



# Predicting Violence in Merseyside: a Network-Based Approach Using No Demographic Information

Paolo Campana<sup>1</sup> · Andrea Giovannetti<sup>2</sup>

Published online: 6 November 2020  
© The Author(s) 2020

## Abstract

**Purpose** We explore how we can best predict violent attacks with injury using a limited set of information on (a) previous violence, (b) previous knife and weapon carrying, and (c) violence-related behaviour of known associates, without analysing any demographic characteristics.

**Data** Our initial data set consists of 63,022 individuals involved in 375,599 events that police recorded in Merseyside (UK) from 1 January 2015 to 18 October 2018.

**Methods** We split our data into two periods: T1 (initial 2 years) and T2 (the remaining period). We predict “violence with injury” at time T2 as defined by Merseyside Police using the following individual-level predictors at time T1: violence with injury; involvement in a knife incident and involvement in a weapon incident. Furthermore, we relied on social network analysis to reconstruct the network of associates at time T1 (co-offending network) for those individuals who have committed violence at T2, and built three additional network-based predictors (associates’ violence; associates’ knife incident; associates’ weapon incident). Finally, we tackled the issue of predicting violence (a) through a series of robust logistic regression models using a bootstrapping method and (b) through a specificity/sensitivity analysis.

**Findings** We found that 7720 individuals committed violence with injury at T2. Of those, 2004 were also present at T1 (27.7%) and co-offended with a total of 7202 individuals. Regression models suggest that previous violence at time T1 is the strongest predictor of future violence (with an increase in odds never smaller than 123%), knife incidents and weapon incidents at the individual level have some predictive power (but only when no information on previous violence is considered), and the behaviour of one’s associates matters. Prior association with a violent individual and prior association with a knife-flagged individual were the two strongest network predictors, with a slightly stronger effect for knife

---

✉ Paolo Campana  
pc524@cam.ac.uk

<sup>1</sup> Institute of Criminology, University of Cambridge, Cambridge, UK

<sup>2</sup> Economics Department, University of Technology Sydney, Sydney, Australia

flags. The best performing regressors are (a) individual past violence (36% of future violence cases correctly identified); (b) associates' past violence (25%); and (c) associates' knife involvement (14%). All regressors are characterised by a very high level of specificity in predicting who will not commit violence (80% or more).

**Conclusions** Network-based indicators *add* to the explanation of future violence, especially prior association with a knife-flagged individual and association with a violent individual. Information about the knife involvement of associates appears to be more informative than a subject's own prior knife involvement.

**Keywords** Violence · Knife incidents · Gun incidents · Co-offending networks · Violence prediction

## Introduction

Violent attacks are one of the rising challenges to the security of urban environments and the well-being of local communities in the UK and elsewhere. For example, England and Wales recorded 43,500 offences involving a sharp instrument (often a knife or a blade) in the year ending March 2019—the highest number since the year 2010/11 (Allen et al. 2019: 4). This is a 42% increase on 2010/11 and an 82% increase on the low point of 2013/14 (Allen et al. 2019: 7).

Such an increase in knife crime in England and Wales is often described as a “knife epidemic”—a label that echoes the “gunshot epidemics” of Boston and Chicago (NYTimes 2015; Benjamin 2015, cf. Koch 2016). Speaking of violence as an epidemic often helps grab media and political attention, yet it risks remaining nothing more than a headline-catching slogan. In this paper, we focus on an often-neglected aspect of the epidemic discourse, i.e. the network perspective underpinning it, and explore its contribution to our understanding of violence.

Epidemics are based on the notion of connectivity, which makes them an inherently relational phenomenon (Jackson 2019, Ch. 3). In the case of contagious diseases, it is the network connecting humans that creates the infrastructure allowing a disease to spread. Public health approaches to reduce violence have a relational element built in—although this may not always be formally spelled out. In this paper, we investigate the role of relationships in generating violence using data from Merseyside Police (Liverpool, UK). Our goal is to offer an assessment of the accuracy of network-based indicators in violence prediction models.

## A Network-Based Approach to Study Violence

In this paper, we define a network as a set of actors (individuals) and their relations among them (following Wasserman and Faust 1994: 20; Campana 2016). The idea of using network analysis to study illegal activities, including gang and organised crime groups, is gaining traction within criminology (see Carrington 2011; Campana and Varese 2012; Morselli 2014; Faust and Tita 2019 and Campana and Varese 2020 for an overview of the studies, challenges, and pitfalls). Papachristos (2009) and Bichler et al.

(2019) have promoted the idea of applying a network perspective to study violence, and termed such perspective ‘networked violence’.

