Table S1 : PRISMA 2020 checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	2-3
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	3
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	3
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	3
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Table S2
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	3，Figure 1
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	3
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points,	3

Table S1 : PRISMA 2020 checklist

Section and Topic	Item #	Checklist item	Location where item is reported
		analyses), and if not, the methods used to decide which results to collect.	
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Table S3,Table S4
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	3-4， Figure 2
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	4
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	4
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	4
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	4
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	3-4
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Table2,Table3
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	3
Reporting bias	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from	3

Table S1 : PRISMA 2020 checklist

Section and Topic	Item #	Checklist item	Location where item is reported
assessment		reporting biases).	
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	4
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	4
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Figure 1
Study characteristics	17	Cite each included study and present its characteristics.	4,Table S3
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	6,Figure 2,Table S5
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	6,Table2,Table3
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	6
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	6,Table 2
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	6-7,Table 2
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	6-8,Table 3
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	6,Figure 2
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Table 2,Table 3

Table S1 : PRISMA 2020 checklist

Section and Topic	Item #	Checklist item	Location where item is reported
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	8-10
	23b	Discuss any limitations of the evidence included in the review.	8-10
	23c	Discuss any limitations of the review processes used.	8-10
	23d	Discuss implications of the results for practice, policy, and future research.	10
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	4,Registration No. CRD42022346896
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	The protocol was registered on PROSPERO
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	NA
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	10
Competing interests	26	Declare any competing interests of review authors.	11
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Table S3,Table S4

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Table S2 Literature search strategy

1.Pubmed

Search number	Query	Results
#1	"Osteoporosis"[Mesh]	62,328
#2	"Osteoporosis"[Title/Abstract] OR "Osteoporoses"[Title/Abstract] OR "bone loss age related"[Title/Abstract] OR "age related bone loss"[Title/Abstract] OR "age related bone losses"[Title/Abstract] OR "bone loss age related"[Title/Abstract] OR ((("bone and bones"[MeSH Terms] OR ("Bone"[All Fields] AND "bones"[All Fields]) OR "bone and bones"[All Fields] OR "Bone"[All Fields]) AND "Losses"[All Fields]) AND "Age-Related"[Title/Abstract])	82,879
#3	"Osteoporosis"[MeSH Terms] OR ("Osteoporosis"[Title/Abstract] OR "Osteoporoses"[Title/Abstract] OR "bone loss age related"[Title/Abstract] OR "age related bone loss"[Title/Abstract] OR "age related bone losses"[Title/Abstract] OR "bone loss age related"[Title/Abstract] OR ((("bone and bones"[MeSH Terms] OR ("Bone"[All Fields] AND "bones"[All Fields]) OR "bone and bones"[All Fields] OR "Bone"[All Fields]) AND "Losses"[All Fields]) AND "Age-Related"[Title/Abstract]))	100,673
#4	"Machine Learning"[Mesh]	55,536
#5	"machine learning"[Title/Abstract] OR "transfer learning"[Title/Abstract] OR "deep learning"[Title/Abstract] OR "prediction model"[Title/Abstract] OR "artificial intelligence"[Title/Abstract] OR "random forest"[Title/Abstract] OR "artificial neural network"[Title/Abstract] OR "ANN"[Title/Abstract] OR "support vector machine"[Title/Abstract] OR "SVM"[Title/Abstract] OR "gradient boosting machine"[Title/Abstract] OR "GBM"[Title/Abstract] OR "Nomogram"[Title/Abstract] OR "XGboost"[Title/Abstract] OR "Logistic"[Title/Abstract] OR "decision tree"[Title/Abstract] OR "external validation"[Title/Abstract]OR "cox"[Title/Abstract]	708,017
#6	"Machine Learning"[MeSH Terms] OR "Machine Learning"[Title/Abstract] OR "transfer learning"[Title/Abstract] OR "deep learning"[Title/Abstract] OR "prediction model"[Title/Abstract] OR "artificial intelligence"[Title/Abstract] OR "random forest"[Title/Abstract] OR "artificial neural network"[Title/Abstract] OR "ANN"[Title/Abstract] OR "support vector machine"[Title/Abstract] OR	869,928

