

<https://doi.org/10.1038/s44294-025-00126-3>

# Leveraging multimodal machine learning for accurate risk identification of intimate partner violence

Check for updates

Jiayi Gu<sup>1,2</sup>, Kimberly Villalobos Carballo<sup>1,2</sup>, Yu Ma<sup>1,2</sup>, Dimitris Bertsimas<sup>2,3,4,6</sup> & Bharti Khurana<sup>1,5,6</sup>

Intimate partner violence (IPV) refers to the abuse from previous or current partners. It is a widespread but underreported public health concern that has a wide range of negative effects on the physical and mental health of those affected. This work presents machine learning models for the early detection of IPV in clinical settings, developed with a dataset of female patients who sought help at a domestic abuse intervention and prevention center of a major hospital in the United States. Utilizing tabular clinical data and unstructured clinical notes, we build single-modality and multimodal models for different data availability scenarios. Our multimodal model can identify patients at risk of IPV with an AUC of 0.88 and years before patients seek help. We validated the model on patients who did not seek help at the intervention center and patients from another hospital in the same integrated network with comparable performance.

Intimate partner violence, or IPV, is characterized by violence or aggression by current or former spouses and partners<sup>1</sup>. IPV is a significant public health challenge that affects people from all socio-economic and geographic backgrounds. The World Health Organization classified IPV as the most widespread form of aggression against women worldwide. According to the Centers for Disease Control and Prevention, more than one-third of women in the United States will experience IPV at some point in their lives<sup>1</sup>. In addition to potentially life-threatening injuries, IPV is associated with negative physical health outcomes such as chronic pain, sexually transmitted infections, menopausal symptoms, and mental health disorders such as depression, PTSD, anxiety, and self-harm<sup>2-4</sup>. Early detection and intervention of IPV are crucial in preventing the worsening of health in those affected. However, victims may hesitate to disclose abusive relationships due to factors such as safety concerns and financial dependency on the abuser<sup>4</sup>, presenting challenges for timely intervention.

Healthcare clinicians have a unique edge in identifying early signs of IPV, as victims present to them frequently before they seek help from law enforcement or social agencies. The United States Preventive Services Task Force recommends routine screening for IPV in all women of childbearing age in healthcare settings<sup>5</sup>. However, the current screening tools capture only a fraction of cases as they simply rely on self-reporting of violence. Given the hesitancy to seek help by IPV victims due to fear, stigma, and safety concerns, many cases go undiagnosed, leading to missed opportunities for timely intervention. As a result, there is an unmet critical medical

and social need for a more objective and comprehensive screening method to detect IPV.

Recent studies have demonstrated that historical clinical and imaging studies can provide valuable information in detecting IPV risk<sup>6-10</sup>. In particular, with access to the imaging history of patients, radiologists have an advantage in recognizing signs of IPV. Patterns in radiology studies, including high frequency of radiology imaging and injuries to the face, neck, and upper extremities, can indicate the likelihood of IPV<sup>11-16</sup>. However, time constraints, silos of subspecialization, cognitive overload, and the urgency to address immediate symptoms often limit the ability of radiologists and clinicians to leverage this information<sup>8</sup>. This motivates the need for an automated decision support tool that can utilize radiology findings and other clinical information to improve IPV identification, as prior work has shown that incorporating clinical information helps improve the accuracy of radiology interpretation and reporting<sup>17</sup>. Such a tool can help healthcare providers identify those at risk of IPV, allowing them to connect patients to appropriate medical and social resources effectively. There exists prior work on building machine learning tools for IPV identification<sup>18-20</sup>. However, they focused on relatively small populations or on predicting intervention response rather than broadly identifying risk in clinical settings.

Recognizing this potential, our work introduces the first multimodal machine-learning approach that incorporates clinical notes (including radiological notes) and structured data from Electronic Medical Records (EMR) to identify patients at risk of IPV in clinical settings. Recent advances in off-the-shelf large language models (LLMs) pre-trained on large corpora

<sup>1</sup>Trauma Imaging Research and Innovation Center, Brigham and Women's Hospital, Boston, MA, USA. <sup>2</sup>Operations Research Center, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. <sup>3</sup>Sloan School of Management, MIT, Cambridge, MA, USA. <sup>4</sup>Dynamic Ideas LLC, Waltham, MA, USA. <sup>5</sup>Harvard Medical School, Boston, MA, USA. <sup>6</sup>These authors contributed equally: Dimitris Bertsimas, Bharti Khurana. e-mail: [bkhurana@bwh.harvard.edu](mailto:bkhurana@bwh.harvard.edu)

of medical texts have demonstrated strong performance in medical language learning tasks<sup>21-24</sup>, allowing for an effective framework for learning from unstructured texts such as radiology notes<sup>6</sup>. Moreover, multimodal machine learning frameworks have been shown to offer superior performance in clinical tasks compared with single-modality approaches. Soenksen et al. proposed a Holistic AI in Medicine (HAIM) framework that utilizes pre-trained feature extractors (such as medical LLMs) to unify learning from different data modalities, including tabular data, language, and imaging data. This multimodal approach consistently outperforms single-modality models for various clinical prediction and diagnostic tasks<sup>25</sup>.

In this work, we propose three distinct models of IPV risk detection: two single modality models utilizing tabular data or clinical notes, and a Holistic AI in Medicine (HAIM) multimodal fusion model. These models are developed with a dataset of female patients enrolled in a domestic abuse intervention and prevention (DAIP) center (cases) at a U.S. academic health center (AHC1), along with age and demographics-matched control patients from the same AHC. The different models are designed to accommodate different clinical data availability scenarios, ranging from settings with clinical note data to those with only tabular data. We evaluate and validate the performance of the models to show their potential for real-world clinical application. Our results show that the multimodal fusion model offers the strongest and most stable performance.

