

In brief

- This is a case study that examines in detail a single application of AI in the public sector, in this instance Durham Constabulary's AI-based Harm Assessment Risk Tool (HART). It is aimed at reducing reoffending rates in the county and easing pressure on the criminal justice system.
- Custody officers need to examine individuals who have been arrested and are in police custody, and assess the likelihood that they will reoffend. Those at medium risk are then offered an alternative to prosecution.
- HART informs officers' decision-making by providing them with a valid second opinion, rather than replace this aspect of their job.
- It was trained on 104,000 case histories of people arrested and processed in custody. There is some concern that the data contained race- and class-based bias, so this aspect is under review.
- The tool went live in 2017 and is still at the experimental stage, but is already thought to be identifying more accurately those at high risk of reoffending.

To learn more about how to proceed along a path towards using AI in the public sector effectively and legitimately, read our other AI case studies and our accompanying paper: [How to make AI work in government and for people](#).

The Challenge

Since the financial crisis of 2008/09, the police in England have had to do more with less. As part of the UK government's austerity programme, and exacerbated by the economic recession in the aftermath of the crisis, police budgets across the country have been cut severely.

In north-east England, Durham Constabulary alone had to absorb a GBP1.3m cut (1 percent of its FY 2009/10 budget) when austerity measures were announced in 2010.^{[1][2]} To stay within budget and meet the targets, Durham police headcount shrank from 1,550 in 2010 to 1,200 in 2015, among other downsizing operations.^[3]

In these years, the Durham force not only looked at headcount to meet savings expectations but also at its efficiency in delivering policing. "It shocked us out of habits," said Chief Constable Michael Barton.^[4] In 2010, the courts were swamped with low-level offences like shoplifting, according to Barton, which in turn meant that significant officers' time was committed to such cases. Given that recidivism was also prevalent among these offenders, Durham Constabulary recognised that it had to find appropriate new ways of dealing with petty criminals, using its more limited resources.^[4] A growing body of research since the 1970s had been suggesting that so-called law-and-order techniques failed to reduce reoffending rates, even though they might satisfy a hunger for punitive sanctions of criminal behaviour held by some sections of society.^[5]

Therefore, Durham Constabulary started experimenting with alternative models for dealing with crime. Police leadership took note of research showing that women were disproportionately harshly treated in the criminal justice system compared to male offenders, "because decision-makers... subliminally considered that women also offended against their femininity".^[6] As a result, the Constabulary launched a new pilot programme which promoted "signposting" and targeted intervention services as opposed to prosecution for women.

The success of this new programme led to its deployment on a much wider array of cases across the entire population of individuals who had been arrested.^[7] It became known as Checkpoint, and was "a voluntary multi-agency adult offender diversion scheme within County Durham and Darlington, targeting low and medium level offenders within the criminal justice process, and offering them a credible alternative to criminal prosecution. It gives service users an opportunity to address the underlying causes of their offending, by encouraging them to engage with services instead of going to court."^[8]

Recognising that prosecution for crimes such as drug possession might fuel reoffending, Checkpoint provides a structured set of interventions that take individuals out of the court system and focuses on future prevention. Once arrested, the police have a six-month window in which to prosecute an individual. Consequently, those who enter the Checkpoint programme receive an intervention (based on nine “pathways” out of offending) within four months of their arrest. These interventions consist of a series of individual commitments made with a caseworker (such as attending alcohol rehabilitation meetings or ensuring that their children arrive at school on time). If these promises are broken, a prosecution is automatically triggered. Less than five percent of individuals have failed the four-month contract with the Checkpoint programme since its launch.^[9]

The Checkpoint programme needs custody officers to be successful at predicting individuals' risk of recidivism – this is crucial to its effectiveness. “The need for good prediction is not just about identifying the dangerous people,” explains the criminologist Professor Lawrence Sherman, “it’s also about identifying people who definitely are not dangerous.”^[10] Neither high-risk nor very low-risk individuals are eligible for the Checkpoint programme. Consequently, Checkpoint required a reliable method of predicting which individuals are likely to commit a non-serious offence within the next two years.^[10]

