

This is a pre-print version of the following publication:

Van Ginneken, E.F.J.C. (2019). The use of risk assessment in sentencing. In J.W. de Keijser, J.V. Roberts, & J. Ryberg (eds.), *Predictive sentencing: Normative and empirical perspectives* (pp. 9-32). Oxford: Hart Publishing.

The Use of Risk Assessment in Sentencing

Esther F.J.C. van Ginneken

A Brief History of Risk Assessment in Sentencing

Over the years, risk, as the likelihood of future criminal behaviour, has played a role in sentencing in different forms and at different stages in the criminal justice process. A general shift can be observed from a welfare-oriented approach to a neoliberal approach, with a declining role for (clinical) experts and a shift of responsibilities from state authorities to the individual (Garland 2001; Feeley and Simon 1992). However, it would be amiss to paint this development with a broad brush only, as there is substantial variation in practice across jurisdictions. This chapter discusses these variations in current and (recent) past practices of risk-based sentencing and also gives an overview of the main debates around risk assessment. Many of these debates will be examined in more detail in the chapters that follow.

The development of risk assessment instruments over time can be aligned with changes in their function in the criminal justice process. The previous century witnessed the popularity and subsequent decline of the rehabilitative ideal, which was followed by a greater prominence of retributive (deserved punishment) and crime control concerns. Today, risk assessment may be said to serve the primary function of efficient offender management, although there is also still attention for rehabilitative, retributive and incapacitative goals.

First Generation: Clinical Judgement and Penal Welfarism

In the heyday of the rehabilitative ideal (roughly the 1950s and 1960s, beginning its descent in the 1970s), it was generally believed that the causes of criminality could be ‘diagnosed’ and ‘treated’ using an individual approach (Phelps 2011). Consequently, many offenders received indeterminate sentences that were terminated upon the apparent rehabilitation (and thus low risk) of the individual. This put a great deal of sentencing discretion at the back-door with parole boards, as judges would only impose a sentencing range or maximum sentence. These judgements by parole boards about rehabilitation (i.e. future dangerousness) were normally informed by clinical assessments of dangerousness, which were not restricted by rules or guidelines about the use of information to reach decisions. Unstructured clinical risk assessments are considered the first generation of risk instruments and they have been

criticised for their lack of reliability and predictive validity (Hanson and Morton-Bourgon 2009; Monahan 1981).

Second Generation: Actuarial Assessment and Selective Incapacitation

In the 1970s and 1980s, dissatisfaction with the rehabilitative potential of prison programmes as well as critiques on indeterminate sentencing contributed to a shift towards determinate sentencing on the basis of retributive and crime control considerations. As part of a large research programme funded by the US National Institute of Justice, a 1982 RAND study claimed that crime could be reduced through selective incapacitation of offenders who are most likely to reoffend (Greenwood and Abrahamse 1982). The study's authors had identified correlates of self-reported offending frequency through a survey of prisoners convicted of robbery and burglary and advocated the use of an actuarial tool that used these correlates, to identify high-rate offenders (for a discussion of methodological and ethical problems, see Auerhahn 1999). The idea of selective incapacitation chimes with the notion that a high proportion of crimes are committed by a small proportion of the most prolific offenders (Wolfgang, Figlio and Sellin 1972). However, even selective incapacitation may incur higher costs than the benefits gained from the crimes it prevents (Blokland and Nieuwebeerta 2007).

A defining characteristic of second generation risk assessment instruments is that they only include static factors: these factors are fixed and unamenable to intervention (e.g. age at first arrest and number of prior convictions). These instruments are actuarial risk assessments in the sense that they calculate risk based on a formula, which may be a regression equation or simple adding of points, and can combine more information than the typical human expert (but see Dressel and Farid 2018). A widely used second generation instrument is the Static-99 for predicting sexual offending. This instrument consists of ten items, including age, offence history, victim typology and whether the offender lives with a partner. A total risk score is calculated by adding points for each risk factor, where a total score of 6 or more (range 0-12) is labelled 'high risk'. The initial validation study found a violent recidivism rate of 59% after 15 years for sex offenders who fell in this high risk category (Hanson and Thornton 1999).

One of the prime criticisms launched at the early actuarial tools was that they had a high false-positive rate, i.e. non-recidivists classified as high risk. To some extent, this was remedied by later actuarial tools that used a probabilistic model to estimate the likelihood of future offending. Actuarial risk assessments compare an individual's characteristics to a reference group and calculated risk scores reflect the degree of similarity between an

individual and a group: a high-risk offender basically shares many characteristics with past recidivists and a low-risk offender shares many characteristics with past non-recidivists. While this conveys a level of uncertainty, these risk scores still inform decisions that have the same consequences as false positives and false negatives associated with dichotomous categorisation.

In the 1990s there was growing attention to danger and risk in society more generally (Giddens 1990; Beck 1992), which manifested itself in the criminal justice domain as a preoccupation with risk assessment and management (Feeley and Simon 1992). Yet, the ‘promise’ of actuarial justice and risk assessment to enhance administrative efficiency (Feeley and Simon 1992; 1994) did perhaps not materialise to the extent envisioned until much more recently (Rothschild-Elyassi, Koehler and Simon forth). The 1990s bore witness to rising crime rates, the introduction of tougher sentences, as well as rapidly growing prison populations. The new generation of risk assessment challenged the dominant focus on punitive responses and incapacitation to some extent by a greater focus on treatment potential.

Third and Fourth Generation: Risk-Needs Assessment and Offender Management

The Risk-Needs-Responsivity (RNR) model (Andrews and Bonta 2010) forms the theoretical foundation from which third and fourth generation of structured professional risk assessments have been developed. Third generation risk assessment instruments include static and dynamic factors, where dynamic factors (sometimes called ‘needs’) can change over time or through intervention. This new generation of instruments targets some of the shortcomings of first and second generation risk assessments, which were less reliable (particularly professional judgements) and, in the case of second generation instruments, did not allow for improvements (i.e. decreases) in risk, because they were based on static factors.

The RNR model outlined the principles for successful interventions in combination with risk assessment. First, the level of treatment should be proportionate to the risk of reoffending (risk principle); this means that offenders with a low risk of reoffending are not unnecessarily targeted for intervention. Secondly, interventions should target criminogenic needs, or dynamic factors that are correlated with recidivism (needs principle). Here, Andrews and Bonta (2010) distinguished between criminogenic needs (e.g. pro-criminal attitudes and substance abuse) and non-criminogenic needs (e.g. self-esteem and major mental disorder). Thirdly, interventions should incorporate effective cognitive social learning strategies and tailor them to a person’s learning style. Briefly, it could be said that risk refers to *who* should

be treated, needs to *what* should be treated, and responsivity to *how* to treat. The RNR-based risk assessment instruments are thus used not only to predict risk, but also to assist in reducing it.

Structured professional risk assessments are a combination of clinical judgement and actuarial risk assessment, although they are often carried out by non-clinical professionals (e.g. probation officers). They provide guidance on which risk factors to take into consideration and how to score them, but also allow a degree of professional discretion in the assessment and weighing of certain indicators of risk (and even overrides in the final recommendation). These structured professional risk assessments are often also used to keep a record of progression throughout the completion of a sentence. The development of third generation instruments into fourth generation instruments is characterised by the combination of risk assessment with offender management, and the incorporation of the responsivity principle. Fourth generation instruments are designed to guide supervision, intervention and release decisions from intake through case closure (Andrews, Bonta and Wormith 2006).