A limited number of works have so far relied on network analysis to study violence, starting from the pioneering work by Kennedy et al. (1997) on gang-level conflicts in Boston. A handful of US-based studies have followed suit, mostly carried out by Papachristos and his colleagues (e.g. Papachristos 2009; Papachristos et al. 2012, 2013, 2015a and 2015b), Tita and Radil (Radil et al. 2010; Tita and Radil 2011), and Randle and Bichler (2017).<sup>1</sup>

Papachristos et al. (2013) looked at Chicago and Boston, and found that prior conflicts have an effect on future gang-related violence. Using data from Chicago, Papachristos et al. (2015a) found that gunshot victimisation increases with exposure to violence in one’s social network. Using data from Newark, NJ, Papachristos et al. (2015b) concluded that the risk of being a victim of a fatal or non-fatal gunshot increases the closer one is to a gang-member in a co-offending network. Green et al. (2017) modelled the idea that a person becomes exposed to gun violence through social interactions with individuals who had been previously subject to gun violence. Their central hypothesis—for which they found some level of support—is that “when someone in your network becomes a subject of gun violence, your risk of becoming a subject of gun violence temporarily increases” (Green et al. 2017: 327). In this study, they used police co-offending data as a proxy for pre-existing social ties (the study site was, once more, Chicago).

The relatively few studies conducted so far have shown the benefit of applying a network approach to study violence. Yet, they are limited in their focus on (a) gang-related events (with the exception of Papachristos et al., 2015a) and (b) gun shootings. Furthermore, they have been conducted in the US context and in cities with an acute violence problem.

In this paper, we advance this line of work in three different ways. First, we move beyond gang and organised crime-related violent events. Second, we move beyond gun-related violence alone and consider all instances of violence with injury. Third, we offer the first analysis of its kind outside the USA and in a setting not characterised by high intensity of generalised violence (Merseyside, UK), where homicide is far less frequent per capita than in US cities like Chicago.

Our study also builds on the pioneering work by Massey et al. (2019) on forecasting homicide locations based on known locations of non-fatal knife assaults in the preceding year. Massey et al. (2019) offer a number of crucial contributions, including shifting the focus to knife-enabled assaults and showing the importance of non-fatal assaults in predicting future fatal homicides. In our study, instead of a spatial approach, we take an individual-based network approach. Future developments might integrate both a spatial and a network perspective.

## Research Questions

Our main question is: how can we best predict violent attacks with injury at Time 2 (T2) using information from Time 1 (T1) on (a) previous violence, (b) previous knife and weapon carrying, and (c) the violence-related behaviour of known associates?

We will explore six sub-questions, in which “violence” denotes attacks with injury:

<sup>1</sup> Campana and Varese (2013) have modelled violence within criminal organisations.

1. What is the predictive value of violence at T1 on committing violence at T2?
2. What is the predictive value of carrying a knife at T1 on committing violence at T2?
3. What is the predictive value of carrying a weapon at T1 on committing violence at T2?
4. What is the predictive value of associating with someone who has committed violence at T1 on committing violence at T2?
5. What is the predictive value of associating with someone who has carried a knife at T1 on committing violence at T2?
6. What is the predictive value of associating with someone who has carried a weapon at T1 on committing violence at T2?

## Data

Our data consist of 63,022 individuals involved in 375,599 police-recorded events in Merseyside and spanning the period from 1 January 2015 to 18 October 2018. The data were collected and made available to us in a fully anonymised form by Merseyside Police, a territorial police force responsible for policing a large area in the northwest of England (UK). Their jurisdiction covers a population of around 1.5 million people, of which roughly half a million reside in its main city, Liverpool.

We have taken a broad approach to our analysis by including all recorded events regardless of their criminal justice outcome: this includes events in which a person was arrested, cautioned, charged, and wanted on warrant, as well as interviewed, suspected, or when no further action was taken. We have, however, excluded from our analysis events classified by the police as domestic incidents and sexual offences.

## Methods

To study the emergence of violence, we split our dataset into two periods: T1 and T2. The first period (T1) runs from 1 January 2015 to 31 December 2016 (2 years). The second period (T2) runs from 1 January 2017 to 18 October 2018 (date of the data extraction, with almost 21 months in T2).

### Model Specification: Dependent and Independent Variables

In this work, we interpret violence (our dependent variable) as “violence with injury” as defined by Merseyside Police. The main offences included in this category are as follows: murder and attempted murder; assault occasioning actual bodily harm (- Section 47); wounding with intent to do grievous bodily harm (Section 18); inflicting grievous bodily harm; malicious wounding; racially or religiously aggravated actual bodily harm; allowing a dog to be dangerously out of control injuring any person (both in a public space or in a non-public space).

