PREDICTIVE FORENSIC PSYCHIATRY AND CRIMINAL LAW: RISK ASSESSMENT WITH THE RIGHTS OF THE ACCUSED

Ms. Gauri Shukla, School of Law, Christ (Deemed to be) University, Pune Lavasa Campus, Lavasa, Pune, Maharashtra, India

Ms. Sarbani Deb, Assistant Professor, School of Law, Christ (Deemed to be) University, Pune Lavasa Campus, Lavasa, Pune, Maharashtra, India

ABSTRACT

The research aims to explore the evidentiary, constitutional and human-rights concerns of using algorithmic and actuarial-based risk assessment instruments in the criminal justice system, and, in particular, the applicability of the instruments in the Indian context. The paper lies at the crossroads between forensic psychiatry, criminology and human rights because it relies on the comparative models in the United States, the United Kingdom and the Netherlands. The Empirical evidence demonstrates that the most widely used tools, including COMPAS, are not opaque and did not offer significant predictive advantage and disproportionately high false positive rates among racial minorities, contradicting the due process guarantees, which are entrenched in the U.S., the contestability criteria, which are established in the jurisprudential dimension of Article 6 ECHR, and the fairness principle, which is stipulated in Article 21 of the Indian Constitution. By comparison, the multi-agency framework of MAPPA used in the U.K. shows how a structured and transparent management can quite effectively align and coordinate to the required human-rights standards, and the Dutch TBS system also offers a gold standard based on ECtHR jurisprudence. Therefore, in order to conform these models and fit them to the Indian context, India must conform to a hybrid regulatory framework which incorporates both transparency in MAPPA and judicial oversight in TBS in order to ensure that this model is within the scope and ambit of the proportionality framework provided in Puttaswamy. This method enhances the use of predictive tools in an operational and validated way that does not prejudice the notions of due process, liberty and equality.

Keywords: Risk assessment, Forensic Psychiatry, algorithmic bias, due process, Predictive Validity.

INTRODUCTION:

The use of risk-assessment tools, such as actuarial tools or sophisticated machine-learning algorithms, has gained prominence in criminal-justice decisions on bail, sentencing, parole, and preventive detention across systems all over the world. These devices are said to be efficient and enhance better safety of the community since they predict the likelihood of a person reoffending. However, they also undermine some main constitutional provisions: equal protection, due process, transparency, non-arbitrariness, proportionality, and the right to a fair hearing.

Literary evidence, such as diverse datasets, reports major issues: low predictive performance, high false-positive, racially/socio-economically unequal effects, and profound obscurity in the methodology of algorithmic creation. Different jurisdictions including the United States, the United Kingdom, and the Netherlands have assumed varied regulatory approaches, including lax invocation of proprietary tools (U.S. COMPAS) to multi-agency risk management structures (U.K. MAPPA) to strict, court-monitored therapeutic preventive detention systems (Netherlands TBS). The systems espouse varied ideas of risk, rights and state power.

In the meantime, India is at a crossroad concerning the constitution. It has one of the most robust substantive due-process and proportionality principles in the world which is expressed in *Maneka Gandhi*[1], *A.K Roy*[2], and *Puttaswamy*[3]. However, Indian preventive detention laws are still shrouded in secrecy, with lax adherence and little conformity to current evidentiary principles. The integration of the algorithmic tools stands a chance of reinforcing structural injustices already present without statutory and procedural reform, and would be a violation of both *Articles 14*[4] and *21*[5].

RESEARCH OBJECTIVES:

This study aims to investigate the possibility of constitutionally incorporating algorithmic and actuarial risk-assessment instruments into the criminal-justice system of India because Article 14 and 21 have high constitutional standards and the doctrine of proportionality as espoused in

[1] Maneka Gandhi v. Union of India (1978) 1 SCC 248.
[2] A.K. Roy v. Union of India (1982) 1 SCC 271.
[3] Justice K.S. Puttaswamy v. Union of India (2017) 10 SCC 1.
[4] INDIA CONST. art. 14.
[5] INDIA CONST. art. 21.

Maneka Gandhi and *Puttaswamy*. The research problem is between the interest of the state in efficient and predictive decision-making and the basic rights of individuals who might be subjected to opaque, unproven, and even discriminative technologies. Current comparative models, such as the U.S use of proprietary software like COMPAS, the MAPPA framework of the United Kingdom, and the judicially regulated TBS regime of the Netherlands, indicate very varied degrees of transparency, control, and protection of rights and suggests the need to develop a constitutionally well-founded regulatory structure suitable to India.

RESEARCH METHODOLOGY:

This research utilises both a doctrinal approach, examining constitutional standards of equality, due process, and proportionality; a comparative approach, comparing MAPPA, TBS, and U.S. algorithm-based practices to determine the best-practice protections; and an interdisciplinary approach, requiring innovative approaches to forensic psychiatry, criminology, machine-learning studies, and human-rights research to determine scientific validity, risks of bias, and procedural consequences to solve the subsequent issue previously discussed. This combined structure supports the creation of a rights-conformity, MAPPA-TBS hybrid framework based on transparency, judicial control, independent authentication, and non-discrimination as the preconditions of the legal use of predictive tools in India.