A typical example of a third generation instrument is the Level of Service Inventory-Revised (LSI-R), which is used on a large scale in Canada and the US, and was developed and tested by Andrews and Bonta (1995). It comprises of 54 items across 10 domains: criminal history, education/employment, financial, family/marital, accommodation, leisure/recreation, companions, alcohol/drug problems, emotional/personal, and attitudes/orientation. Similar instruments are used in England and Wales (Offender Assessment System, OASys) and other European countries (e.g. Risk Assessment Scales, RISC, in the Netherlands). The LSI-R was later revised into a fourth generation instrument, the Level of Service/Case Management Inventory (LS/CMI; Andrews, Bonta and Wormith 2004), which – among other changes – includes additional non-scored sections that provide qualitative information that may be relevant for offender supervision and treatment; it further allows the test administrator to designate a subcomponent as ‘strength’, which could also inform an offender’s case plan.

The LSI-R has strong advocates as well as fierce critics. There are numerous studies in support of the utility of the LSI-R in predicting recidivism (Gendreau, Little and Goggin 1996; Vose, Cullen and Smith 2008). Nevertheless, reliability and validity of lengthy instruments such as the LSI-R have been questioned; only few of the LSI-R items appear to predict recidivism and there are issues with inter-rater disagreement on ratings of items and risk level (Austin et al. 2003; Dowdy, Lacy and Unnithan 2001). Static predictors are more consistently scored and have higher predictive value. This has prompted suggestions to

separate risk predictions from needs assessment and case planning (Baird 2009; Caudy, Durso and Taxman 2013; *cf* Labrecque et al. 2014).

The developments sketched above cannot simply be characterised as an increase in punitiveness: on the one hand, risk assessments are used to identify high-risk offenders and subject them to greater control and often long-term incarceration, on the other hand, they are also used to identify low-risk offenders and divert them from prison. It is recognised that imposing control and interventions on low-risk offenders may actually have adverse effects. Furthermore, later generation instruments explicitly regard the offender as capable of change and are used to inform ‘treatment’ (perhaps not the most fitting term for the dominant cognitive behavioural interventions on offer, see Duguid 2000: 197). It may therefore be more appropriate to speak of a hybrid model of penology (see for example Hannah-Moffat 2005; O’Malley 1999), which combines different aims of sentencing including rehabilitation, rather than of a ‘new penology’ of actuarial justice, as characterised by Feeley and Simon (1992; 1994).

Issues around the Use of Risk Assessment in Sentencing

The use of risk assessment in sentencing has been criticised on various grounds, including empirical concerns about validity and normative concerns about fairness. This section will briefly discuss the main debates, many of which will be reviewed in more detail in following chapters.

Normative Concerns

One the most prominent and elaborately discussed criticisms of risk-based sentencing is its violation of retributive norms of proportionality. Just deserts theory holds that the severity of a punishment should correspond to the seriousness of the offence, which comprises harm and culpability (Von Hirsch 1976). Retributive sentencing is thus backward-looking. Risk assessment brings concerns about the future into calculations of sentence severity, which means that offenders who have committed crimes of equal seriousness may receive differential sentences, based on future risk. Although, arguably, even within a retributive framework it is possible to envisage differential sentences for equally serious crimes, responsive to risk (Ryberg, Chapter 4 in this volume). There is debate among retributivist scholars on whether it is justified to consider prior convictions in sentencing (Lee 2010; Von

Hirsch 2010; Bagaric 2001). Risk-based sentencing is more compatible with utilitarian approaches to punishment, in particular to advance the aim of crime control through incapacitation of high-risk offenders or risk reduction by targeted intervention.

A related concern is that the use of risk assessment in sentencing in effect holds offenders responsible for factors that are static (e.g. age and gender) or are often outside their direct control (e.g. employment). Tonry (2014: 171) notes, ‘there is something fundamentally unethical or immoral about apportioning punishments or other intrusions on liberty on the basis of ascribed characteristics for which no coherent argument can be made that offenders bear personal responsibility for them.’ To the extent that variable risk factors are a choice, it is still problematic to attach penal consequences to lawful life choices.

So far, however, the use of risk assessment in sentencing decisions has survived legal challenge. The fourth generation instrument Correctional Offender Management Profiles for Alternative Sanctions (COMPAS) is used across the United States, to inform sentencing decisions as well as offender management. It assesses risk on the basis of 137 factors and was developed by a commercial organisation, Northpointe, which makes the algorithm proprietary. In *State v Loomis* (2016)¹, Loomis argued that the proprietary nature of COMPAS prevented him from challenging the accuracy and validity of the instrument, and that using the instrument in sentencing is an improper form of gendered assessment. The Wisconsin Supreme Court ruled that the court’s use of the COMPAS risk score as an element (rather than the sole factor) in determining Loomis’s sentence did not violate his due process rights to be sentenced individually and using accurate information, and by taking gender and race into account; a petition to the US Supreme Court to challenge this decision was denied.²

Predictive Validity and the Meaning of ‘Risk’

Another key concern with the use of risk assessment in sentencing is the imperfect predictive validity (see also Fazel, Chapter 11 in this volume). Risk is by definition uncertain: it may or may not materialise into harm. This is reflected in the probability scores that are generated by modern risk assessments, which convey information about the average recidivism rate for a

¹ *State v Loomis*, 881 N.W.2d 749 (Wis. 2016)

² COMPAS does not use race as an explicit risk factor. The instrument received criticism after journalistic platform ProPublica published an article showing that black defendants were more likely to receive false positive (high risk) scores than white defendants, while white defendants were more likely to receive false negative (low risk) scores (Angwin et al. 2016). Overall accuracy was the same for white and black defendants. Others have criticised these findings for their limited assessment of fairness, as other criteria showed no racial disparities (Flores, Bechtel and Lowenkamp 2016; Chouldechova 2017).

group of offenders who share the same characteristics (included in the instrument) as the individual subjected to the assessment. Yet, often, these raw probability scores are reduced to simplistic categories of low, moderate or high risk in communication to the courts. Predictive validity refers to a risk assessment instrument's ability to correctly assess the likelihood of (serious) reoffending.

Any decision based on a risk score has the potential to make a wrong assumption about actual recidivism: the so called false positives (i.e. a prediction of reoffending when this would not have occurred) and false negatives (i.e. a prediction of no reoffending when this will occur). The error rate of risk assessment is substantial; while the most commonly used risk assessment instruments appear fairly accurate at predicting low risk of offending, they have a high rate of false positives (Fazel et al. 2012). The most commonly used measure of predictive accuracy for risk assessment instruments is the area under the Receiver Operating Characteristic (ROC) curve. The area under the curve (AUC) can be interpreted as the probability that a randomly selected recidivist has a higher risk score than a randomly selected non-recidivist. Most risk instruments have an AUC value between 0.66 and 0.78 (Singh, Grann and Fazel 2011), which means that they have moderate discriminatory power.³ Even given the AUC value of a particular instrument, the likelihood of false positives and false negatives varies depending on the cut-off score that is used to make decisions about risk: the higher the cut-off score on an instrument for labelling an offender as high-risk, the greater the specificity (i.e. positive predictive value), but the lower the sensitivity of detecting true positives. In practice, users of risk assessment instruments are likely to cast a wider net, which classifies a relatively substantial proportion of offenders as high risk, in order to maximise sensitivity and avoid that any high risk offenders go undetected; this also means that, inadvertently, a considerable proportion of offenders is inaccurately labelled as high risk (i.e. false positives. Of course, this may have consequences at the sentencing stage (see also Hester, Chapter 12 in this volume).

