



OPEN ACCESS

EDITED BY

Abdul Roudsari,
University of Victoria, Canada

REVIEWED BY

Armita Adily,
University of New South Wales, Australia
Agus Byna,
Universitas Sari Mulia, Indonesia

*CORRESPONDENCE

Giulio Nittari
✉ giulio.nittari@unicam.it

RECEIVED 10 May 2026

REVISED 03 June 2026

ACCEPTED 05 June 2026

PUBLISHED 23 June 2026

CITATION

Bailo P, Nittari G, Spasari T, Gibelli F and Ricci G (2026) Artificial intelligence in intimate partner violence risk pathways: a PRISMA-ScR review of femicide prevention and medico-legal accountability.
Front. Digit. Health 8:1877651.
doi: 10.3389/fdgth.2026.1877651

COPYRIGHT

© 2026 Bailo, Nittari, Spasari, Gibelli and Ricci. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Artificial intelligence in intimate partner violence risk pathways: a PRISMA-ScR review of femicide prevention and medico-legal accountability

Paolo Bailo¹, Giulio Nittari^{2*}, Tommaso Spasari³, Filippo Gibelli¹ and Giovanna Ricci¹

¹Section of Legal Medicine, School of Law, University of Camerino, Camerino, Italy, ²Telemedicine and Telepharmacy Centre, School of Medicinal and Health Products Sciences, University of Camerino, Camerino, Italy, ³Section of Occupational and Legal Medicine and BioLaw, Niccolò Cusano University, Rome, Italy

Introduction: Artificial intelligence (AI), machine learning, natural language processing and related decision-support methods are increasingly studied in intimate partner violence (IPV), domestic-violence and gender-based violence contexts. The key question is not whether AI can predict femicide as an individual lethal event, but whether AI-related methods may help institutions recognise, document, communicate and act on distributed signs of escalation across clinical, legal, police, social-service and digital settings.

Methods: This PRISMA-ScR scoping review, informed by Joanna Briggs Institute guidance and structured using the Population-Concept-Context framework, mapped English-language AI-related literature in IPV, domestic violence, coercive-control and femicide-related risk pathways. Sexual violence was included only when embedded in IPV, domestic-abuse, coercive-control, family-violence, lethality-risk or femicide-related pathways.

Results: Searches identified 4,099 records; after deduplication, 2,906 were screened, 166 reports were assessed at full text and 125 were included in the core evidence map. The evidence was heterogeneous, spanning clinical and electronic health records, police narratives, legal documents, social media or online posts, survey data, linked administrative data and survivor-facing digital tools. AI-related methods were used mainly for detection, classification, record linkage, risk stratification, text mining, triage or decision support rather than for direct evaluation of femicide-prevention interventions. Femicide, lethality and severe escalation were addressed in only part of the corpus, and few studies examined implementation, human oversight, false reassurance, fairness, privacy or downstream institutional action in depth.

Discussion: The findings do not support individual femicide prediction or demonstrate that AI prevents lethal violence. Instead, they support a more defensible role for AI as a bounded component in human-led risk-recognition pathways. The review develops a six-layer conceptual synthesis linking distributed risk signals, AI-assisted signal processing, human contextual review, multi-agency response, legal-ethical governance and medico-legal accountability. AI may support institutional recognition and coordination, but it cannot substitute for professional judgment, survivor-centred practice, due process or adequately resourced prevention systems.

KEYWORDS

artificial intelligence, decision support, digital health, femicide, human oversight, intimate partner violence, medico-legal accountability, risk assessment

1 Introduction

Violence against women, including intimate partner violence (IPV), remains a major public-health, human-rights and medico-legal concern. International and European sources describe IPV as associated with physical injury, mental-health harm, sexual and reproductive consequences, social isolation and economic vulnerability, and they document the continuing burden of gender-based violence across jurisdictions (1–4). At the most severe end of this continuum, intimate partner and family-related femicides represent gender-related killings that frequently arise in contexts of prior abuse, coercive control, threats, separation, stalking, weapon access, previous police contact, healthcare encounters or social-service involvement (5–8). These observations do not make lethal outcomes predictable in an individual deterministic sense. They do, however, raise a medico-legal question that is both narrower and more concrete: when signs of escalation are available across institutional settings, how are they recognised, documented, communicated and acted upon?

The medico-legal relevance of this problem does not lie in claiming that lethal violence can be forecast with certainty. It lies in the relationship between foreseeability, documentation, escalation and response. Femicide should be framed as a possible endpoint of pathways of violence, coercive control and institutional contact, not as a mechanically predictable event. Risk assessment and risk management guidance in the IPV field already emphasises lethality risk, repeated victimisation, escalation and safety planning without collapsing these into deterministic prediction (7, 9). Domestic homicide and death-review sources likewise point to missed opportunities, information silos, incomplete risk management and implementation failures rather than to a total absence of warning signs (10–12). For forensic and legal medicine, the central question is therefore whether danger was reasonably recognisable and whether institutions responded in a proportionate, documented and coordinated way.

Healthcare is one important part of this distributed pathway. World Health Organization (WHO) guidance emphasises survivor-centred first-line support, safe enquiry, confidentiality, referral and health-system organisation rather than actuarial certainty (13–15). Emergency departments, primary care, mental-health services, obstetric and gynaecological settings and injury-related encounters may all generate records relevant to violence, fear or repeated harm. At the same time, police calls, prior complaints, protection-order proceedings, social-service referrals, shelter contact and digital abuse reports may sit in separate institutional systems. The warning signal, if present, is therefore often fragmented. No single professional may see the full chronology, and no single dataset can safely represent the lived dynamics of coercive control.

Artificial intelligence (AI)-related methods have attracted interest because they may process unstructured records, identify patterns in text, support linkage across systems or prompt triage and review. Studies in the included corpus have applied machine learning or natural language processing (NLP) to legal documents concerning intimate partner femicide and to police narrative data (16, 17). Other work has examined clinical notes, electronic health records (EHRs) and broader institutional records relevant to IPV identification (18, 19). These methods may contribute to detection, record organisation or prioritisation of professional review. Their promise, however, is often easier to state than to implement safely. The institutional data available for modelling are incomplete, shaped by under-reporting, affected by structural inequality and sensitive in ways that raise privacy, discrimination and surveillance concerns.

That heterogeneity makes conceptual precision necessary. Prediction, risk assessment, risk stratification, early warning, decision support and prevention are not interchangeable. Prediction implies a forecast that a particular future event will occur. Risk assessment evaluates known risk factors and case context. Risk stratification assigns relative priority categories. Early warning identifies signals requiring professional review. Decision support assists reasoning but does not replace it. Prevention requires action: safety planning, referral, protection, perpetrator management, follow-up and inter-agency coordination. These distinctions matter because a low algorithmic score may create false reassurance, while a high score may prompt intrusive or discriminatory intervention if not embedded in meaningful human review. A low score should therefore be interpreted only as the absence of a model-generated alert under specified data conditions, not as evidence that danger is absent. Conversely, a high score

Abbreviations

AI, artificial intelligence; AUC, area under the curve; EDPB, European Data Protection Board; EHR, electronic health record; EU, European Union; GDPR, General Data Protection Regulation; GREVIO, Group of Experts on Action against Violence against Women and Domestic Violence; IPV, intimate partner violence; JBI, Joanna Briggs Institute; NLP, natural language processing; PCC, Population-Concept-Context; WHO, World Health Organization.

should trigger proportionate professional review rather than automatic coercive, intrusive or discriminatory action.

This review maps AI-related methods studied in IPV, domestic violence, coercive-control and femicide-related risk contexts. It asks not whether AI can predict femicide, but how AI-related methods have been used across risk-recognition pathways and what those uses imply for prevention, governance and medico-legal accountability. Existing AI/IPV literature is too heterogeneous to support a single technical conclusion, and only a subset of studies addresses femicide, intimate partner homicide or lethality directly (20, 21). This supports a scoping-review approach and a cautious conceptual synthesis. In this article, femicide-prevention implications are used in a specific and limited sense: femicide is not treated as an individual prediction endpoint, and AI is not presented as a substitute for risk assessment, clinical judgement, police decision-making, judicial oversight or social-service response. Rather, the analysis examines whether AI-related methods may support earlier institutional stages at which risk signals are recorded, missed, linked or escalated, while preserving professional responsibility and avoiding false reassurance, discriminatory surveillance and technological solutionism.

2 Methods

2.1 Study design, eligibility criteria and population-concept-context (PCC) framework

This study was conducted as a PRISMA-ScR scoping review and was informed by JBI guidance for scoping-review methodology (22, 23). A scoping approach was selected because the field is broad, methodologically heterogeneous and not yet suited to meta-analysis or a narrow effectiveness review. The review protocol was not prospectively registered because the review was exploratory and iterative, as is common in scoping work that aims to map an emerging interdisciplinary field. Nevertheless, the eligibility framework, screening process, data charting and reporting were structured according to PRISMA-ScR and JBI principles throughout. The objective was to map how AI-related methods have been studied in IPV, domestic violence, coercive-control and femicide-related risk contexts, and to examine the implications of this evidence for medico-legal and public-health prevention pathways.

Eligibility was structured using the Population-Concept-Context (PCC) framework, which is recommended in JBI scoping-review methodology for clarifying broad review questions and organising heterogeneous evidence. The population domain was defined inclusively as any person exposed to IPV, domestic violence, domestic abuse, family violence, coercive control, gender-based violence, violence against women, stalking in an IPV/domestic-abuse context, intimate partner homicide, domestic homicide, femicide or lethality risk. The review recognised that most retrieved publications focused on women and girls, but eligibility was not limited to cisgender women. Same-sex relationships, transgender relationships and LGBTQ+ populations were considered within scope where the violence occurred within an intimate-partner or domestic-abuse pathway. For this review, IPV was operationally

understood as violence, abuse, coercive control, stalking or lethality-related risk occurring in current or former intimate relationships, including current or former spouses, partners, cohabiting or non-cohabiting partners, boyfriends/girlfriends and ex-boyfriends/ex-girlfriends.