## Results

### Experiment setup

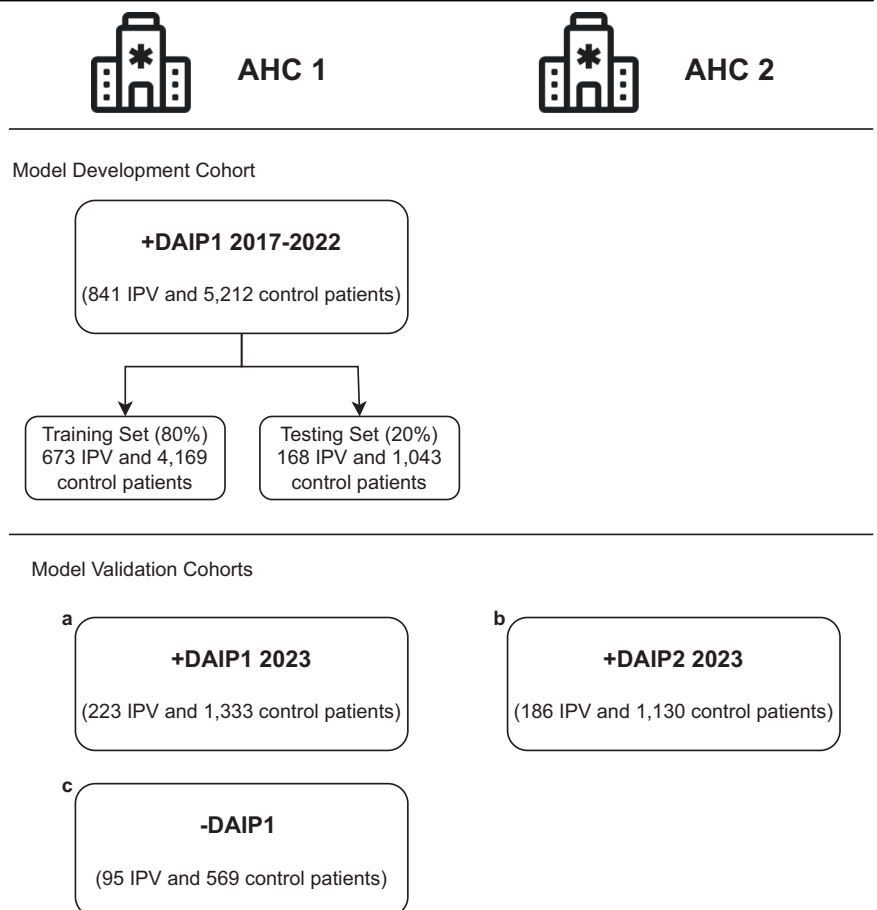
The IPV group used to develop the models consists of 841 female patients enrolled in the domestic abuse intervention and prevention center at a US Academic Health Center (AHC1) between 2017 to 2019 and 2021 to 2022, through either self-reporting or clinician referral. We exclude patients from 2020 due to its unique nature because of COVID-19, marked by a drop in patients reporting IPV and notable shifts in

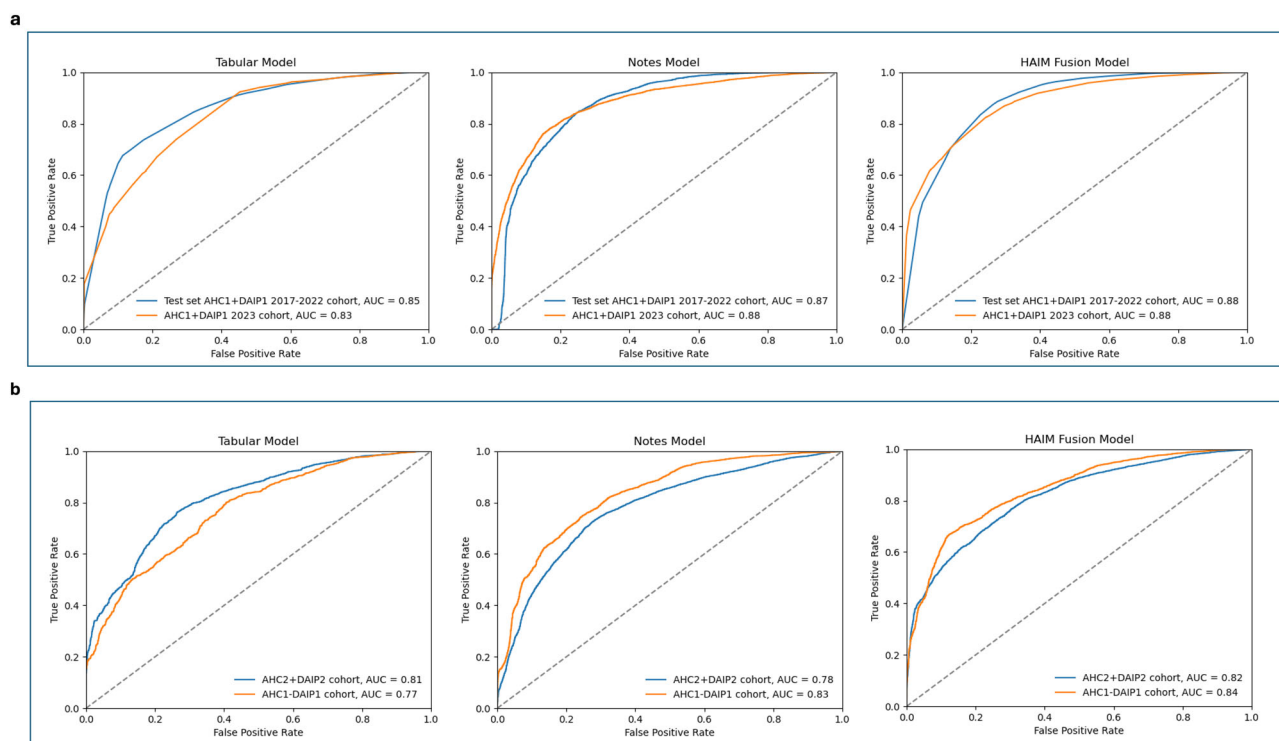
injury patterns<sup>26</sup>. This group is accompanied by 5,212 demographics-matched non-IPV (control) patients. The control group is constructed in a two-step process: (1) random sampling of female patients in the AHC, excluding any patients who carry any ICD code diagnoses of physical and non-physical abuse (2) matching based on demographic characteristics to ensure similarity to the IPV group. Details on the matching process are provided in the Methods section.

We validate our models on three cohorts of patients, each with demographics-matched control patients. The IPV patients in the first validation cohort include AHC1 patients who enrolled in the DAIP in 2023. Recognizing the limitation of using patients motivated to seek help at the DAIP centers, the second validation cohort includes patients who carried diagnoses of IPV at the AHC1 from 2021 to 2023 but did not enroll in any DAIP. This group of IPV patients is defined as carrying ICD-10 code diagnoses that include both physical and non-physical abuse. The last validation cohort includes IPV patients enrolled in DAIP2 in AHC2, a different hospital from the same healthcare network. Patients under 18 were excluded from all cohorts. Figure 1 summarises all cohorts for model development and validation.

