

Est.  
1841

YORK  
ST JOHN  
UNIVERSITY

Olawade, David B., Ayoola, Folashayo I., Ebo, Titus Oloruntoba, Asaolu, Akinsola J., Egbon, Eghosasere and David-Olawade, Aanuoluwapo Clement (2025) Artificial Intelligence in Forensic Mental Health: A Review of Applications and Implications. *Journal of Forensic and Legal Medicine*, 113. p. 102895.

Downloaded from: <https://ray.yorks.ac.uk/id/eprint/12074/>

The version presented here may differ from the published version or version of record. If you intend to cite from the work you are advised to consult the publisher's version:

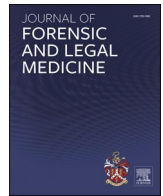
<https://doi.org/10.1016/j.jflm.2025.102895>

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. [Institutional Repositories Policy Statement](#)

# RaY

Research at the University of York St John

For more information please contact RaY at  
[ray@yorks.ac.uk](mailto:ray@yorks.ac.uk)



## Review

# Artificial intelligence in forensic mental health: A review of applications and implications

David B. Olawade<sup>a,b,c,d,\*</sup>, Folasayo I. Ayoola<sup>e</sup>, Titus Oloruntoba Ebo<sup>f</sup>,  
Akinsola J. Asaolu<sup>g</sup>, Eghosasere Egbon<sup>h</sup>, Aanuoluwapo Clement David-Olawade<sup>i</sup>

<sup>a</sup> Department of Allied and Public Health, School of Health, Sport and Bioscience, University of East London, London, United Kingdom

<sup>b</sup> Department of Research and Innovation, Medway NHS Foundation Trust, Gillingham, ME7 5NY, United Kingdom

<sup>c</sup> Department of Public Health, York St John University, London, United Kingdom

<sup>d</sup> School of Health and Care Management, Arden University, Arden House, Middlemarch Park, Coventry, CV3 4FJ, United Kingdom

<sup>e</sup> Department of Health Policy and Management, Faculty of Public Health, College of Medicine, University of Ibadan, Ibadan, Nigeria

<sup>f</sup> Forensic Mental Health Unit, Nottinghamshire Healthcare NHS Foundation Trust, United Kingdom

<sup>g</sup> Center for Research and Development, Kingsway Hospital, Derbyshire Healthcare NHS Foundation Trust, Derby, United Kingdom

<sup>h</sup> Department of Tissue Engineering and Regenerative Medicine, Faculty of Life Science Engineering, FH Technikum, Vienna, Austria

<sup>i</sup> Endoscopy Unit, Glenfield Hospital, University Hospitals of Leicester, NHS Trust, Leicester, United Kingdom

## ARTICLE INFO

## Keywords:

Artificial intelligence  
Forensic mental health  
Risk assessment  
Algorithmic fairness  
Ethical considerations

## ABSTRACT

This narrative review explores the transformative role of artificial intelligence (AI) in forensic mental health, focusing on its applications, benefits, limitations, and ethical considerations. AI's capabilities, particularly in areas such as risk assessment, mental health screening, behavioral analysis, and treatment recommendations, present promising advancements for accuracy, efficiency, and objectivity in forensic evaluations. Predictive models and natural language processing enhance the precision of high-stakes assessments, enabling early intervention and optimized resource allocation. However, AI's integration in forensic mental health also brings significant challenges, particularly regarding data quality, algorithmic bias, transparency, and legal accountability. Limited access to high-quality, representative data can hinder reliability, while biases within AI models risk perpetuating existing disparities. Ethical concerns surrounding data privacy and the "black box" nature of many AI algorithms underscore the need for transparency and accountability. The review highlights future directions for responsible AI use, including improving data standards, fostering interdisciplinary collaboration, and establishing robust regulatory frameworks to safeguard ethical and fair AI applications in forensic settings. Balancing technological innovation with ethical considerations and legal obligations is essential to ensure AI supports justice and upholds public trust. This review calls for ongoing research, policy development, and cautious implementation to harness AI's potential while protecting individuals' rights within the justice system.

## 1. Introduction

Forensic mental health is a specialized field within mental health services that intersects with the criminal justice system, focusing on individuals who have encountered legal issues and whose mental health status plays a critical role in their legal proceedings.<sup>1,2</sup>

This domain encompasses the assessment, treatment, and management of individuals with mental disorders who are also involved in criminal activities or at risk of offending. Professionals within forensic mental health engage in complex evaluations, such as risk assessments for violent behavior, competency evaluations for trial proceedings, and

post-conviction treatment plans aimed at reducing recidivism.<sup>3</sup> These evaluations and interventions hold significant weight in the legal context, often influencing judicial decisions regarding sentencing, rehabilitation, and community reintegration.<sup>4</sup> Given the severe implications of these assessments for individuals and society, there is a continuous need for accuracy, efficiency, and reliability within forensic mental health.<sup>5</sup>

The advent of artificial intelligence (AI) has transformed numerous aspects of healthcare, providing innovative tools for diagnosis, treatment planning, and predictive analytics.<sup>6</sup> AI in healthcare leverages algorithms and machine learning models that process vast amounts of

\* Corresponding author. Department of Allied and Public Health, School of Health, Sport and Bioscience, University of East London, London, United Kingdom.  
E-mail address: [d.olawade@uel.ac.uk](mailto:d.olawade@uel.ac.uk) (D.B. Olawade).

<https://doi.org/10.1016/j.jflm.2025.102895>

Received 6 November 2024; Received in revised form 16 January 2025; Accepted 19 May 2025

Available online 24 May 2025

1752-928X/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

data to identify patterns, make predictions, and generate recommendations with high precision.<sup>7</sup> In mental health, AI applications have advanced significantly, with tools now capable of early diagnosis, suicide risk prediction, and customized treatment pathways based on individual patient data.<sup>8</sup> These advancements are especially promising in cases where human biases or resource limitations might otherwise impede optimal care. For instance, AI-driven platforms utilizing natural language processing can analyze speech patterns to detect mental health issues, while machine learning models can assess vast clinical datasets to forecast the likelihood of psychiatric episodes.<sup>9</sup> As AI technologies evolve, their integration within mental health has created new possibilities for addressing patient needs more effectively, increasing the accessibility and accuracy of care in a domain traditionally marked by diagnostic complexities and subjective assessments.<sup>10</sup>

Within the niche field of forensic mental health, the potential of AI applications is particularly compelling. The ability to automate portions of risk assessment, offender profiling, and behavioral analysis could greatly enhance the efficiency and reliability of forensic evaluations, which are often resource-intensive and prone to human bias.<sup>11</sup> However, the integration of AI into forensic mental health raises unique considerations. Unlike general mental health care, forensic mental health applications must navigate legal, ethical, and human rights concerns due to the often-coercive nature of forensic interventions and the high stakes involved in their outcomes. Furthermore, forensic settings require particular sensitivity to biases, as any misjudgments in risk assessments or mental health evaluations could disproportionately affect marginalized communities, possibly leading to unjust treatment within the criminal justice system.<sup>12</sup> The complex ethical landscape, alongside technical challenges like data availability and algorithmic fairness, underpins the importance of a nuanced examination of AI's role in forensic mental health.<sup>13</sup>

This review aims to provide a comprehensive overview of the application of AI within forensic mental health, focusing on existing tools, their benefits, limitations, and ethical considerations. By analyzing current AI-driven methodologies in risk assessment, diagnostic support, and treatment recommendations, this review highlights both the advancements AI brings to forensic mental health and the ethical and practical challenges that arise. These challenges include concerns about data privacy, algorithmic bias, transparency, and the risk of reinforcing existing systemic biases within the justice system. Additionally, this review explores the regulatory landscape and call for interdisciplinary approaches that involve professionals in AI, law, and mental health to guide responsible AI integration in forensic settings. The ultimate goal of this review is to inform practitioners, policymakers, and technologists of the promise and pitfalls of AI in forensic mental health, encouraging ethical practices and ongoing research in this evolving intersection between technology, healthcare, and justice.