The concept domain included artificial intelligence, machine learning, deep learning, natural language processing, large language models, algorithmic or automated decision-making, predictive analytics, data mining, text mining, risk prediction, decision support, clinical decision support, risk stratification, early-warning systems, record linkage, data linkage and EHR-based identification. The context domain included healthcare, emergency care, primary care, mental health, police or law enforcement, courts or judicial settings, emergency calls, social services, shelters, digital platforms and multi-agency prevention systems. Reports were included in the empirical evidence map only when all three domains were present: a relevant violence context, an AI-related or algorithmic technology, and a function related to detection, screening, classification, risk assessment, risk stratification, triage, decision support, referral, prevention, inter-agency coordination, medico-legal accountability or legal/ethical governance. Records were excluded from the core evidence map when they addressed general violence, offender-risk prediction, child abuse, elder abuse, generic digital health or commercial tools without a clear IPV/femicide risk-pathway and AI-related component. Records addressing sexual violence were included only when sexual violence was situated within intimate partner violence, domestic abuse, coercive control, family violence, lethality risk or femicide-related pathways. Studies focused exclusively on non-partner sexual assault, sexual-violence forensics, child sexual exploitation or sexual-offence detection without an IPV/domestic-abuse risk-pathway component were outside the scope.

2.2 Information sources, search strategy and selection process

The empirical search covered biomedical, public-health, social-science, criminology and computing-oriented databases, including PubMed/MEDLINE, Scopus, Web of Science, PsycINFO, Criminal Justice Abstracts, IEEE Xplore and ACM Digital Library. Searches covered all years available in each database up to 28 April 2026 and were limited to English-language records. Search strings combined terms for IPV, domestic violence, coercive control, femicide, intimate partner homicide and gender-based violence with terms for AI, machine learning, NLP, predictive analytics, decision support, risk stratification, record linkage and related technologies. Database-specific search strategies are provided in [Supplementary Table S4](#). Searches were intentionally sensitive because preliminary testing showed that the narrower expression “AI and femicide” would miss relevant work on IPV risk pathways and institutional data sources. Reference-list and backward citation checking were also used to identify additional potentially relevant records.

After deduplication, 2,906 records were screened at title and abstract level by two reviewers using the eligibility criteria described above. Records were separated into likely

core evidence, background/contextual material, legal/ethical complementary material or exclusions. Full texts were sought for 180 reports. Full-text assessment of the 166 retrieved reports was performed primarily by two reviewers, with support from the author group. Disagreements or uncertain classifications were resolved by consensus among all authors. Fourteen reports could not be retrieved despite legal and institutional access attempts. These reports are listed separately in [Supplementary Table S5](#) on the basis of bibliographic metadata and titles only; because their full texts were unavailable, they were not assessed at full-text level and were not incorporated into the core evidence map. The final core evidence map included 125 reports/studies. Twenty-eight background/contextual sources and seven legal/ethical complementary sources were retained outside the empirical core; three duplicate or variant records were not counted separately. The PRISMA-ScR flow is summarised in [Figure 1](#).

2.3 Data charting, evidence synthesis and complementary legal-policy sources

Data charting captured bibliographic information, PCC classification, country or region where available, study type, violence focus, population, relationship status where reported, institutional setting, data source, AI or digital technology type, intended function, target outcome, validation status, performance metrics when reported, femicide/lethality relevance, coercive-control relevance, repeat-victimisation or severe-escalation relevance, and discussion of ethical or governance issues. One reviewer performed the main data charting, and all records were checked and refined with support from the author group. Governance fields included false positives, false negatives, false reassurance, algorithmic bias, discrimination, surveillance risk, privacy or data protection, explainability, human oversight, data minimisation, auditability, institutional responsibility and

medico-legal accountability. The expanded record-level PCC classification and data-charting matrix is provided in [Supplementary Table S2](#). Categories were not always mutually exclusive; for example, a single study could involve healthcare data, text mining and risk stratification. Consequently, setting and function are reported descriptively and, where counts are provided, as non-mutually-exclusive mappings rather than as mutually exclusive percentages.

A complementary grey/legal-policy corpus was used to support interpretation of governance, institutional responsibility and public-health implications. These sources included official legal instruments and institutional guidance, including General Data Protection Regulation (GDPR), the European Union (EU) AI Act, Directive (EU) 2024/1385 and the Istanbul Convention ([24–27](#)). They also included European Data Protection Board (EDPB) guidance on automated decision-making, data protection impact assessment, privacy by design and AI-model governance ([28–31](#)), together with public-health, Group of Experts on Action against Violence against Women and Domestic Violence (GREVIO) and domestic-homicide review sources relevant to risk management, coordination and institutional learning ([32–34](#)). These sources were not included in the empirical PRISMA core and did not alter the review counts. They informed the legal, ethical, public-health and medico-legal interpretation of the mapped evidence.

2.4 Conceptual synthesis development

The review did not include a formal risk-of-bias or quality appraisal because the purpose was to map the breadth, characteristics and conceptual implications of a heterogeneous evidence base rather than estimate the effectiveness of an intervention. Synthesis was descriptive and thematic. We first mapped technologies, data sources, settings and intended functions; then identified where femicide, lethality, severe escalation, governance and medico-legal issues appeared in the corpus. The initial conceptual synthesis was developed from the

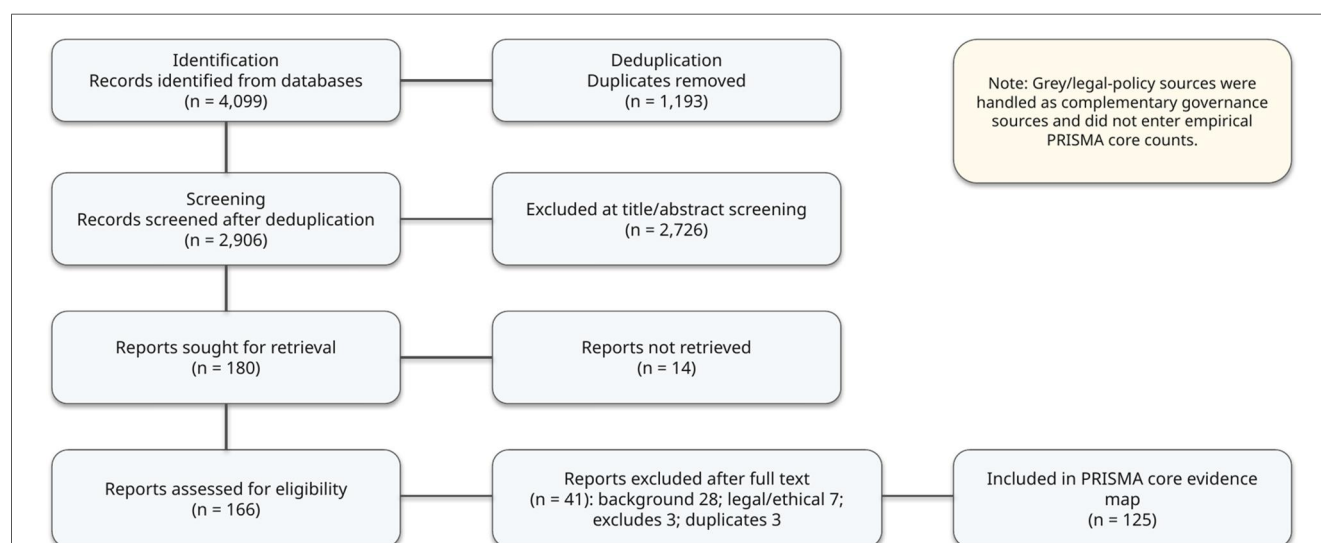


FIGURE 1

PRISMA-ScR flow diagram. The figure summarises identification, deduplication, screening, full-text retrieval and inclusion in the PRISMA core evidence map. Complementary background/contextual and legal/ethical sources are shown outside the empirical core.

charted data and then discussed and refined with all authors. The six layers were identified through thematic analysis of recurring implementation gaps, oversight requirements and accountability points across the empirical corpus, and were then mapped along a risk-recognition pathway from distributed warning signals to institutional response and post-event accountability. This synthesis was interpreted in light of the complementary legal and policy sources. It is not presented as a validated instrument, prediction model, operational checklist, clinical guideline or police tool.

3 Results

3.1 Study selection and characteristics of included records

The search identified 4,099 raw records. After deduplication, 2,906 records were screened. Full texts were sought for 180

reports, 166 were assessed and 125 were included in the PRISMA core evidence map (Figure 1). The included PRISMA-ScR core records are listed in [Supplementary Table S1](#), and the record-level PCC classification and data-charting matrix is reported in [Supplementary Table S2](#). Table 1 provides a compact summary of the main study characteristics using descriptive, non-mutually-exclusive categories. The included corpus covered a wide range of publication types, including model-development studies, validation studies, observational studies using administrative or clinical data, text-mining studies, social-media or online-community analyses, digital intervention evaluations, conference papers, reviews and conceptual or governance-oriented contributions. This breadth confirms that the field is not organised around one technology, one dataset or one outcome.

The included records focused primarily on IPV, domestic violence or family violence rather than femicide as a direct endpoint. A smaller subset addressed intimate partner femicide, intimate partner homicide, lethality risk, severe escalation or domestic homicide review-type concerns. This

TABLE 1 Study characteristics summary of the PRISMA-ScR core evidence map.

