




# Predicting Risk of Heroin Overdose, Remission, Use, and Mortality Using Ensemble Learning Methods in a Cohort of People with Heroin Dependence

Christina Marel<sup>1</sup>  · Mohammad H. Afzali<sup>2</sup> · Matthew Sunderland<sup>1</sup> · Maree Teesson<sup>1</sup> · Katherine L. Mills<sup>1</sup>

Accepted: 12 February 2024  
© The Author(s) 2024

## Abstract

Despite decades of research demonstrating the effectiveness of treatments for heroin dependence, rates of heroin use, dependence, and death have dramatically increased over the past decade. While evidence has highlighted a range of risk and protective factors for relapse, remission, and other outcomes, this presents clinicians with the challenge as to how to synthesise and integrate the evolving evidence-base to guide clinical decision-making and facilitate the provision of personalised healthcare. Using data from the 11-year follow-up of the Australian Treatment Outcome Study (ATOS), we aimed to develop a clinical risk prediction model to assist clinicians calculate the risk of a range of heroin-related outcomes at varying follow-up intervals for their clients based on known risk factors. Between 2001 and 2002, 615 people with heroin dependence were recruited as part of a prospective longitudinal cohort study. An ensemble machine learning approach was applied to predict risk of heroin use, remission, overdose, and mortality at 1-, 5-, and 10+ year post-study entry. Variables most consistently ranked in the top 10 in terms of their level of importance across outcomes included age; age first got high, used heroin, or injected; sexual trauma; years of school completed; prison history; severe mental health disability; past month criminal involvement; and past month benzodiazepine use. This study provides clinically relevant information on key risk factors associated with heroin use, remission, non-fatal overdose, and mortality among people with heroin dependence, to help guide clinical decision-making in the selection and tailoring of interventions to ensure that the ‘*right treatment*’ is delivered to the ‘*right person*’ at the ‘*right time*’.

**Keywords** Heroin dependence · Risk prediction · Outcome · Remission · Relapse · Mortality

---

✉ Christina Marel  
christina.marel@sydney.edu.au

<sup>1</sup> Matilda Centre for Research in Mental Health and Substance Use, University of Sydney, Sydney, NSW, Australia

<sup>2</sup> Department of Psychiatry, Université de Montreal, CHU Ste Justine Hospital, CHU Ste-Justine, Montreal, Canada

Increasing heroin use is a major contributor to the global opioid crisis. Rates of heroin use, dependence, and death have grown dramatically over the past decade, and the burden of disease associated with opioid use disorder is greater than any other illicit drug class (Degenhardt et al., 2013, 2019; National Institute on Drug Abuse, 2021). In the USA, national opioid-related deaths more than doubled between 2010 and 2019 (WONDER, 2020) and more than 28% of overdose deaths in 2019 involved heroin (WONDER, 2021). While the prevalence of opioid-related incidents across Europe are comparatively lower than the USA, opioid-related deaths in Scotland surged between 2009 and 2018, exceeding those in the USA (29% vs. 14% in 2018), and 45% of drug-related deaths involved heroin (van Amsterdam et al., 2021). In Australia, opioids (including heroin) have been the leading cause of drug-related deaths for the past two decades (Australian Institute of Health & Welfare, 2021b). One in four drug-induced Australian deaths in 2019 was due to heroin; the highest since 1997 (Chrzanowska et al., 2021). This global trend has continued despite decades of research demonstrating the effectiveness of a range of treatments available for heroin dependence, supported by numerous high-quality randomised-controlled trials (RCTs) and naturalistic longitudinal studies (Degenhardt et al., 2019; Hser et al., 2015; Mattick et al. 2009a; Mattick et al., 2014).

Research has also revealed a range of risk and protective factors for relapse, remission, and other outcomes (Chen et al., 2019, 2020; Darke, 2011; Mattick et al. 2009b). For example, poorer treatment outcomes have been associated with more severe heroin use, more extensive polydrug use (Chen et al., 2019; Darke et al., 2007; Hser et al., 1999), criminal involvement (Flynn et al., 2003; Strang et al., 2020), psychological distress, and poorer mental health (Hser, 2007; Schaar & Öjehagen, 2001; Strang et al., 2020). In contrast, 'treatment readiness' has been linked to improved client outcomes (Darke et al., 2005; Flynn et al., 2003). Long-term stable treatment, across fewer treatment episodes, has been associated with improved outcomes, regardless of treatment modality (Flynn et al., 2003; Gossop et al., 1999; Teesson et al., 2015), with the exception of detoxification. Due to its nature as a short-term intervention, detoxification is not associated with improved outcomes; on the contrary, repeated episodes of detoxification have been associated with poorer outcomes (Teesson et al., 2015). As with other long-term, stable treatment approaches, the evidence regarding residential rehabilitation demonstrates its efficacy across a range of outcomes (Darke et al., 2006; de Andrade et al., 2019; Gossop et al., 1999).

However, with the expansion in new knowledge comes the complex challenge for clinicians in how to synthesise and integrate the evolving evidence base to guide clinical decision-making in the provision of personalised healthcare—that is, the selection and tailoring of interventions to ensure that the '*right treatment*' is delivered to the '*right person*' at the '*right time*' (Bates, 2010; Primary Health Advisory Group 2016). There is a critical need for easy-to-use clinical tools that synthesise research evidence and provide clinically meaningful information that can be applied in busy practice settings.

Delivering the right treatment to a client based on individual characteristics and personalised decision-making maximises the likelihood of positive outcomes and reduces the costs and negative side effects associated with using inappropriate treatments (Bates, 2010). Personalised healthcare is based on a solid foundation of evidence-based clinical practice that acknowledges variability in patient response to treatments and seeks to incorporate client-specific factors as prognostic indicators (Bates, 2010). Prediction models support clinical decision-making by estimating the risk that a positive (or negative) outcome will occur within a designated timeframe, in a person with particular objective risk factors (Moons et al., 2009). They are fundamental to guiding shared-decision-making between

clinicians and healthcare consumers (Moons et al., 2009; Steyerberg, 2019), in the selection of personalised, well-timed, and targeted interventions, care coordination, and referral.

The past two decades have seen rapid growth in the development and advancement in machine learning, with > 600 risk prediction models across medical fields including outcomes related to cardiovascular disease (e.g., Damen et al., 2016), falls after stroke (e.g., Walsh et al., 2016), diabetes (e.g., Yang et al., 2021; Zhang et al., 2016), cancer (e.g., Lee et al., 2019; Usher-Smith et al., 2016), kidney injury (Wilson et al., 2016), and heart failure (e.g., Adler et al., 2020; Sahle et al., 2017). A small number of risk-prediction models have been developed in relation to opioid use; however, these have largely been restricted to examinations of the risk of overdose among people using prescription opioids for pain (Glanz et al., 2018; Klimas et al., 2019; Lo-Ciganic et al., 2019). Despite the wealth of evidence regarding risk and protective factors that contribute to the maintenance and recovery from heroin dependence, as well as other outcomes including fatal and non-fatal overdose, mortality, and co-occurring psychopathology (Chen et al., 2019, 2020; Darke et al., 2014, 2016; Davoli et al., 2007; Hser et al., 2001, 2017; Jimenez-Treviño et al., 2011; Strang et al., 2020; Teesson et al., 2017), there are currently no existing models to assist clinicians utilise and integrate this critical information into their practice and guide treatment selection.

The present study aimed to address this gap using data from the 11-year follow-up of the Australian Treatment Outcome Study (ATOS), the largest and longest prospective longitudinal study of heroin dependence conducted in Australia, with a view to developing a clinical risk prediction model to assist clinicians calculate the risk of a range of heroin-related outcomes at varying follow-up intervals for their clients, depending on baseline information. This information is critical to guiding clinicians in selecting the *right treatment*, for the *right person* at the *right time*, potentially transforming healthcare for people with heroin dependence.

