# **Sentencing Dangerous Offenders in the Era of Predictive Technologies: new skin, same old snake?**

David Gurnham

# Introduction

Questions about how the state might and ought to respond to former offenders believed to pose a continuing danger are rarely very far from the public agenda.[[1]](#endnote-1) There may be no serious doubt that public protection must be part of that response and has a role to play in sentencing alongside other important aims. However, given the practical and moral implausibility of handing out indeterminate or life sentences to all potential re-offenders, indecision and disagreement persist about precisely what that role is.[[2]](#endnote-2) Technologies that use data for purposes of risk assessment and for predicting re-offending arguably have an important contribution to make on this account, and indeed such technologies already play a role in jurisdictions including England and Wales. For example, in the UK the Offender Assessment System (OASys) has since 2002 been producing algorithmic risk scores by combining various data collected about offenders. The system is now used across the criminal justice system to inform risk-based decisions regarding bail, sentencing and parole as well as offender-management in prisons and in the community.[[3]](#endnote-3) However, public protection and risk considerations also provoke countervailing concerns about ensuring proportionality in sentencing and about preventing unduly draconian, stigmatising and marginalising impacts on particular individuals and communities. These concerns are enduring, and principally philosophical and socio-legal in nature. Indeed, that they are not fundamentally altered by the advent of new and more sophisticated predictive technologies is the chief focus of discussion in this chapter.

The chapter contends therefore that prior to, and more fundamental than, any issues about the transparency,[[4]](#endnote-4) fairness[[5]](#endnote-5) and accountability[[6]](#endnote-6) of predictive technology are questions about whether sentencing informed by risk and future offending can be countenanced *at all* in light of liberal justice values such as desert and moral agency. Liberal notions of just punishment are built upon the principle that the individual offender, as a one time and future potential member of the moral community, should be able to understand his punishment as the communication of moral censure for the injury done by the crime to that community as a whole. This is an essential principle from a liberal retributivist standpoint because it ensures that the offender himself ‘owns’ his punishment: by being made to take responsibility both for the harm he did to the community and for the penitential debt that he consequently and uniquely must repay in order finally to re-join it.[[7]](#endnote-7)

Such a principled approach does make room for future-oriented considerations about the amount of punishment (typically in terms of time to be served in gaol) necessary for the message of censure to be fully understood and to enable reconciliation to take place. However, committing to this principled approach becomes difficult and problematic as soon as we bring to the fore the very different set of priorities represented by the desire to protect the public, and furthermore make the public *feel* protected from the danger of future crimes that convicted offenders *might* commit after their release. As we shall see, these are considerations that pull in entirely different directions: if we are to take seriously the principle of individualised justice as desert in the liberal retributive sense, then we face serious (potentially intractable) difficulties in justifying *any* sort of role for predictive risk profiling and assessment, let alone sentencing based on automated algorithms drawing on big data analytics. In this respect, predictive technologies present us, not with genuinely new problems, but merely a more sophisticated iteration of established *actuarial* risk assessment (ARA) techniques.

This chapter describes some of the reasons why principled and social justice objections to predictive, risk-based sentencing make so elusive any genuinely synthetic resolution or compromise. In this arena, the choice for legal experts is not about whether to be a *facilitator* for, or a *brake* on, the application of predictive technology to sentencing, since both of these roles suggest that any legitimacy problems that arise may be overcome by adjusted expectations or differently framed legal or ethical parameters. The question is rather how it might even be possible to conceive such a thing without seriously undermining fundamental principles of justice and fairness. The chapter leaves the well-known socio-technical issues of transparency, accountability, fairness and accuracy to one side, since these are in any case ably discussed elsewhere in this book. Section 1 begins by describing the relevant legal and technological context, before section 2 moves on to outline the principled liberal philosophical and social justice objections to predictive profiling and risk assessment having any role to play in criminal justice sentencing. The chapter concludes with a brief reflection on what the irresolution of these objections means for productively advancing discussion.

# 2. Sentencing dangerous offenders: the legal and technological context

The Criminal Justice Act (CJA) 2003 provides the legal framework for sentencing dangerous offenders in England and Wales, albeit now in a form revised and amended numerous times since its original inception.[[8]](#endnote-8) Under the current laws the assessment of offender risk makes provision for the use of algorithmic risk assessment systems without explicitly mandating it. Judges sentencing offenders convicted of certain ‘specified offences’ (generally serious offences of a violent or sexual nature) apply statutory extended sentencing provisions ‘for the purpose of protecting members of the public from serious harm’.[[9]](#endnote-9) In designating the dangerous provisions as applying only to certain very *serious* offences but also that any period of extension to the basic sentence reflects *public protection* priorities, the legislation seeks to promote both retributive (desert-based) and incapacitative (risk-based) principles at the same time.