## 2. Needs and challenges in forensic mental health

The forensic mental health field is inherently complex, existing at the intersection of healthcare, criminal justice, and legal ethics. Professionals within this domain deal with a diverse range of challenges that are compounded by the dual need for accurate mental health evaluations and reliable risk assessments to ensure both individual rights and public safety.<sup>14</sup> AI has emerged as a promising tool to assist forensic mental health professionals, yet several unique needs and challenges, including risk assessment, criminal profiling, mental state evaluations, treatment compliance, and data constraints, complicate its implementation. Forensic mental health operates under unique constraints that differentiate it from other mental health domains. The primary areas of focus—risk assessment, criminal profiling, mental state evaluations, and treatment compliance—each come with challenges that AI could help address but also raise concerns about accuracy, reliability, and ethical ramifications.<sup>15</sup>

### 2.1. Risk assessment

Risk assessment is critical in forensic mental health as it helps determine the potential for future violent behavior, recidivism, and self-harm among individuals with mental health issues.<sup>16</sup> These assessments are essential for judicial decisions related to sentencing, parole, and treatment. Traditionally, risk assessments rely on standardized tools and clinical judgment, both of which are prone to subjectivity and variability.<sup>17</sup> AI offers potential here by providing data-driven risk models that can assess risk levels based on large datasets, patterns, and historical data, potentially leading to more accurate predictions.<sup>18</sup> Studies show that machine learning models, such as logistic regression and neural networks, have shown higher predictive validity than traditional methods in violence risk assessment, suggesting their ability to improve accuracy in high-stakes environments. For example, a study showed that a wide variability in the experimental settings and characteristics of the enrolled samples emerged across studies, which probably represented the major cause for the absence of shared common predictors of violence found by the models learned.<sup>19</sup> Nonetheless, a general trend toward a better performance of ML models compared to structured violence risk assessment instruments in predicting the risk of violent episodes emerged with three out of eight studies with an AUC above 0.80. However, the implementation of these AI tools in forensic settings is still met with concerns about over-reliance on algorithmic predictions, which could reduce the nuanced and case-specific judgment required in forensic cases.<sup>20</sup>

### 2.2. Criminal profiling

Criminal profiling is another area where AI is believed to have transformative potential. Profiling involves analyzing behaviors, psychological patterns, and other individual traits to assist in criminal investigations and understand the psychological underpinnings of criminal behavior.<sup>21</sup> AI, particularly natural language processing (NLP), has been leveraged to examine vast amounts of unstructured data, such as text from criminal records, case files, and behavioral reports, to detect patterns and assist in profiling.<sup>22</sup> A recent study demonstrates the effectiveness of AI in creating behavioral profiles by analyzing linguistic markers and personality indicators, showing promise in identifying underlying psychological traits of offenders.<sup>23</sup> Yet, the use of AI in profiling introduces risks of reinforcing stereotypes or biases that may exist in the training data. Moreover, over-reliance on AI for criminal profiling can lead to ethical challenges, as AI-generated profiles might skew judgment and lead to confirmation bias in investigations, further complicating the already complex role of profiling in forensic mental health.<sup>24</sup>

### 2.3. Mental state evaluations

Mental state evaluations in forensic contexts are used to determine an individual's mental capacity and competence, which is essential for legal proceedings, particularly in cases involving insanity defenses or competency to stand trial.<sup>25</sup> The mental state at the time of the offense and during the trial influences sentencing, treatment requirements, and sometimes the entire outcome of the legal process. AI-driven tools, such as those employing speech and facial analysis, have been proposed to aid in assessing mental states by detecting subtle behavioral cues that may indicate distress, deception, or specific mental health symptoms. Another studies conducted showed that AI tools can reliably analyze speech patterns to identify mental health disorders with reasonable accuracy, especially in controlled environments.<sup>26</sup> A systematic review of studies using speech for automated assessments across a broader range of psychiatric disorders.<sup>26</sup> Additionally, the narrative review thoroughly examined the use of AI in mental health care and successfully highlighted the increasing awareness of AI's potential to aid and improve mental well-being.<sup>27</sup> Thus, suggesting a future where AI plays

an integral role in addressing global mental health crisis. However, the application of AI in mental state evaluations is constrained by the variability of human behavior and context-dependent nuances that cannot be fully captured by current algorithms.<sup>28</sup> Additionally, ethical issues arise from using AI-driven evaluations in a legal context, where the stakes are high and any inaccuracies in evaluation could lead to grave legal consequences.<sup>29</sup>

#### 2.4. Treatment compliance

In forensic settings, treatment compliance is essential to reduce recidivism and aid in the rehabilitation of individuals with mental health needs.<sup>30</sup> Forensic mental health practitioners face unique challenges in ensuring compliance, especially when treatment is court-mandated. AI could support treatment compliance monitoring by tracking behavioral patterns, attendance, and adherence to prescribed interventions through mobile applications or wearable devices. A study conducted revealed that digital tools powered by AI can predict treatment adherence based on user data, identifying potential lapses early on.<sup>31</sup> Nevertheless, the use of AI to monitor compliance raises privacy concerns, as continuous surveillance can be perceived as invasive. Furthermore, mandatory AI-based monitoring may encounter resistance from individuals, as the potential for misuse or mishandling of sensitive data could lead to mistrust and affect their willingness to participate fully in treatment.<sup>32</sup>

### 3. Data constraints in forensic mental health

AI's success in forensic mental health depends heavily on the quality, quantity, and diversity of data used to train algorithms.<sup>33</sup> However, the forensic mental health domain faces specific data challenges, including limitations in data availability, potential biases, and concerns about data quality. These constraints pose significant barriers to developing effective, fair, and reliable AI applications.<sup>8</sup>

#### 3.1. Issues around data availability

The forensic mental health field faces unique data constraints because case data is often protected due to confidentiality, legal restrictions, and ethical considerations.<sup>34</sup> Access to high-quality, longitudinal data is limited, restricting the capacity to develop comprehensive models. This limited data availability also hampers the diversity of datasets, leading to less representative models.<sup>35</sup> Without diverse data, AI models risk being undertrained or poorly generalized, which could lead to inaccurate assessments when applied to a broader population. Moreover, since forensic cases are often rare and complex, data is inherently limited, requiring alternative data-gathering methods, such as data synthesis, which might not accurately reflect real-world forensic mental health cases.<sup>36</sup>

#### 3.2. Biases in forensic data

Data biases are especially problematic in forensic mental health, where training data might reflect historical inequalities within the criminal justice system. For instance, if an AI model is trained on historical criminal records, it could inadvertently inherit biases present in those records, which may disproportionately affect certain demographics, particularly marginalized communities.<sup>37</sup> Studies indicate that AI models trained on biased data can perpetuate or even exacerbate existing disparities, leading to unfair risk assessments or profiling.<sup>38</sup> Bias mitigation techniques, such as fairness-aware algorithms and balanced sampling, are being explored; however, they are still not foolproof, and bias remains a persistent issue in forensic AI applications. Ensuring fairness in forensic AI requires ongoing data monitoring, robust bias-detection tools, and consideration of systemic issues that might influence AI-generated outcomes.

#### 3.3. Data quality in forensic cases

For AI models to be reliable, data quality is paramount. In forensic mental health, data quality can be inconsistent, as records might be incomplete, subjective, or contain errors. Human error in data entry, inconsistent reporting practices, and variations in data standards across regions all contribute to this problem.<sup>28</sup> A study conducted found that data quality issues led to significantly lower predictive accuracy in forensic AI models, underlining the importance of data integrity in achieving reliable AI outcomes.<sup>39</sup> This study also looked closely at how AI and ML techniques are used in digital forensics and incident response. Low-quality data not only impacts accuracy but can also lead to erroneous interpretations, which are especially detrimental in forensic settings where AI-driven decisions can have severe consequences. Addressing data quality issues may require standardized protocols for data entry, cross-agency collaboration, and rigorous quality control processes.

### 4. AI applications in forensic mental health

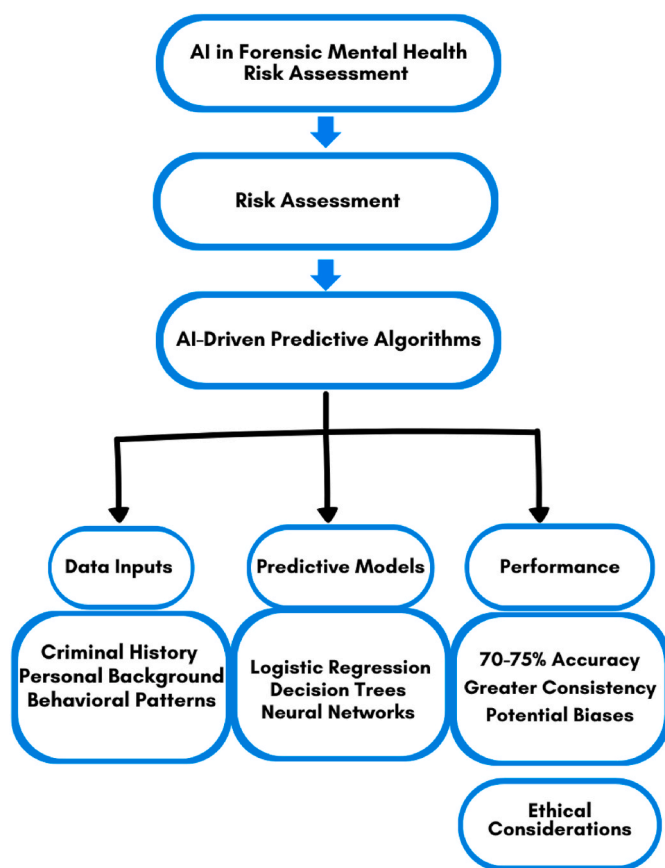
Artificial intelligence is being increasingly applied in forensic mental health, addressing critical areas such as risk assessment, mental health screening, behavioral analysis, and treatment recommendations. These applications offer the potential to improve the accuracy, objectivity, and efficiency of forensic evaluations, thereby supporting judicial decisions and mental health interventions with greater data-driven precision.<sup>28,40</sup> Below, the key AI applications in this field are explored, including specific tools and examples that highlight their utility and challenges in forensic settings.