## Method

### Procedure

The ATOS cohort comprised 615 people with heroin dependence. Baseline data were collected between 2001 and 2002 from 615 people with heroin dependence, 535 of whom were recruited from 19 agencies treating heroin dependence in the greater Sydney region, Australia (201 entering maintenance therapies, 201 entering detoxification, 133 entering residential rehabilitation) (Ross et al., 2005). A comparison group of 80 people who were not entering treatment were recruited from needle and syringe programs. Follow-up interviews were conducted with 549 (89.3%), 495 (80.5%), 469 (76.3%), 429 (69.8%), and 431 (70.2%) people at 3, 12, 24, 36 months and 11 years, respectively, with 94.5% completing at least one follow-up interview. A search of the Australian National Death Index, a statutory register of all deaths in Australia administered by the Australian Institute of Health and Welfare (AIHW), was conducted in July, 2015, and details of deaths that had occurred among participants since 2001 obtained. Participants were matched by full name, sex, and date of birth.

Factors associated with study retention were examined by Teesson et al (2015), and included index treatment modality, age, sex, previous treatment history, number of heroin use days in the preceding month, number of drug types used in the preceding month,

major depression, post-traumatic stress disorder (PTSD), antisocial personality disorder (ASPD), and borderline personality disorder (BPD). The final model revealed that the sample re-interviewed at 11 years was broadly representative of the initial cohort; the only independent predictors of loss to follow-up were male sex (67.8 vs. 74.5%; OR=0.65, 95% CI=0.44, 0.99) and a diagnosis of major depression at baseline (64.9 vs. 71.8%; OR=0.67, 95% CI=0.45, 0.99). ASPD was retained in the final model but was not related significantly to retention at 11 years ( $p=0.074$ ). Ethical approval was obtained from the University of New South Wales Human Research Ethics Committees and participating area health services.

## Measures

### Structured interview

Structured interviews were administered to participants at baseline and at each follow-up utilising measures with established psychometric properties and have been described in detail previously (Ross et al., 2005). In brief, baseline and follow-up interviews included demographic (age, sex, past month main source of income, and past month accommodation), treatment history, drug use history, injection-related health, physical and mental health, and psychopathology. Past-month drug use and criminal involvement were assessed using the opiate treatment index (OTI) (Darke et al., 1992). DSM-IV diagnoses of current heroin dependence and past-month major depression were obtained using the Composite International Diagnostic Interview version 2.1 (CIDI) (World Health Organisation 1993). Lifetime diagnoses of PTSD, ASPD, and BPD were assessed at baseline. DSM-IV diagnoses of ASPD were obtained using a modified version of the Diagnostic Interview Schedule (Robins et al., 1981); participants were screened for ICD-10 BPD using the International Personality Disorders Examination Questionnaire (Loranger et al., 1994); general mental health was measured using the Short Form-12 (SF-12) (Ware et al., 1996); and DSM-IV diagnoses of lifetime PTSD were obtained using the CIDI (World Health Organisation 1993). Demographics, drug use, criminal involvement, general health, and depression were reassessed at each follow-up, in addition to the number of times participants had commenced treatment for heroin dependence and the treatment duration (Teesson et al., 2008). To maximise participant recall, the time-line follow-back (TLFB) method was used with the 11-year interviews, which anchors interview questions to significant events in participants' lives (e.g., significant relationships or separations, births of children, deaths) (Hunt & Andrews, 1995). The life chart was conducted at the beginning of the interview, and participants were referred back to the chart when further questions were asked that required the recall of dates.

### Statistical analyses

Machine-learning algorithms include a set of predictive techniques, which are particularly useful when there are several predictor variables (i.e., high-dimensional) and the focus is on the generalisability and reliability of prediction. Guidelines from the National Institute on Drug Abuse (NIDA) highlight machine learning as 'the most promising approach to identify predictive markers for psychiatric disorders and classify psychiatric populations with high-dimensional data'. For the prediction modelling, data preparation included standardisation of continuous predictors and dummy coding of categorical variables. Moreover,

as several outcomes presented in the current study are considered ‘rare events’ even in a high-risk sample, an adapted strategy is necessary to address the potential imbalance among those with and without the outcome. Hence, to address the class imbalance, an oversampling strategy was implemented for each outcome. Random oversampling involves randomly duplicating examples from the minority class and adding them to the training dataset. Examples from the training dataset are selected randomly with replacement. This means that examples from the minority class can be chosen and added to the new ‘more balanced’ training dataset multiple times; they are selected from the original training dataset, added to the new training dataset, and then returned or ‘replaced’ in the original dataset, allowing them to be selected again. This technique can be effective for those machine learning algorithms that are affected by a skewed distribution and where multiple duplicate examples for a given class can influence the fit of the model. This was followed by a three-step analysis plan to establish the performance of an ensemble learning model using cross-validation and investigate the importance of each predictor variable for each outcome.

In the first step, we trained and fine tuned an ensemble machine learning model for each outcome separately. Ensemble learning is a machine learning method that produces a single prediction function based on a weighted average of all individual algorithms considered for prediction (i.e., ‘weak learners’) with an optimal trade-off between bias and variance (Pavlyshenko, 2018). The ensemble learning model used here was a combination of random forests (RF), support vector machine (SVMs), and elastic-net algorithms. Using the *tidymodels package* in *R statistical software*, a nested iterative cross-validation procedure was used to tune the hyperparameters of each model and to evaluate the predictive power of models on independent observations. The cross-validation procedure partitioned the full sample into subsamples of data, using an iterative process where 25% of data was set aside as the test sample, and a ‘training model’ was fine tuned on the tenfold partitioned observations in the remaining data (training and validation sets). This fine-tuned training model was then used, in the second step, to predict the observations in the set-aside test fold, thereby ensuring the independence of the test fold sample.

In the second step, to measure the predictive performance of each algorithm, a set of prediction indices were calculated based on the model’s ability to predict later adverse outcomes in the test set (Šimundić, 2009). The performance of the machine-learning algorithms was evaluated using *sensitivity*, *specificity*, *F1 prediction score*, and *accuracy (ACC)*. Briefly, sensitivity is defined as the number of true positives divided by the number of true positives plus false positives. Recall is defined as the number of true negatives divided by the number of true negatives plus the number of false negatives. The F1 prediction score is defined as the balance between the precision and the recall. Finally, ACC is defined as the number of true positives plus the number of true negatives divided by the total sample. For all indices, values closer to 1.0 demonstrate better predictive power. Previous cut-points were used to determine strength of prediction with acceptable ranging between 0.6 and 0.75, good ranging between 0.75 and 0.9, and excellent ranging between 0.9 and 1.

To improve the interpretability of our results, the third step focused on determining the relative feature importance of the predictive coefficients for each algorithm. For this, feature importance measures for each of the individual *weak learners*, namely RF, elastic net, and the SVM models, were extracted using the *vip package* in *R*. The Gini importance (or mean decrease impurity), a measure of information gain for each predictor, was used for the RF model, whereas the magnitude of non-shrunked coefficients was used for the elastic net model. As the SVM creates a hyperplane to maximise the distance between the predicted classes, the feature importance can be conceptualised as the vector of coordinates

orthogonal to the hyperplane. Finally, the features were ranked from the least to the most important and the mean rank of feature importance was calculated across the three *weak learners*.

### **Outcome variables: heroin use, remission, overdose, mortality**

Four outcome domains were examined: heroin use, remission from heroin use, non-fatal overdose, and mortality. The risk of experiencing these outcomes over the short (1 year), medium (5 years), and long (10 years) term was modelled for all variables (degree of balance illustrated in Table 1 of the supplementary materials). Due to the availability of data, risk of mortality was additionally modelled over 15 years (Table 1).

Past-month heroin use was defined as any heroin use occurring in the month prior to 1-year post-baseline, 5-year post-baseline, and 10-year post-baseline. Consistent with definitions of remission utilised in other areas of public health and community medicine, remission from heroin use was operationalised into three groups: short-term remission (defined as  $\geq 3$  and  $< 12$  months of continual abstinence from heroin use); medium-term remission (defined as  $\geq 12$  months and  $< 5$  years of continual abstinence from heroin use); and long-term remission (defined as  $\geq 5$  years of continual abstinence from heroin use) (Worley, 2017). Non-fatal overdose was defined as any self-reported heroin-related overdose that occurred between baseline and 1-, 5-, and 10-year post-baseline. Mortality from baseline to 1-, 5-, 10-, and 15-year post-baseline was based on data obtained from the AIHW.

### **Predictor variables**

Informed by the literature, 41 baseline predictor variables were selected and included in the analysis, and included predictors relating to demographics, psychopathology, drug use, and treatment (Table 2). With the exception of age and years of school completed which were continuous, and drug used for first high which was categorical, all variables were binary (yes/no, male/female).