In order to determine the level of potential danger posed by an offender convicted of a ‘specified offence’, the CJA 2003 directs that the court (a) ‘*must* take into account all such information as is available to it about the nature and circumstances of the offence’, (b) ‘*may* take into account any information which is before it about any pattern of behaviour of which [the specified offence] forms part’, and, presumably in order to catch anything that might yet escape notice, also (c) ‘any information about the offender which is before it’.[[10]](#endnote-10) These provisions afford some considerable latitude as to how any information relating to the offender’s risk status may be generated and presented to the court. A psychiatrist may be called upon to give expert evidence after having made a clinical assessment of the offender, or it may be algorithmically produced (OASys, referred to above). The algorithm generates a risk score based on both static variables (age, sex, previous convictions) and dynamic ones (given to change over time).[[11]](#endnote-11) It considers, amongst other things, ‘social and economic factors in relation to the offender including accommodation, employability, education, associates, relationships and drug or alcohol abuse; and the offender's thinking, attitude towards offending and supervision and emotional state.’[[12]](#endnote-12)

In assessing the impact of the use of OASys as a system through which algorithmic prediction is already allowed to contribute to sentencing decisions, there are two things that must be noted at this point. The first of these is that there are reasons to believe that its impact on how sentencing courts go about their business has so far been rather modest and undramatic. In the first place, even where OASys is used to inform a sentencing decision, a clinical expert may in any case be called to give oral testimony to the court as to the correct interpretation of the risk score so produced.[[13]](#endnote-13) Furthermore, research by the Law Commission carried out after the introduction of OASys found that judges in England and Wales seemed still to be highly confident in their own innate, instinctual ability to determine risk – a confidence apparently undimmed by the availability of algorithmic risk assessment.[[14]](#endnote-14) On the other hand however, Hannah-Moffat has published research suggesting (albeit relating to a different jurisdiction) that judges see algorithmically-derived sentencing decisions as a way for them to minimize the risk to themselves of being blamed for being too lenient or too harsh[[15]](#endnote-15) – as a reputational safeguard and for seeming to be responding to the public’s fear of crime.[[16]](#endnote-16)

The second point is that OASys and other systems used in criminal justice systems around the world for assessing offender risk are in some significant ways different from the more truly automated *personalisation* technologies used in marketing and other commercial contexts. The latter class of systems take fuller advantage of ever-increasing computer processing power, the volume of data that can be handled and the variety of applications to which this can be put[[17]](#endnote-17) to scrape massive quantities of data about an individual created by his or her online presence and activity. By contrast, the data that OASys collects is limited to particular pre-selected variables and gathered by old-fashioned methods such as surveys and interviews conducted in prisons. These variables are then fed into and manipulated by an automated algorithm, but (at least as things currently stand) their inclusion is the result of deliberate human design, choices and labour rather than artificial intelligence.

Both of the points referred to in this section (the impact of algorithmic risk assessment on judicial decision-making and the degree to which the systems used in criminal justice truly exploit automation and available data) warrant scrutiny in their own right. Significant change on either of these fronts would indeed take us into new and uncharted territory. However, this chapter is concerned with what automated risk assessment and predictive systems in criminal justice actually look like now, and as we shall see those already in use pose a major challenge to liberal principles of just punishment.

# 3. Old arguments with a new urgency: two principled objections to predictive technology

Having described in broad terms the relevant legal and technical landscape for dangerous offender sentencing in England and Wales, we are in a position to examine how, as regards the most important challenges they pose, predictive systems are actually not really new at all, but merely a more sophisticated iteration of established *actuarial* techniques for assessing risk. In this section therefore, we turn to these hitherto intractable matters relating to the principled basis of risk assessment in criminal justice.

The techniques currently available (and described above in relation to OASys) for automating risk-assessment represent a particular chapter within a larger penological cultural paradigm-shift that took place across a number of jurisdictions (including England and Wales) from the 1970s. This period of criminal justice history saw concerns about punishing and rehabilitating individual offenders giving way at least in part to considerations of how arrays of data might help to develop a more efficient, informed, rational and cost-effective system of managing dangerous populations and public sensitivities towards risk. This paradigm-shift was famously announced by Feeley and Simon as the ‘new penology’, and was made possible by the development of actuarial techniques and technologies for recording and processing population, demographic and socioeconomic data and using it to predict the risk of offending behaviours.[[18]](#endnote-18) In England and Wales as in other jurisdictions around the world, these techniques have long informed risk-based decision-making in a wide variety of contexts including bail decisions, non-custodial dispositions, offender management and parole, decisions to recall to prison an offender out on licence and civil preventive orders as well as pre-sentence reports for sentencing judges.[[19]](#endnote-19)

The use of predictive algorithms in sentencing may represent something new and different in the sense that much greater quantities of data can be processed more quickly than previous technologies allowed, and that the potential for further advances seem now to be more profound as AI systems become more capable and sophisticated. However, these systems are nevertheless fundamentally a species of the already-established practice of actuarial risk assessment (ARA), and thus share both the principled foundations, and also the objections to which that practice has long been exposed.[[20]](#endnote-20) The key philosophical and socio-legal objections to using actuarial or algorithmic techniques to inform risk-based sentencing are that this is incompatible with fundamental principles of just punishment as a response to individual wrongdoing, and that it leads to sentencing decisions for offenders that are arbitrary, marginalising and socially unjust. We turn to those objections now, and to implications these have for the justifiability of a role for predictive technologies.

