

Machine Learning Based Prediction Model for Using Non-steroidal Anti-inflammatory Drugs on Risk of Adverse Events

Juan Lu

University of Western Australia

Ling Wang

University of Western Australia

Mohammed Bennamoun

University of Western Australia

Isaac Ward

University of Western Australia

Senjian An

Curtin University

Ferdous Sohel

Murdoch University

Benjamin JW Chow

University of Ottawa

Girish Dwivedi

Harry Perkins Institute of Medical Research

Frank M Sanfilippo (✉ frank.sanfilippo@uwa.edu.au)

University of Western Australia

Research Article

Keywords: machine learning, non-steroidal anti-inflammatory drugs (NSAIDs), gradient boosting machine

Posted Date: January 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-153020/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

To investigate the usefulness of machine learning approaches on linked administrative health data at the population level in predicting older patients' one-year risk of adverse events after supplied with non-steroidal anti-inflammatory drugs (NSAIDs). Patients from a Western Australian cardiovascular population who were supplied with NSAIDs between 1 Jan 2003 and 31 Dec 2004 were identified from Pharmaceutical Benefits Scheme data. Comorbidities from linked hospital admissions data and medication history were inputs. Admissions for acute coronary syndrome or deaths within one year from the last supply date were outputs. Machine learning classification methods were used to build models to predict ACS and optimise their performance, measured by the area under the receiver operating characteristic curve, sensitivity and specificity. There were 69007 patients in the NSAIDs cohort with mean age 76 years and 54.3% were female. 1882 patients were admitted for acute coronary syndrome and 5405 patients dead within one year after their last supply of NSAIDs. The multi-layer neural network, gradient boosting machine and support vector machine were applied to build various classification models. The models achieved an average AUC-ROC score of 0.72 predicting ACS and 0.84 predicting death. Machine learning models applied to linked administrative data can potentially improve adverse event risk prediction. Further investigation of additional data and approaches are required further to improve the performance for adverse event risk prediction.

Introduction

Machine learning is increasingly common in big data science, experimentally in the electronic medical record (EMR) data^{1–3}. Machine learning models have shown their advantages in risk predictions based on a wide array of patients' EMR data^{4, 5}. The risk prediction of drug response has drawn close attention in public health. Adverse drug reactions (ADRs) have long been recognised as a potential outcome of taking medicines. It is practically impossible for doctors to reduce patients' risks of ADRs through manually screening millions of health records while prescribing. Therefore, we would like to build a set of machine learning models based on a mass of patients' EMR data to predict patients' risk of ADRs.

ADRs are common in older adults with various reactions and a significant proportion of ADRs are responsible for hospital admissions^{6, 7}. Moreover, the older adults are nearly seven times more likely to be hospitalised due to ADR-related problems than their younger counterparts^{6, 8}. Thus, accurate ADR risk prediction models are necessary for clinical practice and helping doctors to reduce the risk of ADRs in the elderly. A large number of surveys aimed to identify the key factors increasing a person's risk of ADR have been proposed^{9, 10}. But they are not suitable for predicting the individual risk of ADRs due to the considerable differences in diseases and drug history between patients. This motivates the machine learning-based risk prediction model design of ADRs based on patients' diseases and drug history in their EMR database. This paper mainly focuses on the use of Non-steroidal Anti-inflammatory Drugs (NSAID) which are reported to be associated with a dose-related increased risk of cardiovascular (CV) events.

NSAIDs are extensively prescribed for the treatment of musculoskeletal disorders, rheumatoid arthritis (RA), osteoarthritis (OA), nociceptive pain, headache, and inflammation^{11, 12}. A large number of structurally diverse NSAIDs with similar therapeutic effects have been developed and NSAIDs have belonged to the most widely used pharmacological drugs, both over the counter (OTC) and doctoral prescription^{13–15}. However, their potential adverse effects are also well known. Multiple previous studies have reported an increased risk of CV events from the use of NSAIDs^{11, 15–21}. For example, Rofecoxib, one of the NSAIDs, was withdrawn from the market in October 2004 after a randomised placebo-controlled trial showed an increased risk of CV among rofecoxib users²¹. Importantly, the population commonly taking NSAIDs is that of elderly individuals who suffer from CV diseases^{11, 15}. Thus, we propose the machine learning-based prediction model for using NSAIDs on the risk of adverse event (AE). Admissions due to acute coronary syndrome (ACS) is one of the very common ADRs of NSAIDs.

The objective of this study was to build a machine learning model to predict the risk of AE for elderly patients who took NSAIDs in Western Australia. We used various patients' comorbidity history and medication history for model development. All the records are from the Pharmaceutical Benefits Scheme (PBS), linked with Hospital Morbidity Data Collection (HMDC) and death register dataset in Western Australia. We compared the performance of different machine learning models and analysed the impact of features on the machine learning model.

Methods

Data sources

The study datasets consisted of public and private hospital admissions for heart disease in Western Australia during 2003–2008 from the HMDC, with linked admission records back to 1980 and forwarded to 2014²². These were linked to matching records from the Western Australian death registry to 2014, and PBS data from mid-2002 to mid-2011 from the Australian Department of Human Services. The HMDC and mortality data are 2 of the core datasets of the Western Australian Data Linkage System²³. The PBS dataset contains patient-level information for medications dispensed from community pharmacies and PBS hospitals, including details such as drug name and strength, quantity supplied, and supply date.