To assess each patient’s evolving risk of IPV, we make daily predictions when patients encounter the hospital system on whether they might be experiencing IPV. To prevent target leakage, we adopt a patient-level splitting strategy; we train with 80% of the patients and test with 20% of the patients. All observations of a specific patient are assigned to either the training or testing group. For self-reported IPV patients, all corresponding observations are labeled as ‘IPV’ (case), while all observations for control patients are labeled as ‘non-IPV’ (control). This labeling scheme maintains consistency in classifying IPV observations, encouraging the model to capture risk factors that could indicate a higher likelihood of IPV, detect early signs of abuse before self-reporting, and allow for continuous monitoring of IPV likelihood over time.

**Fig. 1 | Model development and validation cohorts.** The patients come from two U.S. Academic Health Centers (AHC1 and AHC2) in an integrated care system. The models are developed with patients enrolled in the domestic abuse intervention and prevention center in AHC1 from 2017-2022 (AHC1 + DAIP1 2017-2022), excluding 2020 patients. The patients are separated patient-wise into training (80%) and testing (20%) sets. The models are validated on three cohorts of patients: (a) AHC1 + DAIP1 2023: patients enrolled in DAIP1 at AHC1 during 2023, with optimally-matched control patients, (b) AHC2 + DAIP2 2023: patients enrolled in DAIP2 at AHC2 during 2023, with optimally matched control patients, (c) AHC1-DAIP1: patients with IPV-related diagnoses but did not seek help from DAIP1 at AHC1 from 2021 to 2023, with optimally-matched control-patients.





**Fig. 2 | The receiver operating characteristic (ROC) curves of the tabular model, notes model, and HAIM fusion model on all cohorts. a** On the out-of-sample test set in the model development cohort (AHC1 + DAIP1 2017–2022 cohort), the tabular model has an AUC of 0.85 and the notes model has an AUC of 0.87. The HAIM outperforms both single-modality models with an AUC of 0.88. On the internal validation cohort (AHC1 + DAIP1 2023 cohort), all models maintain an AUC above 0.83, with the HAIM fusion model having a nearly equal AUC of 0.88.

**b** On the IPV patients in AHC2 (AHC2 + DAIP2 2023 cohort), the tabular model has an AUC of 0.81 and the notes model has an AUC of 0.78. On the AHC1 IPV patients who did not seek help at the domestic abuse intervention and prevention center, the tabular model has an AUC of 0.81 and the notes model has an AUC of 0.78. The HAIM fusion model maintains stable and superior performance, with AUCs of 0.82 and 0.84 for the two cohorts respectively. AHC=Academic Health Center, DAIP=domestic abuse intervention and prevention.

### Model performance

We evaluate model performance using the area under the receiver operating characteristic curve (AUC), sensitivity (true positive rate), and specificity (true negative rate), which are well-suited metrics for evaluating unbalanced learning tasks<sup>27</sup>. For IPV identification in practice, it is crucial to balance sensitivity and specificity. False negatives can lead to victims not receiving care or intervention, leaving them vulnerable to continued abuse. False positives can result in individuals being wrongly labeled as victims, leading to unnecessary interventions, emotional distress, and possible harm to the patient-clinician relationship. We report these values for both the test set and the validation sets.

On the out-of-sample test set in the AHC1 + DAIP1 2017–2022 cohort, the tabular models can predict IPV risk with an AUC of 0.85 and the clinical notes model with an AUC of 0.87. The HAIM fusion model yields an AUC of 0.88, outperforming both single-modality models. On the AHC1 + DAIP1 2023 cohort, all models maintain an AUC above 0.82, with the HAIM fusion model having a nearly equal AUC of 0.88. Although the HAIM fusion model consistently outperforms the single-modality models, the tabular data-only model can perform comparably as a standalone model for IPV risk assessment when clinical texts are unavailable or difficult to collect. The performance of the three models on the AHC1 + DAIP1 cohorts is shown in Fig. 2a. The AUC performance of the tabular and fusion model across demographic subgroups is presented in Supplementary Tables 1A and 1B.

We calibrate the predicted probabilities to accurately reflect the likelihood of IPV presence using isotonic regression<sup>28</sup> (results in Supplementary Fig. 1). Given the calibrated probabilities, we evaluate the Tabular and Fusion models' accuracy, sensitivity, and specificity. We compare and select the probability threshold that balances test-set sensitivity and specificity. The comparison and the performance of these thresholds on the internal

validation cohort are presented in Supplementary Table 2. Both models can achieve above 73% sensitivity and specificity with the chosen threshold.

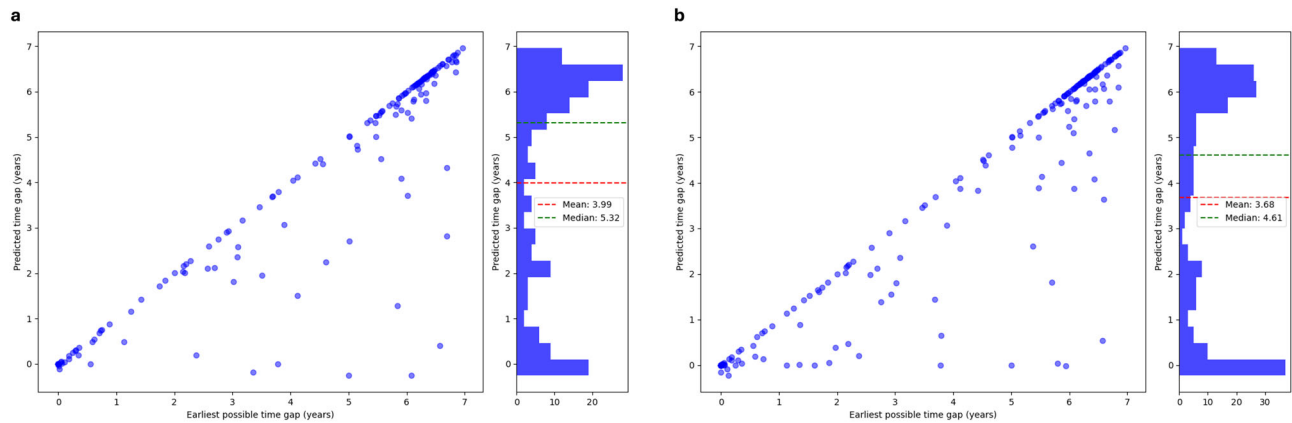
### Prediction-report lead time in self-report patients

Our models can detect IPV risk years before the patients enroll in the DAIP center. In the AHC1 + DAIP1 2023 cohort, the tabular model can detect 68% of IPV cases ahead of time, while the fusion model can predict 80.6% of cases in advance. We note that the dataset only includes data from 2017 onwards to align with the rollout of a new electronic medical record system at both AHCs. This restricts our ability to identify cases at the earliest stages. We calculate a prediction-report lead time by comparing the patient's first self-report date with the earliest date the model detected IPV; a positive result indicates early detection. Figure 3 presents the relationship between the largest possible time gap (the difference between the report date and the date with the earliest available data) and the prediction-report time gap. The tabular model achieves a mean lead time of 3.99 years and a median lead time of 5.32 years, whereas the fusion model has an average lead time of 3.68 years and a median lead time of 4.61 years. While the tabular model has higher mean and median prediction lead times, the fusion model can detect more IPV cases in advance.