Domain	Non-mutually exclusive categories in the core evidence map	n/125 where safely coded	Interpretive note
Population/violence focus	Domestic violence/abuse; intimate partner violence; gender-based violence/violence against women; family violence; femicide/femicide; stalking; intimate partner/domestic homicide; lethality or severe escalation; coercive control	Domestic violence/abuse 81; IPV 54; gender-based violence/violence against women 27; family violence 15; femicide/femicide 9; stalking 7; intimate partner/domestic homicide 5; lethality/severe escalation 4; coercive control 4	Most records addressed IPV or domestic violence pathways rather than femicide as a direct endpoint.
AI/digital method	Machine learning; predictive analytics/modeling; deep learning; AI; NLP/text mining; EHR-based identification; data mining; record/data linkage; LLM/generative AI; decision support/risk stratification	Machine learning 73; predictive analytics/modeling 44; deep learning 37; AI 33; NLP/text mining 31; EHR/electronic health records 15; data mining 14; record/data linkage 8; LLM/generative AI 6; decision support/risk stratification 4	Methods were often combined within a single study.
Intended function	Detection/identification; prevention/intervention/referral; prediction/forecasting; risk assessment/stratification; screening; early warning/surveillance; legal/ethical governance; decision support/triage; inter-agency coordination	Detection/identification 112; prevention/intervention/referral 58; prediction/forecasting 56; risk assessment/stratification 35; screening 19; early warning/surveillance 10; legal/ethical governance 10; decision support/triage 8; inter-agency coordination 2	Functions describe intended or discussed use, not demonstrated prevention effectiveness.
Data source/context	Healthcare and EHR data; police/law-enforcement records; legal/judicial documents; online communities and social-media data; survey or population-level data; linked administrative datasets; survivor-facing digital tools	Detailed record-level coding is provided in Supplementary Table S2 .	Contexts were classified descriptively because a single record could involve multiple data sources or settings.
Femicide/lethality relevance	Femicide/femicide; intimate partner homicide/domestic homicide; lethality or severe escalation	14 records had direct femicide, homicide or lethality/severe-escalation coding.	The remaining records contribute indirectly through IPV/domestic-violence risk-pathway evidence.
Validation/implementation status	Model development/validation; model development/evaluation; clinical or implementation trial; observational/data-linkage; qualitative; review/meta-analysis; survey/cross-sectional; protocol, case study, conceptual/technical or unclear	Model development/validation 59; model development/evaluation 12; clinical/implementation trial 7; observational/data-linkage 4; qualitative 13; review/meta-analysis 11; survey/cross-sectional 5; protocol/case/conceptual/unclear 14	Technical feasibility was more common than downstream implementation or outcome evaluation.

Counts are descriptive and non-mutually exclusive except where study-design categories were coded as mutually exclusive in the supplementary evidence map. Detailed record-level PCC classification is provided in [Supplementary Table S2](#).

TABLE 2 AI functions across IPV risk pathways and femicide-prevention implications.

AI-related function	Typical data sources/settings	Pathway contribution	Key limits	Required human oversight
Detection and screening	EHRs, clinical notes, coded records, emergency or primary-care settings	May help flag possible IPV-related presentations or under-coded abuse histories for professional review.	Risk of under-detection, poor coding, confounding and false reassurance when records are incomplete.	Outputs should prompt contextual assessment, not automatic labelling or action.
NLP and text mining	Police narratives, legal documents, clinical notes, online posts	Can extract patterns, injuries, abuse types, coercive-control indicators or help-seeking signals from unstructured text.	Narratives reflect institutional language, reporting patterns and missing context; semantic accuracy does not equal safety impact.	Professionals must interpret outputs in light of chronology, survivor account and service context.
Record linkage and data linkage	Health, police, registry, administrative or multi-agency data	May reveal distributed contacts and repeated victimisation across systems.	High privacy burden; linkage errors; risks of disproportionate surveillance and unclear lawful basis.	Strict purpose limitation, data minimisation, governance review and accountable access controls.
Risk prediction and stratification	Survey data, police/clinical datasets, risk factors, prior violence indicators	May stratify cases for review or identify combinations associated with severe harm.	Model performance is not equivalent to individual femicide prediction or validated prevention.	Scores must be advisory, auditable and overridden when contextual information indicates risk.
Digital triage and decision support	Clinical workflows, service platforms, safety-planning or referral systems	May structure documentation, triage prompts, referral and follow-up.	A prompt without a response pathway is not prevention; poorly designed systems may increase workload.	Clear escalation protocols, training, referral capacity and accountability for action/inaction.
Survivor-facing digital tools	Chatbots, online support, help-seeking platforms	May support access to information or referral when safe and accessible.	Digital safety, perpetrator monitoring, confidentiality and unequal access are major concerns.	Tools must preserve autonomy, privacy and safe exit options; human support should be available.

IPV, intimate partner violence; EHR, electronic health record; NLP, natural language processing. The table summarises mapped functions and conceptual implications; it does not imply demonstrated femicide-prevention effectiveness.

distribution is important for interpretation: the evidence base is relevant to femicide-prevention implications because femicide may occur at the end of escalating IPV trajectories, but most included studies were not designed to evaluate femicide prevention directly. The corpus therefore supports a risk-pathway interpretation rather than an event-prediction interpretation.

Because scoping reviews aim to map rather than rank evidence, the included records were interpreted as a landscape of uses rather than as a hierarchy of effectiveness. Several records were methodological or model-development contributions, while others were exploratory, implementation-oriented or review-based. The field is consequently best understood as an emerging set of partial applications. Some studies show that relevant signals may be extracted from text or structured records; others show that these signals are difficult to translate into practice without careful attention to workflow, safety, confidentiality and professional responsibility. This distinction is important because a technically successful classifier does not automatically become a prevention mechanism.

3.2 Technologies, data sources and intended functions

The technologies mapped in the review included machine-learning classifiers, deep-learning approaches, NLP pipelines,

large-language-model applications, text and data mining, decision-support tools, record linkage, EHR-based screening or identification, social-media classification, chatbot-supported help-seeking and risk-prediction or risk-stratification models. Police and legal-text studies used NLP or text mining to identify abuse types, injuries, coercive-control behaviours, mental-health mentions, recorded-crime features, future offending or femicide-related case characteristics (16, 17, 35–38). Closely related work examined police-record surveillance outputs, linkage between police and health records, and NLP-assisted improvement of police-recorded domestic-abuse data (39–42). Healthcare-oriented studies used diagnostic codes, clinical notes, electronic health records, emergency-department narratives, dental or injury presentations and linked administrative records to identify IPV or related harms (18, 19, 43, 44). These studies differ substantially in purpose: some make hidden information easier to retrieve, others develop risk or classification models, and others explore digital support or help-seeking needs.

A further pattern concerned the distinction between data extraction and decision support. Some NLP and text-mining studies sought to improve the visibility of IPV-related information already present in narrative records, including police narratives, clinical notes and online disclosures (45–48). This is a relatively bounded use: the technology helps make existing institutional knowledge more searchable or structured. Other studies moved closer to risk prediction, risk stratification or prioritisation, including models for court or legal processes,

police risk assessment, recidivism and domestic-violence vulnerability (48–51). These raise more difficult questions because the output may influence how cases are ranked, reviewed or escalated. Extracting an IPV circumstance from a clinical note, classifying abuse type in a police narrative, and assigning a relative risk category to a person or case have different implications for consent, transparency, error, proportionality and reviewability. The mapped literature did not always distinguish these functions clearly. This conceptual imprecision is itself a finding of the review.

The data sources differed in their proximity to institutional action. Clinical notes and diagnostic records are relevant to screening and referral, but they may not capture coercive control, fear, threats or perpetrator access. EHR-based studies ranged from identification through diagnoses and phone calls to association-rule mining and older-women health correlates (19, 43, 44, 52). Other clinical and emergency-care studies addressed traumatic-brain-injury prescreening, hospital identification, dental or injury presentations and mental-health record classification (53–56). Additional emergency and clinical narratives further showed how much depends on documentation practices and the reason for contact (46, 47, 57, 58). Police narratives may contain more detailed event descriptions, but they reflect reporting and recording practices. Linked police-health data and primary-care or family-law linkage studies show the value of institutional linkage, while also exposing the legal and ethical difficulty of bringing separate systems together (41, 59). Social media and online forums can show help-seeking, disclosure and distress, yet they raise privacy and representativeness problems (60–62). Survivor-facing chatbots and digital platforms may reduce barriers to information, but their safety depends on usability, confidentiality, perpetrator monitoring risk and the availability of human services after disclosure (63–65). Across these sources, the review found no evidence that more data alone produces safer outcomes.

Across these technological groups, validation and interpretability were uneven. Some studies reported conventional discrimination metrics, but fewer addressed calibration, transportability across jurisdictions, or performance in settings different from the development dataset. This matters because IPV-related data are strongly shaped by local documentation practices, legal definitions, service access and reporting behaviour. Predictive-modelling studies in domestic violence and violence-against-women contexts used approaches ranging from multimodal machine learning and ensemble learning to XGBoost and feature-selection approaches (66–68). Other studies examined machine learning for violence-against-women classification, IPV measurement, domestic-violence vulnerability and prevalence or incidence estimation (69–72). Further work developed perpetrator-focused, multidimensional victimisation or feature-explanation models across different populations and jurisdictions (73–76). Such studies are important for mapping technical possibilities, but they do not remove the need for external validation, calibration and implementation testing. A model trained on police narratives in one jurisdiction may not translate to another jurisdiction with different recording rules or risk-assessment procedures. Similarly, EHR-based approaches may perform differently in emergency care, primary care, mental health or obstetric settings because the opportunity

to ask about violence, the language used by clinicians and the coding practices vary. The review therefore supports a cautious interpretation of technical performance: reported accuracy or area-under-the-curve (AUC) values should be read as evidence of feasibility within a defined dataset, not as evidence that a tool is ready for deployment in femicide-prevention pathways.

A related limitation is that validation was often separated from the conditions under which tools would be used. A model may perform adequately on a retrospective dataset and still fail when deployed in a setting where documentation is incomplete, disclosure is unsafe, staff have little time to interpret outputs, or referral options are unavailable. Some studies explicitly engaged with fairness, interpretability, feature selection or risk-factor explanation, but these concerns were not consistently integrated into workflow, survivor safety or accountability analysis (48, 49, 75). Other population-level models estimated domestic violence, family violence or IPV vulnerability across geographical contexts, including Türkiye, Liberia and Iran, and across population or situational contexts, including homelessness among young adults and COVID-19-related conditions (77–81). These records broaden the empirical map but also reinforce the need to separate model-development value from implementation readiness. The review therefore suggests that implementation evidence should be treated as part of the evidentiary standard, not as a later technical detail. In this field, the relevant question is not only whether a system identifies IPV-related information, but whether the resulting prompt can be reviewed, contextualised and acted upon without increasing risk.