Inclusion criteria and selection

In this paper, we sampled patients supplied with NSAIDs at least once between 1 Jan 2003 and 31 Dec 2004 and aged 65 or above, from PBS dataset. All the drugs were identified by the Anatomical Therapeutic Chemical (ATC) code. The sampling period, due to that Rofecoxib was withdrawn from the market in October 2004, ensured that we could capture all the records of NSAIDs. The PBS dataset just recorded medications that enjoy government benefits and did not include records of out-of-pocket payments. Previous research had shown that patients aged 65 or above are mostly concessional beneficiaries, and their dispensing records in the PBS data are mostly complete²⁴. Furthermore, most of the patients taking NSAIDs are also elderly and ARDs are more common and serious for the elderly. Thus,

the age of the patients in the study was restricted to 65 and above. Figure 1 showed the timeline of sampling. The selected study patients are records between 1 Jan 2003 and 31 Dec 2004. The patients' comorbidity history dates back to 10 years, and their drug history goes back to six months. The AEs of NSAIDs are recorded based on the patient's status within one year after taking NSAID. All the features and outcomes of the study cohort will be detailed in the following subsections.

Input features

The features in our model are composed of (1) patients' demographic information, (2) patient's comorbidity history and (3) drug history. Demographic information includes the patients' age, gender, marital status, and indigenous ethnicity. These are very common features in medical records and are considered to be strongly related to the patient's health. Age and gender were defined at the last supply date of the NSAIDs for the study cohort. Marital status and indigenous ethnicity were defined at the last admission before patients' last NSAIDs supply. The comorbidity history and drug history are recorded based on the timeline design, and all-time information will not be recorded in the features. History of comorbidities was determined from the diagnosis codes (both ICD-9-CM and ICD-10-AM) in the HMDC dataset with a 10-year lookback period from the last supply date (Supplementary table 1: detailed ICD code). Comorbidities included 13 features: ischaemic heart disease, hypertension, atrial fibrillation, diabetes, chronic obstructive pulmonary disease, peripheral vascular disease, stroke, chronic kidney disease, cancer, dementia, depression, heart failure, and cardiomyopathy. We used comorbidity history as continuous variables representing the frequency of previous admissions. Drug history was identified using a 6-month look back from the last supply date of the cohort using the PBS data, and drugs were grouped into 16 features corresponding to the first character of the corresponding ATC code²⁵. We also included the history of NSAIDs into 13 features corresponding to 13 NSAIDs. Drug history was presented as continuous variables representing the total number of scripts supplied to the patients.

Outcome

We focused on patients' risk of ACS and all-cause mortality in our study, as previous studies have presented the risks of NSAIDs in CV events.^{11, 15–21} ACS admission was identified from the principal discharge diagnosis field from the HMDC records using International Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification [ICD-10-AM]) code I20.0 for unstable angina and I21 for myocardial infarction. We also classified patients who died due to cardiovascular diseases as ACS. Patients who died before their last supply date were excluded. (Fig. 2). All-cause mortality was identified from the death registry. We also looked at a composite outcome, including both ACS admissions and all-cause mortality. Follow-up of patients began after their last supply date and finished at 365 days after the last supply date. In all the records we obtained, there are some patients with the same input features but different outcomes (with or without the event), which will interfere with the prediction results. Therefore, we excluded these records before training the machine learning models.

Machine Learning Method

We developed three machine learning models for risk prediction: gradient boosting classifier model (GBM)²⁶, multi-layer neural network (MLNN) model and support vector machine (SVM) model. These machine learning models are well-performed in clinical risk prediction^{27–29}. However, there is no literature exploring their performance in risk prediction of taking NSAIDs. All the records are split into the training set and testing set with a ratio of 0.75:0.25. Further details of GBM, MLNN (Supplementary Fig. 1) and SVM are described in the Supplementary File. All analyses and model building were done with Python version 3.7 and relevant libraries, including scikit-learn³⁰, and Keras³¹.

The predictive performance of models was compared by calculating sensitivity, specificity, and AUC-ROC. We used the Youden index³² to identify the optimised threshold for the ML model predictions that would achieve a balanced sensitivity and specificity. For all models, we randomly split the dataset using different random states and calculated their mean performance matrices and their 95% confidence intervals from training and evaluating the models 50 times. Once the outperformed model was identified, we conducted a sensitivity analysis using different NSAIDs testing set. We divided the test set on patients supplied with different NSAIDs. The model was then compared with the Cox regression model based on the same features to validate our modelling and performance. We built two cox regression models, with one of them is using the same continuous variables as we had in machine learning models, the other one is built on the same features, but all features were binary variables. Feature importance plots were generated by GBM for inspection.

Results

Cohort characteristics

Figure 2 shows the results of each step in sampling from the dataset. There were 109,101 patients supplied with NSAIDs during 2003 and 2004, and 40,212 were excluded due to age < 65 years or they died before the last supply (Fig. 1). Therefore, we identified 68,889 patients in the cohort with more than 40% users of Celecoxib and 35% users of Rofecoxib. Table 1 shows patient characteristics for the study groups. The mean age was 76, and more than 50% of the cohort was female. More males developed ACS, and older patients were more likely to develop an adverse outcome. Cardiovascular diseases such as ischaemic heart disease and heart failure were more common among patients who developed ACS than those with no ACS. The frequency of comorbidity history was higher in patients who died during the follow-up.

Table 1
Characteristics of the study cohorts for NSAIDs during 2003–2004

Features	Total Cohort	ACS*	p-value	All-cause death	p-value
	No. of patients (%) n = 68 889 (100)	No. of patients (%) n = 2757 (4.0)		No. of patients (%) n = 5405 (7.8)	
Age (years, mean[SD])	76.0 (7.2)	78.8 (8.0)	< 0.0001	80.9 (7.8)	< 0.0001
Female	37 389 (54.3)	1324 (48.0)	< 0.0001	2733 (54.3)	< 0.0001
Celecoxib	29 774 (43.2)	1204 (43.7)	0.6	2373 (43.9)	0.3
Rofecoxib	24 432 (35.5)	953 (34.6)	0.3	1851 (34.3)	0.051
Indometacin	4088 (5.9)	180 (6.53)	0.2	370 (6.9)	0.004
Sulindac	286 (0.4)	11 (0.4)	0.9	29 (0.5)	0.1
Diclofenac	11 660 (16.9)	422 (15.3)	0.02	705 (13.0)	< 0.0001
Diclofenac, combinations	18 (0.03)	4 (0.15)	< 0.0001	3 (0.06)	0.2
Piroxicam	3414 (5.0)	130 (4.7)	0.6	203 (3.8)	< 0.0001
Meloxicam	12 982 (18.8)	434 (15.7)	< 0.0001	686 (12.7)	< 0.0001
Ibuprofen	4305 (6.3)	161 (5.8)	0.4	319 (5.9)	0.3
Naproxen	5883 (8.5)	214 (7.8)	0.1	435 (8.1)	0.2
Ketoprofen	1951 (2.8)	65 (2.4)	0.1	122 (2.3)	0.008
Tiaprofenic acid	398 (0.6)	10 (0.4)	0.1	34 (0.6)	0.6
Fenamates	46 (0.07)	2 (0.07)	0.9	2 (0.04)	0.4
Comorbidity history					
Ischemic Heart Disease	14 445 (21.0)	1031 (37.4)	< 0.0001	1150 (21.3)	0.6