LITERATURE REVIEW:

Literature on algorithmic risk evaluation shows that there are recurrent issues of predictive validity, bias, and constitutional adequacy. The review starts with the historical developments of forensic risk assessment with its origins in the 1960s criticism of the insanity defence, Morris believed that the legal definition of dangerousness is not based on empirical evidence and that psychiatric resources are being mis-allocated to treatment more than prediction.[6] Initial studies noted the methodological poor quality of actuarial instruments (e.g., VRAG, HCR-20, LSI-R) with limited calibration, external validation and moderate AUCs, validating the call made by Fazel in regards to improved data and clear reporting.[7] Later studies revealed racial bias of popular tools like the PCRA demonstrating that Black and White offenders are equally predictable with different distributions of scores in test outcomes, which explains the necessity

[6] Norval Morris, Psychiatry and the Dangerous Criminal, 41(3) S. Cal. L. Rev. 516 (1968).
[7] Seena Fazel, "The Scientific Validity of Current Approaches to Violence and Criminal Risk Assessment" in Jan W de Keijser, Julian V Roberts and Jesper Ryberg (eds), "Predictive Sentencing: Normative and Empirical Perspectives" (Hart Publishing 2019) (197-212)

of not only testing bias but also calculating disparate impact.[8] Fairness scholarship (Berk et al.) demonstrated that various definitions of fairness are mutually exclusive, and practitioners are compelled to make accuracy versus equity and to accommodate base-rate disparities.[9] Similar progress in machine-learning demonstrated that interpretable models (RiskSLIM, EBM, Additive Stumps) are able to achieve the same performance as black-box performance but maintain fairness, but with difficulty in generalizability, temporal drift, and incorporating fairness constraint.[10] Neuroimaging research demonstrated incremental predictive validity of resting-state rCBF compared to conventional risk factors, but small sample sizes and non-replication make this technology not clinically applicable.[11] The evidence of experimental research indicates that models like COMPAS can only achieve a moderate level of accuracy and contain an overrepresentation of false-positives within marginalized populations, which begs the questions of equality and due process[12]. The study of forensic psychiatry, criminology and machine learning identifies structural lack of transparency, proxy discrimination, and the constraints of actuarial prediction, whilst research on neurolaw raises concerns about the ethical validity of predicting aversion to harm[13].Throughout the literature, there are chronic gaps, namely the lack of integration of protective and dynamic factors, the lack of external validation across jurisdictions, the presence of ethical tensions between the prediction of risks and liberty, and the necessity of interdisciplinary, longitudinal, designs that integrate clinical, criminological, and computational designs.[14] Based on doctrinal analysis, such evidentiary norms as *Daubert*[15] plus ECtHR jurisprudence on contestability and constitutional mandates of non-arbitrariness, substantive due process, and proportionality set forth in India place high restrictions on any liberty-influencing predictive instrument. Contemporary comparative research designates the MAPPA system of the U.K. and the TBS regime of the Netherlands as more rights-protective ideals because of the transparency, multi-agency regulation, and judicial checks and balances, as opposed to the construct of U.S. proprietary algorithms. Collectively,

[8] Jennifer L. Skeem and Christopher T. Lowenkamp, Risk, Race and Recidivism: Predictive Bias and Disparate Impact, 54(4) Criminology 680 (2016).
[9] Richard Berk, Hoda Heidari, Shahin Jabbari, Michael Keams and Aaron Roth,"Fairness in Criminal Justice Risk Assessments: The State Of the Art," University of Pennsylvania(May 30, 2017).
[10] C Wang, Bin Han, Bhrij Patel, et al., "In Pursuit of Interpretable, Fair and Accurate Machine Learning for Criminal Recidivism Prediction," 39 Journal of Quantitatuve Criminology 519(2023).
[11] Delfin C, Krona H, Andine P, Ryding E , Wallinius M, Hofvander B," Prediction of recidivism in a long-term follow-up of forensic psychiatric patients: Incremental effects of neuroimaging data", PLOS ONE, Vol. 14 , No. 5 (2019).
[12] ProPublica: Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine Bias.
[13] Danielle Keats Citron (2016). Technological Due Process.
[14] Richard Rogers,"The Uncritical Acceptance of Risk Assessment in Forensic Practice," 24 Law and Human Behaviour 595 (2000).
[15] Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).

the literature has determined that algorithmic tools can solely be used constitutionally in systems that promote transparency, independent testing, contestability, and constant judicial oversight.

1. MAJOR LEGAL DEBATES:

The use of algorithmic and actuarial risk-assessment tools by criminal justice systems brings about underlying legal arguments in the interface of constitutional rights, evidentiary reliability, proportionality, equality, and human-rights theory. In many jurisdictions, one of the recurring tensions faced by the courts is that, although predictive analytics are claimed to ensure efficiency in administrations and improved safety of citizens, they also threaten many long-established norms, such as the presumption of innocence, individualistic justice delivery, and fair procedures guaranteed by the Constitution. The normative issue at the centre is when and whether predictive tools, whose accuracy, transparency and methodological validity have been contentious, should play any normative role in decisions concerning liberty or punishment.

The first axis of debate has to do with predictive accuracy and its legality. Literature in the fields of forensic psychiatry, criminology and machine-learning research suggests that the average AUC values of criminal-justice risk tools are roughly 0.62 to 0.75, which is only slightly superior to chance and usually lower than structured professional judgment. The rates of false-positive are particularly alarming, particularly when it comes to such tools as COMPAS, when empirical tests are conducted, such as *ProPublica (2016)*[16]*,* which estimated that, in this case, Black defendants were much more commonly falsely predicted. These empirical facts directly bring constitutional doctrine into play: when risk tools predict threats on some populations in a systematic way, their application can be found guilty of breaching Equal Protection in the U.S. or *Article 14*[17] of the Indian Constitution or ECHR *Article 14*[18] in Europe. They also compromise proportionality, which means the deprivation of liberty cannot be based on exaggerated or unsubstantiated risk estimates in the Constitution.