The Initiative

“It’s 3am on Saturday morning. The man in front of you has been caught in possession of drugs. He has no weapons, and no record of any violent or serious crimes. Do you let the man out on police bail the next morning, or keep him locked up for two days to ensure he comes to court on Monday?”^[10] In 2012, Durham Constabulary began developing the Harm Assessment Risk Tool (HART), an AI decision support system designed to help custody officers resolve exactly that dilemma: whether or not a suspect should be kept in custody or referred to Checkpoint. They developed the tool in-house in collaboration with the University of Cambridge’s Centre for Evidence-based Policing. The objectives for Checkpoint, and custody officers’ use of HART within it, were to:

- Decrease reoffending rates
- Reduce harm caused by offenders
- Lower the cost per offender
- Increase victim satisfaction
- Improve offender behaviour.^{[5][11]}

The AI system was trained on 104,000 histories of people previously arrested and processed in custody between 2008 and 2012.^[10] HART also uses data from 34 different categories, such as a suspect’s gender and postcode, to

reach its conclusions, which it then displays to the custody officer. HART classifies suspects at three risk levels of reoffending in the next two years:

- Low (unlikely to commit any offence)
- Medium (likely to commit a non-serious offence)
- High (very likely to commit a serious offence such as murder, aggravated violence, a sexual crime, or robbery).

Only those offenders who are deemed medium risk by the officer are eligible for Checkpoint. The tool was then dry-tested in 2013 with a two-year follow-up period to see whether it returned accurate results on reoffending.

An ongoing randomised controlled trial (RCT) of Checkpoint, which began in 2017, is going to provide empirical evidence on whether these goals are met. Experimental conditions for the RCT are established through the following process:

1. The offender is arrested and taken into custody
2. HART forecasts the reoffending risk for the next two years, and the custody officer reaches a decision, partly influenced by HART’s forecast
3. Custody officers screen those in the moderate risk category for eligibility to participate in Checkpoint – those who are eligible form the “intention to treat” population
4. Individuals are randomly assigned to one of two groups – the actual Checkpoint treatment group, and the Checkpoint control group against which the performance of the programme will be measured
5. Lastly, both groups are processed, either through Checkpoint or the regular criminal justice system.^[5]

HART is meant to inform human decision-making rather than replace it. Custody officers use HART’s prediction as a decision aid, while retaining their autonomy to classify a person differently. “The custody officer must consider other factors in the decision-making process... under both the Policing and Crime Act 2017 and the Bail Act 1976.”^[12] The model “simply does not have all of the information available to it”. Not only it is not connected to databases outside Durham Constabulary’s direct control (e.g. local council data or the Police National Database), it also cannot replicate the “local knowledge” that custody officers can apply to a given situation. These officers “will frequently be aware of other information that overrides the model’s predictions, and they must apply their own judgement in deciding upon the disposition of each offender’s case”.^[13]

When the algorithm was initially tested, “members of the police force were asked to mimic its outcomes by predicting whether a person would be of a low, moderate or high risk of reoffending. Almost two-thirds (63.5 percent) of police officers ranked offenders in the moderate category.”^[14] The model and the officers agreed only 55 percent of the time, with human decision-makers skewed towards moderate and low risk ratings.^{[12][13]} When evaluated against historical data, the model was found to predict low-risk individuals with 98 percent accuracy, and high-risk individuals with 88 percent accuracy.^[15]

As HART was developed against the backdrop of resource constraints, training for custody officers was focused on dealing with immediate issues relating to the use of the tool as a decision aid. For example, “to prevent existing human biases – around race and social prejudices – creeping into use with HART, Durham Constabulary has provided members of its staff with awareness sessions around unconscious bias. The police force also stresses it has not included race within the predictors its algorithm uses and the tool’s output is an aide for humans to help them make decisions.”^[14] Chief Constable Michael Barton also funds two employees every year to attend an Applied Criminology and Police Management master’s programme at Cambridge University to “build a body of knowledge” within the Constabulary, and promote “evidence-based policy rather than policy-based evidence”.^[6]

