LCO Commissioned Report

AI Case Study: Probabilistic Genotyping DNA Tools in Canadian Criminal Courts

June 2021

LAW COMMISSION OF ONTARIO
COMMISSION DU DROIT DE L'ONTARIO

About the Law Commission of Ontario

The Law Commission of Ontario (LCO) is Ontario's leading law reform agency.

The LCO provides independent, balanced, and authoritative advice on complex and important legal policy issues. Through this work, the LCO promotes access to justice, evidence-based law reform and public debate.

The LCO evaluates laws impartially, transparently and broadly. The LCO's analysis is informed by legal analysis; multi-disciplinary research; contemporary social, demographic and economic conditions; and the impact of technology.

The LCO is located at Osgoode Hall Law School, York University, Toronto.

More information about the LCO is available at ***www.lco-cdo.org***.

Law Commission of Ontario Reports

Legal Issues in the Last Stages of Life (Forthcoming 2021)
Indigenous Legal Issues in the Last Stages of Life (Forthcoming 2021)
AI, Algorithms and Government Decision-Making (Forthcoming 2021)
Regulating AI: Critical Issues and Choices (April 2021)
The Rise and Fall of AI and Algorithms in American Criminal Justice: Lessons for Canada (October 2020)
Defamation Law in the Internet Age (March 2020)
Class Actions: Objectives, Experiences and Reforms (July 2019)
Legal Capacity, Decision-Making and Guardianship (March 2017)
Simplified Procedures for Small Estates (August 2015)
Capacity and Legal Representation for the Federal RDSP (June 2014)
Review of the Forestry Workers Lien for Wages Act (September 2013)
Increasing Access to Family Justice (February 2013)
Vulnerable Workers and Precarious Work (December 2012)
A Framework for the Law as It Affects Persons with Disabilities: Advancing Substantive Equality for Persons with Disabilities through Law, Policy and Practice (September 2012)
Curriculum Modules for Law Schools: A Framework for Teaching about Violence Against (August 2012)
A Framework for the Law as It Affects Older Adults: Advancing Substantive Equality for Older Persons through Law, Policy and Practice (April 2012)
Modernization of the Provincial Offences Act (August 2011)
Joint and Several Liability Under the Ontario Business Corporations Act (February 2011)
Division of Pensions Upon Marriage Breakdown (December 2008)
Fees for Cashing Government Cheques (November 2008)

The following individuals contributed to research or drafting of this Commissioned Report:

Primary Authors

Jill R. Presser, Presser Barristers
Kate Robertson, Associate, Markson Law; Research Fellow, The Citizen Lab (Munk School of Global Affairs & Public Policy, University of Toronto)

Law Commission of Ontario Staff

Nye Thomas, Executive Director
Ryan Fritsch, Counsel
Kiana Blake, LCO Student Scholar
Rudra Krishnan, LCO Student Scholar

Citation
Jill R. Presser and Kate Robertson, *AI Case Study: Probabilistic Genotyping DNA Tools in Canadian Criminal Courts,* Law Commission of Ontario (Toronto: June 2021).

Note: Portions of this paper are drawn from Kate Robertson and Jill R. Presser, "Algorithmic Technology and Criminal Law in Canada," Chapter 3 in Jill R. Presser, Jesse Beatson, and Gerald Chan eds., *Litigating Artificial Intelligence* (Emond: May 2021), relied on here with the express permission of Emond Publishing.

Contact

Law Commission of Ontario
2032 Ignat Kaneff Building
Osgoode Hall Law School, York University
4700 Keele Street
Toronto, Ontario, Canada
M3J 1P3

Email: LawCommission@lco-cdo.org
Web: www.lco-cdo.org
Twitter: @LCO_CDO
Tel: (416) 650-8406
Toll-free: 1 (866) 950-8406

Funders

The Law Commission of Ontario is funded by the ***Law Foundation of Ontario***, the ***Law Society of Ontario***, ***Osgoode Hall Law School***, and ***York University***. The LCO also receives significant in-kind support from Osgoode Hall Law School and York University.

ABSTRACT

This paper is the fourth in a series of Law Commission of Ontario (LCO) Issue Papers considering the use of artificial intelligence (AI), automated decision-making (ADM), and algorithms in the Canadian civil, administrative, and criminal justice systems. This paper considers the role and impact of AI-driven probabilistic genotyping technology to generate evidence used in the criminal justice system.

Probabilistic genotyping (PG) is the use of artificial intelligence algorithms to analyze DNA samples collected in police investigations or criminal prosecutions. The overarching concern of this report is to examine whether and how AI-driven technologies like PG can continue to meet the high standards of due process, accountability and transparency, and fundamental legal rights demanded by the *Canadian Charter of Rights and Freedoms* and, by extension, the criminal justice system.

The data and data science used at every stage of AI and algorithmic decision-making have human rights, equity, due process, and access to justice consequences. These tools often have built-in and largely hidden biases and discriminatory inferences in their decision-making. While appearing scientific, objective, and authoritative, they may be unreliable and invalid. Failure to study, understand, and regulate these tools can have significant system-wide and individual repercussions. Absent proper scrutiny, process, and legislation, there is a risk that AI tools, including PG DNA algorithms, will worsen racism in Canada's justice system, and put access to justice further out of reach for many Ontarians. Unequal access to participation in decision-making about data and technology can further entrench existing biases and inequality. Reliance on AI tools, like PG DNA algorithms, may provide what appears to be infallible and reliable scientific evidence against individuals who are factually innocent. Because of the well-established over-representation of low income people and members of racialized and Indigenous communities in the criminal justice system, those most likely to face PG DNA evidence will often be those least likely to be financially resourced to challenge it. Absent proper scientific study, regulation, enforcement of relevant *Charter* rights, and due process protections, PG DNA evidence may lead to wrongful convictions. This concern is likely to disproportionately impact on communities that already suffer from systemic discrimination in the justice system and Canadian society at large.

For these reasons, this report makes the following recommendations:

1. Statutory Amendments focused on the use of PG DNA analysis as evidence:
    a. Prescribing limits on the admission of AI-generated evidence
    b. Codifying presumptive inadmissibility in the absence of strict scrutiny
    c. Amendments to *Criminal Code* provisions applicable to expert evidence
    d. Amendments to *Evidence Act* provisions applicable to electronic documents

2. Statutory Amendments focused on enhancing systemic transparency and accountability:
    a. Modernizing Canada's system of oversight governing the collection, use, retention, and sharing DNA information within police investigations and criminal proceedings
    b. Enacting regulations requiring transparency and accountability with respect to algorithms that will be used for criminal justice purposes, including openness of source code and source code review

3. Practices and Training, including:
    a. Establishing prosecutorial guidelines concerning the use of PG DNA evidence in criminal proceedings
    b. Developing access to training programs for all justice-system participants

4. Review of legal aid programs to identify and remedy gaps in policies and budgetary constraints

5. Further Research and Evaluation, including:
    a. Researching and evaluating PG DNA methods and their potential for bias
    b. Evaluating human rights and ethical concerns surrounding the use of DNA phenotyping in criminal investigations
    c. Evaluating the scope of existing Charter rights to ensure that fundamental rights and freedoms apply with equal force in circumstances where AI-generated evidence is used in criminal investigations and proceedings

CONTENTS

AI Case Study: Probabilistic Genotyping DNA Tools in Canadian Criminal Courts

I. INTRODUCTION

This paper is the fourth in a series of Law Commission of Ontario (LCO) Issue Papers considering the use of artificial intelligence (AI), automated decision-making (ADM), and algorithms in the Canadian civil, administrative, and criminal justice systems. This paper considers the role and impact of AI-driven probabilistic genotyping technology to generate evidence used in the criminal justice system. The LCO's first Issue Paper, *The Rise and Fall of AI and Algorithms in American Criminal Justice: Lessons for Canada*, looks at AI and algorithmic tools that aid criminal courts in pretrial custody or bail decision-making. The LCO has also released Issue Papers addressing AI regulation and government development of AI systems. A fifth LCO Issue Paper, considering the use of AI and algorithms in civil and administrative justice, will be released later in 2021.[1]

What are algorithms, automated decision-making, and AI?

The AI Now Institute defines algorithms and AI as follows:

"An Algorithm is generally regarded as the mathematical logic behind any type of system that performs tasks or makes decisions… Artificial Intelligence (AI) has many definitions, and can include a wide range of methods and tools, including machine learning, facial recognition, and natural language processing. But more importantly, AI should be understood as more than just technical approaches. It is also developed out of the dominant social practices of engineers and computer scientists who design the systems, and the industrial infrastructure and companies that run those systems. Thus, a more complete definition of AI includes technical approaches, social practices and industrial power."[2]

The difficulty of defining AI and algorithms has been noted by many. The focus of this report is on probabilistic genotyping, which uses AI that operates as a predictive model or tool. In this paper, the LCO will use the phrases "algorithmic decision-making" and "AI" to describe this range of technologies.

The LCO's Issue Papers provide an important "first look" at the potential use and regulation of AI and algorithmic technologies in Canadian justice systems. Each paper identifies a series of important legal, policy and practical issues and choices that Canadian policymakers and justice system stakeholders should consider before these and other AI technologies are widely adopted in civil, administrative and criminal law contexts across Canadian jurisdictions.

The LCO's research and consultations in this area draw heavily from the international experience with these systems, particularly in the United States and the European Union. This experience provides both insights and important lessons for Canadian policymakers and justice system stakeholders.

The primary authors of this report, Jill R. Presser and Kate Robertson, are both practicing criminal defence lawyers with active practices and experience litigating at all levels of courts in Canada. Both are actively engaged with, and have experience in, litigating at the intersection of emerging technologies and the law.[3]

The litigation perspective is extremely important to the analysis of probabilistic genotyping (PG) tools in Canadian criminal justice, which is the specific subject of this report. As litigators, the primary authors have endeavoured to bring to it a practical and realistic view of how PG is being (and will likely be) used in Canadian criminal courts.

One of the benefits of litigation is that it can expose the flaws and legal uncertainties of new technologies. This has been evidenced across North America where new AI and algorithmic tools in both criminal and civil justice systems have been heavily litigated with dismaying results: the exposure of systemic flaws including inaccuracy, unreliability, unintended biases, opacity, lack of explainability, data illiteracy, automation bias, etc.[4] The common theme that emerges from all litigation is that AI and algorithmic tools were introduced *prior* to serious discussions about proper guardrails governing their use. Regulation by litigation is not a sufficient answer to the challenges and risks posed by these systems. Proactive regulation is required.

The specific litigation case study of this paper is probabilistic genotyping (PG). PG uses artificial intelligence algorithms to analyze DNA samples collected in police investigations or criminal prosecutions. Such samples may not be sufficiently pristine, intact, or distinct to generate the singular genetic fingerprint needed for a definitive "match." Samples collected from victims, weapons, objects or places often consist of a cocktail of incomplete or fragmentary DNA from any number of different sources. PG is used to analyze DNA samples that could not otherwise be analyzed or would not otherwise yield information that could be used. The ability to identify, differentiate and correlate a specific sample to a specific person is often then a question of likelihood or probability based on more general traits and the relative "signal-to-noise" ratio of the sample itself. PG technologies use sophisticated statistical algorithms and AI-driven pattern recognition in the attempt to better hone legally relevant inferences from these kinds of probabilistic DNA samples. In turn, these inferences may be admitted into evidence as part of a criminal prosecution and may ultimately lead to a conviction. Where PG DNA tools are not scientifically valid or accurate, or the limits of what they can tell us are not well understood by justice system participants, they may contribute to wrongful convictions or exonerations.