The second dispute with significant implications is the admissibility of evidence in the form of an algorithmic prediction and its epistemic validity. Under *Daubert v. Merrell Dow (1993)*[19]

[16] *Supra Note 8 at pg. 3.*
[17] *Supra Note 4 at pg. 2.*
[18] *Ibid.*
[19] *Supra Note 10 at pg. 3.*

In the United States, federal courts need expert evidence to meet the criteria of testability, peer review, known error rate, and standards of application. Most proprietary risk systems, such as those of COMPAS, do not meet these requirements as a result of business secrets that do not allow the methodology and error promotions to be disclosed. *Kumho Tire (1999)*[20] goes even further to explain that Daubert is applied to technical and probabilistic devices thus opaque algorithmic models are presumptively inadmissible unless proven. In spite of the fact that there are still states who stick to *Frye*[21], even the general acceptance test as envisioned by Frye is hard to meet as the scientific community itself is split on the issue of the validity of actuarial prediction in sentencing.

In India, the Indian Evidence Act, 1872, and particularly *Section 45*[22], demands the expert testimony to be based on provable body of knowledge and open methodology. In *State of H.P. v. Jai Lal (1999)*[23], Indian courts have stated that expertise is achieved with "special study, experience, or observation", and that a proprietary, non-transparent algorithm cannot achieve such abilities. This predisposes risk scores especially: in the absence of training data, feature weights, or validation metrics, they are unable to pass any basic evidentiary test. Therefore, based on the U.S. and Indian approaches to law, scientific ambiguity and opacities of algorithmic technologies impose challenges on its use in freedom determination in a legal manner.

A third and more fundamental problem is that of the presumption of innocence and preventive governance. Risk measures are often applied to pre-trial bail, sentencing, parole and preventive custody- settings where future behavior forecasting directly influences the denial of liberty. *United States v. Salerno (1987)*[24] believes in preventive detention strictly subject to due-process protections, in which liberty is the rule and detention is the "carefully limited exception". The predictive tools that exaggerate risk or do not disclose risks transform this exemption into a normal administrative process. Critics like *Sandra Mayson*[25] suggest that algorithmic prediction reworks the assumption of innocence to become an assumption of risk, which does not fit into the normative underpinnings of the criminal law.

[20] Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999).
[21] Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).
[22] Indian Evidence Act § 45.
[23] State of H.P. v. Jai Lal, 7 SCC 280 (1999).
[24] United States v. Salerno, 481 U.S. 739 (1987).
[25] Sandra G. Mayson (2019). "Bias In, Bias Out," Yale Law Journal.

In India, as per *Article 21* and *22* of the Indian Constitution, specifically in the cases of *Maneka Gandhi v. Union of India (1978)* and *A.K. Roy v. Union of India (1982),* requires that deprivation of liberty must be based on fairness, non-arbitrariness, and reasonableness in the procedure. The implementation of a substantive proportionality test by the Supreme Court in *Puttaswamy (2017)* also demands that any limitation of liberty, in order to be enforced, should seek a legitimate end, be reasonably tailored to that end, be the minimum constraint, and be proportionate in its effects. These requirements cannot be the case with risk tools whose error rates are not known or the manner in which they operate is opaque. Specifically, when a predictive score is not explicable to an individual being affected, the right to be heard and the right to reasons are being interfered with (i.e., the right to be heard and the right to reasons are an element of due-process that have been recognized by Articles 14 and 21).

Another constitutional issue is algorithmic transparency. Proprietary risk equipment often cloaks their models with claims to intellectual property, and train their models confidential, such that neither the training data, the statistical weights or the error margins are disclosed. This is contrary to guarantees of due process in all major jurisdictions. U.S. law offers the right to contest supporting evidence and challenge its foundation through the principles of confrontation and fair-trial. *Article 6 ECHR*[26] asks disclosure to be adequate in the UK in order that meaningful participation is possible. Opacity is essentially in opposition to the fairness doctrine of *Maneka Gandhi* and the explainability criteria that arise in *Puttaswamy* in India. Whenever the algorithm involved is a black box the proportionality cannot sensibly be applied to it nor the arbitrariness be judgeable.

Last, is an extensive worry about human-rights restrictions on preventive detention and risk-based restrictions. Under the *ICCPR Article 9*[27], the deprivation of liberty must not be arbitrary and arbitrariness deals with both substantive and procedural flaws. Jurisprudence ECtHR- *M. v. Germany*[28] and *Van der Velden v. Netherlands*[29] – states that preventive detention must be premised on rigorous evidence, and it must be reviewed by a court on a regular basis. In situations where risk instruments depend on statistical relationships between groups of people, and not on individualized relationships, there is a risk of them being collective punishment or

[26] Convention for the Protection of Human Rights and Fundamental Freedoms art. 6, Nov. 4, 1950, 213 U.N.T.S. 221.
[27] International Covenant on Civil and Political Rights art. 9, Dec 16, 1966, 999 U.N.T.S. 171.
[28] M. v. Germany, App. No. 19359/04 (ECtHR, 2009).
[29] Van der Velden v. Netherlands, App. No. 29514/05 (ECtHR, 2006).

discriminatory, particularly in relation to historically over-policed groups. This emphasizes the fact that algorithmic tools tend to replicate the structural biases that occur in the training set, so their blind application will not be in line with international human-rights principles.

In such a way, the principal legal controversies all point towards one direction: even though predictive technologies are presently administratively useful, their scientific frailty, lack of transparency, and biases have significant constitutional challenges. The rule of law demands clear, justified, and personal evidence prior to the limitation of liberty. Devoid of such standards, no law can justify algorithmic prediction as a justification to detain, prosecute, or otherwise coerce the state.