## 3.1. Public safety versus individual liberty, responsibility and autonomy

Liberal notions of justice in the first place regard punishment as necessary in order to acknowledge that a wrong was committed by an individual offender against the community as a whole. They also recognise and accept that the state has a legitimate interest in protecting its citizens from individuals who it has reason to consider pose a severe public risk *if a way can be found to do so justly*.[[21]](#endnote-21) The ‘if’ is important, because neither the proven justifiability of punishing the offender following conviction, nor the potential risk of his future re-offending provides a justification for punishments that are indeterminate or indefinite. The sentence itself requires justification as a ‘deserved response’ to lawfully proven crime,[[22]](#endnote-22) and for liberal retributivists desert represents the only good and reliable reason for displacing the usual assumptions about a person’s harmlessness and their entitlement to liberty.[[23]](#endnote-23)

If we agree then that public safety and individual liberty are *both* important aims of a liberal criminal justice system, then we must be able to explain how the state can pursue a public safety agenda in a way that does not involve disproportionate treatment of the individual or unjust deprivation of liberty.[[24]](#endnote-24) Proportionate sentencing thus means having regard to the costs that criminal justice measures impose on individual liberty, as well as to the positive value of eventual reconciliation between offender and society.[[25]](#endnote-25) Hence, principles of liberty and proportionality in combination assert themselves as a counterweight to the state’s inclination to ‘err on the safe side’ by keeping the offender in gaol and away from *potential* victims. Principles of liberty and proportionality serve to remind that punishing an offender more than is justified by the seriousness of the crime actually committed imposes *known* burdens and costs upon him, and that by contrast any future offending is always a quantity as yet essentially *unknown* and yet to materialise.[[26]](#endnote-26)

At the level of holding offenders responsible for their wrongdoing, liberal retributive rationales for punishment place considerable weight on the function of punishment as a communicative act that recognises an individual’s moral responsibility and autonomy. As matters of principle derived *a priori* rather than empirically, moral responsibility and autonomy serve to justify putting an individual on trial in the first place. They also give sense to the idea of punishment as a moral message of censure from the community (via the criminal justice state) to the offender.[[27]](#endnote-27) From a liberal standpoint, that punishment can be understood as a sort of communicative act is important because it underlines the idea that the punishment serves a justified and principled end: that the infliction by the offender of an injury against the community is something that *he can be held responsible for*, and that through his punishment he can repay a unique debt that has the power to regrow the bond thus severed between him and the community.

None of this is to deny that there are hoped-for prospective consequential gains to be made during the period of punishment. Punishment that ‘works’ in a liberal sense is one in which the offender is brought to understand the injury he has caused, to accept responsibility for having inflicted it, and to complete his penitential and rehabilitative journey back to society. How much of an offender’s sentence is eventually spent inside a prison and how much of it may be spent outside and subject to parole or licence conditions may be determined in part by considerations such as progress during the term of incarceration. But the key point is that the assessment about how much punishment is necessary *in total* is retrospective (ie desert-based), not prospective (risk-based). Only questions of seriousness and desert for the instant offence may legitimately determine that total measure. Prospective assessments of that total measure of punishment cannot be conceived within the liberal retributivist perspective, since the offender is necessarily unable to take responsibility for an offence he has not (yet) committed and may never commit. In other words, the offender’s established guilt for the offence he was convicted for does not entitle the state to assume that he is *already* guilty of a further, future offence. To say otherwise is to undermine the moral premise of punishment itself: that the offender is to be credited as an autonomous individual capable of *choosing not* to reoffend.[[28]](#endnote-28)

This insistence on a desert-based, retrospective calculation of punishment may not prohibit calculations of risk and predictive technologies for decisions about parole or indeed offender management, but there are clearly some serious problems when it comes to setting an overall sentence. At the most basic level, a liberal retributive approach and a risk-based one look in two very different, indeed opposed, directions. On the one hand, liberal retributivism as a philosophy of justice based on individual moral responsibility assumes as a matter of principle that the future is a fundamentally dark continent. Advocates of risk assessment and re-offending predictions in criminal justice on the other hand deny the unknowability of the future; they instead propose that, with sufficient quantities of suitably high-quality and reliable data, sensible predictions about future offending can be made and punishments (as well as other criminal justice interventions, such as policing) determined at least in part on that basis. As we have seen in the previous section, the ‘dangerousness’ sentencing provisions of the CJA 2003 try to accommodate both of these principles at the same time.

