Expert Systems with Applications 42 (2015) 7511–7529

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Risk assessment and risk management of violent reoffending among prisoners

Anthony Costa Constantinou [a,*], Mark Freestone [b], William Marsh [a], Norman Fenton [a], Jeremy Coid [b]

[a] Risk and Information Management Research Group, School of Electronic Engineering and Computer Science, Queen Mary, University of London, London E1 4NS, UK
[b] 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 EC1A 7BE, UK

ARTICLE INFO

Article history:
Available online 28 May 2015

Keywords:
Bayesian networks
Belief networks
Causal intervention
Forensic medicine
Mental health
Released prisoners
Violent offence

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.

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 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, Gregory, Lam, Binder, & Sullivan, 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, Yeomans, & Boer, 2005). Accurate prediction for violence, even from the same data, can be heavily influenced by the analytical method (Elbogen & Johnson, 2009; Van Dorn, Volavka, & Johnson, 2012), suggesting that the true underlying causes of violence are yet to be fully understood.

* Corresponding author.
E-mail addresses: anthony@constantinou.info (A.C. Constantinou), m.c.freestone@qmul.ac.uk (M. Freestone), d.w.r.marsh@qmul.ac.uk (W. Marsh), n.fenton@qmul.ac.uk (N. Fenton), j.w.coid@qmul.ac.uk (J. Coid).

http://dx.doi.org/10.1016/j.eswa.2015.05.025
0957-4174/

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 (Douglas, Hart, Webster, & Belfrage, 2013; Webster, Douglas, Eaves, & Hart, 1997) 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, 2011) and that, with some offender populations, predictive efficacy is no better than chance (Coid, Ullrich, & Kallis, 2013a).

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 (Heckerman, Mamdani, & Wellman, 1995; Jensen, 1996; Pearl, 1988). 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, Linial, Nachman, & Pe'er, 2000; Hohenner, Wachsmuth, & Sagerer, 2005; Jiang, Neapolitan, Barmada, & Visweswaran, 2011), gaming (Lee & Park, 2010), computer science and artificial intelligence (de Campos, Fernández-Luna, & Huete, 2004; Fenton & Neil, 2012; Pourret, Naim, & Marcot, 2008), medicine (Heckerman, Breese, & Nathwani, 1992; Diez et al., 1997; Nikovski 2000), and law (Fenton, Lagnado, & Neil, 2013; Fenton & Neil, 2011; 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, Park, & Shin, 2009), systematic development of causal interventional systems for prognostic decision support (Yet, Perkins, Marsh, & Fenton, 2011), qualitative examination and evaluation of service offered by the loan departments of Greek Banks (Tarantola, Vicard, & Ntzoufras, 2012), safety control decision support in dynamic complex project environments (Zhang, Wu, Ding, Skibniewski, & Yan, 2013), football match prediction (Constantinou, Fenton, & Neil, 2012; Constantinou, Fenton, & Neil, 2013) and inference of referee bias (Constantinou, Fenton, & Pollock, 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, Guglielmetti Mugion, Musella, Renzi, & Vicard, 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. 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