3.3 Femicide-prevention implications, governance issues and evidence gaps

Only part of the corpus addressed femicide, intimate partner homicide, lethality or severe escalation directly. These studies are important because they show that AI-related methods can be applied to legal documents, homicide-related data, pregnancy-associated homicide, attempted femicide, news alert detection or multilevel risk-factor modelling (16, 82). Related femicide- or homicide-oriented records included systematic, data-mining and participatory-alert approaches (21, 83–85). Reviews and applied studies also show growing interest in machine learning for femicide detection or IPV-prevention claims, but this literature remains heterogeneous and unevenly connected to real-world safety outcomes (20, 86). The corpus also includes risk-related work on antepartum IPV, stalking and indicators in coded records, which are relevant to escalation pathways but should not be treated as femicide-prediction instruments (87–89). These studies do not establish that individual femicide events can be predicted reliably. Nor do they demonstrate that algorithmic outputs prevent lethal violence. The more consistent finding is that AI-related methods may help organise large volumes of fragmented information, identify patterns in unstructured text, support case-linkage or prompt further professional review. These are pathway functions rather than proof of prevention.

The governance content of the empirical corpus was often less developed than the modelling content. The main ethical, legal and

TABLE 3 Ethical, legal and medico-legal safeguards for AI-supported IPV risk pathways.

Issue	Why it matters in IPV/femicide-prevention pathways	Safeguard or implication
False negatives and false reassurance	A low-risk output may discourage escalation despite coercive control, threats, stalking or separation risk.	Low-risk scores should never substitute for contextual assessment or professional judgment.
False positives and intrusive intervention	Over-classification can stigmatise survivors, trigger unwanted intervention or increase surveillance burdens.	Use proportionate triage, documented review and survivor-centred response options.
Algorithmic bias and discrimination	Data reflect reporting inequalities, policing practices, healthcare access and socioeconomic vulnerability.	Assess fairness, under-protection and over-intervention across groups before deployment.
Privacy and data minimisation	IPV pathways may involve health, sexual, family, location and justice-related data.	Limit processing to necessary variables, purposes and access rights; conduct DPIA where required.
Human oversight	Automated outputs may shape institutional priority or protective responses.	Oversight must be meaningful, informed and able to challenge, override or ignore the output.
Explainability and auditability	Professionals need to understand why an alert was generated and what action followed.	Maintain interpretable outputs, logs, escalation records and reviewable decision trails.
Institutional accountability	Responsibility cannot be displaced to a model, vendor or score.	Post-event review should examine recognition, documentation, communication, escalation and action/inaction.
Legal governance	GDPR, the EU AI Act, Directive (EU) 2024/1385 and Istanbul Convention/GREVIO materials constrain use.	Treat AI as bounded support within lawful, proportionate, human-led and coordinated service pathways.
Automation bias	Professionals may over-trust or passively accept AI outputs because they appear objective or technically authoritative.	Training, workflow design and audit should require active contextual review, documentation of disagreement and the ability to override both high-risk and low-risk outputs.
Automated decisions and high-risk AI governance	Outputs may affect access to referral, shelter assessment, protection, policing priority or healthcare triage.	Avoid solely automated significant decisions; ensure lawful basis, meaningful human review, contestability and, where applicable, EU AI Act high-risk safeguards.

The safeguards are derived from the review synthesis and complementary legal/policy corpus. They are not a validated implementation protocol.

medico-legal safeguards for human-governed AI-supported IPV risk pathways are summarised in Table 3. Many studies reported technical performance but said less about how an output should be interpreted, who should review it, how disagreement should be handled, what level of uncertainty is acceptable, or how survivors' autonomy and safety should be protected. Few studies examined how professionals respond to AI prompts in real settings, whether alerts are ignored, whether low-risk outputs reduce vigilance, or whether high-risk outputs produce intrusive action. False positives and false negatives were often implicit rather than analysed as institutional risks. False negatives may create reassurance where contextual assessment should continue, while false positives may expose survivors or families to unnecessary intrusion, stigma or coercive interventions. Bias and fairness were inconsistently examined, despite the likelihood that IPV reporting, healthcare access, policing, language, migration status, disability, socioeconomic position and race or ethnicity may influence recorded data. Studies of police risk assessment, interpretable recidivism modelling and disengagement from legal proceedings illustrate that fairness, explanation and procedural consequences are not optional add-ons when algorithmic outputs may affect institutional action (48–50).

The evidence gaps were especially pronounced for outcome evaluation. Very few records connected AI-supported identification, triage or help-seeking tools to measurable changes in referral, safety planning, protection, perpetrator

management, follow-up, survivor-reported outcomes, severe escalation or lethal harm. Survivor-facing digital studies, chatbot evaluations, large-language-model applications and GraphRAG studies shift attention beyond model performance, but they remain early applications rather than validated response pathways (45, 60, 63–65, 90, 91). The corpus therefore cannot support claims that AI prevents femicide. It shows instead where AI-related methods may be studied more responsibly: recognition of distributed signals, structuring of unstructured records, linkage of repeated contacts, prioritisation for human review, documentation of reasoning and auditability of institutional response. For a medico-legal and public-health readership, the key distinction is that prevention is not achieved when a record is classified; it begins only when classification is translated into safe enquiry, professional review, referral, protection, follow-up or coordinated action.

4 Conceptual synthesis: from AI tools to accountable risk pathways

The included corpus does not point toward a single dominant technology, outcome definition or implementation pathway. Instead, it shows a fragmented field in which AI-related methods are used for detection, classification, linkage, triage, decision support or risk stratification across different

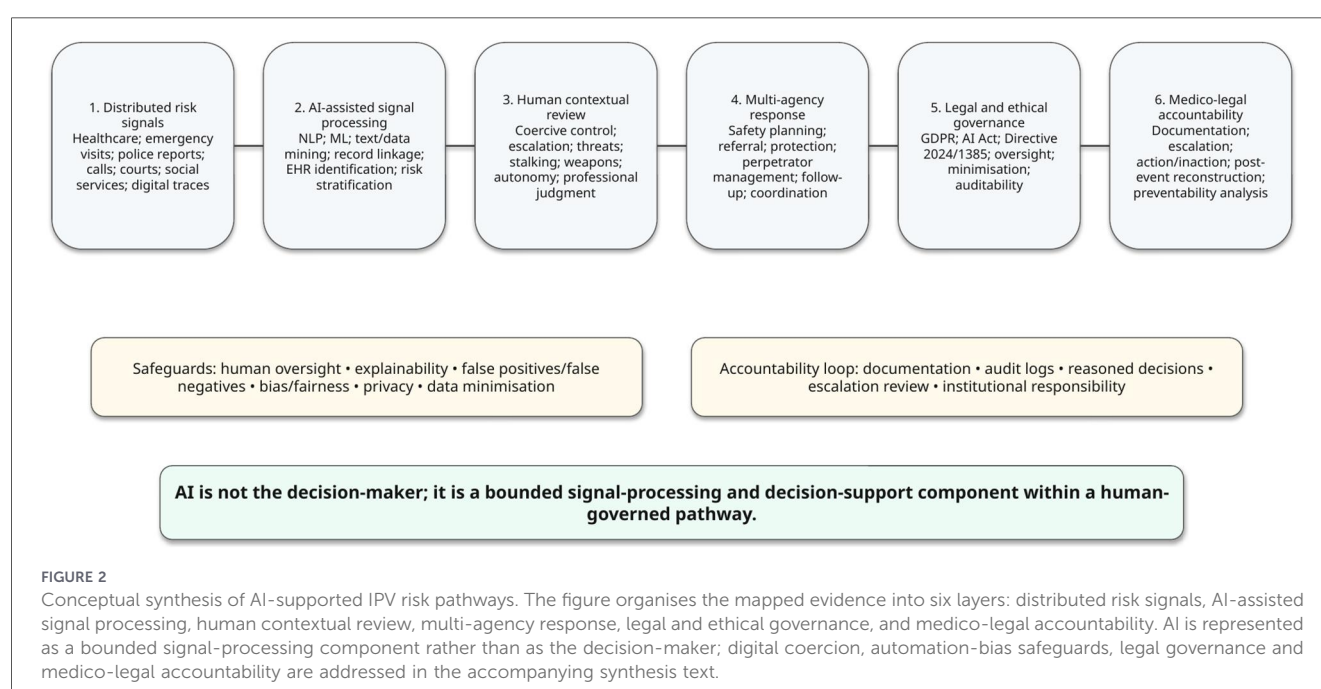
institutional contexts. A conceptual synthesis is therefore useful not to create a linear pipeline or a new risk score, but to locate these methods within the wider sequence of risk recognition, documentation, escalation, response and accountability. This is particularly important because IPV cases are dynamic and institutional information may be partial, delayed or contradictory. A survivor may disclose in one setting and deny abuse in another because of fear, dependence, migration status, disability, children, financial control, digital coercion, spyware, perpetrator monitoring or the risk that the perpetrator will discover the disclosure or associated digital traces. A police record may describe an incident without capturing the longer pattern of coercion, while a healthcare record may document injuries without recording the relationship context. Distributed signals are therefore fragments requiring contextual interpretation rather than data points that automatically produce a decision.

We interpret the evidence through a six-layer conceptual synthesis of AI-supported IPV risk pathways. This synthesis is not a validated intervention, risk score, clinical guideline, police tool or operational checklist. Its purpose is to organise the mapped evidence and clarify where AI-related methods may assist, and where human interpretation, institutional responsibility and legal safeguards remain decisive. The first layer consists of distributed risk signals. These may appear in healthcare encounters, emergency visits, police reports, emergency calls, judicial documents, social-service contacts, shelter referrals or digital contexts. Such signals may include repeated injuries, fear, threats, stalking, coercive control, separation, pregnancy-related vulnerability, child-related risk, weapon access, repeated calls or fragmented prior complaints. The scoping review suggests that AI methods may be relevant because many of these signals are embedded in unstructured or dispersed records rather than in one clean dataset. The six-layer synthesis is summarised in Figure 2.

The second layer is AI-assisted signal processing. This is the only layer in which algorithmic methods are centrally located. NLP, machine learning, text mining, data linkage, EHR-based identification and risk stratification may help process large volumes of information or prompt review. Yet this layer remains bounded. It does not determine what happened, what risk means in context, or what response is proportionate. Studies using police narratives, EHRs, linked datasets, social-media posts and survivor-facing digital tools show the range of potential signal-processing functions (36, 40, 59, 92). Other model-development studies show that risk stratification can be technically explored across populations and contexts, but they also demonstrate why a model output cannot be equated with a preventive decision (69–72).