Can this gulf between these approaches to doing justice be bridgeable somehow? In cases involving (say) very great harm and also very little prospect of the offender becoming less of a risk (for example a perpetrator of mass murder who remains convinced of the rightness and necessity of his cause and actions), it would seem that the requirements both of retribution and public protection point in the same direction. In such a case, the problems described above are arguably obviated since a very long sentence would be justifiable by both retributive *and* public protection aims working together, or at least in parallel. This was effectively the reasoning of Lord Steyn in the House of Lords’ landmark ruling in *Hindley* that confirmed that ‘whole life’ tariffs for the most serious cases were lawful and justifiable.[[29]](#endnote-29) At a political level, it is not uncommon to find retribution and public protection fused together as the components necessary to ensure public confidence, for example as in the December 2019 Queen’s speech.[[30]](#endnote-30)

The problem with that sort of approach however is that it is entirely contingent on the facts of particular cases, and provides no assistance in instances in which retributive and public protection priorities pull in *different* directions (such as in the case of the convicted terrorist murderer who genuinely and permanently renounces his extreme cause, or the petty arsonist who, unrelatedly, happens to be a committed (say) anti-Semite and harbours a strong desire to kill Jews).

A more generally applicable (and more explicitly utilitarian) defence of using predictive risk assessment in sentencing to promote public safety from future offences is to argue that in cases involving dangerous offenders, ordinary retributive principles might be displaced for the ‘greater good’, ie in the interests of all those people who have *not* committed any crime and may be threatened by the offender’s release.[[31]](#endnote-31) This ‘choice of principles’ has obvious attractions from a public policy point of view, but it is also problematic from a principled one in that it denies the *a priori* status of the dangerous offender as a responsible moral agent and as such the possibility of his reconciliation with the society he has injured.

## 3.2. Can predictive risk-based sentencing ever satisfy objections about arbitrariness, marginalisation and social injustice?

In addition to the concerns about compatibility with philosophical principles of individual justice, critics of predictive technologies point out that serious social justice problems flow from its reliance on actuarial risk assessment (ARA) techniques. It is true that systems for predicting future potential reoffending rely, not only on information about the offender’s *own* past and present behaviours, attitudes and health (eg past offending, hostility to law enforcement or other people, alcohol or drug-dependency), but also on data about the re-offending rates amongst communities sharing the same demographic attributes and socio-economic status as the offender.[[32]](#endnote-32) The problematic implications for policy and social justice have been explored by other scholars (for example the chapter by Pamela Ugwudike in this book), but this chapter emphasises that the use of demographic information to inform individual punishment is intrinsically morally problematic. In the first place, the use of such data is suggestive, not of *individual* justice, but of a judgement about dangerous populations and communities more generally. For this reason, a sentencing decision so based arguably further undermines the idea of punishment as an expression of censure issued from the community to an individual responsible offender.[[33]](#endnote-33) In the case of both ‘dynamic’ variables (such as drug use, place of residence and attitudes, that the offender might be able to do something to change over time) and static ones (beyond any intervention, eg sex and age),[[34]](#endnote-34) the actuarial approach adopted by OASys and systems like it that calculate risk according to how populations of *other* people with particular characteristics and histories have generally behaved leaves it prone to two particular objections.

The first of these objections is that of *arbitrariness*, since the individual offender could in effect be any number of people who happen to share the same demographic characteristics and historical or current circumstances.[[35]](#endnote-35) The second is that it *exacerbates social inequality*, since an individual offender categorised within a marginalised and vulnerable population is more likely to produce a high risk score than an offender who is not so categorised.[[36]](#endnote-36) This means that a policy of imposing longer sentences on account of a higher risk score will tend also to punish more harshly, for offending of equal seriousness, individuals who are socially and economically disadvantaged. For the same reason, harsher sentences may be expected to follow those with family and mental health-related problems[[37]](#endnote-37) and furthermore carry negative implications about communities that may already be stigmatised or marginalised due to perceptions of dangerousness.[[38]](#endnote-38)

There are four rejoinders to these objections. Assessing predictive risk-assessment systems in their best light (ie assuming good quality data and the removal from consideration of any variables that are insignificant or known to be corrupted by bias), these are plausible enough to entertain. However, they are ultimately unpersuasive since they fail to counter the underlying moral objection that actuarial and predictive technologies undermine individualised liberal justice principles.

First, advocates of predictive risk assessment contend that its alleged arbitrariness is tolerable so long as it ‘works’ in the sense that it can be shown to be generally more accurate than alternative predictive techniques: in that case penal resources can be directed accordingly, and public safety promoted in a rational way.[[39]](#endnote-39)

Second, because it works by deciding in advance which variables to factor in (as opposed to the more truly automated ‘big data-scraping’ algorithms that we noted above are increasingly utilised in marketing and other commercial contexts), the actuarial principles upon which systems like OASys are built can ensure that cases that are relevantly alike can be treated alike. This ensures that no one offender is treated more leniently or harshly on account of variations in the application of, and weight given by, different judges to the available risk variables.[[40]](#endnote-40)

Third, demographic data can be and are in fact combined with a more individualised assessment of the offender. There is no reason why demographic information should not be used in conjunction with personal information or indeed clinical assessment, and indeed as we have seen above, OASys and the dangerousness sentencing provisions in England and Wales do accommodate both such approaches.[[41]](#endnote-41)

Fourth, in response to the accusation that predictive risk assessment exacerbates social inequality, it might be suggested that there is nothing *intrinsic* to the correlation between high risk scores and socio-economic marginalisation and vulnerability. In practice, social harm is not inevitable, since a penal policy can be pursued that determines that a low risk score will be treated as a reason to reduce a sentence rather than a high risk score as a reason to increase it.[[42]](#endnote-42) In any case, there is no reason why sentencing judges should be obliged (*de jure* or *de facto*) to follow the algorithmically-produced risk assessment if there are reasons to take a different view in a particular case, again as noted above in the England and Wales context.