	"SVM"[Title/Abstract] OR "gradient boosting machine"[Title/Abstract] OR "GBM"[Title/Abstract] OR "Nomogram"[Title/Abstract] OR "XGboost"[Title/Abstract] OR "Logistic"[Title/Abstract] OR "decision tree"[Title/Abstract] OR "external validation"[Title/Abstract]OR "cox"[Title/Abstract]	
#7	"fractures, bone"[MeSH]	299,700
#8	"fractures bone"[Title/Abstract] OR "broken bones"[Title/Abstract] OR "bone broken"[Title/Abstract] OR "bones broken"[Title/Abstract] OR "broken bone"[Title/Abstract] OR "Fractures"[Title/Abstract] OR "Fracture"[Title/Abstract]	343,051
#9	"fractures, bone"[MeSH Terms] OR "fractures bone"[Title/Abstract] OR "broken bones"[Title/Abstract] OR "bone broken"[Title/Abstract] OR "bones broken"[Title/Abstract] OR "broken bone"[Title/Abstract] OR "Fractures"[Title/Abstract] OR "Fracture"[Title/Abstract]	325,361
#10	#3AND #6 AND #9	2,409

2.Cochrane

Search number	Query	Results
#1	MeSH descriptor: [Osteoporosis] explode all trees	5,754
#2	(Osteoporosis):ti,ab,kw OR (Osteoporoses):ti,ab,kw OR (Bone Loss, Age-Related):ti,ab,kw OR (Age-Related Bone Loss):ti,ab,kw OR (Age-Related Bone Losses):ti,ab,kw	11,868
#3	(Bone Loss, Age Related):ti,ab,kw OR (Bone Losses, Age-Related):ti,ab,kw	549
#4	#1 OR #2 OR #3	12,188
#5	MeSH descriptor: [Machine Learning] explode all trees	866
#6	(machine learning):ti,ab,kw OR (Transfer Learning):ti,ab,kw OR (Deep learning):ti,ab,kw OR (Prediction model):ti,ab,kw OR (artificial intelligence):ti,ab,kw	10,742
#7	(random forest):ti,ab,kw OR (artificial neural network):ti,ab,kw OR (ANN):ti,ab,kw OR (Support vector machine):ti,ab,kw OR (SVM):ti,ab,kw	3,194
#8	(Gradient Boosting Machine):ti,ab,kw OR (GBM):ti,ab,kw OR (Nomogram):ti,ab,kw OR (XGboost):ti,ab,kw OR (Logistic):ti,ab,kw	32,161

#9	(Decision tree):ti,ab,kw OR (External validation):ti,ab,kw	2,025
#10	#5 OR #6 OR #7 OR #8 OR #9	43,070
#11	MeSH descriptor: [Fractures, Bone] explode all trees	8,166
#12	(Fractures, Bone):ti,ab,kw OR (Broken Bones):ti,ab,kw OR (Bone, Broken):ti,ab,kw OR (Bones, Broken):ti,ab,kw OR (Broken Bone):ti,ab,kw	9,471
#13	(Fractures):ti,ab,kw OR (Fracture):ti,ab,kw	27,252
#14	#11 OR #12 OR #13	27,386
#15	#4 AND #10 AND #14	170

3.Embase

Search number	Query	Results
#1	'osteoporosis'/exp	152,054
#2	'osteoporosis':ab,ti OR 'osteoporoses':ab,ti OR 'bone loss, age-related':ab,ti OR 'age-related bone loss':ab,ti OR 'age-related bone losses':ab,ti OR 'bone loss, age related':ab,ti OR 'bone losses, age-related':ab,ti	121,472
#3	#1 OR #2	176,124
#4	'machine learning'/exp	377,384
#5	'machine learning':ab,ti OR 'transfer learning':ab,ti OR 'deep learning':ab,ti OR 'prediction model':ab,ti OR 'artificial intelligence':ab,ti OR 'random forest':ab,ti OR 'artificial neural network':ab,ti OR ann:ab,ti OR 'support vector machine':ab,ti OR svm:ab,ti OR 'gradient boosting machine':ab,ti OR gbm:ab,ti OR nomogram:ab,ti OR xgboost:ab,ti OR logistic:ab,ti OR 'decision tree':ab,ti OR 'external validation':ab,ti OR 'cox':ab,ti	1,265,200
#6	#4 OR #5	1,485,765
#7	'fracture'/exp	383,399
#8	'fractures, bone':ab,ti OR 'broken bones':ab,ti OR 'bone, broken':ab,ti OR 'bones, broken':ab,ti OR 'broken bone':ab,ti OR 'fractures':ab,ti OR 'fracture':ab,ti	363,644