Our independent variables are a set of binary variables (0/1) defined as follows:

- Violence T1: whether the individual has committed violence with injury at time T1;

- Weapon T1: whether the individual has been flagged by the police for a weapon incident at time T1. (To this end, we relied on the flag ‘gun\_involved’ included in the dataset.);
- Knife T1: whether the individual has been flagged by the police for a knife incident at time T1. (To this end, we relied on the flag ‘knife\_involved’ included in the dataset.)

This set of variables captures the behaviour of a single individual *i* but does not take into consideration the behaviour of *i*’s associates. In this work, we expanded our analysis to encompass the network of associates of *i*. We consider as an “associate” any person with whom individual *i* has committed a crime (of any type minus domestic violence and sexual offences) during time T1. In other words, we built the network of co-offenders at T1 using all the crime types included in the datasets. A co-offender is a person who has shared a criminal event with individual *i* (i.e. both are associated to the same police record of one or more specific events; we remind the reader that police records are event-based: see Campana and Varese 2020 for a discussion on the structure of police-recorded events). This is a loose interpretation of “criminal association” as it does not require co-membership in the same organised crime group/gang nor repeated criminal interactions. Such a broad view is the best proxy for *i*’s pre-existing social ties given the evidence available—in line with Green et al. (2017).

The network element is captured by three binary variables (0/1):

- Violence associates T1: whether any of individual *i*’s associates has committed violence with injury at time T1, with or without *i*;
- Weapon associates T1: whether any of individual *i*’s associates has been flagged by the police for a weapon incident at time T1, with or without *i*. (To this end, we relied on the flag ‘gun\_involved’ already included in the dataset);
- Knife associates T1: whether any of individual *i*’s associates has been flagged by the police for a knife incident at time T1, with or without *i*. (To this end, we relied on the flag ‘knife\_involved’ included in the dataset).

We first identified the individuals who have committed violence at time T2. To create the co-offending network, we dropped those individuals who were only present at T2, and focussed on individuals who had been present at both T1 and T2. The network of their associates was then built based on offences recorded at T1. Our model is intentionally simplified as it does not require information on socio-demographic characteristics of the individuals or spatial information on the individuals and/or criminal events. Future development of the model can integrate such additional perspectives.

In this work, we tackle the issue of predicting violence using two different approaches: (a) a series of “bootstrapping” logistic regressions and (b) a “sensitivity/specificity” analysis.

### Bootstrapping Logistic Regression Method

We ran seven logistic regression models: four are presented in the main text (Models 1–4) and three in the [Appendix](#) (Models A1–A3). To provide robust

estimations and confirm the validity of our models, we used a bootstrapping method—a procedure that tests the consistency of results across different samples from the same data set. We statistically tested our estimates using the variance-covariance matrix obtained from a bootstrapped model ( $N = 500$  repetitions) as opposed to the standard variance-covariance matrix. The small divergence between estimates allows us to rule out extreme skewness in categorical data distribution. The bootstrapping method also allows us to obtain tighter confidence intervals for the estimates, when it is not possible to rule out potential estimation bias, a technical challenge to assessing the validity of the predictors.

### Sensitivity/Specificity Analysis

Sensitivity and specificity analyses are carried out in medicine to understand the performance of a medical test in identifying individuals who carry a specific disease. *Sensitivity* indicates the extent to which, on a scale from 0 to 1, actual positive cases are correctly classified: high sensitivity points to the ability of a test to correctly identify true positives and minimise false negatives.<sup>2</sup> *Specificity* indicates, also on a scale from 0 to 1, the extent to which negative occurrences are correctly classified: high specificity points to the ability of a test to correctly identify actual negatives and minimise false positives.<sup>3</sup> In the cases we study, high specificity means that the test is most often correct when it predicts that  $i$  will not commit violence at T2.

### Findings

We found that a population of 7720 individuals committed violence (with injury) during T2. Of those, 2004 are also present in the population identified at T1 (27.7%). In this paper, we focus on this latter set of individuals (*target* individuals) and reconstruct their network of associates at time T1: 497 individuals (out of the 2004) have an associate who has committed violence at T1 (24.8%); 279 have an associate who has carried a knife at T1 (13.9%); and 119 have an associate who has carried a weapon at T1 (23.9%). Of those target individuals, 712 have also committed violence at T1 (35.5%). Figure 1 shows the network of the target individuals (large dark grey dots) and their associates (small light grey dots).