All four rejoinders, however, provoke further objections, all of which suggesting that the problems are not so easily overcome. To take the first one described above (that the arbitrariness of actuarial predictive technologies is tolerable so long as it is more accurate than the alternative), we may observe that testing the accuracy of any given prediction is complicated by contingencies such as the particular impact of the sentencing disposition (say, a prison sentence) on the individual offender. In studies that have nevertheless sought to test the accuracy of re-offending predictions based on actuarial assessments, these have never been shown to get it right more than about 70% of the time. This has remained the case in the era of OASys and other comparable systems with their many co-mingled variables.[[43]](#endnote-43) It is of course possible that future predictive systems that are truly automated and algorithmic in the selection of variables and collection of data might in time prove more ‘accurate’, but there remains to be seen if indeed such a system could be applied at all in the context of sentencing. In any case, because risk labels as generated in systems now in use represent a *group*-level probabilistic forecast, the re-offending prediction is not subject to proof or falsification by the individual offender.

This is very much a principled objection to prediction. Even putting aside the question of whether the offender is in prison or on release, it is necessarily impossible for that person to show that, after all, it was wrong to label him or her ‘high risk’ since the risk label itself indicates only what *proportion* of people sharing those characteristics may be expected to commit further crimes, not that *he or she* will.[[44]](#endnote-44) Short of a prediction of either zero or 100% certainty of re-offending, this can never be disproved at an individual level. It is also a matter of some debate as to whether the crime-reduction achieved by the greater incapacitation (and the at least theoretical possibility for rehabilitation) effected by longer sentences for dangerous criminals outweighs the unintended consequences of such a policy.[[45]](#endnote-45) It is plausible to suspect that at least in some cases, offenders will only be further hardened or damaged by having to endure an extended prison term. In these cases, the longer sentences for dangerous offenders may have some value in terms of incapacitation, but they may also serve to *increase* the risk of reoffending after eventual release.[[46]](#endnote-46)

Regarding the second rejoinder (predictive risk assessment enables courts to treat like cases alike), a question may be raised with respect to what constitutes a ‘relevant’ variable and what justifies us in concluding that ‘Case A’ is relevantly *like* ‘Case B’ but *unlike* ‘Case C’. If the variable that separates offender C from offenders A and B is a socio-economic one such as indebtedness or employment status (and hence not one relating to the offender’s own *criminal* responsibility), then to insist that he therefore be treated differently is no less unjust than the more ‘haphazard’ variations of treatment between different judges attaching weight inconsistently to the relevant factors, or even applying wholly divergent factors. If anything, the automated approach is almost certainly the more unjust, given that harder treatment of the economically disadvantaged offender is systematically and deliberately embedded in the algorithm.

As for the third rejoinder (group-level risk-factors can be strengthened by being combined with individual ones), it is indeed true that in many cases factors relating to the offender’s personal profile and characteristics may appear to lend supporting weight to a calculation of risk otherwise based on demographic or socio-economic data. But what is it that the group-level data actually tell us about *the individual being sentenced*, such that this information can meaningfully be *combined* into the sentencing decision? That an offender may have a string of past relevant convictions arguably bears relevantly on the risk that he or she may re-offend. Notwithstanding that it is not part of the offence for which the offender is being asked in the instant case to take responsibility for, past offending nonetheless at least affords information about how the offender him or her*self* has exercised or failed to exercise moral agency in the past. On the other hand, the fact that the offender may happen to share with a certain section of the population characteristics such as being of a relatively low socio-economic status, does not so reference that individual’s moral agency at any point at all. Rather, it simply identifies the offender with a certain number of *other* people, and asks how those other people have behaved in the past. Consequently, that latter set of characteristics cannot simply be welded onto considerations about the individual being sentenced without sacrificing the principles of individual justice and moral autonomy.[[47]](#endnote-47)