## **Results**

### **Cohort characteristics**

The baseline characteristics of the ATOS cohort have been described in detail previously and are provided in Table 2 of the Supplementary Materials (Ross et al., 2005). In brief, the mean age at baseline was 29.3 years (standard deviation [SD] 7.8) and 66.2% were male. Participants reported completing a mean of 10.0 years (SD 1.7) of school education, and just under half (45.7%) reported government allowances as their main source of income. More than half (54.6%) had been involved in past month crime and 40.8% had spent time in prison. Participants reported a mean length of heroin-using career of 9.6 years (SD 7.4), and the mean number of drugs used in the past month was 9.0 (SD 1.7). There were high rates of psychopathology, with 24.6% meeting criteria for current major depression, 41.1% for lifetime PTSD, 71.5% for ASPD, and 45.5% screening positive for BPD.

**Table 1** Outcomes: heroin use, remission, overdose, and death

	Short term	Medium term	Long term
<i>Heroin use</i>			
Past-month heroin use	1-year post-baseline	5-year post-baseline	10-year post-baseline
<i>Remission</i>			
Short-term: $\geq 3$ and $< 12$ months of continual abstinence from heroin use	1-year post-baseline	5-year post-baseline	10-year post-baseline
Medium-term: $\geq 12$ months and $< 5$ years of continual abstinence from heroin use	-	5-year post-baseline	10-year post-baseline
Long-term: $\geq 5$ years continual abstinence from heroin use	-	-	10-year post-baseline
<i>Overdose</i>			
Non-fatal heroin overdose	1-year post-baseline	5-year post-baseline	10-year post-baseline
<i>Mortality</i>			
Mortality	1-year post-baseline	5-year post-baseline	10-year post-baseline 15-year post-baseline

**Table 2** Summary of predictor variables ( $n=41$ ) included in analysis

Category and predictor variables	Response options
<i>Sociodemographic</i>	
Age	Continuous
Australian born	Yes/no
Main source of income in past month government benefit	Yes/no
Past month criminal involvement	Yes/no
Prison history	Yes/no
Sex	Male/female
Unstable housing	Yes/no
Years of school completed	Continuous
<i>Psychopathology</i>	
Antisocial personality disorder (ASPD)	Yes/no
Attempted suicide in past 12 months	Yes/no
Borderline personality disorder (BPD)	Yes/no
Experienced interpersonal trauma	Yes/no
Major depression	Yes/no
Post-traumatic stress disorder (PTSD)	Yes/no
Severe mental health disability	Yes/no
Severe physical health disability	Yes/no
Experienced sexual trauma	Yes/no
Experienced trauma	Yes/no
<i>AOD use</i>	
First got high > 13 years old	Yes/no
Age when first used heroin < 17	Yes/no
Drug used for first high	Categorical
Ever overdosed	Yes/no
First injected > 17 years old	Yes/no
Overdosed past 12 months	Yes/no
Past-month alcohol use	Yes/no
Past-month amphetamine use	Yes/no
Past-month antidepressant use	Yes/no
Past-month benzodiazepine use	Yes/no
Past-month cannabis use	Yes/no
Past-month cocaine use	Yes/no
Past-month daily heroin use	Yes/no
Past-month hallucinogen use	Yes/no
Past-month other opioid use	Yes/no
Past-month tobacco use	Yes/no
Polysubstance use	Yes/no
<i>AOD treatment</i>	
Detoxification	Yes/no
Methadone maintenance therapy	Yes/no
No treatment	Yes/no
Previous treatment	Yes/no
Residential rehabilitation	Yes/no



## Model performance

The model performance for all outcomes is illustrated in Table 3 (area under the curves [AUC] and receiver operating characteristic curves [ROC] are provided in the Supplementary Materials). For the past-month heroin use outcome, the ensemble learning model showed an acceptable performance (ACC=0.73, F1=0.76) for 1-year post-baseline, a good performance (ACC=0.83, F1=0.84) for 5-year post-baseline, and a good performance (ACC=0.85, F1=0.86) for 10-year post-baseline. For the short-term remission outcome, the ensemble learning model showed a good performance (ACC=0.87, F1=0.87) for 1-year post-baseline, an excellent performance (ACC=0.91, F1=0.91) for 5-year post baseline, and a good performance (ACC=0.88, F1=0.89) for 10-year post baseline. For the medium-term remission outcome, the ensemble learning model showed an acceptable performance (ACC=0.75, F1=0.78) for 5-year post-baseline and a good performance (ACC=0.82, F1=0.85) for 10-year post-baseline. For the long-term remission outcome, the ensemble learning model showed an acceptable performance (ACC=0.75, F1=0.75) for 10-year post baseline. For the overdose outcome, the ensemble learning model showed an excellent performance (ACC=0.91, F1=0.91) for 1-year post-baseline, a good

**Table 3** Details of the model performance measures for each outcome<sup>†</sup>

Outcome	Measure	1 year	5 years	10 years	15 years
Past-month heroin use	Accuracy	0.73	0.83	0.85	
	Sens	0.86	0.83	0.94	
	Spec	0.62	0.82	0.76	
	F1	0.76	0.84	0.86	
Remission (short term)	Accuracy	0.87	0.91	0.88	
	Sens	0.90	0.90	0.95	
	Spec	0.84	0.91	0.81	
	F1	0.87	0.91	0.89	
Remission (medium term)	Accuracy	-	0.75	0.82	
	Sens		0.83	0.98	
	Spec		0.67	0.64	
	F1		0.78	0.85	
Remission (long term)	Accuracy	-	-	0.75	
	Sens			0.78	
	Spec			0.72	
	F1			0.75	
Overdose	Accuracy	0.91	0.86	0.76	
	Sens	0.87	0.85	0.78	
	Spec	0.96	0.88	0.75	
	F1	0.91	0.86	0.77	
Mortality	Accuracy	-	-	0.85	0.89
	Sens			0.87	0.89
	Spec			0.83	0.90
	F1			0.85	0.89

<sup>†</sup>Acceptable strength of prediction: between 0.6 and 0.75; good: between 0.75 and 0.9; excellent: between 0.9 and 1

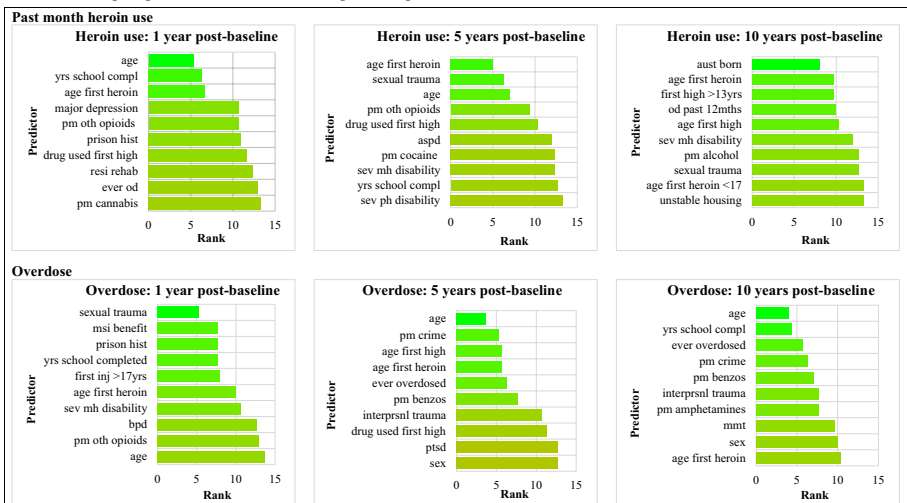
performance (ACC=0.86, F1=0.86) for 5-year post-baseline, and an acceptable performance (ACC=0.76, F1=0.77) for 10-year post-baseline. For the mortality outcome, the ensemble learning model showed a good performance (ACC=0.85, F1=0.85) for 10-year post baseline and a good performance (ACC=0.89, F1=0.89) for 15-year post-baseline.

