Written evidence submitted by Jamie Grace, Sheffield Hallam University (MAC0001)

The Macpherson Report twenty-one years on

Summary

1. Machine learning technologies and data analytics used by UK police forces will carry risks of entrenching racial bias when used to inform operational policing approaches to dealing with crime. While this risk can be mitigated, namely through the use of proper oversight mechanisms, and allowing scrutiny from community groups and civil society, there remains particular problem with police use live facial recognition (LFR) and potential racial discrimination in the policing of public places.

2. Some tech companies working with police forces around the world are starting to introduce moratoria on the deployment of LFR tools by police forces while the risks of potential bias in their use is investigated and better understood. But the profit imperative for companies, and the desire to innovate and make the most of resources on the part of UK police forces, will mean that private-sector tech involvement will be a permanent presence in this element of the UK justice system.

3. For these reasons, the Home Affairs Select Committee should go one step beyond the findings of the recent Committee on Standards in Public Life report on *Artificial Intelligence and Public Standards*, and recommend that government and Parliament create a new statutory regulator for the use of developing technologies in the UK criminal justice system. This new independent statutory regulation for police technology and criminal justice data ethics would also help to keep the use of machine learning and data analytics by UK law enforcement agencies under sufficient scrutiny.

Written evidence - *Police powers and practice involving data analytics*

4. I am an academic researcher, examining human rights impacts stemming from the use of artificial intelligence and machine learning/data analytics, particularly in criminal justice settings. You have called as a Committee for evidence on the development of police powers and practice, since July 2019, and on the way that these developments have had impacts on BME communities. I would like to offer up some evidence on two key points: firstly, the use of machine learning/data analytics tools that require careful oversight to guard against bias of different kinds; and secondly, the way that live facial recognition technology used in public places by police forces, to identify individuals, needs similar careful oversight.

Safeguarding against bias in police use of machine learning/data analytics tools

5. The Information Commissioner's Office picked up on a key issue of intersectionality - specifically an increased impact on the statutory 'protected characteristics' of race and age, and a likely breach of the Equality Act 2010 - when it issued an Enforcement Notice under the Data Protection Act in relation to the Gangs Matrix operated by the Metropolitan Police (ICO 2018a). The Gangs Matrix had not been used in a way that was sufficiently transparent or open to challenge by the (disproportionately) young black men and teenage males that it 'scored' for gang connections in the London area (MOPAC 2018). The ICO has produced a checklist for compliance for police forces using gang intelligence databases (ICO 2018b) but arguably the best confirmation of the impact of greater scrutiny arising from the Enforcement Notice against this algorithmic (in)justice came when the Mayor's Office for Policing and Crime (MOPAC) in London prioritised overhauling the workings of the Gangs Matrix (MOPAC 2020).

6. But it should be noted that the Gang Matrix operated by the Metropolitan Police is not as sophisticated as machine learning-based tools, where there is an even greater risk that policing approaches are swayed, and possibly biased, by the powerful insights the latter can offer into the justice landscape that police forces must navigate. Fundamentally, it must not be forgotten that algorithmic tools used for advanced data analytics and 'predictive policing' are 'trained' or built as models from data that is societally skewed because of, in turn, biased policing practices. This could create a real risk of a discriminatory 'feedback loop' in some contexts for policing practices.

7. Efforts to promote algorithmic accountability across the UK police service have begun, however. At the time of writing, a number of forces within the UK police service use a self-regulation framework in relation to machine learning and data analytics, aimed at police forces that are adopting greater data science approaches in their intelligence analysis processes. Known as 'ALGO-CARE', this regulatory framework is a checklist of key considerations in legal, ethical and data science best-practice, to be used by police forces in guiding their innovation and their adoption of capabilities around data analytics and machine learning applications (Oswald & Ors, 2018)

8. ALGO-CARE requires police forces to use predictive analytics in an advisory (not determinative) way, with control over their intellectual property in the algorithm concerned, and in a way that is lawful; granular; challengeable; accurate; responsible and explainable. The research developing ALGO-CARE was a co-authored evaluation of the legalities of the 'Harm Assessment Risk Tool' (HART), used currently by Durham Constabulary (Oswald & Ors 2018). The HART tool is a leading application of machine learning technology as used in intelligence analysis and risk management practices by police in the UK. HART was the first such police machine learning project in the UK to be

open to early academic scrutiny; and as a result was the first which has led to the development of a model regulatory framework, in the form of ALG-CARE, for algorithmic decision-making in policing.