Regarding the fourth rejoinder (that social injustice is not inevitable, depending on the policy framework adopted), we may accept that innovative sentencing policy can potentially go some way to countering the socially divisive implications of algorithmically produced assessments. However, it remains the case that if high risk correlates positively with socioeconomic marginalisation and vulnerability (which it generally does) then the *contribution* of such data to sentencing considerations will be to provide a reason to punish marginalised, vulnerable people more harshly.[[48]](#endnote-48) It is after all the principled basis for predictive technology that concerns us here. In any case, we know that in sentencing jurisdictions where such technology is used in the interests of public safety, high risk *does* generally mean a longer sentence, which makes the argument that policy *might in theory* reverse this state of affairs at best disingenuous, and at worst cynical. In England and Wales, there is nothing in the CJA 2003 to prompt judges to treat the marginalising impact for socio-economically disadvantaged offenders of longer sentences as a mitigating factor. Furthermore, experience in this jurisdiction of the ill-fated Indeterminate Sentence for Public Protection (IPP) between 2009 and 2012 (during which time sentencing judges readily exercised their statutory discretion to impose an indeterminate sentence for public protection) suggests that it would be unwise to expect judges to exercise their discretion to make good this lack even if permitted to do so.[[49]](#endnote-49) Both in principle and in practice then, it is legitimate to worry that risk-based sentencing tends to be divorced from broader considerations of the social impact of punishment, namely the vulnerability of those on which it falls most heavily.[[50]](#endnote-50)

# 4. Concluding remarks: problems with no solution?

The ever-present demand for efficient ways to allocate limited carceral resources and to offer protection to the public from dangerous offenders means that there will likely always be a demand for smart systems that aid decision-making by making use of available data. By extension, it may be anticipated that there will also always be a role for legal and other relevant experts who can supply a figurative bridge between a rapidly evolving technology, problems of transparency, openness, fairness and bias thrown up by its application and routes to socio-technical solutions and legitimating normative and legal frameworks.[[51]](#endnote-51) But as discussed above, the apparent progress enabled by this marriage of legal, ethical and technical expertise and policy has not resolved some of the most obstinate difficulties facing the justifiability of algorithmic predictive technologies and the risk-based sentencing they support.

Those difficulties are no mere matters of academic scruple, since they engage fundamental principles concerning the basic relationship between the offender and the state: about the basis on which the former’s hard treatment and deprivation of liberty by the latter is justifiable at all, how much of that treatment can legitimately be meted out without further justification, and the broader social costs that follow it. At the same time, it would be difficult to overstate the problems involved in actually overcoming this impasse, since the liberal principles in question – retrospective and responsibility-based – are quite separate from, and run directly counter to, those prospective and probabilistic principles of risk assessment.

On the one hand then, this chapter has identified some significant and important problems regarding the justifiability of predictive risk assessment as a source of information for sentencing dangerous offenders. On the other hand, laying out a ‘solution’ to this, even assuming such a thing exists, is a much larger task for another time and place. I conclude therefore with a plea to colleagues who may be keen to use their legal expertise to assist the development and application of new and apparently world-changing technologies not to lose sight of the issues discussed above: either the core liberal-philosophical principles underlying individual justice or the social justice and marginalisation implications. The nature of these concerns provides compelling evidence that any solution we might come to is unlikely to be found through technological advances. For, if by ‘advances’ we mean further automation at the data-collection end in order to feed into the process a greater quantity and variety of information, this would inevitably mean undermining the extent to which our current offender risk-assessment systems do guard against bad quality, misleading or discriminatory data. Leaving aside for now those more distant challenges, we may say that data-driven risk-based sentencing is not a source of truly novel problems, but merely a new skin for an old snake whose dimensions and character are already well known.

1. See Jamie Doward, ‘Usman Khan was freed. Then he went on a killing spree. How did this happen?’ *The Guardian* (London, 1 Dec 2019) <www.theguardian.com/uk-news/2019/dec/01/khan-attack-will-put-sentencing-and-release-of-terrorists-under-scrutiny>. [↑](#endnote-ref-1)
2. Indeterminate sentences for Public Protection (IPP) were introduced by the Criminal Justice Act 2003 but then abolished by the Legal Aid and Sentencing of Offenders (LASPO) Act 2012. See Harry Annison, *Dangerous Politics* (OUP 2015). [↑](#endnote-ref-2)
3. See Home Office, *First-tier Tribunal bail: completing the bail summary* (20 January 2020); HM Prison Service Order, *Offender Assessment and Sentence Management - OASys* (Prison Service Order 2205, 20 April 2005). [↑](#endnote-ref-3)
4. Lyria Bennett Moses and Janet Chan, ‘Using Big Data for Legal and Law Enforcement Decisions: Testing the New Tools’ (2014) 37(2) University of New South Wales Law Journal 643; Emre Bayamlıoğlu and Ronald Leenes, ‘The “rule of law” implications of data-driven decision-making: a techno-regulatory perspective’ (2018) 10(2) Law, Innovation and Technology 295, 311; Kelly Hannah-Moffat, ‘Actuarial Sentencing: An “Unsettled” Proposition’ (2013) 30(2) Justice Quarterly 270, 284-87; Vincent Chiao, ‘Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice’ (2019) 15 International Journal of Law in Context 126, 134-35. [↑](#endnote-ref-4)
5. See Bernard E Harcourt, ‘Risk as a Proxy for Race: The Dangers of Risk Assessment’ (2015)

