Risk assessment and risk management of violent re-offending among prisoners

Constantinou, AC; Freestone, M; Marsh, W; Fenton, NE; Coid, J

For additional information about this publication click this link.
http://qmro.qmul.ac.uk/xmlui/handle/123456789/15804

Risk assessment and risk management of violent reoffending among prisoners

Anthony Costa Constantinou[*,1], Mark Freestone[2], William Marsh[3], Norman Fenton[4], Jeremy Coid[5].

1. Risk and Information Management Research Group, School of Electronic Engineering and Computer Science, Queen Mary, University of London, London, UK, E1 4NS. E-mail: anthony@constantinou.info
2. Violence Prevention Research Unit, Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary, University of London, London, UK, EC1A 7BE. E-mail: m.c.freestone@qmul.ac.uk
3. Risk and Information Management Research Group, School of Electronic Engineering and Computer Science, Queen Mary, University of London, London, UK, E1 4NS. E-mail: d.w.r.marsh@qmul.ac.uk
4. Risk and Information Management Research Group, School of Electronic Engineering and Computer Science, Queen Mary, University of London, London, UK, E1 4NS. E-mail: n.fenton@qmul.ac.uk
5. Violence Prevention Research Unit, Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary, University of London, London, UK, EC1A 7BE. E-mail: j.w.coid@qmul.ac.uk

ABSTRACT

Forensic medical practitioners and scientists have for several years sought improved decision support for determining and managing care and release of prisoners with mental health problems. Some of these prisoners can pose a serious threat of violence to society after release. It is, therefore, critical that the risk of violent reoffending is accurately measured and, more importantly, well managed with causal interventions to reduce this risk after release. The well-established predictors in this area of research are typically based on regression models or even some rule-based methods with no statistical composition, and these have proven to be unsuitable for simulating causal interventions for risk management. In collaboration with the medical practitioners of the *Violence Prevention Research Unit* (VPRU), *Queen Mary University of London*, we have developed a Bayesian network (BN) model for this purpose, which we call DSVM-P (Decision Support for Violence Management - Prisoners). The BN model captures the causal relationships between risk factors, interventions and violence and demonstrates significantly higher accuracy (cross-validated AUC score of 0.78) compared to well-established predictors (AUC scores ranging from 0.665 to 0.717) within this area of research, with respect to whether a prisoner is determined suitable for release. Even more important, however, the BN model also allows for specific risk factors to be targeted for causal intervention for risk management of future re-offending. Hence, unlike the previous predictors, this makes the model useful in terms of answering complex clinical questions that are based on unobserved evidence. Clinicians and probation officers who work in these areas would benefit from a system that takes account of these complex risk management considerations, since these decision support features are not available in the previous generation of models used by forensic psychiatrists.

Keywords: Bayesian networks, belief networks, causal intervention, forensic medicine, mental health, released prisoners, violent offence

1 INTRODUCTION

Violence is a major global public health and social concern. While violence can generally be described as an extreme form of aggression, the many different types of violence in conjunction with the limited understanding of their links with certain mental states make violent behaviour difficult to assess and predict. Previous research in criminology, forensic psychology and psychiatry has discovered both weak and strong associations between violence and various other demographic, environmental and individual factors; often referred to as 'risk factors'. Some of the factors that predict violence most strongly are 'static' or

---

[*] Corresponding author.

unchangeable measures of past behaviour, such as personality disorder, previous convictions for violence or violence at a young age (Monahan, 1984); other factors such as criminal networks, substance use/misuse, or serious mental illness, may be amenable to treatment or resolve over time and are therefore considered 'dynamic' (Hanson & Harris, 2000). Yet some factors, such as active symptoms of mental illness, or intoxication, are subject to minute-to-minute or hour-to-hour fluctuations and may be considered as 'acute' factors, that influence violent outcome but remain relatively unpredictable (McNeil et al., 2003).

Accuracy in risk assessment plays a major role in identifying the small group of individuals thought to pose a very high risk of harm to society and in monitoring their level of risk during and after treatment (Douglas et al., 2005). Accurate prediction for violence, even from the same data, can be heavily influenced by the analytical method (Elbogen & Johnson, 2009; van Dorn et al., 2012), suggesting that the true underlying causes of violence are yet to be fully understood.

Prediction of violence by individuals in psychiatric and criminal justice services has evolved from simple *unstructured* estimation of risk based on clinical knowledge and intuition, through an 'actuarial' approach based on static predictors of violence, to *structured professional judgement* (SPJ), in which a list of static risk factors is considered alongside dynamic factors as well as idiosyncratic factors specific to the individual to provide a guided formulation of an individual's risk of violence. There are many SPJ tools following this template available to the clinician, including the HCR20 (Webster et al., 1997; Douglas et al., 2013) or Violence Risk Scale (VRS; Wong & Gordon, 2003). Although intended as guides to clinical practice, accurate validation of these risk assessment tools requires summation of the values assigned to each item and the use of the resulting numerical scale to create a 'predictive' model of future violence (e.g. Doyle et al, 2014). However, any large scale analysis of these predictive models finds that, on aggregate, neither SPJ measures nor actuarial lists of static factors perform above a 'threshold' AUC (Area Under Curve) value of 0.70 (Fazel, Singh & Grann, 2012; Yang, Wong & Coid, 2010) or correctly classifying only 60% of cases (Troquete et al, 2014), suggesting that the evidence base for such predictive models is not compelling. Additional research has also raised concerns that involvement in these studies by original authors of the risk assessment tools may have led to inflated estimates of accuracy (Singh, Grann & Fazel, 2010); and that, with some offender populations, predictive efficacy is no better than chance (Coid, Ullrich & Kallis, 2013).

Further, while previous research may aid clinical decision-makers, who are responsible for future detention or release of prisoners, in formulating possible specific risk scenarios, none of the previous studies take explicit account of the underlying causal factors of violence, and the dependencies between these and any interventions. Instead, they mostly rely on the association between variables of interest, and checklists with no statistical composition. As a result, the previously used modelling techniques are inadequate when it comes to risk management, whereby repeated and frequently updated assessment of an individual must take into consideration the effectiveness of causal interventions, thereby going beyond a classification and regression framework, and into causal analysis for simulating potential interventions. Clinicians and probation officers who work in these areas would benefit from a decision support system that takes account of these complex risk management considerations, and this can be achieved, as we show in this paper, by the use of causal Bayesian networks (BNs).

BNs, sometimes also called *belief networks* or *causal probabilistic networks*, can be applied to model complex problems, where variables and knowledge from different sources need to be integrated within a single causal framework (Pearl, 1988; Heckerman et al., 1995; Jensen, 1996). The use of BNs for risk assessment and risk management of violent behaviour has not previously been studied in this area of research, yet it bears similarities with other

areas of critical risk assessment and decision making where properly developed BNs have provided transformative improvements (Fenton & Neil, 2012). For instance, BNs have been employed for analysis and knowledge representation with success in diverse domains such as computational biology and bioinformatics (Friedman et al., 2000; Hohenner et al., 2005; Jiang et al., 2011), gaming (Lee & Park, 2010), computer science and artificial intelligence (de Campos et al., 2004; Pourret et al., 2008; Fenton & Neil, 2012), medicine (Heckerman et al., 1992; Diez et al., 1997; Nikovski 2000), and law (Fenton & Neil, 2011; Fenton et al., 2013; Taroni et al.,2014). Especially relevant recent use of BNs include management of project maintenance delays based on expert judgments (de Melo & Sanchez, 2008), risk analysis in large projects to extend their understanding of project risks within the Korean shipbuilding industry (Lee et al., 2009), systematic development of causal interventional systems for prognostic decision support (Yet et al., 2011), qualitative examination and evaluation of service offered by the loan departments of Greek Banks (Tarantola et al., 2012), safety control decision support in dynamic complex project environments (Zhang et al., 2013), football match prediction (Constantinou et al., 2012; 2013) and inference of referee bias (Constantinou et al., 2014), detection of problems in software development project processes (Perkusich, 2015), and jointly monitoring internal and external performance of a Master's programme of an Italian University in a holistic approach (Di Pietro et al., 2015).

Despite the significant benefits demonstrated, BNs are still under-exploited in clinical assessment. Experts may be challenged to express their knowledge in probabilistic form, and for complex problem domains elicitation of expert knowledge may require an extensive iterative process to ensure that the experts a) agree on the structure of the model and the variables to be considered for inference; and b) are comfortable with the nodes, states, and conditional dependences before they make any statements of probability.

In this paper, we present a BN model, which we call DSVM-P, for risk assessment and risk management of violent reoffending for released prisoners. The paper contributes to forensic psychiatry research with a novel causal probabilistic model that challenges the well-established regression and rule-based predictors (which currently represent the state-of-the-art in violence prevention) with higher predictive accuracy, superior decision support, and superior risk management via the simulation of causal interventions. The paper also contributes to expert systems research by showing how an expert-constructed BN model that *learns* from complex questionnaire and interviewing data (that was never intended for causal analysis) is still capable of outperforming the relevant state-of-the-art predictors, in terms of whether a mentally ill prisoner is determined suitable for release, by assessing the risk of violence over a specified time period after release.

The paper is organised as follows: Section 2 describes the data and methodology behind the development of DSVM-P; Section 3 describes the model; Section 4 discusses the results; Section 5 discusses model benefits and limitations relative to the current state-of-the-art; Section 6 provides our concluding remarks and direction for future research.

2 DATA AND METHODOLOGY

Extensive statistical analysis of cohort data, primarily focusing on classification, has previously been carried out by the research team, leading to the development of a conceptual staged assessment and management model for individual patients and released prisoners (Coid et al., 2009; Ullrich & Coid, 2011). While this statistical analysis has identified useful predictors for violent behaviour, it has also shown that none have sufficient predictive accuracy for a purely statistical approach to be effective for decision support.

The data used is the Prisoner Cohort Study (PCS) dataset (Coid et al., 2009) which consists of interview and assessment data on 1717 prisoners serving sentences of at least 2

years for sexual or violent offences (Coid et al., 2007). Interviews were performed over two phases; phase 1 interviews took place during prison sentence approximately 2 years before release, and phase 2 interviewing approximately 2 years after release. However, only 1004 of these cases were interviewed at phase 2, of whom 13 cases could not be matched to the criminal records of the Police National Computer (PNC), and a further 38 were lost to follow-up. Therefore, 953 individual cases were considered for parameter learning; 778 males and 175 females.

2.1. Expert-driven causal structure

BNs provide us with the flexibility to construct the causal structure of the model purely by expert judgment. DSVM-P was built by combining data and knowledge. The development of DSVM-P was supported by two clinically active experts in forensic psychiatry (JC) and forensic psychology (MF), each with at least 8 years' experience in forensic mental health research, having published widely on: criminal justice outcomes (Fox & Freestone, 2008; Coid et al., 2011; Coid et al., 2013), psychopathy and personality disorder (Coid et al., 2012; Freestone et al., 2013), and mental illness (Coid et al., 2013). Overall, the model development process first determined the structure and then the parameters of the model. The structure was mainly based on expert knowledge while the parameters were learnt from data. We consider these two stages in turn.

The primary steps for model development were: a) expert driven identification of model variables which were considered to be important for estimating the risk of violence, and b) expert constructed causal model structure based on the variables identified at step (a). The model structure was divided into a number of key model components which we explain in detail later in Section 3.

Much of the current research on BN construction assumes that sufficient data may be available to make the experts' input redundant. However, while 953 cases may seem like a lot of data, it is actually insufficient for constructing a causal structure as complex as that presented in Figure A.1 (for reasons explained in (Fenton, 2015)). Also, relying purely on data-driven solutions in such problems can miss explanatory or intervention information. To understand this, consider Figure 1 which presents three different models for *head injury,* and demonstrates how an expert constructed BN model that incorporates both expert knowledge and data (Figure 1c), can be more sensible than both a causal model learnt purely from data (Figure 1b; Sakellaropoulos & Nikiforidis, 1999) and a standard data-driven statistical regression predictor (Figure 1a).

The motivation behind Figure 1 is to demonstrate that purely data-driven models are bound to fail in generating a sensible causal structure when important factors (e.g. in this example "seriousness of injury", and "treatment") are absent from the dataset. When an observation is provided into the causal network, the model informs us about factors that are directly or indirectly causally linked to the observed event. This highlights the importance of the causal structure for inference, and perhaps the requirement in spending effort with the domain experts in ensuring that causality between factors in the BN flows sensibly.