9. The National Police Chiefs' Council (NPCC) took the decision in November 2018 to promote the use of ALGO-CARE as a model for best practice in the self-regulation by UK police forces of their development of machine learning/algorithmic tools (Grace 2020). In the summer of 2019, it was confirmed by the NPCC that West Midlands Police (WMP) were incorporating the ALGO-CARE checklist or framework in internal development processes in relation to new intelligence analysis tools. West Midlands Police now host the National Data Analytics Solution (NDAS) for the UK police service as a whole. ALGO-CARE is built into the project initiation process for NDAS, and has been used to provide ethical oversight for data analytics projects concerning identifying risks factors around vulnerability to modern slavery, and the perpetration of knife crime (West Midlands Police and Crime Commissioner, 2020). Importantly, Essex Police have also drawn on the ALGO-CARE framework in setting up the oversight processes for their data analytics partnership with Essex County Council (Essex Centre for Data Analytics 2019). I also know from my own recent research using freedom-of-information requests that forces in North Wales, Wiltshire, Lancashire, Avon & Somerset and Kent have also built the ALGO-CARE model into their initiation processes and governance procedures for data analytics and AI projects. This growing adoption of self-regulation is proof of a respect for the professional ethics in the use of machine learning and data analytics in policing. But this self-regulation is only a start - police forces need to be subject to scrutiny from a new, specialist and independent regulator in this key space.

Safeguarding against systemic bias in the use of live facial recognition in public places

10. In paragraph 11 of their revised draft of their General Comment No. 37 on the right to peaceful assembly, the UN Human Rights Committee explains quite rightly that:

> *"Surveillance technologies can be used to detect threats of violence and thus to protect the public, but they could also infringe on the privacy and other rights of participants and bystanders."*

11. Where these surveillance technologies are empowered by artificial intelligence, or machine learning and data analytics, this can lead to biases across different groups in the levels of enjoyment of different human rights, including the right to non-discrimination, since the datasets used to train these tools are often, for example, skewed along lines of race and/or gender. This is no doubt part of the reason why IBM and Amazon have been reported as temporarily abandoning their provision of support for live facial recognition technology used by police forces (BBC News, 2020; Ng, 2020).

12. It must be noted that there is a consistent and important concern about poorer accuracy rates for facial recognition technology with regard to the real-time identification of non-white persons (Harwell 2019). In the UK, there has been a (so far unsuccessful) challenge to the use of live facial recognition in the case of *R (Bridges) v South Wales Police* (2019), but this decision is being appealed. As a result, the lawfulness of the legal and regulatory framework around live facial recognition is something that we must still with caution. The Metropolitan Police have published a document (Metropolitan Police Service 2020) that sets out what they term their legal mandate for Live Facial Recognition (LFR), and this acknowledges the need for compliance with the Public Sector Equality Duty (PSED) in deploying the controversial technology in public spaces, but has few details as to how decision-makers in this regard would ensure they were 'properly informed', as required by the PSED, about the equality issues inherent in deploying LFR in various areas of London with a differing prevalence of people of different ethnicities, for example.

13. The Met also have an interesting pair of issues in their own recently published *guidance* document on live facial recognition technology (as opposed to their purported 'legal mandate' document). First they seem to make a policy commitment to public notification prior to deployments; second they committed to only deploy the technology overtly. The Met say their watch lists of suspect photographs used in the deployment of LFR are not marked with data about ethnicity, which might, it seems, means that accuracy rates for 'hits' or 'flags' in each LFR will be harder to determine. This seems to undermine PSED compliance, in either the spirit or the letter of the law. The Met claim it is because they should only process ethnicity data when strictly necessary for policing purposes, and that this is not a strictly necessary purpose under the terms of Part 3 of the Data Protection Act 2018. But this disregards the self-monitoring the PSED requires. The PSED is a statutory duty, just as the requirement for minimal data processing under the Data Protection Act 2018 is a statutory duty. Perhaps in an evaluation of LFR deployments, 'hits' or matches by ethnicity can be added back in to the watch list data - but if this is the case, the Met's claim that they remove ethnicity data from watch lists seems pointless. Efforts to engage with the public over LFR must be more genuine than this sort of dry, data protection-driven detailing in response to valid concerns (Yesburg & Ors 2020).

14. The operational guidance from the Met concerning LFR would be more reassuring for public confidence on the issue of bias if there was a clearer commitment and explanation as to the overall purpose of the use of LFR by the police in London in meeting their duties under the PSED, and actually reducing bias in street-level policing over time. The force falsely claimed in an equalities impact assessment that the use of the technology was supported by the UK Biometrics

Commissioner under current governance arrangement, risking public confidence in the integrity of their use of LFR (Gayle 2020).

