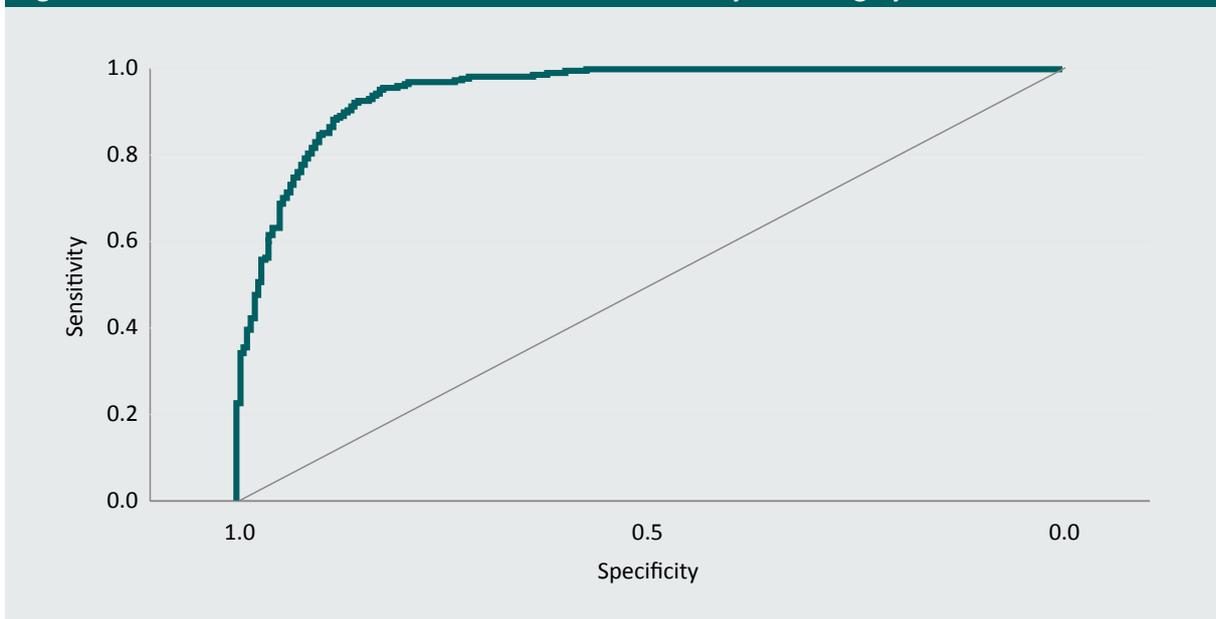
	True negative	True positive	Classification error
Predicted negative	532	16	2.91%
Predicted positive	6	120	4.76%
Classification error	1.11%	11.76%	674

Source: NSW OMCG offending database [computer file]

Model predicting any offending

Similar to the model developed for high-harm offending, the random forest analysis predicting any offending among OMCG members produced a robust model. This model featured prediction accuracy of 91.6 percent, as demonstrated by Figure 5 (AUROC=0.916).

Figure 5: ROC curve for random forest model trained on any offending by OMCG members



Source: NSW OMCG offending database [computer file]

Table 2 identifies where the model successfully predicted, and where it failed to predict, offending by OMCG members. This model was particularly good at identifying those who would commit any offence (which was the majority of OMCG members), correctly classifying 93.1 percent of members who offended within the five-year period (a false negative rate of 6.9%). This model was less accurate in predicting which individuals would not commit any offence within this time frame, with a false positive rate of 29.6 percent (or specificity of 70.4%). Overall, the misclassification rate for this model was 15.0 percent, which is higher than the misclassification rate of the high-harm offending model.