#### 4.1. Risk assessment and prediction

One of the primary applications of AI in forensic mental health is in risk assessment, which involves evaluating the likelihood of violence, recidivism, or self-harm among individuals involved in the criminal justice system. AI-driven predictive algorithms are employed to enhance the accuracy of these assessments, utilizing vast datasets to identify risk factors and behavioral patterns. Predictive models such as logistic regression, decision trees, and neural networks have demonstrated promising results in assessing risk. For example, the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) tool uses machine learning to predict recidivism risk based on a variety of individual data points, such as criminal history and personal background.<sup>41</sup> Some studies did also revealed that COMPAS and similar algorithms often perform with greater consistency than traditional assessment methods, with predictive accuracy rates around 70–75%.<sup>42</sup> Fig. 1 presents a simplified flowchart illustrating the application of AI in forensic mental health risk assessment. However, these tools have also been critiqued for potential biases, particularly regarding racial disparities, which can impact their fairness and acceptance in forensic contexts. Thus, while AI can enhance the efficiency and standardization of risk assessments, it also necessitates careful consideration of ethical implications to prevent exacerbating existing inequalities.<sup>43</sup>

#### 4.2. Mental health screening and diagnosis

AI is also being applied to aid in the psychological evaluations required in legal cases, offering tools that assist with mental health screening and diagnosis. Traditional diagnostic procedures can be time-consuming and subjective, relying heavily on clinician interpretation.<sup>44</sup> AI-based tools provide a more objective analysis by leveraging data from psychological tests, clinical records, and even direct interactions to assess mental health conditions more systematically. Automated tools that analyze responses from structured interviews or digital questionnaires have shown significant potential in supporting diagnostic decisions. For example, the AI platform MindStrong utilizes machine



**Fig. 1.** Schematic representation of AI application in forensic mental health risk assessment. This figure illustrates the flow from concept to implementation, key components, and ethical considerations. This diagram showcases the integration of data inputs, predictive models, and performance metrics in AI-driven risk assessment tools, while also emphasizing the critical role of ethical considerations in their development and use.

learning algorithms to analyze mobile device interactions—such as typing patterns and screen navigation—to detect cognitive patterns associated with various mental health disorders.<sup>45</sup> By identifying deviations in these patterns, MindStrong can signal the presence of conditions like depression or anxiety with high sensitivity. Furthermore, automated analysis tools that track verbal and non-verbal responses during interviews can offer insights into emotional states and potential behavioral risks.<sup>46</sup> Although promising, the use of AI in diagnosis must be approached with caution, as nuances in individual cases may not always be accurately captured by algorithms, especially in the complex and high-stakes context of forensic evaluations.

#### 4.3. Natural language processing for behavioral insights

Natural language processing (NLP), a branch of AI focused on analyzing language patterns, is particularly valuable in forensic mental health for gaining insights into an individual's cognitive and emotional state. NLP algorithms can analyze written or spoken language to detect signs of mental health issues, as well as identify behavioral patterns relevant to forensic assessments.<sup>47</sup> For instance, AI-driven NLP models can process court transcripts, interview recordings, or written reports to identify language markers of psychological distress, deception, or cognitive distortions. Research indicates that NLP tools can identify cues of depression, anxiety, and other mental health conditions based on language patterns, such as tone, sentiment, and complexity of speech.<sup>48</sup> Moreover, these models have been used to detect deception by analyzing linguistic inconsistencies, speech hesitations, or changes in sentence

structure during testimony or interrogations. For example, IBM's Watson system has been used to assess the veracity of statements by analyzing language for patterns indicative of lying or cognitive stress. While effective, NLP tools in forensic settings must be carefully calibrated to account for individual differences in language use, as failing to do so could lead to inaccurate interpretations and potential miscarriages of justice.<sup>49</sup>

#### 4.4. Computer vision in behavioral assessment

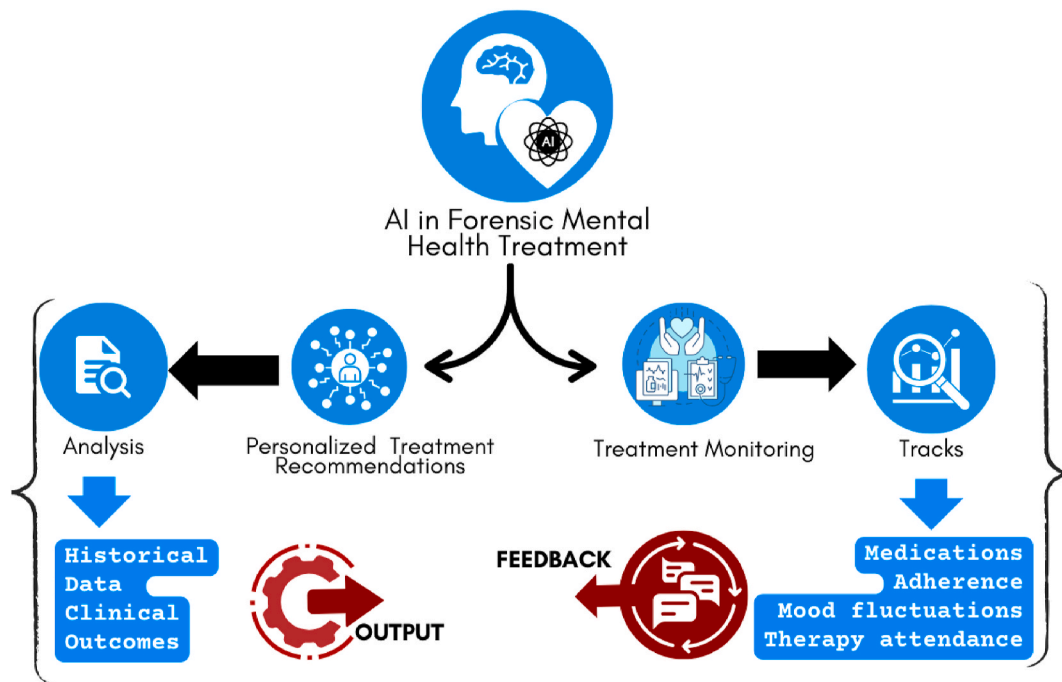
Computer vision is another powerful AI technology used to interpret visual cues relevant to forensic mental health assessments. By analyzing facial expressions, eye movements, and body language, computer vision algorithms can provide objective measures of emotional and behavioral states that are crucial in evaluating mental health conditions.<sup>50</sup> Facial recognition software, for example, can detect micro-expressions associated with emotions like anger, fear, or sadness, which may indicate underlying psychological states or intentions. Eye-tracking technology is also utilized to monitor gaze patterns, which can provide insights into attentional focus, arousal levels, and even deception. Additionally, movement analysis, which tracks body posture and gestures, can assist in understanding behavioral tendencies that might relate to aggression or risk of harm. In a study conducted by a group of scholars, computer vision models achieved an accuracy of over 80 % in detecting signs of agitation or distress during interviews, proving useful in assessing individuals with potentially volatile behavior.<sup>51</sup> Despite these advancements, ethical and privacy concerns are significant, as continuous surveillance and interpretation of personal behaviors raise questions about consent and the right to privacy, especially in mandatory assessment situations.<sup>52</sup>