<u>Conclusions</u>

15. The Committee on Standards in Public Life have had their say in a report on AI and standards in public life (2020), and there has been a report of a Parliamentary Committee on AI technology implications for civil society in the UK (Lords Select Committee, 2017). Furthermore, the Information Commissioner's Office has published its own consultation on a draft AI auditing framework (2020). The UK Centre for Data Ethics and Innovation is also to undertaking to develop a code of practice for policing in the UK with regard to the use of data analytics and machine learning (Macdonald 2020). I would argue that to this mix of contributions to the public discourse on these issues should come a recommendation from the Home Affairs Select Committee for a new independent statutory regulator for police technology and criminal justice data ethics which would help to keep the use of live facial recognition, machine learning and data analytics by UK law enforcement agencies under sufficient scrutiny.

<u>Recent pieces of my own underpinning research</u>

J. Grace and R. Bamford, ''AI Theory of Justice': Using Rawlsian approaches to better legislate on machine learning in government', May 2020, accepted for publication, *Amicus Curiae*, available online as a draft paper: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3588256

J. Grace, ' Human rights, regulation and the development of algorithmic police intelligence analysis tools in the UK', available online as a draft paper at: http://ssrn.com/abstract=3303313

J. Grace, 'Machine learning technologies and their inherent human rights issues in criminal justice contexts', a chapter in a forthcoming collection on technology and human rights challenges, edited by Alexandra Moore and James Dawes, final draft available at: http://ssrn.com/abstract=3487454

J. Grace, 'Algorithmic impropriety' in UK policing contexts: A developing narrative?', in John McDaniel and Ken Pease (eds) *Policing and Artificial Intelligence*, forthcoming in 2020, available at: http://ssrn.com/abstract=3487424

J. Grace, 'Algorithmic impropriety in UK policing?', (2019) *Journal of Information Rights, Policy and Practice*; Vol, 3 Issue 1: https://jirpp.winchesteruniversitypress.org/articles/abstract/23/

Wider references to other research

BBC News (2020), 'IBM abandons 'biased' facial recognition tech', 9th June 2020, from: https://www.bbc.co.uk/news/technology-52978191

CSPL (Committee on Standards in Public Life) (2020), 'Artificial Intelligence and Public Standards'.

ECDA (Essex Centre for Data Analytics) (2019) 'Transparency and trust' Essex Partnership https://www.essexfuture.org.uk/ecda/collaborative-learning/transparency-and-trust/

Gayle, Damien (2020) 'Watchdog rejects Met's claim that he supported facial recognition', *The Guardian* 12 February 2020

Grace, Jamie (2020) 'Fresh, fair, and smart: data reliability in predictive policing', about:intel, available at https://aboutintel.eu/predictive-policing-data-reliability/

Harwell, Drew (2019) 'Federal study confirms racial bias of many facial-recognition systems, casts doubt on their expanding use', *Washington Post* 19 December 2019

ICO (Information Commissioner's Office) (2018a) *Metropolitan Police Service Gangs Matrix Enforcement Notice*, Wilmslow: Information Commissioner's Office.

ICO (Information Commissioner's Office) (2018b) *Processing gangs information: A checklist for police forces*, Wilmslow: Information Commissioner's Office.

ICO (Information Commissioner's Office) (2020) *ICO consultation on the draft AI auditing framework guidance for organisations*, Wilmslow: Information Commissioner's Office.

Lords Select Committee on Artificial Intelligence (2018) *Artificial Intelligence in the UK: Ready, Willing and Able?*, HL Paper 100

Macdonald, Lara (2020) 'What next for police technology and ethics?', Centre for Data Ethics and Innovation https://cdei.blog.gov.uk/2020/02/26/what-next-for-police-technology-and-ethics/

Metropolitan Police Service (2020) *Live Facial Recognition: Legal Mandate*, London: Metropolitan Police Service.

MOPAC (Mayor's Office for Police and Crime) (2020), 'Mayor's intervention results in overhaul of Met's Gangs Matrix', MOPAC, available at https://www.london.gov.uk//press-releases/mayoral/mayors-intervention-of-met-gangs-matrix

Ng, Kate (2020), 'Amazon bans police use of facial recognition technology over racial bias fears', *The Independent*, 11th June 2020, from: https://www.independent.co.uk/news/world/americas/amazon-police-facial-recognition-ban-racism-a9560016.html

Oswald, Marion & Ors (2018) 'Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality', 27:2 *Information & Communications Technology Law*, 223-250

WMPCC (West Midlands Police and Crime Commissioner) (2020), 'Ethics Committee minutes' (January 2020) Birmingham: Office of the Police and Crime Commissioner for the West Midlands

Yesburg, Julia & Ors (2020) 'Public support for Live Facial Recognition and implications for COVID-19 policing', London School of Economics Politics and Policy Blog, available at: https://blogs.lse.ac.uk/politicsandpolicy/covid-19-lfr/

June 2020