SD, standard deviation; COPD, chronic obstructive pulmonary disease; PVD, peripheral vascular disease; ACS, acute coronary syndrome. Drug groups: A, alimentary tract and metabolism; B, blood and blood forming organs; C, cardiovascular system; D, dermatologicals; G, genito urinary system and sex hormones; H, systemic hormonal preparations; J, anti-infectives for systemic use; L, antineoplastic and immunomodulating agents; M, musculo-skeletal system; N, nervous system; P, antiparasitic products; R, respiratory system; S, sensory organs; V, various.

Features	Total Cohort	ACS*		All-cause death	
Hypertension	1091 (1.6)	52 (1.9)	0.2	113 (2.1)	0.002
Atrial Fibrillation	3468 (5.0)	198 (7.2)	< 0.0001	390 (7.2)	< 0.0001
Diabetes	3779 (5.5)	186 (6.8)	0.003	409 (7.6)	< 0.0001
COPD	3659 (5.3)	255 (9.3)	< 0.0001	624 (11.5)	< 0.0001
PVD	3226 (4.7)	277 (10.1)	< 0.0001	474 (8.8)	< 0.0001
Stroke	2545 (3.7)	164 (6.0)	< 0.0001	425 (7.9)	< 0.0001
Chronic Kidney Disease	1337 (1.9)	78 (2.8)	< 0.0001	159 (2.9)	< 0.0001
Cancer	3049 (4.4)	98 (3.6)	0.02	747 (13.8)	< 0.0001
Dementia	386 (0.6)	27 (1.0)	0.003	140 (2.6)	< 0.0001
Depression	831 (1.2)	43 (1.6)	0.08	94 (1.7)	0.0002
Heart Failure	2474 (3.6)	335 (12.2)	< 0.0001	629 (11.6)	< 0.0001
Cardiomyopathy	136 (0.2)	9 (0.3)	0.1	19 (0.4)	0.008
Drug history					
Drug group A	42 495 (61.7)	1868 (67.8)	< 0.0001	4032 (74.6)	< 0.0001
Drug group B	22 848 (33.2)	1214 (44.0)	< 0.0001	2352 (43.5)	< 0.0001
Drug group C	57 953 (84.1)	2389 (86.7)	0.0002	4496 (83.2)	0.05
Drug group D	13 710 (19.9)	546 (19.8)	0.9	1128 (20.9)	0.06
Drug group G	6924 (10.1)	243 (8.8)	0.03	536 (9.9)	0.7
Drug group H	12 275 (17.8)	520 (19.6)	0.01	1494 (27.6)	< 0.0001

SD, standard deviation; COPD, chronic obstructive pulmonary disease; PVD, peripheral vascular disease; ACS, acute coronary syndrome. Drug groups: A, alimentary tract and metabolism; B, blood and blood forming organs; C, cardiovascular system; D, dermatologicals; G, genito urinary system and sex hormones; H, systemic hormonal preparations; J, anti-infectives for systemic use; L, antineoplastic and immunomodulating agents; M, musculo-skeletal system; N, nervous system; P, antiparasitic products; R, respiratory system; S, sensory organs; V, various.

Features	Total Cohort	ACS*		All-cause death	
Drug group J	30 857 (44.8)	1336 (48.5)	< 0.0001	3151 (58.3)	< 0.0001
Drug group L	3192 (4.6)	131 (4.8)	0.8	585 (10.8)	< 0.0001
Drug group M	29 698 (43.1)	1152 (41.8)	0.2	1935 (35.8)	< 0.0001
Drug group N	46 410 (67.4)	2075 (75.3)	< 0.0001	4517 (83.6)	< 0.0001
Drug group P	5829 (8.5)	290 (10.5)	< 0.0001	593 (11.0)	< 0.0001
Drug group R	13 747 (20.0)	640 (23.2)	< 0.0001	1456 (26.9)	< 0.0001
Drug group S	21 652 (31.4)	933 (33.8)	0.005	1919 (35.5)	< 0.0001
Drug group V	2554 (3.7)	109 (4.0)	0.5	261 (4.8)	< 0.0001

SD, standard deviation; COPD, chronic obstructive pulmonary disease; PVD, peripheral vascular disease; ACS, acute coronary syndrome. Drug groups: A, alimentary tract and metabolism; B, blood and blood forming organs; C, cardiovascular system; D, dermatologicals; G, genito urinary system and sex hormones; H, systemic hormonal preparations; J, anti-infectives for systemic use; L, antineoplastic and immunomodulating agents; M, musculo-skeletal system; N, nervous system; P, antiparasitic products; R, respiratory system; S, sensory organs; V, various.