27(4) Federal Sentencing Reporter 237, on the history of criminal justice (focusing on US experience). [↑](#endnote-ref-5)
6. Caryn Devins and others, ‘The Law and Big Data’ (2017) 27 Cornell Journal of Law and Public Policy 357, see especially 363-65, 370-85; Bayamlioglu and Leenes (n 4) 303-7, 307-9. [↑](#endnote-ref-6)
7. RA Duff, *Punishment, Communication, and Community* (OUP 2000). [↑](#endnote-ref-7)
8. The most significant changes for present purposes being introduced by the Criminal Justice and Immigration Act 2008 and the LASPO Act 2012. The Counter Terrorism (Sentencing and Release) Bill and the Sentencing Bill set out in the December 2019 Queen’s Speech will bring about further changes to the CJA dangerousness provisions (Prime Minister’s Office, 2019) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\_data/file/853886/Queen\_s\_Speech\_December\_2019\_-\_background\_briefing\_notes.pdf>. [↑](#endnote-ref-8)
9. Schedule 15 CJA 2003 (as amended) lists the relevant ‘specified’ offences. Quoted section: s 226A(7), as added by LASPO, s 124. [↑](#endnote-ref-9)
10. s 229(2) subsections (a), (b) and (c). Emphases added. [↑](#endnote-ref-10)
11. Prison Reform Trust, *Offender Management and Sentence Planning* (2019) <www.prisonreformtrust.org.uk/ForPrisonersFamilies/PrisonerInformationPages/OffenderManagementandsentenceplanning>. [↑](#endnote-ref-11)
12. CPS, ‘Sentencing Dangerous Offenders’ (06 November 2019) <www.cps.gov.uk/legal-guidance/sentencing-dangerous-offenders>. [↑](#endnote-ref-12)
13. Lucia Zedner, ‘Erring on the Side of Safety: risk assessment, expert knowledge, and the criminal court’ in GR Sullivan and Ian Dennis (eds), *Seeking Security: Pre-Empting the Commission of Criminal Harms* (Hart Publishing 2012). [↑](#endnote-ref-13)
14. Law Commission, *The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales: A New Approach to the Determination of Evidentiary Reliability* (Law Com Consultation Paper 190, 2009) 43-46. For commentary, see Zedner (n 13) 240 who suggests that the LC research undermines claims to AI-assisted sentencing being particularly ‘scientific’. [↑](#endnote-ref-14)
15. Kelly Hannah-Moffat, ‘The uncertainties of risk assessment: partiality, transparency, and just decisions’ (2015) 27(4) Federal Sentencing Reporter 244. [↑](#endnote-ref-15)
16. Kelly Hannah-Moffat (n 4) 270. See also Annison (n 2) ch 8 on populism and penal policy. [↑](#endnote-ref-16)
17. Kelly Hannah-Moffat, ‘Algorithmic risk governance: Big data analytics, race and information activism in criminal justice debates’ (2019) 24(4) Theoretical Criminology 453, 456-57; Moses and Chan (n 4) 650-51. [↑](#endnote-ref-17)
18. Malcom M Feeley and Jonathan Simon, ‘The new penology: notes on the emerging strategy of corrections and its implications’ (1992) 30(4) Criminology 449. [↑](#endnote-ref-18)
19. Andrew Ashworth and Lucia Zedner, *Preventive Justice* (OUP 2014) 127-29. [↑](#endnote-ref-19)
20. Kelly Hannah-Moffat and Paula Maurutto, ‘Re-contextualizing pre-sentence reports: Risk and race’ (2010) 12(3) Punishment and Society 262; Sonja B Starr, ‘Evidence-Based Sentencing and the Scientific Rationalization of Discrimination’ (2014) 66 Stanford Law Review 803; John Monahan and Jennifer L Skeem, ‘Risk Assessment in Criminal Sentencing’ (2016) 12 Annual Review of Clinical Psychology 489. For a review of its use internationally, see Gwen van Eijk, ‘Socioeconomic marginality in sentencing: The built-in bias in risk assessment tools and the reproduction of social inequality’ (2017) 19(4) Punishment and Society 463, 465-6; Annison (n 2) ch 8. [↑](#endnote-ref-20)
21. Indeed, states have a general duty to take steps to promote citizens’ Article 2 rights (ECHR). [↑](#endnote-ref-21)
22. Duff (n 7) 170. [↑](#endnote-ref-22)
23. Ashworth and Zedner (n 19) ch 6. [↑](#endnote-ref-23)
24. Andrew von Hirsch and Andrew Ashworth, *Proportionate Sentencing* (OUP 2005) 56; Ashworth and Zedner (n 19) 120, 162. [↑](#endnote-ref-24)
25. Duff (n 7) 166. See also von Hirsch and Ashworth (n 24). [↑](#endnote-ref-25)
26. Zedner (n 13) 220-21. See also Ashworth and Zedner (n 19) 150. [↑](#endnote-ref-26)
27. Duff (n 7). [↑](#endnote-ref-27)
28. Ashworth and Zedner (n 19) 149-150. [↑](#endnote-ref-28)