2. CROSS-JURISDICTIONAL LEGAL PRINCIPLES:

In the modern criminal-justice systems, the control over the risk-assessment devices and risk-preventive restricts illustrates sharp disparities in the requirements in the statutory design, the requirements of transparency of the jury and the assurances of the judicial checks. Comparative legal analysis of the United States, the United Kingdom (MAPPA), the Netherlands (TBS), and India suggests that although the four jurisdictions face almost identical pressures: the demand of the people for the predictive control, to efficiency in the administration and the constitutional limitations on deprivation of liberty, they differ in the institutional reaction sharply. This deviation offers an abundant analysis of how India can develop a protection of rights regulatory strategy.

2.1. United States: Procedural Due Process and Reliability of Evidence.

Due Process Clauses of Fifth[30] *and Fourteenth*[31] *Amendments* are the constitutional basis in the U.S. The cornerstone to preventive detention is *United States v. Salarno (1987)*[32]*:* the Supreme Court approved the Bail Reform Act because it subjected it to severe procedural protection, made the government bear the burden of it and restricted its application to situations which were strictly circumscribed. *Salarno*[33], hence determines that the only way preventive deprivation of liberty is constitutional under the law is where it is backed by credible,

[30] U.S. Const. amend. V.
[31] U.S. Const. amend. XIV, § 1
[32] *Supra note 19, at pg. 5.*
[33] *Ibid.*

personalized evidence and through a sound judicial process.

These safeguards are however challenged by algorithmic tools like COMPAS. In *Loomis v. Wisconsin (2016)*[34], The Wisconsin Supreme Court allowed sentencing based on the COMPAS, but with conditions: the judges were not allowed to rely on the score as definitive, and they had to consider its disadvantages, such as the lack of transparency and group aspects. *Loomis*[35] reveals one of the active structural tensions: proprietary algorithms, the methodology of which is safeguarded by the trade-secret statutes, are not disclosed, which deprives the accused of the opportunity to challenge the evidence. Under *Daubert and Kumho Tire*[36], courts are expected to scrutinize expert evidence, measured in testability, peer review, notorious error rates, and standards, which most algorithm tools cannot satisfy, especially when the originators are uncooperative to disclosure.

Hence, although the U.S. system is procedurally endowed, it condones algorithmic secrecy compared to the European ones, where judicial warnings are a way of living instead of mandatory disclosure. This establishes an asymmetry in procedures: guilty parties experience the effects of the predictive technologies that they cannot test or scrutinize

2.2. United Kingdom: MAPPA in a Human-Rights Framework

The UK system of dealing with those who present serious risk known as MAPPA is regulated by the *Criminal Justice Act of 2003*[37] and explained by national MAPPA Guidance. The design of MAPPA is essentially administrative and multi-agency that encompasses police, probation, prisons and other bodies. In contrast to the U.S. practice, which focuses on evidentiary admissibility, MAPPA prioritizes the organization of risks, information sharing between agencies, and the review on a regular basis.

MAPPA engages with the *Human Rights Act 1998*[38] which includes the ECHR, such as Article 5 (liberty), 6 (fair trial) and 8 (privacy). The applicable ecotrhorous jurisprudence- e.g. in *Guzzardi v. Italy*[39] (control measures could be a deprivation of liberty depending on its force)

[34] State v. Loomis, 881 N.W.2d 749 (Wis. 2016).
[35] *Ibid*
[36] *Supra note 15, at pg 5.*
[37] Criminal Justice Act 2003 (UK), Part II (MAPPA).
[38] Human Rights Act 1998, c. 42 (U.K.)
[39] Guzzardi v. Italy, App. No. 7367/76 (ECtHR, 1980).

and *James, Wells and Lee v. UK*[40] (detention without meaningful rehabilitative review is contrary to Article 5) This means that any restriction must be necessary, proportionate and open to reconsideration.

In spite of the fact that MAPPA does not represent the algorithmic risk-assessment system per se, the regulatory framework in the program can be studied as the lesson of how the processes of technologically-assisted decision-making should be managed. MAPPA states that all supervisory decisions that will be made must be documented with a proper explanation, and there should be an audit trail that can withstand legal or administrative review. Its multi-agency protocols of information sharing introduce disclosure requirement within it, which brings up transparency and ensures that one authority does not exercise unfettered discretion. In addition, MAPPA requires periodic and systematic review of risk, and there are procedural protective measures that increase with a person ascending the levels of supervision. Collectively, these characteristics depict how the arbitrariness can be alleviated and the rights-compliance improved with the help of a sound regulatory frame, which is based on documentation, transparency, and regularity without using any algorithmic-based devices.

In comparison to the U.S., MAPPA is disclosed more clearly and more clearly related to human-rights proportionality, but explicitly still permits a broad discretion and restricted judicial intervention unless liberty-restricting actions are interrogated.

2.3. Netherlands: TBS and European Gold Standard of Judicial Oversight

The Dutch Terbeschikkingstelling (TBS) regime, which was created based on *Articles 37a-38g of the Dutch Criminal Code*[41], is commonly acclaimed as among the most regulated and rights-respectful risk-based detention regimes in modern criminal trials. Importantly, TBS can be ordered by no other body than a court and then only in two conditions, both of which are strict statutory conditions: the commission of a serious offence involving a clinically-determined mental disorder, and the existence of a proven serious danger to society, which is evident. The difference is that TBS has a very institutionalised judicial oversight and proportionality review. The maximum length of the first term of detention is rigid--usually a year or two and more prolongation of such a term is a subject of new judicial approval, in which case there must be new psychiatric evaluations and disorganised risk-analysis notes. Those under TBS do not lose

[40] James, Wells & Lee v. United Kingdom, App. No. 25119/09 (ECtHR, 2012).
[41] Wetboek van Strafrecht arts. 37a – 38g(Neth.)

the valuable procedural rights such as the right to attend hearings, attack expert conclusions and seek outside judgments so that the deprivation of liberty is individualised and on-going justified. Such a trans--layered probationary system of judicial oversight and evidentiary review has rendered TBS a paradigmatic instance of the rights-compatible preventive detention.