The third layer is human contextual review. At this stage, algorithmic outputs must be interpreted against chronology, survivor voice, coercive-control context, missing information, professional knowledge and proportionality. Human review is therefore not a procedural formality after signal processing, but the point at which uncertainty, context and possible action are assessed. This layer must also guard against automation bias: without training and institutional safeguards, professionals may over-trust low-risk or high-risk outputs because they appear objective, technical or authoritative.

The fourth layer is multi-agency response. Risk recognition has preventive value only if it connects to action: safe enquiry, referral, safety planning, protection measures, perpetrator management, follow-up and coordination across health, police, courts and social services. A signal without a response pathway is not prevention. This layer is supported more by public-health, risk-management and legal-policy sources than by AI-performance studies, because it concerns organisational capacity, not model accuracy (26, 27). Public-health and monitoring sources likewise emphasise implementation,



coordination and risk management as institutional duties rather than algorithmic outputs (32–34). It is also where survivor autonomy and digital safety become central. A tool that prompts disclosure without safe referral, confidentiality or follow-up may increase documentation without increasing protection.

The fifth layer is legal and ethical governance: lawfulness, transparency, data minimisation, privacy, explainability, fairness, auditability and safeguards against discriminatory or disproportionate use define the conditions under which AI-supported pathways may be acceptable. The sixth layer is medico-legal accountability. This layer concerns how warning signs, AI-generated alerts, missed alerts, suppressed or overridden prompts, uncertainty, missing information, professional reasoning, escalation decisions and institutional responses are documented in a way that supports continuity, institutional learning and post-event reconstruction. It is the layer at which AI methods become relevant not as accountable actors, but as traceable elements within a reviewable human and institutional decision pathway.

This synthesis deliberately resists technological solutionism. AI is only one bounded component within a pathway between distributed warning signs and human, multi-agency, legally accountable action. The conceptual value of the synthesis lies in separating signal processing from contextual judgment, judgment from response, and response from accountability. It also requires any AI-supported proposal to specify which layer it addresses and which layers remain outside the model. A text-mining tool may help identify abuse-related information, but it does not decide whether a survivor is safe. A linkage model may show repeated contacts across systems, but it does not create a multi-agency response. A chatbot may provide information or classify help-seeking needs, but it cannot replace confidential advocacy or emergency protection (63, 64, 90). A predictive classifier may rank cases, but it cannot determine whether an intervention is lawful, proportionate or acceptable to the survivor (49, 51). The same framework disciplines the language of prevention: most AI-related work in the included corpus belongs, at most, to secondary or tertiary prevention, supporting recognition of existing or escalating violence, but it does not address the structural determinants of gender-based violence and does not by itself provide shelter, legal protection, economic support or perpetrator accountability. AI-supported outputs must therefore remain contestable, interpretable and subordinate to professional reasoning, particularly in coercive-control contexts where visible injury may be absent and disclosure may be constrained by fear or surveillance.

5 Discussion

This scoping review maps a growing but heterogeneous body of literature on AI-related methods in IPV and domestic-violence risk contexts. The principal finding is not that AI can predict femicide, nor that AI-supported tools have been shown to prevent lethal violence. Rather, the evidence shows that AI-related methods are being applied to parts of the institutional ecology in which IPV risk is recorded: clinical notes, EHRs, police narratives, legal documents, linked administrative data,

social media and digital help-seeking environments. These applications are relevant because femicide-prevention work often depends on earlier recognition of escalation, but they remain several steps removed from demonstrating prevention. The strongest contribution of the field is therefore organisational rather than deterministic: AI may help institutions process, link or prioritise information that would otherwise remain fragmented.

A useful way to interpret these findings is to separate epistemic value from preventive value. AI may have epistemic value when it helps reveal patterns in records that are otherwise difficult to review at scale. This is particularly relevant for narrative police records, EHR notes and linked administrative data. Preventive value requires more. It depends on whether the signal reaches a trained person, whether that person has time and authority to act, whether services are available, whether survivor safety is protected, and whether the action taken is proportionate. Much of the existing literature stops at the epistemic level. It asks whether a model can detect, classify or stratify. The harder question is whether such detection changes what institutions do next.

Individual femicide prediction is a scientifically fragile and ethically dangerous frame because femicide is a low-base-rate, context-dependent and socially embedded outcome. A model trained on partial institutional data may appear technically impressive while still being vulnerable to under-reporting, missingness, selection bias and proxy discrimination. More importantly, even a statistically plausible risk output does not answer the practical questions on which prevention depends: who reviews the case, what information is missing, how survivor autonomy is protected, what response is available, and how the perpetrator's behaviour is managed. Risk-pathway support is more defensible because it asks AI to assist with limited tasks—such as flagging records for review, extracting IPV-related circumstances from text, linking repeated contacts or supporting triage—while leaving contextual interpretation and action to accountable human systems.

False reassurance is a particularly important medico-legal risk. A low or absent risk flag may be misread as absence of danger, especially in busy clinical, police or social-service environments. This risk is not merely technical; it concerns institutional behaviour. If professionals rely on a low-risk output despite threats, stalking, strangulation, separation, weapon access, repeated contacts or survivor fear, the problem is not only model error but also a failure of contextual judgment. False positives also matter. Over-identification can expose survivors to unwanted interventions, unnecessary surveillance, loss of trust, or escalation by perpetrators if digital or institutional actions are not handled safely. A responsible system must therefore treat both false negatives and false positives as safety and rights issues, not as abstract performance metrics.

The analysis also cautions against a victim-only risk lens. Femicide prevention cannot be reduced to identifying high-risk victims. It must also address perpetrator behaviour, coercive control, threats, stalking, weapon access, breaches of orders, separation dynamics and institutional capacity to intervene. Tools that focus only on the victim's records may miss the relational and behavioural nature of danger. This has medico-legal significance because preventability is rarely a property of

one data point. It concerns the chronology of warning signs, the behaviour of the perpetrator, the response of institutions and the availability of protection. Studies of perpetrator-focused prediction, IPV perpetration among homeless young adults, stalking and legal-process disengagement illustrate why the unit of analysis should often be the pathway or case context rather than the victim alone (50, 74, 78, 88). AI-supported approaches should therefore be evaluated not only for their ability to classify victim vulnerability, but also for their capacity to support broader case understanding, perpetrator management and inter-agency coordination.

Bias, surveillance and discrimination are similarly central. IPV data are not neutral. They reflect who reports, who is believed, who accesses healthcare, how police narratives are written, which injuries are coded, and which communities are over- or under-policed. Algorithmic systems trained on such data may reproduce under-protection of some groups and over-intervention in others. The use of digital traces, online posts or linked institutional datasets may further increase the risk of disproportionate monitoring, particularly for migrants, racialised women, disabled women, LGBTQ+ survivors or those economically dependent on perpetrators. Social-media and online-community studies are useful for detecting expressions of distress or help-seeking, but they also make clear that digital data are context-dependent, unevenly representative and ethically sensitive (93, 94). Deep-learning studies using online posts, Facebook-based datasets, Persian-language social-media content and other online-text corpora extend this map but do not resolve the underlying representativeness and privacy concerns (95–97). These risks do not mean that AI-related methods should never be studied; they mean that any implementation must be narrow, transparent, proportionate and accountable.

The legal and policy corpus reinforces this conclusion. GDPR principles limit indiscriminate data aggregation and require lawfulness, fairness, transparency, purpose limitation, minimisation, accuracy and accountability (24). Special-category data and high-risk processing raise additional safeguards (24, 29, 30). Depending on design and effect, automated decision-making safeguards may be relevant, particularly if outputs influence access to protection, policing priority or institutional action (28). The EU AI Act adds a risk-based governance frame, including human oversight and, where applicable, fundamental-rights impact assessment logic (25). EDPB guidance on AI models and data protection by design further supports a cautious approach to lawfulness, fairness, minimisation and downstream use (30, 31). Directive (EU) 2024/1385, the Istanbul Convention and GREVIO materials shift the focus toward prevention, protection, support, training, coordination and risk management rather than technological prediction alone (26, 27, 33, 34). These sources do not prohibit every AI-supported pathway, but they make clear that deployment in this field would require strict governance and cannot be justified by general claims about innovation or efficiency.

A concrete example illustrates why these safeguards matter. If a hospital AI-supported triage system classified a survivor as “low risk” and that output automatically reduced access to specialist IPV referral, shelter assessment, protection planning or professional review, the issue would no longer be mere technical

prioritisation. It would concern automated decision-making with potentially legal or similarly significant effects, requiring a clear lawful basis, meaningful human intervention, contestability, documentation and safeguards against false reassurance. GDPR Article 22 may be relevant where decisions are based solely on automated processing and produce legal or similarly significant effects, while broader GDPR principles remain relevant even when a human remains in the loop (24, 28). Depending on intended purpose, use context and effect, some law-enforcement triage, victim-risk assessment, emergency-call prioritisation or healthcare triage systems may also fall within the EU AI Act’s high-risk governance logic, including human oversight and fundamental-rights safeguards where applicable (25). Not every IPV-related AI system will automatically fall into these categories, but systems that shape access to protection, policing priority or medical prioritisation require particularly strict justification and oversight.

A further implication concerns the relationship between digital health and legal accountability. Digital systems can create records of alerts, overrides, missing information and decisions. Properly governed, such logs may support continuity and institutional learning. Poorly governed, they may create defensive documentation, automation bias or confusion about who is responsible. In medico-legal reconstruction, an ignored alert may be relevant, but so may blind reliance on a reassuring output. Studies of linked police-health data, clinical-record mining and text-mined police narratives show how digital traces can make institutional knowledge more visible, but visibility is not the same as accountability (39, 41, 42, 54). Accountability should therefore be understood as a design requirement, not only as a post-event process. Systems should be built so that professionals can explain why a signal was generated, how it was reviewed, what contextual information was considered, and why a particular action was or was not taken.

The medico-legal contribution of this review lies in reframing foreseeability. Foreseeability is not deterministic prediction. In post-event reconstruction, the relevant questions are whether warning signs were available, how they were documented, whether information was communicated, whether escalation was recognised, whether actions were proportionate and whether institutional procedures were followed. AI-generated alerts, logs or risk flags may become evidentially relevant in this reconstruction, but they cannot become the accountable actor. Responsibility remains with institutions and professionals who design, deploy, interpret and act—or fail to act—within the pathway. This is why the conceptual synthesis links prevention to accountability: an early-warning signal is meaningful only when it is embedded in a documented, reviewable and resourced response system.