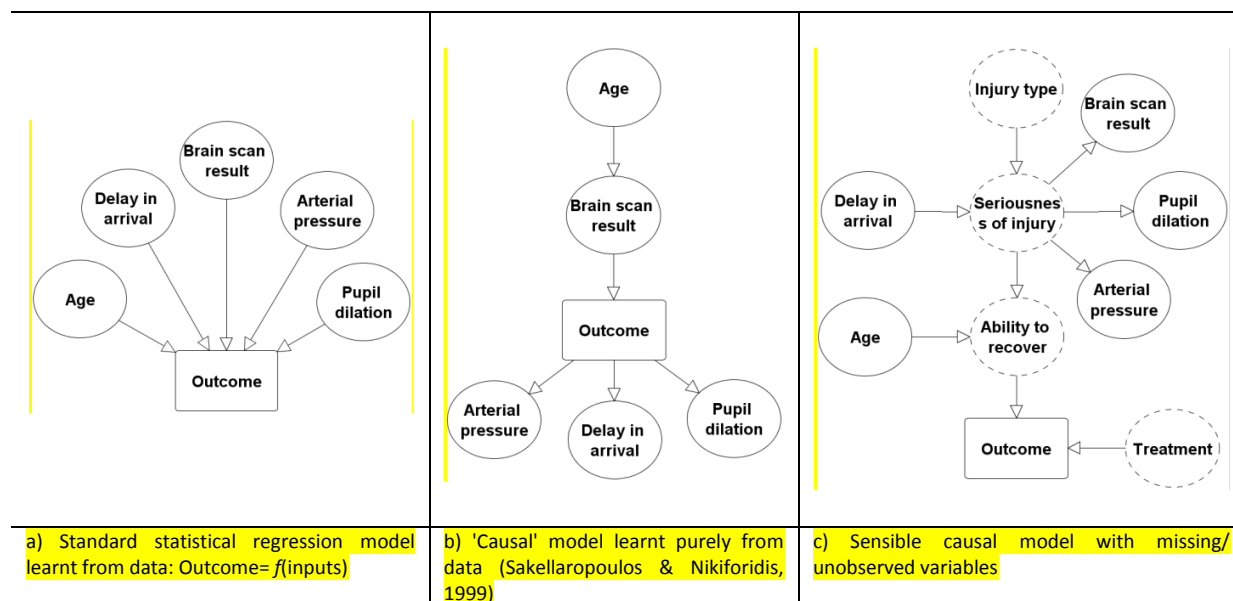|		
a) Standard statistical regression model learnt from data: Outcome= *f*(inputs)	b) 'Causal' model learnt purely from data (Sakellaropoulos & Nikiforidis, 1999)	c) Sensible causal model with missing/ unobserved variables

Figure 1. The problem with data-driven models (Fenton, 2014).

2.2. Parameter learning

The model is parameterised using data from the PCS. The first step is to link relevant questionnaire data to model variables, with the help of the experts, and a BN variable is linked to one or more relevant questionnaire answers. For example, in the case of the variable *Financial difficulties*, the sources of information for learning were answers provided to questions "*Are you behind paying bills?*", "*Have you recently had any services cut off?*", and "*What is your average weekly income*". We assume *p(Financial difficulties=Yes)* if evidence of financial difficulties are observed for at least one of those responses.

The next step is to learn the model parameters. To deal with missing data we use the Expectation Maximisation (EM) algorithm (Lauritzen, 1995). The experts were then asked to review the model (by playing with the model in AgenaRisk), in terms of inferred outcomes at different parts of the model, and suggest further revisions where necessary. In particular, after model reviewing, revisions were normally suggested (or had to be performed) in cases where:

- disagreements between experts initially existed about the inclusion or not of one or more variables in the model;
- disagreements between experts initially existed about the link (or the direction of a link) between model variables;
- disagreements between experts initially existed about the formulation of one or more model variables from questionnaire data;
- data indicated no effect between causally defined model variables;
- causal model links that were initially creating an endless loop for a set of variables (cycles are not allowed in BNs);
- further analysis revealed very strong correlation between non-linked model variables.

Since many of the steps were expert driven, disagreements between experts about both model the structure and the variable identification were encountered due to the high complexity of the domain. Extensive iterative process for expert knowledge elicitation

ensured eventual agreement between experts on both the structure of the model and the data variables considered for inference.

3   THE MODEL

DSVM-P was built using the AgenaRisk BN tool (AgenaRisk, 2012). As well as the standard discrete variables, AgenaRisk also supports continuous state variables which are approximated using dynamic discretisation (Neil et al., 2010); we make use of this unique feature for a number of variables as described later in this section.

The model is constructed on the basis of six generic factors: *Criminal attitude*, *Personality disorder*, *Socioeconomic factors*, *Mental illness*, *Substance misuse*, and *Treatment responsivity*. There are model components corresponding to each of the six factors. A seventh component called *Violence and other static risk factors* links dynamic and static risk factors for assessing violence. Figure 2 demonstrates a simplified model component topology of the overall BN, and the complete BN model is presented in Figure A.1. Table B.1 provides detailed description of all the model variables. Note that, although at this schematic level (Figure 2) there is a cycle, no cycles exist in the full model.
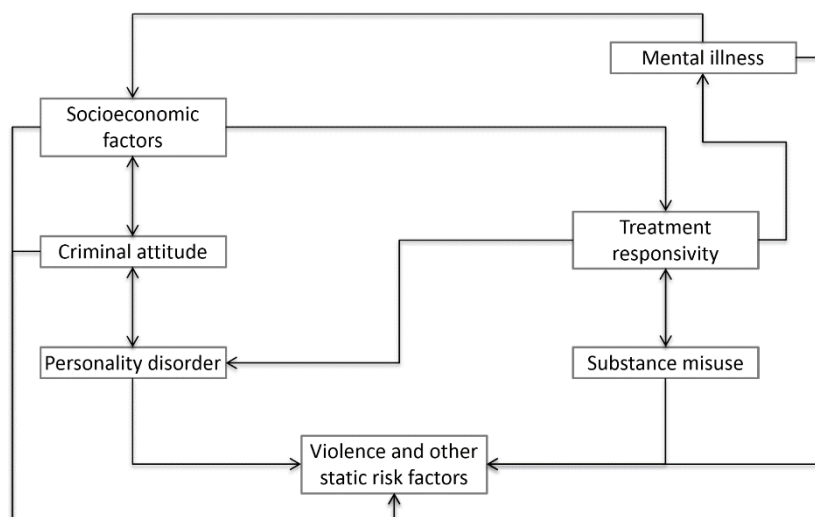

Figure 2. Simplified model component topology of the overall BN. Dual-directed links between components indicate multiple dependencies between variables of one component to another.

We provide a brief description for each of the six model components and demonstrate their direct interactions with child/parent nodes from other components in the subsections that follow. In addition, we also provide a detailed description on the design of the seventh component, which is responsible for linking all of the parts of the model for future violence estimation. There are four categories of nodes/variables:

- Oval nodes with solid border representing observable variables;
- Oval nodes with dashed border representing definitional or any other latent variables;
- Square nodes with solid border representing interventions (i.e. treatments or therapies);
- Square nodes with dashed border representing latent unobservable variables as a result of an intervention (i.e. post-treatment effect).

3.1. Model component: Criminal attitude

Involvement in crime and a criminal lifestyle has been long known to be associated with violence (Andrews & Bonta, 1994), either through the instrumental use of force by criminals to obtain goals (e.g. in robbery) or through a tendency for criminal activities that may not be violent in themselves (e.g. sale of illegal drugs) to be associated with a more violent lifestyle due to operating outside the scope of the law (White, 1997). Involvement in criminal activities is hypothesised to be positively influenced by the presence of criminal activity in familial or peer groups, which may in turn lead causally to the development of attitudes supportive of crime in an individual (Patterson et al., 1989).
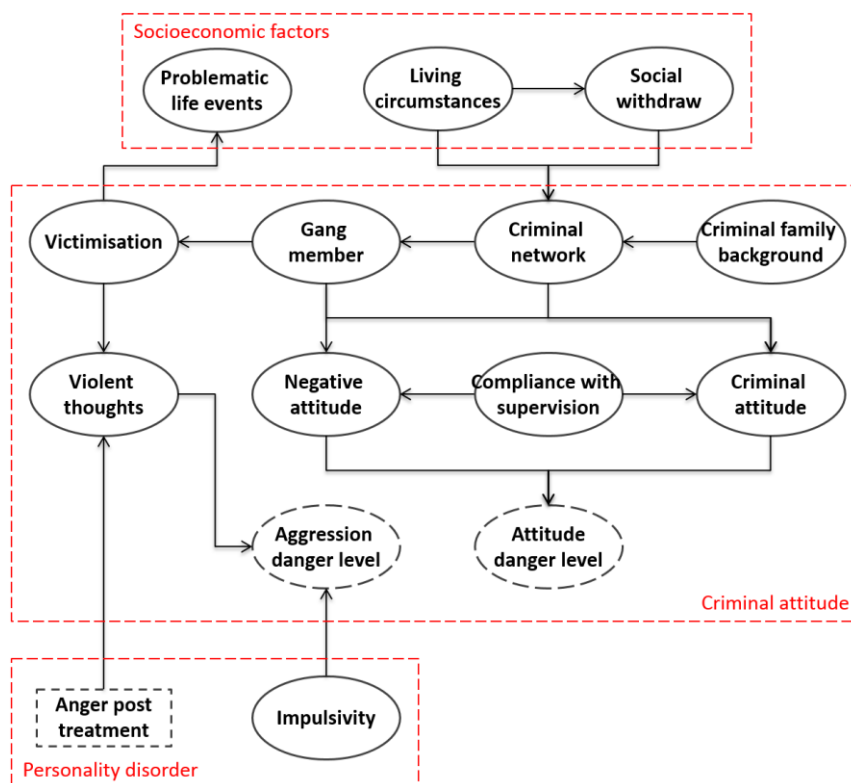

Figure 3. *Criminal attitude* component and its direct interactions with child/parent nodes from other components.

In the development of this component (Figure 3) we reasoned that markers of criminal attitude (expression of criminal attitude; criminal network; criminal family background) would be more causally related to violence if they were accompanied temporally by acute risk factors such as anger, victimisation, the presence of violent thoughts, or gang membership. Similarly but conversely, social withdrawal (e.g. due to symptoms of mental illness as demonstrated later in Figure 6) may work protectively in this regard as it would remove offenders from a context in which they may act violently as part of a criminal group.

3.2. Model component: Personality disorder

Personality disorders are chronic mental disorders which are characterised by a pervasive pattern of disturbed thought and behaviour persisting from early adulthood (APA, 2013), some of which have links with thought and behavioural patterns associated with violence. For example, *borderline personality disorder* (BPD) is characterised by disturbed identity,

impulsive behaviour and self-harm; *antisocial personality disorder* (AsPD) is characterised by high levels of anger and aggression, deception failure to obey social norms – for instance, through criminality – and a lack of remorse. Arguably another form of personality disfunction, *psychopathy* is not currently a medical diagnosis, but is an accepted condition within forensic services measured by a 20-item checklist called the *Psychopathy Checklist-Revised* (Hare, 2003) and comprising two separate but correlated factors each consisting of two 'facets': Factor One is characterised by the absence of empathy and remorse ('affective' facet) together with interpersonally manipulative traits ('interpersonal' facet); and Factor Two comprises mostly behavioural dysfunction relating partly to impulsivity ('impulsive' facet, or Facet Three) or the tendency to act without thinking; and criminality ('antisocial' facet). Some traits indicative of antisocial personality disorder - particularly impulsivity - are shared by those comprising Factor Two of psychopathy (Coid & Ullrich, 2010). Where the presence of Factor Two traits have been found to correlate directly with criminal violence (Skeem & Mulvey, 2001), Factor One traits predict violence only weakly (Skeem et al., 2002), but has a strong negative influence on treatment outcome (Olver et al., 2013).
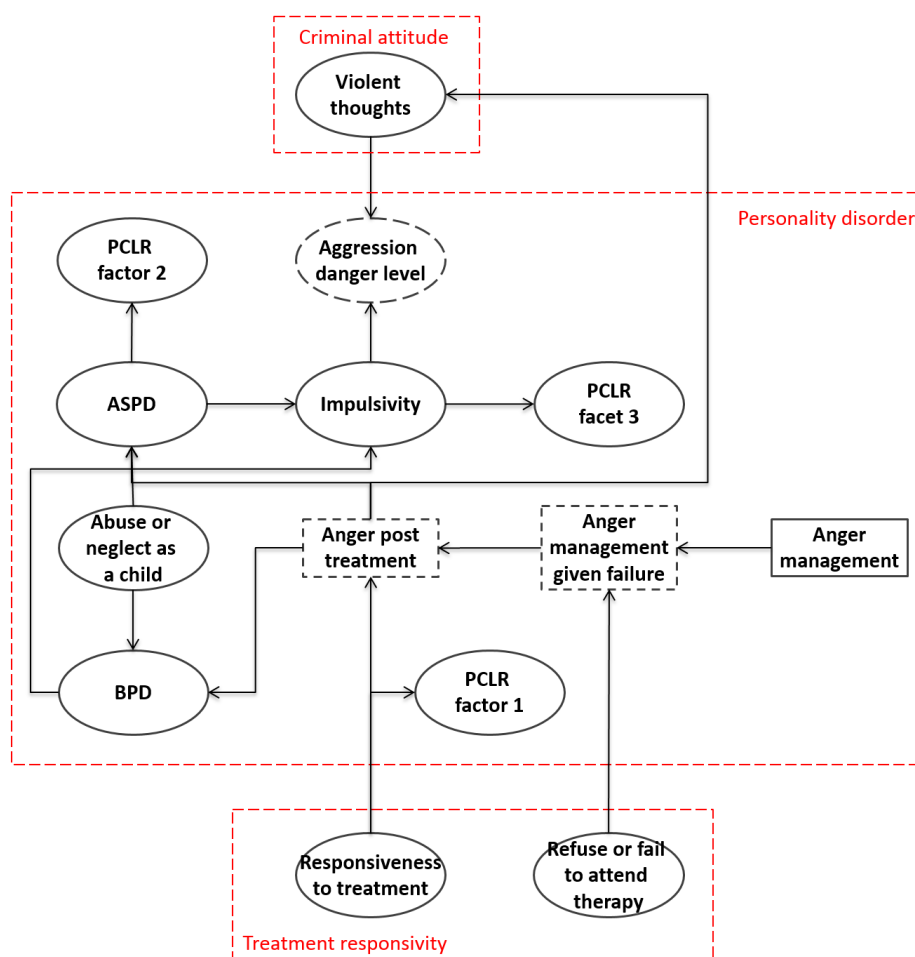
Figure 4. *Personality disorder* component and its direct interactions with child/parent nodes from other components.