In ECtHR cases such as *Van der velden v. Netherlands (2006)*[42] and *M. v. Germany (2009)*[43], preventive detention should not be punitive, as well as it should not lose contact with therapeutic objectives and must be supported with the current threat independently evaluated. The Dutch courts are often known to question expert reports about their methodological soundness- much exaggerated than the U.S. courts when looking at algorithmic scores. It has led to a system where transparency, accountability and judicial control cannot be mangled out of the legitimacy of risk-based measures. Among the analyzed systems, TBS provides the best architecture of judicial safeguards.

India: The Review of the Constitution in the Light of the Weakness of the Statutes

Learning India has both wide statutory power of preventive detention combined with robust constitutional judicial review on fairness and proportion ability. *Articles 14 and 21* stipulate that detention of liberty must go through the non-arbitrariness, fairness, and reasonability tests. The due process was constitutionalised by *Maneka Gandhi (1978)* and preventive detention scrutinized by *A.K. Roy (1982)*. *Puttaswamy (2017)* proposed a formal proportionality test that entailed: A legitimate aim, a Rational connection, Necessity/least restrictive means and balancing of rights and state interests. Nevertheless, statutory preventive-detention regime (e.g., NSA, PSA) has several weak procedural rights, little disclosure, and little judicial review, which are incompatible with algorithmic decision-making, which requires an even more significant level of transparency due to the inherent opaque nature and probabilistic logic of the process.

The constitutional doctrine of India is much more demanding than its statutory practice: on the event that algorithmic tools were brought in without changing the statute, they would tend to breach *Article 14* and *Article 21* on the basis of lack of transparency, lack of explanation, and

[42] *Supra note 31, at pg. 7.*
[43] *Supra note 30, at pg. 7.*

lack of ability of the detainees to appeal to the methodology.

2.5 Comparative Doctrinal Synthesis

Four dimensions of law, identified by a cross-jurisdictional comparison demonstrate four critical legal principles:

2.5.1 The Constitutional Prerequisite of Transparency

Transparency would be a prerequisite in the legal application of risk-assessment tools across jurisdictions. Both the United Kingdom and the Netherlands have regulatory frameworks to ensure that the evidentiary basis, decision-making documentation, and periodic review is disclosed, which allows the courts and the persons impacted to question the rationale of the liberty-restricting measures. In comparison, the United States permits the usage of proprietary algorithms, including COMPAS, the logic of which is not subject to trade-secret safeguard, and defendants have no chance to either analyze or contest the technology behind unfavorable categorizations. The constitutional jurisprudence of India - especially under *Article 14 and Article 21* - stretches further as it states that any lack of transparency in state action characterized by deprivation of liberty is in se arbitrary and as such unconstitutional. Devoid of procedural value, transparency in this landscape is a substantive constitutional principle that lies at the core of fairness, equality and accountability.

2.5.2 Judicial Oversight as the Determinant of Legitimacy

Judicial review forms the main protection that prevents risk based or preventative decisions that are not constitutional. The Dutch TBS regime is the strictest variant: each original imposition and renewal of detention must be accompanied by new judicial decisions with new expert evidence and open to adversarial challenge. MAPPA is administrative but incorporates the multi-agency review, and documented rationale and it presents a tiered mechanism of oversight, which is not judicial but instills institutional checks on discretion. This is a hybrid case: on the one hand, the constitutional doctrine requires extreme scrutiny of the actions limiting the freedoms of citizens; on the other hand, statutory preventive-detention schemes offer limited judicial participation, which creates a structural contradiction between principle and practice. In the meantime, the

United States is largely a country of adversarial hearings but limited substantive judicial questioning of the algorithms themselves. It has a comparative direction: the closer and more fundamental the judicial supervision, the greater the constitutional legitimacy of the risk-governance system.

2.5.3 Universality Proportionality and Necessity as Universal Constitutional Tests

An analysis of the ECHR jurisprudence, ICCPR regulations, and the Indian constitution law illustrates the international alignment of reasons of proportionality and necessity as the tenets that should be applied to any law restraining liberty. In the European Court of Human Rights, preventive detention or other measures of supervision must demonstrate a legitimate purpose, reasonableness, and as weakening, and it must be reasonable in view of a changing risk (e.g. James, Wells and Lee; M.v. Germany). Likewise, *ICCPR Article 9*[44] understands arbitrariness to contain a procedural and substantive disproportionality. Puttaswamy in India establishes the proportionality as a four-step examination that binds all actions of state including an algorithmically informed decision-making. Collectively, these doctrines affirm that risk-based measures cannot be merely justified using the predictive efficiency claim, they should be the least restrictive of rights, based on individually tailored evidence and can withstand judicial test proportionality.

2.5.4 Contestability and Right to Challenging Evidence

Lastly, a cross-jurisdictional study highlights that procedural fairness is focus on the right to question unfavorable evidence. In those systems where opaque algorithmic apparatus may be used- such as with COMPAS in the United States- those individuals systematically cannot access, challenge, or even refute the logic behind their risk determination. This refusal of contests contravenes the adversarial system in the United States court process, the guarantee of fair hearing enshrinement of concerning India and its constitutional duty that all decisions that restrict liberty are justified, subject to review, and challenge under *Article 14* and *Article 21*. This issue is exactly what jurisdictions like the Netherlands and the United Kingdom are striving to prevent, as their systems are crafted in such a way as to maintain the ability of the evidentiary

[44] ICCPR art. 9.

foundation of risk assessments, be it clinical, actuarial or multi-agency, to be open, open to review, and rebuting. This issue leads to the appearance of contestability as a universal procedural protection without which risk-based governance cannot qualify under constitutional or human-rights considerations.