Future research should move beyond model performance alone and test whether AI-supported prompts can be understood, governed and acted upon in real service pathways. Studies need external validation, fairness assessment, calibration, human-factors evaluation, survivor acceptability, digital-safety analysis and implementation research. Outcome evaluation should examine whether AI-supported tools change professional behaviour, referral, case conferencing, safety planning, perpetrator management or continuity of care, while also measuring false reassurance, intrusive false positives, privacy harms,

discrimination, workload, interpretability and institutional response (67, 70, 75, 77). For digital health, the implication is that this field should not be framed as algorithmic prediction of rare lethal events, but as responsible infrastructure for safer, more consistent, better documented and less fragmented human action. Work on IPV surveillance, prevalence estimation and coded-record indicators shows that measurement and identification remain necessary research tasks, but these tasks must be connected to service pathways if they are to have preventive significance (73, 80, 89). Without these conditions, AI may increase documentation without improving safety, or produce alerts that institutions cannot meaningfully act upon.

The policy implications are consequently modest but important. Institutions considering AI-supported IPV tools should begin with the response pathway, not with the model. They should define what problem the tool addresses, who will review its outputs, what safeguards apply, what actions are available, how survivor autonomy will be protected and how errors will be audited. They should also avoid introducing systems that merely shift responsibility to frontline workers without providing resources, training or service capacity. In this field, the ethical test of a digital system is not whether it can produce a score, but whether it contributes to safer, more accountable and less fragmented care and protection. Integrative and systematic reviews of AI in IPV prevention already point to similar concerns about heterogeneity, limited implementation evidence and the need for human oversight; the present review extends this by foregrounding femicide-prevention implications and medico-legal accountability (20, 86).

6 Strengths and limitations

This review has several strengths. It maps an interdisciplinary evidence base across health, digital health, criminology, policing, legal-document analysis, online support and AI governance. It uses a transparent PRISMA-ScR process, retains all included core studies in [Supplementary Material](#), separates empirical evidence from complementary legal and policy sources, and adds a record-level PCC classification matrix to improve traceability of the corpus. It also brings a medico-legal perspective to a field often framed primarily as technical prediction or digital innovation. By citing representative studies across clusters while retaining the complete set of 125 records in [Supplementary Tables S1, S2](#), the review avoids both under-documenting the corpus and turning the article into a catalogue.

The limitations are equally important. The included records were highly heterogeneous in design, data source, outcome definition and technical maturity, precluding meta-analysis. The predominance of studies focused on IPV or domestic violence generally, rather than femicide, lethality or severe escalation specifically, should be interpreted as a characteristic of the corpus with important inferential limits rather than as evidence that the review can support femicide-event prediction. The review supports a risk-pathway interpretation: femicide-prevention implications are derived from literature on recognition, documentation,

linkage and escalation of IPV-related signals, not from direct evidence that AI predicts or prevents lethal outcomes. The review also does not map the broader AI literature on sexual violence, sexual-assault forensics or non-partner sexual victimisation. Its focus is narrower: AI-related methods in IPV and domestic-abuse risk pathways, where sexual violence may be one component of coercive, escalating or lethality-relevant abuse. The review relies on published and retrievable records. Fourteen reports could not be retrieved despite legal and institutional access attempts and are listed in [Supplementary Table S5](#) using metadata-level information only. Because these full texts were unavailable, their detailed PCC, methodological, validation and implementation characteristics could not be assessed; their non-retrieval may have introduced retrieval bias if unavailable reports differed systematically from retrieved records in publication type, technical detail, local implementation context or reporting completeness. Complementary legal and policy sources helped contextualise governance but do not constitute empirical evidence of AI effectiveness. Finally, the conceptual synthesis has not been validated as an intervention, risk-assessment tool or operational protocol. It should be treated as an interpretive structure for future research and policy discussion. Terms such as prediction, risk assessment, risk stratification and prevention were used inconsistently across the literature, so the synthesis imposed conceptual distinctions that improve interpretability but may not always match the terminology used by the original authors. Not every included study is discussed individually in the main text because the review synthesises a heterogeneous scoping corpus through representative clusters; all 125 records remain traceable in [Supplementary Tables S1, S2](#).

7 Conclusion

AI-related methods are increasingly being studied across IPV and domestic-violence risk contexts, but the evidence does not support the prediction of individual femicide events or demonstrate that AI prevents lethal violence. The more defensible contribution lies in bounded support for recognition, linkage, documentation and escalation of risk signals within human-led institutional pathways. Such support is valuable only when it is connected to survivor-centred practice, professional judgment, multi-agency response, legal safeguards and accountable documentation. A conceptual synthesis of these pathways can help clarify where AI may assist and where medico-legal accountability must remain with institutions and professionals. In femicide-prevention contexts, a signal without a lawful, proportionate and resourced response is not prevention.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#), further inquiries can be directed to the corresponding author.

Author contributions

PB: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. GN: Conceptualization, Supervision, Writing – review & editing. TS: Formal analysis, Writing – review & editing. FG: Formal analysis, Writing – review & editing. GR: Supervision, Writing – review & editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author GR declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. The authors used ChatGPT

(OpenAI) only for language refinement, editorial checking, formatting support and preparation of submission materials. No generative AI tool was used to generate data, perform the scoping-review search, screen records, extract or chart data, analyse results, or draw conclusions. All content was reviewed, corrected and approved by the authors, who remain fully responsible for the manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fdgth.2026.1877651/full#supplementary-material>

References

- World Health Organization. Violence against women fact sheet. (2024). Available online at: <https://www.who.int/news-room/fact-sheets/detail/violence-against-women> (Accessed May 4, 2026).
- World Health Organization. Violence against women prevalence estimates, 2023. (2025). Available online at: <https://www.who.int/publications/i/item/9789240116962> (Accessed May 4, 2026).
- FRA/Eurostat/EIGE. EU gender-based violence survey: key results. (2024). Available online at: <https://fra.europa.eu/en/publication/2024/eu-gender-violence-survey-key-results> (Accessed May 4, 2026).
- FRA/EIGE/Eurostat. EU gender-based violence survey: evidence for policy and practice. (2026). Available online at: <https://fra.europa.eu/en/publication/2026/eu-gender-based-violence-survey-evidence> (Accessed May 4, 2026).
- UN Women/UNODC. Femicides in 2024: Global estimates of intimate partner/family member femicides. (2025). Available online at: <https://www.unwomen.org/sites/default/files/2025-11/femicides-in-2024-global-estimates-of-intimate-partner-family-member-femicides-en.pdf> (Accessed May 4, 2026).
- UN Women/UNODC. Statistical framework for measuring the gender-related killings of women and girls (femicide/feminicide). (2022). Available online at: <https://data.unwomen.org/publications/statistical-framework-measuring-gender-related-killings-women-and-girls-also-referred> (Accessed May 4, 2026).
- EIGE. A guide to risk assessment and risk management of intimate partner violence in the EU. (2019). Available online at: https://eige.europa.eu/sites/default/files/documents/20191701_mh0119277enn.pdf.pdf (Accessed May 4, 2026).
- EIGE. Combating coercive control and psychological violence against women in the EU Member States. (2022). Available online at: <https://eige.europa.eu/publications-resources/publications/combating-coercive-control-and-psychological-violence-against-women-eu-member-states> (Accessed May 4, 2026).
- EIGE. Risk assessment and risk management toolkit pages. (2019). Available online at: https://eige.europa.eu/gender-based-violence/risk-assessment-risk-management/step-1-define-purpose-and-objectives-police-risk-assessment?language_content_entity=en (Accessed May 4, 2026).
- UK Home Office. Key findings from analysis of domestic homicide reviews. (2024). Available online at: <https://www.gov.uk/government/publications/key-findings-from-analysis-of-domestic-homicide-reviews/key-findings-from-analysis-of-domestic-homicide-reviews> (Accessed May 4, 2026).
- Domestic Abuse Commissioner. Learning from Loss: Ensuring the lessons from domestic homicide reviews lead to change. (2025). Available online at: <https://domesticabusecommissioner.uk/wp-content/uploads/2025/07/Learning-from-Loss.pdf> (Accessed May 4, 2026).
- Office of the Chief Coroner for Ontario. Domestic violence death review committee: 2019–2020 annual report. (2024). Available online at: <https://www.ontario.ca/document/domestic-violence-death-review-committee-2019-2020-annual-report> (Accessed May 4, 2026).
- World Health Organization. Responding to intimate partner violence and sexual violence against women: WHO clinical and policy guidelines. (2013). Available online at: <https://www.who.int/publications/i/item/9789241548595> (Accessed May 4, 2026).
- World Health Organization. Health care for women subjected to intimate partner violence or sexual violence: a clinical handbook. (2014). Available online at: <https://www.who.int/publications/i/item/WHO-RHR-14.26> (Accessed May 4, 2026).
- World Health Organization. Strengthening health systems to respond to women subjected to intimate partner violence or sexual violence: a manual for health managers. (2017). Available online at: <https://www.who.int/publications/i/item/9789241513005> (Accessed May 4, 2026).
- Garcia-Vergara E, Almeda N, Fernández-Navarro F, Becerra-Alonso D. Artificial intelligence extracts key insights from legal documents to predict intimate partner femicide. *Sci Rep.* (2023) 13(1):18212. doi: 10.1038/s41598-023-45157-5
- Karystianis G, Adily A, Schofield PW, Wand H, Lukmanjaya W, Buchan I, et al. Surveillance of domestic violence using text mining outputs from Australian police records. *Front Psychiatry.* (2021) 12:787792. doi: 10.3389/fpsy.2021.787792
- Lenert L, Rheingold AA, Simpson KN, Scherbakov D, Aiken M, Hahn C, et al. Electronic health record-based screening for intimate partner violence: a cluster