### Feature importance

The top 10 most important predictors for all outcomes identified by the ensemble learning model are shown in Figs. 1, 2, and 3. As illustrated, for heroin use occurring in the month prior to 1-, 5-, and 10-year post-baseline, the feature importance analysis highlighted the role of being Australian born, age, age when first got high, daily use of heroin (past month), and overdosed in past 12 months at baseline. Similarly, for short-term remission occurring 1-, 5-, and 10-year post-baseline, the feature importance analysis identified age, age of first heroin use, age when first got high, receiving a government benefit as main source of income, past month cocaine use, past month criminal involvement, and prison history. For medium-term remission occurring 5- and 10-year post-baseline, the feature importance analysis highlighted the important role of age first used heroin, drug used for first high, age when first experienced a high, and residential rehabilitation. Unstable housing and age when first used heroin were identified as prominent features associated with long-term remission.

For non-fatal heroin-related overdose that occurred up to 1-, 5-, and 10-year post-baseline, the feature importance analysis identified age, age when first used heroin,

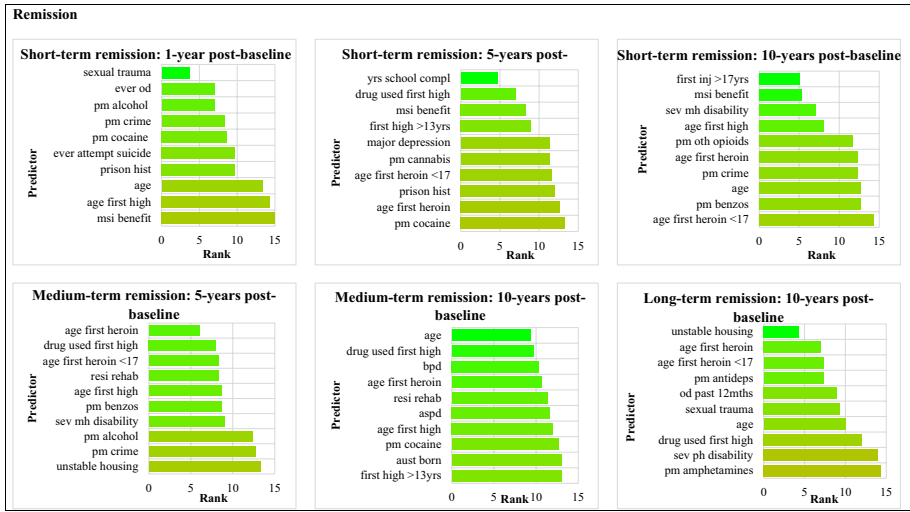
Panel 1: Top 10 performance and feature importance predictors of heroin use<sup>a</sup> and overdose<sup>b</sup>



<sup>a</sup>Past month heroin use defined as any heroin use occurring in the month prior to 1-, 5- and 10-years post-baseline. <sup>b</sup>Non-fatal overdose defined as any self-reported heroin-related overdose that occurred up to 1-, 5- and 10-years post-baseline. Age = age at baseline; yrs school completed = years of school completed; age first heroin = age when first used heroin; major depression = past month major depression; pm oth opioids = past month other opioid use; prison hist = prison history at baseline; drug used first high = drug used for first high; resi rehab = in residential rehabilitation; ever od = ever overdosed; pm cannabis = past month cannabis use; age first heroin = age when first used heroin; sexual trauma = experienced sexual trauma; aspd = antisocial personality disorder; pm cocaine = past month cocaine use; sev mh disability = severe mental health disability; sev ph disability = severe physical health disability; aust born = born in Australian; first high >13yrs = first got high when aged over 13 years; od past 12mths = overdosed in the past 12 months; pm alcohol = past month alcohol use; age first heroin <17 = aged less than 17 years when first used heroin; msi benefit = received government benefit as main source of income; first inj >17yrs = aged over 17 years when first injected; age first heroin = age when first used heroin; bpd = borderline personality disorder; intersprnl trauma = experienced interpersonal trauma; ptsd = post traumatic stress disorder; pm benzos = past month benzodiazepine use; pm amphetamines = past month amphetamine use; mmt = in methadone maintenance therapy.

Fig. 1 Top 10 performance and feature importance predictors of heroin use<sup>a</sup> and overdose<sup>b</sup>

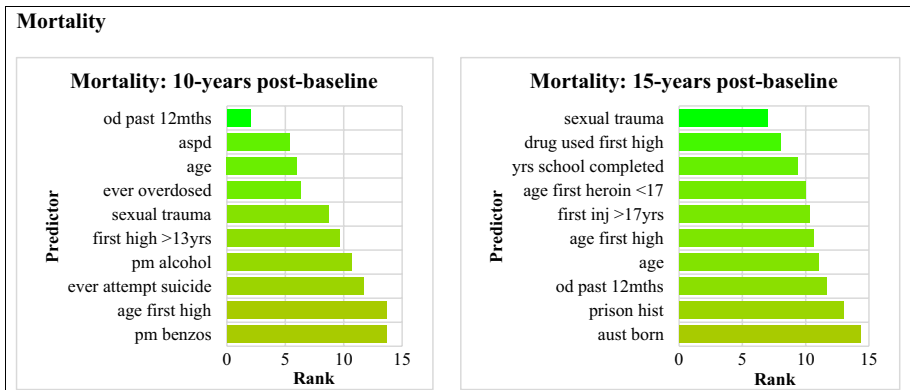
Panel 2: Top 10 performance and feature importance predictors of remission<sup>e</sup>



<sup>e</sup>Short-term remission was defined as  $\geq 3$  and <12 months of continual abstinence from heroin use; medium-term remission was defined as  $\geq 12$  months and <5 years of continual abstinence from heroin use; and long-term remission was defined as  $\geq 5$  years of continual abstinence from heroin use. As with heroin use, remission was modelled over the short (1-year), medium (5-years) and long (10-years) term. Pm crime = involved in past month criminal activity; pm cannabis = past month cannabis use; age first high = age when first experienced a high; pm cocaine = past month cocaine use; pm antidepressants = past month antidepressant use.

Fig. 2 Top 10 performance and feature importance predictors of remission<sup>e</sup>

Panel 3: Top 10 performance and feature importance predictors of mortality<sup>e</sup>



<sup>e</sup>Mortality based on data obtained from the AIHW, and modelled up to 1-, 5-, 10- and 15-years post-baseline.

Fig. 3 Top 10 performance and feature importance predictors of mortality<sup>e</sup>

overdose history, experienced interpersonal trauma, past-month use of benzodiazepines, past-month criminal involvement, sex, and years of school completed.

In regard to mortality, the feature importance analysis highlighted age, age of first high, overdosed in the 12 months prior to baseline, and experienced sexual trauma. More detailed graphs illustrating the performance of all predictors in the model across all outcomes are included in the Supplementary Materials.

## Discussion

Using 11-year follow-up data from ATOS, the largest and longest prospective longitudinal study of heroin dependence conducted in Australia, this study developed machine learning models to predict the risk of heroin use, remission from heroin use, non-fatal heroin-related overdose, and mortality, at varying follow-up intervals. Despite decades of research and significant economic investment, high levels of health service utilisation (Australian Institute of Health & Welfare, 2021a; Darke et al., 2015), and a large body of research regarding evidence-based treatment (Degenhardt et al., 2019; Hser et al., 2015; Mattick et al. 2009a; Mattick et al., 2014), many people with heroin dependence do not receive the help they need.

One of the complex challenges for clinicians is the synthesis and integration of the rapidly evolving evidence base into information that can guide clinical decision-making, yet many clinicians lack time, resources, guidance, or procedures on how to best synthesise, analyse, or utilise this information in practice to deliver personalised care (DeRubeis et al., 2014; Grol & Grimshaw, 2003; Oxman & Flottorp, 2001). Easy-to-use clinical tools can synthesise vast amounts of research evidence and provide meaningful information to guide clinicians in selecting the *right treatment*, for the *right person* at the *right time*.

This study builds on the strengths of the stacking approach to predictive modelling, which is combining different predictive models into one ensemble model. Ensemble modelling is a recently developed yet a well-established and widely used technique to improve model performance that enables averaging out modelling error from diverse models and thereby enhances the generalisable signal detection (Gunes et al., 2017). The model performances across all outcomes over all follow-up intervals were predominantly good. The exceptions were past-month heroin use at 1-year post-baseline, which was of acceptable performance; short-term remission that occurred 5-year post-baseline, and non-fatal overdose that occurred 1-year post-baseline, which were both of excellent performance. The stacking ensemble learning approach combines predictions from multiple machine learning algorithms and uses these predictions as inputs to a higher-level learning model. Recent studies highlighted that stacking-based predictive modelling is highly effective for time series forecasting and classification problems with highly imbalanced data (Pavlyshenko, 2018).