### Validation results

We further validate the models on two additional validation cohorts: IPV patients enrolled in the domestic abuse intervention and prevention center at AHC2 from the same integrated care system (AHC2 + DAIP2) and AHC1 IPV patients who did not seek support from DAIP (AHC1-DAIP1). The performance of the models on the validation cohorts is shown in Fig. 2b.

For AHC2 patients, the tabular model yields an AUC of 0.81, and the clinical notes model an AUC of 0.78 while the merged model achieves an AUC of 0.82. On the AHC1-DAIP1 cohort, the tabular model yields an



**Fig. 3 | Prediction-report lead time in self-report patients.** The prediction-report lead time is calculated as the difference between the patient's first self-report date and the earliest date when the model detected IPV; a positive result indicates early detection. We compare this value to the earliest possible time gap: the difference between the first time the patient encounters the healthcare system and when they

report IPV. Both the tabular model (a) and the fusion model (b) can detect IPV risk on average more than three years in advance. While the tabular model has slightly higher mean and median prediction lead times, the fusion model can detect more IPV cases in advance.

AUC of 0.77, and the notes model has an AUC of 0.83. The HAIM fusion model outperforms both models with an AUC of 0.84. The enrollment of IPV victims in domestic abuse intervention and prevention centers can change patterns in healthcare utilization, introducing potential feature variations when compared with the patients who disclosed IPV but did not seek help from DAIP. This could account for the variation in the tabular model's performance. Overall, the validation results demonstrate the models' generalizability to external populations and their potential to identify unreported IPV cases.

### Predictive features

Since the clinical texts are encoded using a transformer-based language model, it is challenging to determine which risk factors in clinical texts inform the predictions of the notes and the fusion model. Nonetheless, we can approximate the feature importance of the tabular model using the SHapley Additive exPlanations (SHAP)<sup>29</sup>. Using the SHAP Python library, we identify the most important features in the tabular model and use a SHAP summary plot to visualize the overall impact of these features on model predictions<sup>30</sup>. The SHAP summary plot is presented in Fig. 4.

We observe that specific clinical features, such as mental health disorders, chest pain, and painkiller use are correlated with a greater likelihood of IPV. Additionally, social factors such as high social deprivation are observed with higher levels of IPV risk. Aside from capturing social factors and trauma-related health outcomes, we note that the model captures injury and healthcare utilization patterns related to IPV. Notably, we observe a correlation between the high utilization of radiology tests for the upper extremity and unclassified locations (often seen in ED setting) and a higher probability of IPV presence, consistent with previous research findings<sup>15,31</sup>. On the contrary, high utilization of mammography and cervical cancer screening could indicate a lower likelihood of IPV presence. This observation may be explained by varied access to healthcare and preventive services across different patient populations. Patients who attend regular screening likely have better access to healthcare and support services. Additionally, their involvement in preventive healthcare services may suggest a lower fear of seeking medical attention, in contrast to the reluctance seen in IPV patients<sup>4</sup>.

The SHAP summary plots for language-based models (the clinical notes model and the fusion model) are provided in Supplementary Figs. S2 and S3. Since each note is represented using the language model embedding of the note, the SHAP plot shows which embedding features contribute the most to the predictions, rather than words or phrases directly from the notes. However, they provide information on which portions (type of note) contribute most to the predictions of the model, allowing the

possibility to trace back to the original note when further explanation is needed.