When constructing this component (Figure 4), we considered personality disoders and psychopathy to be static, lifetime constructs (in the manner suggested by (Douglas et al., 2013) with potential antecedents in childhood abuse or neglect (Johnson et al., 1999) that increase vulnerability to impulsive and aggressive behaviour – which in turn increase risk of violence – and can interfere with treatment response. Anger within the component is

modelled as if it was a trait in personality disorder (what is known as 'trait' anger; Spielberger & Sydeman, 1994); however in the dataset used for validation we only had access to information about 'state' anger, which details the individual's feelings of anger at the time of interview. Using 'state' as a proxy for 'trait' anger may lead to some inaccuracies as the individual may have been angry at the time of interview for legitimate reasons (length of the interview; victimisation in prison, etc) unrelated to personality.

3.3. Model component: Socioeconomic factors

Low or unstable socioeconomic status has been shown to be associated with violent crime, but only causally in the case of acute stress (i.e., hour-to-hour fluctuations in status such as being made homeless) or in the context of a general 'stain theory', by which violence can be explained as the product of multiple overlapping stressors upon an individual (Agnew, 1992).

In this model component (Figure 5) our intention was to model social stresses upon an individual that might lead to violence in an attempt to cope – e.g. through robbery or displaced aggression against family or friends – and to see how an individual's social resources – education, intelligence, social network – might counteract the effects of the stress. Mental disorders such as anxiety or depression, which may also negatively influence an individual's ability to cope, were linked in from component 3.4 (below).

Figure 5. *Socioeconomic factors* component and its direct interactions with child/parent nodes from other components.

In terms of individual resources, higher – or more stable - socioeconomic status, including both high intelligence and higher levels of educational attainment, may act protectively in terms of preventing an individual's involvement in crime (de Vogel et al., 2011). Evidence suggests that stable intimate relationships and appropriate, supervised living circumstances for prisoners and patients nearing discharge are important factors in preventing violence (Ullrich & Coid, 2011; de Vries Robbé et al., 2013).

3.4. Model component: Mental illness

Mental illness in this component refers to a specific set of mental disorders – mood disorders or psychoses – that may differentially affect risk of violence. Mental illness and violence have long been stereotypically linked in Western culture through archaic representations of the 'mental patient' but the reality is that they have been said to have an 'intricate link' which may be explained by other risk factors such as substance misuse (Elbogen & Johnson, 2009) or may depend upon specific markers for mental illness such as childhood abuse or neglect (van Dorn et al., 2012). In either case, effective treatment for mental illness is widely understood to be critical in preventing violence in individuals with such a condition.

In constructing this component (Figure 6), our approach was to build nodes relating to individual symptoms or traits of mental illness, rather than diagnostic categories. Diagnostic categories can be difficult to ascertain to all but the best-trained of clinicians; and even then reliability of diagnosis between clinicians can be very poor (McGorry, 2013). Further, recent research has demonstrated that specific symptoms, rather than the clusters of symptoms represented by diagnoses, may have links to violence, particularly when mediated by affective states such as anger. Examples of this include: a subset of delusional beliefs being causally linked to violence (Coid et al., 2013; Keers et al., 2013); or command hallucinations directing the patient to harm others (McNiel et al., 2000).
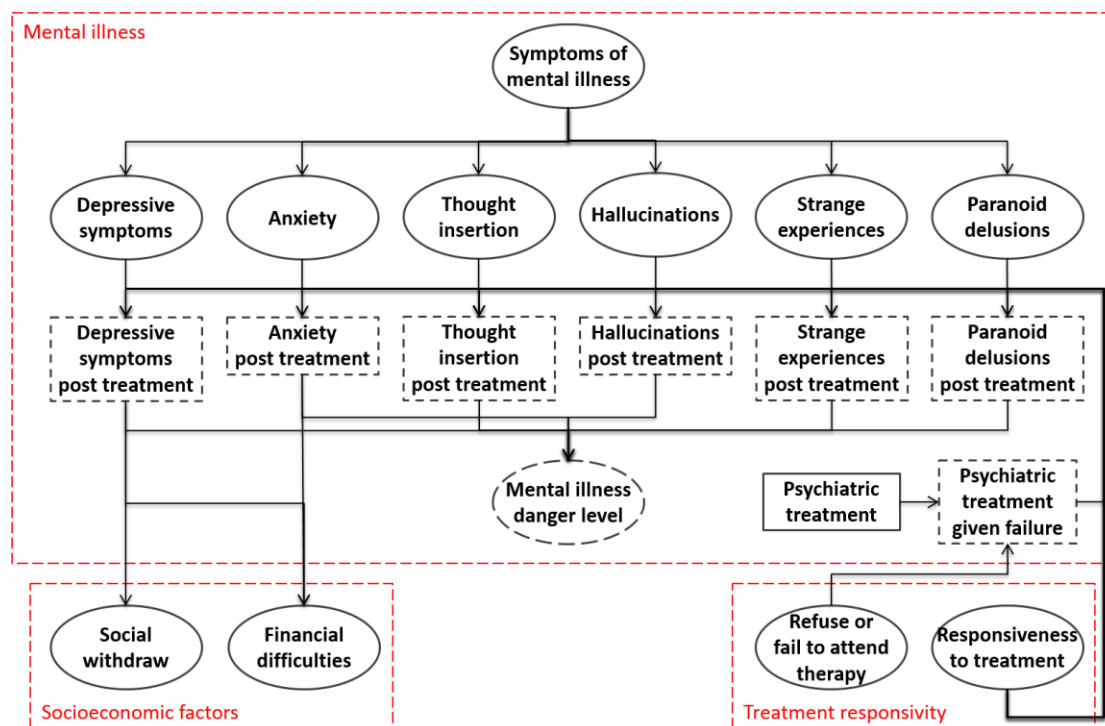

Figure 6. *Mental illness* component and its direct interactions with child/parent nodes from other components.

3.5. Model component: Substance misuse

The substance misuse component (Figure 7) assesses the risk level for violent re-offending based on the misuse of a number of drugs and/or hazardous alcohol consumption. Substance abuse is a clinically identified psychiatric disorder characterised by distress caused to an individual due to the use of a psychoactive drug (APA, 2013), including alcohol, that may also manifest in extreme cases as substance dependence where it leads to increased need for the drug. The relationship between substance abuse and violence is complex: it may be causative in the sense that some stimulants directly increase aggressive or violent behaviour through their psychopharmacological action (e.g. Davis, 1996); that substance abuse or dependency stimulates acquisitive violence to fund addiction ('economic compulsive violence'; (Goldstein, 1985)) or it may be that use of illegal substances implies involvement in social systems where violence is more likely ('systemic violence'; Boles & Miotto, 2003). Whatever the case, substance misuse has been found to increase risk of violence by up to four times in most populations, particularly in individuals suffering from existing mental illness (Steadman et al., 1998; Elbogen & Johnson, 2009).
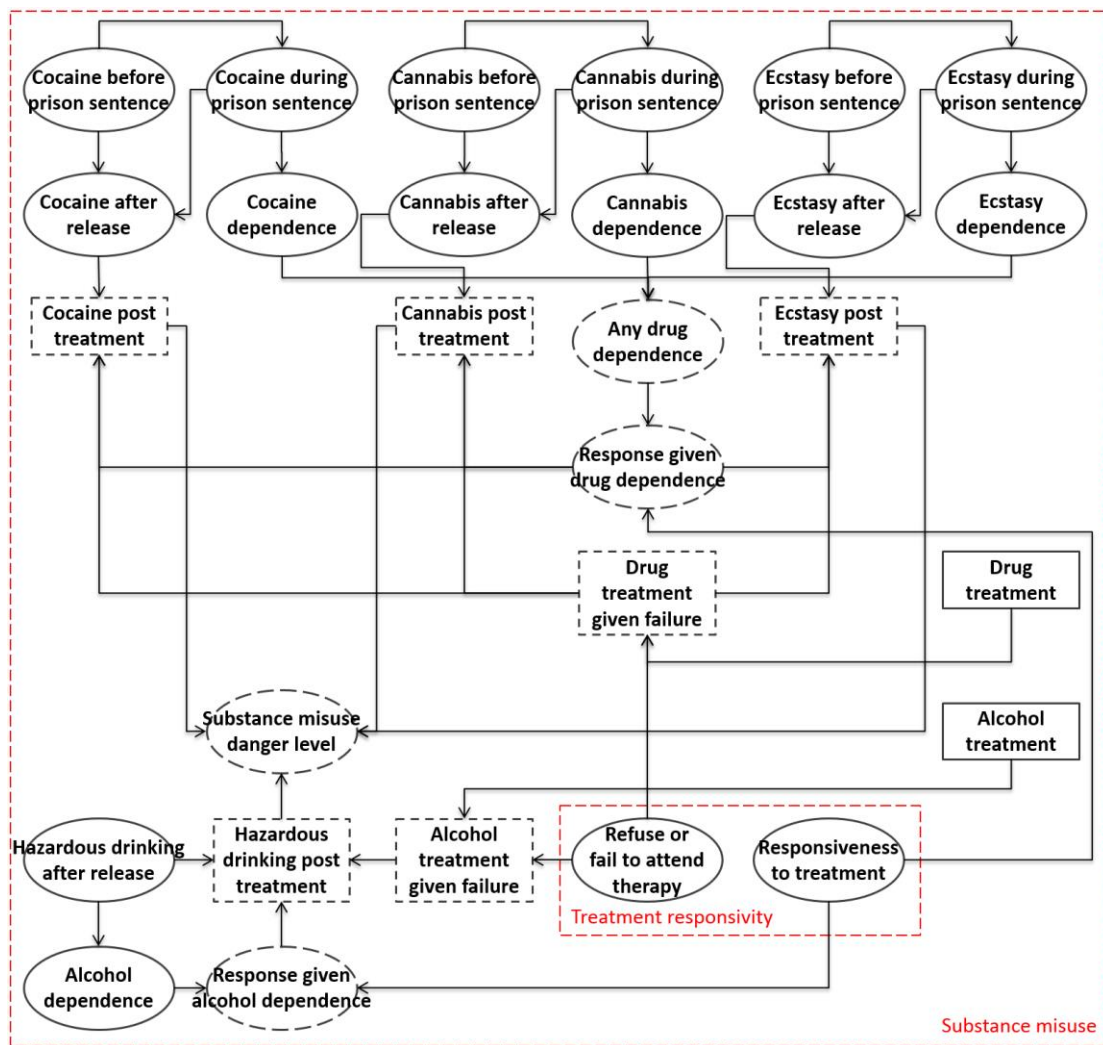

Figure 7. *Substance misuse* component and its direct interactions with child/parent nodes from other components.

3.6. *Model component: Treatment responsivity*

Poor adherence or response to treatment in individuals with severe mental illness, are known risk factors for violence (Witt et al., 2013). Equally, the effect of successful treatment on either substance misuse or symptoms of mental illness may be to nullify the relationship of these disorders to violence by removing the underlying cause (addiction compulsion; command hallucinations, etc).