#### 4.5. Treatment recommendations and monitoring

In the realm of forensic mental health, treatment is often directed toward reducing recidivism and aiding rehabilitation, requiring that treatment plans be closely tailored to each individual. AI is employed in creating personalized treatment recommendations by analyzing historical treatment data, clinical outcomes, and individual characteristics to propose intervention plans.<sup>53</sup> Machine learning algorithms can predict treatment responses, suggesting interventions that align with an individual's unique psychological profile and treatment history. For example, tools like Quartet Health employ predictive models to match patients with appropriate treatment resources based on their needs, increasing the likelihood of treatment adherence and positive outcomes.<sup>54</sup> AI is also instrumental in treatment monitoring, where wearable devices and mobile applications can track physiological and behavioral data to gauge compliance with treatment plans. For instance, AI-enabled applications can monitor medication adherence, mood fluctuations, and participation in therapy sessions, providing real-time feedback to both the individual and the clinician.<sup>55</sup> A study by Bain et al., 2017 indicate that AI-based monitoring can improve treatment adherence rates by 20–30 % by identifying lapses and intervening promptly.<sup>56</sup> However, such monitoring raises ethical considerations, as individuals may perceive constant surveillance as intrusive, potentially impacting their engagement with the treatment process. Fig. 2 presents a simple illustration for the application of AI in forensic mental health treatment. Ensuring transparent communication about AI's role in treatment monitoring is essential to maintain trust and cooperation in forensic settings.<sup>57</sup>

### 5. Benefits of AI in forensic mental health

The integration of artificial intelligence in forensic mental health has introduced transformative benefits, allowing for more accurate, efficient, and equitable decision-making processes in a complex field where high-stakes outcomes are the norm. AI tools can analyze vast datasets,



**Fig. 2.** AI applications in forensic mental health treatment. This figure illustrates the dual role of AI in generating personalized treatment recommendations and monitoring treatment progress. It highlights the key inputs, processes, and outputs for each application, demonstrating how AI integrates various data sources to enhance treatment efficacy and patient monitoring in forensic mental health contexts.

reduce subjective biases, enable early intervention, and optimize resource allocation, ultimately enhancing the quality of forensic mental health services and their impact on the criminal justice system.<sup>58</sup>

### 5.1. Increased accuracy and efficiency

One of the primary benefits of AI in forensic mental health is its ability to analyze large datasets rapidly and accurately, improving the precision of predictions regarding mental health and risk assessments. Traditional methods rely on clinical interviews and standardized assessment tools, which can be limited by human error and subjective interpretation.<sup>28,59</sup> AI, on the other hand, can process complex datasets, such as behavioral records, criminal history, and psychological assessments, identifying patterns that may not be immediately evident to clinicians. For instance, machine learning algorithms applied in forensic risk assessment, such as those used in predictive models for recidivism or violent behavior, have demonstrated enhanced predictive accuracy compared to traditional risk assessment tools.<sup>60</sup> This heightened accuracy allows forensic mental health practitioners to make more reliable determinations, thereby contributing to fairer and more informed judicial decisions.

### 5.2. Objective analysis

The use of AI introduces a level of objectivity in forensic mental health evaluations that traditional methods often lack. Forensic assessments are frequently subject to biases stemming from human factors, such as clinician experiences, cultural perceptions, and implicit biases.<sup>13</sup> AI mitigates this issue by relying on data-driven algorithms that can reduce the influence of subjective judgments, offering a more standardized approach to evaluations. By using AI algorithms trained on diverse datasets, forensic assessments can avoid some of the common biases that may affect human evaluators. A recent study showed that AI-based risk assessment tools, when properly calibrated, provided less biased outcomes in recidivism predictions across different demographic groups.<sup>61</sup> While not entirely free from biases—since AI models reflect

the data they are trained on—the structured and consistent application of these tools still reduces some of the more overt subjective influences, leading to more equitable and balanced assessments.

### 5.3. Early detection and intervention

AI has shown significant promise in the early detection of mental health issues and high-risk behaviors, facilitating timely interventions that could reduce the likelihood of crime or escalation of mental health crises.<sup>33</sup> By leveraging predictive analytics, AI can identify individuals at high risk of violent behavior, self-harm, or other dangerous behaviors earlier than traditional methods. This capability is crucial in forensic settings, where timely intervention can prevent potential harm to individuals and society. For example, predictive models developed to monitor behavioral patterns have been used to flag individuals with a high likelihood of reoffending or violating parole terms.<sup>19</sup> Early detection mechanisms can enable preemptive actions, such as targeted mental health interventions, enhanced supervision, or placement in rehabilitative programs. A study demonstrated that AI-powered early intervention systems reduced reoffending rates by 15 % among high-risk individuals, highlighting the potential of AI to contribute to crime prevention.<sup>62</sup> By identifying high-risk cases sooner, forensic practitioners can implement preventive measures more effectively, promoting public safety and reducing the demands on the criminal justice system.

### 5.4. Resource allocation

AI-driven insights allow for the strategic allocation of resources, helping forensic mental health practitioners and legal authorities prioritize cases based on risk and urgency. Forensic mental health resources are often limited, with caseworkers, treatment programs, and monitoring systems under substantial strain.<sup>33</sup> AI's ability to categorize and prioritize cases according to assessed risk levels enables agencies to allocate their resources more effectively, ensuring that high-risk individuals receive timely attention. This optimization can be particularly beneficial in settings where resources are stretched thin, allowing

practitioners to address the most critical cases first and reducing case backlogs. For example, AI tools that assess mental health risk and treatment needs can direct resources toward individuals requiring intensive intervention while recommending less-intensive support for lower-risk cases. This tiered approach not only optimizes resource use but also enhances the outcomes of forensic mental health programs by ensuring that interventions are appropriately matched to individual needs. Fig. 3 provides an overview of the benefits of AI in forensic mental health. It highlights four main advantages: increased accuracy and efficiency, objective analysis, early detection of issues, and better resource allocation. It was reported that AI-assisted resource allocation has improved efficiency in forensic case management by 20 %, underscoring its potential to streamline case workflows and improve outcomes.<sup>63</sup>

## 6. Future directions and research opportunities

As AI continues to evolve, its application in forensic mental health offers new pathways for innovation, requiring robust frameworks and multidisciplinary research to address current challenges and enhance its efficacy. The future of AI in this field hinges on improving data quality, fostering interdisciplinary collaboration, ensuring fairness and transparency, advancing predictive tools, and establishing ethical regulatory standards. Each of these areas presents substantial research opportunities to build a responsible, effective, and fair AI ecosystem in forensic mental health.<sup>34</sup>

### 6.1. Improving data quality and accessibility

Data quality is critical for AI's success in forensic mental health, yet the sector often faces issues with incomplete or biased data that can undermine the reliability of AI applications. Improving data quality involves curating datasets that are accurate, representative, and ethically sourced. One promising avenue is the promotion of open-source data initiatives, which facilitate access to a diverse range of high-quality datasets while maintaining ethical standards.<sup>64</sup> Open-source

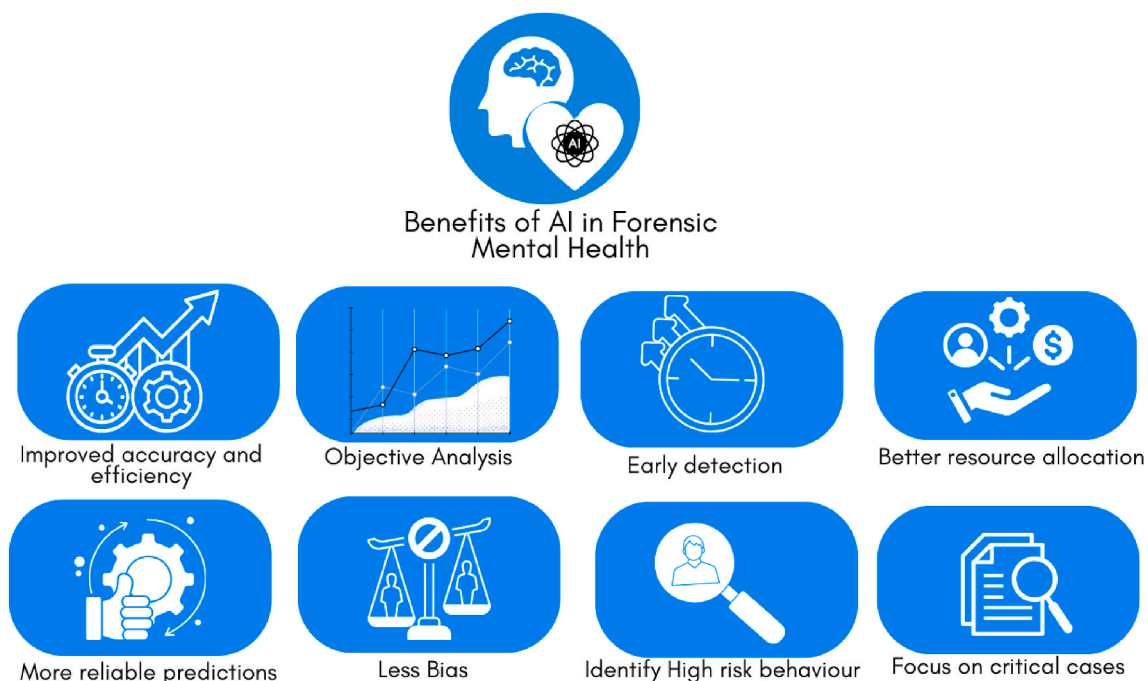
data can democratize AI research, allowing a wider pool of researchers to refine algorithms and improve model robustness across varying demographic and situational contexts. Additionally, more ethical data collection practices, such as obtaining informed consent and anonymizing sensitive information, are essential to protect individual rights and uphold privacy. These approaches will require ongoing research into secure, responsible data-sharing methods that ensure data integrity while minimizing risks of misuse. By focusing on quality and accessibility, future AI tools can be better trained to produce more reliable, fair outcomes, ultimately leading to more trustworthy applications in forensic settings.<sup>65</sup>