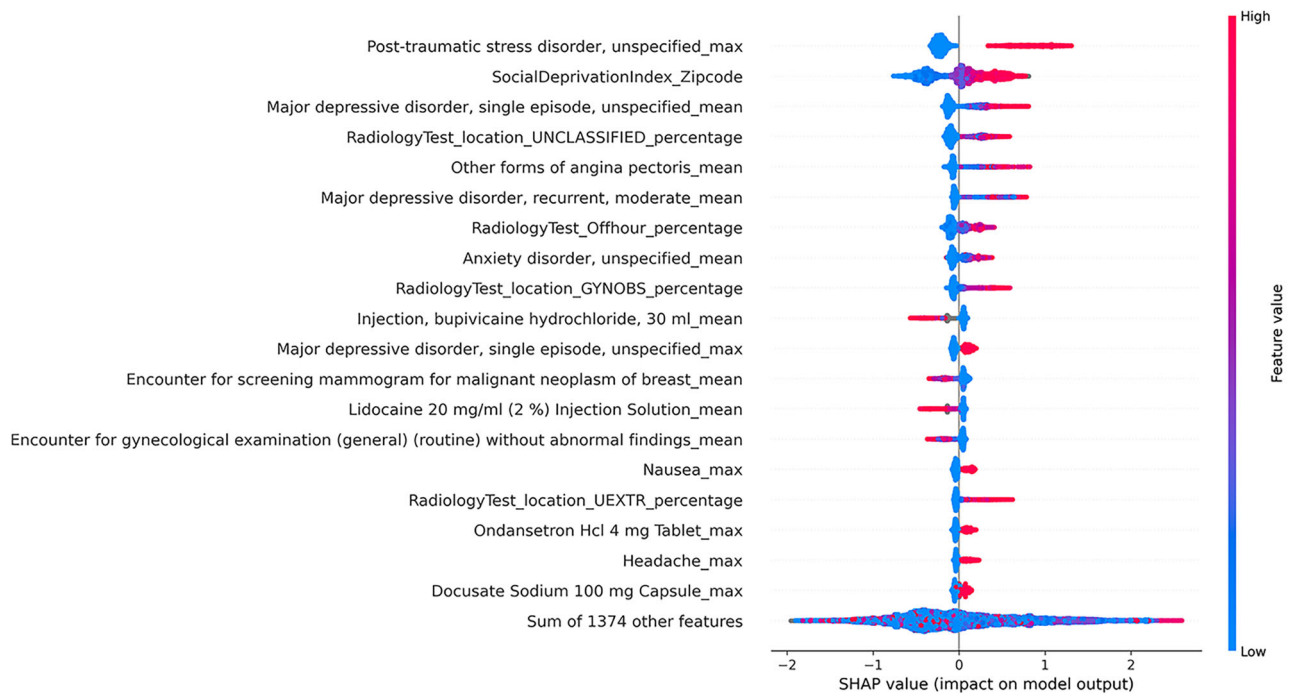
### Discussion

This work introduces three machine learning models for the detection of intimate partner violence (IPV) in clinical settings. We developed models tailored to different data availability scenarios, including models that use only tabular data and models that incorporate clinical notes (such as radiology reports) when available. All three models demonstrate strong out-of-sample performance. Notably, the HAIM fusion model outperforms single modality models, achieving an out-of-sample area under the receiver operating characteristic curve (AUC) of over 0.8 in all cohorts. The single modality tabular model yields comparable performance, making it useful in low-resource hospitals where only limited medical data is recorded. Additionally, the tabular model can be used with interpretability frameworks like SHAP to provide insights into the risk factors of IPV, which can help clinicians better assess IPV cases and inform intervention decisions. Moreover, the models can identify risk of IPV years before patients seek help, demonstrating their potential to support proactive screening.

The HAIM fusion model performs consistently when validated on IPV patients who did not seek support from the domestic abuse intervention and prevention centers and patients from another AHC in the same network, achieving AUCs above 0.8. Even when a decline in performance is observed in either the tabular or the notes model in these validation groups, the fusion model maintains a strong performance. This result demonstrates that compared to traditional machine learning approaches that utilized a single modality of data, the multimodal modeling approach improved stability in model performance even with modality-specific feature variability when applied to an unseen group of patients. In such cases, it can be difficult to know in advance which modality's predictions will be more reliable; by combining both, the fusion model achieves more stable performance than relying on either modality alone. In the HAIM framework, the different modalities are processed separately and only merged at the prediction stage. As a result, feature changes in one modality do not hinder learning from the other modalities. This advantage of the HAIM framework is particularly relevant in healthcare, where variations in data availability and in the recording of unstructured data (notes, images, etc.) are common across different hospitals.

The models proposed have significant clinical implications, transforming how healthcare providers identify and support patients at risk of IPV. Traditional screening methods rely only on self-reporting, often leading to underreporting due to fear, stigma, and safety concerns. A prediction tool utilizing the models proposed can address this gap by providing





**Fig. 4 | SHAP summary plot for the tabular model.** The SHAP summary plots visualize the SHAP values of the 20 most important features in the tabular model. Each dot on the plot represents a prediction instance, with the color indicating the value of the feature. The position on the horizontal axis depicts the direction and magnitude contribution to the individual prediction (right for a higher likelihood of

IPV, left for a lower likelihood of IPV). The SHAP summary plot gives a global understanding of feature importance and how individual features impact predictions. Features with high SHAP values for red dots indicate a potential positive correlation between high feature values and increased IPV risk, while high SHAP values for blue dots indicate the opposite.

an objective, evidence-based risk assessment based on historical clinical data available in EMR. By enabling early identification of at-risk patients, the tool can facilitate timely, trauma-informed interventions that can prevent escalation into severe injuries, chronic health conditions, and even IPV-related homicides.

While our models demonstrate strong performance in validation cohorts, several limitations and considerations should be taken into account for their practical use. First, the model is developed using IPV patients who sought help at DAIP, which can introduce selection bias. Although we validate the models on patients who did not seek help from DAIP but carry IPV-related diagnoses, IPV victims who are reluctant to disclose their experience or seek help are underrepresented in the model, potentially limiting the models' generalizability. Second, due to the underreported nature of IPV, the true population distribution of IPV and non-IPV patients is unknown. As a result, the distribution of patients in our validation cohorts may not accurately reflect real-world deployment settings. Careful performance evaluation on the more general population of female patients should be conducted before implementing the models for clinical use. Third, our models rely on EMR data, which may be incomplete, inconsistent, or subject to documentation biases. Lastly, while the tabular model offers an approximate understanding of predictive factors of IPV, it does not statistically attribute or quantify the impact of each factor. An important direction of future work is to conduct a rigorous statistical analysis of contributing factors. In this work, individual-level demographic variables of the patients were excluded to avoid learning a biased predictor; however, prior works have shown their value in IPV identification<sup>18,19</sup>. Future iterations of the model should carefully include socio-demographic variables to aid in IPV identification and evaluate their significance.

It is also important to emphasize that the models are not intended to provide a definitive diagnosis of IPV, but rather to serve as a tool to assist healthcare providers in having patient-centered conversations around IPV. By generating individualized risk probability scores, the models enable timely screening and help providers identify patients who may benefit from patient-centered, sensitive conversations and the offering of support

resources, without requiring patient disclosure or assigning a formal diagnosis. Previous research has identified that intrusion and loss of autonomy during the disclosure process can hinder IPV victims' willingness to seek help<sup>32-34</sup>. Therefore, when utilizing the models in any IPV screening support system, it is crucial to implement safeguards to ensure that IPV-related conversations are conducted sensitively without compromising patient autonomy.

In conclusion, this work represents the first effort towards a practical multimodal machine learning approach to IPV risk detection in healthcare settings. We plan to use our models to develop a decision support tool embedded in electronic medical record systems to provide real-time IPV risk evaluations in clinical settings. When used in a patient-centered manner, this tool can serve as a key component of a proactive approach to IPV intervention, enabling timely and effective support and ultimately leading to improved long-term health outcomes for at-risk patients.

## Methods

### Patient cohorts control demographics matching

We constructed control (non-IPV) cohorts that closely mirrored the demographic characteristics of IPV patients by optimally matching the age and demographic information of both the IPV and control patients, an approach inspired by recent work by Bertsimas et al. on mirroring randomized control trials using observational data<sup>35</sup>. Since the IPV group has a different demographic composition than the general patient population, this discourages the model from overfitting to demographic patterns and producing biased predictions. The initial pool of control patients provided by the hospital consists of ten randomly selected female patients within a  $\pm 10$ -year age range and approximately similar racial backgrounds to each IPV patient. A logistic regression model was trained to classify patients into IPV and non-IPV groups based on demographics (race and ethnicity features). This simple approach provides an approximation of the strength of the association between each demographic variable and IPV risk via the coefficients of the logistic regression model. These coefficients are incorporated in a mathematical optimization model to encourage close matching

**Table 1 | Model Development Cohort Patient Characteristics**

	# patients	Total 6053	Control 5212	IPV 841
Age	<30	16.50%	16.42%	17.00%
	30–39	30.48%	30.58%	29.85%
	40–49	23.03%	23.00%	23.19%
	50–59	15.35%	15.33%	15.46%
	60+	14.64%	14.66%	14.51%
Race	White	37.67%	38.05%	35.32%
	Black	28.65%	28.61%	28.89%
	Asian	3.65%	3.78%	2.85%
	Alaskan/Pacific	0.73%	0.58%	1.66%
	Other	20.06%	19.49%	23.54%
	Unknown	9.25%	9.50%	7.73%
	Multi-race	1.88%	1.80%	2.38%
Ethnicity	Hispanic	19.02%	19.19%	17.95%
	Non-Hispanic	71.40%	70.49%	77.05%
	Unknown	9.58%	10.32%	4.99%

Summary demographics statistics of the model development cohort, in percentages of each patient group.

in demographic variables that may have significantly different distributions in the IPV and control groups. Specifically, the coefficients are used to compute a weighted distance metric that quantifies the demographic similarity between each IPV and potential control patients. To achieve a closer match in age, we discretize age into 5-year age groups; each IPV patient is matched with six control patients within the same 5-year age group while minimizing the difference in their demographic features.