A further consideration is the reliability of the information that is fed into the measure: 'no risk-assessment device can be better than the data from which it is constructed' (Gottfredson and Moriarty 2006: 183). Instruments with low levels of transparency into how scores are calculated can be criticised because the validity cannot be checked or challenged; this is one of the greatest problems with machine learning approaches to risk assessment (see also 'The Future of Risk Assessment' below and Hannah-Moffat, Chapter 10 in this volume).

³ A systematic review on the reporting of predictive validity in violence risk assessment studies found that AUC results were often misinterpreted (Singh, Desmarais and Van Dorn 2013).

Other validity problems are that risk estimates do not include an assessment of the potential effect of specific sentences (or interventions as part of a sentence) on recidivism and that the sample used for developing an instrument may not be representative for the population on which it is used (Starr 2014). Fazel et al. (2012: 5) conclude, ‘risk assessment tools in their current form can only be used to roughly classify individuals at the group level, and not to safely determine criminal prognosis in an individual case.’ The use of group data to make predictions about individual cases is not only empirically flawed, but also ethically dubious (Hart, Michie and Cooke 2007; Netter 2007; Starr 2014; *cf.* Skeem and Monahan 2011), as it holds individuals responsible for group-based tendencies.

Finally, one should also consider the type of outcome that is predicted and its practical relevance. ‘Risk’ can have a variety of meanings: it can mean the likelihood of parole violation, re-arrest, a criminal charge, reconviction, re-imprisonment or any specific type or seriousness of re-offending. Instruments that were developed using official statistics (particularly arrests and parole violations) partly reflect police and probation practices rather than actual (harmful) behaviour (Harcourt 2007). Regardless, validation studies use a wide variety of outcome measures so it needs to be questioned to what extent they measure the intended construct. It is also important to understand that someone who falls into a ‘high risk’ category does not necessarily pose a great risk of *serious* offending (or causing serious harm); in other words, risk does not equal dangerousness. Many conceive of ‘dangerousness’ as a function of risk (in terms of likelihood) and harm.⁴ The meaning of dangerousness may vary across jurisdictions and is fluid over time, depending partly on public and political concerns about crime at any given time (Pratt 1995). Post 1970, dangerousness has come to be understood as mainly (risk of) violent and sexual offending, which is also reflected in current legislation in England and Wales, discussed under ‘Assessments of Dangerousness and Indeterminate Sentences’ below. Given the low prevalence rate of serious violent offending, it is inherently difficult to predict its occurrence (but much easier to predict its non-occurrence, see Hester, Chapter 12 in this volume). Many instruments therefore are validated to predict general or any violent offending (including threats or attempts of violence without bodily harm) (Singh, Grann and Fazel 2011; Yang, Wong and Coid 2010). Overall, then, ‘high risk’ can be easily misinterpreted if decision-makers are not given information on error rates and the nature of the outcome that is predicted.

⁴ But see Slobogin (2012), who defines dangerousness as simply the likelihood of reoffending.

Discrimination and Responsibilisation

Some argue in favour of risk assessment for reasons of fair treatment: it can – to some extent – eliminate judicial bias, enhance transparency in sentencing, and prevent unnecessary intervention and the infliction of harm on low-risk offenders. However, there are also concerns that risk-based sentencing results in sentencing disparity and reproduces or even exacerbates social inequality. There are various possible explanatory mechanisms for such an effect. First, risk assessment instruments may have mixed predictive validity for different groups, such as men and women, or individuals with a different ethnic background (Skeem, Monahan and Lowenkamp 2016; Raynor and Lewis 2011; *cf.* Skeem and Lowenkamp 2016). Risk and need factors may vary across groups, operate differently and be experienced differently (e.g. Hannah-Moffat 2016; Van Voorhis et al. 2010). Secondly, the predictive fairness of risk-based sentencing may be compromised because factors that are correlated with recidivism are also correlated with ethnicity and social class (e.g. employment status, marital status, education level, and substance abuse), which has to be understood in a socio-economic and historical context (Holtfreter, Reisig and Morash 2004; Van Eijk 2017). For instance, the correlation between criminal history and race may not only reflect differential participation of ethnic minorities in crime, but also differential selection through disproportionately high arrest, prosecution and conviction rates (Harcourt 2007).

Use of marital status, employment, education, family status, and residential stability as factors in prediction instruments systematically disadvantages minority defendants. The social and economic disadvantages that disproportionately afflict blacks and Hispanics in America are partly the products of historic and ongoing discrimination and bias. (Tonry 2014: 173)

If risk assessment is employed to send high-risk offenders to prison, it is likely to simply reinforce destructive cycles of incarceration and associated social disadvantage (Van Eijk 2017; Hannah-Moffat 2016). As a result, the risk assessments will help to sustain the inequalities that are root causes of many forms of delinquent behaviour, including exclusion from society and economic disadvantage.

Finally, risk assessment may be criticised for depoliticising issues of social justice. Forward-looking, risk-based sentencing diverts attention away from underlying causes of crime and instead holds individuals fully responsible not only for past crimes, but also for

their future risk (known as ‘responsibilisation’, which transfers the responsibility for solving problems from state authorities to individuals). In other words, one might easily confuse the causes of crime with the causes of individual differences in crime. In risk assessment, the terms ‘needs’ and ‘dynamic risk’ are used interchangeably, which devalues the importance of non-criminogenic needs (alterable circumstances which are not directly, empirically linked to recidivism). Valid needs, then, are only those that are potential targets of risk reduction, because they have a statistically proven association with offending. While these needs are often connected to structural constraints (including poverty), they are presented as individual deficits, which are also an individual’s responsibility to overcome (Hannah-Moffat 2016; Holtfreter et al. 2004; Van Eijk 2017). This bypasses political debate about what should be done about crime – apart from incarcerating high-risk offenders – and to what extent it is a political and societal responsibility to *collectively* reduce risk factors of crime.

The Role of Risk Assessment in Sentencing

We can roughly distinguish two functions of risk assessment in sentencing: risk assessment can be used either to limit (judicial) discretion, or to inform how discretion in sentencing may be used. Instances of the former function are primarily observed in jurisdictions that make use of sentencing guidelines, which can contain an implicit form of risk assessment in themselves, or alternatively can contain provisions that require judges to make a risk assessment and apply it in a certain way. Risk assessment can also be used to inform judges (as well as parole boards) in more discretionary decisions regarding type and length of sentences. These various uses of risk assessment are discussed in turn, and illustrated with typical examples.

Limiting Discretion: Sentencing Guidelines

Sentencing guidelines have a more or less prescriptive function in sentence calibration, not only in relation to offence seriousness, but also in relation to risk. They can be regarded as a form of risk assessment in themselves as they normally incorporate at least offence history in the recommendation of a sentence (range); they may also – additionally – explicitly require or advise the consideration of risk in sentencing decisions. Most guidelines states in the US employ a sentencing grid, which recommends sentence options based on two criteria: the seriousness of the current offence and an offender’s prior record. The length of a prison sentence can vary greatly depending on the number of prior convictions (Frase and Hester

2015). Sentencing grids, in particular, limit judicial discretion in the sense that prior convictions are uniformly associated with higher recommended sentences; therefore, the presence of this particular risk factor is not subject to much interpretation (see also Roberts and Frase, Chapter 9 in this volume).⁵ It has been argued that such sentencing grids function more to standardise judicial practice in line with logics of control and incapacitation than as a form of risk assessment or actuarial justice (Rothschild-Elyassi, Koehler and Simon in press). Matters are more complicated when sentencing guidelines include the recommendation to consider risk assessment in the sentencing decision, primarily in order to divert low-risk offenders from prison.