This report is intended, in part, to answer the following related questions about PG:

1. How are *Charter*, due process, and common law rights affected by these tools and what do the *Charter*, due process, and the common law say about these tools?
2. What statutory, regulatory, evidentiary, or other changes are necessary to ensure that any use of PG DNA tools complies with law and to protect access to justice?

It is important to ask these questions now because AI-driven technologies are spreading rapidly in the legal world.

PG tools are in widespread use across North America and in other parts of the world.[5] They are extremely significant for the purposes of this report because they are the first artificial intelligence tools in regular use in the criminal courts in a number of Canadian jurisdictions.[6] They also purport to generate positive evidence that goes to guilt or innocence. As a result, ensuring the accuracy and reliability of evidence that comes of out of PG is essential to ensuring that justice is done in individual cases.[7]

For these reasons, PG tools present an imminent need for critical review and challenge. As explored in LCO's forward-looking companion paper on AI and bail, algorithmic risk assessment tools are not yet in widespread use in criminal justice in this country. They purport to generate risk prediction profiles, but not actual evidence that goes directly to whether an accused should be convicted or acquitted. This stands in sharp contrast with PG tools that are in widespread and regular use in this country, and which do aim to produce potentially very inculpatory evidence. And yet, while algorithmic risk prediction in bail and sentencing has been studied and challenged, PG as been largely overlooked.[8]

The novel introduction of AI-driven and probabilistic technologies introduces a new set of procedural and evidentiary ambiguities which may be too easily glossed over or simply unknown. This can undermine the integrity of the criminal justice system if issues unique to these technologies are not clearly identified, problematized and effectively communicated to—and understood by—accused, advocates, prosecutors, and adjudicators alike.

Some of the issues discussed in this paper will likely be familiar, such as well-known "black box" criticisms of algorithmic decision-making, the challenge of explicability, and concerns for biases ingrained in datasets. This paper also considers how legal, technical and operational choices about AI and algorithms can affect bias, due process, access to justice and the fairness of government and justice system decision-making.

Other topics are likely to be new or unfamiliar. Any new technology that purports to generate potentially inculpatory evidence should be carefully reviewed for reliability and accuracy. PG is no exception to this rule. All the more so because it may actually raise particular reliability and accuracy concerns:[9]

For example, it is difficult for defence counsel to obtain meaningful disclosure of how these tools work, which limits counsel's ability to evaluate or challenge them. This is because in most cases, developers/owners assert that their PG programs are proprietary trade secrets that they cannot or will not disclose.[10]

Moreover, the DNA evidence generated by PG is often confidently received by judges and juries, possibly because it benefits from the positive halo associated with the infallibility of science. It appears to be methodical, evidence-based, and immune from the individual and unpredictable judgments of subjective human actors. In other words, PG tools appear to be objective and authoritative.[11] This appearance can be deceiving. PG, like all tools built and operated by humans, is susceptible to error.[12] Cognitive bias may affect the reliability of forensic scientific evidence.[13] Subjectivity can also get baked into the way an algorithmic tool is built, or may come in through the assumptions of the particular analyst operating it.[14]

Finally, as discussed at length in LCO's companion paper on bail, all AI tools are prone to algorithmic bias. Unreliable, inaccurate, or biased PG DNA evidence that goes unchallenged or is uncritically received in evidence may work great injustice. Litigators would be well advised to avoid such miscarriages of justice by seeking meaningful disclosure, educating themselves, retaining relevant experts, and where appropriate, challenging the admission of PG DNA evidence.

II. ABOUT THE LCO

The Law Commission of Ontario (LCO) is Ontario's leading law reform agency. The LCO provides independent, balanced and authoritative advice on complex and important legal policy issues. Through this work, the LCO promotes access to justice, evidence-based law reform and public debate.

LCO reports are a practical and principled long-term resource for policymakers, stakeholders, academics and the general public. LCO's reports have led to legislative amendments and changes in policy and practice. They are also frequently cited in judicial decisions, academic articles, government reports and media stories.

This report is part of the LCO's ongoing ***AI, ADM and the Justice System*** project. The first phase of this project brings together policymakers, legal professionals, technologists, NGOs and community members to discuss the development, deployment, regulation and impact of AI and algorithms on access to justice, human rights, and due process. The LCO's project considers this technology in both the criminal and civil/administrative law justice systems. Completed initiatives within this project include:

- LCO/Ontario Digital Service Workshop;[15]
- *The Rise and Fall of Algorithms in the American Justice System: Lessons for Canada*;[16]
- *Regulating AI: Critical Issues and Choices*.[17]
- LCO Forum on AI and ADM in the Civil and Administrative Justice System;[18]
- LCO Forum on AI in Ontario's Criminal Justice System (with The Citizen Lab, Criminal Lawyers Association and the International Human Rights Program, Faculty of Law, University of Toronto);[19]
- ***AI, Automated Decision-Making: Impact on Access to Justice and Legal Aid***;
- ***AI for Lawyers: A Primer on Artificial Intelligence in Ontario's Justice System*** with Element AI and Osgoode Hall Law School; and,
- ***Roundtable on Digital Rights and Digital Society*** with the Mozilla Foundation.

The LCO is also undertaking projects respecting Protection Orders, the Last Stages of Life, the Indigenous Last Stages of Life, and Environmental Accountability. ***Protection Orders***, ***Last Stages of Life***, ***Indigenous Last Stages of Life***, and ***Environmental Accountability***.

More information about the LCO is available at ***www.lco-cdo.org***.

III. DNA EVIDENCE AND WHAT'S NEW ABOUT PG IN CANADIAN CRIMINAL COURTROOMS

DNA evidence has been received in evidence in the justice system for a relatively long time (as new forensic tools go), to great effect in securing convictions and exonerations. It has been routinely accepted as reliable and authoritative in our courts.[20] But the DNA evidence we are accustomed to receiving in the justice system, "old school" DNA evidence, is not the same as the DNA evidence that results from analysis of samples using PG.[21]

1. Random Match Probability

The usual method of analyzing DNA is called Random Match Probability ("RMP"). This is the DNA evidence with which most justice system actors are familiar. It is typically and most effectively used to analyze uncomplicated crime scene samples: DNA that came from a single source, with relatively high quantities of DNA, predominantly from bodily fluids like saliva, semen, or blood. In this process, a forensic lab will go through several steps to deduce a reliable DNA profile from a sample, and then report on the rarity of the profile (likelihood of a coincidental match) and on whether two profiles are a match.[22] In this manner, RMP measures the probability that a DNA match was random or coincidental.[23] This method of analyzing DNA evidence has been called the "gold standard" of forensic evidence. It presents a non-evaluative fact: a description that corresponds to something real, empirical, objectively true in the world.[24]

Unfortunately, RMP works reliably only for uncomplicated DNA samples. It does not work as well in analyzing DNA samples where there are multiple contributors, degraded samples, or small samples.[25]

As police began swabbing objects at or near crime scenes for cells that may have been left behind, they started obtaining "trace" DNA evidence. Trace evidence often contains small amounts of DNA that may only provide an incomplete or fragmentary DNA profile. It often contains mixtures of more than one person's DNA, because DNA may be transferred to an object by anyone and everyone who has contact with it.[26] You can readily imagine how many sources of DNA may be found on a light switch or doorknob in a public (or private) building.[27]

Importantly, "trace" DNA samples often only consist of incomplete fragments of DNA. This means that, unlike the RMP "gold standard," a trace sample may not be able to be linked to a specific individual with a high degree of probability. Instead, the "trace" sample may only indicate a profile of the kinds of generally inherited sequences of DNA that a person may have. Further, this profile itself may be incomplete, and it may be difficult to separate and interpret the DNA profile "signals" from the "noise." In "trace" samples that are incomplete, and which may include DNA from more than one contributor (or possibly even from multiple contributors), it can be especially difficult to make sense of the DNA present.

2. Probabilitistic Genotyping

In contrast to RMP tools, PG tools are algorithmic programs that interpret the results of complex DNA mixtures. They were developed in an attempt to draw out some of the inferences that may be difficult to interpret in "trace" fragmented, mixed, and complex samples. There are a number of different PG programs, some of which are open source while others are proprietary commercial tools.[28] They all use statistical methods and mathematical algorithms in DNA profiling.

Unlike RMP, PG profiles cannot confirm that a particular person is a source of the DNA sample. Rather, PG profiles may reveal the *kind* of suspect you might be looking for. This is because PG tools compare the probability of two different hypotheses or theories of the case that seek to explain the source of the DNA in the sample. The tools then determine which hypothesis is more probable. Typically, one hypothesis would be the police/Crown hypothesis, namely that the suspect is the source of some of the DNA in the mixture. The other hypothesis is usually a defence hypothesis (or what the forensic analyst assumes the defence hypothesis would be), namely that the suspect is not the source of some of the DNA in the mixture. Importantly, this process does not include defense input: Rather, the hypotheses tested are determined by the technician operating the PG tool, usually a forensic analyst working at a government-operated forensic lab.[29] As a result, the hypothesis tested as purporting to represent the defence hypothesis may not, in fact, be the actual defence theory.[30]

The two hypotheses are compared, with a view to determining which is more probable. The AI tool runs the two hypotheses through its algorithm to produce a "Likelihood Ratio" or "LR." The LR is an expression of which of the two hypotheses is said to be more likely.[31]

The LR *appears* to answer the question a jury is asked to determine, namely, whether the defendant was the source of some of the DNA at issue. However, LRs do not actually answer this question. Instead, they only "weigh the relative likelihood of two very specific hypotheses."[32] No matter how unlikely two hypotheses are in the real world, when compared and contrasted, one will always be more statistically likely than the other. As a result, there is an implicit false binary in the way that the LR is phrased. The LR that results from PG analysis expresses a relationship between two hypotheses in relation to their relative likelihood, but it does not tell us how objectively likely either hypothesis is in the real world.[33]

Bess Stiffelman provides the following non-DNA-evidence-related example of how LRs express which of two hypotheses is more likely, without expressing the relationship of either hypothesis to objective truth:

> *One could propose any two hypotheses to answer a question, and theoretically come up with a likelihood ratio. Suppose you returned home to find your dog sitting next to a torn-up pillow, and feathers are everywhere. You could compare the hypothesis that your home was ransacked by burglars with the hypothesis that the pillows you bought at Ikea were designed to explode after six months. After gathering all the relevant data, like that your door was locked when you came home, and nothing else was broken, and the strength of the pillow fibres, you could come up with a likelihood ratio. Let's assume you determine the likelihood ratio that, given the evidence, it was 10,000 times more likely that the pillow just fell apart than it was likely your home was ransacked by burglars. Neither of these hypotheses are in fact the correct explanation for the evidence before you. Of course, your dog played with your pillow like a chew toy and tore it to shreds. So, although it was 10,000 times more likely the pillow just fell apart than it was likely your home was ransacked, this doesn't prove your pillow fell apart on its own. Both of the hypotheses were in fact incorrect.*[34]

The LR that results from PG analysis can only ever report a relationship between two hypotheses and say which is the more likely. It cannot tell us whether either of them is true. As a result, LRs do not express the probability that a suspect is the source of some of the crime scene DNA. Recognizing this, the National Research Council has written that "[t]he likelihood ratio is still one step removed from what a judge or jury truly seeks—an estimate of the probability that a suspect was the source of a crime sample, given the observed profile of the DNA extracted from samples."[35] This is a subtle but crucial distinction that is not easy to understand, and for this reason, LRs are easily and likely misunderstood as evidencing the probability of the defendant's guilt.[36] The actual meaning of the evidence produced by PG algorithmic tools is both difficult for expert witnesses to meaningfully convey and for judges and juries to understand.[37]

In order to convert what the trier of fact actually needs and wants to know (which is the probability that the defendant is the source of some of the DNA), a further step is required. This involves complex statistical analysis, known as Bayesian probability analysis. Bayes' theorem requires the making of certain assumptions about the probability of an event occurring. It describes the process by which information relevant to the *prior* probability of an event is combined with new information to produce a *posterior* probability.[38] In other words, to get at whether the defendant was probably a contributor to the DNA sample from the crime scene (the posterior probability), the trier has to decide the prior likelihood of the defendant having contributed to the sample (the prior probability), considered in light of the new information offered by the PG tool through the LR.[39] Statisticians calculate posterior probability (or the probability that the suspect did actually contribute some DNA to the sample) mathematically by multiplying the prior probability by the LR. This analysis is complex, but the conclusion is clear: PG tools or reports cannot stand alone as evidence of probability of guilt or the probability in the real world of the defendant having contributed to the DNA.