The *Treatment responsivity* component is represented by two factors: 1) the responsiveness to any given treatment, and 2) the risk of refusing or failing to attend any given therapy. We have already demonstrated in the previous subsections how treatment responsivity is individually linked to the components of mental illness, personality disorder, substance misuse, and criminal attitude. Figure 8 demonstrates these links collectively.
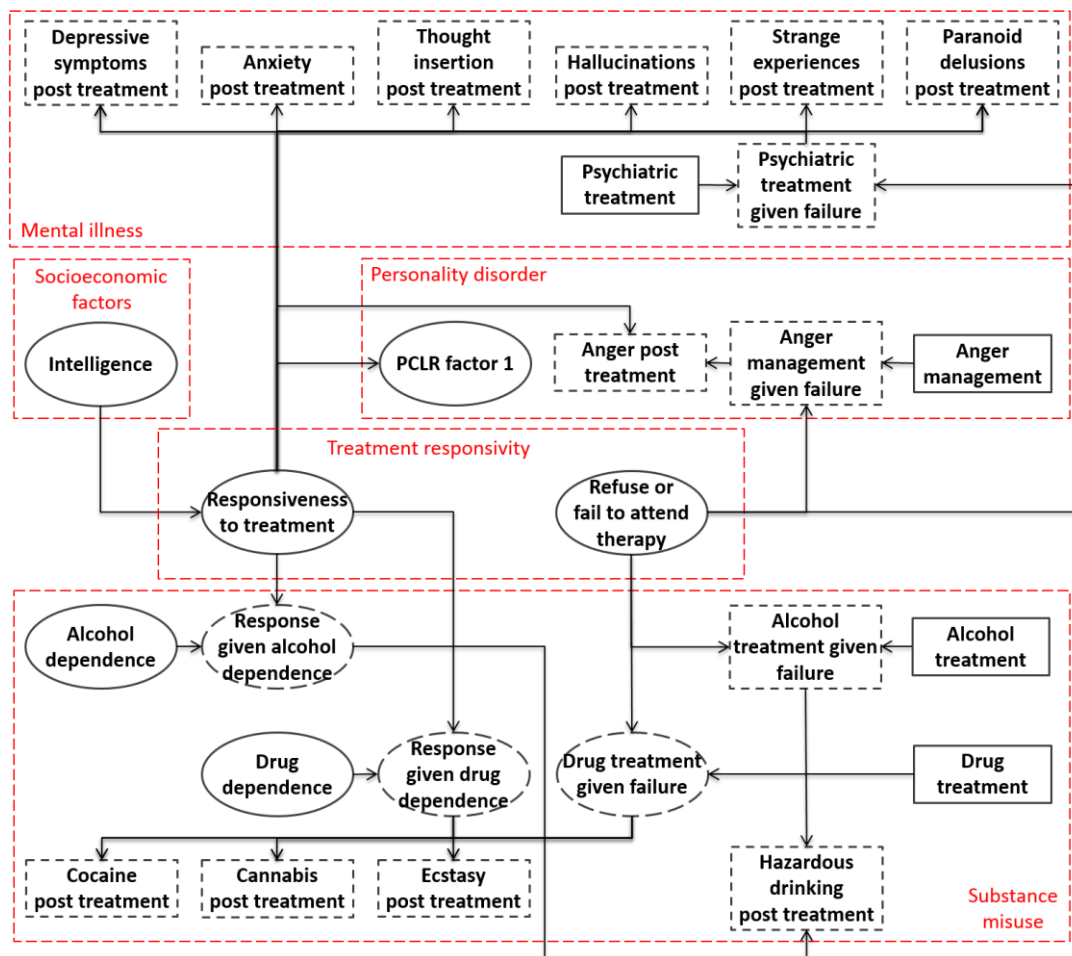

Figure 8. *Treatment responsivity* component and its direct interactions with child/parent nodes from other components.

3.7. *Model component: Violence and other static risk factors*

In the previous six subsections we have demonstrated the six model components corresponding to each of the six dynamic factors. Four danger level variables and one protective level variable are associated with these components. The four danger levels and the protective level variables are binary defined with states *Low* and *High*, indicating relative low and high risks for violent re-offend based on key-variables within those components. Specifically, the danger indication will be *High* for the combination of observed key-

variables for which the highest rate of violence is observed, and vice versa. For any other combination of observations, for which the rate of violence is between the minimum (*Low*) and maximum (*High*), the probabilities for *Low* and *High* will adjust relatively (i.e. *Low=High*=50% when the rate of violence is equal to the average rate).
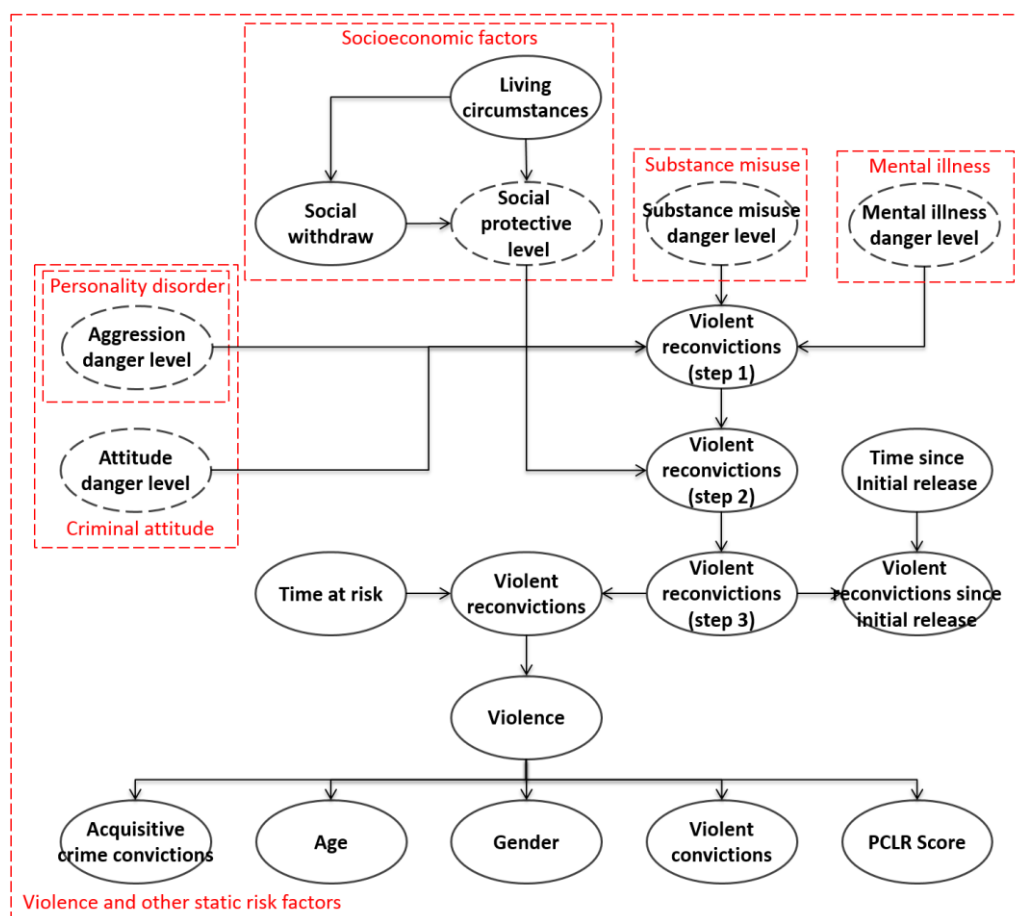

Figure 9. *Violence and other static risk factors* component and its direct interactions with child/parent nodes from other components.

The component presented in Figure 9 can be described in five steps. In brief, from steps 1 to 3 the *Violent convictions rate* is inferred hierarchically and respectively for each step, based on a) the danger levels, b) the protective level, and c) the number of days the released prisoner has already spent out of prison (with or without evidence of violent re-offence). Further, at step 4 the revised *Violent convictions rate (step 3)* is considered for predicting the expected number of violent reconvictions over a specified period of time in the future, before this information is revised at step 5 on the basis of the five static risk factors. Each of the five steps is described in detail as enumerated below:

1. We assume that the *Violent reconviction rate (step 1)* follows a $Beta(\alpha, \beta)$ distribution which is estimated on the basis of the four danger levels; where hyperparameter α is the number of violent convictions observed over the observation period and hyperparameter β is the observation period (in days) minus α.

   While the *beta* distribution assumes the combinations of *Low* and *High* for the danger levels, we have instead provided the combinations of ¬*High* and *High* as demonstrated in Appendix C. This was done to ensure that sufficient data points are generated for a reasonably well informed prior for *p(Violent reconviction rate (step 1))*; if we were to

follow the proper set of combinations no prior information would had been available for many of those combinations due to an insufficient number of instances in the dataset. Further, the high complexity behind the definition of each danger level made conditional probabilities between danger levels highly uncertain and not feasible for expert probability elicitation.

2. A revised *beta* distributed *Violent reconviction rate (step 2)* is generated based on the social protective level.

3. A revised *beta* distributed *Violent reconviction rate (step 3)* is generated based on *Time since initial release* (assessed in number of days), and *Violent reconvictions since initial release*. We assume that the three variables follow a *Beta-binomial* approach such that the *Beta* distribution *Violent reconviction rate (step 3)* serves as conjugate distribution of the $\sim Binomial(n, p)$ distribution *Violent reconvictions since initial release*, formulating a compound distribution such that the p parameter is randomly drawn from the *Beta* distribution. The variable *Time since initial release* serves as the input n (in days) for the Binomial distribution. Consequently, the process assumes constant probability[†] for violence over each trial (day).

So, for example, if we are monitoring an individual over a period of two years and we observe no evidence of violent re-offence, then our belief for that individual becoming violent in the future diminishes (in comparison to what it was immediately after release). Figure 10 demonstrates the reduction in the risk of violent re-offence over a period of 2,000 days with no evidence of violent re-conviction (and the prediction given assumes further 2,000 days in the future; i.e. *p(Time at risk=2000)*). The reduction effect is subject to exponential decay. For example, after 1,000 days out of prison, without evidence of violence, the reduction is approximately 10 absolute percentile points (i.e. down to ~18% from ~28%), whereas after further 1,000 days the risk is further reduced by 4.5 absolute percentile points (i.e. down to ~13.5% from ~18%).
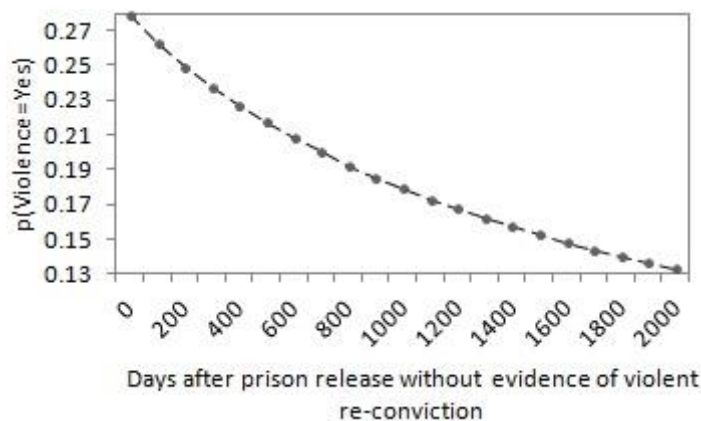

Figure 10. Risk reduction for *p(Violence=Yes)* over the specified number of days out of prison with no evidence of violent reconviction. This assessment assumes *p(Time at risk= 2,000)*.

Alternatively, Figure 11 demonstrates how the risk of violent re-offence would have increased had we observed violent reconvictions for that individual and over the same period (and with the same assumption for *Time at risk*). In this case, the increase in the

---

[†] Time-series analysis was not possible with our dataset, and no other relevant published research study has attempted to answer this question.

risk of violence follows a logarithmic growth. For example, when we observe 2 violent reconvictions (after 2,000 days spent in the community) the risk of violence over the next 2,000 days follows an increase of a massive 56 absolute percentile points (as opposed to observing 0 violent reconvictions), whereas in the case 4 violent reconvictions the risk of violence is increased by an additional 19.5 absolute percentile points (which is still a significant increase, but considerably lower than the increase in the first scenario).
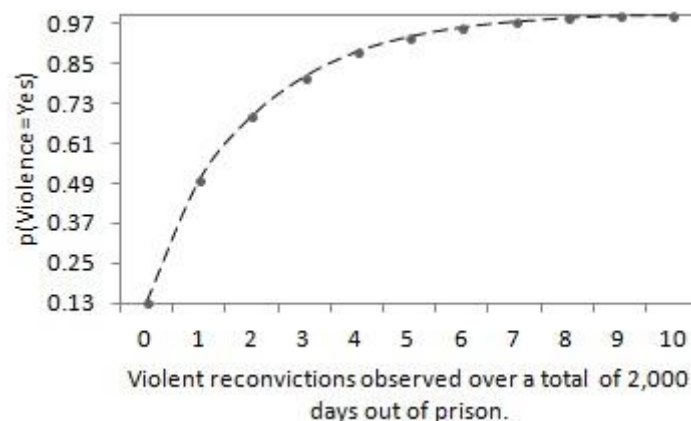

Figure 11. Risk increase for *p(Violence=Yes)* over the specified number of violent reconvictions observed over 2,000 days out of prison. This assessment assumes *p(Time at risk=2,000)*.