- randomized clinical trial. *JAMA Netw Open*. (2024) 7(8):e2425070. doi: 10.1001/jamanetworkopen.2024.25070
19. Tabaie A, Wyand SA, Cao F, Losonczy L, Bennett SS, Blackall LM, et al. Identifying latent patterns of intimate partner violence using electronic health records and association rule mining. *AMIA Annu Symp Proc*. (2024) 2024:1229–38.
20. Li Y, Emezue C, Song C, Udmuangpia T. The role of artificial intelligence for intimate partner violence prevention: a systematic review. *J Clin Nurs*. (2026): 1–25. doi: 10.1111/jocn.70275
21. da Silva BBA, de Oliveira Rodrigues CM, Filho WFT, de Araújo Fagundes RA. Application of machine learning in the detection and prevention of femicide: a systematic review. *Lect Not Netw Syst*. (2026) 1751:127–39. doi: 10.1007/978-3-032-12882-9_11
22. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D. PRISMA Extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. (2018) 169(7):467–73. doi: 10.7326/M18-0850
23. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Scoping reviews. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. *JBIM Manual for Evidence Synthesis*. Adelaide: JBI (2024). doi: 10.46658/JBIMES-24-09
24. European Union. Regulation (EU) 2016/679 (GDPR). (2016). Available online at: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32016R0679> (Accessed May 4, 2026).
25. European Union. Regulation (EU) 2024/1689 (EU AI Act). (2024). Available online at: <https://eur-lex.europa.eu/legal-content/EN-FR/TXT/?from=EN&uri=CELEX%3A32024R1689> (Accessed May 4, 2026).
26. European Union. Directive (EU) 2024/1385 on combating violence against women and domestic violence. (2024). Available online at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX%3A32024L1385> (Accessed May 4, 2026).
27. Council of Europe. Convention on preventing and combating violence against women and domestic violence (Istanbul Convention, CETS No. 210). (2011). Available online at: <https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/210> (Accessed May 4, 2026).
28. European Data Protection Board. Automated decision-making and profiling. (2018). Available online at: https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/automated-decision-making-and-profiling_en (Accessed May 4, 2026).
29. European Data Protection Board. Data protection impact assessments: high-risk processing. (2018). Available online at: https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/data-protection-impact-assessments-high-risk-processing_en (Accessed May 4, 2026).
30. European Data Protection Board. Guidelines 4/2019 on article 25 data protection by design and by default. (2020). Available online at: https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-42019-article-25-data-protection-design-and_en (Accessed May 4, 2026).
31. European Data Protection Board. Opinion 28/2024 on certain data protection aspects related to AI models. (2024). Available online at: https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_en (Accessed May 4, 2026).
32. WHO Regional Office for Europe. Care, courage, change: health-sector leadership in tackling violence against women and girls. (2025). Available online at: <https://www.who.int/europe/publications/i/item/WHO-EURO-2025-12846-52620-81536> (Accessed May 4, 2026).
33. GREVIO. 5 general report on GREVIO's activities. (2025). Available online at: <https://edoc.coe.int/en/violence-against-women/12076-5th-general-report-on-grevio-activities.html> (Accessed May 4, 2026).
34. GREVIO/Council of Europe. Thematic perspectives on the implementation of the Istanbul Convention. (2025). Available online at: <https://edoc.coe.int/en/violence-against-women/12132-thematic-perspectives-on-the-implementation-of-the-istanbul-convention.html> (Accessed May 4, 2026).
35. Karystianis G, Adily A, Schofield PW, Greenberg D, Jorm L, Nenadic G, et al. Automated analysis of domestic violence police reports to explore abuse types and victim injuries: text mining study. *J Med Internet Res*. (2019) 21(3):e13067. doi: 10.2196/13067
36. Karystianis G, Chowdhury N, Sheridan L, Reutens S, Wade S, Allnutt S, et al. Text mining domestic violence police narratives to identify behaviours linked to coercive control. *Crime Sci*. (2024) 13(1):2. doi: 10.1186/s40163-024-00200-2
37. Karystianis G, Adily A, Schofield P, Knight L, Galdon C, Greenberg D, et al. Automatic extraction of mental health disorders from domestic violence police narratives: text mining study. *J Med Internet Res*. (2018) 20(9):e11548. doi: 10.2196/11548
38. Karystianis G, Simpson A, Adily A, Schofield P, Greenberg D, Wand H, et al. Prevalence of mental illnesses in domestic violence police records: text mining study. *J Med Internet Res*. (2020) 22(12):e23725. doi: 10.2196/23725
39. Karystianis G, Cabral RC, Han SC, Poon J, Butler T. Utilizing text mining, data linkage and deep learning in police and health records to predict future offenses in family and domestic violence. *Front Digit Health*. (2021) 3:602683. doi: 10.3389/fdgth.2021.602683
40. Cook D, Weir R, Humphreys L. Improving police recorded crime data for domestic violence and abuse through natural language processing. *Front Sociol*. (2025) 10:1686632. doi: 10.3389/fsoc.2025.1686632
41. Karystianis G, Cabral RC, Adily A, Lukmanjaya W, Schofield P, Buchan I, et al. Mental illness concordance between hospital clinical records and mentions in domestic violence police narratives: data linkage study. *JMIR Form Res*. (2022) 6(10):e39373. doi: 10.2196/39373
42. Adily A, Karystianis G, Butler T. Text mining police narratives to identify types of abuse and victim injuries in family and domestic violence events. *Trends Issues Crime Crim Justice*. (2021) 630:1–12. doi: 10.52922/ti04923
43. Bhargava R, Temkin TL, Fireman BH, Eaton A, McCaw BR, Kotz KJ, et al. A predictive model to help identify intimate partner violence based on diagnoses and phone calls. *Am J Prev Med*. (2011) 41(2):129–35. doi: 10.1016/j.amepre.2011.04.005
44. Yilmaz S, Gunay E, Lee DH, Whiting K, Silver K, Koyutürk M, et al. Adverse health correlates of intimate partner violence against older women: mining electronic health records. *PLoS One*. (2023) 18(3):e0281863. doi: 10.1371/journal.pone.0281863
45. Guan S, Hui V, Stiglic G, Constantino RE, Lee YJ, Wong AKC. Classifying the information needs of survivors of domestic violence in online health communities using large language models: prediction model development and evaluation study. *J Med Internet Res*. (2025) 27(1):e65397. doi: 10.2196/65397
46. Tabaie A, Zeidan AJ, Evans DP, Smith RN, Kamaleswaran R. A novel technique to identify intimate partner violence in a hospital setting. *West J Emerg Med*. (2022) 23(5):781–8. doi: 10.5811/westjem.2022.7.56726
47. Botelle R, Bhavsar V, Kadra-Scalzo G, Mascio A, Williams MV, Roberts A, et al. Can natural language processing models extract and classify instances of interpersonal violence in mental healthcare electronic records: an applied evaluative study. *BMJ Open*. (2022) 12(2):e052911. doi: 10.1136/bmjopen-2021-052911
48. Ögüztüzün Ç, Koyutürk M, Karakurt G. Interpretable machine learning to identify risk factors for recidivism in intimate partner violence. *AMIA Annu Symp Proc*. (2024) 2024:1285–94.
49. Turner E, Brown G, Medina-Ariza J. Predicting domestic abuse (fairly) and police risk assessment; (la predicción (equitativa) de la violencia doméstica y la evaluación policial de riesgo). *Psychosoc Interv*. (2022) 31(3):145–57. doi: 10.5093/pi2022a11
50. Escobar-Linero E, García-Jiménez M, Trigo-Sánchez ME, Cala-Carrillo MJ, Sevillano JL, Domínguez-Morales M. Using machine learning-based systems to help predict disengagement from the legal proceedings by women victims of intimate partner violence in Spain. *PLoS One*. (2023) 18(6):e0276032. doi: 10.1371/journal.pone.0276032
51. Berk RA, Sorenson SB, Barnes G. Forecasting domestic violence: a machine learning approach to help inform arraignment decisions. *J Empirical Legal Stud*. (2016) 13(1):94–115. doi: 10.1111/jels.12098
52. Johnson RD, Griffiths LJ, Cowley LE, Broadhurst K, Bailey R. Risk factors associated with primary care-reported domestic violence for women involved in family law care proceedings: data linkage observational study. *J Med Internet Res*. (2023) 25:e42375. doi: 10.2196/42375
53. Liu LY, Bush WS, Koyutürk M, Karakurt G. Interplay between traumatic brain injury and intimate partner violence: data driven analysis utilizing electronic health records. *BMC Womens Health*. (2020) 20(1):269. doi: 10.1186/s12905-020-01104-4
54. Karakurt G, Patel V, Whiting K, Koyutürk M. Mining electronic health records data: domestic violence and adverse health effects. *J Fam Violence*. (2017) 32(1):79–87. doi: 10.1007/s10896-016-9872-5
55. Halpern LR, Dodson TB. A predictive model to identify women with injuries related to intimate partner violence. *J Am Dent Assoc*. (2006) 137(5):604–9. doi: 10.14219/jada.archive.2006.0255
56. Sprecher AG, Muellemann RL, Wadman MC. A neural network model analysis to identify victims of intimate partner violence. *Am J Emerg Med*. (2004) 22(2):87–9. doi: 10.1016/j.ajem.2003.12.005
57. Sachdeva AS, Bell A, Furst DJ, Kozlowski DA, Crabtree-Nelson S, Raicu D. A traumatic brain injury prescreening tool for intimate partner violence patients using initial clinical reports and machine learning. *AMIA Jt Summits Transl Sci Proc*. (2024) 2024:401–8.
58. Barboza-Salerno GE, Watson-Grace AC, Harrington TR, Shockley McCarthy KJ, Yang H, Alhajjaj Z, et al. Sex-based differences in injury patterns and hospitalization from emergency department narratives of intimate partner violence, United States, 2013–2024. *Am J Emerg Med*. (2026) 103:181–9. doi: 10.1016/j.ajem.2026.02.008
59. Kennedy N, Win TL, Bandyopadhyay A, Kennedy J, Rowe B, McInerney C, et al. Insights from linking police domestic abuse data and health data in south Wales, UK: a linked routine data analysis using decision tree classification. *Lancet Public Health*. (2023) 8(8):e629–38. doi: 10.1016/S2468-2667(23)00126-3
60. Hui V, Eby M, Constantino RE, Lee H, Zelazny J, Chang JC, et al. Examining the supports and advice that women with intimate partner violence experience received in online health communities: text mining approach. *J Med Internet Res*. (2023) 25(1):e48607. doi: 10.2196/48607