This analysis gives rise to several important issues and challenges to the use of PG tools in criminal proceedings.

The Presumption of Innocence

The use of PG tools can violate the presumption of innocence.

The essential variable of the prior probability of a defendant's contribution of DNA to a sample cannot be known. It must be assumed. If we assume *zero* prior probability of the defendant contributing DNA

to a sample, consistent with the presumption of innocence, the result will be zero posterior probability. Any number multiplied by zero equals zero. If, however, we assume prior probability that the defendant contributed DNA to the sample, we are presuming guilt. This violates the presumption of innocence.[40]

Without Bayesian probability analysis, the LRs produced by PG do not provide evidence of the probability that an individual's DNA is contained in a sample. In order for the Bayesian probability analysis to provide evidence of such probability, the presumption of innocence and the Crown's burden of proof must be violated. In criminal cases in Canada and the USA, only LRs are introduced as evidence—not the subsequent Bayesian probability analysis.

Bias and the Black Box

PG tools raise serious issues regarding algorithmic bias and the difficulty of understanding how these tools work.

Although PG algorithmic systems appear to be objective, and represent a step forward from purely subjective approaches to the interpretation of complex DNA mixtures,[41] there are many ways in which assumptions by the tool's developer, the lab, or the analyst factor into their results. Researcher Itiel Dror has said that it is a mistake to think that PG systems are able to objectively analyze DNA mixtures. Like humans, algorithmic tools must make assumptions about what is and is not useful or relevant information. These assumptions are reflected in the statistical, biological, and computer models coded into each PG tool. Only a tool's developers know what those assumptions are.[42]

In the absence of rigorous, independent, and comparative scientific validation of PG systems, no one knows what the impacts of those assumptions are on the reliability and accuracy of the tool's results. The 2016 report of the President's Council of Advisors on Science and Technology (PCAST) recognized this in its addendum to its report:

> *While likelihood ratios are a mathematically sound concept, their application requires making a set of assumptions about DNA profiles that require empirical testing. Errors in the assumptions can lead to errors in the results. To establish validity with a range of parameters, it is thus important to undertake empirical testing with a variety of samples in the relevant range.*[43]

Scientific Validity

To date, there has been very little in the way of independent or comparative studies establishing the scientific validity of PG.[44]

The independent research studies that have been completed have only tested PG tools for a limited range of mixture types (in relation to the number of contributors of DNA to the sample, the ratio of minor contributors, and the total amount of DNA available for testing). There has been little research into the reliability or accuracy of PG analysis of more complex mixtures.

In 2016, the PCAST report concluded that market leading PG tools STRmix and TrueAllele appeared to be reliable only in limited circumstances. These tools had only been established to be reliable for the analysis of three person DNA mixtures in which at least 20% of the intact DNA came from the minor contributor and in which the DNA amount exceeded the minimum level for the method.[45] This means that PG tools have not been established to be reliable for DNA mixtures where more than three people have contributed DNA, where the minor contributor contributed less than 20% of the DNA, or where there was a low amount of DNA collected.[46]

Analysis Based on Assumptions

In addition to the impact of unknown assumptions coded into the algorithmic tool on the reliability and accuracy of PG DNA evidence, the assumptions of the lab or technician operating the tool can also strongly influence the results. These assumptions are reflected in the hypotheses the technician asks the PG tool to test for. The choice of hypotheses tested will impact on the LR that results. The hypotheses include assumptions about the number of contributors to a sample, and whether the contributors were related family-members or members of a small genetically similar racial or ethnic community.

A DNA mixture assumed to have three contributors (itself an assumption that may or may not be correct)[47] has four possible explanations or hypotheses for who those contributors could have been: (i) the complainant, the defendant, and an unknown person; (ii) the complainant and two unknown persons; (iii) the defendant and two unknown persons; (iv) three unknown persons.[48] The PG algorithm will generally only be used to compare two hypotheses to develop a LR for which of the two is more probable. The choice of hypotheses compared will determine the LR for that pair of hypotheses, but will say nothing at all about the possible hypotheses that are not tested.

In the real world, there many alternative potential hypotheses that could explain any given DNA mixture. Some of the contributors of DNA could be related family members or members of a genetically similar racial or ethnic group. LRs are produced, in part, based on statistics reflecting the frequency with which certain DNA markers are present in the population.[49] Some genes present more often within families or racial or ethnic communities than they do in the population at large. This means that assumptions about whether contributors to a sample are or are not related family members or share race or place or origin will affect the LR that is produced:

> *Family members will see different frequencies of certain genes, just as any discrete population will see different frequencies of certain genes—just as Norwegians will have a higher frequency of blue eyes than, say, Koreans, those of Caribbean descent will see a higher frequency of certain genes than those who are entirely unrelated. Thus, an LR will be different if the technician hypothesizes that the source of the sample was the individual suspect and two unknown individuals of Caribbean descent. We don't know how different.*[50]

Data Literacy and Complexity

For the reasons described above, there are real limits on what PG DNA algorithms can tell us. But these limitations are hard to understand. There is a high risk that triers will not understand LRs and their limitations, mistaking them as speaking to the probability of whether the defendant was, in fact, the source of some of the DNA in the sample.[51] In fact, the limitations in what the LR can tell us are so conceptually challenging that even many of the experts charged with giving PG DNA evidence in court may not fully comprehend them.[52] And, if they do, they may have difficulty conveying them meaningfully to triers of fact. Two experts at the National Institute of Science and Technology ("NIST") who studied the use of LRs in the courtroom concluded that the value of probability statements (like the LR) breaks down where such information must be conveyed from one person to another, as in courtroom testimony.[53] They reported that using LRs to describe the weight of evidence risks conclusions that are driven more by unsubstantiated assumptions than by actual data.[54]

Disclosure and Due Process

Many of the issues described above present themselves as challenges to disclosure and due process.

PG algorithmic tools have not been established to be reliable or accurate in independent comparative studies, beyond the parameters described above. Their uncertain accuracy and reliability is exacerbated by the fact that it is, practically speaking, very difficult to challenge PG tools in court. This is at least in part because it is difficult for defence counsel to obtain meaningful disclosure of how these tools work to be able to evaluate or challenge them. The developers claim that their PG systems are proprietary trade secrets that they cannot or will not disclose.[55] And PG tools will often exhibit the lack of transparency inherent in machine-learning methods, known as the "black box" problem.[56] Without such disclosure, it can be impossible to evaluate how a system does or doesn't work, whether it is or isn't reliable or accurate.

In addition to making it impossible to test PG tools for reliability and accuracy, inadequate or non-existent disclosure may also violate due process and constitutional rights. The right to disclosure facilitates essential procedural and substantive rights for accused people protected under s. 7 of the *Charter*. These include the right to know the case against them, and the right to make full answer and defence.

In addition, the DNA evidence generated by PG is often confidently received by judges and juries, possibly because it benefits from the high degree of confidence typically associated with both scientific evidence and artificial intelligence. It appears to be methodical, evidence-based, and immune from the individual and unpredictable judgments of subjective human actors. It appears to be objective and authoritative.[57] But PG, like all tools built and operated by humans, is susceptible to error.[58] Cognitive bias may affect the reliability of forensic scientific evidence.[59] Subjectivity may be baked into the way an algorithmic tool is built, or may come in through the assumptions of the particular analyst operating it.[60] And, like all algorithmic tools, PG is prone to algorithmic bias.[61]

Unreliable, inaccurate, or biased PG DNA evidence that goes unchallenged or is uncritically received in evidence may work injustice. It may result in misidentification of the contributors to crime scene DNA, and thereby lead to wrongful convictions or wrongful exonerations.

Despite these shortcomings, PG evidence is widely admitted in US court proceedings and increasingly in Canada. John Buckleton, the developer of one of the market-leading PG tools, STRmix, reports that as at the time of writing, his tool had been used in 123,315 criminal cases around the world; STR-mix-generated evidence had been presented in court in 3,655 of these; but there had been only 32 admissibility hearings prior to its reception in evidence.[62]

Similarly, notwithstanding anecdotal evidence from many criminal defence lawyers in Ontario, of many cases in which STRmix was employed by the Centre for Forensic Sciences (CFS) in their clients' cases, there are very few reported decisions that explicitly refer to PG evidence. None of them report challenges to its admissibility. QuickLaw and CanLII searches of the terms "STRmix," "TrueAllele," and "probabilistic genotyping" in Canadian cases reveal only five reported decisions. None of these involve contested *Mohan voir dires* to consider the admissibility of this evidence.[63] In Canada, to date, there appears to have been only one case in which the admissibility of PG DNA evidence was considered, the unreported decision of *R. v. Klimowicz*.[64] Mr. Klimowicz was unrepresented. He asked no questions of the CFS DNA analyst. There was no defence expert called. Perhaps unsurprisingly, the PG DNA evidence was admitted in Mr. Klimowicz's case.

IV. UNIQUE CHALLENGES RAISED BY PG'S USE OF AI: LESSONS FOR THE CANADIAN CRIMINAL JUSTICE SYSTEM

This section summarizes a number of important issues and observations regarding the use of PG DNA evidence in criminal proceedings in Canada's justice system. These issues and challenges include unique challenges raised by PG's use of AI, which pose critical challenges and obstacles to the critical due process safeguards that are paramount in criminal law proceedings.