4. A prediction for *Violent convictions* is generated on the basis of a repeated *Beta-binomial* process, such that the *Beta* distribution *Violent reconviction rate (step 3)* serves as conjugate distribution of the $\sim Binomial(n, p)$ distribution *Violent convictions*, formulating a compound distribution such that the p parameter is randomly drawn from the *Beta* distribution. The variable *Time at risk* serves as the input n for the *Binomial* distribution. Specifically, when we provide information for *Time at risk* DSVM-P generates *Binomial* distributed prediction for the number of violent convictions expected over the specified (time at risk) period.

5. The variable *Violence* indicates a binary prediction for future violent reconviction, and which is translated from *Violent reconvictions* such that *0* violent reconvictions indicate *p(Violence=No)* and *1≥* violent reconvictions indicate *p(Violence=Yes)*. The prediction for violence (and consequently violent reconvictions) is then revised based on the five static risk factors of *Age*, *Gender*, *PCLR Score*, *Prior violent convictions* and *Prior acquisitive crime convictions*. All the of above static risk factors serve as strong predictors for violence but none of which serves as an underlying cause of violence.

4 RESULTS AND DISCUSSION

In this section we assess the performance of DSVM-P and comment on the results. Specifically, Section 4.1 assesses the predictive accuracy of DSVM-P, Section 4.2 analyses interventions, and Section 4.3 analyses the danger levels.

4.1. Assessment of predictive accuracy

While there are several scoring functions available to assess the predictive accuracy of a probabilistic model of violent reoffending, the area under the curve (AUC) of a receiver operating characteristic (ROC) is the standard method in this domain for binary predictive

distributions. Hence, we use the AUC of ROC to compare the predictive performance of DSVM-P against other well-established probabilistic predictors in this area.

Some advantages, such as independence of both base rate and selection ratio, over other measures are appreciated in this field (Hanley & McNeil, 1982a, 1982b; Rice & Harris, 1995), and in (Rice & Harris, 2005) the authors outlined why the AUC is the preferred measure of predictive or diagnostic accuracy in forensic psychology or psychiatry. However, the AUC has also been subject to criticism. Singh (2013) explains why AUCs do not capture how well a risk assessment model's predictions of risk agree with actual observed risk, indicating that the AUCs provide an incomplete portrayal of predictive validity. While there is a long debate in the literature (Lobo et al., 2007) on how to interpret AUCs, still more than half of violence risk assessment validation studies report only the AUC (Singh, 2013) since there is no other agreed measure for violence accuracy in this domain.

Typically, the AUCs are either reported based on the whole development sample or based on a cross-validated sample. An AUC score of 0.5 indicates forecasting capability no better than chance, whereas a score of 1 (or 0) corresponds to a perfect predictive model.

Evidently, AUCs reported on the whole development sample are likely to be optimistic, especially when the model is optimised for the sample upon which they were developed in which case running the danger of overfitting the model. DSVM-P generates an AUC score of 0.79 (95% CI: 0.7552-0.8215. Performing a *10-fold cross-validation* the AUC score only drops to 0.78 (95% CI: 0.7449-0.8149). This suggests no danger for model overfitting and that the predictive accuracy of DSVM-P is expected to be very good for other similar data samples.

Table 1 shows how the cross-validated performance of the DSVM-P compares against the three well established predictors within this area of research when employed with the same dataset. All three predictors are used in clinical practice in the UK and internationally and have been previously validated through the use of AUC statistics calculated against the 'sum' of the items as described above. These predictors are:

a) **Violence Risk Appraisal Guide** (VRAG, Quinsey 2006): is a model developed in Canada for predicting reoffending by mentally ill offenders on the basis of 12 variables linked to violence risk − such as maladjustment at school or a history of alcohol use − that correlated best with reoffending as determined by multiple regression analysis (Harris et al., 1993). It has the advantages of brevity and significant use in, and validation with, clinical justice populations, but does not consider dynamic, or changeable items, associated with violence risk, and which can be targeted for intervention.

b) **Health-Clinical-Risk 20** (HCR-20; Webster et al, 1997): is a 20-item checklist of static and dynamic risk factors associated with violence in psychiatric patients, such as acute symptoms of mental illness or stable relationships after discharge (Webster et al., 2005). The HCR-20 is an SPJ measure used primarily by clinicians seeking to assess readiness for discharge amongst patients whose mental disorder is linked to their offending: it is relatively quick to complete but requires training and extensive clinical experience.

c) **Psychopathy Checklist-Revised** (PCL-R; Hare, 2003): is a checklist of 20 variables measuring psychopathy, a disorder characterised by a callous lack of empathy and remorse, shown to be strongly associated with offending behaviour in prisoner populations, although some meta-analyses have drawn its performance as a risk assessment measure into question (Singh, Grann & Fazel, 2011). Like the HCR20, it requires extensive training and experience to use accurately.

Table 1. Comparison of AUC scores between the DSVM-P and the three well established predictors for violence risk assessment, when employed with the same dataset.

Model	AUC
DSVM-P	0.78
VRAG	0.7171
HCR-20v2	0.665
PCL-R	0.6648

DSVM-P provides a significant increase in predictive accuracy for violent recidivism, over the three predictors discussed above. Figure 12 presents the partial AUCs for DSVM-P (left graph) generated at 100-90% *specificity* and *sensitivity*, and superimposed ROC curves for DSVM-P (95% CI), VRAG, HCR-20, and PCL-R predictors; indicating the significance levels between DSVM-P and the other three predictors, as well as trade-off between *sensitivity* and *specificity*.

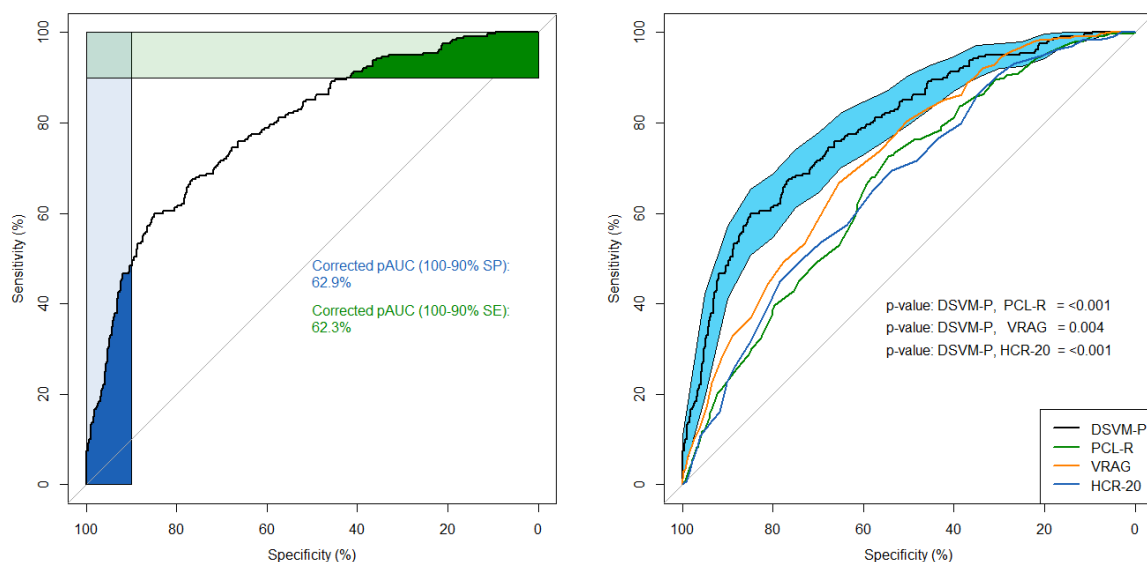Figure 12. Partial AUCs for DSVM-P (left graph), and superimposed ROC curves (right graph) for DSVM-P(95% CI), PCL-R, VRAG, and HCR-20 predictors.

4.2. Analysis of the interventions

Table 2 below demonstrates the expected reduction in the risk of violence for each intervention introduced in the model. Over each iteration, the what-if analysis (or sensitivity analysis) assumes *p(Violence=Yes)* for five years forward (i.e. *p(Time at risk=1,825)*), observable active symptoms for the intervention under analysis, and observable inactive symptoms for the remaining three interventions (with all of the other model factors unknown).

Assuming no intervention (i.e. no treatment/therapy), the results show that psychotic symptoms generate a considerable higher risk for violence (i.e. 42.85%) over hazardous drinking, drugs and anger. When intervention is advised, the results suggest that there is not much difference between partial and full responsiveness to treatment over all four interventions, and show that psychiatric treatment can be very effective with 42.88% relative reduction in the risk of violent re-offence, followed by alcohol treatment with a relative risk reduction of 24.43%, but drug treatment and anger management less effective. However, as

stated in Section 3.2, results relating to anger management should be interpreted with caution due to the temporal unreliability of the 'state' model used to measure anger levels in sample participants.

The same experiment is repeated, but this time the assumption is that the symptoms associated with the remaining three interventions (over each iteration) are also unknown (instead of inactive). As expected, the results (Table 3) demonstrate a decrease in intervention effectiveness for all cases, but the relative impact between interventions remains similar to that presented in Table 2. Repeating the experiment for a third time, but with all symptoms associated with each of the interventions being active over all iterations, the results demonstrated that none of the treatments was capable of individually providing any meaningful reduction in the risk of violent re-offence; implying that the active symptoms associated with the remaining three interventions (over each iteration) were strong enough to maintain the risk for future re-offending at the same high risk level.

Table 2. Sensitivity analysis for *p(Violence=Yes)* assuming 5 years forward, with sensitivity variables each of the four interventions assessed individually and relative to the specified treatment responsiveness. The analysis assumes observable active symptoms for the intervention under analysis, and observable inactive symptoms for the remaining three interventions, over each iteration.

Intervention for:	Treatment="No" (i.e. no intervention)	Responsiveness to treatment ="Partly"	Responsiveness to treatment ="Yes"	Sensitivity to Violence	Reduction rate for p(Violence=Yes)
Alcohol	0.1392	0.0945	0.0789	0.0633	24.43%
Drugs	0.1028	0.0871	0.0822	0.0209	9.49%
Anger	0.1444	0.1388	0.1382	0.0062	4.29%
Psychiatric	0.4285	0.2655	0.2448	0.1838	42.88%

Table 3. Sensitivity analysis for *p(Violence=Yes)* assuming 5 years forward, with sensitivity variables each of the four interventions assessed individually and relative to the specified treatment responsiveness. The analysis assumes observable active symptoms only for the intervention under analysis, over each iteration.

Intervention for:	Treatment="No" (i.e. no intervention)	Responsiveness to treatment ="Partly"	Responsiveness to treatment ="Yes"	Sensitivity to Violence	Reduction rate for p(Violence=Yes)
Alcohol	0.4183	0.3656	0.3483	0.0700	16.73%
Drugs	0.3610	0.3445	0.3395	0.0215	5.96%
Anger	0.3448	0.3374	0.3369	0.0079	2.29%
Psychiatric	0.4406	0.3505	0.3388	0.1018	23.10%

4.3. Analysis of the danger levels

Table 4 demonstrates the impact for each of the danger levels, when are individually and collectively observed, for *p(Violence=Yes)*, again assuming five years forward. The results clearly demonstrate that the risk for future re-offending is extremely low when all of the four danger levels indicate *Low* danger. When only one of the danger levels is observed as being *High*, the substance misuse appears to be most dangerous with 32.44% probability for future re-offence, whereas aggression the least dangerous with 17.38% probability. Combining two *High* danger levels, the combination of aggression and mental illness appears to be significantly more dangerous than residual combinations (with 63.55% probability for future re-offence), whereas the combination of aggression and attitude (with 30.49% probability for future re-offence) appears to be the least dangerous. Combining three *High* danger levels the risk for future re-offence is increased under all scenarios; but for the combination of

aggression, mental illness and substance misuse the risk drops considerably. This result needs further exploration, but could be due to the cluster of symptoms representing a disturbed but non-criminal group of individuals whose aggression was associated with mental illness and substance use but who mostly lacked motive or capacity for violence.

Table 4. Danger level analysis for *p(Violence=Yes)* over five years forward. A √ indicates *High* observable danger level, and ¬*High* (or *Low*) otherwise (Appendix C provides more details on these combinations).