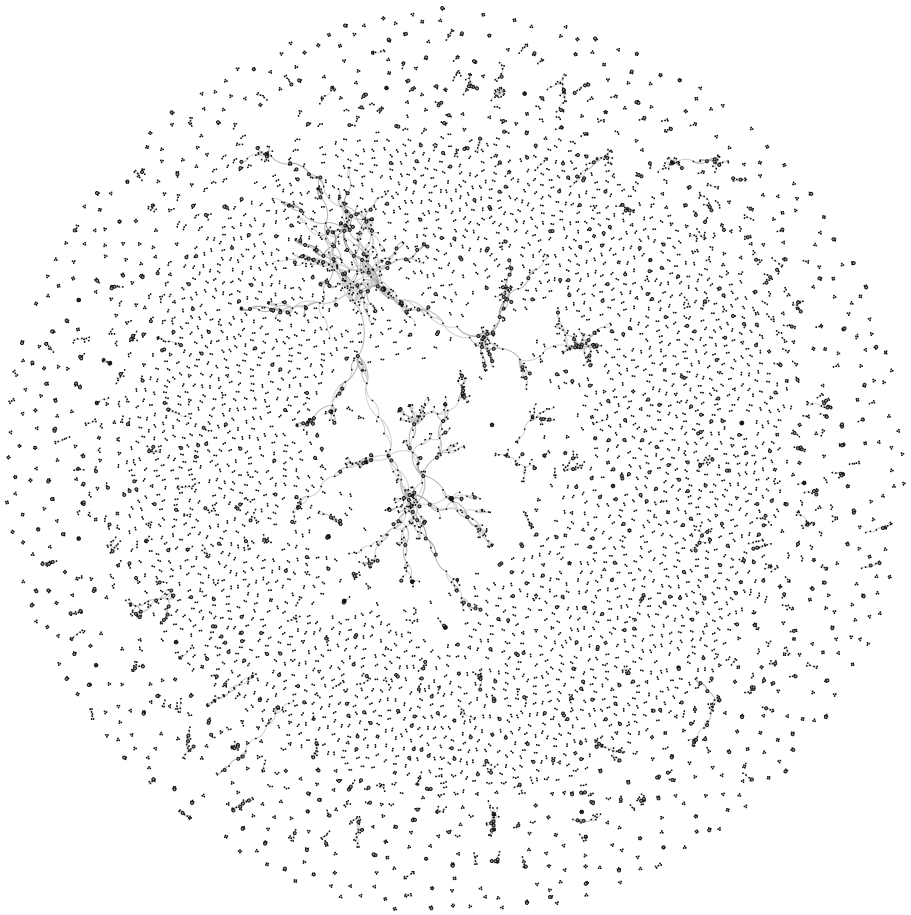
### Estimating the Determinants of Violence Through Logistic Regression Models

Next, we look at the determinants of violence at time T2 through a series of logistic regression models capturing behaviour at time T1.<sup>4</sup> We start by considering only the violence-related behaviour of target individuals (Table 1).

<sup>2</sup> The formula for sensitivity is as follows: true positive/(true positive + false negative).

<sup>3</sup> The formula for specificity is as follows: true negative/(true negative + false positive).

<sup>4</sup> A note on how to interpret logistic coefficients: if the coefficient *violence T1* is equal to  $-x$ , it means that a decrease of  $x$  is expected in the log odds of *violence T2* with a one-unit increase in *violence T1* (with other variables kept constant); if positive, we expect an increase in the log odds. Coefficients can be converted into odds ratios by exponentiating them. The odds of an event happening is defined as the probability that the event occurs divided by the probability that the event does not occur. Odds ratio is the ratio of two odds.



**Fig. 1** The network of target individuals and their associates. Note: large dark grey = target individuals; small light grey = associates (co-offenders)

Model 1 is akin to a baseline model: we know from the literature that previous violence tends to be associated with higher chances of committing violence in subsequent periods, and we found support for this relationship also in our data. The odds of observing violence at T2 for an individual who has committed violence at T1 are equal to 2.37. In terms of percentage changes, the odds for an individual who has committed violence at T1 to also commit violence at T2 are 137.5% higher than for those who have not committed violence at T1 (we remind the reader that in our models we are comparing against non-violent offenders, not against the general population).