Issue #1 No longer the gold standard: The need for a new legal framework for PG DNA evidence

The admissibility and reliability of DNA probability statistics has been litigated a number of times in the context of "old school" RMP DNA evidence.[65] Some of this litigation focused on the potentially overwhelming impact of DNA probability statistics, and how difficult they may be for jurors to understand in context. The jurisprudence generally stands for the proposition that these concerns go to weight, rather than to admissibility. Limiting instructions have generally been held to suffice to address any prejudice that might otherwise arise from expert and statistical RMP DNA evidence.[66] As noted above, PG DNA evidence has yet to be litigated in Canada. But when it is, the existing jurisprudence governing RMP DNA evidence likely should not be simply applied by analogy to PG DNA evidence. RMP is highly tested and the validation approach is well known, but PG DNA tools are neither well tested nor validated in the same way.[67] In addition to validation for their biological DNA analysis, PG tools must also be validated for their algorithmic software, a feature that RMP DNA analysis does not share and that will not have been validated for RMP DNA analysis methods. For these

reasons, the existing legal framework that governs the admissibility of RMP DNA evidence may not be appropriate for courts' determination of the admissibility of PG DNA evidence. A new legal framework, one that specifically addresses the unique legal challenges and scientific complexity of PG DNA evidence may be required.

Issue #2 Data (il)literacy, limitations on probative value, and risk of prejudicial effect

As described earlier, in Part 3, there are unique and complex issues associated with PG methods, and the difficulty that humans have in comprehending the (limited) value of LR evidence. This difficulty carries with it a prejudicial danger that triers may incorrectly assume it to be evidence of actual contribution of DNA to the crime scene sample, and therefore incorrectly understand it to be probative of guilt. As further noted above, without Bayesian probability analysis, the LRs produced by PG do not provide evidence of the probability that an individual's DNA is contained in a sample. In order for the Bayesian probability analysis to provide evidence of such probability, the presumption of innocence and the Crown's burden of proof must be violated.

Litigation surrounding the admissibility and weight of PG-generated LRs will likely be required to assess the compatibility of PG evidence with existing common law frameworks applicable to the assessment of evidence that protect the presumption of innocence.[68] LRs are problematic because they present the relative likelihood that the prosecution's hypothesis is correct, as compared to the defendant's, on the issue of the originating identity of a DNA sample. However, disbelief of a hypothesis does not prove its opposite.[69] In criminal trials, triers are prohibited from putting a defence hypothesis into a contest with the prosecution's case, as such an approach is inconsistent with the presumption of innocence. The Crown prosecution must establish its case beyond a reasonable doubt. It is not a question of which hypothesis as to DNA identity is more likely to be true. For example, in *R. v. A.P.*,[70] the Court of Appeal for Ontario held that the trial judge made an error by focusing on which "version of events" was "most probable in all of the existing circumstances."[71] The Court of Appeal wrote that "by setting up the whole case as a choice between two competing versions of events and stating that the version that is 'most probable in all of the existing circumstances' will be selected as 'true', the trial judge came dangerously close to deciding the ultimate issue at trial on a balance of probabilities."

Issue #3 Automation Bias

Justice system participants may be inclined to accept the evidence generated by PG DNA tools, without adequate consideration of their scientific validity or accuracy. This is in part because of the well-documented tendency of lay people to be favourably impressed by expert scientific evidence.[72] Studies reveal that members of the public see DNA evidence, in particular, as reliable, accurate, and authoritative.[73] Many studies have found that jurors see DNA evidence as more than 90% accurate.[74] Because it is perceived as incredibly reliable and accurate, grounded in the hard sciences and objective, DNA evidence is usually particularly damning or exonerating. Without necessarily understanding the evidence, the science that underpins it, or its potential for error, jurors are likely to consider DNA evidence as infallible evidence of guilt.[75]

In the case of PG AI tools, the tendency of lay people to be impressed by DNA evidence is exacerbated by automation bias. This is the cognitive tendency to be impressed by, and defer to, the outputs of

apparently scientifically and technologically advanced artificial intelligence tools.[76] Since PG purports to analyze DNA evidence using artificial intelligence, it marries two types of tools that are regarded as highly authoritative by lay people. Triers approach PG DNA evidence already predisposed to find it determinative. The danger here is that the seemingly objective and scientific nature of PG DNA evidence will cause judges and jurors to place undue weight on it, without considering evidence which casts doubt on its reliability.

Issue #4 Algorithmic Bias

All AI tools are prone to algorithmic bias. Unreliable, inaccurate, or biased PG DNA evidence that goes unchallenged or is uncritically received in evidence endangers Canada's justice system and the vulnerable individuals who are affected by it.[77] Further study is required to examine AI-generated DNA evidence, and the existing potential for algorithmic bias and data bias in the algorithmic processing of DNA data sets and PG algorithms. Without careful attention and study to these issues, PG AI might exacerbate data biases because of the inferential nature of the method, and its presented hypotheses. Evaluating the potential for bias in the DNA data sets is one means of guarding against the potential for continuing or worsened discriminatory impacts within the criminal justice system on racialized and Indigenous individuals, and other individuals with protected characteristics under section 15 of the *Canadian Charter of Rights and Freedoms*.

Issue #5 Barriers to transparency and due process arising from private sector involvement

The inner workings of PG AI tools are difficult to discover, evaluate, or litigate because their developers typically resist making meaningful disclosure. They do so on the basis that their tools are proprietary trade secrets. Market leading PG tool TrueAllele has forcefully resisted efforts to require it to disclose its source code or algorithms. As of January 2017, defendants in at least seven US states had sought disclosure of TrueAllele's code in their criminal cases. Disclosure was denied in all cases on the basis of the developer's right to protect its intellectual property.[78] In one recent case, a New Jersey appellate court allowed the defence application for disclosure of the source code of TrueAllele.[79] It remains to be seen whether TrueAllele will appeal this decision.

The other market leading PG tool, STRmix, will make its underlying information, including source code and foundational validation research, available to defence experts for review upon signing a non-disclosure and confidentiality agreement.[80] However, the restrictions placed on this defence access are so extensive that one is left wondering whether there is much utility in it at all.[81] Ontario's Centre of Forensic Sciences ("CFS"), which uses the STRmix software, advises that it does not have access to STRmix's source code. All disclosure must come from the company that owns the tool, under the terms of its access policy and pursuant to the terms of its nondisclosure agreement. In addition, CFS does not routinely disclose its internal validation information. It has, however, disclosed such information in isolated cases in the past, in response to specific requests.[82] To date, one defendant in Quebec has also sought to obtain disclosure in a criminal case in respect of the STRMix, including the validation study performed by Quebec's forensic science laboratory. The Court rejected the defendant's application for disclosure.[83]

Issue #6 Barriers to access to justice presented by complex, multi-disciplinary methods

PG AI tools involve complex and inter-connected statistical, biological, and algorithmic mathematical models. In the US case of *People v Gissantaner*, a Michigan court commented on the multidisciplinary nature of the PG tool at issue in that case: "[i]t is a combination of forensic DNA techniques, mathematical theory, statistical methods . . . decisional theory, computer algorithms, interpretation, and subjective opinions . . . "[84]

The complexity and multi-disciplinary nature of PG algorithms has implications for the validation of these tools. Independent scientific study must validate their biological DNA analysis, their statistical modelling, and their algorithmic software. In addition, the complexity and multi-disciplinary nature of PG algorithms has implications for the justice system. It may be difficult to find relevant independent experts (meaning experts outside of government forensic DNA labs) to assist defendants and the court.[85]

Most litigators and adjudicators don't understand the technology well enough to know what to challenge or question when PG tools are used. Even forensic experts can struggle to be able to meaningfully review and analyze PG evidence. PG tools are designed to process DNA samples that are too difficult for human analysis. The tools rely on mathematical formulas that are understood and reviewable only by a relatively small pool of experts. Even if disclosure is received, it still remains for counsel and the Courts to interpret very technical material such as source code, large aggregates of data, statistical modelling and the like. Jurists must develop some expertise, or at least a working familiarity with DNA statistical modelling, as well as with the PG DNA AI technology and the issues it raises. Lawyers litigating PG tools, and adjudicators presiding over such litigation, must overcome technological illiteracy to become conversant with how the technology works.[86] Relevant experts can be of assistance, if they and the resources to retain them are available.

Issue #7 Critical challenges to due process protections

Individuals embarking on litigation in criminal cases regarding the admissibility and reliability of PG evidence will face unique challenges that undermine access to due process safeguards. We have already examined some of the dangers associated with the admission of PG evidence in criminal trials, having regard to questions concerning the reliability of PG evidence and the court's truth-seeking function. Access to justice and due process concerns pose related challenges when it comes to the use of PG evidence by the prosecution. Just as the complexity and lack of transparency of PG evidence can create a danger that unreliable, misleading, or inaccurate evidence may be used to inculpate a defendant, so too can the complexity and lack of transparency of PG evidence make it particularly difficult for defendants and their counsel to mount a meaningful challenge to this class of evidence.

Issues relating to due process and access to justice in criminal litigation involving PG evidence include the following:

- **Difficulty obtaining meaningful disclosure:** As described above, it is difficult to access meaningful disclosure of PG tools and how they work in the face of claims that the technology is intellectual property protected from disclosure as proprietary trade secret. This, in turn, makes it difficult for a defendant to challenge the tools for accuracy and reliability, and interferes with the ability to make full answer and defence.

- **Navigating the tools' lack of transparency:** Even if the use of PG were to be disclosed as a matter of course, the workings of the AI is prone to lack of transparency as a "black box." This too makes it difficult for defendants to know or understand exactly how a PG tool used in their case worked, how it came to its conclusions, to test those conclusions for accuracy, or otherwise challenge the tool's outputs in court.
- **Limited access to independent experts:** As described above, one of the distinct features of PG algorithmic tools is that they are used in circumstances where it is too difficult for the human mind to analyze DNA samples. This tends to pose problems from an access to justice perspective in criminal litigation, as there is only a small pool of experts who are capable of meaningfully analyzing and challenging the mathematical processing through PG of information gleaned from complex DNA mixtures. Even the availability of these experts may be limited by geographic distance, limited availability in busy schedules, or the existence of a conflict of interest in the litigation if previously retained by the prosecution.
- **Funding litigation:** Another challenge in litigating PG algorithmic tools arises from the need to raise adequate funding to mount legal challenges, including the cost of time required by legal counsel to gain technological literacy and mount these challenges properly, and the cost of hiring knowledgeable experts. Litigating PG DNA tools and the evidence they generate is bound to be complex, lengthy, and costly. Automation bias (the tendency to defer to technological outputs), and the opacity of most algorithmic tools, will aggravate this problem because they necessitate the hiring of experts.[87] Disputing PG tools and their results will thus be out of reach for those who cannot afford to hire counsel, let alone the required experts.[88] Research reveals that members of historically disadvantaged groups like low-income people and members of BIPOC communities are significantly over-represented in the criminal justice system.[89] For this reason, members of these communities are disproportionately likely to face PG DNA evidence in criminal court, and they are typically the least resourced financially to challenge the tools and the evidence they generate.[90]
- **Limited access to opportunity for cross-examination:** In cases where the prosecution seeks to adduce PG evidence, the robust right to make full answer and defence may be undermined. Cross-examination of witnesses "forms the foundation of the adversarial trial process for there is no better vehicle in which to test the veracity and weight of an adversary's evidence than through cross-examination."[91] Opportunity for cross-examination relating to the reliability of PG evidence may be limited if no human witness with direct knowledge is proximate or available. As noted, because of the complexity and opacity of PG tools, even forensic analysts who operate them may not understand or be able to explain core issues about how they operate, limitations in what the LR means and does not mean, what assumptions were built into the code, and what impact those assumptions had on the LR that results. The opacity or 'black box' nature of PG tools may mean that no human may be privy to the inner workings of the algorithmic processes they employ. Cross-examination may be limited by assertions of trade secrecy if forensic analysts are unwilling or contractually unable to testify or be examined on matters covered by claims of trade secrecy.