### 6.2. Interdisciplinary collaboration

The development of AI applications for forensic mental health cannot be confined to the field of AI alone. Integrating insights from psychology, ethics, law, and criminology is vital to creating tools that address the complexities inherent in this field. Interdisciplinary collaboration fosters a holistic approach, enabling the creation of AI solutions that are scientifically rigorous, ethically sound, and legally viable.<sup>66</sup> For example, psychologists and mental health professionals can contribute insights into behavior patterns and mental health diagnostics, ensuring AI tools are contextually relevant and accurate. Ethical scholars can provide guidance on developing tools that respect individual rights and avoid stigmatization, while legal professionals can assist in aligning AI tools with judicial standards.<sup>67</sup> Establishing forums, conferences, and collaborative research projects that bring together experts from these diverse fields can foster the innovation needed to address complex forensic mental health issues. Research in interdisciplinary methodologies, such as how to effectively integrate psychological and legal insights into AI models, will be essential for developing nuanced solutions that meet the demands of forensic applications.<sup>68</sup>

### 6.3. Focus on algorithmic fairness and transparency

Algorithmic fairness and transparency are central to the ethical use



**Fig. 3.** Overview of the benefits of AI in Forensic Mental Health. This diagram illustrates four key advantages: increased accuracy and efficiency, objective analysis, early detection of high-risk behaviors, and improved resource allocation. Each benefit leads to specific positive outcomes that enhance decision-making and service delivery in forensic mental health settings.

of AI in forensic mental health, particularly given the high stakes involved in forensic assessments, such as risk predictions and criminal profiling.<sup>69</sup> As biases in data can lead to unfair treatment of certain groups, ongoing research must focus on developing frameworks that ensure algorithmic fairness. This involves identifying and mitigating biases in training data, as well as adopting fairness-aware algorithms that can adjust for disparities among demographic groups. Transparency is equally important, as the “black box” nature of many AI algorithms can obscure how decisions are made, leading to mistrust and resistance in legal contexts. Developing interpretable models that allow forensic professionals to understand and validate AI-driven assessments is a priority for responsible AI.<sup>70</sup> Transparent AI frameworks will also improve accountability, ensuring that practitioners can trace decision-making processes and verify the soundness of assessments. Future research should explore explainability techniques that strike a balance between transparency and model complexity, as well as methodologies for continuously auditing AI systems to uphold fairness over time.

#### 6.4. Expansion of AI-driven predictive tools

AI-driven predictive tools offer considerable promise in forensic mental health, particularly for assessing risks and anticipating behavioral patterns.<sup>71</sup> However, current models have limitations in accuracy and generalizability, highlighting the need for further development of more precise predictive tools. Future research can explore advanced machine learning techniques, such as deep learning and ensemble methods, to enhance the predictive capabilities of AI in this context. These tools could be designed to evaluate not only static factors, such as criminal history, but also dynamic factors, such as recent behavior or changes in mental health status. The incorporation of real-time data from wearable devices or mobile applications also presents new research opportunities for continuous monitoring and prediction of high-risk behaviors. Improved predictive models can lead to better-tailored interventions and more precise risk assessments, ultimately contributing to more effective rehabilitation and public safety efforts.<sup>72</sup> As predictive tools evolve, however, researchers must remain vigilant about ethical considerations, particularly the risks of over-surveillance and the potential for infringing on individual freedoms.

#### 6.5. Policy development for ethical AI use

The adoption of AI in forensic mental health necessitates comprehensive policies and regulations to ensure ethical use, protect individual rights, and promote accountability. Given the sensitive nature of forensic assessments, policy frameworks must address issues such as data privacy, informed consent, and the appropriate scope of AI applications.<sup>15</sup> Regulations should specify when and how AI can be used in forensic settings, particularly in high-stakes applications like risk assessment and criminal profiling. Policy development should involve stakeholders across various fields, including AI researchers, legal experts, ethicists, and mental health professionals, to establish balanced guidelines that safeguard both individuals and the integrity of the judicial system. Standardizing AI use in forensic mental health can also facilitate public trust, as clear policies provide transparency on how AI tools are employed and outline protocols for addressing potential errors or biases.<sup>73</sup> Researchers and policymakers alike should consider how international regulations, such as the EU’s GDPR or forthcoming AI legislation, could inform national policies to create a coherent framework for AI ethics in forensic mental health.

### 7. Regulatory frameworks and the need for harmonized international standards

Artificial intelligence (AI) applications in forensic mental health are developing rapidly, yet the regulatory landscape remains fragmented

across jurisdictions, creating challenges for consistent, ethical, and effective implementation.<sup>28</sup> Existing regulatory frameworks vary widely in their focus, comprehensiveness, and enforceability, leaving gaps in oversight that could compromise the responsible use of AI in forensic settings.<sup>74</sup> This section provides a brief comparative analysis of key regulatory frameworks in different jurisdictions and underscores the urgent need for harmonized international standards.

#### 7.1. Comparative analysis of regulatory frameworks

In the United States, AI regulation in forensic and healthcare contexts largely falls under existing laws, such as the Health Insurance Portability and Accountability Act (HIPAA) and the Federal Trade Commission (FTC) Act.<sup>75</sup> HIPAA provides guidance on data privacy and security but does not specifically address AI’s unique challenges, such as algorithmic transparency and bias. The FDA has begun to issue guidelines for AI/ML-based medical devices, but these are limited to healthcare applications and do not comprehensively cover forensic uses.<sup>76</sup> State-level initiatives, like California’s AI accountability measures, highlight growing awareness but also contribute to a patchwork regulatory approach.

The European Union (EU) has been proactive in establishing comprehensive regulations for AI. The EU’s General Data Protection Regulation (GDPR) sets strict rules for data processing, consent, and individual rights, which have implications for AI use in forensic contexts.<sup>77</sup> Additionally, the proposed AI Act aims to establish a risk-based framework for AI applications, including stringent requirements for high-risk systems like those used in forensic mental health. While these measures are robust, their implementation across member states can result in inconsistencies.

China has introduced sector-specific guidelines, such as the AI Ethics Governance Guidelines and regulations under its Cybersecurity Law.<sup>78</sup> These frameworks emphasize data security, algorithmic transparency, and accountability. However, China’s regulatory focus is often criticized for prioritizing state interests over individual rights, raising concerns about privacy and ethical AI use in forensic applications. Countries like Canada and Australia have also introduced frameworks for AI governance, such as Canada’s Directive on Automated Decision-Making and Australia’s AI Ethics Framework.<sup>79</sup> While these guidelines address general principles like fairness, accountability, and transparency, they often lack enforceable regulations specific to AI in forensic settings.