#9	#7 OR #8	471,174
#10	#3 AND #6 AND #9	4,387

4.Web of science

Search number	Query	Results
#1	Osteoporosis (Topic) or Osteoporoses (Topic) or Bone Loss, Age-Related (Topic) or Age-Related Bone Loss (Topic) or Age-Related Bone Losses (Topic) or Bone Loss, Age Related (Topic) or Bone Losses, Age-Related (Topic)	210,210
#2	machine learning (Topic) or Transfer Learning (Topic) or Deep learning (Topic) or Prediction model (Topic) or artificial intelligence (Topic) or random forest (Topic) or artificial neural network (Topic) or ANN (Topic) or Support vector machine (Topic) or SVM (Topic) or Gradient Boosting Machine (Topic) or GBM (Topic) or Nomogram (Topic) or XGboost (Topic) or Logistic (Topic) or Decision tree (Topic) or External validation (Topic) or Cox (Topic)	3,698,410
#3	Fractures, Bone (Topic) or Broken Bones (Topic) or Bone, Broken (Topic) or Bones, Broken (Topic) or Broken Bone (Topic) or Fractures (Topic) or Fracture (Topic)	1,302,805
#4	#1 AND #2 AND #3	5,502

Table S3 Characteristics of included studies in meta-analysis

Author	Year	Country	Data source	Sample population type	Mean age,years	Fracture site	Total sample, n	Validation Method	ML method	Model evaluation metrics
Wu, Q[15]	2020	USA	gene database	men	74.8	multiple	5130	internal	LR ANN RF BT	AUC Sensitivity Specificity Accuracy
Villamor, E[16]	2020	Spain	clinical hospital	women	81.4	hip	137	internal	LR SVM ANN RF	Accuracy
Van Geel, Tacm[17]	2011	Netherlands	questionnaire collection	women	62	vertebral	2372	-	SM	AUC
Ulivieri, F. M[18]	2021	Italy	clinical hospital	patient	48.5	vertebral	90	-	ANN	Sensitivity Specificity Accuracy
Yoda, T[19]	2021	Japan	clinical hospital	patient	77.6	vertebral	97	internal	CNN	ROC AUC Sensitivity Specificity
Jiang, X. Z[20]	2013	USA	clinical hospital	women	61.4	multiple	615	-	LR	AUC Sensitivity Specificity Accuracy
Schousboe, J. T[21]	2014	USA	clinical hospital	women	75	vertebral	7233	-	LR	AUC ROC
Sandhu, S. K[22]	2010	Australia	electronic health record	patient	74	multiple	200	-	LR	AUC
Rubin, K. H[23]	2018	Denmark	administrative	subjects	61.4	multiple	2495339	internal	LR	AUC ROC Accuracy PPV NPV
Pluskiewicz, W[24]	2010	Poland	osteoporosis registry	women	68.5	multiple	2012	-	LR	ROC AUC
Jang, E. J[25]	2016	Korea	questionnaire collection	subjects	61	multiple	768	-	LR	C-statistics