A strength of the current study is the ability to examine a wide range of predictors. The variables most consistently ranked in the top 10 in terms of their level of importance across outcomes included age; age first got high, used heroin, or injected; sexual trauma; years of school completed; prison history; severe mental health disability; past-month crime; and past-month benzodiazepine use. Somewhat surprisingly given the consistent association of major depression and poor outcome in analysis of the ATOS data and other longitudinal studies of heroin dependence (Malik et al., 2019; Marel et al., 2023), severe mental health disability and the experience of sexual trauma were the only psychiatric variables identified among the most predictive factors in this study. In contrast, the absence of a mood disorder was found to reduce the likelihood of prescription opioid use disorder in a recent systematic review among those commencing prescription opioids for pain (Klimas et al., 2019).

These findings are of particular clinical importance for identifying those most at risk across a number of outcome domains, with implications for prevention and early intervention (for example, school-based programs aimed at delaying age of onset (Gardner et al., 2023), early intervention for people who have experienced sexual trauma).

Findings from this study are broadly consistent with the few studies that have undertaken similar modelling procedures to assess opioid-related risk. The majority of this work has been undertaken to estimate non-fatal and/or fatal overdose risk, using administrative and linked datasets to develop models, such as Medicare or prescription claims (Chang et al., 2019; Heo et al., 2022; Lo-Ciganic et al., 2019), insurance plans (Sun et al., 2020), pharmacy-based opioid therapy (Glanz et al., 2018), prescription claims (Chang et al., 2019; Heo et al., 2022), hospital or public health services (Nguyen et al., 2023; Ripperger et al., 2021; Saloner et al., 2020), the criminal justice system (Ripperger et al., 2021), and death certificates (Nguyen et al., 2023; Saloner et al., 2020). Age, mental health disorder, substance abuse/dependence, tobacco use/abuse/dependence, and long-acting or extended-release opioid formulation were associated with the risk of two-year fatal and non-fatal opioid overdose in a US study utilising data from patients enrolled in pharmacy-based opioid therapy (Glanz et al., 2018). Similarly, being male, younger, and a greater number of benzodiazepine prescriptions were associated with higher odds of non-fatal opioid overdose in a retrospective cohort study, which utilised prescription and patient data in Maryland (Chang et al., 2019). However, the exclusive use of administrative data in these studies prevented the capturing of non-fatal overdose that did not result in hospitalisation, and the risk of non-prescribed opioid use on overdose was not included.

To our knowledge, this is the first study to predict a range of heroin-related outcomes across different follow-up intervals. While recent advances in modelling and machine learning have led to the development of models to assess single outcomes associated with the use of opioids, these have predominantly been limited to prescription opioid use (Glanz et al., 2018), chronic pain-care patients (Klimas et al., 2019), or exclusively based on administrative data (Lo-Ciganic et al., 2019). While these studies have been able to generate large sample sizes from constructed cohorts and incorporate other health outcomes depending on the data sources, the exclusive use of administrative data can be problematic. The sole reliance on administrative data for investigating outcomes associated with highly disadvantaged and vulnerable population groups is likely to result in systematic bias, particularly when data are dependent on self-disclosure in environments that are typically highly stigmatising and marginalising toward people with mental and substance use disorders (Calderwood & Lessof, 2009; Marel & Mills, 2018). These cohorts are also subject to Berkson's bias (Westreich, 2012), whereby the combination of exposure to a risk (e.g., heroin dependence) and occurrence of disease make it more likely that an individual will present to health services, leading to elevated associations between the exposure and the health outcomes under investigation relative to that which would be found among the broader population of people with heroin dependence.

The findings from this study should be interpreted with several caveats in mind. Firstly, the study was based on self-report measures. While there has been debate surrounding self-report as a tool for measuring substance use, there has been extensive literature consistently demonstrating its validity and reliability in research settings (Jackson et al., 2005; Napper et al., 2010; Teesson et al., 2015). However, participants were asked to report on events up to a 10- to 11-year period, and findings may be subject to recall bias. Secondly, while the demographic characteristics and substance use histories of the ATOS cohort were consistent with previous international and Australian studies of people with heroin dependence (Darke et al., 2002; Gossop et al., 2000; Hubbard et al., 1997), care should be taken in generalising these results to other treatment systems outside Australia. Despite the strengths of using the modelling techniques in this study, one limitation is that direction of effects cannot be attributed to individual variables as they are derived based on two- to three-level interactions across the myriad

variables and averaged over three models in the ensemble. As such the ‘direction’ of each predictor cannot be ascertained independently of the numerous other predictors. In addition, the current study used between-group data to assume individual level predictions. While these results might provide some indication of outcomes for at-risk individuals determined at one time point, they do not reveal the potential impact of time-varying covariates and whether within-person trajectories might result in differing outcomes or more accurate predictions. Furthermore, variables used in this model were examined as static predictors. Future research should examine the effects of time-varying risk and protective factors in model development that may further inform the timing of interventions.

In spite of these limitations, the current study provides clinically relevant information on key risk factors associated with heroin use, remission, non-fatal overdose, and mortality among people with heroin dependence. The ability to identify such clinically meaningful outcomes has vital implications for healthcare practitioners and policymakers who lack access to accurate information on which to base decisions on the delivery and timing of targeted interventions.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11469-024-01257-5>.

**Acknowledgements** CM, MT, and KLM received NHMRC fellowships. The authors thank all the ATOS study participants, as well as the NSW Ministry of Health, the NSW Pharmacy Guild, and the Australian Institute of Health and Welfare. We would also like to thank the ATOS research assistants, Joanne Bennett (White), Sarah Ellis, Phillipa Ewer, Sandra Fairbairn, Evelyn Harvey (Wilhelm), Alys Havard, Kate Hetherington, Sonja Memedovic, Louise Mewton, and Anna Williamson.

**Author contribution** CM conceptualised, designed the study, and drafted the manuscript. MHA and MS contributed to the design, carried out the analyses, critically reviewed the results, and revised the manuscript. MT critically reviewed the results and contributed to manuscript revision. KLM contributed to the analyses, reviewing of results, and manuscript drafting and revision. All authors approved the final manuscript as submitted.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions. The study was funded by the Australian NHMRC.

**Data availability** Data from this study are currently unavailable.

## Declarations

**Consent to participate** All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000. Informed consent was obtained from all patients for being included in the study. Ethical approval for the 11-year follow-up was obtained from the University of New South Wales Human Research Ethics Committees (HC10203).

**Conflict of interest** The study was funded by an Australian National Health and Medical Research Council (NHMRC) Project Grant, with NHMRC Fellowships to Christina Marel, Maree Teesson, and Katherine Mills, and supported by Matilda Centre funding.