The final model development cohort contains 841 IPV and 5,212 control patients. Table 1 presents summary statistics of the demographics of the IPV and the optimally matched control cohort. The demographic distribution of the optimally matched control group closely mirrored that of the IPV group. Summary demographics statistics of the validation cohorts are included in Supplementary Tables 3–5.

**Data Sources**

We consider data from the EMRs between January 2017 and December 2023. The structured data include demographics, diagnosis history, medication records, past radiology studies, past hospital visits, and vital signs (e.g., blood pressure, weight, pain scores). Since we don’t have patient-level data on social determinants of health, we approximate this with public datasets such as social deprivation and household income at the zip code level<sup>36,37</sup>. The dataset also contains clinical notes and reports, including radiology reports, progress notes, physical exam notes, social worker notes, and Emergency Department notes. This study was approved by the Institutional Review Board of Mass General Brigham (Protocol # 2016P002096). Informed consent was waived due to the use of retrospective data from a large cohort collected over an extended period of years. This study was conducted in accordance with the Declaration of Helsinki.

**Data preprocessing and feature engineering**

In preprocessing the tabular data, since the hospital transitioned the diagnosis coding system (from ICD-9 to ICD-10) during the period considered in this study, we converted all diagnosis records in the ICD-9 to ICD-10 format. This standardization ensures consistency in representing the same diagnosis across data recorded in different periods. To prevent target leakage, we remove diagnoses that are directly indicative of IPV (e.g., adult physical/sexual abuse, encounter for mental health services for the victim and perpetrator of abuse). We include the top 5% of most frequent diagnoses and medications in the model, which results in a subset of 119

diagnoses and 187 medications. In addition, we retain specific diagnoses (e.g., head injury, diabetes) relevant to IPV based on previous research findings<sup>2,12,13</sup>. Additional tabular features include the frequency of emergency department (ED) visits, as prior research suggests a correlation between IPV and frequent ED visits<sup>6</sup>. Lastly, we extracted injury patterns, such as the injury site and time of radiology study from radiology tests. The time-related features are relevant as IPV-related visits tend to be more frequent during weekends, holidays, and outside of regular business hours<sup>38</sup>.

In preprocessing the clinical notes, text irrelevant to the prediction task (headers and footers) was removed. In addition, we remove sections in radiology reports that could directly contain the prediction target to prevent target leakage. For instance, the Indication/History/Reason for exam sections of the radiology report describe the clinical reason for the imaging study (e.g., assault by head strike), which can provide direct clues about the prediction task. Consequently, the model may learn to rely on this text to make predictions rather than capturing more generalized patterns that signal IPV-related injuries.

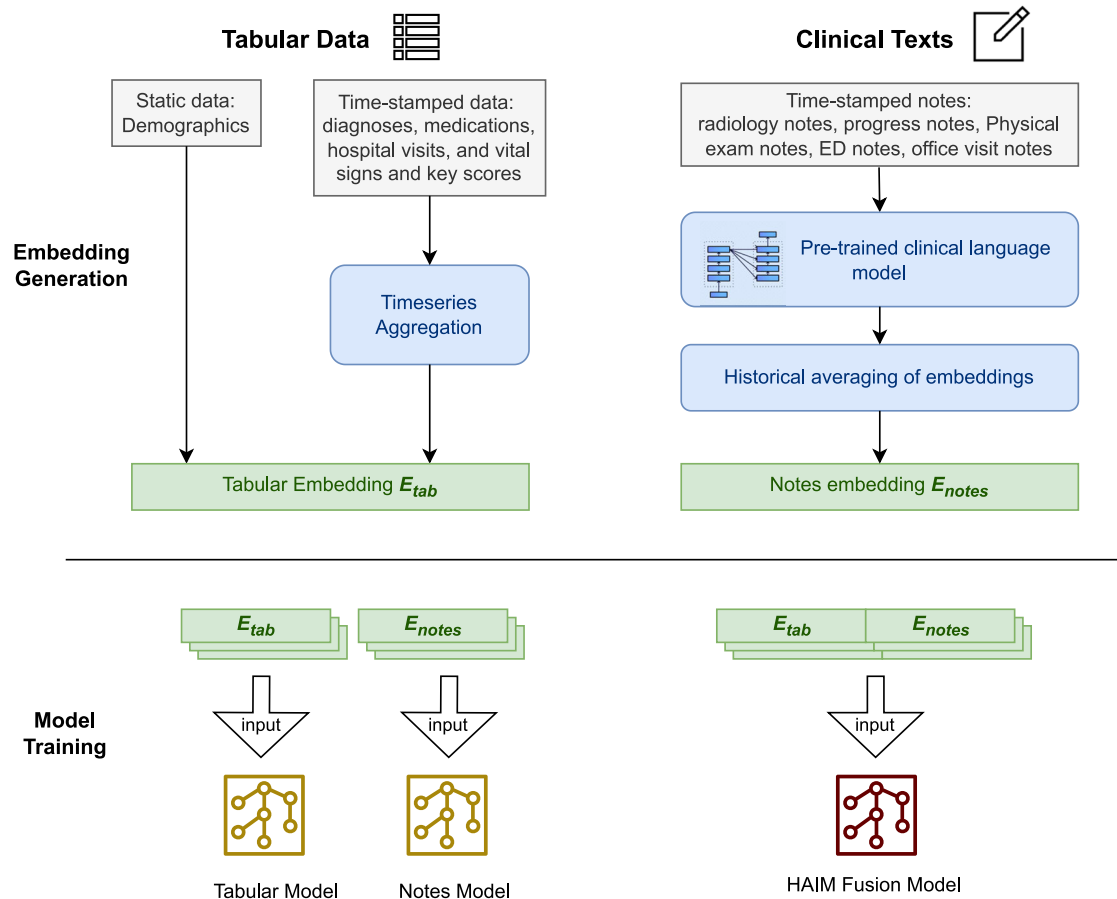
**Models**

For the tabular data-only model, we leverage time-series processing techniques to aggregate and summarize the historical timestamped measurements such as diagnoses, medications, vitals, visits, etc. This involves computing statistics such as the minimum, maximum, average values, and most recent values of the measurements over the time window up to the prediction day. Meanwhile, demographic values remain static and are not subject to time-series aggregation. By capturing the temporal trends and patterns in the data, we can assess whether changes or patterns in health indicators over time indicate the presence of IPV.