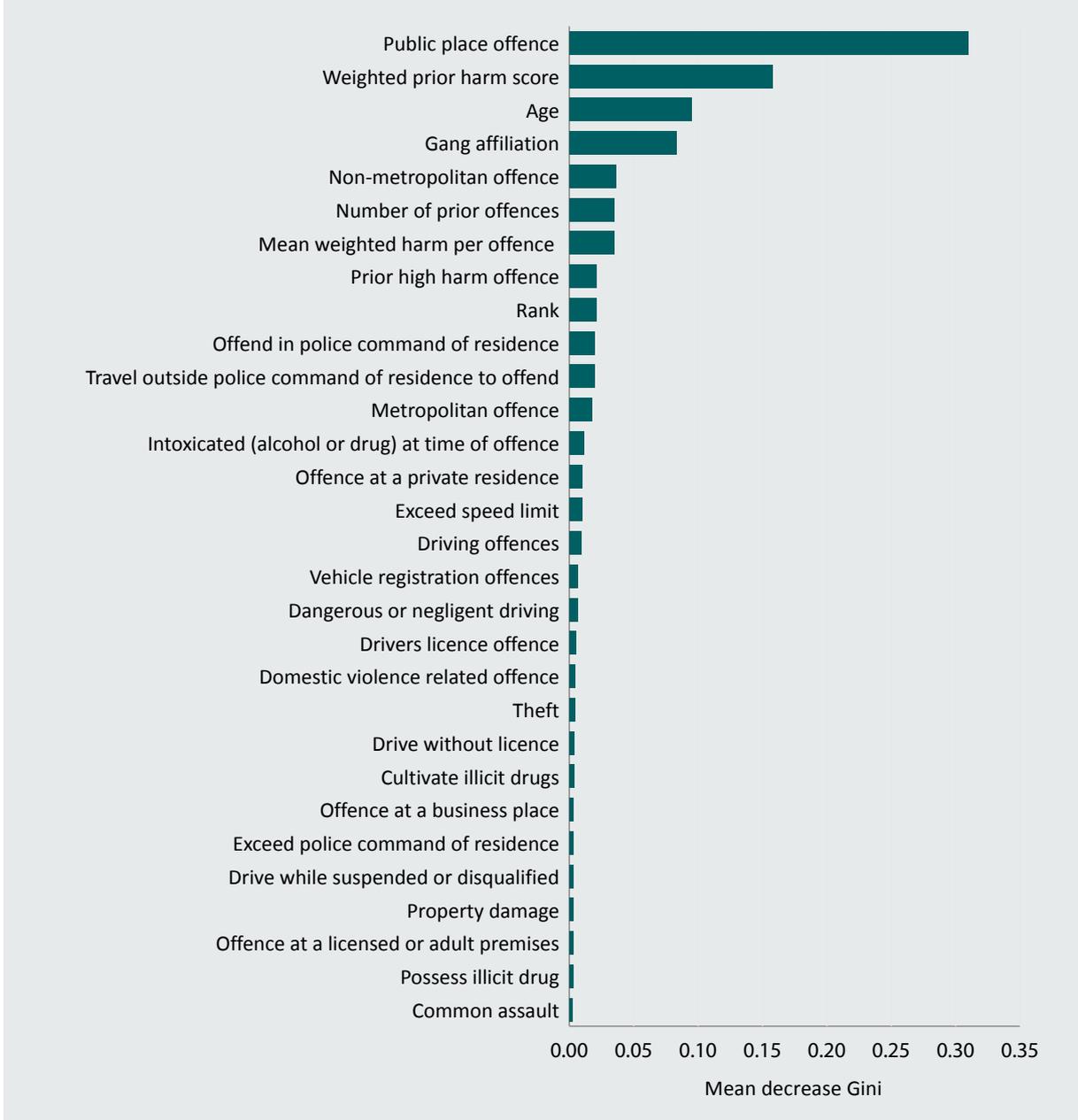
	True negative	True positive	Classification error
Predicted negative	169	30	15.07%
Predicted positive	71	404	14.95%
Classification error	29.58%	6.91%	674

Source: NSW OMCG offending database [computer file]

Additionally, the variables that were most important in making these predictions were different to those in the model developed to predict high-harm offending. While the variables themselves were similar, their relative importance changed considerably (Figure 6). For example, the most important variable was not the weighted prior harm score but the number of recorded public place offences.

The age and gang affiliation of a member were also more important in predicting any offending than they were in predicting high-harm offending. The types of prior offences were less important in this model, while the prominence of driving offences among the list of variables was noted (although still comparatively unimportant). The finding regarding prior public place offences, and the similar effect sizes of different prior offence types, suggested a possible surveillance effect, whereby members who had committed higher numbers of offences in public places were more prone to further recorded offending, possibly as a function of proactive enforcement targeting these more visible offences.