61. Zhang Z. Probing social support and self-disclosure within Chinese online domestic violence support groups: leveraging multiple machine learning approaches. *J Community Psychol.* (2025) 53(5):e70027. doi: 10.1002/jcop.70027
62. Salehi M, Ghahari S, Hosseinzadeh M, Ghalichi L. Domestic violence risk prediction in Iran using a machine learning approach by analyzing Persian textual content in social media. *Heliyon.* (2023) 9(5):e15667. doi: 10.1016/j.heliyon.2023.e15667
63. De Filippo A, Bellatin P, Tietz N, Grant E, Whitefield A, Nkopane P, et al. Effects of digital chatbot on gender attitudes and exposure to intimate partner violence among young women in South Africa. *PLOS Digit Health.* (2023) 2(10):e0000358. doi: 10.1371/journal.pdig.0000358
64. Vogt C, Giljohann S, Köpsel N, González Cabezas S, Houtsonen J, Izaguirre Choperena A, et al. From needs assessment to usability testing: evaluating the AinoAid™ chatbot for domestic violence support. *BMC Womens Health.* (2026) 26(1):31. doi: 10.1186/s12905-025-04202-3
65. Hui V, Tian L, Feng X, Chen Q, Wong AKC, Bloom TL. Perspectives and preferences of domestic violence survivors regarding digital platform and AI chatbot for help-seeking: a qualitative study. *PLoS One.* (2026) 21(2):e0342453. doi: 10.1371/journal.pone.0342453
66. Gu J, Carballo KV, Ma Y, Bertsimas D, Khurana B. Leveraging multimodal machine learning for accurate risk identification of intimate partner violence. *NPJ Womens Health.* (2026) 4(1):15. doi: 10.1038/s44294-025-00126-3
67. Saudagar S, Jadhav A, Shelke M. Towards safer homes: ai-driven predictions for domestic violence prevention. *Commun Comput Inform Sci.* (2025) 2491:206–20. doi: 10.1007/978-3-031-90577-3_19
68. Alnuaimi AFAH, Albaldawi Thk. An empirical ensemble machine learning approach for the classification and prediction of domestic violence crimes in Iraq. In: *2025 3rd International Conference on Business Analytics for Technology and Security (ICBATS).* IEEE Conferences (2025). doi: 10.1109/icbats66542.2025.11258312
69. Cruz-Mendoza MC, Melendez-Armenta RA, Canul-Reich J, Muñoz-Benítez J. Machine learning applied to improve prevention of, response to, and understanding of violence against women. *Informatics.* (2025) 12(2):40. doi: 10.3390/informatics12020040
70. Warnars HLHS, Sunge AS, Suzanna, Bevyadi B, Muyebe MK. Prediction of violence against women using ensemble learning models: a comparative study of LightGBM, XGBoost, and others. *Int J Saf Secur Eng.* (2025) 15(4):727–35. doi: 10.18280/ijse.150409
71. González-Sanabria J-S, Pinto C, Zuñiga J, Ordoñez H, Blanco X. XGBoost classifier-based model to predict the nature of gender-based violence. Case study: santander, Colombia. *J Univ Comput Sci.* (2025) 31(8):758–87. doi: 10.3897/jucs.129515
72. Shashidhara S, Mamidi P, Vaidya S, Daral I. Using machine learning prediction to create a 15-question IPV measurement tool. *J Interpers Violence.* (2024) 39(1-2):11–34. doi: 10.1177/08862605231191187
73. Hossain MM, Asadullah M, Rahaman A, Miah MS, Hasan MZ, Paul T, et al. Prediction on domestic violence in Bangladesh during the COVID-19 outbreak using machine learning methods. *Appl Syst Innov.* (2021) 4(4):77. doi: 10.3390/asi4040077
74. Choi EY, Hyun HJ. A predictive model of domestic violence in multicultural families focusing on perpetrator. *Asian Nurs Res.* (2016) 10(3):213–20. doi: 10.1016/j.anr.2016.04.004
75. Warnars HLHS, Sunge AS, Suzanna, Bevyadi B, Muyebe M. The effect of feature selection based on CatBoost, LIME, SHAP, and random forest in identifying the risk of violence against women. *Stud Comput Intell.* (2026) 1227:331–47. doi: 10.1007/978-3-032-01133-6_23
76. Le BL, Trafford L, Firtala S, Williams B. Beyond physical violence: a machine learning framework for predicting IPV victimisation using multidimensional predictors. *Eur J Criminol.* (2026):1–28. doi: 10.1177/14773708251412637
77. Başaran F, Duru P. Determining domestic violence against women using machine learning methods: the case of türkiye. *J Eval Clin Pract.* (2025) 31(3):e14180. doi: 10.1111/jep.14180
78. Um MY, Manikonda L, Eapen DJ, Ferguson KM, Maria DMS, Narendorf SC, et al. Predicting intimate partner violence perpetration among young adults experiencing homelessness in seven U.S. Cities using interpretable machine learning. *J Interpers Violence.* (2025) 40(7-8):1727–51. doi: 10.1177/08862605241263588
79. Rahman R, Khan MNA, Sara SS, Rahman MA, Khan ZI. A comparative study of machine learning algorithms for predicting domestic violence vulnerability in Liberian women. *BMC Womens Health.* (2023) 23(1):542. doi: 10.1186/s12905-023-02701-9
80. Chen Z, Ma W, Li Y, Guo W, Wang S, Zhang W, et al. Using machine learning to estimate the incidence rate of intimate partner violence. *Sci Rep.* (2023) 13(1):5533. doi: 10.1038/s41598-023-31846-8
81. Todorovic K, O'Leary E, Ward KP, Devarasetty PP, Lee SJ, Knox M, et al. Prevalence, increase and predictors of family violence during the COVID-19 pandemic, using modern machine learning approaches. *Front Psychiatry.* (2022) 13:883294. doi: 10.3389/fpsyt.2022.883294
82. Verrey J, Ariel B, Harinam V, Dillon L. Using machine learning to forecast domestic homicide via police data and super learning. *Sci Rep.* (2023) 13(1):22932. doi: 10.1038/s41598-023-50274-2
83. Peddireddy S, Yan S, Kim S, Xiao R. Intimate partner homicide among women of childbearing age: identifying multilevel risk factors with machine learning. *AMIA Annu Symp Proc.* (2024) 2024:1003–12.
84. More S, Ticona W. A proposal of data mining model for the classification of an act of violence as a case of attempted femicide in the Peruvian scope. *Lect Not Netw Syst.* (2023) 724:756–72. doi: 10.1007/978-3-031-35314-7_63
85. Dogan A. Participatory machine learning models in femicide news alert detection. In: *Thirty-Sixth AAAI Conference on Artificial Intelligence/Thirty-Fourth Conference on Innovative Applications of Artificial Intelligence/Twelfth Symposium on Educational Advances in Artificial Intelligence.* Assoc Advancement Artificial Intelligence (2022). Vol. 36. p. 13134–5. doi: 10.1609/aaai.v36i11.21703
86. Hui V, Constantino RE, Lee YJ. Harnessing machine learning in tackling domestic violence—an integrative review. *Int J Environ Res Public Health.* (2023) 20(6):4984. doi: 10.3390/ijerph20064984
87. Prakash J, Fay K, Gujrathi R, Rosner B, Nour N, Khurana B. Antepartum intimate partner violence: development of a risk prediction model. *J Womens Health.* (2024) 33(9):1259–66. doi: 10.1089/jwh.2024.0038
88. Jung S, Himmen MK, Velupillai N, Buro K. Does stalking behavior improve risk prediction of intimate partner violence? *Vict Offenders.* (2022) 17(4):553–70. doi: 10.1080/15564886.2021.1963899
89. Syed S, Ashwick R, Schlosser M, Gonzalez-Izquierdo A, Li L, Gilbert R. Predictive value of indicators for identifying child maltreatment and intimate partner violence in coded electronic health records: a systematic review and meta-analysis. *Arch Dis Child.* (2021) 106(1):44–53. doi: 10.1136/archdischild-2020-319027
90. Hui V, Law J, Guan S, Feng X. Leveraging GraphRAG with large language models to identify help-seeking information among domestic violence survivors from qualitative interviews. *Stud Health Technol Inform.* (2025) 329:1794–5. doi: 10.3233/SHTI251218
91. Vowels LM, Vohra P, Li D, Zeinoddin P, Elswick A, Marcantonio T, et al. Large language models for psychosocial risk assessment: a multi-method evaluation across suicide, intimate partner violence, and substance misuse. *PLOS Digit Health.* (2026) 5(4):e0001352. doi: 10.1371/journal.pdig.0001352
92. Harris DR, Clements BP, Anthony N, Bhullar M, Quesinberry D, Delcher C. Addressing diagnostic code variability in intimate partner violence surveillance through natural language processing: evidence from substance use disorder populations. *Drug Alcohol Depend.* (2026) 284:113166. doi: 10.1016/j.drugalcdep.2026.113166
93. Xue J, Chen J, Chen C, Hu R, Zhu T. The hidden pandemic of family violence during COVID-19: unsupervised learning of tweets. *J Med Internet Res.* (2020) 22(11):e24361. doi: 10.2196/24361
94. Castorena CM, Abundez IM, Alejo R, Granda-Gutiérrez EE, Rendón E, Villegas O. Deep neural network for gender-based violence detection on twitter messages. *Mathematics.* (2021) 9(8):807. doi: 10.3390/math9080807
95. Subramani S, Michalska S, Wang H, Du J, Zhang Y, Shakeel H. Deep learning for multi-class identification from domestic violence online posts. *IEEE Access.* (2019) 7:46210–24. doi: 10.1109/ACCESS.2019.2908827
96. Subramani S, Wang H, Vu HQ, Li G. Domestic violence crisis identification from Facebook posts based on deep learning. *IEEE Access.* (2018) 6. doi: 10.1109/access.2018.2871446
97. Salehi M, Ghahari S. Classification of domestic violence Persian textual content in social media based on topic modeling and ensemble learning. *Heliyon.* (2024) 10(22):e39953. doi: 10.1016/j.heliyon.2024.e39953