3 COMPARATIVE RISK-ASSESSMENT INSTRUMENTS: COMPAS, MAPPA, TBS, INDIA:

The criminal justice systems of modern society are increasingly dependent on the use of risk-assessment tools to inform their decisions on the setting of bail, sentencing, parole, and even preventive detention. However, the design, openness, precision and regulatory framework of these tools differ dramatically in jurisdiction. The comparative study of COMPAS (U.S.), MAPPA (U.K.) and TBS (Netherlands) and the new Indian environment indicate a strong disparity in scientific validity, legal protection, and human-rights adherence.

COMPAS (United States): Accuracy Arguments, Opacity and Equal-Protection Arguments

COMPAS is a proprietary sentencing facility and parole algorithmic risk-assessment instrument prevalent in state courts throughout the United States. Its methodological obscurity is well-known. In *State v. Loomis (2016)*[45], The Wisconsin Supreme Court based the ruling on the fact that defendants had no rights to access the COMPAS algorithm due to trade-secret regulations, and this fact posed risks to the right to dispute unfair evidence. The court allowed the use of COMPAS only with warnings of its restriction, such as that it operated based on group-based statistic correlations.

As an empirical model, COMPAS exhibits average predictive performance. Independent reviews indicate that the most common AUC scores are approximately 0.65-0.70 which is slightly higher than chance[46]. According to the *ProPublica*[47] investigation, false-positive rates were significantly higher in the cases of the black defendants as they had more chances to be misclassified as high-risk despite the lack of reoffending. This poses an Equal Protection

[45] State v. Loomis, 881 N.W.2d 749 (Wis. 2016).
[46] *Supra Note 8 at pg. 3.*
[47] *Ibid.*

concern under the *Fourteenth Amendment*[48] which may result in unjustifiable differences in outcome based on race as well.

Admissible expert tools by Daubert should be capable of giving evidence of testability, peer-reviewed, and known error rates[49]. COMPAS does not meet most of these criteria due to the fact that the model, the training data, and weighting of the model are not disclosed. This contradicts the due-process in the *United States v. Salerno*[50],Individualized and reliable evidence to warrant detention is needed [Salerno 1987]. COMPAS can therefore be viewed as the least transparent model in the surveyed systems and how the algorithmic transparency is exploited to compromise contestability, accuracy assessment, and constitutional review.

MAPPA (United Kingdom): Administrative, Multi-Agency Risk Management

In contrast to COMPAS, MAPPA is not a single algorithmic tool but a statutory multi-agency framework created by the *Criminal Justice Act 2003*[51] in order to handle sexual and violent offenders. It incorporates professional judgment and systematic risk tools, and inter-agency communication, as opposed to being based on a proprietary algorithm. MAPPA cases concern police, probation, prisons, and local communities, who mutually determine the degree of risk and subject people to the control levels of Tier 1-3 [*MAPPA Guidance 2022]*[52].

MAPPA has its legitimacy because the transparency requirements are mandatory, and records are meticulously kept; it also requires the competent authorities involved to explain decision rationale. This establishes an audit trail of accountability in compliance with Article 5 and Article 8 ECHR that necessitates necessity and proportionality in liberty limiting or privacy invasive actions [HRA 1998; ECHR Art. 5][53]. The ECtHR in *James, Wells & Lee v. UK*[54] stressed that prolonged detention of suspects who are not actively reviewed is in contravention of Article 5; the annual (and occasionally quarterly) reviews MAPPA undertakes are direct responses to such concerns.

[48] U.S. Const. amend. XIV, § 1.
[49] *Supra Note 10 at pg. 3.*
[50] Salerno, *supra note 19, at pg. 5.*
[51] Criminal Justice Act 2003 (UK).
[52] MAPPA Guidance (Ministry of Justice, UK).
[53] Human Rights Act 1998
[54] *Supra note 42, at pg. 9.*

Although MAPPA delivers with the means of structured tools (e.g., OASys, RM2000), they remain transparent and testable and are never decided by an algorithm that cannot be called proprietary. The right-based structure of the UK guarantees that the decisions made in MAPPA can be challenged and the person has the right to question the misplaced decisions. MAPPA demonstrates greater procedural fairness than COMPAS but detractors observe that there can be a discrepancy in application between regions.

TBS (Netherlands): Preventive Detention under judicial Control and with High Evidential Standards

The system of preventive-detention in Europe most strict and rigorous is the Dutch TBS (Terbeschikkingstelling) system. Under Articles *37a-38g of the Dutch Criminal Code*[55], TBS is both the mandate of the court and of a clinically determined mental disorder, must serve the two-fold objects of social security and rehabilitation [*TBS Code Art. 37a*][56]. Notably, risk-evaluation instruments of the TBS (such as HKT-30, STATIC-99) are clear, peer-reviewed, and regularly debated by the courts.

The structural safety of the TBS system is a judicial control whereby the deprivation of any way of liberty is heavily justified and closely followed. First TBS incarcerations are set to a short term, usually a year or two, based on the principle that preventive imprisonment should be time bound and open to renewal. The automatic or administrative lengthening of the detention is not possible as any extension in the relation to the original term needs new expert appraisals and new judicial permission. Victims of TBS also enjoy the strong participatory rights where they have a right of access to independent psychiatric assessments as well as frequent hearings where they can challenge the evidentiary foundations of their further imprisonment. More importantly, TBS detention has to always be maintained not by administrative whim but based on evidence-based judicially challenged conclusions of present, not past, dangerousness. This stratified procedural structure renders TBS among the most rights-tilting preventive regimes of detention in the world.