Risk assessments integrate more information than prior records into a calculation of recidivism risk and judges have the risk scores available in addition to sentencing guidelines. The *2011 Guiding Principles report* by a National Working Group on the use of offender risk and needs assessment information at sentencing outlines how risk assessments should be used, with the ultimate aim to reduce recidivism in a cost-effective way (Casey, Warren and Elek 2011). Importantly, the guiding principles explicitly separate the punitive aim of sentencing from crime reduction aims and state that risk and needs assessments (RNAs) should only be used for the latter purpose. That is, the outcomes of RNAs should not be used to make sentences more or less punitive, which is in line with the judgement in *Malenchik v Indiana* (2010)⁶, in which the Supreme Court of Indiana found that risk assessment scores should not be interpreted as mitigating or aggravating factors in determination of the length of a sentence. Rather, the Court recommended that RNAs should be used primarily for deciding whether to suspend (part of) a sentence and for assigning programmes, treatment or specific requirements. Therefore, the guiding principles only apply to offenders who are probation-eligible according to state sentencing guidelines. While (additional) probation or treatment requirements may not be *intentionally* punitive, it should be kept in mind that they may certainly be experienced as such (see, for example, Durnescu 2011; Van Ginneken and Hayes 2017).

The state of Virginia has gone further than most other states in its use of risk assessment at the sentencing stage, which they introduced to offset the anticipated increase in the occupation of prison beds resulting from truth-in-sentencing legislation and the abolition of parole. The Virginia Criminal Sentencing Commission (2011) developed a Nonviolent

⁵ In *United States v Booker*, 543 U.S. 220 (2005) the US Supreme Court ruled that federal mandatory sentencing guidelines were unconstitutional, which consequently made sentencing guidelines advisory.

⁶ *Malenchik v Indiana*, 928 N.E.2d 564 (Ind. 2010)

Risk Assessment (NVRA) that has been used since 2002 to identify drug and property offenders with the lowest risk of committing further crimes and recommend them for diversion from prison. A few interesting observations can be made about Virginia's experience with risk assessment at sentencing. First, the instrument calculates a score primarily based on offence history and current offence characteristics. The only offender characteristics are age (lower age yields higher risk score) and gender (males receive higher scores). Employment and marital status were included initially, but excluded from the most recent versions due to difficulty with verification and negligible loss of predictive validity (Ostrom and Kauder 2012). The NVRA is therefore a rather straightforward actuarial instrument. Secondly, there is substantial discrepancy in the extent to which recommendations from the NVRA are followed by judges. Under half of the eligible offenders who scored 'low risk' actually received an alternative sanction (Garrett, Jakubow and Monahan 2018), including diversion from prison to jail. The same study found that there was substantial variation across the circuit court districts in the imposition of alternative sentences. The majority of judges rated the availability of alternative interventions within their jurisdiction as inadequate (Monahan, Metz and Garrett 2018). Thus, while Virginia achieved its aim of diverting at least 25% of the lowest risk drug and property offenders from prison, there is no consistent application (Garrett, Jakubow and Monahan 2018).

Increasingly, other states also report risk and needs information at the sentencing stage, typically as part of the pre-sentence investigation (PSI) report. Most states that conduct RNA prior to sentencing use state-specific instruments, but the LSI-R, LS/CMI and COMPAS are also used (Elek, Warren and Casey 2015; Vera Institute of Justice 2011). Most of these instruments are much more elaborate than Virginia's NVRA. For example, the Indiana Risk Assessment System requires probation officers to score criminal history; education, employment and financial situation; family and social support; neighbourhood problems; substance abuse; peer associations; and criminal attitudes and behavioural patterns (Elek, Warren and Casey 2015). Overall, the US does not attach binding consequences to specific levels of risk, but sentencing guidelines recommend diversionary sanctions for offenders with low risk scores, while (lengthy) prison sentences are reserved for offenders with high risk scores. Yet, despite clear directions on how judges should interpret and apply risk scores, the research discussed above shows that judges are still likely to use their discretion to deviate from the recommendations on using risk scores.

The serious consequences of being deemed a ‘high-risk offender’ are most aptly illustrated with a discussion of the practice of indeterminate sentencing in England and Wales. Sentencing guidelines in England and Wales are developed by the Sentencing Council and must be followed by judges, unless it is not in the interests of justice to do so (Coroners and Justice Act 2009, s.125(1)). They tend to offer a fairly wide range of penalty options, the choice of which is guided by considerations of culpability, harm, aggravating factors (including relevant previous offences), and mitigating factors. Of particular interest regarding these guidelines, is the assessment of dangerousness that judges *must* make for serious specified offences (Criminal Justice Act 2003, s.224), which can result in an indeterminate sentence of life imprisonment.

In cases where judges have to make a discretionary assessment of whether to impose a life sentence or extended sentence for certain violent or sexual offences, they do so under the dangerousness provisions outlined in the Criminal Justice Act (CJA) 2003 (s.229). In effect, this appraisal of dangerousness constitutes a risk assessment. A life sentence must be imposed by the court if it is available for the committed offence and ‘the court considers that the seriousness of the offence, or of the offence and one or more offences associated with it, is such as to justify the imposition of a sentence of imprisonment for life’ (s.225(2)). The judge assesses dangerousness by consideration of: the nature and circumstances of the offence, the nature and circumstances of previous offences, any pattern of behaviour including the aforementioned offence(s), and any information about the offender (s.229). This information should normally be included in a pre-sentence report prepared by the National Probation Service (NPS). Such a report also includes an analysis of the likelihood of reoffending and risk of harm.⁷ The courts may also consider psychiatric reports and other expert opinion.

England and Wales have – for Western European standards – an exceptionally high proportion of offenders sentenced to indeterminate prison sentences. At the end of 2017, 7,144 prisoners were serving a life sentence, which accounts for 11% of their prison population (Ministry of Justice 2018).⁸ Life imprisonment is the mandatory sentence for offenders over 21 convicted of murder, it is a discretionary maximum sentence for offenders

⁷ See PI 04/2016 ‘Determining Pre Sentence Reports’. In practice, this involves a combination of actuarial and structured risk assessments (e.g. Offender Group Reconviction Scores, OGRS, and Offender Assessment System, OASys). The NPS also allocates cases based on risk for future management purposes, using the (actuarial) Risk of Serious Recidivism tool (Robinson 2017).

⁸ Included in this number are 60 prisoners with a whole life order, which means they will not be eligible for parole at any point, with the exception of release on compassionate grounds by the Home Secretary.

over 21 convicted of serious offences such as manslaughter, rape, and armed robbery, and an automatic sentence for a second listed offence under conditions outlined in s.122 in the Legal Aid, Sentencing and Punishment of Offenders Act (LASPOA) 2012 (inserted as s.224A into CJA 2003). At sentencing, the judge also imposes a minimum term that the offender has to serve in prison before they can be considered for parole. At the end of 2017, there were 2,014 prisoners (28%) with a life sentence with an expired tariff (Ministry of Justice 2018). Released life prisoners remain on licence for the remainder of their natural life, which means they can be recalled to prison if this is deemed necessary to protect the public.