Aggression	Attitude	Mental illness	Substance misuse	*p(Violence=Yes)*
-	-	-	-	0.0242
-	-	-	√	0.3244
-	-	√	-	0.2600
-	√	-	-	0.2468
√	-	-	-	0.1738
-	-	√	√	0.3798
-	√	-	√	0.4901
√	-	-	√	0.5405
-	√	√	-	0.4318
√	-	√	-	0.6355
√	√	-	-	0.3049
-	√	√	√	0.6625
√	-	√	√	0.2569
√	√	-	√	0.7374
√	√	√	-	0.6578
√	√	√	√	0.7784

Figure 13 presents a sensitivity analysis for target node *p(Violence=Yes)* based on the nine specified sensitivity nodes. The analysis assumes that all treatments are instantiated to "No". The tornado graph reveals three apparent clusters of impact on future violence, based on this BN structure. In the highest impact cluster we observe the factors of age, prior violent convictions and PCL-R; in the second highest impact cluster we observe prior acquisitive crime convictions and all four danger levels; whereas gender appears to be the least significant factor of the nine considered. The tornado graph also demonstrates which state corresponds to what increase/decrease for *p(Violence=Yes)*. For example, when it comes to the variable *Age* the state which results into the highest probability for *p(Violence=Yes)* is "18-19", whereas the state "60+" generates the lowest probability.
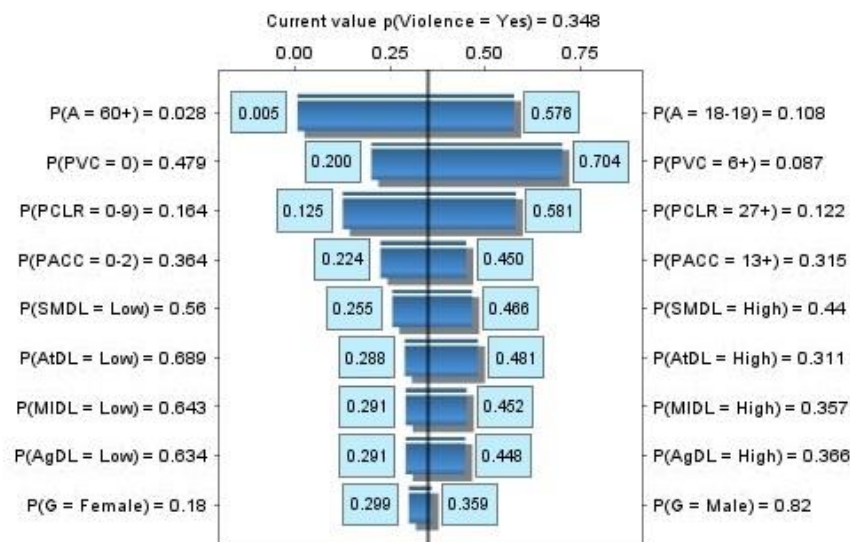
Figure 13. Sensitivity analysis for target node *p(Violence=Yes)* on the basis of the nine specified sensitivity nodes, where *A* is age, *PVC* is prior violent convictions, *PACC* is prior acquisitive crime convictions, *SMDL* is substance misuse danger level, *AtDL* is attitude danger level, *MIDL* is mental illness danger level, *AgDL* is aggression danger level, and *G* is gender. The analysis assumes that all four treatments are instantiated to "No".

5   MODEL BENEFITS & LIMITATIONS

This section provides a review of the benefits and limitations of the BN approach (and the DSVM-P model in particular), and how these compare against the well-established classical regression and rule-based predictors and methods within the domain of forensic psychiatry, which represent the current state-of-the-art.

The practical and methodological benefits can be summarised as follows:

1. ***Improved accuracy***: DSVM-P demonstrates significant improvements in the predictive accuracy (cross-validated AUC score of 0.78), compared to the current predictors that have been employed with the same dataset (AUC scores ranging from 0.665 to 0.717), with respect to whether a prisoner is determined suitable for release.

2. ***Interventional analysis***: The BN approach generally allows for specific risk factors to be targeted for causal intervention for risk management. In the specific case of DSVM-P this is done by examining whether the risk of future re-offending can be managed to acceptable levels as a result of one or more interventions, and this makes the model useful in terms of answering complex clinical questions that are based on unobserved evidence. This allows for analysis that goes beyond the predictive accuracy and into improving risk management and decision support.

3. ***Inverse inference***: In contrast to the current predictors, inference in BN models can be performed from effect to cause rather than just from cause to effect. This unique capability provides radically improved decision-support, since it enables extensive what-if analysis and provides the decision maker with the ability to examine surprising evidence. Notably, the capability of *explaining away* evidence of violence and other related risk factors can be used in DSVM-P by decision makers to

investigate the reasons as to why a particular individual has been violent, when the model was indicating otherwise.

4. ***More informative predictions:*** While, like other predictors, DSVM-P generates binary predictions for violence (i.e. Yes/No), it also provides multinomial predicted distributions for future violence (i.e. expected number of violent convictions). DSVMP also provides predictions for future violent re-offence over a specified time forward (i.e. for as little as one day, up to many years forward).

5. ***Handles missing evidence:*** While current predictors only consider what information is available for predictive analysis, and hence ignore factors for which information is unavailable, DSVM-P allows flexibility with model inputs due to the BN framework. Further, when assessing a prisoner, missing evidence (i.e. when the prisoner does not respond to specific questions) are not ignored but rather inferred from evidence provided to other factors within the model, and which are linked (directly or indirectly) to important unobserved variables;

6. ***Structural integrity:*** If required (i.e. in future studies, or when DSVM-P is learnt with different datasets), expert knowledge can be easily incorporated for factors that are important for prediction but which the historical database fails to capture. This allows the model to retain its structure for future relevant studies, regardless how limited the dataset might be in terms of the number of variables.

The practical and methodological limitations can be summarised as follows:

1. ***Extensive effort required for development:*** Developing a causal expert-driven BN, such as DSVM-P, not only requires collaboration with domain experts, but also an extensive iterative development process. Although recent research in BN modelling (Fenton & Neil, 2012) has introduced a range of techniques for reducing the burden of expert elicited models, this up-front development effort remains the primary barrier to more widespread adoption of BNs.

2. ***Necessary use of subjectivity:*** Relying on expert judgment implies inevitable subjectivity and also possible bias. This is partly addressed by using multiple experts. This, in turn, may lead to disagreements between experts; especially with regards to the causal structure of the model in such a highly complex domain.

3. ***Complexity:*** Because a causal expert-driven BN model incorporates additional variables beyond what is in any dataset, the models are typically more 'complex' mathematically (although not conceptually) than the regression and rule-based predictors. In DSVM-P, the model complexity results in a number of practical limitations. Specifically:

   a) some model variables could have been modelled with a higher number of states, and others with a higher number of *parent* nodes, but this option was not feasible due to insufficient data size;
   b) the combination of the danger levels is modelled sub-optimally (Appendix C) to ensure that sufficient data points are generated for a reasonably well informed prior, and this approach is expected to generate slightly overestimated violent reconviction rates;

c) DSVM-P assumes that there is a constant (daily) risk rate of violent re-offending that does not vary with time.

The limitations 3(a) and 3(b) can be overcome with a sufficiently larger dataset, whereas limitation 3(c) can be overcome when relevant data becomes available to allow time-series analysis for the risk rate of future violent re-offence. Having appreciated the impact the data size has on such a large and complex BN model, a richer dataset also promises even higher forecasting capability and hence, superior decision support.

6 CONCLUDING REMARKS AND FUTURE WORK

We have presented a novel Bayesian network model, which we call DSVM-P, for risk assessment and risk management of future violent re-offending for released prisoners who suffer from mental illness. Specifically, in terms of risk assessment, the model can be used to assess the risk of violence for a given individual, and over a specified time period assuming release. In terms of risk management, the model can be used to examine whether the risk of violence for the given individual can be managed to acceptable levels after release.

The need for such a system was evident by forensic medical practitioners and scientists who work in this area of research and who, over a period of several years, remained unimpressed by the decision support offered by the classical statistical, regression and rule-based systems that still dominate this area of research (Coid et al., 2014). As a result, forensic medical practitioners have identified the need to examine new ways of modelling that include the representation of causal relationships. Hence, it was felt that causal BN models could improve on the state-of-the-art. To our knowledge, this is the first BN system developed for violence prevention management in forensic psychiatry. As a result, the implications of this paper, even though it is simply a BN application of practical use, expand to both areas of research (forensic medical sciences and expert and intelligent systems).

In terms of implications in forensic psychiatry and violence prevention research, the resulting BN model presented in this paper provides an important step forward for decision support and risk management. Clinicians and probation officers who work in these areas would benefit from such a decision support system that handles the underlying complexity and that is able to properly quantify uncertainty to improve risk management and decision making by simulating the effect of potential interventions (e.g. treatments and therapies) for prisoners who are about to be released.

In terms of implications in expert and intelligent systems research, we have shown how an expertly constructed BN model, with parameter learning performed based on complex questionnaire and interviewing patient data with missing values (data that was not really suitable for causal analysis) is still capable of significantly outperforming the state-of-the-art predictors within this area of research with respect to whether a mentally ill prisoner is determined suitable for release. Specifically, the BN model demonstrates a cross-validated AUC score of 0.78, and this compares well against well-established predictors such as the VRAG, HCR20v2 and PCL-R, which demonstrate AUC scores ranging from 0.665 to 0.717 when employed with the same dataset (details in sections 4 and 5). The implications are extended to the interventional modelling case in the sense that the BN demonstrates how actions are supported by the model, with respect to determining whether a prisoner's risk of violence can be managed to acceptable levels after release on the basis of some causal intervention, such as treatment, therapy and/or medication. The outcome of this paper is in general agreement with many other studies that demonstrate decision support benefits, in various other domains, using probabilistic graphical models (de Melo & Sanchez, 2008; Pourret et al., 2008; Lee et al., 2009; Lee & Park, 2010; Fenton & Neil, 2011; 2012;

Tarantola et al., 2012; Constantinou et al., 2012; 2013; Zhang et al., 2013; Di Pietro et al., 2015; Perkusich, 2015).

In terms of AI and decision support research, the problem addressed in this paper is typical of many critical decision-making scenarios (especially in medicine, forensics, and transport safety assessment); specifically decision-makers are seeking improved methods for prediction and risk assessment, but have either little relevant data, or have to rely on poorly structured data. In such scenarios pure data-driven machine learning methods will not produce models that provide the necessary accuracy and insights. However, in combination with expert judgment, causal BNs provide the potential to do better. The method described in this paper contributes to a new research framework for building BN models in such situations. Extensive further research in this area is being carried out in the BAYES-KNOWLEDGE project (Fenton, 2014).

Other planned research extensions will determine the usefulness of DSVM-P through expert validation by carrying out pilot studies with clinicians and a qualitative assessment on a graphical user interface which is planned for future development. The capability of BNs as decision support tools will also be evaluated in individuals with serious mental health problems who are about to be discharged from Medium Secure Services.

ACKNOWLEDGEMENTS

APPENDIX A

Figure A.1. The complete Bayesian network model.

APPENDIX B: Description of model variables

Table B.1. Description of the model variables.