Figure 6: Variable importance associated with any offending by OMCG members



Note: The larger the mean decrease Gini coefficient, the stronger the association of that variable with any offending

Source: NSW OMCG offending database [computer file]

Discussion

Recent evidence has suggested that OMCG members are prolific in violent and criminal enterprise offending (Morgan, Dowling & Voce 2020), but that recorded offending and related harm, as well as criminal mobility, is concentrated among a relatively small proportion of individuals and chapters, and among younger OMCG members (Dowling & Morgan 2021; Morgan, Dowling & Voce 2020; Voce, Morgan & Dowling 2021). Further, there are considerable differences between clubs. Some clubs have a much higher proportion than others of patched members—office bearing and not—with a history of serious offending, and some clubs have none at all (Morgan, Dowling & Voce 2020; Morgan & Payne 2021). These findings are not unique to Australian OMCGs (see von Lampe & Blokland 2020).

Given the high proportion of OMCG members involved in crime, including serious crime, predicting which members will commit high-harm offences offers significant potential benefits in terms of disruption. The model presented in this report was able to predict high-harm offending among OMCG members with a high degree of accuracy. Using custodial data to produce a weighted harm score for prior offending was important in the predictive accuracy of this model, as was the range of variables derived from these data for analysis. These findings suggest that high-harm offending among OMCG members was relatively predictable. More specifically, it was possible to predict high-harm offending using the police, custody and gangs data available for this study.

Further, the results suggest that the risk of high-harm offending by OMCG members was related to a history of repeated low-level or moderately harmful offences. Though it appears prior violence was still an important predictor of future high-harm offending (which includes further violence), the findings suggest that the model was relatively successful at identifying individuals prior to the point at which they escalated to more serious offences. It is possible that the findings here are influenced by police efforts targeting OMCG members. Once an offender has been detected for a high-harm offence, the tactics used by law enforcement may limit the likelihood of repeated high-harm offending. That is, OMCG members who commit high-harm offences become known to law enforcement, are subject to higher levels of surveillance and, as a result, have less opportunity to commit further high-harm offences. Nevertheless, the ability to accurately predict which OMCG members are most likely to commit future high-harm offences is an important finding.

There were two reasons for focusing this research on high-harm offending. First, the model is designed to help inform efforts to target and disrupt the offences—and OMCG members—that are most harmful to the community. These findings may be used to develop more accurate identification and disruption techniques for policing high-risk members of OMCGs. This model could be used by intelligence analysts to tailor disruption efforts and resource distribution among specialist units responsible for policing OMCGs. There are potential benefits in efficiently targeting the small proportion of OMCG members responsible for the most harmful crimes, but targeted strategies are also likely to be more effective. This paper does not attempt to recommend disruption activities that may be suitable, but they may include efforts to encourage disaffiliation from gangs (see Boland et al. 2021), or enforcement or regulatory measures designed to reduce the opportunity for further offending.

The second reason for focusing on high-harm offending was to address potential criticisms of applying machine learning techniques to OMCG members or high-risk offenders more generally. There are important ethical implications to consider when applying risk assessment (McSherry 2013) and machine learning specifically (McKay 2019) in a real-world context. Both models had low misclassification rates and demonstrate the value of machine learning models—despite their relative infancy in criminology—in minimising the risk of false positives and false negatives, both of which have important ramifications in a criminal justice context. However, focusing on high-harm offending improved the accuracy of the model and, most importantly from an ethical perspective, reduced the rate of false positives.

Of course, risk assessment among groups of offenders is a complex task and, while this study produced a robust model, it should not be considered an absolute measure of whether an individual will or will not commit a high-harm offence. But, given this represents a smaller group of individuals than OMCG members who committed any offence over the five-year reference period, and high-harm offending by definition represents those offences that cause the greatest harm to victims and the wider community, this research points to the value of using machine learning methods and focusing on high-harm offending to make policing of OMCGs more efficient and effective.