Imprisonment for Public Protection: A Failed Experiment

The costs of guideline-mandated indeterminate sentences for high-risk offenders are high, which was particularly evident from the brief experiment in England and Wales with another type of indeterminate sentence specifically for ‘dangerous offenders’: Imprisonment for Public Protection (IPP). The IPP sentence was introduced with the CJA 2003 (s.225) and abolished with LASPOA 2012 (s.123). In December 2017, there were still 3,029 prisoners serving an IPP sentence, most of whom (87%) with an expired tariff (Ministry of Justice 2018).

The IPP experiment aptly demonstrates the difficulties with operationalising risk. Initially, there was no minimum offence seriousness threshold for imposing an IPP sentence (apart from commission of a listed offence) and courts had to make the judgement whether an offender was ‘dangerous’, i.e. ‘a significant risk to members of the public of serious harm occasioned by the commission by him [sic] of further such offences’ (CJA 2003, s. 229(1)). The original legislation contained a presumption of dangerousness, stating that ‘the court must assume that there is such a risk’ if an offender had committed one of the listed offences. In their risk assessment, courts were obliged to take into account information about the offence and were allowed to also take into account information about the offender and patterns of behaviour including the offence. The courts would normally have information from the structured risk assessment OASys at their disposal, but there were no guidelines or case law on how to derive conclusions about dangerousness from these scores (Ashworth and Zedner 2014: 126). Furthermore, the validity of OASys scores was found questionable by a report from the Chief Inspector of Prisons and Probation (HMCIP 2008). As a consequence of the wide net of risk cast with the CJA 2003, offenders were given relatively low commensurate tariffs but nevertheless faced extended, indeterminate sentences (Jacobson and Hough 2010).

The Criminal Justice & Immigration Act 2008 sought to remedy some of the problems with proportionality as well as expansion of the prison population by removing the presumption of dangerousness and introducing a minimum tariff of two years before an IPP sentence could be imposed. LASPOA 2012 abolished the IPP sentence altogether and reduced judicial discretion by introducing the automatic life sentence for a second serious violent or sexual offence.

Informing Discretion

In jurisdictions with greater judicial discretion at sentencing, the use of risk assessment is less transparent. In civil law countries, judges are not normally bound by guidelines, but instead have to abide by statutory maximum (and sometimes minimum) sentences. As a result, there are no instructions on applying risk scores in a specified manner in sentencing; so if and how outcomes from risk assessments should be used for sentencing purposes is left to judicial discretion. Nevertheless, judges often consider risk assessments in some form in their sentencing decisions. While specific preventive policies and measures betray the political preoccupation with risk, we can observe a persistence of rehabilitative ideals among judges and criminal justice professionals (McNeill et al. 2009).

Common practice is to convey judgements of risk in the form of pre-sentence (investigation) reports, which are presented in court. Pre-sentence investigation (PSI) reports are widely used and tend to convey information about risk and criminogenic needs of offenders. In these reports, risk assessments tend to be based on structured and clinical risk assessments, which may incorporate actuarial elements. Research on PSI reports in Belgium, Sweden and Denmark suggests that professional and clinical judgements are most important in determining risk; there is little evidence of the use of actuarial risk assessment at the sentencing stage (Wandall 2010; Persson and Svensson 2012; Beyens and Scheirs 2010). In Canada, on the other hand, structured actuarial risk assessment is a dominant feature of PSIs (Hannah-Moffat and Maurutto 2010).

Even when risk is clearly communicated at the sentencing stage in PSI reports, it is not necessarily interpreted in the same straightforward manner as is the case with sentencing grids and guidelines (i.e. sentence enhancements for prior records or diversion from prison in the case of low risk). In the Netherlands, for example, judges appear to respond to risk with a more rehabilitative than incapacitative approach, which may even suggest that criminogenic needs are interpreted as mitigating rather than aggravating factors. A quasi-experimental study (Van Wingerden, Van Wilsem and Moerings 2014) found that Dutch judges were,

contrary to expectations, more likely to sentence offenders identified as ‘high-risk’ by a structured risk-based PSI report to less controlling sentences (i.e. a suspended sentence without special conditions) than offenders with the same risk who did not have a PSI report at the time of sentencing. In line with expectations, offenders identified as low-risk prior to sentencing were more likely to receive diverting types of sentences (i.e. suspended sentence or community sentence) than their counterparts without risk assessment.

Beyond Punishment: Preventive Detention and Safety Measures

There are provisions for risk-based sentencing that avoid the engagement with normative issues of punishment through deployment of what Hart (1968) called a *definitional stop*: if it is not called punishment, it does not have to be justified as such. This is problematic because moral questions and issues cannot be simply ‘defined out of existence’ (Kleinig 1973: 13). Many continental European jurisdictions have risk-related sanctions or measures that can be imposed at the sentencing stage in addition to or instead of purely offence-related penalties (Van der Wolf and Herzog-Evans 2015). The integration of such measures in penal law makes them technically distinct from safety measures created within civil law and they also differ from provisions for safety measures after a sentence has been served: but both of these suffer from some of the same problems discussed below, particularly with respect to the proportionality of suffering inflicted. For example, many US states have the option to detain serious sexual offenders indefinitely under *civil* commitment laws, when they have a ‘mental abnormality’ that predisposes them to sexual violence (*Kansas v Hendricks* [1997]).⁹ The penal safety measures imposed at sentencing require a *prospective* judgement of risk rather than an assessment of *current* risk to determine safety of release and therefore do not take into account a possible reduction of risk over time.

As these sanctions are not technically punishments – even though they are imposed only when an offence has been committed – they bypass retributive concerns of proportionality; instead they are intended primarily for public protection. This means that individuals can be deprived of their liberty for much longer than would normally be considered proportionate to the seriousness of the offence. These types of measures should not be confused with those for offenders who are not held criminally responsible for their acts (not discussed in this chapter). De Keijser (2011) has argued elsewhere that the difference

⁹ *Kansas v Hendricks*, 521 U.S. 346 (1997)

between punishment and measure is merely semantic in relation to the infliction of suffering: while punishment involves the intended infliction of suffering, any suffering inflicted by a measure is unintended. Regardless, the deprivation of liberty – unlikely to be the only deprivation resulting from a measure – is inevitably accompanied by pain. There are many examples of such measures which are preventive in name and punitive in terms of subjective experience (see Ashworth and Zedner [2014] for examples and critiques). The section below discusses two examples from the Netherlands of penal measures involving liberty deprivation, which provides an interesting case study given that the measures target two very different categories of high-risk offenders: (1) serious violent and sexual offenders and (2) persistent (low-level) offenders.

The first measure is *terbeschikkingstelling* (TBS, entrustment order). Of particular interest for the purpose of this discussion is the version of TBS which can be imposed in addition to a regular term of imprisonment, meaning the offender is held (at least partially) criminally responsible.¹⁰ TBS extends a sentence beyond the severity proportionate to the offence, although a TBS order will be served in a TBS-institution rather than regular prison; this is not the case in France, for example, where safety detention is executed in prison (Van der Wolf and Herzog-Evans 2015). The TBS order incapacitates the offender for its duration, but also entails treatment. The offender may decline treatment, but this would considerably reduce their chances of release. TBS may only be imposed for crimes with a maximum statutory punishment of at least four years, when a psychiatric disorder (to some extent) led to the commission of the offence, and when there is a risk of recidivism. Its duration is limited to a maximum of four years for non-violent offenders, although in practice the vast majority of offenders given a TBS order are sentenced for violent offences. For violent offenders, TBS can be extended indefinitely with two-year increments upon review. The decision to impose a TBS order relies primarily on a psychiatric report provided to the court, which is still heavily based on clinical diagnosis and risk assessment. Risk assessments during a TBS order, for example to determine temporary release, increasingly make use of structured risk assessment tools, such as the Dutch HKT-R and the international HCR20V3.