Variable No.	Node name	Model component	Node type	Node category	Node states
1	Victimisation	Criminal attitude	Labelled	Observable	No/Yes
2	Gang member		Labelled	Observable	No/Yes
3	Criminal network		Labelled	Observable	No/Yes
4	Criminal family background		Labelled	Observable	No/Yes
5	Criminal attitude		Labelled	Observable	No/Yes
6	Violent thoughts		Labelled	Observable	No/Yes
7	Compliance with supervision		Labelled	Observable	No/Partial/Yes
8	Negative attitude		Labelled	Observable	No/Partial/Yes
9	Attitude danger level		Labelled	Latent	Low/High
10	Aggression danger level	Criminal attitude/ Personality disorder	Labelled	Latent	Low/High
11	ASPD	Personality disorder	Labelled	Observable	No/Yes
12	BPD		Labelled	Observable	No/Yes
13	Abuse or neglect as a child		Labelled	Observable	No/Yes
14	Anger		Labelled	Observable	No/Yes
15	Impulsivity		Labelled	Observable	No/Partial/Yes
16	PCLR factor 1		$\sim \text{TNormal}(\mu, \sigma^2, 0, 16)$	Observable	0-16
17	PCLR factor 2		$\sim \text{TNormal}(\mu, \sigma^2, 0, 18)$	Observable	0-18
18	PCLR facet 3		$\sim \text{TNormal}(\mu, \sigma^2, 0, 10)$	Observable	0-10
19	Anger management		Labelled	Observable intervention	No/Yes
20	Anger management given failure		Labelled	Latent	No/Yes
21	Anger management post-treatment		Labelled	Latent	No/Yes
22	Intelligence	Socioeconomic factors	Labelled	Observable	Extremely Low/ Borderline/ Low Average/Average/ High Average/Superior
23	Living circumstances		Labelled	Observable	Homeless/Bail Hostel or Shelter/Living alone/ Living with partner/ Living with family or friends/Other
24	Education		Labelled	Observable	No/GCSE or O'Level/ A'Level+/Other
25	Stress		Labelled	Observable	No/Yes
26	Financial difficulties		Labelled	Observable	No/Yes
27	Employment or training		Labelled	Observable	No/Yes
28	Problematic life events		Labelled	Observable	No/Yes

29	Social withdraw		Labelled	Observable	No/Yes
30	Ability to cope		Labelled	Observable	Low/High
31	Domestic stability		Labelled	Observable	Low/High
32	Social protective level		Labelled	Latent	Low/High
33	Symptoms of mental illness		Labelled	Observable	No/Partial/Yes
34	Depressive symptoms		Labelled	Observable	No/Yes
35	Anxiety		Labelled	Observable	No/Yes
36	Thought insertion		Labelled	Observable	No/Yes
37	Hallucinations		Labelled	Observable	No/Yes
38	Strange experiences		Labelled	Observable	No/Yes
39	Paranoid delusions		Labelled	Observable	No/Yes
40	Psychiatric treatment		Labelled	Observable intervention	No/Yes
41	Depressive symptoms post-treatment	Mental illness	Labelled	Latent	No/Yes
42	Anxiety post-treatment		Labelled	Latent	No/Yes
43	Thought insertion post-treatment		Labelled	Latent	No/Yes
44	Hallucinations post-treatment		Labelled	Latent	No/Yes
45	Strange experiences post-treatment		Labelled	Latent	No/Yes
46	Paranoid delusions post-treatment		Labelled	Latent	No/Yes
47	Psychiatric treatment given failure		Labelled	Latent	No/Yes
48	Mental illness danger level		Labelled	Latent	Low/High
49	Cocaine before prison sentence		Labelled	Observable	No/Yes
50	Cannabis before prison sentence		Labelled	Observable	No/Yes
51	Ecstasy before prison sentence		Labelled	Observable	No/Yes
52	Cocaine during prison sentence		Labelled	Observable	No/Yes
53	Cannabis during prison sentence		Labelled	Observable	No/Yes
54	Ecstasy during prison sentence		Labelled	Observable	No/Yes
55	Cocaine after release		Labelled	Observable	No/Yes
56	Cannabis after release		Labelled	Observable	No/Yes
57	Ecstasy after release		Labelled	Observable	No/Yes
58	Hazardous drinking after release		Labelled	Observable	No/Yes
59	Cocaine dependence		Labelled	Observable	No/Yes
60	Cannabis dependence	Substance	Labelled	Observable	No/Yes
61	Ecstasy dependence		Labelled	Observable	No/Yes

62	Alcohol dependence	misuse	Labelled	Observable	No/Yes
63	Drug treatment		Labelled	Observable	No/Yes
64	Alcohol treatment		Labelled	Observable intervention	No/Yes
65	Cocaine post-treatment		Labelled	Latent	No/Yes
66	Cannabis post-treatment		Labelled	Latent	No/Yes
67	Ecstasy post-treatment		Labelled	Latent	No/Yes
68	Hazardous drinking post-treatment		Labelled	Latent	No/Yes
69	Drug treatment given failure		Labelled	Latent	No/Yes
70	Alcohol treatment given failure		Labelled	Latent	No/Yes
71	Any drug dependence		Labelled	Definitional	No/Yes
72	Response given drug dependence		Labelled	Latent	No/Yes
73	Response given alcohol dependence		Labelled	Latent	No/Yes
74	Substance misuse danger level		Labelled	Latent	Low/High
75	Responsiveness to treatment	Treatment responsivity	Labelled	Observable	No/==Partly==/Yes
76	Refuse or fail to attend therapy		Labelled	Latent	No/Yes
77	Prior acquisitive crime convictions	Violence and other static risk factors	Labelled	Observable	0-2/3-12/13+
78	Prior violent convictions		Labelled	Observable	0/1/2-5/6+
79	PCLR total score		Labelled	Observable	0-9/10-16/17-26/27+
80	Age		Labelled	Observable	18-19/20-21/22-25/26-29/30-34/35-39/40-49/50-59/60+
81	Gender		Labelled	Observable	Female/Male
82	Time at risk		$\sim$Uniform(a, b)	Observable	0-5000
83	Violent reconvictions rate (step 1)		$\sim$Beta(a, b)	Latent	0-1
84	Violent reconvictions rate (step 2)		Revised $\sim$Beta	Latent	0-1
85	Violent reconvictions rate (step 3)		Revised $\sim$Beta	Latent	0-1
86	Time since initial release		$\sim$Uniform(a, b)	Observable	0-5000
87	Violent ==re==convictions since initial release		$\sim$Binomial(n, p)	Observable	0-inf
88	Violent convictions		$\sim$Binomial(n, p)	Latent	0-inf
89	Violence		Labelled	Latent	No/Yes

APPENDIX C: Combining danger level indications

Let us assume the combination $\{H, H, H, L\}$ where L is the input for the danger level of substance misuse. Revisiting Figure C.4 above we notice that the substance misuse danger level is low (i.e. High$= 0\%$) when we observe N (i.e. no substance use) and high (i.e. High$= 100\%$) when we observe AD (i.e. both alcohol and drug use), whereas for combinations of A and D the danger level is uncertain. Hence, by providing the prior information of combination $\{H, H, H, \neg H\}$ for combination $\{H, H, H, L\}$, the model considers all the combinations between N, A and D (i.e. $\neg High$) iteratively, instead of simply N (i.e. *Low*), against the other three component danger levels.

This sub-optimal approach was only introduced due to insufficient number of instances in our dataset; it can be safely ignored for datasets with sufficiently larger number of instances. It should also be noted that while the naive Bayesian classification could have also been introduced to effectively deal with the insufficient sample size, it was considered inappropriate (due to its naive independence assumptions) for this case, since we were only interested in modelling the violence rate based on the combinations of those danger levels.

Table C.1. Danger level combinations provided for the $\sim Beta(\alpha, \beta)$ distribution *Violence reconvictions rate(step 1)*.

Combinations assumed by the model	Combinations provided to the model
L,L,L,L	L,L,L,L
H,L,L,L	H,¬H, ¬H, ¬H
L,H,L,L	¬H,H, ¬H, ¬H
L,L,H,L	¬H, ¬H,H, ¬H
L,L,L,H	¬H, ¬H, ¬H,H
H,H,L,L	H,H, ¬H, ¬H
H,L,H,L	H, ¬H,H, ¬H
H,L,L,H	H, ¬H, ¬H,H
L,H,H,L	¬H,H,H, ¬H
L,H,L,H	¬H,H, ¬H,H
L,L,H,H	¬H, ¬H,H,H
H,H,H,L	H,H,H, ¬H
H,H,L,H	H,H, ¬H,H
H,L,H,H	H, ¬H,H,H
L,H,H,H	¬H,H,H,H
H,H,H,H	H,H,H,H

REFERENCES

AgenaRisk. (2002). Bayesian Network and Simulation Software for Risk Analysis and Decision Support. Retrieved February 5, 2015, from http://www.agenarisk.com

Agnew, R. (1992). Foundation for a General Strain Theory. *Criminology* 30(1), 47-87

American Psychiatric Association (APA). (2013*). Diagnostic and Statistical Manual of Mental Disorders: 5th edition (DSM-5)*. Arlington. VA: Amer Psychiatric Pub Incorporated.

Boles, S. M., & Miotto, K. (2003). Substance abuse and violence: A review of the literature. *Aggression and Violent Behavior*, 8(2): 155-174.

Boylan, J. F., Kavanagh, B. P., & Armitage, J. (2011). Randomised controlled trials: important but overrated? *Journal of the Royal College of Physicians*, 41: 126–31.

Cartwright, N., & Munro, E. (2010). The limitations of randomized controlled trials in predicting effectiveness. *Journal of Evaluation in Clinical Practice,* 16: 260-266.

Coid, J.W., Yang, M., Ullrich, S., Zhang, T., Sizmur, S., Roberts, C., & Rogers, R. D. (2009). Gender differences in structured risk assessment: Comparing the accuracy of five instruments. *Journal of Consulting and Clinical Psychology*, 77(2): 337-348.

Coid, J. W., & Ullrich, S. (2010). Antisocial personality disorder is on a continuum with psychopathy. *Comprehensive Psychiatry*, Vol. 51, Iss. 4: 426-433.

Coid, J. W., Yang, M., Ullrich, S., Zhang, T., Sizmur, S., Farrington, D., et al. (2011). Most items in structured risk assessment instruments do not predict violence. *The Journal of Forensic Psychiatry & Psychology, 22*(1), 3-21.

Coid, J., Ullrich, S., & Kallis, C. (2013). Predicting future violence among individuals with psychopathy. *British Journal of Psychiatry*, 203: 387–388. doi: 10.1192/bjp.bp.112.118471.

Coid, J. W., Ullrich, S., Kallis, C., Keers, R., Barker, D., Cowden, F., et al. (2013). The relationship between delusions and violence: findings from the East London first episode psychosis study. *JAMA Psychiatry, 70*(5), 465-471.

Coid, J., Freestone, M. & Ullrich, S. (2012) Subtypes of psychopathy in the British household population: findings from the national household survey of psychiatric morbidity. *Social Psychiatry and Psychiatric Epidemiology*, 47(6): 879-891.

Coid, J. W., Ullrich, S., Kallis, C., Freestone, M., Gonzalez, R., Bui, L., Constantinou, A. C., Fenton, N., Marsh, W., Yang, M., DeStavola, B., Hu, J., Shaw, J., Doyle, M., Archer-Power, L., Davoren, M., Osumili, B., McCrone, P., Barrett, K., Hindle, D., Bebbington P. (2014). Improving Risk Management in Mental Health Services. *British National Institute for Health Research (NIHR)*, July 2014.

Constantinou, A. C., Fenton, N. E. & Neil, M. (2012). pi-football: A Bayesian network model for forecasting Association Football match outcomes. *Knowledge-Based Systems*, 36: 322, 339.

Constantinou, A. C., Fenton, N. E. & Neil, M. (2013). Profiting from an inefficient Association Football gambling market: Prediction, Risk and Uncertainty using Bayesian networks. *Knowledge-Based Systems,* 60-86.

Constantinou, A. C., Fenton, N. E., & Pollock, L. J. H. (2014). Bayesian networks for unbiased assessment of referee bias in Association Football. Journal of Psychology of Sport and Exercise, Vol. 15, 5: 538-547.

de Campos, L. M., Fernández-Luna, J. M., & Huete, J. F. (2004). Bayesian networks and information retrieval: an introduction to the special issue. *Information Processing & Management*, 40 (5): 727–733.

de Melo, A. C. V., & Sanchez, A. J. (2008) Software maintenance project delays prediction using Bayesian Networks. *Expert Systems with Applications*, 34: 908-919.

de Vogel, V., de Vries Robbé, M., de Ruiter, C., & Bouman, Y.H.A. (2011). Assessing protective factors in forensic psychiatric practice. Introducing the SAPROF. *International Journal of Forensic Mental Health*, 10, 171-177.

de Vries Robbé, M., de Vogel, V., & de Spa, E. (2011). Protective factors for violence risk in forensic psychiatric patients. A retrospective validation study of the SAPROF. *International Journal of Forensic Mental Health*, 10, 178-186.

Davis. W.M. (1996) Psychopharmacologic violence associated with cocaine abuse: kindling of a limbic dyscontrol syndrome? *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 20 (8) pp. 1273–1300

Di Pietro, L., Guglielmetti Mugion, R., Musella, F., Renzi, M. F., & Vicard, P. (2015). Reconciling internal and external performance in a holistic approach: A Bayesian network model in higher education. *Expert Systems with Applications*, 42: 2691-2702.

Díez,, F. J., Mira, J., Iturralde, E., & Zubillaga, S. (1997). DIAVAL, a Bayesian expert system for echocardiography. *Artificial Intelligence in Medicine*, 10 (1): 59–73.

Douglas, K. S., Yeomans, M. & Boer, D. P. (2005). Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behaviour*, 32: 479-510.

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20 (Version 3): Assessing risk of violence – User guide*. Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

Doyle, M., Archer-Power, L., Coid, J., Kallis, C., Ullrich, S. & Shaw, J. (2014) Predicting Post-Discharge Community Violence in England and Wales Using the HCR-20v3. *International Journal of Forensic Mental Health*, 13: 140-147. DOI: 10.1080/14999013.2014.906517

Elbogen, E. B., & Johnson, S. C. (2009). The intricate link between violence and mental disorder: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*, 66(2): 152-161.