#### 7.2. Challenges of fragmented regulations

The variability in regulatory frameworks across jurisdictions creates several challenges. First, inconsistencies in data privacy standards, algorithmic accountability, and ethical oversight make it difficult for multinational AI developers to ensure compliance across markets. Second, these differences can lead to unequal protections for individuals undergoing forensic assessments, potentially exacerbating existing disparities in justice systems. Finally, the lack of global standards complicates efforts to validate and benchmark AI systems, hindering cross-border collaboration and research.<sup>76</sup>

#### 7.3. The need for harmonized international standards

Given the global nature of AI development and its widespread applications, harmonized international standards are essential to address the challenges posed by fragmented regulatory frameworks. Consistent rules for data privacy and security are crucial to protect sensitive forensic data, particularly in cases where individuals’ mental health information is subject to scrutiny.<sup>80</sup> A unified approach to algorithmic transparency is also needed to ensure AI-driven decisions in forensic contexts are understandable and can be challenged when necessary. Standards for bias mitigation should require rigorous detection and correction protocols to promote fairness and equity, reducing the risk of

exacerbating existing disparities in justice systems. Additionally, clear guidelines for accountability and oversight are essential to define the roles and responsibilities of developers, users, and regulators, ensuring that all stakeholders uphold ethical practices. International standards must also prioritize validation and benchmarking processes to establish consistent criteria for evaluating the effectiveness and reliability of AI systems in forensic settings. Efforts such as the OECD's AI Principles and UNESCO's Recommendation on the Ethics of AI provide valuable foundations, but these initiatives must be expanded and tailored specifically to the complexities of forensic applications. By fostering collaboration between governments, international organizations, AI developers, and forensic professionals, it is possible to create actionable and enforceable standards that support responsible AI use while safeguarding individual rights and public trust.<sup>81</sup>

Overall, the lack of harmonized international standards for AI regulation in forensic mental health poses significant ethical, legal, and operational challenges. While existing frameworks in jurisdictions like the EU, US, and China provide valuable starting points, their inconsistencies highlight the need for a globally coordinated approach. By establishing unified standards, stakeholders can ensure that AI systems are used responsibly, equitably, and effectively in forensic settings, safeguarding both individual rights and public trust.

## 8. Challenges, limitations, and ethical considerations

The application of artificial intelligence (AI) in forensic mental health holds significant promise, but its real-world implementation comes with a range of challenges, limitations, and ethical considerations. While AI can enhance accuracy, efficiency, and objectivity, these benefits are accompanied by practical difficulties, inherent pitfalls, and complex ethical and legal concerns that need to be addressed for successful adoption.<sup>82</sup>

### 8.1. Practical challenges in real-world implementation

The integration of AI tools into forensic mental health settings faces several practical barriers, starting with resistance from practitioners. Many forensic evaluators are accustomed to traditional methods and may lack familiarity with AI technologies. This resistance often stems from concerns about the reliability of AI tools, their interpretability, and the potential erosion of professional judgment.<sup>33</sup> For example, practitioners may question the validity of AI-driven risk assessments if the underlying algorithms are not transparent or easily understood. To overcome this challenge, training programs and workshops tailored to forensic professionals can help bridge the knowledge gap, fostering trust and confidence in AI tools.

Legal constraints also pose significant hurdles. In many jurisdictions, the admissibility of AI-driven evidence in court remains ambiguous. Questions about the reliability of AI outputs, the opacity of algorithms, and the potential for bias can lead to reluctance among legal professionals to rely on AI-generated insights. Clear legal frameworks that define the admissibility criteria for AI-driven evaluations are necessary to build trust in these tools and ensure their ethical application. Moreover, integrating AI tools into existing forensic workflows requires significant infrastructural adjustments. Many organizations lack the technological infrastructure or financial resources to adopt AI solutions. Establishing partnerships with technology providers and securing funding through government or private grants could alleviate this issue, making AI adoption more feasible.

### 8.2. Limitations and pitfalls

Despite its potential, AI in forensic mental health is not without limitations. One critical limitation is the issue of data quality. AI models rely heavily on large datasets for training and validation, but forensic mental health data is often scarce, fragmented, or biased. Inconsistent

data collection practices, incomplete records, and demographic disparities can lead to models that perform poorly or reinforce existing biases.<sup>13</sup> Another significant limitation is the risk of over-reliance on AI tools. While these tools can provide valuable insights, they are not infallible and may produce false positives or negatives. For instance, an AI system might incorrectly label an individual as high-risk due to errors in the data or algorithmic misinterpretation. Such errors can have severe consequences in forensic contexts, affecting legal outcomes and individual rights.

Additionally, AI systems often lack contextual understanding. Unlike human practitioners, AI cannot consider the nuanced, case-specific factors that are critical in forensic evaluations. For example, an algorithm might overlook cultural or environmental factors that influence behavior, leading to incomplete or inaccurate assessments. Addressing these limitations requires ongoing validation and refinement of AI models, as well as their use as complementary tools rather than replacements for human judgment.<sup>83</sup>

### 8.3. Ethical and legal challenges

The ethical challenges of using AI in forensic mental health are multifaceted. One of the most pressing concerns is the potential for algorithmic bias. If AI systems are trained on biased datasets, they can perpetuate or even amplify existing disparities, particularly for marginalized communities.<sup>15</sup> For example, an AI model trained on historical data from a justice system with systemic racial bias might produce risk assessments that disproportionately affect minority populations. To mitigate this, developers must prioritize bias detection and correction during the design and testing phases of AI systems.

Privacy is another significant ethical concern. Forensic mental health evaluations often involve highly sensitive personal information, raising questions about how data is collected, stored, and used. Ensuring compliance with data protection regulations such as GDPR and HIPAA is essential, but additional safeguards may be necessary to address the unique sensitivities of forensic data. Transparency and accountability are also critical. Many AI systems function as "black boxes," producing outputs without clear explanations of how decisions were reached. This lack of transparency can undermine trust in AI tools, particularly in legal settings where decision-makers require clear justifications for their actions. Developing explainable AI systems that provide interpretable outputs is essential to address this issue.<sup>84</sup>

From a legal perspective, the question of liability is particularly contentious. If an AI-driven evaluation produces an erroneous risk assessment, it is unclear who should be held accountable—the developer, the user, or the organization deploying the tool. Establishing clear legal guidelines for AI accountability is crucial to ensure fairness and prevent misuse. Finally, the use of AI in forensic mental health raises broader societal questions about the balance between public safety and individual rights. While AI tools can improve public safety by identifying high-risk individuals, they must not infringe on the rights of those undergoing forensic evaluations. Ethical guidelines and legal frameworks should ensure that AI applications prioritize human dignity and fairness.<sup>70</sup>

To navigate these challenges and limitations, several strategies can be employed. First, fostering interdisciplinary collaboration among AI developers, forensic professionals, ethicists, and legal experts can ensure that AI tools are designed with a holistic understanding of forensic contexts. Second, implementing robust regulatory frameworks that address data quality, algorithmic fairness, and accountability is essential for responsible AI use. Third, ongoing training and education for practitioners can help reduce resistance and improve the integration of AI into forensic workflows. Lastly, adopting a cautious and iterative approach to AI implementation—starting with pilot programs and rigorous validation studies—can help identify and address potential issues before broader deployment.

## 9. Conclusion

AI has emerged as a transformative tool in advancing forensic mental health, providing enhanced accuracy, efficiency, and objectivity in assessments that shape critical outcomes within the criminal justice system. From predictive risk assessment and mental health diagnostics to behavioral analysis and treatment recommendations, AI applications have introduced data-driven precision into a field traditionally limited by subjective interpretation and resource constraints. The potential benefits of AI in improving early intervention, streamlining resource allocation, and promoting fairer, more consistent evaluations underscore its significant contribution to forensic mental health practices.

However, the integration of AI in forensic contexts must be approached with a balanced perspective, recognizing both the technological possibilities and the ethical and legal responsibilities inherent in its use. The high stakes associated with forensic mental health evaluations amplify the need for AI models that are both reliable and equitable, as decisions based on these models can significantly impact individuals' lives and public safety. Ethical considerations, including data privacy, algorithmic fairness, and transparency, are essential to prevent bias, ensure individual rights, and uphold public trust in AI-driven evaluations. Legal frameworks must also adapt to govern AI use responsibly, establishing guidelines that protect against misuse and ensure accountability within forensic practices.