  For these reasons, the prosecution may not be able to produce witnesses who can provide complete or meaningful evidence about PG DNA tools and the LRs they produce. Cross-examination of the witnesses who are produced may be hampered by the limitations in the witnesses' knowledge, understanding, and ability to communicate in light of the complexity and lack of transparency of the tool, or as a result of claims of trade secrecy. It may be impossible for the defence to elicit evidence that is key to making full answer and defence.

- **Access to training:** Justice-system participants will generally lack the baseline knowledge and technological literacy necessary to mount and conduct a meaningful analysis of the reliability of PG DNA evidence. Both lawyers and adjudicators will need to undergo training (either formal or through self-study) to enable a comprehensive understanding of the nature of PG evidence and its attendant risks. Hiring knowledgeable experts may fill some or even all of the gaps in lawyers' own tech literacy. Similarly, court-appointed experts may assist in identifying and remedying gaps in adjudicators' tech literacy (as in the *Gissantaner* case)[92]. However, as noted above, where funding challenges do not allow for the hiring of defence experts or courts cannot or do not appoint their own independent experts, jurists' own knowledge will be essential.

V. CHOICES AHEAD: KEY LAW REFORM CONSIDERATIONS

The LCO's first AI Issue Paper, *The Rise and Fall of AI and Algorithms in American Criminal Justice*, identified a series concerning the use of algorithms in Canada's justice system, the LCO identified a series of issues and options that should be addressed prior to the widespread implementation of any AI or algorithmic system in the Canadian criminal justice system.[93] These lessons and observations included, for example:

- identifying challenges arising from algorithmic bias (referred to as the 'bias in, bias out' problem);
- challenges associated with measuring and defining fairness, having regard to non-discrimination principles;
- hinderances to algorithmic transparency; challenges associated with measuring and achieving reliability, accuracy, and validity in algorithmic methods; and
- the complex and contested intersection between law and policy that is implicated when human actors attempt to generate algorithmic predictions.[94]

The LCO identified in its earlier report that the first step forward is for policymakers and stakeholders to collectively agree to address these overarching issues and identify an appropriate process for doing so. Further, arising from the LCO's review of these issues, it is clear that the systemic legal issues raised by algorithms cannot be addressed through individual litigation, best practices, or piecemeal legislation. Comprehensive law reform is required.[95]

These lessons and observations are similarly cogent in the context of PG DNA evidence. All AI tools are prone to algorithmic bias. Unreliable, inaccurate, or biased PG DNA evidence that goes unchallenged or is uncritically received in evidence endangers Canada's justice system and the vulnerable individuals who are affected by it. As emphasized in the LCO's earlier paper, the LCO strongly supports broad participation in the design, development, deployment and governance of AI and algorithmic systems in the Canadian justice system. This participation must include technologists, policy makers, law makers, and, crucially, the communities who are likely to be most affected by this technology.

Building upon the overarching recommendations in *The Rise and Fall of AI and Algorithms in American Criminal Justice*, what follows below are further recommendations tailored to the unique challenges arising from PG DNA methods. When it comes to PG DNA evidence, given that this particular form of algorithmic tool is already being used on the ground in criminal trials, the need for action now is pressing. To avoid miscarriages of justice, comprehensive action is required to respond to the risks and

due process concerns identified in this report, and to ensure that any continued use of PG DNA evidence in Canadian criminal courts is consistent with the *Charter* and human rights at stake.

When considering risks surrounding the use of PG DNA evidence (or any AI tool), it is also important to highlight that Canada's criminal justice system is still a work-in-progress when it comes to implementing the recommendations from previous commissions of inquiry in the areas of forensic science and expert evidence. For example, in 2018, the Federal/Provincial/Territorial Heads of Prosecutions Subcommittee on the Prevention of Wrongful Convictions published its report, *Innocence at Stake: The Need for Continued Vigilance to Prevent Wrongful Convictions in Canada*,[96] describing the need for further reform in the area of forensic science. The Subcommittee described how commissions of inquiry continue to note "the dangers of ignoring the lessons of the past and the need for constant and continued vigilance."[97] The Subcommittee conducted a review of the adequacy of the implementation of past recommendations arising from commissions of inquiry.[98] In conducting its review, the Subcommittee identified a number of priority areas of continued reform in the area of forensic evidence.[99] The recommendations set out below are in keeping with the lessons and recommendations set out in these previous inquiries.

The recommendations set out below address the following four areas:

1. Statutory Reform: Amendments focused on the use of PG DNA analysis as evidence
2. Statutory Reform: Amendments focused on enhancing systemic transparency and accountability
3. Practices and Training
4. Legal Aid Plans
5. Areas for Further Research and Evaluation

1. Statutory Reform: Amendments focused on the use of PG DNA analysis as evidence

We recommend that the federal and provincial governments consider statutory amendments to prescribe the use of PG systems, to promote the accountability of PG systems, and to ensure PG systems comply with appropriate *Charter*, due process and evidential requirements. These areas of statutory reform are grouped into two broad categories: 1) amendments focused on the use of PG DNA analysis as evidence; and, 2) statutory amendments focused on enhancing systemic transparency and accountability surrounding the use of algorithms in criminal justice system.

a. Prescribed limits on the admission of AI-generated evidence: Amendments to the *Criminal Code* should be considered to prescribe limits on the admissibility of AI-generated evidence, such as PG DNA evidence, in criminal proceedings.[100] The merit of establishing presumptive limits on the admission of AI-generated evidence is high, given the unique challenges associated with AI-evidence and the imbalance of power, knowledge, resources, and expertise between individual defendants in the criminal justice system and the developers of AI-based tools. In balancing the interests of justice, relevant factors include probative value, prejudicial effects, and other circumstances that may impair the reliability of the PG DNA evidence.[101]

b. Presumptive inadmissibility in the absence of strict scrutiny: In *Criminal Code* reform applicable to AI-generated evidence, the presumptive inadmissibility of novel forms of forensic evidence should be entrenched through the codification of common law

principles.[102] The party tendering such evidence must bear a burden of proving reliability on the rigorous standard applicable to novel forms of expert evidence. Full adherence to placing the burden of proof on the moving party is particularly important to ensure that vulnerable defendants in the criminal justice system are not required to bear a heavy persuasive burden of showing the need for caution and strict scrutiny of novel, AI-based forensic methods.[103]

Strict scrutiny requires adherence to the Supreme Court's caution against placing undue weight on whether the technique has been previously accepted as legitimate.[104] Private vendors who sell PG DNA analysis tools tend to emphasize the number of previous cases where PG DNA evidence has been admitted into criminal proceedings. However, there are still no peer-reviewed, independent, or comparative studies establishing the reliability of the tools outside very limited parameters. In Canada, numerous reported cases document the admission of PG DNA evidence into evidence, but to date, there is no reported instance of any close consideration of the exigencies of interpreting complex DNA mixtures, and what those challenges may mean for the reliability and accuracy of PG DNA evidence.[105]

For example, the authors have been unable to locate any reported cases in Canada involving PG DNA evidence where the reasons for judgment note the characteristics of the originating DNA sample (including the number of DNA contributors to the mixture that was subject to analysis). The authors have also been unable to find any reported judgments involving PG DNA evidence that advert to the parameters for reliability according to existing validation studies, and whether these were met. Equally, the authors have been unable to locate any reported decisions involving PG DNA evidence that advert to the particular forensic laboratory's validation studies, and whether their own internal standards were met in the particular case.

c. **Amendments to *Criminal Code* provisions applicable to expert evidence:** The *Criminal Code* includes certain regulations governing the admission of expert testimony in criminal proceedings.[106] In the recent report by the University of Toronto's Citizen Lab and the International Human Rights Program, the authors highlight that AI-generated evidence may itself be considered as a form of expert opinion evidence. The analogy between AI-generated opinions and human expert opinions highlights an existing framework that may be drawn upon to develop robust due process protections applicable to AI-generated evidence.[107] To this end, *Criminal Code* amendments relating to notice requirements, access to disclosure, and requirements regarding the availability of human witness testimony in the courtroom should be carefully considered to address the unique challenges associated with AI-generated evidence such as PG DNA analysis.

d. **Amendments to *Evidence Act* provisions applicable to electronic documents:** Amendments should address what distinctions should be drawn, where necessary and appropriate, between electronic documents stored on traditional information management and computer-processing systems, and new complex AI-based forensic methods.[108] This review should consider the common law interpretation of these provisions, including jurisprudence that has addressed the authentication of electronic documents generated by automated processes.[109] Particular attention should be paid to the circumstances in which parties would be able to adduce electronic records in the absence of an expert human witness to introduce them.[110]

2. Statutory Reform: Amendments focused on enhancing systemic transparency and accountability

a. **Modernizing of Canada's system of oversight governing the collection, use, retention, and sharing DNA information within police investigations and criminal proceedings:** Comprehensive review is required to keep pace with advances in forensic investigative methods, to protect *Charter* rights, and to enable accountability surrounding the use of sensitive techniques.[111] Reforms should aim to facilitate public access to information regarding what DNA databases law enforcement authorities and forensic laboratories obtain access to; where DNA profiles are sourced from; and the relationship between private companies (such as PG vendors and genetic ancestry services) and public sector actors regarding the collection, sharing, and use of DNA profiles. Additional attention is warranted to establish best practices surrounding the relationship between criminal investigators and forensic laboratories when it comes to PG DNA analysis, given the potential for human error and confirmation bias in the use of PG tools and LRs.[112] The Citizen Lab and the IHRP recommended that provincial and territorial governments enact Ministerial directives to address gaps in existing oversight mechanisms for law enforcement's use and experimentation with algorithmic technologies. The use of Ministerial Directives within a broader and comprehensive system of oversight governing the responsible development and use of algorithms is consistent with findings set out in the *Report of Ipperwash Inquiry*.[113]

b. **Enacting regulations requiring transparency and accountability with respect to algorithms that will be used for criminal justice purposes, including openness of source code and source code review:** In *To Surveil and Predict*, the Citizen Lab and IHRP set out recommendations directed at ensuring algorithmic accountability and transparency, openness and review of source code, and maintenance of public sector control over algorithms that will be used for criminal justice purposes.[114] We likewise endorse this recommendation, set out in the report as follows:

> *Government and law enforcement agencies that develop their own algorithmic… technologies—whether in-house or with private vendors—must make the source code and related details of such technologies publicly available in machine-readable and human-readable forms. Applying open software licensing in this context may be appropriate on the basis that such technologies are publicly funded and should be available to members of the public for purposes of review or research*
>
> *If extenuating circumstances engage the public interest to an extent that outweighs the importance of disclosing the source code and related details of algorithmic…technologies to the public, the information must still be accessible to oversight bodies and independent experts on a confidential basis (where the source code itself remains confidential, but findings and rulings are public) in order to complete impact assessments, procurement reviews, security testing, auditing, investigations, or court proceedings. Governments and law enforcement agencies that acquire off-the-shelf algorithmic policing technologies from commercial vendors must include open-source code review as a condition of all procurement contracts, including the waiving of trade secret and related proprietary rights where the review, testing, oversight, and auditing mechanisms listed above are concerned. Possible agencies that could play an oversight role with respect to ensuring source code transparency, review, testing, and auditing include the Office of the Privacy Commissioner and the Canadian Centre for Cyber Security.* [115]

3. Practices and Training

In order to ensure that PG tools are used appropriately and well-understood by all criminal justice system participants, we recommend:

a. **Establishing prosecutorial guidelines concerning the use of PG DNA evidence in criminal proceedings:** Guidelines may incorporate directions to mandate commitment to robust and full disclosure surrounding PG DNA methods to defendants and their counsel. Where appropriate, consideration should also be given to constructing safeguards to limit the circumstances in which the Crown will attempt to adduce PG DNA evidence in criminal cases, having regard to complexity and quality of the DNA mixture at issue and the critical need to guard against wrongful convictions. For example, as noted above, PCAST has found that PG tools have only been established to be reliable for the analysis of three person DNA mixtures in which at least 20% of the intact DNA came from the minor contributor and in which the DNA amount exceeded the minimum level for the method. Crown policy should mandate that prosecutors not rely on PG DNA evidence that results from samples that do not meet established parameters for validity and reliability.

b. **Developing access to training programs for all justice-system participants:** As noted above, both lawyers and adjudicators will need to undergo training (either formal or through self-study) to enable a comprehensive understanding of the nature of PG evidence and its attendant risks. Provincial law societies, Crown agencies, law enforcement agencies, and the National Judicial Institute should facilitate access to training. Where appropriate, regulators should also consider applicable professional standards for lawyers to ensure professional competency guidelines are modified as required to address the need for technological literacy when litigating constitutional laws, human right laws, and other evidentiary issues in criminal proceedings involving the potential admission of algorithmic evidence.