**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/>.

## References

- Adler, E. D., Voors, A. A., Klein, L., Macheret, F., Braun, O. O., Urey, M. A., Zhu, W., Sama, I., Tadel, M., Campagnari, C., Greenberg, B., & Yagil, A. (2020). Improving risk prediction in heart failure using machine learning. *European Journal of Heart Failure*, 22(1), 139–147.
- Australian Institute of Health and Welfare (2021a) *Alcohol and other drug treatment services in Australia annual report*, AIHW, Australian Government. Accessed 18 Feb 2023
- Australian Institute of Health and Welfare (2021b) *Alcohol, tobacco & other drugs in Australia*, AIHW, Australian Government. Accessed 20 Feb 2023
- Bates, S. (2010). Progress towards personalized medicine. *Drug Discovery Today*, 15(3–4), 115–120. <https://doi.org/10.1016/j.drudis.2009.11.001>
- Calderwood, L., & Lessof, C. (2009). Enhancing longitudinal surveys by linking to administrative data. In P. Lynn (Ed.), *Methodology of Longitudinal Surveys* (pp. 55–71). John Wiley & Sons.
- Chang, H. Y., Krawczyk, N., Schneider, K. E., Ferris, L., Eisenberg, M., Richards, T. M., Lyons, B. C., Jackson, K., Weiner, J. P., & Saloner, B. (2019). A predictive risk model for nonfatal opioid overdose in a statewide population of buprenorphine patients. *Drug and Alcohol Dependence*, 201, 127–133.
- Chen, T., Du, J., Zhong, N., Zhao, Y., Sun, H., Chen, Z., Jiang, H., & Zhao, M. (2020). Trajectories of heroin use predict relapse risk among heroin-dependent patients: A 5-year follow-up study. *Journal of Clinical Neuroscience*, 76, 134–139.
- Chen, T., Zhong, N., Du, J., Li, Z., Zhao, Y., Sun, H., Chen, Z., Jiang, H., & Zhao, M. (2019). Polydrug use patterns and their impact on relapse among heroin-dependent patients in Shanghai China. *Addiction*, 114(2), 259–267.
- Chrzanowska, A., Man, N., Sutherland, R., Degenhardt, L., & Peacock, A. (2021). *Trends in Drug-Induced Deaths in Australia, 1997–2019*. <https://doi.org/10.26190/g2bk-t998>
- Damen, J. A. A. G., Hooft, L., Schuit, E., Debray, T. P. A., Collins, G. S., Tzoulaki, I., Lassale, C. M., Siontis, G. C. M., Chiochia, V., Roberts, C., Schlüssel, M. M., Gerry, S., Black, J. A., Heus, P., Van Der Schouw, Y. T., Peelen, L. M., & Moons, K. G. M. (2016). Prediction models for cardiovascular disease risk in the general population: Systematic review. *BMJ (Online)*, 353. <https://doi.org/10.1136/bmj.i2416>
- Darke, S. (2011). *The Life of the Heroin User: Typical Beginnings*. Cambridge University Press.
- Darke, S., Hall, W., Wodak, A., Heather, N., & Ward, J. (1992). Development and validation of a multidimensional instrument for assessing outcome of treatment among opiate users: The opiate treatment index. *Addiction*, 87(5), 733–742.
- Darke, S., Marel, C., Mills, K. L., Ross, J., Slade, T., Burns, L., & Teesson, M. (2014). Patterns and correlates of non-fatal heroin overdose at 11-year follow-up: Findings from the Australian Treatment Outcome Study. *Drug and Alcohol Dependence*, 144, 148–152. <https://doi.org/10.1016/j.drugalcdep.2014.09.001>
- Darke, S., Marel, C., Mills, K. L., Ross, J., Slade, T., & Teesson, M. (2016). Years of potential life lost amongst heroin users in the Australian Treatment Outcome Study cohort, 2001–2015. *Drug and Alcohol Dependence*, 162, 206–210. <https://doi.org/10.1016/j.drugalcdep.2016.03.010>
- Darke, S., Marel, C., Ross, J., Mills, K. L., Slade, T., & Teesson, M. (2015). Health service utilization among heroin users: 11-year follow-up of the Australian Treatment Outcome Study Cohort. *Addictive Disorders and Their Treatment*, 14(3), 159–166. <https://doi.org/10.1097/ADT.0000000000000069>
- Darke, S., Ross, J., Mills, K. L., Williamson, A., Havad, A., & Teesson, M. (2007). Patterns of sustained heroin abstinence amongst long-term, dependent heroin users: 36 months findings from the Australian Treatment Outcome Study (ATOS). *Addictive Behaviors*, 32(9), 1897–1906. <https://doi.org/10.1016/J.ADDBEH.2007.01.014>
- Darke, S., Ross, J., Teesson, M., Ali, R., Cooke, R., Ritter, A., & Lynskey, M. (2005). Factors associated with 12 months continuous heroin abstinence: Findings from the Australian Treatment Outcome Study (ATOS). *Journal of Substance Abuse Treatment*, 28(3), 255–263. <https://doi.org/10.1016/J.JSAT.2005.01.006>
- Darke, S., Topp, L., Kaye, S., & Hall, W. (2002). Heroin use in New South Wales, Australia, 1996–2000: 5 year monitoring of trends in price, purity, availability and use from the Illicit Drug Reporting System (IDRS). *Addiction*, 97(2), 179–186. <https://doi.org/10.1046/J.1360-0443.2002.00032.X>

- Darke, S., Williamson, A., Ross, J., & Teesson, M. (2006). Residential rehabilitation for the treatment of heroin dependence: Sustained heroin abstinence and drug-related problems 2 years after treatment entrance. *Addictive Disorders and Their Treatment*, 5(1), 9–18. <https://doi.org/10.1097/01.ADT.0000210702.63165.81>
- Davoli, M., Bargagli, A. M., Perucci, C. A., Schifano, P., Belleudi, V., Hickman, M., Salamina, G., Diecidue, R., Vigna-Taglianti, F., & Faggiano, F. (2007). Risk of fatal overdose during and after specialist drug treatment: The VEdeTTE study, a national multi-site prospective cohort study. *Addiction*, 102(12), 1954–1959. <https://doi.org/10.1111/j.1360-0443.2007.02025.x>
- de Andrade, D., Elphinston, R. A., Quinn, C., Allan, J., & Hides, L. (2019). The effectiveness of residential treatment services for individuals with substance use disorders: A systematic review. *Drug and Alcohol Dependence*, 201, 227–235.
- Degenhardt, L., Grebely, J., Stone, J., Hickman, M., Vickerman, P., Marshall, B. D. L., Bruneau, J., Altice, F. L., Henderson, G., Rahimi-Movaghar, A., & Larney, S. (2019). Global patterns of opioid use and dependence: Harms to populations, interventions, and future action. *The Lancet*, 394(10208), 1560–1579. [https://doi.org/10.1016/S0140-6736\(19\)32229-9](https://doi.org/10.1016/S0140-6736(19)32229-9)
- Degenhardt, L., Whiteford, H. A., Ferrari, A. J., Baxter, A. J., Charlson, F. J., Hall, W. D., Freedman, G., Burstein, R., Johns, N., Engell, R. E., Flaxman, A., Murray, C. J. L., & Vos, T. (2013). Global burden of disease attributable to illicit drug use and dependence: Findings from the Global Burden of Disease Study 2010. *The Lancet*, 382(9904), 1564–1574. [https://doi.org/10.1016/S0140-6736\(13\)61530-5](https://doi.org/10.1016/S0140-6736(13)61530-5)
- DeRubeis, R. J., Cohen, Z. D., Forand, N. R., Fournier, J. C., Gelfand, L. A., & Lorenzo-Luaces, L. (2014). The personalized advantage index: Translating research on prediction into individualized treatment recommendations A Demonstration. *Plos ONE*, 9(1), e83875. <https://doi.org/10.1371/journal.pone.0083875>
- Flynn, P. M., Joe, G. W., Broome, K. M., Simpson, D. D., & Brown, B. S. (2003). Recovery from opioid addiction in DATOS. *Journal of Substance Abuse Treatment*, 25(3), 177–186. [https://doi.org/10.1016/S0740-5472\(03\)00125-9](https://doi.org/10.1016/S0740-5472(03)00125-9)
- Gardner, L. A., Stockings, E., Champion, K. E., Mather, M., & Newton, N. C. (2023). Alcohol initiation before age 15 predicts earlier hazardous drinking: A survival analysis of a 7-year prospective longitudinal cohort of Australian adolescents. *Addiction*. <https://doi.org/10.1111/add.16376>
- Glanz, J. M., Narwaney, K. J., Mueller, S. R., Gardner, E. M., Calcaterra, S. L., Xu, S., Breslin, K., & Binswanger, I. A. (2018). Prediction model for two-year risk of opioid overdose among patients prescribed chronic opioid therapy. *Journal of General Internal Medicine*, 33(10), 1–8. <https://doi.org/10.1007/S11606-017-4288-3/TABLES/5>
- Gossop, M., Marsden, J., Stewart, D., & Rolfe, A. (1999). Treatment retention and 1 year outcomes for residential programmes in England. *Drug and Alcohol Dependence*, 57(2), 89–98. [https://doi.org/10.1016/S0376-8716\(99\)00086-1](https://doi.org/10.1016/S0376-8716(99)00086-1)
- Gossop, M., Marsden, J., Stewart, D., & Rolfe, A. (2000). Patterns of improvement after methadone treatment: 1 year follow-up results from the National Treatment Outcome Research Study (NTORS). *Drug and Alcohol Dependence*, 60(3), 275–286. [https://doi.org/10.1016/S0376-8716\(00\)00109-5](https://doi.org/10.1016/S0376-8716(00)00109-5)
- Grol, R., & Grimshaw, J. (2003). From best evidence to best practice: Effective implementation of change in patients' care. In *Lancet* (Vol. 362, Issue 9391, pp. 1225–1230). Elsevier. [https://doi.org/10.1016/S0140-6736\(03\)14546-1](https://doi.org/10.1016/S0140-6736(03)14546-1)
- Gunes F, Wolfinger R, Tan P-Y (2017) *Stacked ensemble models for improved prediction accuracy*. Presented at the SAS Global Forum Conference, Cary, NC: SAS Institute Inc
- Heo KN, Lee JY, Ah YM (2022) Development and validation of a risk-score model for opioid overdose using a national claims database. *Sci Rep*. 12(1):4974. <https://doi.org/10.1038/s41598-022-09095-y>
- Hser, Y. I. (2007). Predicting long-term stable recovery from heroin addiction: Findings from a 33-year follow-up study. *Journal of Addictive Diseases*, 26(1), 51–60. [https://doi.org/10.1300/J069v26n01\\_07](https://doi.org/10.1300/J069v26n01_07)
- Hser, Y.-I., Evans, E., Grella, C., Ling, W., & Anglin, D. (2015). Long-term course of opioid addiction. *Harvard Review of Psychiatry*, 23(2), 76–89.
- Hser, Y. I., Grella, C. E., Hsieh, S. C., Anglin, M. D., & Brown, B. S. (1999). Prior treatment experience related to process and outcomes in DATOS. *Drug and Alcohol Dependence*, 57(2), 137–150. [https://doi.org/10.1016/S0376-8716\(99\)00081-2](https://doi.org/10.1016/S0376-8716(99)00081-2)
- Hser, Y.-I., Hoffman, V., Grella, C. E., & Anglin, M. D. (2001). A 33-year follow-up of narcotics addicts. *Archives of General Psychiatry*, 58(5), 503–508. <https://doi.org/10.1001/ARCHPSYC.58.5.503>
- Hser, Y.-I., Huang, D., Saxon, A. J., Woody, G., Moskowitz, A. L., Matthews, A. G., & Ling, W. (2017). Distinctive trajectories of opioid use over an extended follow-up of patients in a multi-site trial on buprenorphine+naloxone and methadone. *Journal of Addiction Medicine*, 11(1), 63–69.