References

URLs correct as at November 2021

Albanese JS 2008. Risk assessment in organized crime: Developing a market and product-based model to determine threat levels. *Journal of Contemporary Criminal Justice* 24(3): 263–273. <https://doi.org/10.1177/1043986208318225>

Australian Federal Police 2021. *AFP-led Operation Ironside smashes organised crime*. Media release, 8 June. <https://www.afp.gov.au/news-media/media-releases/afp-led-operation-ironside-smashes-organised-crime>

Barnes GC & Hyatt JM 2012. *Classifying adult probationers by forecasting future offending*. <https://nij.ojp.gov/library/publications/classifying-adult-probationers-forecasting-future-offending>

Bennett Moses L & Chan J 2018. Algorithmic prediction in policing: Assumptions, evaluation, and accountability. *Policing and Society* 7: 806–822. <https://doi.org/10.1080/10439463.2016.1253695>

Berk R 2021. Artificial intelligence, predictive policing, and risk assessment for law enforcement. *Annual Review of Criminology* 4: 209–237. <https://doi.org/10.1146/annurev-criminol-051520-012342>

Berk R 2019. *Machine learning risk assessments in criminal justice settings*. New York: Springer. <https://doi.org/10.1007/978-3-030-02272-3>

Berk R 2013. Algorithmic criminology. *Security Informatics* 2(1): 1–14. <https://doi.org/10.1186/2190-8532-2-5>

Berk R & Bleich J 2014. Forecasts of violence to inform sentencing decisions. *Journal of Quantitative Criminology* 31(1): 79–96. <https://doi.org/10.1007/s10940-013-9195-0>

- Berk R, Heidari H, Jabbari S, Kearns M & Roth A 2018. Fairness in criminal justice risk assessment: The state of the art. *Sociological Methods & Research* 50(1): 3–44. <https://doi.org/10.1177/0049124118782533>
- Berk R, Sherman L, Barnes G, Kurtz E & Ahlman L 2009. Forecasting murder within a population of probationers and parolees: A high stakes application of statistical learning. *Statistics in Society* 172(1): 191–211. <https://doi.org/10.1111/j.1467-985X.2008.00556.x>
- Berk R, Sorenson SB & Barnes G 2016. Forecasting domestic violence: A machine learning approach to help inform arraignment decisions. *Journal of Empirical Legal Studies* 13(1): 94–115. <https://doi.org/10.1111/jels.12098>
- Bjørge T 2019. Preventing organised crime originating from outlaw motorcycle clubs. *Trends in Organized Crime* 22: 84–122. <https://doi.org/10.1007/s12117-017-9322-7>
- Blokland A, van Hout L, van der Leest W & Soudijn M 2019. Not your average biker: Criminal careers of members of Dutch outlaw motorcycle gangs. *Trends in Organized Crime* 22: 10–33. <https://doi.org/10.1007/s12117-017-9303-x>
- Boland D et al. 2021. Effects of outlaw motorcycle gang membership and the support needs of former members. *Trends & issues in crime and criminal justice* no. 614. Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/ti78078>
- Coglianesi C & Lehr D 2017. Regulating by robot: Administrative decision making in the machine learning era. *Georgetown Law Journal* 105: 1147–1223. <https://ssrn.com/abstract=2928293>
- Couronné R, Probst P & Boulesteix AL 2018. Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinformatics* 19(270): 1–14. <https://doi.org/10.1186/s12859-018-2264-5>
- Cubitt TIC, Wooden KR & Roberts KA 2020. A machine learning analysis of serious misconduct among Australian police. *Crime Science* 9(22): 1–14. <https://doi.org/10.1186/s40163-020-00133-6>
- Dowling C et al. 2021. The changing culture of outlaw motorcycle gangs in Australia. *Trends & issues in crime and criminal justice* no. 615. Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/ti78054>
- Dowling C & Morgan A 2021. Criminal mobility of outlaw motorcycle gangs in Australia. *Trends & issues in crime and criminal justice* no. 619. Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/ti04992>
- Grogger J, Gupta S, Ivandic R & Kirchmaier T 2021. Comparing conventional and machine-learning approaches to risk assessment in domestic abuse cases. *Journal of Empirical Legal Studies* 18(1): 90–130. <https://doi.org/10.1111/jels.12276>
- Hong H, Xiaoling G & Hua Y 2016. Variable selection using mean decrease accuracy and mean decrease Gini based on random forest. Paper presented at the 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS). <https://doi.org/10.1109/ICSESS.2016.7883053>