A continued focus on interdisciplinary research, ethical oversight, and cautious implementation is necessary to realize the full potential of AI in forensic mental health. Collaboration among AI developers, mental health professionals, ethicists, and legal experts will be key to developing tools that not only advance the field but also respect the rights of individuals and communities. As AI technology progresses, so too must our commitment to responsible use, ensuring that AI applications in forensic mental health are guided by both innovation and integrity. This careful balance will be essential to foster trust, improve justice outcomes, and create an ethical framework that supports both technological advancement and human dignity.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- van Es RMS, Kunst MJJ, de Keijser JW. Forensic mental health expert testimony and judicial decision-making: a systematic literature review. *Aggress Violent Behav.* 2020 Mar 1;51, 101387.
- Fallon-Kund M, Bickenbach J. Strengthening the voice of persons with mental health problems in legal capacity proceedings. *Laws.* 2016;5:29, 2016 Jun 29 [Internet] [cited 2024 Nov 10];5(3):29 <https://www.mdpi.com/2075-471X/5/3/29/html>.
- De Vries Robbé M, De Vogel V, Douglas KS, Nijman HLI. Changes in dynamic risk and protective factors for violence during inpatient forensic psychiatric treatment: predicting reductions in postdischarge community recidivism. *Law Hum Behav.* 2015;39(1):53.
- Berryessa CM, Balavender A. *The Value of Remorse as a Therapeutic Tool for Probation Officers in Sentencing.* 2021.
- Fazeli S, Mohammadi Zeidi I, Lin CY, et al. Depression, anxiety, and stress mediate the associations between internet gaming disorder, insomnia, and quality of life during the COVID-19 outbreak. *Addict Behav Reports.* 2020 Dec 1;12, 100307.
- Bhattamisra SK, Banerjee P, Gupta P, Mayuren J, Patra S, Candasamy M. Artificial intelligence in pharmaceutical and healthcare research. *Big Data Cogn Comput.* 2023; 7(1):10.
- Chattu VK. A review of artificial intelligence, big data, and blockchain technology applications in medicine and global health. *Big Data Cogn Comput.* 2021;5(3):41.
- Graham S, Depp C, Lee EE, et al. Artificial intelligence for mental health and mental illnesses: an overview. *Curr Psychiatry Rep.* 2019;21:1–18.
- Kerz E, Zanwar S, Qiao Y, Wiechmann D. Toward explainable AI (XAI) for mental health detection based on language behavior. *Front psychiatry.* 2023;14, 1219479.
- Dawoodbhoy FM, Delaney J, Cecula P, et al. AI in patient flow: applications of artificial intelligence to improve patient flow in NHS acute mental health inpatient units. *Heliyon.* 2021;7(5), e06993.
- Ribeiro RAB, de Matos Soeiro CBB. Analysing criminal profiling validity: underlying problems and future directions. *Int J Law Psychiatr.* 2021;74.
- Wangmo T, Seaward H, Pageau F, Hiersemenzel L-P, Elger BS. Forensic-psychiatric risk evaluations: perspectives of forensic psychiatric experts and older incarcerated persons from Switzerland. *Front Psychiatr.* 2021;12, 643096.
- Timmons AC, Duong JB, Simo Fiallo N, et al. A call to action on assessing and mitigating bias in artificial intelligence applications for mental health. *Perspect Psychol Sci.* 2023;18(5):1062–1096.
- Marks P, Thornton H, Westbury B, Rix KJB, Shelley S. Response from the Faculty of Forensic and Legal Medicine to editorial on mental health services failing the criminal justice system. *BMJ (Pract Obs Ed).* 2022;376.
- Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J Med Internet Res.* 2019;21(5), e13216.
- Watts D, Moulden H, Mamak M, Upfold C, Chaimowitz G, Kapczynski F. Predicting offenses among individuals with psychiatric disorders - a machine learning approach. *J Psychiatr Res.* 2021 Jun;138(12):146–154.
- Kirchbner J, Günther M, Sonnweber M, King A, Lau S. Factors and predictors of length of stay in offenders diagnosed with schizophrenia - a machine-learning-based approach. *BMC Psychiatry.* 2020. <https://doi.org/10.1186/s12888-020-02612-1>.
- Noriega JP, Rivera LA, Herrera JA. Machine learning for credit risk prediction: a systematic literature review. *Data (Kb, 1971).* 2023;8(11):169.
- Parmigiani G, Barchielli B, Casale S, Mancini T, Ferracuti S. The impact of machine learning in predicting risk of violence: a systematic review. *Front psychiatry.* 2022; 13, 1015914.
- Piraiyanu A-I, Fulga A, Musat CL, et al. Enhancing the evidence with algorithms: how artificial intelligence is transforming forensic medicine. *Diagnostics.* 2023;13(18): 2992.
- Guo J. Research on the application of criminal psychology profiling technology in crime scene investigation. *Des Eng.* 2020:137–150.
- Siddiqui T, Amer AYA, Khan NA. Criminal activity detection in social network by text mining: comprehensive analysis. In: *2019 4th International Conference on Information Systems and Computer Networks (ISCON).* IEEE; 2019:224–229.
- Mello FL de, Souza SA de. Decision maker profiling using their mental behavior pattern. *Front Psychol.* 2021;12, 667255.
- Tollenaar N, van der Heijden PGM. Optimizing predictive performance of criminal recidivism models using registration data with binary and survival outcomes. *Stiglic G, ed. PLoS One.* 2019 Mar 8;14(3), e0213245.
- Guo Ningzhi. Identification of violent patient with schizophrenia using a hybrid machine learning approach at the individual level. *Psychiatry Res.* 2021;306(12), 114294.
- Low DM, Bentley KH, Ghosh SS. Automated assessment of psychiatric disorders using speech: a systematic review. *Laryngoscope Investig Otolaryngol.* 2020;5(1): 96–116.
- Olawade DB, Wada OZ, Odetayo A, David-Olawade AC, Asaolu F, Eberhardt J. Enhancing mental health with Artificial Intelligence: current trends and future prospects. *J Med Surg Public Health.* 2024 Aug;3, 100099.
- Tornero-Costa R, Martínez-Millana A, Azzopardi-Muscat N, Lazari L, Traver V, Novillo-Ortiz D. Methodological and quality flaws in the use of artificial intelligence in mental health research: systematic review. *JMIR Ment Heal.* 2023;10(1), e42045.
- Dutta BM. The ethics of artificial intelligence in legal decision making: an empirical study. *Psychol Educ.* 2018;55(1):292–302.
- Han W, Redlich AD. The impact of community treatment on recidivism among mental health court participants. *Psychiatr Serv.* 2016;67(4):384–390.
- Zhang S, Dieciuc M, Dilanchian A, Boot W. The development of adherence support messages for an ai-based reminder system. *Innov Aging.* 2022;6(Suppl 1):71.
- Mooghal M, Stroud AM, Yoo DW, et al. Barriers and facilitators to trustworthy and ethical AI-enabled medical care from patient and healthcare provider perspectives: a literature review. *medRxiv.* 2023:2010–2023.
- D'Alfonso S. AI in mental health. *Curr Opin Psychol.* 2020;36:112–117.
- Gooding P, Kariotis T. Ethics and law in research on algorithmic and data-driven technology in mental health care: scoping review. *JMIR Ment Heal.* 2021;8(6), e24668.
- Ernst AF, Albers CJ, Timmerman ME. A comprehensive model framework for between-individual differences in longitudinal data. *Psychol Methods.* 2023;29(4): 748–766. <https://psycnet.apa.org/doi/10.1037/1089-2888-020-02612-1>.
- Schick A, Rauschenberg C, Ader L, et al. Novel digital methods for gathering intensive time series data in mental health research: scoping review of a rapidly evolving field. *Psychol Med.* 2023;53(1):55–65.
- Meyer LF, Valença AM. Factors related to bias in forensic psychiatric assessments in criminal matters: a systematic review. *Int J Law Psychiatr.* 2021;75, 101681.
- DeAlcala D, Serna I, Morales A, Fierrez J, Ortega-García J. *Measuring bias in AI models: an statistical approach introducing N-Sigma.* IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC). IEEE; 2023:1167–1172, 2023.
- Dunsin D, Ghanem MC, Ouazzane K, Vassilev V. A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response. *Forensic Sci Int Digit Investig.* 2024;48, 301675.
- Zhou S, Zhao J, Zhang L. Application of artificial intelligence on psychological interventions and diagnosis: an overview. *Front Psychiatr.* 2022;13, 811665.
- Dressel J, Farid H. The accuracy, fairness, and limits of predicting recidivism. *Sci Adv.* 2018;4(1), eaa05580.