- Hubbard, R. L., Craddock, S. G., Flynn, P. M., Anderson, J., & Etheridge, R. M. (1997). Overview of 1-year follow-up outcomes in the Drug Abuse Treatment Outcome Study (DATOS). *Psychology of Addictive Behaviors, 11*(4), 261–278. <https://doi.org/10.1037/0893-164X.11.4.261>
- Hunt, C., & Andrews, G. (1995). Comorbidity in the anxiety disorders: The use of a life-chart approach. *Journal of Psychiatric Research, 29*(6), 467–480.
- Jackson, C. T., Covell, N. H., Frisman, L. K., & Essock, S. M. (2005). Validity of self-reported drug use among people with co-occurring mental health and substance use disorders. *Journal of Dual Diagnosis, 1*(1), 49–63. [https://doi.org/10.1300/J374V01N01\\_05](https://doi.org/10.1300/J374V01N01_05)
- Jimenez-Treviño, L., Saiz, P. A., García-Portilla, M. P., Díaz-Mesa, E. M., Sánchez-Lasheras, F., Burón, P., Casares, M. J., Marina, P., Gutiérrez, E., & Bobes, J. (2011). A 25-year follow-up of patients admitted to methadone treatment for the first time: Mortality and gender differences. *Addictive Behaviors, 36*(12), 1184–1190. <https://doi.org/10.1016/j.addbeh.2011.07.019>
- Klimas, J., Gorfinkel, L., Fairbairn, N., Amato, L., Ahamad, K., Nolan, S., Simel, D. L., & Wood, E. (2019). Strategies to identify patient risks of prescription opioid addiction when initiating opioids for pain: A systematic review. *JAMA Network Open, 2*(5), e193365. <https://doi.org/10.1001/JAMANETWORKOPEN.2019.3365>
- Lee, A., Mavaddat, N., Wilcox, A. N., Cunningham, A. P., Carver, T., Hartley, S., Babb de Villiers, C., Izquierdo, A., Simard, J., Schmidt, M. K., Walter, F. M., Chatterjee, N., Garcia-Closas, M., Tischkowitz, M., Pharoah, P., Easton, D. F., & Antoniou, A. C. (2019). BOADICEA: A comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genetics in Medicine, 21*(8), 1708–1718.
- Lo-Ciganic, W. H., Huang, J. L., Zhang, H. H., Weiss, J. C., Wu, Y., Kwoh, C. K., Donohue, J. M., Cochran, G., Gordon, A. J., Malone, D. C., Kuza, C. C., & Gellad, W. F. (2019). Evaluation of machine-learning algorithms for predicting opioid overdose risk among medicare beneficiaries with opioid prescriptions. *JAMA Network Open, 2*(3), e190968–e190968. <https://doi.org/10.1001/JAMANETWORKOPEN.2019.0968>
- Loranger, A. W., Sartorius, N., Andreoli, A., Berger, P., Buchheim, P., Channabasavanna, S. M., Coid, B., Dahl, A., Diekstra, R. F. W., Ferguson, B., Jacobsberg, L. B., Mombour, W., Pull, C., Ono, Y., & Regier, D. A. (1994). The International Personality Disorder Examination: The World Health Organization/Alcohol, Drug Abuse, and Mental Health Administration International Pilot Study of Personality Disorders. *Archives of General Psychiatry, 51*(3), 215–224.
- Malik, E., Adelson, M., Sason, A., Schreiber, S., & Peles, E. (2019). Outcome of patients with high depressive symptoms on admission to methadone maintenance treatment. *Journal of Dual Diagnosis, 15*(4), 281–290.
- Marel, C., & Mills, K. L. (2018). *Current practices and support needs of healthcare providers in relation to comorbidity in CESPNN in relation to addressing patients' co-occurring mental health and alcohol and other drug issues*. National Drug and Alcohol Research Centre Technical Report No. 338, UNSW.
- Marel, C., Wilson, J., Darke, S., Ross, J., Slade, T., Haber, P. S., Haasnoot, K., Visontay, R., Keaveny, M., Tremonti, C., Mills, K. L., & Teesson, M. (2023). Patterns and predictors of heroin use, remission, and psychiatric health among people with heroin dependence: Key findings from the 18–20-year follow-up of the Australian Treatment Outcome Study (ATOS). *International Journal of Mental Health and Addiction. https://doi.org/10.1007/S11469-022-01006-6*
- Mattick, R. P., Ali, R., & Lintzeris, N. (2009a). Pharmacotherapies for the treatment of opioid dependence: Efficacy, cost-effectiveness and implementation guidelines. In *Pharmacotherapies for the Treatment of Opioid Dependence: Efficacy, Cost-Effectiveness and Implementation Guidelines*. New York: Informa Healthcare; Boca Raton: CRC Press. <https://doi.org/10.3109/9780203414088>
- Mattick, R. P., Breen, C., Kimber, J., & Davoli, M. (2009b). Methadone maintenance therapy versus no opioid replacement therapy for opioid dependence. *Cochrane Database of Systematic Reviews, 3*. <https://doi.org/10.1002/14651858.CD002209.PUB2>
- Mattick, R. P., Breen, C., Kimber, J., & Davoli, M. (2014). Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. *Cochrane Database of Systematic Reviews, 2*. <https://doi.org/10.1002/14651858.CD002207.PUB4>
- Moons, K. G. M., Royston, P., Vergouwe, Y., Grobbee, D. E., & Altman, D. G. (2009). Prognosis and prognostic research: What, why, and how? In *BMJ (Online)* (Vol. 338, Issue 7706, pp. 1317–1320). British Medical Journal Publishing Group. <https://doi.org/10.1136/bmj.b375>
- Napper, L. E., Fisher, D. G., Johnson, M. E., & Wood, M. M. (2010). The reliability and validity of drug users' self reports of amphetamine use among primarily heroin and cocaine users. *Addictive Behaviors, 35*(4), 350–354. <https://doi.org/10.1016/J.ADDBEH.2009.12.006>