For clinical notes and reports, we leverage a transformer-based bidirectional encoder model pre-trained on a large corpus of clinical texts to generate contextual embeddings of the text. The resulting fixed-length embeddings of the notes serve as inputs to the downstream models. In contrast with traditional NLP techniques such as bag-of-words or TF-IDF, this approach captures contextual nuance in the texts and the medical knowledge learned from the pre-training of the language model. Given the extended length of clinical notes and reports, we encode the text with Clinical-Longformer, a compact encoder model suitable for encoding long sequences of medical<sup>21,39</sup>. With the recent success of generative LLMs and their domain-specific fine-tuned variants, we also evaluate MedAlpaca 7B, a medical domain fine-tuned LLaMA 7B model<sup>40</sup>. We build our prediction models using both Clinical-Longformer and MedAlpaca 7B to compare their performance as contextual encoders for our task. We utilize MedAlpaca 7B for its smaller model size and extended context window, as larger models pose significant challenges in deployment. For both note-only and fusion models, MedAlpaca achieved similar performance to the Clinical Longformer with a marginal AUC difference of ± 0.02. Detailed results and comparison are presented in Supplementary Table 6. This demonstrates that a smaller model can be used without sacrificing accuracy, allowing for deployment in clinical environments with limited computational resources.

For both the tabular and text-based models, we train downstream machine learning models using either the tabular data or the embedding representations of the clinical texts. The machine learning models compared include logistic regression, decision trees, random forest, and gradient-boosted trees. In preliminary experiments, XGBoost consistently outperformed the other three methods. Therefore, we opt to use XGBoost for our final models, using the xgboost Python library<sup>41</sup>. We select the best hyperparameters based on Grid Search with 5-fold cross-validation on the training set and report the performance on the test set. The hyperparameter tuning is done with the GridSearchCV module in the scikit-learn Python library<sup>42</sup>.

The HAIM fusion model concatenates the embeddings extracted from the tabular data with those of clinical texts for model training and prediction. This concatenation enables the model to capture interactions between different modalities, such as the relationships between demographic features, clinical measurements, and textual information from the clinical texts. The pipeline of the models is outlined in Fig. 5.



**Fig. 5 | Single modality and fusion model pipeline.** The tabular data includes static demographics data and time-stamped data like diagnoses, medications, and vitals. The timestamped data are processed using time-series techniques to capture historical feature variations (such as the most recent value, mean, max, etc.). The unstructured clinical texts are processed using a transformer-based clinical Large Language Model, and the

resulting fixed-length vector embeddings of the clinical texts are used as inputs to the downstream models. For single-modality models, we train classification models using either the tabular data or the embedding representations of the clinical texts. The HAIM fusion model concatenates the embeddings extracted from the tabular data with those of clinical texts as inputs to the model.

### Data availability

The clinical data in this study contain protected health information (PHI), hence not publicly available. Researchers interested in collaboration should contact the corresponding authors and apply for data access through the AHC’s Institutional Review Board.

### Code availability

Due to the risk of exposing protected health information and the highly sensitive nature of IPV risk detection, the code for data analysis and modeling is not publicly available. However, details on data preparation and modeling implementation are available in the methods section. Researchers interested in more details on the methodologies can contact the corresponding author.

Received: 22 November 2024; Accepted: 24 December 2025;

Published online: 13 March 2026

### References

- Basile, K. C., Smith, S. G., Kresnow, M., Khatiwada, S. & Leemis, R. W. The National Intimate Partner and Sexual Violence Survey: 2016/2017 Report on Intimate Partner Violence. *National Center for Injury Prevention and Control, Centers for Disease Control and Prevention* <https://www.cdc.gov/nisvs/documentation/nisvsReportonSexualViolence.pdf> (2022).
- Othman, S., Dillon, G., Hussain, R., Loxton, D. & Rahman, S. Mental and Physical Health and Intimate Partner Violence against Women: A Review of the Literature. *Int. J. Fam. Med.* **2013**, 313909 (2013).
- Stubbs, A. & Szoek, C. The Effect of Intimate Partner Violence on the Physical Health and Health-Related Behaviors of Women: A Systematic Review of the Literature. *Trauma Violence Abuse* **23**, 1157–1172 (2022).
- Hien, D. & Ruglass, L. Interpersonal partner violence and women in the United States: An overview of prevalence rates, psychiatric correlates and consequences and barriers to help seeking. *Int. J. Law Psychiatry* **32**, 48–55 (2009).
- US Preventive Services Task Force Screening for Intimate Partner Violence, Elder Abuse, and Abuse of Vulnerable Adults: US Preventive Services Task Force Final Recommendation Statement. *JAMA* **320**, 1678–1687 (2018).
- Reis, B. Y., Kohane, I. S. & Mandl, K. D. Longitudinal histories as predictors of future diagnoses of domestic abuse: modelling study. *BMJ* **339**, b3677 (2009).
- Prakash, J. et al. Antepartum Intimate Partner Violence: Development of a Risk Prediction Model. *J. Womens Health* 2002 <https://doi.org/10.1089/jwh.2024.0038> (2024).
- Khurana, B., Seltzer, S. E., Kohane, I. S. & Boland, G. W. Making the ‘invisible’ visible: transforming the detection of intimate partner violence. *BMJ Qual. Saf.* **29**, 241–244 (2020).
- Tabaie, A., Zeidan, A., Evans, D., Smith, R. & Kamaleswaran, R. A Novel Technique to Identify Intimate Partner Violence in a Hospital Setting. *West. J. Emerg. Med.* **23**, 781–788 (2022).
- Chen, I. Y. et al. Intimate Partner Violence and Injury Prediction From Radiology Reports. *Pac Symp Biocomput* **26**, 55–66 (2021).