The Durham process has been characterised by a higher degree of openness than comparable tools. Information about HART, including the internal weights and structures of the model it uses, has been published by the designers and by Durham Constabulary as part of parliamentary evidence,^[12] academic papers, and conference presentations.^{[5][13]} An audit trail to show “how the system arrived at any given decision should scrutiny be required later” was also promised.^[16]

Importantly, the initiative also acquires legitimacy through the way it is set up as an experiment which gathers data to test specific hypotheses. The concept of “experimental proportionality” has been applied to permit the use of unproven algorithms in a controlled and time-limited way, as part of “a combination of approaches to combat algorithmic opacity”.^[13] Opacity in the machine learning context arises from “the interplay between [data inputs and the written code] in the mechanism of the algorithm”, even though taken separately the data inputs and the written code are straightforward and clear.^[13]

To introduce checks and balances for opacity, Durham Constabulary, in collaboration with the University of Winchester, has developed an ethical framework and process for the analysis of algorithmic systems in policing called ALGO-CARE (“Advisory; Lawful; Granularity; Ownership; Challengeable; Accuracy; Responsible; Explainable”).^[17] There are important questions about the risk of employing such algorithms in police systems, but there has been relatively little discussion in the public domain about how these principles have been concretely applied in the Durham case.

Having research partners appears to have made the HART trial more agile and responsive to insights. The University of Cambridge not only contributed to the initiative with the kind of technical programming skills that police forces do not tend to possess but also introduced scientific review methods. This encouraged the ongoing maintenance and adaptation of HART.

The Impact

As HART is currently at the experimental stage, any definitive evidence of its impact will be delivered only once the RCT has been analysed in full.^[18] At this stage, data suggests that HART is not only very accurate but has also reduced the likelihood of making dangerous errors (where the risk of reoffending is underestimated) as against cautious errors (where it is overestimated). Comparing two cohorts of offenders, where the second cohort actually committed more offences after their initial custody episode, the rate of all the offenders forecast as low risk but actually displaying high-risk behaviour remained at only 2.4 percent. This means that the model has become more cautious over time, which is in the interest of public safety.^[13]

Accurate forecasting “could also lead to more ‘effective triage’ and therefore better results in terms of reoffending and cost” – addressing the original challenge facing the Durham police force.^[19] Although this is unlikely to be solely attributable to the use of AI in the Checkpoint programme, “Durham Constabulary was rated outstanding for managing its resources in the last independent inspection and [Police Crime Commissioner Ron Hogg] said he hopes the force will remain at the ‘pinnacle’ of policing”.^[3] Chief Constable Barton, reflecting on Checkpoint’s impact on wider capacities in the criminal justice system, said: “When magistrates first heard of the scheme in 2010, they thought we were quasi-judicial. Now they deal with the people who should be there, focusing on the more serious offences.”^[3] Courts are no longer overwhelmed by the number of low-level cases.



Despite “being so open [and acknowledging] that this approach is new to policing and is therefore also new to communities”, HART attracted intense criticism in early 2018 that it was a black box. Its use of potentially prejudicial marketing data was exposed by Big Brother Watch – a UK thinktank concerned with the protection of privacy rights – and publicised by the online magazine Wired,^[14] among other publications.^{[13][20]} Big Brother Watch accused Durham Constabulary of having input into HART marketing data that uses race- and class-based stereotypes.

Public records revealed that the Constabulary had bought datasets from Experian, which links people to social categories, using occupational, demographic, and health data, but also postcodes and names. For example, in the dataset purchased by the Durham force, “people called ‘Stacey’ are likely to fall under ‘Families with Needs’ who receive ‘a range of benefits.’”^[21] Sheena Urwin, head of criminal justice at Durham Constabulary, has defended the use of this data, arguing that the Experian data helped the force “to better understand our communities and to improve our engagement.”^[21]

In September 2017, Urwin coauthored a research article with scientists from the universities of Cambridge, Sheffield and Winchester, drawing lessons from HART for other public sector applications of AI decision support tools. In the article, the authors explore the dynamic changes of the model that occur through machine learning and the impact this could have on custody officers’ interaction with HART. They go on to explain that, as a result, the tool was refreshed with more recent data.^[13] This indicates that the way the experiment was set up allowed officials to be responsive to emerging insights and public criticism.