Barret A. Monchka[26]	2021	Canada	osteoporosis registry	subjects	75	vertebral	12742	internal	CNN	AUC Sensitivity Specificity Accuracy PPV NPV
Mehta, S. D[27]	2020	USA	clinical hospital	patient	69	vertebral	307	internal	SVM	AUC ROC Sensitivity Specificity Accuracy PPV NPV
Langsetmo, L[28]	2011	Canada	questionnaire collection	subjects	67.6	multiple	5758	internal	SM	C-Statistics ROC
Ioannidis, G[29]	2017	Canada	electronic health record	subjects	61	multiple	29848	internal	DT LR	C-statistics
K. K. Nishiyama[30]	2013	Canada	questionnaire collection	women	73	multiple	116	internal	SVM	ROC AUC Sensitivity Specificity Accuracy
Kruse, C[31]	2017	Denmark	administrative	subjects	60.8	hip	7252	internal	DT NB	ROC AUC Sensitivity Specificity
Kolanu, N[32]	2021	Australia	electronic health record	patient	73.4	multiple	5416	external	ANN	ROC AUC Sensitivity Specificity
Kim, H. Y[33]	2016	Korea	administrative	subjects	60	multiple	718508	internal	SM	C-statistics
Hsieh, C. I[34]	2021	China	clinical hospital	patient	72.2	hip	36279	external	Other DL	AUC ROC Sensitivity Specificity Accuracy PPV NPV
Hong, N[35]	2021	Korea	clinical hospital	women	73	hip	2462	internal	SM	C-statistics

Ho-Le, T. P[36]	2017	Australia	osteoporosis registry	women	69.1	hip	1167	external	BT SVM ANN LR kNN SVM	AUC Sensitivity Specificity
Henry, M. J[37]	2011	Australia	osteoporosis registry	women	74	multiple	600	-	LR	AUC ROC Sensitivity Specificity
Galassi, A[38]	2020	Spain	electronic health record	women	81.4	hip	137	internal	DT LR RF SVM	Sensitivity Specificity Accuracy
FitzGerald, G[39]	2014	California	questionnaire collection	women	67	multiple	47429	-	SM	C-statistics
Ferizi, U[40]	2019	USA	osteoporosis registry	women	62	multiple	92	-	LR BT kNN SVM NB	AUC ROC Sensitivity Specificity
Enns-Bray, W. S[41]	2019	USA	clinical hospital	women	77.2	hip	254	-	LR	AUC ROC
Engels, A[42]	2020	Germany	administrative	patient	75.6	hip	78074	internal	SVM RF LR Ensemble learning	AUC ROC
De Vries, B. C. S[43]	2021	The Netherlands	clinical hospital	patient	68	multiple	9348	internal	BT ANN RF SM	C-statistics
Cheung, E. Y[44]	2012	China	electronic health record	women	62	multiple	2266	-	SM	AUC ROC Sensitivity Specificity
Chanplakorn, P[45]	2021	Thailand	osteoporosis registry	women	68.5	vertebral	617	-	SM	AUC ROC
Bredbenner, T.	2014	USA	clinical hospital	men	65	hip	922	internal	LR	AUC

L[46]										ROC
Beyaz, S[47]	2020	Turkey	osteoporosis registry	patient	74.9	multiple	2106	-	ANN	AUC ROC Sensitivity Specificity Accuracy
Berry, S. D[48]	2018	USA	administrative	subjects	84	hip	1278304	external	SM	C-statistics
Beaudoin, C[49]	2021	Canada	administrative	subjects	75.1	multiple	581281	internal	SM	C-statistics
Baleanu, F[50]	2022	Belgium	clinical hospital	women	70.1	multiple	3560	-	LR	AUC ROC
Almog, Y. A[51]	2020	USA	electronic health record	patient	50	vertebral	9806205	internal	ANN	AUC ROC Sensitivity Specificity
Zagorski, P[52]	2021	Poland	questionnaire collection	women	65.2	hip	389	-	LR	AUC ROC Sensitivity Specificity PPV NPV
Diez-Perez, A[53]	2007	Spain	questionnaire collection	women	72.3	multiple	5201	-	SM	AUC ROC
Lix, L. M[54]	2018	Canada	osteoporosis registry	women	65.6	multiple	31999	-	LR	AUC ROC
Li, Q. J[55]	2021	China	clinical hospital	patient	70	multiple	562	internal and external	LR	C-statistics
Lee, S[56]	2008	Korea	osteoporosis registry	women	65	multiple	94	-	SVM	Sensitivity Specificity
Jacobs, J. W. G[57]	2010	Portugal	questionnaire collection	subjects	66	vertebral	314	-	LR	AUC ROC Sensitivity Specificity
Eller-Vainicher, C[58]	2011	Italy	questionnaire collection	women	68	vertebral	372	-	ANN LR	AUC ROC Sensitivity Specificity