The second measure is *Inrichting Stelselmatige Daders* (ISD, Institution for Persistent Offenders), which is a determinate custodial measure of a maximum of two years, imposed instead of a regular prison sentence. It targets offenders who frequently commit less serious crimes, and are seen to be responsible for a disproportionately high number of offences. ISD-

¹⁰ TBS can also be imposed on offenders not held criminally responsible due to mental illness at the time of the offence.

prisoners are incarcerated on dedicated ISD-wings, within regular prison facilities. The ISD measure is best understood as a measure to protect society of nuisance, rather than danger. Similar to TBS, ISD imposes a sanction that is more severe than would be justified solely on retributive grounds. It can be imposed at the request of the public prosecutor, if a defendant has been convicted at least three times in the preceding five years. The judicial decision is further informed by a pre-sentence report, including risk assessment, provided by the Probation Service. Apart from the public protection rationale underlying ISD, it is also thought that the duration of the measure is potentially more effective at addressing problems of addiction than repetitive short sentences; there is some evidence that the ISD-measure is effective at reducing crime and recidivism (Tollenaar, Van der Laan and Van der Heijden 2014; 2018).

The Future of Risk Assessment

The use of risk assessment in sentencing is not a new development, but its function and form have changed in the last few decades. While risk assessment is still used at the back-door of sentencing to determine safety of release, it is increasingly used at the front-door of sentencing. Here, it has two primary (and to some extent, mutually exclusive) functions: (1) to limit judicial discretion by attaching prescribed or advised sentences to offenders' risk and (2) to inform judicial discretion, mainly in the form of pre-sentence investigation reports containing risk/needs assessments. There is no uniformity in how risk assessment is used and even the consequences attached to it. It can serve to identify low-risk offenders who can be safely diverted from prison; to identify high-risk offenders who should be given indeterminate sentences or safety measures; or to determine the most appropriate type of sentence or conditions. Actuarial assessment has gained some ground but the remaining influence of professional judgement should not be underestimated. Nevertheless, policies and legislation for dangerous offenders are illustrative of increasing risk aversion and, in effect, limit judicial discretion. This final section reflects on two recent developments that may signal further changes in the shape and methods of risk-based sentencing: (1) the criminalisation of risk and (2) technological advancements, including machine learning and neurological risk assessment.

While the preceding discussion has focused on risk assessment as part of the sentencing process and in determining release decisions, there is a trend towards an assumption of risk with certain activities. This concerns primarily pre-inchoate or preparatory offences, thought to indicate risk of further offending that has the potential to cause great

harm. Pre-inchoate offences mainly concern activities associated with terrorism (e.g. encouragement of terrorism; engaging in any conduct in preparation for acts of terrorism; see the UK's Terrorism Act 2006).¹¹ Characteristic of these offences is that activities that are not harmful *an sich* are seen – in combination with circumstances and a degree of interpretation about intent – as carrying enough risk for harm as to warrant severe punishment (the maximum penalty for the preparatory activities under the Terrorism Act 2006 is life imprisonment). This fits with a shift from post-crime to pre-crime (Zedner 2007), whereby criminal justice resources and responses are increasingly oriented towards the pre-emption of crime, rather than the reaction to crime. Such a criminalisation of risk obscures assumptions about risk factors and how they are related to actual harm: the problem of imperfect predictive validity is completely disregarded when behaviour indicative of a high risk of harm is itself illegal.

Technological advancements, similarly, raise new questions about fairness, even though they may improve the accuracy of predictions. There are two particularly noteworthy developments in relation to the technology of risk prediction: machine learning algorithms and neurological assessments. Machine learning techniques have enabled the search for instruments that can detect more complex patterns than traditional regression-based instruments (Berk and Bleich 2013; Berk et al. 2009). This means that the algorithm that turns input information into a risk score is highly complex (a 'black box'), and consequently cannot be used to inform interventions and monitor progress in the way that third and fourth generation instruments can be used. There is not yet agreement on whether machine learning techniques are more accurate at predicting recidivism (Tollenaar and Van der Heijden 2013; Berk and Bleich 2013; for a discussion of the impact of machine learning on the predictive fairness of risk assessment, see Hannah-Moffat, Chapter 10 in this volume).

A second development is the potential of neurological assessment to inform assessments of dangerousness. While the study of brain function in relation to offending is still in its infancy and has not yet widely penetrated the court room (although, see Catley and Claydon 2016; Denno 2015; Gaudet and Marchant 2016), recent brain imaging developments could potentially be used for the purpose of 'neuropsyching' violence and offending more generally (Aggarwal 2009; Gkotsi and Gasser 2016; Glenn and Raine 2014; Nadelhoffer and

¹¹ Although, as Ashworth and Zedner (2014) note, the offence of 'meeting a child following sexual grooming' in English law can be seen as a pre-inchoate offence. It penalises a person who intentionally meets or travels to meet a minor with the intention of doing something that would constitute the commission of an offence (Sexual Offences Act 2003, s.15).

Sinnot-Armstrong 2012). For example, reduced functioning in the frontal lobe is associated with antisocial and violent behaviour (Yang and Raine 2009). While most evidence is cross-sectional in nature, there are some studies that have successfully used neuro-assessment to predict re-arrest and violent offending (Aharoni et al. 2013; Pardini et al. 2014). Nevertheless, the brain is a product of complex biosocial interactions and there is a danger of reductionism and stigmatisation if we rely on ‘brain data’ to predict future offending (Gkotsi and Gasser 2016). So far, however, research suggests that neurological evidence of brain abnormalities has mostly been interpreted as diminished culpability in sentencing decisions, rather than as an indication of future risk (Catley and Claydon 2016; Denno 2015). This, again, suggests that the *discretionary* use of risk assessment does not at this time have a uniformly ‘aggravating’ impact on sentencing decisions. Yet, given the volatile nature and political sensitivity of ‘risk’, the future may not be kind to such discretion.

References

- Aggarwal, NK (2009) ‘Neuroimaging, culture, and forensic psychiatry’ *Journal of the American Academy of Psychiatry and the Law*, 37, 239.
- Aharoni, E, Vincent, GM, Harenski, CL, Calhoun, VD, Sinnott-Armstrong, W, Gazzaniga, M S and Kiehl, KA (2013) ‘Neuroprediction of future rearrest’ *Proceedings of the National Academy of Sciences*, 110, 6223.
- Andrews, DA and Bonta, J (1995) *The level of supervision inventory - Revised* (Toronto, Multi-Health Systems).
- _____ (2010) *The psychology of criminal conduct*, 5th edn (Cincinnati, OH, Anderson Publishing Company).
- Andrews, DA, Bonta, J and Wormith, JS (2004) *The Level of Service/Case Management Inventory (LS/CMI)* (Toronto, Multi-Health Systems).
- _____ (2006) ‘The recent past and near future of risk and/or need assessment’ *Crime & Delinquency*, 52, 7.
- Angwin, J, Larson, J, Mattu, S and Kirchner, L (2016) ‘Machine bias: There’s software used across the country to predict future criminals. And it’s biased against Blacks’ *ProPublica*, available online at www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [last accessed at 16 March 2018].
- Ashworth, A and Zedner, L (2014) *Preventive justice* (Oxford, Oxford University Press).