Fazel, S., Singh, J.P., & Grann, M. (2012) Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: systematic review and meta-analysis. *BMJ*, 345. doi: http://dx.doi.org/10.1136/bmj.e4692

Fenton, N.E. & Neil, M. (2011). Avoiding Legal Fallacies in Practice Using Bayesian Networks. *Australian Journal of Legal Philosophy*, 36: 114-151.

Fenton, N.E. & Neil, M. (2012). *Risk Assessment and Decision Analysis with Bayesian Networks*. CRC Press.

Fenton, N. (2014). Effective Bayesian Modelling with Knowledge Before Data. Retrieved April 29, 2015, from BAYES_KNOWLEDGE: https://www.eecs.qmul.ac.uk/~norman/projects/B_Knowledge.html

Fenton, N. (2015). The problem with big data and machine learning. Retrieved April 29, 2015 from Probability and Risk: http://probabilityandlaw.blogspot.co.uk/2015/03/the-problem-with-big-data-and-machine.html

Fenton, N. E., Lagnado D., & Neil M. (2013). A General Structure for Legal Arguments Using Bayesian Networks. *Cognitive Science,* 37: 61-102.

Fox, A. & Freestone, M (2008) An evaluation of the Good Lives and Discipline (GLAD) Programme at a UK DSPD Unit. *Prison Service Journal*, 151

Freestone, M., Howard, R., Coid, J.W. & Ullrich, S. (2013) Adult antisocial syndrome with comorbid borderline pathology: association with antisocial and violent outcomes. *Personality and Mental Health,* 7(1): 11-21.

Friedman, N.F., Linial, M., Nachman, I., & Pe'er, D. (2000). Using Bayesian Networks to Analyze Expression Data. *Journal of Computational Biology*, 7 (3/4): 601–620.

Goldstein, P.J. (1985). The Drugs/Violence Nexus: A tripartite conceptual framework. *Journal of Drug Issues*, (39): 493-506.

Grann, M., Langstrom, N., Tengstrom, A., & Kuligren, G. (1999). Psychopathy (PCL-R) Predicts Violent Recidivism Among Criminal Offenders with Personality Disorders in Sweden. *Law and Human Behavior*, 23: 2, 205-218.

Hanley, J.A. & McNeil, B.J. (1982a). Maximum attainable discrimination and the utilization of radiologic examinations. *Journal of Chronic Disorders* 35: 601–611.

Hanley, J.A. & McNeil B.J. (1982b). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143: 29–36.

Hanson, R. K., & Harris, A. J. (2000). Where Should We Intervene?: Dynamic Predictors of Sexual Offence Recidivism. *Criminal Justice and Behavior*, 27(6): 6-34.

Harris, G.T., Rice, M.E., & Quinsey, V.L. (1993). Violent recidivism of mentally disordered offenders: the development of a statistical prediction instrument. *Criminal Justice and Behavior* 20: 315–335.

Hare, R. D. (2003). *Manual for the Hare Psychopathy Checklist-Revised: 2nd Edition*. Toronto, Canada: Multi-Health Systems.

Heckerman, D., Breese, J., & Nathwani, B. (1992). Toward normative expert systems I: the PATHFINDER project. *Methods of Information Medicine*, 31: 90–105.

Heckerman D., Mamdani, A. & Wellman, M. (1995). Real-world applications of Bayesian networks. *Communications of the ACM*, vol. 38, no. 3, pp. 25-26.

Hohenner, M., Wachsmuth, S., & Sagerer, G. (2005). Modelling expertise for structure elucidation in organic chemistry using Bayesian networks, *Knowledge-Based Systems*, 18 (4–5): 207–215.

Jensen, F. V. (1996). *An introduction to Bayesian networks*. New York Springer.

Jiang, X., Neapolitan, R.E., Barmada, M.M., & Visweswaran, S. (2011). Learning Genetic Epistasis using Bayesian Network Scoring Criteria. *BMC Bioinformatics*, 12: 89.

Johnson, J.G., Cohen, P., Brown, J., Smailes, E.M. & Bernstein, D.P. (1999) Childhood Maltreatment Increases Risk for Personality Disorders During Early Adulthood. *Archives of General Psychiatry* 56 (7): 600-606.

Lauritzen, S. L. (1995). The EM algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis*, 19: 191-201.

Lee, E., Park, Y., & Shin, J. G. (2009). Large engineering project risk management using a Bayesian belief network. Expert Systems with Applications, 36: 5880-5887.

Lee, K. C., & Park, B. (2010). A general Bayesian network approach to analyzing online game item values and its influence on consumer satisfaction and purchase intention. Communications in Computer and Information Science, 114: 53-62.

Lobo, J. M., Jimenez-Valverde, A., & Real, R. (2007). AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography, 1-7.

Kroner, C., Stadtland, C., Eidt, M. & Nedopil, N. (2007). The validity of the Violence Risk Appraisal Guide (VRAG) in predicting criminal recidivism. *Criminal Behaviour and Mental Health*, 17: 89−100.

McGorry, P. D. (2013), The next stage for diagnosis: validity through utility. *World Psychiatry*, 12: 213–215.

McNeil D, Gregory A, Lam J, Binder, R.L. & Sullivan, G.R (2003) Utility of decision support tools for assessing acute risk of violence. *Journal of Consulting and Clinical Psychology* 71:945—95

McNeil, D., Eisner, J.P. & Binder, R.L. (2000) The Relationship Between Command Hallucinations and Violence. *Psychiatric Services,* 51(10): 1288-1292.

Mears, D. P. & Cochran, J. C. (2013). What is the effect of IQ on offending? *Criminal Justice and Behavior*, doi: 10.1177/0093854813485736.

Monahan, J. (1984). The prediction of violent behaviour: Toward a second generation of theory and policy. *The American Journal of Psychiatry*, 141, 10-15.

Neil, M., Marquez, D. & Fenton, N. E. (2010). Improved Reliability Modeling using Bayesian Networks and Dynamic Discretization. *Reliability Engineering & System Safety*, 95(4), 412-425.

Nikovski, D. (2000). Constructing Bayesian networks for medical diagnosis from incomplete and partially correct statistics. *IEEE Transaction on Knowledge and Data Engineering*, 12, 4: 509-516.

Olver, M., Lewis, K. & Wong, S.C. (2013) Risk reduction treatment of high-risk psychopathic offenders: the relationship of psychopathy and treatment change to violent recidivism. *Journal of Personality Disorders*

Pearl, J. (1988). *Probilistic reasoning in intelligent systems: networks of plausible inference*. San Mateo, California, Morgan Kaufmann Publishers.

Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge University Press.

Pearl, J. (2009). *Causality: Models, Reasoning and Inference, 2nd Edition*. CUP.

Perkusich, M., Soares, G., Almeida, H., & Perkusich, A. (2015). A procedure to detect problems of processes in software development projects using Bayesian networks. Expert Systems with Applications, 42: 437-450.

Pourret, O., Naim, P., & Marcot, B. (2008). *Bayesian Networks: A Practical Guide to Applications*. Chichester, UK: Wiley.

Quinsey, V.L. (1998) *Violent Offenders – Appraising and Managing Risk*. Washington, DC: American Psychological Association.

Rice, M. E. & Harris, G. T. (1995). Violent recidivism: assessing predictive validity. *Journal of Consulting and Clinical Psychology* 63: 737–748.

Rice, M. E., & Harris, G. T. (2005). Comparing Effect Sizes in Follow-Up Studies: ROC Areas, Cohen's *d*, and *r*. Law and Human Behavior, 29: 5, 615-620

Sakellaropoulos, G. C., & Nikiforidis, G. C. (1999). Development of a Bayesian Network for the prognosis of head injuries using graphical model selection techniques. *Methods Inf Med*, 38(1): 37-42.

Sansone, R. A., Songer, D. A., & Gaither, G. A. (2000). Medically self-harming behavior and its relationship to borderline personality among psychiatric inpatients. *The Journal of Nervous and Mental Disease*, 188: 384-386.

Singh JP, 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 Reviews,* 31(3):499-513. doi: 10.1016/j.cpr.2010.11.009

Singh, J. P. (2013). Predictive Validity Performance Indicators in Violence Risk Assessment: A Methodological Primer. *Behavioral Sciences and the Law*, 31: 8-22.

Singh JP, Grann M, Fazel S (2013) Authorship Bias in Violence Risk Assessment? A Systematic Review and Meta-Analysis. *PLoS ONE* 8(9): e72484. doi:10.1371/journal.pone.0072484

Skeem, J., Monahan, J., & Mulvey, E. (2002).Psychopathy, treatment involvement, and subsequent violence among civil psychiatric patients. *Law and Human Behavior,* 26, 577-603

Skeem, J. & Mulvey, P. (2001) Psychopathy and community violence among civil psychiatric patients: Results from the MacArthur Violence Risk Assessment Study. *Journal of Consulting & Clinical Psychology,* 69, 358-374

Smith, G. C., & Pell, J. P. (2003). Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised control trials. *British Medical Journal*, 327: 1459-1461.

Spiegelhalter, D. J., Abrams, K. R., & Myles, J. P. (2004). *Bayesian Approaches to Clinical Trials and Health-care Evaluation*. John Wiley and Sons.

Spielberger, C. D., & Sydeman, S. J. (1994). State-Trait Anxiety Inventory and State-Trait Anger Expression Inventory. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment*. (pp. 292-321). Hillsdale, NJ: Erlbaum.

Stadtland, C., Kleindienst, N., Kröner, C., Eidt, M. & Nedopil, N. (2005). Psychopathic traits and risk of criminal recidivism in offenders with and without mental disorders. *International Journal of Forensic Psychiatry* 4: 89–97.

Stadtland, C. & Nedopil, N. (2004). Vergleichende Anwendung heutiger Prognoseinstrumente zur Vorhersage krimineller Rückfälle bei psychiatrisch begutachteten Probanden. *Monatsschrift für Kriminologie und Strafrechtsreform* 2: 77–85.

Steadman, H., Mulvey, E., Monahan, J., Robbins, P., Appelbaum, P., Grisso, T., Roth, L., & Silver, E. (1998). Violence by people discharged from acute psychiatric inpatient facilities and by others in the same neighborhoods. *Archives of General Psychiatry*, 55, 393-401.

Tarantola, C., Vicard, P., & Ntzoufras, I. (2012). Monitoring and improving Greek banking services using Bayesian Networks: An analysis of mystery shopping data. *Expert Systems with Applications*, 39: 10103-10111.

Taroni, F., Aitken, C., Garbolino, P., & Biedermann, A. (2014). Bayesian Networks and Probabilistic Inference in Forensic Science (2nd ed.). Chichester, UK: John Wiley & Sons.

Troquete, N.A., van den Brink, R.H., Beintema, H., Mulder, T., van Os, T.W., Schoevers, R.A., & Wiersma, D. (2014) Predictive Validity of the Short-Term Assessment of Risk and Treatability for Violent Behavior in Outpatient Forensic Psychiatric Patients. *Psychological Assessment*, Dec 15 2014. doi: 10.1037/a0038270

Ullrich, S., & Coid, J. W. (2011). Protective factors for violence among released prisoners - Effects over time and interactions with static risk. *Journal of Consulting Clinical Psychology*, 79(3): 381-390.

Van Dorn, R., Volavka, J., & Johnson, N. (2012). Mental disorder and violence: is there a relationship beyond substance use? *Social Psychiatry and Psychiatry Epidemiology*, 47(3): 487-503.

Webster, C.D., Douglas, K.S., Eaves, D.E. & Hart, S.D. (1997) *HCR-20: Assessing Risk for Violence, Version 2*. Vancouver, BC: Simon Fraser University.

Witt, K., van Dorn, R. & Fazel, S. (2013) Risk Factors for Violence in Psychosis: Systematic Review and Meta-Regression Analysis of 110 Studies. *PLoS ONE*, 8(2), doi:10.1371/journal.pone.0055942

Wong, S. & Gordon, A. (2003) *Violence Risk Scale (VRS) Version 2*. Saskatoon, CA: University of Saskatoon.

Yang, M., Wong, S.C., & Coid, J. (2010) The efficacy of violence prediction: a meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, 136 (5): 740-767. doi: 10.1037/a0020473.

Yet, B., Perkins, Z., Marsh, W., & Fenton, N. (2011). Towards a Method of Building Causal Bayesian Networks for Prognostic Decision Support. *Probabilistic Problem Solving in Biomedicine*, Vol. 107.

Zhang, L., Wu, X., Ding, L., Skibniewski, M. J., & Yan, Y. (2013). Decision support analysis for safety control in complex project environments based on Bayesian Networks. *Expert Systems with Applications*, 40: 4273-4282.