Model 2 adds the effect of having been flagged at time T1 for (a) a weapon-involved incident and (b) a knife-involved incident. Having committed previous violence remains the strongest predictor; both knife and weapon have a positive effect, but only having been flagged for a weapon-involved incident is statistically significant. In terms of percentage changes, a weapon incident at T1 increases the odds of violence at T2 by 43%. (Model A1 in the Appendix Table 3 looks at the effect of knife flagging and weapon flagging when no information on previous violence is considered; in sum, both

**Table 1** Predicting violence in Merseyside: target individuals

	Log odds	Std. error	Probability	Odds ratio
<b>Model 1</b>				
Violence T1	0.865	0.055	0.000***	2.37
Constant	- 1.509	0.030	0.000***	0.22
<i>Pseudo R2</i>	0.0241			
<i>Log likelihood</i>	- 4707.10			
<b>Model 2</b>				
Violence T1	0.842	0.053	0.000***	2.32
Knife T1	0.112	0.107	0.295	1.12
Weapon T1	0.361	0.183	0.049*	1.43
<i>Pseudo R2</i>	0.0247			
<i>Log likelihood</i>	- 4704.24			

Note: Dependent variable = violence at T2. Bootstrap estimations with 500 replications. N obs. = 9206. Significance: \*\*\* < 0.001; \*\* < 0.01; \* < 0.05; + < 0.1

have a positive effect and show a percentage increase in the odds of violence of 87% for weapon and 43% for knife).

Next, we assess the effect of network-based indicators (Table 2).

**Table 2** Predicting violence in Merseyside: target individuals and their associates

	Log odds	Std. error	Probability	Odds ratio
<b>Model 3</b>				
Violence T1	0.804	0.059	0.000***	2.23
Violence associates T1	0.248	0.065	0.000***	1.28
Constant	- 1.544	0.032	0.000***	
<i>Pseudo R2</i>	0.0257			
<i>Log likelihood</i>	- 4699.68			
<b>Model 4</b>				
Violence T1	0.804	0.057	0.000***	2.23
Knife T1	- 0.004	0.120	0.969	0.99
Weapon T1	0.200	0.210	0.342	1.22
Violence associates T1	0.150	0.082	0.068+	1.16
Knife associates T1	0.18	0.104	0.083+	1.20
Weapon associates T1	0.021	0.154	0.891	1.02
Constant	- 1.54	0.033	0.000	0.21
<i>Pseudo R2</i>	0.0263			
<i>Log likelihood</i>	- 4696.73			

Note: Dependent variable = violence at T2. Bootstrap estimations with 500 replications. N obs. = 9206. Significance: \*\*\* < 0.001; \*\* < 0.01; \* < 0.05; + < 0.1



Model 3 shows that the behaviour of an offender's associates *does* matter in explaining future violence. This holds true also when controlling for the past violent behaviour of a target individual. In other words, the information about the associates *adds* to our understanding of the phenomenon. Prior association with a violent individual at time T1 increases the odds of committing violence at time T2 by 28%. (In Model A2 in the Appendix Table 4, we estimate the model considering only the information on the associates' behaviour without any information on the past violent behaviour of a target individual: the positive effect holds true, and points to an increase in the odds of committing violence at T2 of 61%).

In Model 4, we jointly consider the full set of indicators: individual-based and network-based. Previous violence by a target individual remains a strong predictor of future violence: there is a +124% increase in the odds of committing violence at T2 compared with an offender who has not committed violence at T1. Secondly, the behaviour of the associates continues to matter: prior association with an individual who has committed violence at T1 increases the odds of committing violence at T2 by 16%; prior association with an individual who has been flagged for knife incident increases the odds by slightly more (+20%).

We do not find, however, any effect in Model 4 for prior association with an individual flagged for a weapon incident. Finally, when adding the associates' behaviour to the model, the fact that a target individual has been flagged for a knife incident loses predictive power—showing no effect in Model 4. Weapon flagging at time T1 still shows a positive effect (+22% increase in odds), but loses its statistical significance. This is partially due to the relative small number of weapon-related incidents in the datasets (we return on this point in the sensitivity/specificity analysis below).<sup>5</sup>

### Sensitivity and Specificity Analysis of the Indicators

Next, we contextualise the relevance of each predictor under consideration through a sensitivity/specificity analysis. We ask to what extent the fact that an individual is flagged or *not* with a given regressor at T1 can help a practitioner assess whether such individual will commit or *not commit* violence at T2.

We remind the reader that sensitivity indicates the extent to which actual positive cases are correctly classified: true positives are correctly identified and false negatives minimised. Specificity indicates the extent to which negative occurrences are correctly classified: actual negatives are identified and false positives minimised. In our context, a true positive is when a knife flag at T1 is associated with violence at T2; conversely, we expect no knife flag at T1 to be associated with no violence at T2 (true negative). A false positive is when a knife flag at T1 is associated with an individual who will not commit violence at T2; a false negative is when no knife flag at T1 is associated with violence at T2.

As there is normally a trade-off between the two measures, we need to look at them jointly to draw conclusions on the strength of each indicator relative to the predicted behaviour of individuals. Figure 2 reports the results of both sensitivity and specificity for the six predictors under consideration.