- Auerhahn, K (1999) 'Selective incapacitation and the problem of prediction' *Criminology*, 37, 703.
- Austin, J, Coleman, D, Peyton, J and Johnson, KD (2003) *Reliability and validity study of the LSI-R risk assessment instrument* (Washington, DC, Institute on Crime, Justice, and Corrections at the George Washington University).
- Bagaric, M (2001) *Punishment and sentencing: A rational approach* (Sydney, Cavendish).
- Baird, C (2009) *A question of the evidence: A critique of risk assessment models used in the justice system* (Madison, WI, National Council on Crime and Delinquency).
- Beck, U (1992) *Risk society: Towards a new modernity* (London, Sage Publications).
- Berk, RA and Bleich, J (2013) 'Statistical procedures for forecasting criminal behavior' *Criminology & Public Policy*, 12, 513.
- Berk, R, Sherman, L, Barnes, G, Kurtz, E and Ahlman, L (2009) 'Forecasting murder within a population of probationers and parolees: A high stakes application of statistical learning' *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172, 191.
- Beyens, K and Scheirs, V (2010) 'Encounters of a different kind: Social enquiry and sentencing in Belgium' *Punishment & Society*, 12, 309.
- Blokland, AA and Nieuwbeerta, P (2007) 'Selectively incapacitating frequent offenders: Costs and benefits of various penal scenarios' *Journal of Quantitative Criminology*, 23, 327.
- Casey, PM, Warren, RK and Elek JK (2011) *Using offender risk and needs assessment information at sentencing: Guidance for courts from a national working group* (Williamsburg, VA, National Center for State Courts).
- Catley, P and Claydon, L (2016) 'The use of neuroscientific evidence in the courtroom by those accused of criminal offenses in England and Wales' *Journal of Law and the Biosciences*, 2, 510.
- Caudy, MS, Durso, JM and Taxman, FS (2013) 'How well do dynamic needs predict recidivism? Implications for risk assessment and risk reduction' *Journal of Criminal Justice*, 41, 458.
- Chouldechova, A (2017) 'Fair prediction with disparate impact: A study of bias in recidivism prediction instruments' *Big data*, 5, 153.
- De Keijser, JW (2011) 'Never mind the pain; It's a measure! Justifying measures as part of the Dutch bifurcated system of sanctions' in M Tonry (ed), *Retributivism has a past: Has it a future?* (Oxford, Oxford University Press).

- Denno, DW (2015) 'The myth of the double-edged sword: An empirical study of neuroscience evidence in criminal cases' *Boston College Law Review*, 56, 493.
- Dowdy, ER, Lacy, MG and Unnithan, NP (2002) 'Correctional prediction and the Level of Supervision Inventory' *Journal of Criminal Justice*, 30, 29.
- Dressel, J and Farid, H (2018) 'The accuracy, fairness, and limits of predicting recidivism' *Science Advances*, 4, eaao5580.
- Duguid, S (2000) *Can prisons work? The prisoner as object and subject in modern corrections* (Toronto, University of Toronto Press).
- Durnescu, I (2011) 'Pains of probation: Effective practice and human rights' *International Journal of Offender Therapy and Comparative Criminology*, 55, 530.
- Elek, JK, Warren, RK and Casey, PM (2015) *Using Risk and Needs Assessment Information at Sentencing: Observations from Ten Jurisdictions* (Williamsburg, VA, National Center for State Courts).
- Fazel, S, Singh, JP, Doll, H and 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' *British Medical Journal*, 345, e4692.
- Feeley, MM and Simon, J (1992) 'The new penology: Notes on the emerging strategy of corrections and its implications' *Criminology*, 30, 449.
- _____ (1994) 'Actuarial justice: The emerging new criminal law' in D Nelken (ed), *The futures of criminology* (London, Sage).
- Flores, AW, Bechtel, K and Lowenkamp, CT (2016) 'False positives, false negatives, and false analyses: A rejoinder to "Machine bias: There's software used across the country to predict future criminals. And it's biased against Blacks."' *Federal Probation*, 80, 38.
- Frase, RS and Hester, R (2015) 'Magnitude of criminal history enhancements' in RS Frase et al, *Criminal history enhancements sourcebook* (Minneapolis, MN, Robina Institute of Criminal Law and Criminal Justice).
- Garland, D (2001) *The culture of control* (Oxford, Oxford University Press).
- Garrett, BL, Jakubow, A and Monahan, J (2018) *Nonviolent risk assessment in Virginia sentencing: The sentencing commission data* (University of Virginia School of Law).
- Gaudet, LM and Marchant, GE (2016) 'Under the radar: Neuroimaging evidence in the criminal courtroom' *Drake Law Review*, 64, 577.
- Gendreau, P, Little, T and Goggin, C (1996) 'A meta-analysis of the predictors of adult offender recidivism: What works!' *Criminology*, 34, 575.

- Gkotsi, GM and Gasser, J (2016) 'Neuroscience in forensic psychiatry: From responsibility to dangerousness. Ethical and legal implications of using neuroscience for dangerousness assessments' *International Journal of Law and Psychiatry*, 46, 58.
- Glenn, AL and Raine, A (2014) 'Neurocriminology: Implications for the punishment, prediction and prevention of criminal behaviour' *Nature Reviews Neuroscience*, 15, 54.
- Giddens, A (1990) *The consequences of modernity* (Stanford, CA, Stanford University Press).
- Gottfredson, SD and Moriarty, LJ (2006) 'Statistical risk assessment: Old problems and new applications' *Crime & Delinquency*, 52, 178.
- Greenwood, PW and Abrahamse, AF (1982) *Selective incapacitation* (Santa Monica, CA, Rand Corporation).
- Hannah-Moffat, K (2005) 'Criminogenic needs and the transformative risk subject: Hybridizations of risk/need in penalty' *Punishment & Society*, 7, 29.
- _____ (2016) 'A conceptual kaleidoscope: contemplating 'dynamic structural risk' and an uncoupling of risk from need' *Psychology, Crime & Law*, 22, 33.
- Hannah-Moffat, K and Maurutto, P (2010) 'Re-contextualizing pre-sentence reports: Risk and race' *Punishment & Society*, 12, 262.
- Hanson, RK and Morton-Bourgon, KE (2009) 'The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies' *Psychological Assessment*, 21, 1.
- Hanson, RK and Thornton, D (1999) *Static 99: Improving actuarial risk assessments for sex offenders*, vol 2 (Ottawa, Solicitor General Canada).
- Harcourt, BE (2007) *Against prediction: Profiling, policing and punishing in the actuarial age* (Chicago, IL, University of Chicago Press).
- Hart, HLA, 1968, *Punishment and responsibility* (Oxford, Oxford University Press).
- Hart, SD, Michie, C and Cooke, DJ (2007) 'Precision of actuarial risk assessment instruments: Evaluating the 'margins of error' of group versus individual predictions of violence' *The British Journal of Psychiatry*, 190(Sup. 49), 60.
- HMCIP (2008) *The indeterminate sentence for public protection: A thematic review* (London, HMCIP)
- Holtfreter, K, Reisig, MD and Morash, M (2004) 'Poverty, state capital, and recidivism among women offenders' *Criminology & Public Policy*, 3, 185.
- Jacobson, J and Hough, M (2010) *Unjust deserts: Imprisonment for public protection* (London, Prison Reform Trust).