11. George, E. et al. Radiologic Findings in Intimate Partner Violence. *Radiology* **291**, 62–69 (2019).
12. Watane, G. V. et al. Imaging Findings on Head Computed Tomography Scans in Victims of Intimate Partner Violence. *J. Comput. Assist. Tomogr.* **47**, 307–314 (2023).
13. Russo, A. et al. Imaging of Violence Against the Elderly and the Women. *Semin. Ultrasound CT MRI* **40**, 18–24 (2019).
14. Khurana, B. et al. Recognizing Isolated Ulnar Fractures as Potential Markers for Intimate Partner Violence. *J. Am. Coll. Radiol.* **18**, 1108–1117 (2021).
15. Khurana, B. et al. Upper extremity fractures due to intimate partner violence versus accidental causes. *Emerg. Radiol.* **29**, 89–97 (2022).
16. Tang, A., Wong, A. & Khurana, B. Imaging of Intimate Partner Violence, From the AJR Special Series on Emergency Radiology. *Am. J. Roentgenol.* **220**, 476–485 (2023).
17. Castillo, C., Steffens, T., Sim, L. & Caffery, L. The effect of clinical information on radiology reporting: A systematic review. *J. Med. Radiat. Sci.* **68**, 60–74 (2021).
18. Um, M. Y. et al. Predicting Intimate Partner Violence Perpetration Among Young Adults Experiencing Homelessness in Seven U.S. Cities Using Interpretable Machine Learning. *J. Interpers. Violence* **40**, 1727–1751 (2025).
19. Chen, Z. et al. Using machine learning to estimate the incidence rate of intimate partner violence. *Sci. Rep.* **13**, 5533 (2023).
20. Benjumed Wynhoven, I. M. & Lepe, C. C. Analysis of IPV success treatment from an AI approach. *PLoS One* **20**, e0323945 (2025).
21. Li, Y., Wehbe, R. M., Ahmad, F. S., Wang, H. & Luo, Y. Clinical-Longformer and Clinical-BigBird: Transformers for long clinical sequences (2022).
22. Lee, J. et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* **36**, 1234–1240 (2020).
23. Alsentzer, E. et al. Publicly Available Clinical BERT Embeddings. In *Proceedings of the 2nd Clinical Natural Language Processing Workshop 72–78* (Association for Computational Linguistics, Minneapolis, Minnesota, USA, <https://doi.org/10.18653/v1/W19-1909> 2019).
24. Luo, R. et al. BioGPT: generative pre-trained transformer for biomedical text generation and mining. *Briefings in Bioinformatics* **23**, bbac409 (2022).
25. Soenksen, L. R. et al. Integrated multimodal artificial intelligence framework for healthcare applications. *Npj Digit. Med.* **5**, 149 (2022).
26. Gosangi, B. et al. Exacerbation of Physical Intimate Partner Violence during COVID-19 Pandemic. *Radiology* **298**, E38–E45 (2021).
27. Thölke, P. et al. Class imbalance should not throw you off balance: Choosing the right classifiers and performance metrics for brain decoding with imbalanced data. *NeuroImage* **277**, 120253 (2023).
28. Zadrozny, B. & Elkan, C. Transforming Classifier Scores into Accurate Multiclass Probability Estimates. *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.* <https://doi.org/10.1145/775047.775151> (2002).
29. Lundberg, S. M. & Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems* (eds. Guyon, I. et al.) vol. 30 (Curran Associates, Inc., 2017).
30. Lundberg, S. M. et al. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2**, 56–67 (2020).
31. Lenehan, P. J. et al. Using Radiology as a Screening Tool to Identify Intimate Partner Violence. *Screening* **22**, 395–404 (2025).
32. Catallo, C., Jack, S. M., Ciliska, D. & MacMillan, H. L. Minimizing the risk of intrusion: a grounded theory of intimate partner violence disclosure in emergency departments. *J. Adv. Nurs.* **69**, 1366–1376 (2013).
33. Tarzia, L. et al. Women’s experiences and expectations after disclosure of intimate partner abuse to a healthcare provider: A qualitative meta-synthesis. *BMJ Open* **10**, e041339 (2020).
34. Feder, G. S., Hutson, M., Ramsay, J. & Taket, A. R. Women Exposed to Intimate Partner Violence: Expectations and Experiences When They Encounter Health Care Professionals: A Meta-analysis of Qualitative Studies. *Arch. Intern. Med.* **166**, 22–37 (2006).
35. Bertsimas, D., Koulouras, A. G. & Margonis, G. A. The R.O.A.D. to precision medicine. *Npj Digit. Med.* **7**, 307 (2024).
36. Robert Graham Center. Social deprivation index (SDI). *Robert Graham Center - Policy Studies in Family Medicine & Primary Care* <https://www.graham-center.org/rgc/maps-data-tools/social-deprivation-index.html> (2018).
37. U.S. Census Bureau. Income in the Past 12 Months (in 2022 Inflation-Adjusted Dollars). *American Community Survey, ACS 5-Year Estimates Subject Tables, Table S1901, 2022*, <https://data.census.gov/table/ACSST5Y2022.S1901> (2022).
38. Khurana, B., Prakash, J. & Loder, R. Assault related injury visits in US emergency departments: An analysis by weekday, month and weekday-by-month. *Chronobiol. Int.* **39**, 1–10 (2022).
39. Beltagy, I., Peters, M. E. & Cohan, A. Longformer: The Long-Document Transformer. (2020).
40. Han, T. et al. MedAlpaca – An Open-Source Collection of Medical Conversational AI Models and Training Data. (2025).
41. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM)*, <https://doi.org/10.1145/2939672.2939785> (2016).
42. Pedregosa, F. et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

## Acknowledgements

The funding from the National Institute of Biomedical Imaging and Bioengineering (NIBIB) and the Office of the Director, National Institutes of Health (1R01EB032384-01; PI: Bharti Khurana, MD, MBA) provided support for every stage of the research, including study design, data collection, analysis, interpretation, manuscript preparation, and the open access publication of the findings. We are grateful to Helen Eng, Application Analyst, and Stacey Duey, Senior Project Specialist, from the Digital Research Operations at Mass General Brigham, for their assistance with the MGB Research Patient Data Registry. We are grateful to Sarah Ingeman and Celina Li for their assistance in collecting data from the Domestic Violence Intervention and Prevention programs at Mass General Brigham.

## Author contributions

J.G. planned and performed experiments, wrote code, analyzed the data, and wrote the manuscript. K.V.C. and Y.M. planned and performed experiments, wrote code, analyzed the data, and edited the manuscript. B.K. and D.B. directed the overall research and edited the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s44294-025-00126-3>.

**Correspondence** and requests for materials should be addressed to Bharti Khurana.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026