Performance of machine learning models

Table 2 shows the performance of different ML models as averages of the model sensitivity, specificity and AUC-ROC from training and evaluating the models 50 times. Among the algorithms examined, we found that GBM using features including age, sex, marital status, indigenous ethnicity, patients comorbidity history and drug history as continuous variables has achieved the best performance on predicting patients' risk of ACS (AUC 0.72, 95% CI 0.71–0.73). It slightly outperformed MLNN (AUC 0.71, 95% CI 0.70–0.71) and SVM(AUC 0.710, 95% CI 0.707, 0.712). The GBM had an average sensitivity of 61% (95% CI 60%-63%) and an average specificity of 72% (95% CI 70%-73%) using cutoffs selected by the Youden index. Machine learning models achieved similar performance on predicting all-cause mortality (AUC 0.84) and composite outcome (AUC 0.78) using the same features. We also compared machine learning models with a Cox regression model based on the same features. The Cox regression model had a lower average AUC. (0.659 95% CI 0.656–0.662).

Table 2

Performance of machine learning models and Cox regression measured by sensitivity, specificity, and AUC-ROC

Models	Performance Metrics	ACS (95% CI)	All-cause death (95% CI)	ACS or All-cause death (95% CI)
GBM	Sensitivity	0.61 (0.60, 0.63)	0.78 (0.78, 0.79)	0.68 (0.67, 0.69)
	Specificity AUC-ROC	0.72 (0.70, 0.73)	0.74 (0.73, 0.75)	0.75 (0.74, 0.75)
		0.72 (0.71, 0.72)	0.837 (0.836, 0.839)	0.780 (0.778, 0.781)
MLNN	Sensitivity	0.61 (0.60, 0.63)	0.76 (0.75, 0.76)	0.69 (0.68, 0.70)
	Specificity AUC-ROC	0.70 (0.69, 0.71)	0.76 (0.75, 0.77)	0.75 (0.74, 0.75)
		0.70 (0.70, 0.71)	0.834 (0.833, 0.836)	0.778 (0.776, 0.780)
SVM	Sensitivity	0.61 (0.60, 0.62)	0.74 (0.73, 0.75)	0.70 (0.69, 0.71)
	Specificity AUC-ROC	0.72 (0.71, 0.73)	0.75 (0.74, 0.75)	0.73 (0.72, 0.74)
		0.710 (0.707, 0.712)	0.813 (0.812, 0.814)	0.777 (0.776, 0.779)
Cox Regression	Sensitivity	0.62 (0.60, 0.64)	0.71 (0.70, 0.72)	0.66 (0.66, 0.67)
	Specificity AUC-ROC	0.63 (0.61, 0.65)	0.69 (0.67, 0.70)	0.66 (0.65, 0.66)
		0.659 (0.656, 0.662)	0.76 (0.75, 0.76)	0.711 (0.710, 713)
Cox Regression*	Sensitivity	0.638 (0.602, 0.674)	0.726 (0.710, 0.742)	0.653 (0.641, 0.665)
	Specificity AUC-ROC	0.66 (0.625, 0.694)	0.729 (0.712, 0.746)	0.728 (0.718, 0.739)
		0.695 (0.688, 0.702)	0.795 (0.793, 0.797)	0.750 (0.745, 754)
*Built on binary variables				

Table 3 shows the performance of GBM on predicting patients supplied with different NSAIDs. It achieved the highest AUC on patients supplied with Sulindac while predicting their risk of ACS (AUC 0.84). Its performance on predicting the risk of ACS was lower on patients supplied with Piroxicam (AUC 0.66). We found similar average AUC between different NSAIDs on all-cause mortality risk prediction, with a slightly lower AUC (0.79) on patients supplied with Ketoprofen. The AUC was higher while predicting the risk of the composite outcome on patients supplied with Sulindac and Tiaprofenic acid.

Table 3
Risk prediction performance (AUC-ROC 95% CI) for different NSAIDs

	ACS	All-cause death	ACS or All-cause death
Indometacin	0.71 (0.70, 0.72)	0.81 (0.80, 0.81)	0.77 (0.77, 0.78)
Sulindac	0.84 (0.78, 0.89)	0.82 (0.80, 0.84)	0.82 (0.80, 0.84)
Diclofenac	0.67 (0.66, 0.68)	0.80 (0.79, 0.80)	0.74 (0.74, 0.75)
Piroxicam	0.66 (0.65, 0.68)	0.80 (0.79, 0.81)	0.73 (0.72, 0.74)
Meloxicam	0.70 (0.69, 0.70)	0.80 (0.80, 0.81)	0.75(0.74,0.75)
Ibuprofen	0.71 (0.70, 0.73)	0.82 (0.81, 0.82)	0.76(0.75,0.77)
Naproxen	0.71 (0.70, 0.72)	0.82 (0.81, 0.82)	0.77(0.77,0.78)
Ketoprofen	0.71 (0.69, 0.73)	0.79 (0.78, 0.80)	0.73(0.72,0.74)
Tiaprofenic acid	0.77 (0.72, 0.82)	0.85 (0.83, 0.87)	0.83 (0.80, 0.85)
Rofecoxib	0.710 (0.705, 0.714)	0.821 (0.819, 0.823)	0.78 (0.77, 0.78)
Celecoxib	0.72 (0.71, 0.72)	0.811 (0.809, 0.813)	0.772 (0.770, 0.774)

Feature Importance

Figure 3 showed the ranked feature importance while predicting AE by GBM. Age was the most important predictor among all the features. Previous cardiovascular diseases such as ischaemic heart disease and heart failure were ranked at top while predicting ACS, which were followed by drug group Cardiovascular system (C) and Nervous system (N) (Fig. 3(a)). Cancer and heart failure history were important features associated with death, as well as drug group (N), Musculo-skeletal system (M) (Fig. 3(b)).

Cyclooxygenase-2 (COX-2) inhibitors were ranked highest among all NSAIDs while predicting patients' risk of ACS. Naproxen, ibuprofen and Ketoprofen were ranked lower comparing with other NSAIDs.

Similar results were found for composite outcome (Fig. 3(c)).