4. Legal Aid Plans

In order to ensure due process and access to justice for criminal defendants and the integrity and fairness of PG-involved criminal proceedings, we recommend:

a. **Review of legal aid programs to identify and remedy gaps in policies and budgetary constraints**: Legal aid programs, and the governments that oversee and fund their mandates, should be required to do a review of how PG DNA tools, and AI in general, will impact upon the low-income individuals that they are mandated to service, and should develop plans for how to provide adequate service to these clients as they will be affected by PG DNA and other AI tools in justice.

5. Further Research and Evaluation

In order to fully uncover sources of risk to *Charter* rights and wrongful convictions in Canada, we recommend further evaluation-through-litigation by justice system participants (including courts and litigators), as well as governmental funding to enable continued research by academics and experts:

a. **Researching and evaluating PG DNA methods and their potential for bias:** As described earlier in this report and the LCO's earlier work, algorithmic bias is a continuing and ever present danger surrounding the use of AI in the criminal justice system. Governments should make funding available to support both forensic laboratories and

independent academic research into PG DNA analysis. Further scientific study is critical to evaluate the benefits and risks of PG DNA evidence. As noted in *To Surveil and Predict*, if

> *algorithmic policing technologies continue to be used in the criminal justice system, due process and remedial rights will likely be hindered in practice by the lack of available expertise to test and challenge the validity of algorithmic methods.*[116]

Developing sources of independent expertise is critical in relation to issues such as the human perception of PG DNA evidence and automation bias; algorithmic bias and accuracy issues that may be present in DNA processing methods; and the accuracy and reliability of PG DNA analysis. Evaluating the potential for bias in the DNA data sets is one means of guarding against the potential for continuing or worsened discriminatory impacts within the criminal justice system on racialized and Indigenous individuals, and other individuals with protected characteristics under section 15 of the *Canadian Charter of Rights and Freedoms*.

b. Evaluating human rights and ethical concerns surrounding the use of DNA phenotyping in criminal investigations: In evaluating the potential for human rights risks such as privacy harms and discrimination in the use of DNA information in criminal justice proceedings, particular attention should be directed towards the use of DNA phenotyping in investigations and forensic laboratories. The use of AI to conduct DNA phenotyping, particularly with regards to profiling on the basis of characteristics that are protected under non-discrimination laws in Canada, must be the subject of careful consultation and review due to Canada's constitutional obligations towards Indigenous persons and under the *Charter*, and due to the human rights and ethical concerns associated with these sensitive techniques.[117]

c. Evaluating the scope of existing *Charter* rights to ensure that fundamental rights and freedoms apply with equal force in circumstances where AI-generated evidence is used in criminal investigations and proceedings: The use of big data and AI in criminal justice present unique and systemic risks. The existing jurisprudential development of *Charter* rights should be interpreted and applied contextually to ensure that all *Charter* rights operate meaningfully in the context of AI, and with equal force for all individuals including racialized and Indigenous persons. For example, if disclosure of AI tools falls to be decided under an overly restrictive disclosure regime, it will be extremely difficult to make full answer and defence in respect to the use of these tools.[118] Adherence to a robust right to disclosure assists to reduce the danger that flawed methods will remain hidden for too long under a scientific veneer of reliability advertised by AI vendors.

VI. CONCLUSION

The data and data science used at every stage of AI and algorithmic decision-making have human rights, equity, due process, and access to justice consequences. These tools often have built-in and largely hidden biases and discriminatory inferences in their decision-making. While appearing scientific, objective, and authoritative, they may be unreliable and invalid. Failure to study, understand, and regulate these tools can have significant system-wide and individual repercussions. Absent proper scrutiny, process, and legislation, there is a risk that AI tools, including PG DNA algorithms, will worsen racism in Canada's justice system, and put access to justice further out of reach for many Ontarians. Unequal access to participation in decision-making about data and technology can further entrench existing biases and inequality. Reliance on AI tools, like PG DNA algorithms, may provide what appears to be infallible and reliable scientific evidence against individuals who, by reason of the complexity and lack of transparency of the tools, lack of meaningful disclosure, and lack of access to resources to mount a legal challenge, are factually innocent. Because of the well-established over-representation of low income people and members of racialized and Indigenous communities in the criminal justice system, those most likely to face PG DNA evidence will often be those least likely to be financially resourced to challenge it. Absent proper scientific study, regulation, enforcement of relevant *Charter* rights, and due process protections, PG DNA evidence may lead to wrongful convictions. This concern is likely to disproportionately impact on communities that already suffer from systemic discrimination in the justice system and Canadian society at large.

VII. NEXT STEPS AND HOW TO GET INVOLVED

The LCO believes that successful law reform depends on broad and accessible consultations with individuals, communities and organizations across Ontario. As a result, the LCO is seeking comments and advice on this report. There are many ways to get involved:

- Learn about the project on the LCO website (***www.lco-cdo.org***);
- Contact us to ask about the project; or,
- Provide written submissions or comments on the Final Report and recommendations.

The LCO can be contacted at:

Law Commission of Ontario
Osgoode Hall Law School, York University
2032 Ignat Kaneff Building
4700 Keele Street
Toronto, Ontario, Canada
M3J 1P3

Email: **LawCommission@lco-cdo.org**
Web: **www.lco-cdo.org**
Twitter: **@LCO_CDO**
Tel: **(416) 650-8406**
Toll-free: **1 (866) 950-8406**

ENDNOTES

1 The complete series of LCO Issue Papers and more information about the LCO's AI, ADM and Justice System project is available at **https://www.lco-cdo.org/en/our-current-projects/ai-adm-and-the-justice-system/**.

2 AI Now Institute, Algorithmic Accountability Policy Toolkit, (October 2018) at 2, online: **https://ainowinstitute.org/aap-toolkit.pdf**.

3 As noted above, Jill R. Presser and Kate Robertson are also co-authors of "Algorithmic Technology and Criminal Law in Canada," Chapter 3 in Jill R. Presser, Jesse Beatson, and Gerald Chan eds., *Litigating Artificial Intelligence,* (Toronto: Emond: May 2021). We have drawn extensively on the portion of that chapter relating to PG tools in the writing of this report, with the generous permission of the book's publisher.

4 Law Commission of Ontario, *The Rise and Fall of AI and Algorithms In American Criminal Justice: Lessons for Canada*, (October 2020) [LCO Criminal AI Paper], online: **https://www.lco-cdo.org/wp-content/uploads/2020/10/Criminal-AI-Paper-Final-Oct-28-2020.pdf**; *Ewert v. Canada*, 2018 SCC 30; *R v. Mills*, 2019 SCC 22; *R v. Rogers*, 2016 ONSC 70; *R v. Spencer*, 2014 SCC 43; *State v. Loomis,* 881 N.W.2d 749 (Wisc. 2016).

5 John Buckleton, the developer of market-leading proprietary PG tool STRmix, reports on his website that STRmix is in use at 56 labs in the United States, all nine state and territorial labs in Australasia, and in 14 other labs elsewhere. This includes three labs in Canada, and labs in England, Ireland, Finland, Denmark, Switzerland, China, and Hong Kong. See **https://johnbuckleton.files.wordpress.com/2020/06/labs-live-ii.pdf**._

6 PG tool STRmix has been in use to analyze complex DNA mixtures the Centre of Forensic Sciences ("CFS") in Ontario since August of 2016: "... the Centre of Forensic Sciences ("CFS") has implemented an improved method for resolving mixed DNA profiles. The method, known as probabilistic genotyping, employs a specialized software application called STRmix™," CFS Technical Information Sheet, *Improving the Interpretation of Complex DNA Mixtures with Probabilistic Genotyping - A Guide to STRmix™ for Clients* (March 2017), online: **https://www.mcscs.jus.gov.on.ca/sites/default/files/content/mcscs/docs/STRmix%20Client%20Guide%20-%20March%202017.pdf**.

STRmix has also been in use since February of 2018 in Quebec at the forensic lab in Montreal, the Laboratoire de Sciences Judiciares et de Medecine Legale du Quebec, and since September of 2018 at the British Columbia Institute of Technology. See Buckleton 2020 at 2.

7 Meghan J. Ryan, *Secret Conviction Programs* (March 23, 2020). Washington and Lee Law Review (2020), Vol. 77, SMU Dedman School of Law Legal Studies Research Paper No. 454 at 2-3, online:
SSRN: **https://ssrn.com/abstract=3496055**.

8 Secret Conviction Programs 2020 at 17.

9 President's Council of Advisors on Science and Technology ("PCAST"), *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (September 2016). Report to the President at 8, 14, 75-82, online: **https://obamawhitehouse.archives.gov/blog/2016/09/20/pcast-releases-report-forensic-science-criminal-courts**; Secret Conviction Programs 2020 at 2, 28-29; James Foy, *DNA Analysis in Secret: Proprietary Algorithms and Forensic Science* (February 2019). For the Defence, Vol. 39, No. 3 at 28; Bess Stiffelman, *No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials* (November 26 2019). Berkeley Journal of Criminal Law, Vol. 24, (2019) at 125, 130, online: **https://lawcat.berkeley.edu/record/1129186?ln=en**.

10 DNA Analysis in Secret: Proprietary Algorithms and Forensic Science 2019 at 28, 31-32.

11 Secret Conviction Programs 2020 at 8, 17-18.

12 Secret Conviction Programs 2020 at 3, 8-9, 11-12.

13 Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 2016 at 31.

14 Amy Jeanguenat, Bruce Budowle, and Itiel Dror, *Strengthening Forensic DNA Decision-Making Through a Better Understanding of the Influence of Cognitive Bias* (July 29, 2017), online:
SSRN: **https://ssrn.com/abstract=3010738**.