Data has shown that, during the test phase, custody officers agreed with only 10 percent of HART’s predictions of offenders being at high risk of reoffending.^[12] This disagreement might lead one to conclude that AI systems cannot replace, or even augment, human decision-making in the criminal justice sector. An alternative view is that these systems may encourage better monitoring and evaluation of the effectiveness of interventions than is currently the case. Whatever these interim conclusions, the assessment of HART is ongoing and its eventual verdict is awaited with great interest.

BIBLIOGRAPHY

- [1] [Concern over Durham and Cleveland Police budget cuts, BBC News, 28 May 2010](#)
- [2] [Annual Report 2009-2010 & Policing Plan 2010-2013, Durham Police Authority and Durham Constabulary, 2010](#)
- [3] [Laura Hill, Fewer police officers in Durham as force faces £400,000 budget cut, Chronicle Live, 21 December 2015](#)
- [4] [Jeremy Armstrong, 'It shocked us out of habits' Durham Police chief on the positive side of funding cuts, Chronicle Live, 16 October 2017](#)
- [5] [Gillian Routledge, A Protocol and Experimental Trial: The Checkpoint Desistance Programme in Durham, University of Cambridge, December 2015](#)
- [6] [Technology and the Law Policy Commission – Algorithms in the Justice System, The Law Society, 11 September 2018](#)
- [7] [Policing and young adults: developing a tailored approach, The Police Foundation, July 2018](#)
- [8] [Community Awards Winners, The Howard League for Penal Reform, 2016](#)
- [9] [Michael Cross, Police chief explains 'justice by algorithm' tool, The Law Society Gazette, 28 July 2018](#)
- [10] [Louise Walsh, Helping police make custody decisions using artificial intelligence, University of Cambridge, 26 February 2018](#)
- [11] [Our Commands, Durham Constabulary website 2018](#)
- [12] [Written evidence submitted, Sheena Urwin, Durham Constabulary, 20 February 2018](#)
- [13] [Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality, Marion Oswald et al., Information & Communications Technology Law, 2018](#)
- [14] [Matt Burgess, UK police are using AI to inform custodial decisions – but it could be discriminating against the poor, Wired, 1 March 2018](#)
- [15] [Jonathan Pugh & Tom Douglas, Using AI to Predict Criminal Offending: What Makes it 'Accurate', and What Makes it 'Ethical', University of Oxford Practical Ethics Blog, 16 June 2017](#)
- [16] [Christian Baraniuk, Durham Police AI to help with custody decisions, BBC News, 10 May 2017](#)
- [17] [Marion Oswald and Sheena Urwin, Written evidence, University of Sheffield and Durham Constabulary, 2017](#)
- [18] [John Cooper, Checkpoint Offender Management, College of Policing, 31 July 2018](#)
- [19] [Marion Oswald, Algorithmic risk assessment policing models – lessons from Durham Constabulary's HART model, University of Winchester, 9 October 2017](#)
- [20] [Durham police criticised over 'crude' profiling, BBC News, 9 April 2018](#)
- [21] [Police use Experian Marketing Data for AI Custody Decisions, Big Brother Watch, 6 April 2018](#)

OTHER SOURCES

- [A Closer Look at Experian Big Data and Artificial Intelligence in Durham Police, Big Brother Watch, 6 April 2018](#)
- [Algorithms in decision-making, Science and Technology Committee, House of Commons, 23 May 2018](#)
- [Edwards L and Veale M, Slave to the Algorithm? Why a 'Right to an Explanation' is Probably Not The Remedy You Are Looking For, Duke Law & Technology Review 16\(1\), pp.18-84, 2017](#)
- [Technology and the Law Policy Commission – Algorithms in the Justice System, The Law Society of England and Wales, 2018](#)



Centre for
Public Impact
A BCG FOUNDATION

Join the conversation

ai@centreforpublicimpact.org

Follow us @CPI_foundation: #FindingLegitimacy
#AllLegitimacy



October 2018
© Centre for Public Impact 2018