Discussion

This study presents a set of machine learning models for predicting the risks of AE after taking NSAIDs using data from PBS and HMDC in Western Australia. We focused specifically on elderly patients (\geq age 65) who took at least one NSAID. The prediction is based on the features including age, sex, medication history and disease history, which are widely concerned and counted in clinical practice. This approach encompasses a wide array of patients to truly reflect the population of patients taking NSAIDs in Western Australia. The machine learning based predictive models for AE showed greater sensitivity, specificity and AUC-ROC versus the classical cox-regression approach and GBM presented the best predictive performance in machine learning models we tested.

Several studies have reported the risk of AE with NSAIDs and the Rofecoxib was withdrawn from the market due to its increased risk of CV. The models are built to predict CV-related AE, death and overall AE. The performance of predicting death is the best with AUC-ROC values range from *0.67 to 0.81*. This does not mean that the death was caused by NSAIDs, but this demonstrates that the predictive models built based on PBS and HMDC work well and can predict the risk of death.

NSAIDs include a series of medicines. Experimental data includes all NSAIDs with more than *100* patients. The AUC-ROC values of risk prediction for different NSAIDs range from *0.60 to 0.88*. The proposed GBM model can be used to predict the risk of AE after taking any NSAID, especially for the *Rofecoxib and Celecoxib* whose AUC-ROC is more than 0.8.

Machine learning models have been widely used on EMR for prediction purpose, such as nationwide cohort predicting suicide death³³, prediction of graft survival in kidney transplant recipients³⁴, risk prediction of AEs following spine surgery⁵. These studies found that the machine learning approach did not show better performance than a classical generalised regression approach. However, in our data machine learning models tend to perform better over the cox-regression model. This could be because most of the input feature in our model are continuous variables and machine learning models turn out to be outperformed on complex variables.

To our knowledge, there is no literature that explores the predictive models for AE taking NSAIDs. Our current study is the first to realise risk prediction in patients with NSAIDs. This study has several strengths. The risk prediction model we developed can be used to avoid some ADRs of NSAIDs. This model can inform patients counselling by enabling doctors to prescribe NSAID with the lowest risk based on individual patient's medication history and disease history. Moreover, this model can also be used in EMR dataset to pick up the patients with a high risk of AE and help the hospital to pay close attention to these patients.

Our study found that the inclusion of demographic features such as marital status, indigenous ethnicity from linked HMDC improved the performance of AE prediction. The average AUC was similar while predicting ACS (AUC 0.71). But the performance was higher while predicting patients' risk of all-cause mortality (AUC 0.81 vs 0.84) and composite outcome (AUC 0.77 vs 0.78), and no overlap between their confidence intervals. Previous studies have confirmed marital status is associated with cardiovascular outcomes and mortality was higher in unmarried population^{35, 36}. Studies have also shown that indigenous Australian had a greater rate of cardiovascular disease and death. ^{37, 38}

We extracted additional features from HMDC, including patients' previous length of stay (days) of each comorbidity in the hospital and the number of days patients spent in intensive care units (ICU) before their last supply. This set of features were presented as continuous variables. We included this set of features to test whether it would improve the AE risk prediction. However, there were no performance gains by adding continuous variables such as length of hospital stay of previous comorbidities and days in ICU before their last NSAIDs supply. The AUCs of all different outcomes were similar. We dropped these features to reduce model complexity.

In our study, we observed minimal performance improvement when using binary variables indicating the presence or absence of previous comorbidities or the use of specific drugs. However, ML models achieved better performance than cox regression when we were using continuous variables to present patients use of different medications and their comorbidity history. This may be because machine learning approaches do not assume linearity for a predictor-outcome association; they are more adept at generating predictions based on continuous variables³⁹.

Our machine learning model ranked COX-2 inhibitors higher among NSAIDs in ACS risk prediction. Multiple previous studies have reported an increased risk of CV events from the use of selective COX-2 inhibitors. 11, 15, 17–21. Rofecoxib was withdrawn from markets based on evidence that showed an increased risk of ACS. 21 Previously study has also combined heart failure has substantially increased the risk mortality. 40 This verifies that our machine learning model is reliable in ranking feature importance.

Despite the merits of this study, there are some limitations. As with all administrative database studies, this study relies on the accuracy of administrative coding of procedures, diagnoses, and records. The PBS dataset did not include all dispensing supplies of NSAIDs such as ibuprofen, as it is also available at the counter. Moreover, PBS dataset did not contain information about the actual drug dosage. Hence, in our study, we calculated the total number of supplied scripts rather than the dose used. In this study, we used state-level linked data to predict patients AE after their NSAIDs supply. The models can be further extended on national linked data in the future. Also, for general applicability, the models can be potentially extended to other drugs or drug groups on different outcomes, and this can be tested in future studies.

Implementing ML models on linked administrative data, including pharmacy claims (e.g. PBS), morbidity, and mortality has the potential to identify patients supplied with NSAIDs that may have a high risk of adverse outcomes. These can then be monitored closely by humans. Further investigation of additional data is required to validate the ML prediction performance on patients' risk of CV adverse events using population-level linked data.

Declarations

Acknowledgements

We thank Dr. Kevin Murray from the School of Population and Global Health (University of Western Australia) for statistical advice. We also thank the following institutions for providing the data used in this study. Staff at the WA Data Linkage Branch and data custodians of the WA Department of Health Inpatient Data Collections and Registrar General for access to and provision of the State linked data. The Australian Department of Health for the cross-jurisdictional linked PBS data. This study is part-funded by Health and Medical Science Faculty Research Grant, University of Western Australia, Australia.

Author contributions statement

Contributors: FM.S., G.D. conceived the experiments. J. L. and L.W conducted the experiments. J.L. analysed the data and together with L.W. wrote the first draft of the manuscript. M.B. contributed to the machine learning design and interpretation and supervised J.L. and L.W. I.W. contributed to fortnightly discussions and planning of the analysis. F.S. and S.A. supervised the machine learning analysis and contributed to the interpretation. B.C. provided critical review and clinical interpretation. G.D. contributed to study design, analysis plan, and clinical interpretation, and supervised J.L. and L.W. FM.S. contributed to the conception, study design, data acquisition, funding, and analysis plan, and supervised J.L. and L.W. All co-authors critically revised the manuscript and gave final approval and agree to be accountable for all aspects of the work ensuring integrity and accuracy.