15 This is a multidisciplinary workshop organized by the LCO and the Ontario Digital Service, a branch of the provincial government's Treasury

Board Secretariat, online: **https://www.lco-cdo.org/wp-content/uploads/2021/03/LCO-Govt-AI-Workshop-Report-%E2%80%94-March-2021.pdf**.

16 Law Commission of Ontario, *The Rise and Fall of AI and Algorithms in American Criminal Justice: Lessons for Canada* (October 2020), online: **https://www.lco-cdo.org/wp-content/uploads/2020/10/Criminal-AI-Paper-Final-Oct-28-2020.pdf**.

17 Law Commission of Ontario, *Regulating AI: Critical Issues and Choices* (April 2021), online: **https://www.lco-cdo.org/wp-content/uploads/2021/04/LCO-Regulating-AI-Critical-Issues-and-Choices-Toronto-April-2021-1.pdf**.

18 In December 2019, the LCO organized Canada's first multidisciplinary forum considering the use of AI and algorithms in regulatory investigations, government benefit determinations, and to support decision-making in the civil and administrative justice systems. This event brought together almost 40 policy makers, lawyers, jurists, technologists, academics, and community organizers to share experiences, discuss issues and consider law reform options in civil and administrative law applications. Read about this event **here**.

19 In March 2019, the LCO organized Canada's first multidisciplinary forum on AI in Canada's criminal justice system. The event brought together more than 50 policymakers, Crown Attorneys, defence counsel, jurists, technologists, academics and community organizers to discuss predictive policing, algorithmic risk assessments, how to "litigate algorithms" and related human rights and due process issues. Material for this event is available **here**.

20 See *R v. Terceira*, 1993 3 SCR 866.

21 No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 111.

22 No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 114-115, 127-128; For a more in-depth description of analysis of single-source DNA samples through RMP, see Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 2016 at 68-71 and Cecilia Hageman, *Forensic Biology and DNA* in Caitlin Pakosh, ed., *The Lawyer's Guide to the Forensic Sciences,* (Canada: Irwin Law: 2016).

23 Cecilia Hageman, "Understanding Genotyping for Lawyers" presented to the Law Commission of Ontario (June 19, 2020).

24 Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 2016 at 70.

25 Forensic Biology and DNA 2016 at 396-398.

26 For a good discussion of DNA transfer and secondary transfer, see No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 115-116; See also Forensic Biology and DNA at 412-413.

27 DNA Analysis in Secret: Proprietary Algorithms and Forensic Science 2019 at 39.

28 As of March 2014, there were at least eight different PG tools, including LRmix, Lab Retriever, like LTD, FST, Armed Xpert, TrueAllele, STRmix, and DNA View Mixture Solution. See Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 2016 at 78. It should be noted that FST has since been largely discredited for being unreliable, see *People v. Collins*, 49 Misc. 3d 595(2015), 15 N.Y.S. 3d 564, 2015 N.Y. Slip Op. 25227; Lauren Kirchner, *Thousands of Criminal Cases in New York Relied on Disputed DNA Testing Techniques*, ProPublica (September 4, 2017), online: **https://www.propublica.org/article/thousands-of-criminal-cases-in-new-york-relied-on-disputed-dna-testing-techniques**; and Lauren Kirchener, *Federal Judge Unseals NY Crime Lab's Software for Analyzing DNA Evidence*, ProPublica (October 20, 2017), online: **https://www.propublica.org/article/federal-judge-unseals-new-york-crime-labs-software-for-analyzing-dna-evidence**.

29 Understanding Genotyping for Lawyers 2020.

30 Understanding Genotyping for Lawyers 2020.

31 Improving the Interpretation of Complex DNA Mixtures with Probabilistic Genotyping – A Guide to STRmix™ for Clients 2017 at 4; DNA Analysis in Secret: Proprietary Algorithms and Forensic Science 2019 at 30-31; No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 112, 118.

32 No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 112.

33 No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 118-119.

34 No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 119-120.

35 National Research Council, *The Evaluation of Forensic DNA Evidence: An Update 201* (1996) at 201.

36 No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 118.

37 Jonathan J. Koehler, *On Conveying the Probative Value of DNA Evidence: Frequencies, Likelihood Ratios, and Error Rates* (1996). University of Colorado Law Review, Vol. 67, at 877-879, online: SSRN: **https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1432149**. See also Ronald Jay Allen and Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence* (January 2007). Journal of Legal Studies, Vol. 36, Northwestern Public Law Research Paper No. 900910, U of Alabama Public Law Research Paper No. 900910, online: SSRN: **https://ssrn.com/abstract=900910**.

38 Evaluation of Forensic DNA Evidence: An Update 201 1996 at 31-32, 132-133; No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 119, 136.

39 No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 133-139.

40 No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 133-139.

41 No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 133-139.

42 Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 2016 at 79.

43 Itiel Dror, quoted in Lael Henterly, "The Troubling Trial of Emanuel Fair," *Seattle Weekly*, (January 11, 2017), online: **https://www.seattleweekly.com/news/the-troubling-trial-of-emanuel-fair/**. It should also be noted that, in addition to assumptions coded into the algorithm, assumptions made by the human analysts who work in forensic labs and operate the tools may also affect or compromise a PG tool's results. In 2010, Dror and Hampikian gave 17 forensic scientists the same DNA mixture to analyze. They found that the results varied greatly. Those scientists who were given background or contextual information about the criminal case were more likely to find evidence that implicated the accused.

44 PCAST, *An Addendum to the PCAST Report on Forensic Science in Criminal Courts* (January 2017) at 8-9, online: **https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_addendum_finalv2.pdf**.

45 Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 2016 at 79-81. K. Kwong notes that the fact that validation studies have been performed by individuals associated with the developers of PG tools makes the validation "questionable in the eyes of the broader scientific community" in Katherine Kwong, *The Algorithm Says You Did It: The Use of Black Box Algorithms to Analyze Complex DNA Evidence*, (2017). Harvard Journal of Law & Technology, Vol. 31 at 289, online: **https://jolt.law.harvard.edu/assets/articlePDFs/v31/31HarvJLTech275.pdf**.

46 Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 2016 at 8, 79-82.

47 The Algorithm Says You Did It: The Use of Black Box Algorithms to Analyze Complex DNA Evidence 2017 at 288-289. See also, *People v Daniel Gissantaner*, 417 F Supp 3d 857 (WD Mich 2019). It should be noted that the trial decision in *Gissantaner* was overturned on appeal: *United States v Daniel Gissantaner,* No. 19-2305 (6th Cir. Mar 5, 2021). The appellate court admitted the PG DNA evidence, without ever considering – or even adverting to – PCAST's determination that PG results are only validated for three person samples where the minor contributor contributed a minimum of 20% of the sample. This was surprising in *Gissantaner*, where the minor contributor had only contributed 7% of the DNA in the sample.

48 There is no way to determine for certain the number of contributors from a DNA sample itself: No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 142.

49 Understanding Genotyping for Lawyers 2020.

50 The Evaluation of Forensic DNA Evidence: An Update 201 1996 at 89-102.

51 No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials 2019 at 122. Internal citation omitted in the quotation, but cites to The Evaluation of Forensic DNA Evidence: An Update" 201 1996 at 102-27, examining subpopulation statistics and issues of relatedness. It is not known whether PG tools are tested on or include genetic frequency statistics for insulated and genetically homogeneous communities, which may further impact on the validity of PG DNA evidence. Failure to account for genetic frequency statistics for insulated and genetically homogeneous populations may be one way in which PG Tools replicate bias. See The Algorithm Says You Did It: The Use of Black Box Algorithms to Analyze Complex DNA Evidence 2017 at 293; and Lauren Kirchner, "Traces of Crime: How New York's DNA Techniques Became Tainted," *The New York Times*, (September 4, 2017), online: **https://www.nytimes.com/2017/09/04/nyregion/dna-analysis-evidence-new-york-disputed-techniques.html?mcubz+3**.

52 On Conveying the Probative Value of DNA Evidence: Frequencies, Likelihood Ratios, and Error Rates 1996 at 877-879. See also The Problematic Value of Mathematical Models of Evidence 2007.

53 On Conveying the Probative Value of DNA Evidence: Frequencies, Likelihood Ratios, and Error Rates 1996 at 876.

54 "NIST Experts Urge Caution in Use of Courtroom Evidence Presentation Method: Use of 'Likelihood Ratio' not consistently supported by scientific reasoning approach, authors state" article on the NIST website (October 12, 2017), online: **https://www.nist.gov/news-events/news/2017/10/nist-experts-urge-caution-use-courtroom-evidence-presentation-method**. See also the article referred to: Steven P. Lund, and Hari Iyer, *Likelihood Ratio as Weight of Forensic Evidence: A Closer Look* (2017). Journal of Research of National Institute of Standards and Technology, Vol. 122, Article No. 27, online: **https://doi.org/10.6028/jres.122.027**.

55 NIST Experts Urge Caution in Use of Courtroom Evidence Presentation Method: Use of 'Likelihood Ratio" not consistently supported by scientific reasoning approach, authors state 2017. See also the article referred to: Likelihood Ratio as Weight of Forensic Evidence: A Closer Look 2017.

56 DNA Analysis in Secret: Proprietary Algorithms and Forensic Science 2019 at 28, 31-32.

57 LCO Criminal AI Paper 2020.

58 Secret Conviction Programs 2020 at 8, 17-18.

59 Secret Conviction Programs 2020 at 3, 8-9, 11-12.

60 Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 2016 at 31.

61 Strengthening Forensic DNA Decision-Making Through a Better Understanding of the Influence of Cognitive Bias 2017.

62 LCO Criminal AI Paper 2020 at p 20-22.

63 See **https://johnbuckleton.wordpress.com/strmix/**.

64 *R v. Ali*, [2019] O.J. No. 5049 (SCJ); *R v. Pereira*, [2019] O.J. No. 3682 (SCJ); *R v. Pereira*, [2019] O.J. No. 2411 (SCJ); *R v. Gautreau*, [2018] O.J. No. 4338 (SCJ); *R v. Dosanjh*, 2019 ONSC 1320.

65 *R v. Allister Klimowicz*, unreported decision of Ebbs. J of the Ontario Court of Justice, June 13, 2017.

66 See for example, *R v. Terceira*, 1993 3 SCR 866, *R v. Fisher*, 2003 SKCA 90, *R v. Eatman* (1999), 222 NBR (2d) 120 (CA).

67 For a good summary of the existing jurisprudence dealing with challenges to the admissibility of statistical probability evidence in the context of RMP DNA, see Forensic Biology and DNA 2016 at 428-430.

68 Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 2016 at 8, 79-82.

69 See, *R v. W. (D.)*, [1991] 1 SCR 742.

70 For example, there must be independent evidence of fabrication before an accused's statement at trial can used be as evidence of guilt: *R v. O'Connor* (2002), 62 O.R. (3d) 263 (C.A.); *R v. Coutts*, [1998] O.J. No. 2555 at para 13 (C.A.); *R v. St. Pierre*, 2017 ONCA 241 at paras. 5-

7. In *R v. MacIsaac* (2017 ONCA 172), Trotter J.A. held that a trier of fact errs in failing to provide reasons as to whether, or why, the accused's statements may be treated as positive evidence of guilt (at para 49).

71 2013 ONCA 344 at paras 38-40.

72 2013 ONCA 344 at para 40 [emphasis in original].