This is a structure that is in line with ECtHR jurisprudence. In *Van der Velden v. Netherlands*[57], The Court emphasized the necessity of the regular control on the basis of revised psychological

[55] Wetboek van Strafrecht arts. 37a-38g, *supra note 43*
[56] Wetboek van Strafrecht arts. 37a (Neth.)
[57] *Supra note 31, at pg. 7.*

testing [Van der Velden 2006]. In *M. v. Germany*[58], the ECtHR pointed to the fact that preventive detention should be therapeutic, as opposed to punitive, and should have an evidence-based case of current or present, other than past risk [M. v. Germany 2009]. TBS is adherent to these principles by incorporating the aspect of risk-assessment in a court-controlled system of therapeutic approach.

Amongst the systems reviewed, TBS offers the highest level of transparency, scientific rigor and legal control - which is why it should be the gold standard of rights-compatible risk governance.

India: Promising, but No Tested or Marked Approach to Risk-Assessment

India has yet to have a national, standardized risk-assessment tool used in sentencing, bail or preventive detention. Preventive detention instead is subject to the assessments of statutes like the *National Security Act (NSA)*[59] and the State Public Safety Acts. Such laws allow administrative detention that has no strong disclosure requirements and little judicial oversight - something that cannot be used with algorithms.

There is constitutional doctrine, however, which is strong. *Maneka Gandhi* demands that the state action must be non-arbitrary and fair [*Maneka Gandhi 1978*]. *A.K. Roy* requires a high level of scrutiny of preventive detention [*A.K. Roy 1982*]. To determine the test of privacy and liberty limits, *Puttaswamy* defines proportionality [*Puttaswamy 2017*]. By these constitutional norms, any risk-assessment instrument being used in the Indian criminal-justice system has to meet a group of rigorous, rights-preserving conditions. The tool needs to be transparent in the methodology it is based on (variables, weighting, and data) first so that it can be competitively challenged by those who are subject to it, and their counsel - a key element of procedural fairness as required by *Article 14* and *Article 21*. Second, it should be shown to be scientifically valid with the reliability levels of the Daubert model by validating its ability to predict using independent studies, the known error rates and reporting of the predictive performance. Third, the tool should not have discriminatory effect, in line with the Article 14 that precludes arbitrariness and indirect discrimination: anything that duplicates structural discrimination in law enforcement or social-economic statistics cannot be explained by the Constitution. Lastly, the system should produce justifiable, verifiable outputs which enable any court of law to

[58] *Supra note 30, at pg. 7.*
[59] National Security Act, No. 65 of 1980 (India)

examine the foundations of any liberty-limiting decision and therefore meets the *Article 21* requirement that state action that impacts personal liberty should be just, fair and reasonable. All of these requirements create a constitutional minimum that any algorithmic or actuarial risk tool has to fulfill to be able to lawfully impact decision-making in India.

Comparative Synthesis

System	**Transparency**	**Oversight**	**Scientific validity**	**Rights compatibility**
COMPAS (U.S.)	Opaque, proprietary	Judicial- However it can't inspect algorithm	Moderate AUC; Bias concerns	Week due to opacity
MAPPA (UK)	High administrative transparency	Multi agency and HRA constraints	Structured tools, documented	Moderate-Strong
TBS (NL)	Very high	Judicial, periodic, expert-driven	Transparent tools, strong review	Highest
INDIA	Low	Limited judicial role	No validated tools	Weak unless reformed

We can find a clear gradient in the comparison: COMPAS is the worst in terms of problem, MAPPA has more structural transparency than an algorithm, TBS provides the most robust protection, and India has a solid constitutional doctrine with weak statutory one.

4 RECOMMENDATIONS BASED ON GLOBAL RISK ASSESSMENT TOOLS:

A comparative, legal and rights-based analysis of risk-assessment devices in the United States, United Kingdom, Netherlands, and India demonstrates one basic fact: predictive instruments cannot be legitimate unless integrated in an open, responsible and constitutionally checked

system. The above-mentioned sections proved that algorithmic instruments, in particular, opaque, proprietary models like COMPAS, present a serious threat of discrimination [Angwin et al. 2016], breach due-process rights, due to their opacities [Loomis 2016], and fall short of the stated evidentiary reliability requirements, such as those established in Daubert [Daubert 1993]. Simultaneously, the existence of comparative systems, including, but not limited to, MAPPA (UK) and TBS (Netherlands), indicates that risk-assessment can be effectively exploited to protect the populace and not undermine the core principles of remaining rights, so long as the law is robust and well-enforced.

The American Trap of Opaque Predictive Tools:

The American case demonstrates how dangerous it can be to make use of the algorithms that cannot be transparent, which are not always the same. Courts do not forbid proprietary risk instruments where defendants cannot check or object to underlying reasoning, as occurs in *State v. Loomis* [Loomis 2016]. This is contradictory of the foundations that were made in the *United States v. Salerno*, that demand that every preventive detention should be based on personalized, good, and challengeable evidence [Salerno 1987]. Empirical studies indicate that COMPAS creates racial bias and false positives [Angwin et al. 2016], ruining efforts that it is a neutral tool and according to the value of Equal Protection. In the case of India, the U.S. example is most cautionary: using unverified or commercially owned tools would blatantly be in violation of *Article 14 and Article 21*.

The English Multi-Agency Transparency and Proportionality Principles:

MAPPA framework exhibits a rights-adhering manner that is not based on opaque algorithmic instruments. Accountability provided under direction of multi-agency decision-making, full documentation and reviewing are protective measures in MAPPA in accordance with the requirements of *ECHR Art. 5 and 8*[60] of proportionality and transparency in liberty-restrictive interventions [ECHR Art. 5; MAPPA Guidance 2022]. Acts of ECtHR, such as those of James, Wells, and Lee. UK, emphasizes that arrest, or supervision by restraint or detention, should always be justified [James, Wells and Lee 2012]. Article 14 and 21 include fairness requirements that the MAPPA system offers: the reasons are recorded, the reviews are done at

[60] Convention for the Protection of Human Rights and Fundamental Freedoms art. 5 & 8, Nov. 4, 1950, 213 U.N.T.S. 221.

multiple levels, and disclosure is mandatory, they help with fairness provisions [Maneka Gandhi 1978].