Zhong, B. Y[59]	2017	China	clinical hospital	patient	72	vertebral	421	internal	SM	Accuracy C-statistics
Xiao, X[60]	2021	USA	gene database	women	64.5	hip	699	-	SM	AUC
Du,J[61]	2022	China	clinical hospital	subjects	71	femur	120	-	SVM RF DT AdaBoost ANN XGBoost	Accuracy Specificity Recall Precision
Wang,M[62]	2022	China	clinical hospital	subjects	73.4	vertebral	7906	-	SM	AUC
Dong,Q[63]	2022	USA	clinical hospital	men	73.7	vertebral	3792	internal	Other DL	AUC ROC Sensitivity Specificity PPV NPV FDR F1 score Accuracy
Wen,Z[64]	2022	China	clinical hospital	patient	73.5	vertebral	270	internal	LR	AUC ROC Specificity Sensitivity PPV NPV Diagnostic efficiency
Pluskiewicz,W [65]	2023	Poland	questionnaire collection	women	66.4	multiple	640	-	LR	AUC
Kong,X[66]	2022	China	clinical hospital	patient	55.1	multiple	1730	-	SM	AUC NRI IDI
Agarwal,A[67]	2023	Canada	electronic health record	women	70.7	multiple	9716	external	SM	AUC ROC

*LR: Logistic Regression;ANN:artificial neural network;SVM = support-vector machine; CNN:convolutional neural network; kNN: k-nearest neighbors; RF: random forests;DT:decision tree;NB:Naive Bayes;BT:Boosted tree;SM:Survival model;DL:deep learnimg model;AUC:area under the receiver operating characteristic curve;ROC:receiver operating characteristic;PPV:positive predictive value;NPV:negative predictive value;FDR:R false discovery rate;NRI:net reclassifcation index; IDI:integrated discrimination improvement.

Table S4 Methodological characteristics of machine learning models developed for outcome prediction in patients with Osteoporosis

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
Wu, Q	2020	Train	M	Multiple fractures	361	4104	LR	10-fold cross validation	Median interpolation				
Wu, Q	2020	Train	M	Multiple fractures	361	4104	RF	10-fold cross validation	Median interpolation				
Wu, Q	2020	Train	M	Multiple fractures	361	4104	BT	10-fold cross validation	Median interpolation				
Wu, Q	2020	Train	M	Multiple fractures	361	4104	ANN	10-fold cross validation	Median interpolation				
Wu, Q	2020	Test	M	Multiple fractures	90	1026	LR	10-fold cross validation	Median interpolation	0.6410	0.7610	0.4420	0.6980
Wu, Q	2020	Test	M	Multiple fractures	90	1026	RF	10-fold cross validation	Median interpolation	0.7005	0.7000	0.4670	0.7590
Wu, Q	2020	Test	M	Multiple fractures	90	1026	BT	10-fold cross validation	Median interpolation	0.7100	0.5650	0.6930	0.8840
Wu, Q	2020	Test	M	Multiple fractures	90	1026	ANN	10-fold cross validation	Median interpolation	0.6910	0.7120	0.5980	0.8390
Villamor, E	2020	Train	F	Hip fracture	65	101	LR	10-fold cross validation					0.7669
Villamor, E	2020	Train	F	Hip fracture	65	101	SVM	10-fold cross validation					0.7569