42. Zhang SX, Roberts REL, Farabee D. An analysis of prisoner reentry and parole risk using COMPAS and traditional criminal history measures. *Crime Delinquen*. 2014;60(2):167–192.
43. Hogan NR, Davidge EQ, Corabian G. On the ethics and practicalities of artificial intelligence, risk assessment, and race. *J Am Acad Psychiatry Law*. 2021;49(3):326–334.
44. Jang H, Oh Y-T, Kim A, Kim SK. User guiding information supporting application for clinical procedure in traditional medicine. In: *Human Interface and the Management of Information Information and Interaction for Health, Safety, Mobility and Complex Environments: 15th International Conference, HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part II 15*. Springer; 2013:100–109.
45. Naim I, Tanveer MI, Gildea D, Hoque ME. Automated analysis and prediction of job interview performance. *IEEE Trans Affect Comput*. 2016;9(2):191–204.
46. Priya K, Roomi SMM, Shanmugavadivu P, Sethuraman MG, Kalaivani P. An automated system for the assessment of interview performance through audio & emotion cues. In: *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*. IEEE; 2019:1049–1054.
47. Rezaii N, Wolff P, Price BH. Natural language processing in psychiatry: the promises and perils of a transformative approach. *Br J Psychiatry*. 2022;220(5):251–253.
48. Wright-Berryman J, Cohen J, Haq A, Black DP, Pease JL. Virtually screening adults for depression, anxiety, and suicide risk using machine learning and language from an open-ended interview. *Front psychiatry*. 2023;14, 1143175.
49. Wei Z, Chen Y, Zhao Q, et al. Implicit perception of differences between NLP-produced and human-produced language in the mentalizing network. *Adv Sci*. 2023;10(12), 2203990.
50. Tziomaka M, Kallipolitis A, Tsanakas P, Maglogiannis I. Evaluating mental patients utilizing video analysis of facial expressions. In: *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer; 2021:182–193.
51. Khan SS, Mishra PK, Javed N, et al. Unsupervised deep learning to detect agitation from videos in people with dementia. *IEEE Access*. 2022;10:10349–10358.
52. Nandy D. Human rights in the era of surveillance: balancing security and privacy concerns. *J Curr Soc Polit Issues*. 2023;1(1):13–17.
53. Parekh A-DE, Shaikh OA, Manan S, Al Hasibuzzaman M. Artificial intelligence (AI) in personalized medicine: AI-generated personalized therapy regimens based on genetic and medical history. *Ann Med Surg*. 2023;85(11):5831–5833.
54. Chow CK, Atkins ER, Billot L, et al. Ultra-low-dose quadruple combination blood pressure-lowering therapy in patients with hypertension: the QUARTET randomized controlled trial protocol. *Am Heart J*. 2021;231:56–67.
55. Chidambaram S, Maheswaran Y, Patel K, et al. Using artificial intelligence-enhanced sensing and wearable technology in sports medicine and performance optimisation. *Sensors*. 2022;22(18):6920.
56. Bain EE, Shafner L, Walling DP, et al. Use of a novel artificial intelligence platform on mobile devices to assess dosing compliance in a phase 2 clinical trial in subjects with schizophrenia. *JMIR mHealth uHealth*. 2017;5(2), e7030.
57. Morley J, Machado CCV, Burr C, et al. The ethics of AI in health care: a mapping review. *Soc Sci Med*. 2020;260, 113172.
58. Vo A, Plachkinova M. Investigating the role of artificial intelligence in the US criminal justice system. *J Information, Commun Ethics Soc*. 2023 Nov 7;21(4):550–567.
59. Tai AMY, Albuquerque A, Carmona NE, et al. Machine learning and big data: implications for disease modeling and therapeutic discovery in psychiatry. *Artif Intell Med*. 2019 Aug 1;99, 101704.
60. Travaini GV, Pacchioni F, Bellumore S, Bosia M, De Micco F. Machine learning and criminal justice: a systematic review of advanced methodology for recidivism risk prediction. *Int J Environ Res Publ Health*. 2022;19, 10594, 2022 Aug 25 [Internet] [cited 2024 Nov 11];19(17):10594 <https://www.mdpi.com/1660-4601/19/17/10594/htm>.
61. Skeem J, Lowenkamp C. Using algorithms to address trade-offs inherent in predicting recidivism. *Behav Sci Law [Internet]*. 2020 May 1;38(3):259–278 [cited 2024 Nov 11] <https://onlinelibrary.wiley.com/doi/full/10.1002/bsl.2465>.
62. Sukhodolov AP, Bychkova AM. Artificial intelligence in crime counteraction, prediction, prevention and evolution. *Russ J Criminol*. 2018;12(6):753–766.
63. Saravana Kumar G, Ramachandran KK, Sharma S, Ramesh R, Qureshi K, Ganesh KVB. AI-assisted resource allocation for improved business efficiency and profitability. *3rd Int Conf Adv Comput Innov Technol Eng ICACITE*. 2023;2023:54–58, 2023.
64. Wang AX, Chukova SS, Simpson CR, Nguyen BP. Data-centric AI to improve early detection of mental illness. *IEEE Work Stat Signal Process Proc.* 2023;2023-July:369–373.
65. Christoforaki M, Beyan O. AI ethics—a bird’s eye view. *Appl Sci*. 2022;12:4130 [Internet]. 2022 Apr 20 [cited 2024 Nov 11];12(9):4130. <https://www.mdpi.com/2076-3417/12/9/4130/htm>
66. Starke G, D’Imperio A, Ienca M. Out of their minds? Externalist challenges for using AI in forensic psychiatry. *Front Psychiatr*. 2023 Aug 24;14, 1209862.
67. Fukuda-Parr S, Gibbons E. Emerging consensus on ‘ethical AI’: human rights critique of stakeholder guidelines. *Glob Policy*. 2021;12(S6):32–44. Jul 1 [cited 2024 Nov 11] <https://onlinelibrary.wiley.com/doi/full/10.1111/1758-5899.12965>.
68. Landers RN, Behrend TS. Auditing the AI auditors: a framework for evaluating fairness and bias in high stakes AI predictive models. *Am Psychol*. 2022 Feb 14;78(1):36–49.
69. Chiao V. Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice. *Int J Law Context*. 2019;15(2):126–139.
70. Garrett BL, Rudin C. Interpretable algorithmic forensics. *Proc Natl Acad Sci India*. 2023 Oct 10;120(41), e2301842120 [cited 2024 Nov 11] <https://www.pnas.org/doi/abs/10.1073/pnas.2301842120>.
71. Tay JL, Li Z, Sim K. Effectiveness of artificial intelligence methods in personalized aggression risk prediction within inpatient psychiatric treatment settings—a systematic review. *J Personalized Med*. 2022;12:1470 [Internet]. 2022 Sep 7 [cited 2024 Nov 11];12(9):1470 <https://www.mdpi.com/2075-4426/12/9/1470/htm>.
72. Melnikov A, Kordzanganeh M, Alodjants A, Lee RK. Quantum machine learning: from physics to software engineering. *Adv Phys*. 2023 Dec 31;X [Internet] [cited 2024 Nov 11];8(1). Available from: <https://www.tandfonline.com/doi/abs/10.1080/23746149.2023.2165452>.
73. Higgins O, Short BL, Chalup SK, Wilson RL. Artificial intelligence (AI) and machine learning (ML) based decision support systems in mental health: an integrative review. *Int J Ment Health Nurs [Internet]*. 2023 Aug 1;32(4):966–978 [cited 2024 Nov 11] <https://onlinelibrary.wiley.com/doi/full/10.1111/inm.13114>.
74. Lorch B, Scheler N, Riess C. Compliance challenges in forensic image analysis under the artificial intelligence Act. In *2022 30th European Signal Processing Conference (EUSIPCO)*. 29. IEEE; 2022 Aug 29:613–617. <https://doi.org/10.23919/EUSIPCO.055093.2022.9909723>.
75. Vokinger KN, Gasser U. Regulating AI in medicine in the United States and Europe. *Nat Mach Intell*. 2021 Sep;3(9):738–739.
76. Yekaterina K. Challenges and opportunities for AI in healthcare. *International Journal of Law and Policy*. 2024 Jul 30;2(7):11–15.
77. Clarke R. Regulatory alternatives for AI. *Comput Law Secur Rep*. 2019 Aug;35(4):398–409.
78. Kingston J. Using artificial intelligence to support compliance with the general data protection regulation. *Artif Intell Law*. 2017 Sep 1;25(4):429–443.
79. Timmers P. Ethics of AI and cybersecurity when sovereignty is at stake. *Minds Mach*. 2019 Oct 11;29(4).
80. Radu R. Steering the governance of artificial intelligence: national strategies in perspective. *Policy and Society*. 2021 Apr 3;40(2):178–193.
81. Di Noia T, Tintarev N, Fatourou P, Schedl M. Recommender systems under European AI regulations. *Commun ACM*. 2022 Apr;65(4):69–73.
82. Etzioni A, Etzioni O. AI assisted ethics. *Ethics and Information Technology*. 2016 May 5;18(2):149–156.
83. Burnell R, Schellaert W, Burden J, et al. Rethink reporting of evaluation results in AI. *Science*. 2023 Apr 14;380(6641):136–138.
84. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell*. 2019 May;1(5):206–215.