The European Gold Standard of Evidence-Based, Judicially Supervised Risk Management Stemming from Netherlands:

The Dutch TBS is the most instructive regime to use in India. TBS incorporate risk evaluation into a therapeutic structure of judicial exertion that is handled under *Article 37a-38g of the Dutch Criminal Code*[61] [TBS Code Art. 37a]. The expert reports are transparent, peer reviewed and subject to scrutiny to the courts at each and every step of the detention. ECtHR opinions - those of *Van der Velden v. Netherlands*[62] *and M. v. Germany*[63]--uphold the need to have a regular review, personalized evidence, and the non-punitive character of the preventive detention [Van der Velden 2006; M. v. Germany 2009]. Of all the systems involved in the comparison, TBS is the most compliant with international human-rights standards in *ICCPR Article 9*[64] *and ECHR Article 5*. Its judicial strictness mirrors the constitutional demands required by India within Puttaswamy of necessity, minimal impairment, and proportionality [Puttaswamy 2017].

The Indian Protective Rights: From Law to Practice the Transformation Needed

India has one of the best protective rights on paper. *Maneka Gandhi* believed that state intervention which prevents liberty must be just and fair and reasonable, which is a substantive due process [Maneka Gandhi 1978]. The use of preventive detention is subject to close examination by *A.K. Roy* [A.K. Roy 1982], whereas a systematic system of proportionality is established by *Puttaswamy* and must be met by any technological or administrative intervention [Puttaswamy 2017]. Furthermore, Article 14 prevents arbitrariness in any way, such as the algorithmic decision-making process, which generates discriminative outcomes.

Nevertheless, the present-day preventive-detention laws in India include NSA and state PSAs that preventive-detainee laws act with low levels of transparency, superficial review procedures, and low evidentiary thresholds. Such gaps in the statutes make India unprepared to implement algorithmic tools without a radical legal and structural change. In the absence of

[61] Wetboek van Strafrecht arts. 37a-38g, *supra note 43*
[62] *Supra note 31 at pg. 7.*
[63] *Supra note 30 at pg. 7.*
[64] S*upra note 29 at pg. 7.*

reform, the incorporation of ML-based predictions is likely to increase arbitrariness and deliver results that are not compatible with *Article 14, 21 and 22.*

CONCLUSION:

The lessons of both comparative jurisdictions and Indian constitutional doctrine indicate a clear avenue, though: opaque, proprietary and unvalidated instruments to determinations in the form of COMPAS must not be used in India, as it is likely to breach constitutional principles of equality, due process and reasoned decision making [Maneka Gandhi 1978; Loomis 2016; Angwin et al. 2016].

As an alternative to incorporating opaque or proprietary algorithmic tools, India ought to come up with a hybrid regulatory framework that integrates the best practices of its comparative models with its constitution-specific needs. Among others, the strengths that MAPPA brings to India include multi-agency review of risks, decision mandatory documentation, periodic re-evaluation, and strong transparency requirement, all contributing to accountability and ensuring that a state cannot act unilaterally or without examination [MAPPA Guidance 2022]. The Dutch system of TBS regime should be adopted in India with judicially based protective measures permanently entrenched, such as a restriction ordered by a court and controlled by a court, a strict judicial review, with a judicial review grounded on new evidence, transparent and expert reviews, and individual and therapeutic-purpose conditions imperative to detention being strictly circumscribed and constantly justified [TBS Code Art. 38; Van der Velden 2006]. Lastly, this approach should be based on the constitutional promises of India: the taboo of arbitrariness and discrimination as it is in *Article 14*, the imperative of procedural fairness and substantive due process as in *Articles 21*, and the structural test of proportionality and necessity of *Puttaswamy*[65] and *Maneka Gandhi*[66]. These collectively would form a transparent, scientifically viable, judicially presided over and completely aligned system which is highly consistent with the constitutional ethos in India.

By using a hybrid model, it is necessarily the case that India can gain the benefits of structured and evidence-based risk management without damaging constitutional values. In contrast to COMPAS, a hybrid system of MAPPA-TBS would be based on the use of open methodologies, judiciary, frequent reviews, and external validation. This would be consistent with the

[65] *Supra note 3, at pg 2.*
[66] *Supra note 1, at pg 2.*

international human-rights requirements in the ICCPR and ECHR, and it would balance with the Indian constitutional construction.

A rights-based approach draws a number of conclusive points on the constitutional regulation of risk-assessment tools in India. To begin with, opaque or proprietary instruments like COMPAS should be absolutely banned in any instance concerning individual freedom because the inherent element of the methodologies being self-restricted without being reviewed automatically breaches the provisions of equality and fairness in *Article 14* and *Article 21*. Second, the constitutional demands of procedural fairness, reasoned decision making and proportionality of India can be met using only transparent, empirically accurate, and entirely contestable tools. Three, India must pursue a MAPPA-TBS hybrid model, which is based on statutory accuracy, robust judicial scrutiny, periodic re-examination, verifiable scientific reliability and safeguarding accountability and non-discrimination by the populace. Lastly, a rights respecting system of the kind, would not just improve the safety of people; it would impose anew the constitutional morality of the individual man, his dignity, and the initial undertakings of the rule of law upon which any state action, in a democratic state, should be based. Essentially, constitutional structure of India simply does not allow a risk-governance human rights oriented- it mandates it.