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
Villamor, E	2020	Train	F	Hip fracture	65	101	ANN	10-fold cross validation					0.7642
Villamor, E	2020	Train	F	Hip fracture	65	101	RF	10-fold cross validation					0.6940
Villamor, E	2020	Test	F	Hip fracture	65	101	LR	10-fold cross validation					0.7309
Villamor, E	2020	Test	F	Hip fracture	65	101	SVM	10-fold cross validation					0.7835
Villamor, E	2020	Test	F	Hip fracture	65	101	ANN	10-fold cross validation					0.6940
Villamor, E	2020	Test	F	Hip fracture	65	101	RF	10-fold cross validation					0.7334
van Geel, Tacm	2011	Train	F	Vertebral fracture	382	2372	SM	Bootstrapping					
Ulivieri, F. M	2021	Train	F	Vertebral fracture	56	90	ANN			0.8300	0.7500	0.8372	
Yoda, T	2021	Train	M+F	Vertebral fracture	28	50	CNN	5-fold cross validation		0.9670	0.9250	0.9490	0.9380
Yoda, T	2021	Test	M+F	Vertebral fracture	21	47	CNN	5-fold cross validation		0.9840	0.9810	0.9490	0.9640
Jiang, X. Z	2013	Train	F	Multiple fractures	15	615	LR			0.7600	0.8100	0.4700	0.5100
Schousboe, J. T	2014	Train	F	Vertebral fracture	2883	7233	LR			0.6790			

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
Sandhu, S. K	2010	Train	F	Multiple fractures	47	144	LR			0.8400	0.7800	0.8000	
Sandhu, S. K	2010	Train	M	Multiple fractures	18	56	LR			0.7600	0.7400	0.8000	
Rubin, K. H	2018	Train	F	Multiple fractures	11898	647103	LR			0.7500	0.7520	0.5650	
Rubin, K. H	2018	Train	M	Multiple fractures	11851	647103	LR			0.7520	0.6450	0.6090	
Rubin, K. H	2018	Test	F	Multiple fractures	4762	600567	LR			0.8740	0.6000	0.6990	
Rubin, K. H	2018	Test	M	Multiple fractures	4776	600566	LR			0.8510	0.6300	0.5840	
Pluskiewicz, W	2010	Train	F	Hip fracture	1599	2012	LR			0.850	0.7590	0.7370	
Pluskiewicz, W	2010	Train	F	Multiple fractures	1704	2012	LR			0.8790	0.7390	0.5980	
Jang, E. J	2016	Train	M	Multiple fractures	36	363	LR			0.7390			
Jang, E. J	2016	Train	F	Multiple fractures	50	405	LR			0.7180			
Barret A. Monchka	2021	Train	M+F	Vertebral fracture	1470	8920	CNN			0.9500	0.8240	0.9430	0.9230
Mehta, S. D	2020	Train	M+F	Vertebral fracture	86	246	SVM	10-fold cross validation		0.9258	0.8950	0.9560	0.9350
Mehta, S. D	2020	Test	M+F	Vertebral fracture	22	61	SVM	10-fold cross validation		0.8963	0.8180	0.9740	0.9180
Langsetmo, L	2011	Test	M	Multiple fractures	139	1606	SM			0.7000			

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
Langsetmo, L	2011	Test	F	Multiple fractures	672	4152	SM			0.6900			
Ioannidis, G	2017	Train	M+F	Multiple fractures	3858	22386	DT			0.6690			
Ioannidis, G	2017	Test	M+F	Multiple fractures	1294	7462	DT			0.6870			
K. K. Nishiyama	2013	Train	F	Multiple fractures	44	88	SVM	10-fold cross validation		0.6800	0.5280	0.7970	0.6890
K. K. Nishiyama	2013	Test	F	Multiple fractures	14	28	SVM	10-fold cross validation		0.8000	0.6880	0.8850	0.8100
Kruse, C	2017	Train	F	Hip fracture	293	4722	NB	5-fold cross validation	random forest imputation	0.9200	0.8800	0.8100	
Kruse, C	2017	Train	M	Hip fracture	47	717	DT	5-fold cross validation	random forest imputation	0.8900	1.0000	0.6900	
Kolanu, N	2021	Train	M+F	Multiple fractures	433	5089	ANN				0.9900	0.9950	
Kolanu, N	2021	Test	M+F	Multiple fractures	97	327	ANN				0.6960	0.9500	
Kim, H. Y	2016	Train	M	Multiple fractures	4889	185127	SM			0.6800			
Kim, H. Y	2016	Train	F	Multiple fractures	14951	174126	SM			0.6500			
Kim, H. Y	2016	Test	M+F	Multiple fractures	19915	359255	SM			0.6650			
Hsieh, C. I	2021	Train	M+F	Hip fracture	2254	5164	Other DL	4-fold cross validation		0.9700	0.8820	0.9140	0.9000