Competing interests

Ben Chow reports non-financial support from Siemens, grants from CV Diagnostix, AusculSciences and TD Bank, and an equity interest in GE, all of which are outside the submitted work. Girish Dwivedi reports personal fees from Pfizer, Amgen, Astra Zeneca and Artrya Pty Ltd, all of which are outside the submitted work. No other competing interests were declared.

Ethics approval

Human Research Ethics Committee approval was obtained from the University of Western Australia (RA/4/1/8065), the WA Department of Health (2014/11), and the Australian Department of Health (XJ-16). We were granted a waiver of informed consent.

Regulation statement

All methods were carried out in accordance with relevant guidelines and regulations.

References

1. Darcy, M., Louie, A. K. & Roberts, L. W. Machine learning and the profession of medicine. *Jama* **315**, 551–552 (2016).
2. Parikh, R. B., Kakad, M. & Bates, D. Integrating predictive analytics into high-value care: the dawn of precision delivery. *Jama* **315**, 651–652 (2016).
3. Shickel, B., Tighe, J., Bihorac, A. & Rashidi, P. Deep ehr: a survey of recent advances in deep learning techniques for electronic health record (ehr) analysis. *IEEE journal biomedical health informatics* **22**, 1589–1604 (2017).

4. Ehlers, A. *et al.* Improved risk prediction following surgery using machine learning algorithms. *eGEMs* **5**, 3 (2017).
5. Han, S., Azad, T. D., Suarez, P. A. & Ratliff, J. K. A machine learning approach for predictive models of adverse events following spine surgery. *The Spine J.* **19**, 1772–1781 (2019).
6. Lavan, H. & Gallagher, P. Predicting risk of adverse drug reactions in older adults. *Ther. advances drug safety* **7**, 11–22 (2016).
7. Stevenson, M. *et al.* Predicting adverse drug reactions in older adults; a systematic review of the risk prediction models. *Clin. interventions aging* **9**, 1581 (2014).
8. Cohen, L. *et al.* National surveillance of emergency department visits for outpatient adverse drug events in children and adolescents. *The J. pediatrics* **152**, 416–421 (2008).
9. Mangoni, A. A. & Jackson, S. H. Age-related changes in pharmacokinetics and pharmacodynamics: basic principles and practical applications. *journal clinical pharmacology* **57**, 6–14 (2004).
10. Davies, E. C. *et al.* Adverse drug reactions in hospital in-patients: a prospective analysis of 3695 patient-episodes. *PLoS one* **4**, e4439 (2009).
11. Zingler, G., Hermann, B., Fischer, & Herdegen, T. Cardiovascular adverse events by non-steroidal anti-inflammatory drugs: when the benefits outweigh the risks. *Expert. Rev. Clin. Pharmacol.* **9**, 1479–1492 (2016).
12. Patrono, & Baigent, C. Low-dose aspirin, coxibs, and other nsoids: a clinical mosaic emerges. *Mol. interventions* **9**, 31 (2009).
13. Brune, & Patrignani, P. New insights into the use of currently available non-steroidal anti-inflammatory drugs. *J. pain research* **8**, 105 (2015).
14. Bindu, S., Mazumder, S. & Bandyopadhyay, U. Non-steroidal anti-inflammatory drugs (nsoids) and organ damage: A current perspective. *Pharmacol.* **180**, 114147, DOI: <https://doi.org/10.1016/j.bcp.2020.114147> (2020).
15. Ungprasert, , Srivali, N., Wijarnpreecha, K., Charoenpong, P. & Knight, E. L. Non-steroidal anti-inflammatory drugs and risk of venous thromboembolism: a systematic review and meta-analysis. *Rheumatology* **54**, 736–742 (2015).
16. Zhou, , Xiao, Y., Li, C., Liu, T. & Wang, W. Association between use of non-steroidal anti-inflammatory drugs and risk of myocardial infarction in patients with spondyloarthritis and osteoarthritis. *Annals rheumatic diseases* **78**, e78–e78 (2019).
17. Lindhardsen, *et al.* Non-steroidal anti-inflammatory drugs and risk of cardiovascular disease in patients with rheumatoid arthritis: a nationwide cohort study. *Annals rheumatic diseases* **73**, 1515–1521 (2014).
18. Huerta, , Varas-Lorenzo, C., Castellsague, J. & Rodriguez, L. G. Non-steroidal anti-inflammatory drugs and risk of first hospital admission for heart failure in the general population. *Heart* **92**, 1610–1615 (2006).