29. *R v Secretary of State for the Home Department ex parte Hindley* [2000] All ER (D) 430 per Lord Steyn: ‘there are cases where the crimes are so wicked that even if the prisoner is detained until he or she dies it will not exhaust the requirements of retribution and deterrence.’ [↑](#endnote-ref-29)
30. Harcourt (n 5) 66. The first of two purposes of the proposed Sentencing Bill are to: ‘Ensure that the most serious violent and sexual offenders spend time in prison that matches the severity of their crimes, protecting victims and giving the public confidence.’ For discussion of the combination of retributive justice and incapacitation in punishment, see von Hirsch and Ashworth (n 24). [↑](#endnote-ref-30)
31. J Floud, ‘Dangerousness and criminal justice’ (1982) Brit J of Crim 22, 213-28. [↑](#endnote-ref-31)
32. See Feeley and Simon (n 18). Again, it must be emphasized that the data referred to here concern past data about the re-offending risk of particular sections of the population, *not* personalised data about the individual’s own past behaviours. [↑](#endnote-ref-32)
33. Duff (n 7) 171. [↑](#endnote-ref-33)
34. ibid. The availability of such treatment programmes that would enable a prisoner to reduce his level of risk during the course of his incarceration is not a given however, and indeed it was on account of the poor provision of such services that the UK was found to be in breach of its Article 5 duties in *James, Wells and Lee v UK* (2012) 56 EHRR 159 – for further discussion see Ashworth and Zedner (n 19) 163. [↑](#endnote-ref-34)
35. Starr (n 20) 842; Eric Silver, ‘Actuarial Risk Assessment: Reflections on an Emerging Social-Scientific Tool’ (2000) 9(1/2) Critical Criminology 123, 134-37. [↑](#endnote-ref-35)
36. On actuarial risk assessment techniques consistently reporting correlation between ‘high risk’ and high levels of vulnerability and marginalisation, see van Eijk (n 20). [↑](#endnote-ref-36)
37. See eg Alison Leibling and Shadd Maruna (eds), *The Effects of Imprisonment* (Routledge 2013); Marie Hutton and Dominique Moran (eds), *The Palgrave Handbook of Prison and the Family (*Palgrave Macmillan 2019). [↑](#endnote-ref-37)
38. See Bernard E Harcourt, ‘Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age’ (University of Chicago Public Law & Legal Theory Working Paper No 94, 2005) on the problem of the self-fulfilling prophecy and the impact of the symbolic message in targeting populations. [↑](#endnote-ref-38)
39. Anthony W Flores, Kristin Bechtel and Christopher T Lowenkamp, ‘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’ (2016) 80(2) Federal Probation 38. [↑](#endnote-ref-39)
40. Chiao (n 4) 131-32. [↑](#endnote-ref-40)
41. Ric Simmons, ‘Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in our Criminal Justice System’ (2016) Michigan State Law Review 947, 984-96. [↑](#endnote-ref-41)
42. See Richard Couzens, ‘Evidence-Based Practices: Reducing Recidivism to Increase Public Safety; a Cooperative Effort by Courts and Probation’ (Cow County Judges Institute, Administrative Office of the Courts Education Division 2011) 10 <www.courts.ca.gov/partners/documents/may11-04-evidence.pdf>. [↑](#endnote-ref-42)
43. See Julia Dressler and Hany Farid, ‘The accuracy, fairness and limits of predicting recidivism’ (2018) 4(1) Science Advances <https://advances.sciencemag.org/content/4/1/eaao5580.full>. [↑](#endnote-ref-43)
44. Ashworth and Zedner (n 19) 232-33. [↑](#endnote-ref-44)
45. See A Ashworth, *Sentencing and Criminal Justice* (CUP 2010) 207. [↑](#endnote-ref-45)
46. See Starr (n 20) who adapts Harcourt’s notion of ‘elasticity’ to argue that calculations of the incapacitatory effect of imprisonment needs to be weighed against the damaging effects on prisoners (and their risk of re-offending) actually brought about *by* extended sentences. We return to the observation that imprisonment is itself a source of social harm below. [↑](#endnote-ref-46)
47. As accepted by Monahan and Skeem (n 20). [↑](#endnote-ref-47)
48. Starr (n 20) 840-1. [↑](#endnote-ref-48)
49. Ashworth and Zedner (n 19) 162; Annison (n 2). [↑](#endnote-ref-49)
50. Starr (n 20) 839. [↑](#endnote-ref-50)
51. Research on AI technologies and their application for government and criminal justice is listed as one of a limited number of key strategic priorities for the Strategic Priorities Fund, UKRI. UK Research and Innovation, ‘The Strategic Priorities Fund’ (*UKRI.org*) <www.ukri.org/research/themes-and-programmes/strategic-priorities-fund/>. [↑](#endnote-ref-51)