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
Hsieh, C. I	2021	Test	M+F	Hip fracture	922	2060	Other DL	4-fold cross validation		0.9600	0.8990	0.9200	0.9100
Hsieh, C. I	2021	Train	M+F	Vertebral fracture	530	57662	Other DL	4-fold cross validation		0.9700	0.6960	0.9790	0.9500
Hsieh, C. I	2021	Test	M+F	Vertebral fracture	922	3346	Other DL	4-fold cross validation		0.9400	0.7400	0.9730	0.9480
Hong, N	2021	Train	F	Hip fracture	143	433	RF			0.7840			0.7300
Hong, N	2021	Train	F	Hip fracture	143	433	BT			0.7680			0.7200
Hong, N	2021	Train	F	Hip fracture	143	433	SVM			0.7590			0.7400
Hong, N	2021	Train	F	Hip fracture	143	433	BT			0.7580			0.7300
Hong, N	2021	Test	F	Hip fracture	34	2029	SM			0.8400			
Ho-Le, T. P	2017	Train	F	Hip fracture	54	700	ANN	5-fold cross validation			0.8890	0.8610	0.8630
Ho-Le, T. P	2017	Train	F	Hip fracture	54	700	LR	5-fold cross validation			0.9070	0.8640	0.8670
Ho-Le, T. P	2017	Train	F	Hip fracture	54	700	KNN	5-fold cross validation			1.0000	0.8330	0.8460
Ho-Le, T. P	2017	Train	F	Hip fracture	54	700	SVM	5-fold cross validation			0.9240	0.9690	0.9660
Ho-Le, T. P	2017	Test	F	Hip fracture	36	467	ANN	5-fold cross			0.8330	0.8770	0.8730

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
								validation					
Ho-Le, T. P	2017	Test	F	Hip fracture	36	467	LR	5-fold cross validation			0.7780	0.8180	0.8150
Ho-Le, T. P	2017	Test	F	Hip fracture	36	467	KNN	5-fold cross validation			0.8060	0.7930	0.7940
Ho-Le, T. P	2017	Test	F	Hip fracture	36	467	SVM	5-fold cross validation			0.8060	0.8160	0.8150
Henry, M. J	2011	Train	F	Multiple fractures	125	600	LR			0.7000	0.6420	0.6620	
Galassi, A	2020	Train	F	Hip fracture	62	96	LR				0.7033	0.7146	0.7081
Galassi, A	2020	Train	F	Hip fracture	62	96	SVM				0.9367	0.6292	0.8077
Galassi, A	2020	Train	F	Hip fracture	62	96	DT				0.5967	0.7446	0.6587
Galassi, A	2020	Train	F	Hip fracture	62	96	RF				0.8330	0.9231	0.8710
FitzGerald, G	2014	Train	F	Multiple fractures	2638	47429	SM			0.6670			
Ferizi, U	2019	Train	F	Multiple fractures	32	92	BT	23-fold cross validation		0.6200	0.5880	0.6670	0.6390
Ferizi, U	2019	Train	F	Multiple fractures	32	92	LR	23-fold cross validation		0.6200	0.5600	0.7010	0.6510
Ferizi, U	2019	Train	F	Multiple fractures	32	92	LR	23-fold cross validation		0.6190	0.5400	0.7010	0.6420