19. Ray, A., Stein, C. M., Hall, K., Daugherty, J. R. & Griffin, M. R. Non-steroidal anti-inflammatory drugs and risk of serious coronary heart disease: an observational cohort study. *The Lancet* **359**, 118–123 (2002).
20. Arfè, *et al.* Non-steroidal anti-inflammatory drugs and risk of heart failure in four european countries: nested case-control study. *bmj* **354**, i4857 (2016).
21. Bresalier, S. *et al.* Cardiovascular events associated with Rofecoxib in a colorectal adenoma chemoprevention trial. *New Engl. J. Medicine* **352**, 1092–1102 (2005).
22. Gunnell, A. S. *et al.* Long-term use and cost-effectiveness of secondary prevention drugs for heart disease in western australian seniors (wamach): a study protocol. *BMJ open* **4**, e006258 (2014).
23. Holman, D. J., Bass, A. J., Rouse, I. L. & Hobbs, M. S. Population-based linkage of health records in western australia: development of a health services research linked database. *Aust. New Zealand journal public health* **23**, 453–459 (1999).
24. Paige, E., Kemp-Casey, A., Korda, R., Banks, E. *et al.* Using australian pharmaceutical benefits scheme data for pharmacoepidemiological research: challenges and approaches. *Public Heal. Res. & Pract.* **25**, e2541546 (2015).
25. Organisation, H. *et al.* Atc/DDD index 2018. [EB/OL] (2018). https://www.whocc.no/atc_ddd_index/ Accessed 20 May, 2018.
26. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Annals statistics* **29**, 1189–1232 (2001).
27. Jalali, A. *et al.* Deep learning for improved risk prediction in surgical outcomes. *Reports* **10**, 1–13 (2020).
28. Hao, , Usama, M., Yang, J., Hossain, M. S. & Ghoneim, A. Recurrent convolutional neural network based multimodal disease risk prediction. *Futur. Gener. Comput. Syst.* **92**, 76–83 (2019).
29. Wu, -H. *et al.* Risk assessment of hypertension in steel workers based on lvq and fisher-svm deep excavation. *IEEE Access* **7**, 23109–23119 (2019).
30. Pedregosa, *et al.* Scikit-learn: Machine learning in python. *J. machine Learn. research* **12**, 2825–2830 (2011).
31. Ketkar, Introduction to keras. In *Deep learning with Python*, 97–111 (Springer, Berkeley, CA, 2017).
32. Fluss, , Faraggi, D. & Reiser, B. Estimation of the youden index and its associated cutoff point. *Biom. Journal: J. Math. Methods Biosci.* **47**, 458–472 (2005).
33. Choi, B., Lee, W., Yoon, J.-H., Won, J.-U. & Kim, D. W. Ten-year prediction of suicide death using cox regression and machine learning in a nationwide retrospective cohort study in south korea. *J. affective disorders* **231**, 8–14 (2018).
34. Yoo, D. *et al.* A machine learning approach using survival statistics to predict graft survival in kidney transplant recipients: a multicenter cohort study. *Sci. reports* **7**, 1–12 (2017).
35. Schultz, M. *et al.* Marital status and outcomes in patients with cardiovascular disease. *J. Am. Hear. Assoc.* **6**, e005890 (2017).

36. Wong, W. *et al.* Marital status and risk of cardiovascular diseases: a systematic review and meta-analysis. *Heart* **104**, 1937–1948 (2018).
37. Bradshaw, J., Alfonso, H. S., Finn, J., Owen, J. & Thompson, P. L. A comparison of coronary heart disease event rates among urban Australian Aboriginal people and a matched non-Aboriginal population. *J. Epidemiol. & Community Heal.* **65**, 315–319 (2011).
38. Katzenellenbogen, J. M. *et al.* Incidence of and case fatality following acute myocardial infarction in Aboriginal and non-Aboriginal Western Australians (2000–2004): a linked data *Heart Lung Circ.* **19**, 717–725 (2010).
39. Desai, J., Wang, S. V., Vaduganathan, M., Evers, T. & Schneeweiss, S. Comparison of machine learning methods with traditional models for use of administrative claims with electronic medical records to predict heart failure outcomes. *JAMA Netw. Open* **3**, e1918962–e1918962 (2020).
40. Bui, L., Horwich, T. B. & Fonarow, G. C. Epidemiology and risk profile of heart failure. *Nat. Rev. Cardiol.* **8**, 30 (2011).

Figures

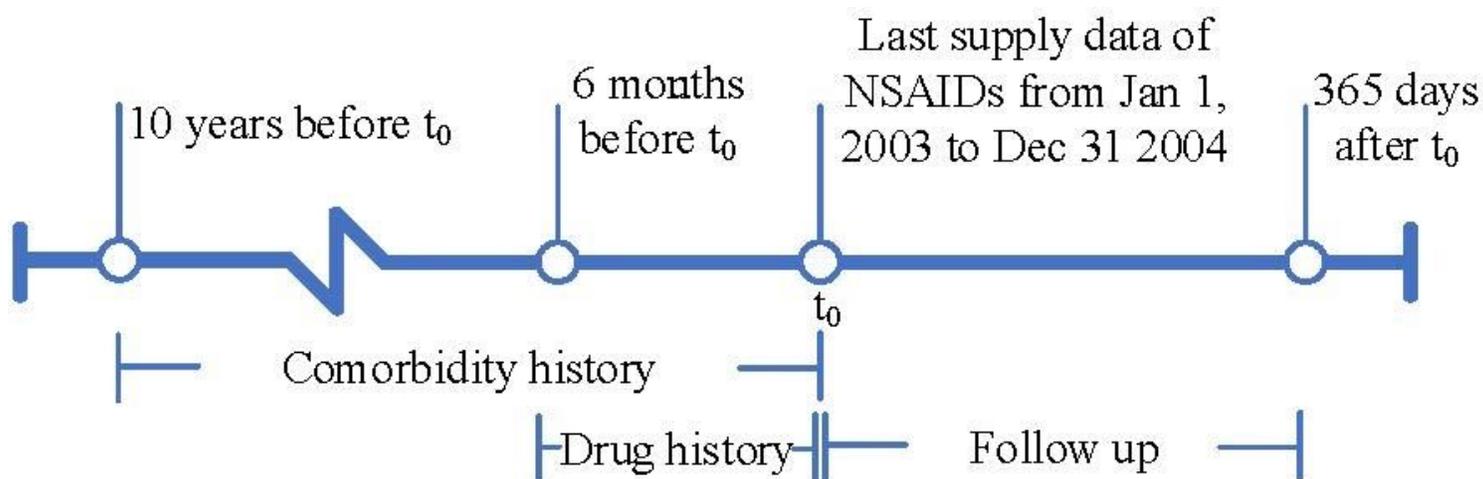


Figure 1

The timeline for study cohort identification. The first supply date for the COX-2 inhibitors or ibuprofen within years 2003 and 2004 was defined as t_0 .