- National Institute on Drug Abuse. (2021). Overdose death rates: Trends and statistics. <https://www.drugabuse.gov/drug-topics/trends-statistics/overdose-death-rates>. Accessed 20 Feb 2023.
- Nguyen, A. P., Glanz, J. M., Narwaney, K. J., Zeng, C., Wright, L., Fairbairn, L. M., & Binswanger, I. A. (2023). Update of a multivariable opioid overdose risk prediction model to enhance clinical care for long-term opioid therapy patients. *Journal of General Internal Medicine*, 2023, 1–8.
- Oxman, A., & Flottorp, S. (2001). An overview of strategies to promote implementation of evidence-based health care. In C. Silagy & A. Haines (Eds.), *Evidence-based practice in primary care* (2nd ed., p. 101). BMJ Books.
- Pavlyshenko, B. (2018). Using stacking approaches for machine learning models. In *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)* (pp. 255–258). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/DSMP.2018.8478522>
- Primary Health Advisory Group. (2016). *Better outcomes for people with chronic and complex health conditions*. Australian Government Department of Health. <https://apo.gov.au/node/62407>
- Ripperger, M., Lotspeich, S. C., Wilimitis, D., Fry, C. E., Roberts, A., Lenert, M., Cherry, C., Latham, S., Robinson, K., Chen, Q., Mcpheeters, M. L., Tyndall, B., & Walsh, C. G. (2021). Ensemble learning to predict opioid-related overdose using statewide prescription drug monitoring program and hospital discharge data in the state of Tennessee. *Journal of the American Medical Informatics Association*, 29(1), 22–32.
- Robins, L. N., Helzer, J. E., Croughan, J., & Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. *Archives of General Psychiatry*, 38(4), 381–389.
- Ross, J., Teesson, M., Darke, S., Lynskey, M., Ali, R., Ritter, A., & Cooke, R. (2005). The characteristics of heroin users entering treatment: Findings from the Australian Treatment Outcome Study (ATOS). *Drug and Alcohol Review*, 24(5), 411–418. <https://doi.org/10.1080/09595230500286039>
- Sahle, B. W., Owen, A. J., Chin, K. L., & Reid, C. M. (2017). Risk prediction models for incident heart failure: A systematic review of methodology and model performance. *Journal of Cardiac Failure*, 23(9), 680–687. <https://doi.org/10.1016/j.cardfail.2017.03.005>
- Saloner, B., Chang, H.-Y., Krawczyk, N., Ferris, L., Eisenberg, M., Richards, T., Lemke, K., Schneider, K. E., Baier, M., & Weiner, J. P. (2020). Predictive modeling of opioid overdose using linked statewide medical and criminal justice data. *JAMA Psychiatry*, 77(11), 1155–1162.
- Schaar, I., & Öjehagen, A. (2001). Severely mentally ill substance abusers: An 18-month follow-up study. *Social Psychiatry and Psychiatric Epidemiology*, 36(2), 70–78. <https://doi.org/10.1007/S001270170066>
- Šimundić, A.-M. (2009). Measures of diagnostic accuracy: Basic definitions. *EJIFCC*, 19(4), 203–211.
- Steyerberg, E. W. (2019). *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating* (2nd ed.). Springer Nature Switzerland. [https://doi.org/10.1111/j.1751-5823.2009.00085\\_22.x](https://doi.org/10.1111/j.1751-5823.2009.00085_22.x)
- Strang, J., Volkow, N. D., Degenhardt, L., Hickman, M., Johnson, K., Koob, G. F., Marshall, B. D. L., Tyndall, M., & Walsh, S. L. (2020). Opioid use disorder. *Nature Reviews Disease Primers*, 6(1), 1–28.
- Sun, J. W., Franklin, J. M., Rough, K., Desai, R. J., Hernandez-Diaz, S., Huybrechts, K. F., & Bateman, B. T. (2020). Predicting overdose among individuals prescribed opioids using routinely collected healthcare utilization data. *PLoS ONE*, 15(10), e0241083. <https://doi.org/10.1371/JOURNAL.PONE.0241083>
- Teesson, M., Marel, C., Darke, S., Ross, J., Slade, T., Burns, L., Lynskey, M., Memedovic, S., White, J., & Mills, K. L. (2015). Long-term mortality, remission, criminality and psychiatric comorbidity of heroin dependence: 11-year findings from the Australian treatment outcome study. *Addiction*, 110(6), 986–993. <https://doi.org/10.1111/add.12860>
- Teesson, M., Marel, C., Darke, S., Ross, J., Slade, T., Burns, L., Lynskey, M., Memedovic, S., White, J., & Mills, K. L. (2017). Trajectories of heroin use: 10–11-year findings from the Australian Treatment Outcome Study. *Addiction*, 112(6), 1056–1068.
- Teesson, M., Mills, K., Ross, J., Darke, S., Williamson, A., & Havard, A. (2008). The impact of treatment on 3 years' outcome for heroin dependence: Findings from the Australian Treatment Outcome Study (ATOS). *Addiction*, 103(1), 80–88. <https://doi.org/10.1111/j.1360-0443.2007.02029.x>
- Usher-Smith, J. A., Walter, F. M., Emery, J. D., Win, A. K., & Griffin, S. J. (2016). Risk prediction models for colorectal cancer: A systematic review. In *Cancer Prevention Research* (Vol. 9, Issue 1, pp. 13–26). American Association for Cancer Research. <https://doi.org/10.1158/1940-6207.CAPR-15-0274>
- van Amsterdam, J., van den Brink, W., & Pierce, M. (2021). Explaining the differences in opioid overdose deaths between Scotland and England/Wales: Implications for European Opioid Policies. *European Addiction Research*, 27(6), 399–412.
- Walsh, M. E., Frances Horgan, N., Walsh, C. D., & Galvin, R. (2016). Systematic review of risk prediction models for falls after stroke. In *Journal of Epidemiology and Community Health* (Vol. 70, Issue 5, pp. 513–519). BMJ Publishing Group Ltd. <https://doi.org/10.1136/jech-2015-206475>

- Ware, J. E. J., Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care*, *34*(3), 220–233.
- Westreich, D. (2012). Berkson's bias, selection bias, and missing data. *Epidemiology*, *23*(1), 159–164.
- Wilson, T., Quan, S., Cheema, K., Zarnke, K., Quinn, R., De Koning, L., Dixon, E., Pannu, N., & James, M. T. (2016). Risk prediction models for acute kidney injury following major noncardiac surgery: Systematic review. *Nephrology Dialysis Transplantation*, *31*(2), 231–240. <https://doi.org/10.1093/ndt/gfv415>
- CDC WONDER. (2020). Wide-ranging online data for epidemiologic research. US Centers for Disease Control and Prevention. <http://wonder.cdc.gov>. Accessed 18 Feb 2022
- CDC WONDER. (2021). Drug Overdose Deaths Remain High. US Centers for Disease Control and Prevention. <http://wonder.cdc.gov>. Accessed 18 Feb 2022
- World Health Organisation. (1993). Composite international diagnostic interview, version 1.1. Geneva: World Health Organisation
- Worley, J. (2017). Recovery in substance use disorders: What to know to inform practice. *Issues in Mental Health Nursing*, *38*(1), 80–91. <https://doi.org/10.1080/01612840.2016.1245375>
- Yang, H., Luo, Y., Ren, X., Wu, M., He, X., Peng, B., Deng, K., Yan, D., Tang, H., & Lin, H. (2021). Risk prediction of diabetes: Big data mining with fusion of multifarious physical examination indicators. *Information Fusion*, *75*, 140–149.
- Zhang, R., Borisenko, O., Telegina, I., Hargreaves, J., Ahmed, A. R., Sanchez Santos, R., Pring, C., Funch-Jensen, P., Dillemans, B., & Hedenbro, J. L. (2016). Systematic review of risk prediction models for diabetes after bariatric surgery. In *The British journal of surgery* (Vol. 103, Issue 11, pp. 1420–1427). Oxford Academic. <https://doi.org/10.1002/bjs.10255>

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