<sup>5</sup> In Model A3 in the Appendix, Table 4 we estimate the explanatory power of prior association with a weapon-flagged and a knife-flagged individual when no other information is fed into the model: knife flagging remains the stronger predictor and points to an increase in the odds of 48%.

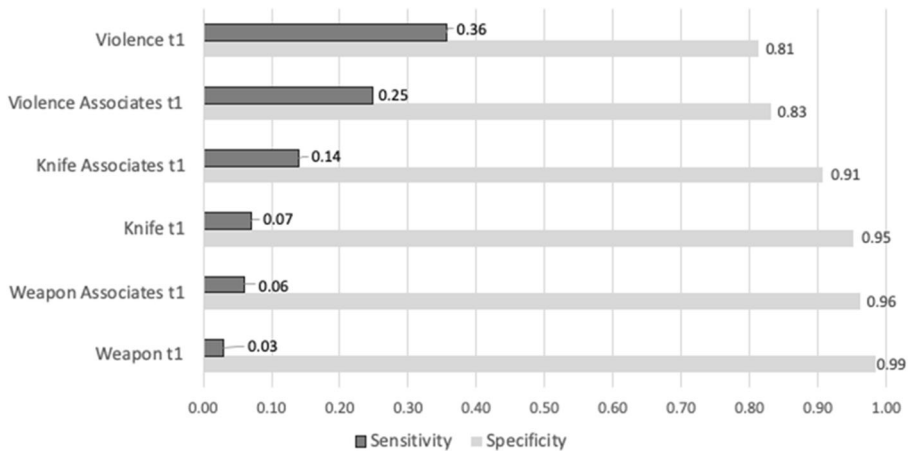


Fig. 2 Sensitivity and specificity of violence predictors

For our purposes, the key measure of interest is sensitivity as we seek to answer the following question: if we rely on, say, knife flagging at T1, what is our ability to correctly identify individuals who will commit violence at T2? To put it in another way, we are trying to minimise false negatives (e.g. no knife at T1 and violence at T2) while capturing as many true positives as possible.

The indicators show a striking heterogeneity in the degree of sensitivity. The best performer is past violence (violence T1): 36% of all individuals observed at T1 who then commit violence at T2 are flagged with violence T1. This is followed by prior association with a violent individual, which correctly identifies 25% of cases. Prior association with a knife-flagged individual correctly identifies 14% of cases. Knife flagging at T1, prior association with a weapon-flagged individual, and weapon flagging at T1 show the lowest level of sensitivity (this is in line with the results of the regression models discussed above). Weapon flagging at T1 (sensitivity level 3%) performs 91% worse than the best available regressor (violence T1). On the other hand, the sensitivity levels of network-based counterparts of these regressors are on average 98% higher.<sup>6</sup>

As for the question of who will *not* commit violence with injury, all regressors considered in our analysis record a very high level of specificity. In other words, for all regressors, individuals for which violence at T2 = 0 (i.e. they do not commit violence at T2) are in most cases correctly identified by a flag (regressor) equal to 0. Overall, the maximum deviation among indicators is just 18%. While this result is partially driven by the fact that the population of interest is denoted by a relatively small number of individuals committing violence at T2 and, similarly, a relatively small number of individuals for which the flags (regressors) are positive, it is still a reassuring dimension of police legitimacy. For police to miss so few people who will commit violence with injuries when the prediction says they will not, they can justify limiting the scope of proactive policing efforts to cases in which evidence does predict future violence.

<sup>6</sup> The very high discrepancy in sensitivity and specificity tests for weapon-flagged and knife-flagged indicators is supportive of our findings relative to the weak statistical strength of such regressors in the logit analysis.

It is also important to note that the network-based indicators still perform well on this measure vis-à-vis individual-based indicators, recording a level of specificity never lower than 83%.

In sum, the joint specificity and sensitivity analysis shows that a better understanding of an individual's co-offending network allows a practitioner to cast a more robust judgement relative to future violent behaviour. In particular, we remark that violence associates at T1 does a good job in both sensitivity and specificity tests, thus efficiently complementing the individual violence T1. For these two variables, both the presence *and the absence* of the flag are informative of future behaviour. Prior association with knife-flagged individuals performs slightly less strongly, but it can still offer some potentially useful operational guidance.

## Conclusions

In this paper, we offered an assessment of a new class of network-based indicators to predict future violence vis-à-vis individual-based indicators. We have purposely used a limited set of information on (a) past violence, (b) weapon incidents, and (c) knife incidents. We relied on two different approaches to explore the emergence of violence and assess the strength of our indicators: a series of robust logistic regression models and a sensitivity/specificity analysis. In this work, we relied on evidence from Merseyside Police spanning the period from January 2015 to October 2018.