- House PD & Neyroud PW 2018. Developing a crime harm index for Western Australia: The WACHI. *Cambridge Journal of Evidence-Based Policing* 2(1–2): 70–94. <https://doi.org/10.1007/s41887-018-0022-6>
- Hughes CE, Chalmers J & Bright DA 2020. Exploring interrelationships between high-level drug trafficking and other serious and organised crime: An Australian study. *Global Crime* 21(1): 28–50. <https://doi.org/10.1080/17440572.2019.1615895>
- Hyndman RJ & Anthanasopoulos G 2014. *Forecasting: Principles and practice*. Melbourne: Monash University
- Klement C 2016. Outlaw biker affiliations and criminal involvement. *European Journal of Criminology* 13(4): 453–72. <https://doi.org/10.1177/1477370815626460>
- Lauchs M, Bain A & Bell P 2015. *Outlaw motorcycle gangs: A theoretical perspective*. London: Palgrave Macmillan. <https://doi.org/10.1057/9781137456298>
- McKay C 2019. Predicting risk in criminal justice procedure: Actuarial tools, algorithms, AI and judicial decision-making. *Current Issues in Criminal Justice* 32(1): 22–39. <https://doi.org/10.1080/10345329.2019.1658694>
- McSherry B 2013. Throwing away the key: The ethics of risk assessment for preventative detention schemes. *Psychiatry, Psychology and Law* 21(5): 779–790. <https://www.doi.org/10.1080/13218719.2014.893551>
- Monterosso S 2018. From bikers to savvy criminals. Outlaw motorcycle gangs in Australia: Implications for legislators and law enforcement. *Crime, Law and Social Change* 69(5): 681–701. <https://doi.org/10.1007/s10611-018-9771-1>
- Morgan A, Dowling C & Voce I 2020. Australian outlaw motorcycle gang involvement in violent and organised crime. *Trends & issues in crime and criminal justice* no. 586. Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/ti04282>
- Morgan A & Payne J 2021. Organised crime and criminal careers: Findings from an Australian sample. *Trends & Issues in crime and criminal justice* no. 637. Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/ti78337>
- Ratcliffe JH & Kikuchi G 2019. Harm-focused offender triage and prioritisation: A Philadelphia case study. *Policing: An International Journal* 42(1): 59–73. <https://doi.org/10.1108/PIJPSM-08-2018-0118>
- Ratcliffe JH, Strang SJ & Taylor RB 2014. Assessing the success factors of organized crime groups: Intelligence challenges for strategic thinking. *Policing: An International Journal of Police Strategies and Management* 37(1): 206–227. <https://doi.org/10.1108/PIJPSM-03-2012-0095>
- Ridgeway G 2013. The pitfalls of prediction. *National Institute of Justice Journal* 271: 34–40. <https://nij.ojp.gov/topics/articles/pitfalls-prediction>
- Stevenson M & Doleac JL 2021. *Algorithmic risk assessment in the hands of humans*. <https://doi.org/10.2139/ssrn.3489440>

United Nations Office on Drugs and Crime (UNODC) 2010. *The globalisation of crime: A transnational organized crime threat assessment*. Vienna: UNODC. <https://www.unodc.org/unodc/en/data-and-analysis/tocta-2010.html>

Voce I, Morgan A & Dowling C 2021. Early-career offending trajectories among outlaw motorcycle gang members. *Trends & issues in crime and criminal justice* no. 625. Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/ti78030>

von Lampe K & Blokland A 2020. Outlaw motorcycle clubs and organized crime. *Crime and Justice* 49. <https://doi.org/10.1086/708926>

Zoutendijk AJ 2010. Organised crime threat assessments: A critical review. *Crime, Law and Social Change* 54: 63–86. <https://doi.org/10.1007/s10611-010-9244-7>

Dr Timothy Cubitt is a Principal Research Analyst in the Australian Institute of Criminology's Serious and Organised Crime Research Laboratory (SOCR-Lab).

Anthony Morgan is the Research Manager of the SOCR-Lab.

General editor, *Trends & issues in crime and criminal justice* series: Dr Rick Brown, Deputy Director, Australian Institute of Criminology. Note: *Trends & issues in crime and criminal justice* papers are peer reviewed. For a complete list and the full text of the papers in the *Trends & issues in crime and criminal justice* series, visit the AIC website at: aic.gov.au

ISSN 1836-2206 (Online) ISBN 978 1 922478 52 8 (Online)
<https://doi.org/10.52922/ti78528>

©Australian Institute of Criminology 2022

GPO Box 1936
Canberra ACT 2601, Australia
Tel: 02 6268 7166

Disclaimer: This research paper does not necessarily reflect the policy position of the Australian Government

aic.gov.au