- Kleinig, J (1973) *Punishment and desert* (The Hague, Martinus Nijhoff).
- Labrecque, RM, Smith, P, Lovins, BK and Latessa, EJ (2014) 'The importance of reassessment: How changes in the LSI-R risk score can improve the prediction of recidivism' *Journal of Offender Rehabilitation*, 53, 116.
- Lee, Y (2010) 'Repeat offenders and the question of desert' in JV Roberts and A von Hirsch (eds), *Previous convictions at sentencing: Theoretical and applied perspectives* (Oxford, Hart Publishing).
- McNeill, F, Burns, N, Halliday, S, Hutton, N and Tata, C. (2009) 'Risk, responsibility and reconfiguration: Penal adaptation and misadaptation' *Punishment & Society*, 11, 419.
- Ministry of Justice (2018) *Offender management statistics quarterly, England and Wales. Quarter: July to September 2017, prison population: 31 December 2017* (London, Ministry of Justice).
- Monahan, J (1981) *The clinical prediction of violent behavior* (Northvale, NJ, Jason J Aronson).
- Monahan, J, Metz AL and Garrett BL (2018) *Nonviolent risk assessment in Virginia sentencing, report 2: A survey of circuit court judges* (University of Virginia School of Law).
- Nadelhoffer, T and Sinnott-Armstrong, W (2012) 'Neurolaw and neuroprediction: Potential promises and perils' *Philosophy Compass*, 7, 631.
- Netter, B (2007) 'Using group statistics to sentence individual criminals: An ethical and statistical critique of the Virginia risk assessment program' *The Journal of Criminal Law and Criminology*, 97, 699.
- O'Malley, P (1999) *The risk society: Implications for justice and beyond* (Victoria, Department of Justice).
- Ostrom, BJ and Kauder, NB (2012) 'The evolution of offender risk assessment in Virginia' *Federal Sentencing Reporter*, 25, 161.
- Pardini, DA, Raine, A, Erickson, K and Loeber, R (2014) 'Lower amygdala volume in men is associated with childhood aggression, early psychopathic traits, and future violence' *Biological Psychiatry*, 75, 73.
- Persson, A and Svensson, K (2012) 'Shades of professionalism: Risk assessment in pre-sentence reports in Sweden' *European Journal of Criminology*, 9, 176.
- Phelps, MS (2011) 'Rehabilitation in the punitive era: The gap between rhetoric and reality in US prison programs' *Law & Society Review*, 45, 33.

- Pratt, J (1995) 'Dangerousness, risk and technologies of power' *Australian & New Zealand Journal of Criminology*, 28, 3.
- Raynor, P and Lewis, S (2011) 'Risk–need assessment, sentencing and minority ethnic offenders in Britain' *British Journal of Social Work*, 41, 1357.
- Robinson, G (2017) 'Stand-down and deliver: Pre-Sentence Reports, quality and the new culture of speed' *Probation Journal*, 64, 337.
- Rothschild-Elyassi, G, Koehler, J and Simon J (in press) 'Actuarial justice' in M Deflem (ed), *The handbook of social control* (Malden, MA, Wiley-Blackwell).
- Singh, JP, Desmarais, SL and Van Dorn, RA (2013) 'Measurement of predictive validity in violence risk assessment studies: A second-order systematic review' *Behavioral Sciences & the Law*, 31, 55.
- Singh, JP, Grann, M and Fazel, S (2011) 'A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants' *Clinical psychology review*, 31, 499.
- Skeem, JL and Lowenkamp, CT (2016) 'Risk, race, and recidivism: Predictive bias and disparate impact' *Criminology*, 54, 680.
- Skeem, JL and Monahan, J (2011) 'Current directions in violence risk assessment' *Current Directions in Psychological Science*, 20, 38.
- Skeem, JL Monahan, J and Lowenkamp, C (2016) 'Gender, risk assessment, and sanctioning: The cost of treating women like men' *Law and Human Behavior*, 40, 580.
- Slobogin, C (2012) 'Risk assessment' in J Petersilia and KR Reitz (eds), *The Oxford Handbook of Sentencing and Corrections* (Oxford, Oxford University Press).
- Starr, SB (2014) 'Evidence-based sentencing and the scientific rationalization of discrimination' *Stanford Law Review*, 66, 803.
- Tollenaar, N and Van der Heijden, PGM (2013) 'Which method predicts recidivism best? A comparison of statistical, machine learning and data mining predictive models' *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176, 565.
- Tollenaar, N, Van der Laan, AM and Van der Heijden, PGM (2014) 'Effectiveness of a prolonged incarceration and rehabilitation measure for high-frequency offenders' *Journal of Experimental Criminology*, 10, 29.
- _____ (2018). 'Correction to: effectiveness of a prolonged incarceration and rehabilitation measure for high-frequency offenders' *Journal of Experimental Criminology*, 14, 121.
- Tonry, M (2014) 'Legal and ethical issues in the prediction of recidivism' *Federal Sentencing Reporter*, 26, 167.

- Van der Wolf, MJ and Herzog-Evans, M (2015) 'Mandatory measures: 'Safety measures'. Supervision and detention of dangerous offenders in France and the Netherlands: A comparative and human rights' perspective' in M Herzog-Evans (ed), *Offender release and supervision: The role of courts and the use of discretion* (Nijmegen, Wolf Legal Publishers).
- Van Eijk, G (2017) 'Socioeconomic marginality in sentencing: The built-in bias in risk assessment tools and the reproduction of social inequality' *Punishment & Society*, 19, 463.
- Van Ginneken, EFJC and Hayes, D (2017) "'Just' punishment? Offenders' views on the meaning and severity of punishment' *Criminology & Criminal Justice*, 17, 62.
- Van Voorhis, P, Wright, EM, Salisbury, E and Bauman, A (2010) 'Women's risk factors and their contributions to existing risk/needs assessment: The current status of a gender-responsive supplement' *Criminal Justice and Behavior*, 37, 261.
- Van Wingerden, S, Van Wilsem, J and Moerings, M (2014) 'Pre-sentence reports and punishment: A quasi-experiment assessing the effects of risk-based pre-sentence reports on sentencing' *European Journal of Criminology*, 11, 723.
- Vera Institute of Justice (2011) *Risk and needs assessments*, Memorandum to the Delaware Justice Reinvestment Task Force, available from: <http://www.ma4jr.org/wp-content/uploads/2014/10/vera-institute-memo-on-risk-assessment-for-delaware-2011.pdf> [last accessed 24 May 2018].
- Virginia Criminal Sentencing Commission (2001) *Annual report* (Richmond, VA, Virginia Criminal Sentencing Commission).
- Von Hirsch, A (1976) *Doing justice: The choice of punishments* (New York, NY, Hill and Wang).
- _____ (2010) 'Proportionality and the progressive loss of mitigation: Some further reflections' in JV Roberts & A von Hirsch (eds), *Previous convictions at sentencing: Theoretical and applied perspectives* (Oxford, Hart Publishing).
- Vose, B, Cullen, FT and Smith, P (2008) 'The empirical status of the Level of Service Inventory' *Federal Probation*, 72, 22.
- Wandall, RH (2010) 'Resisting risk assessment? Pre-sentence reports and individualized sentencing in Denmark' *Punishment & Society*, 12, 329.
- Wolfgang, M, Figlio, R and Sellin, T (1972) *Delinquency in a birth cohort* (Chicago, IL, University of Chicago Press).

- Yang, Y and Raine, A (2009) 'Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: A meta-analysis' *Psychiatry Research: Neuroimaging*, 174, 81.
- Yang, M, Wong, SC and Coid, J (2010) 'The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools' *Psychological Bulletin*, 136, 740.
- Zedner, L (2007) 'Pre-crime and post-criminology?' *Theoretical Criminology*, 11, 261.