What we did not rely on was any information about the demographic characteristics of the subjects, the communities in which they reside, or any other information besides the variables we have displayed in this report. Compared to previous models predicting high vs. low-risk of offending (e.g. Berk et al. 2009), these models are far less controversial. By stripping from the data any characteristics that could be associated with ethnic or racial bias, the model offers predictions of future behaviour based solely on data about past behaviour—of individuals and their co-offenders.

The robust logistic regression models with these limited, demographic-free data pointed to three main findings. Firstly, information on previous violence is the strongest predictor of future violence across all model specifications, with an increase in odds never smaller than 123%. Previous involvement in a weapon incident increases the odds by 43% when only individual-level information is considered, but stops being informative when network-level information is considered. Individual involvement in a knife incident has little, if any, predictive power.

Secondly, network information on one's associates matters and *adds* to our understanding of the phenomenon. Two predictors emerged as the strongest: prior association with a violent individual and prior association with an individual flagged for a knife incident. Surprisingly, the latter has a slightly stronger effect than the former (20% increase in odds vs. +16% of previous violence in the most complete model).

The importance of prior violence has been confirmed by a sensitivity/specificity analysis of the indicators. The same analysis has also shown that network-based measures may help practitioners cast a more robust judgement relative to future violent behaviour. Prior association with a violent individual and, to a lesser extent, prior association with a knife-flagged individual are identified as good performers.

**Limitations** As this is a preliminary work, there are a number of limitations. Firstly, we used ‘weapon’ and ‘knife’ flags as recorded by the police, but future works might rely on different ways to assess knife and weapon carrying. Secondly, the boundaries of what constitutes violence with injury might be redefined, moving away from the statutory police definition. Thirdly, we implemented a fixed separation between T1 and T2, but future developments might rely on moving windows or individual-based time windows. Also, this would allow to increase the number of individuals for which it is possible to build a co-offending network (in our work, this is limited to 27.7% of those who have committed violence at time T2). Finally, it is possible to expand our approach by integrating network information on associates with spatial information on where the crime events took place.

**Benefits** This approach can enable police to limit the time they invest in preventing violence by people who are already unlikely to commit violence with injuries. Like solvability factors, which predict that a past crime will not be solved, these prediction factors can guide police in what cases *not* to assign. With high levels of true negatives, police can both save scarce resources, and reassure the public that they are rarely ignoring dangerous people by using this evidence-based approach.

**Acknowledgements** We are extremely grateful to Larry Sherman for his intellectual generosity and for motivating us to carry out this study; he has also offered insightful comments on an earlier version of this paper. We would also like to express our sincere gratitude to Merseyside Police and particularly to Chris Gibson and the evidence-based team that have made this research possible. A special thank you goes to Louise Kane and Christopher Wells for extracting the data from the police records management system and preparing the original dataset. Our thanks also go to Alyssa Knisley for her comments on an earlier version of the paper.

**Funding** PC received financial support from Leverhulme Trust through the Research Grant RPG-2018-119; AG reports support from the Australian Research Council through Discovery Project DP170100429.

## Appendix

**Table 3** Predicting violence in Merseyside: targeted individuals

	Log odds	Std. error	Probability	Odds ratio
Model A1				
Knife T1	0.359	0.103	0.000***	1.43
Weapon T1	0.626	0.174	0.000***	1.87
Constant	-1.313	0.026	0.000***	0.27
<i>Pseudo R2</i>	0.0028			
<i>Log likelihood</i>	-4810.22			

Note: Dependent variable = violence at T2. Bootstrap estimations with 500 replications. N obs. = 9206. Significance: \*\*\* < 0.001; \*\* < 0.01; \* < 0.05; + < 0.1

**Table 4** Predicting violence in Merseyside: associates only

	Log odds	Std. error	Probability	Odds ratio
Model A2				
Violence associates T1	0.4796	0.058	0.000***	1.62
Constant	− 1.378	0.029	0.000***	0.25
<i>Pseudo R2</i>	0.0063			
<i>Log likelihood</i>	− 4793.41			
Model A3				
Knife associates T1	0.395	0.085	0.000***	1.48
Weapon associates T1	0.214	0.127	0.093+	1.23
Constant	− 1.335	0.025	0.000***	0.26
<i>Pseudo R2</i>	0.0038			
<i>Log likelihood</i>	− 4805.3709			