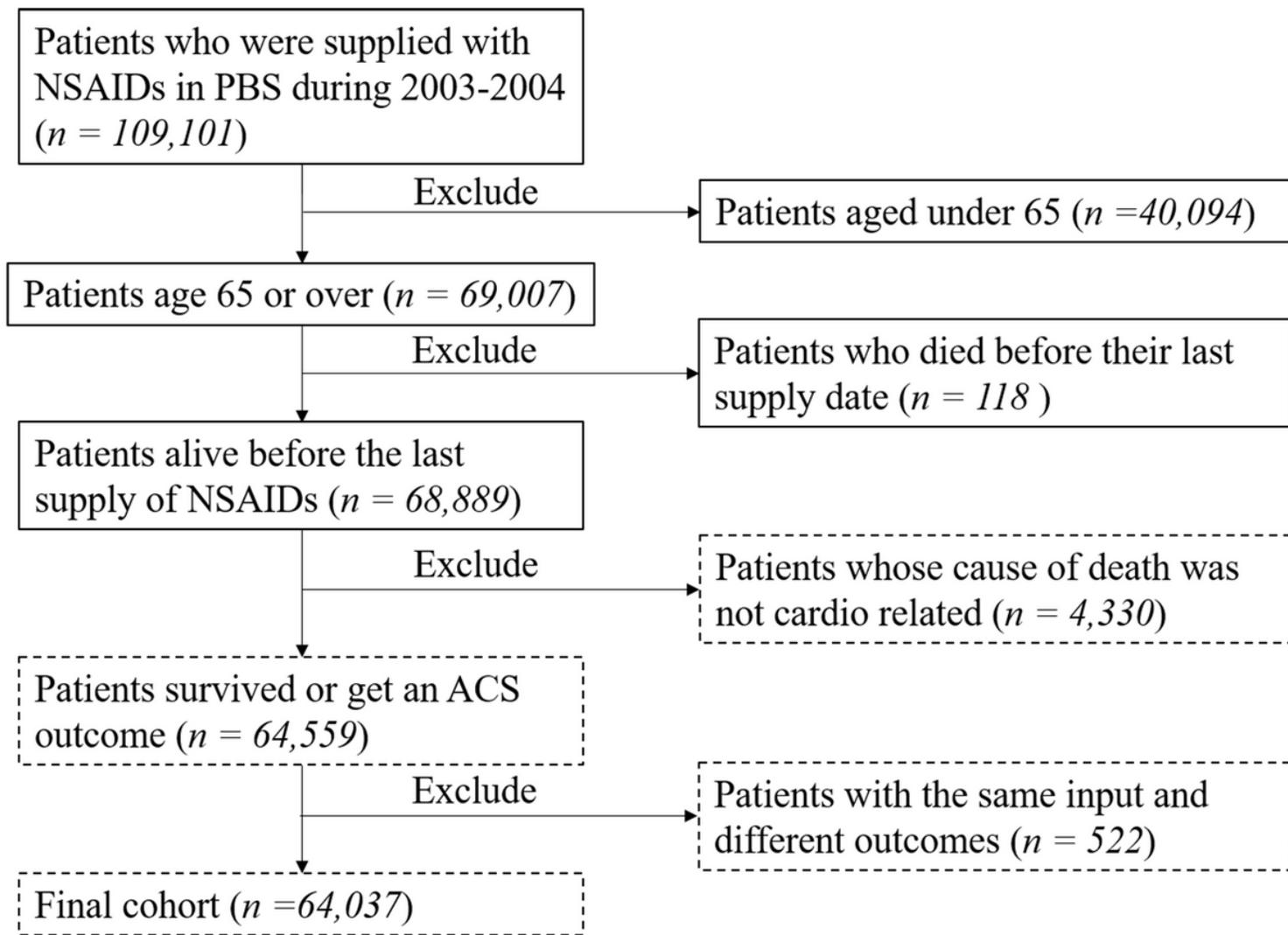


Figure 2

Flowchart showing identification of the study cohort. ACS, acute coronary syndrome.

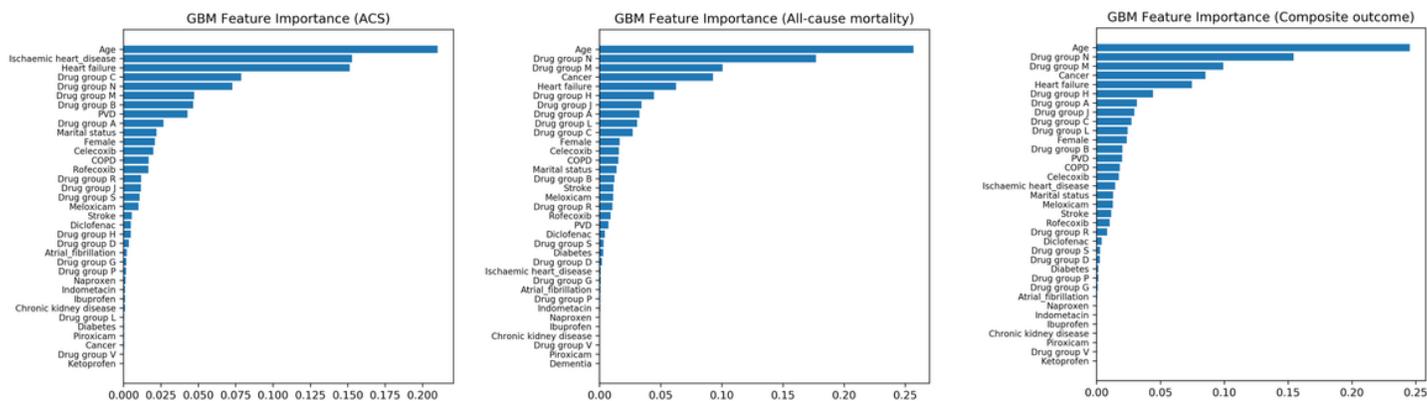


Figure 3

Feature importance of the prediction model for AE. (a) Feature importance for ACS; (b) Feature importance for All-cause death; (c) Feature importance for the composite outcome (ACS or all-cause

death)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [pharmacovigilancescientificreportssupplementary.pdf](#)