73 This phenomenon was recognized by the Supreme Court of Canada in *R v. Mohan*, [1994] 2 SCR 9.

74 The Algorithm Says You Did It: The Use of Black Box Algorithms to Analyze Complex DNA Evidence 2017 at 276.

75 Katie Worth, *The Surprisingly Imperfect Science of DNA Testing,* Frontline (June 24, 2015), cited in The Algorithm Says You Did It: The Use of Black Box Algorithms to Analyze Complex DNA Evidence 2017 at 276.

76 Secret Conviction Programs 2020 at 17-19, 34.

77 Lindsey Barrett, *Reasonably Suspicious Algorithms: Predictive Policing at the United States Border* (2017) New York University Review of Law & Social Change, Vol. 41:327 at 342-43, online: **https://socialchangenyu.com/wp-content/uploads/2017/09/barrett_digital_9-6-17.pdf**.

78 LCO Criminal AI Paper 2020 at 3, 22-23.

79 The Troubling Trial of Emanuel Fair Seattle Weekly 2017. See also The Algorithm Says You Did It: The Use of Black Box Algorithms to Analyze Complex DNA Evidence 2017 at 286-287.

80 *State of New Jersey v. Pickett,* Docket No. A-4297-19T4, Superior Court of New Jersey Appellate Division (February 3, 2021); See also *United States v. Ellis*, 19-369 (Wd. Pa. Oct. 23, 2020) following *Pickett* and ordering disclosure of source code.

81 *STRmix Access Policy* (April 2020), online: **https://www.strmix.com/assets/Uploads/Access-to-STRmix-Software-by-Defence-Legal-teams-April-2020.pdf**. See also *STRmix Full Nondisclosure and Confidentiality Agreement* (April 2020), online: **https://www.strmix.com/assets/STRmix/STRmix-PDFs/NonDisclosure-and-Confidentiality-Agreement-Full-NDA-Template-FINAL-April-2020.pdf**.

82 The restrictions on defence expert access imposed by STRmix include that defence experts will be afforded one-time access, in person, at the offices of legal counsel to the company in Akron, Ohio, under the direct supervision of a representative fo the company on a standalone inspection computer that is not internet-enabled. See STRmix Access Policy 2020.

83 Email exchange between Jill R. Presser and Jack Laird, Section Head – Biology, Centre of Forensic Schiences, May 3, 6, 7, 2020.

84 *Aithaqi c. R.*, 2020 QCCS 870.

85 *United States v. Gissantaner*, 417 F. Supp. 3d 857 (W.D. Mich. 2019), at 46-49.

86 It should be noted that the court in *Gissantaner*, retained independent experts to assist the court of its own motion. While there are some independent experts internationally, there appear to be relatively few in Canada outside of government forensic DNA labs, and these government experts will typically be retained by, and provide evidence for, the prosecution.

87 Sarah Valentine, *Impoverished Algorithms: Misguided Governments, Flawed Technology and Social Control* (April 29, 2019). Fordham Urban Law Journal, Vol. 46, 2019, online **https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3374005**.

88 Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control 2019 at 397-398.

89 mpoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control 2019 at 397-398.

90 See Dr. Ivan Zinger, Correctional Investigator of Canada, "Indigenous People in Federal Custody Surpasses 30% Correctional Investigator Issues Statement and Challenge" Public Safety Canada (January 21, 2020), online: **https://www.canada.ca/en/public-safety-canada/news/2020/01/indigenous-people-in-federal-custody-surpasses-30-correctional-investigator-issues-statement-and-challenge.html**; Kate Roberton, Cynthia Khoo, and Yolanda Song, To Surveil and Predict, A Human Rights Analyis of Algorithmic Policing in Canada, Citizen Lab and International Human Rights Program, Unversity of Toronto Faculty of Law, (September 2020) at 15-18, online: https://citizenlab.ca/2020/09/to-surveil-and-predict-a-human-rights-analysis-of-algorithmic-policing-in-canada/; *R v. Le*, **2019 SCC 34**; Ontario Human Rights Commission, *Under*

Suspicion: Research and Consultation Report on Racial Profiling in Ontario (April 2017), online: *Ontario Human Rights Commission* **http://www.ohrc.on.ca/en/under-suspicion-research-and-consultation-report-racial-profiling-ontario**; The Honourable Michael H Tulloch, *Report of the Independent Street Checks Review* (2018) at 44, online: **https://www.mcscs.jus.gov.on.ca/sites/default/files/content/mcscs/docs/StreetChecks.pdf**; Ontario Human Rights Commission, *A Disparate Impact: Second interim report on the inquiry into racial profiling and racial discrimination of Black persons by the Toronto Police Service* (10 August 2020), online: **http://www.ohrc.on.ca/en/disparate-impact-second-interim-report-inquiry-racial-profiling-and-racial-discrimination-black**.

91 These include *The Lamer Commission of Inquiry Pertaining to the Cases of: Ronald Dalton, Gregory Parsons, Randy Druken*, (June 2006) [The Lamer Inquiry]; *Report of the Commission of Inquiry into Certain Aspects of the Trial and Conviction of James Driskell*, (January 2007) [The Driskell Inquiry]; *Inquiry into Pediatric Forensic Pathology in Ontario*, (September 2008) [The Goudge Inquiry]; *Report of the Motherisk Hair Analysis Independent Review*, (December 2015); and the Motherisk Commission that was established on January 15, 2016.

92 *R v. Fife*, 2004 BCCA 261 at para 33.

93 STRmix Access Policy 2020; *STRmix Full Nondisclosure and Confidentiality Agreement* 2020.

94 LCO Criminal AI Paper 2020 at 3-4.

95 LCO Criminal AI Paper 2020 at 20-37.

96 LCO Criminal AI Paper 2020 at 38.

97 Federal/Provincial/Territorial Heads of Prosecutions Subcommittee on the Prevention of Wrongful Convictions, "Chapter 2: Forensic Evidence and Expert Testimony,"*Innocence at Stake: The Need for Continued Vigilance to Prevent Wrongful Convictions in Canada* (2018), online: **https://www.ppsc-sppc.gc.ca/eng/pub/is-ip/is-ip-eng.pdf** [*Innocence at Stake*].

98 Innocence at Stake 2018.

99 These include the *Lamer Inquiry* (2006), the *Driskell Inquiry* (2007), the *Goudge Inquiry* (2008), the *Motherisk Hair Analysis Independent Review* (2015), and the *Motherisk Commission* (2016).

100 The Subcommittees recommendations included, for example, the creation of a permanent national multidisciplinary group to study and make recommendations concerning aspects of forensic sciences in Canada; legislative reform to codify, clarify and enhance rules concerning the admissibility of expert opinion evidence; and, reform to Practice Directions at all levels of court focussing on the requirements for the admissibility of expert opinion evidence: Innocence at Stake 2018 at 130.

101 While the common law of evidence is in many circumstances a sufficient and robust safeguard to protect the right to a fair trial, in some circumstances, Parliament has enacted *Criminal Code* provisions to expressly set out preconditions to admissibility of certain types of evidence. See, e.g., *Criminal Code*, RSC 1985, c. C-46, section 276, section 269.1, section 277, section 278.92, section 320.31, and section 320.32.

102 Such circumstances might include: whether the limitations of the LR and what it actually conveys has been, or is capable of being, expressed to the trier; whether access to disclosure has been limited by assertions of trade secrecy by a private vendor of a PG DNA tool or other barriers; whether the AI method has been the subject of publication to enable peer review by independent experts; whether the proposed PG DNA evidence falls within the parameters of any existing validation studies; and whether the defendant is either unrepresented or under-represented due to limitations in the scope of legal aid coverage.

103 *Mohan* at 25; see also *R v Dimitrov*, **2003 CanLII 50104 (ON CA) at para 37**.

104 A growing body of research and case law also shows a range of contextual and historical factors demonstrating disproportionate and systemic challenges faced by racialized and Indigenous people in the criminal justice system. This includes the phenomenon of false guilty pleas that represent an ongoing source of wrongful convictions in Canada: Innocence at Stake 2018. It also includes disproportionate representation of racialized, Indigenous, and low income people in the criminal justice and correctional systems: Report of the Independent Street Checks Review 2018.

105 In *R v. J-LJ*, **2000 SCC 51** at para 34, the Supreme Court of Canada provided further guidance regarding assessing the admissibility of novel forensic science. The Court cautioned against placing undue weight on whether the technique is generally accepted as legitimate.

106 *Pereira*.

107 The *Code* enacts, for example, notice requirements and obligations to provide disclosure of the qualifications of the proposed expert: *Criminal Code*, RSC 1985, c. C-46, section 657.3.

108 This includes the safeguards and procedural protections that have been developed and recommended through commissions of inquiry in relation to forensic expert evidence.

109 Canada's and Ontario's *Evidence Act* enact provisions specifically applicable to electronic documents and information stored computer systems. The provisions pertain to the best evidence rule and the authentication of electronic records: *Canada Evidence Act*, RSC 1985, c C-5, s 31.1-31.9; *Evidence Act*, RSO 1990, c E.23, s. 34.1.

110 *R v. Dennis James Oland*, 2015 NBQB 245 at paras 99-106. The purpose of the traditional functional approach to the authentication of electronic records obtained form routine and non-controversial computing systems may be misplaced in the context of AI, particularly given the unique challenges that lie ahead with examining the reliability of AI-based evidence and problems associated with bias, accountability, and transparency.

111 In this review, consideration should be given to the imbalance of knowledge and expertise between vendors of AI-tools and individual defendants in criminal court, and the importance of examining the circumstantial indicators of reliability in evaluating the reliability and accuracy of opaque algorithms.

112 Close consideration may be given to the identification by forensic analysts of the total number of potential contributors in trace DNA samples, including insulating forensic analysts from information from criminal investigators regarding the theory of the case under investigation. Establishing silos between the forensic laboratory and criminal investigators may assist in reducing the risk that human error or other forms of confirmation biases undermine the accuracy and reliability of PG methods. For example, "human error can occur though sample mix-ups, contamination, incorrect interpretation and errors in reporting": Innocence at Stake *2018*.

113 The authors of The Citizen Lab's and the IHRP's report highlight that "[t]he Report of the Ipperwash Inquiry in Ontario supports the use of Ministerial directives as a method of introducing regulatory oversight for law enforcement activities on areas of particular concern to human rights": To Surveil and Predict, A Human Rights Analyis of Algorithmic Policing in Canada 2020 at 163-164, citing Sidney B. Linden, *Report of the Ipperwash Inquiry* (Toronto: Published by Ministry of the Attorney General, Queen's Printer for Ontario, 2007).

114 See generally, *To Surveil and Predict*.

115 To Surveil and Predict at 169.

116 To Surveil and Predict at 168.

117 For an interesting discussion of this issue in an international context, see New Zealand Law Commission, *The Use of DNA in Criminal Investigations* (October 2020), chapters 14 and 15.

118 Recent litigation demonstrates the risks inherent in overly restrictive interpretation of disclosure rights. In *Aithaqi,* the Court rejected the defendant's request for access to validation studies conducted by Quebec's forensic science laboratory, because the Court found that the defendant had not shown why the validation studies were relevant to the case. Even in traditional frameworks relating to access to records outside of the defendant's possession, the Supreme Court has cautioned that "it is important to recognize that the accused will be in a very poor position to call evidence given that he has never had access to the records [sought]": *R v. O'Connor*, [1995] 4 SCR 411 at para 19.