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Allen, G., Audickas, L., Loft, P., & Bellis, A. (2019). Knife crime in England and Wales. In *Briefing paper, SN4303*. London: House of Commons Library.
- 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. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(1), 191–211.
- Benjamin G. (2015). “Gun violence is an epidemic. It is time for a public health response”. *The Guardian*, December 4<sup>th</sup>.
- Bichler, G., Norris, A., Dmello, J. R., & Randle, J. (2019). The impact of civil gang injunctions on networked violence between the Bloods and the Crips. *Crime & Delinquency*, 65(7), 875–915.
- Campana, P. (2016). Explaining criminal networks: strategies and potential pitfalls. *Methodological Innovations*, 9.
- Campana, P., & Varese, F. (2012). Listening to the wire: criteria and techniques for the quantitative analysis of phone intercepts. *Trends in Organized Crime*, 15(1), 13–30.
- Campana, P., & Varese, F. (2013). Cooperation in criminal organizations: kinship and violence as credible commitments. *Rationality and Society*, 25(3), 263–289.
- Campana, P., and Varese, F. (2020). “Studying organized crime networks: data sources, boundaries and the limits of structural measures”. *Social Networks*, online first.
- Carrington, P. J. (2011). Crime and social network analysis. In J. Scott & P. J. Carrington (Eds.), *The SAGE handbook of social network analysis* (pp. 236–255). Thousand Oaks: SAGE.

- Faust, K., & Tita, G. E. (2019). Social networks and crime: pitfalls and promises for advancing the field. *Annual Review of Criminology*, 2, 99–122.
- Green, B., Horel, T., & Papachristos, A. V. (2017). Modeling contagion through social networks to explain and predict gunshot violence in Chicago, 2006 to 2014. *JAMA Internal Medicine*, 177(3), 326–333.
- Jackson, M.O. (2019). *The human network: how your social position determines your power, beliefs, and behaviors*. Vintage.
- Kennedy, D. M., Braga, A., & Piehl, A. M. (1997). The (un) known universe: mapping gangs and gang violence in Boston. In D. Weisburd & T. McEwen (Eds.), *Crime mapping and crime prevention* (pp. 219–237). Monsey: Criminal Justice Press.
- Koch, P. (2016). “Is ‘epidemic’ the right metaphor for gun violence?”. *The Atlantic*, January 4<sup>th</sup>.
- Massey, J., Sherman, L. W., & Coupe, T. (2019). Forecasting knife homicide risk from prior knife assaults in 4835 local areas of London, 2016–2018. *Cambridge Journal of Evidence-Based Policing*, 3(1–2), 1–20.
- Morselli, C. (Ed.). (2014). *Crime and networks*. New York: Routledge.
- NYTimes (2015). “End the gun epidemic in America”. Editorial, December 4<sup>th</sup>.
- Papachristos, A. V. (2009). Murder by structure: dominance relations and the social structure of gang homicide. *The American Journal of Sociology*, 115(1), 74–128.
- Papachristos, A. V., Braga, A. A., & Hureau, D. M. (2012). Social networks and the risk of gunshot injury. *Journal of Urban Health*, 89(6), 992–1003.
- Papachristos, A. V., Hureau, D. M., & Braga, A. A. (2013). The corner and the crew: the influence of geography and social networks on gang violence. *American Sociological Review*, 78(3), 417–447.
- Papachristos, A. V., Wildeman, C., & Roberto, E. (2015a). Tragic, but not random: the social contagion of nonfatal gunshot injuries. *Social Science & Medicine*, 125, 139–150.
- Papachristos, A. V., Braga, A. A., Piza, E., & Grossman, L. S. (2015b). The company you keep? The spillover effects of gang membership on individual gunshot victimization in a co-offending network. *Criminology*, 53(4), 624–649.
- Radil, S. M., Flint, C., & Tita, G. E. (2010). Spatializing social networks: using social network analysis to investigate geographies of gang rivalry, territoriality, and violence in Los Angeles. *Annals of the Association of American Geographers*, 100(2), 307–326.
- Randle, J., & Bichler, G. (2017). Uncovering the social pecking order in gang violence. In B. LeClerc & E. U. Savona (Eds.), *Crime prevention in the 21st century: Insightful approaches for crime prevention initiatives* (pp. 165–186). Cham: Springer.
- Tita, G. E., & Radil, S. M. (2011). Spatializing the social networks of gangs to explore patterns of violence. *Journal of Quantitative Criminology*, 27(4), 521–545.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Paolo Campana** , Ph.D. is a University Lecturer in Criminology and Complex Networks at the Institute of Criminology, University of Cambridge, where is the leader of the Crime & Networks Group.

**Andrea Giovannetti** , Ph.D. has been a post-doctoral research associate in the Department of Economics at the University of Technology–Sydney since 2017, and since 2020 has been a member of the [Crime & Networks Group](#) at the [Institute of Criminology](#) of the University of Cambridge.