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
Ferizi, U	2019	Train	F	Multiple fractures	32	92	SVM	23-fold cross validation		0.5910	0.4490	0.7440	0.6410
Ferizi, U	2019	Train	F	Multiple fractures	32	92	kNN	23-fold cross validation		0.5060	0.2690	0.7420	0.5760
Ferizi, U	2019	Train	F	Multiple fractures	32	92	NB	23-fold cross validation		0.5650	0.4520	0.6790	0.6020
Enns-Bray, W. S	2019	Train	F	Hip fracture	95	254	LR			0.7270			
Engels, A	2020	Train	M+F	Hip fracture	6115	20456	LR	10-fold cross validation		0.7140			
Engels, A	2020	Train	M+F	Hip fracture	6115	20456	RF	10-fold cross validation		0.6860			
Engels, A	2020	Train	M+F	Hip fracture	6115	20456	SVM	10-fold cross validation		0.6600			
Engels, A	2020	Train	M+F	Hip fracture	6115	20456	BT	10-fold cross validation		0.7110			
Engels, A	2020	Train	M+F	Hip fracture	6115	20456	Ensemble learning	10-fold cross validation		0.7220	1.0000		
Engels, A	2020	Train	M+F	Hip fracture	6115	20456	BT	10-fold cross validation		0.7250			
Engels, A	2020	Test	M+F	Hip fracture	1529	57618	LR	10-fold cross validation		0.6950	1.0000		

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
Engels, A	2020	Test	M+F	Hip fracture	1529	57618	RF	10-fold cross validation		0.6850			
Engels, A	2020	Test	M+F	Hip fracture	1529	57618	SVM	10-fold cross validation		0.6500			
Engels, A	2020	Test	M+F	Hip fracture	1529	57618	BT	10-fold cross validation		0.7020			
Engels, A	2020	Test	M+F	Hip fracture	1529	57618	Ensemble learning	10-fold cross validation		0.6980			
Engels, A	2020	Test	M+F	Hip fracture	1529	57618	BT	10-fold cross validation		0.7030			
de Vries, B. C. S	2021	Train	M+F	Multiple fractures	805	7578	SM			0.6970			
de Vries, B. C. S	2021	Train	M+F	Multiple fractures	805	7578	ANN			0.6700			
de Vries, B. C. S	2021	Train	M+F	Multiple fractures	805	7578	RF			0.6870			
de Vries, B. C. S	2021	Test	M+F	Multiple fractures	165	1770	SM			0.6250			
de Vries, B. C. S	2021	Test	M+F	Multiple fractures	165	1770	ANN			0.5880			
de Vries, B. C. S	2021	Test	M+F	Multiple fractures	165	1770	RF			0.5930			
Cheung, E. Y	2012	Train	F	Multiple fractures	106	2266	SM			0.7300	0.8080	0.5170	
Chanplakorn, P	2021	Train	F	Vertebral fracture	179	617	LR			0.6500	0.4300	0.8600	

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
Bredbenner, T. L	2014	Train	M	Hip fracture	45	472	LR	10-fold cross validation		0.9300			
Beyaz, S	2020	Train	M+F	Multiple fractures	235	2106	CNN	5-fold cross validation			0.8250	0.6930	0.7770
Berry, S. D	2018	Train	M	Hip fracture	3541	119874	SM			0.6922			
Berry, S. D	2018	Train	F	Hip fracture	11012	299794	SM			0.7106			
Berry, S. D	2018	Test	M+F	Hip fracture	28050	858636	SM			0.6800			
Beaudoin, C	2021	Train	M+F	Multiple fractures	57678	307909	SM			0.6810			
Beaudoin, C	2021	Test	M+F	Multiple fractures	21809	273372	SM			0.6790			
Baleanu, F	2022	Train	F	Multiple fractures	410	3560	LR			0.7300			
Almog, Y. A	2020	Train	M+F	Vertebral fracture	2468694	6329986	ANN			0.8120	0.8120		0.1920
Almog, Y. A	2020	Test	M+F	Vertebral fracture	295479	3476219	ANN			0.6680	0.7070		0.1140
Zagorski, P	2021	Train	F	Hip fracture	49	389	LR			0.8840	0.9390	0.7120	
Diez-Perez, A	2007	Train	F	Multiple fractures	363	5201	SM			0.6720			
Lix, L. M	2018	Train	F	Multiple fractures	749	31999	LR			0.7060			
Li, Q. J	2021	Train	F	Multiple fractures	49	403	LR			0.8820			
Li, Q. J	2021	Test	F	Multiple fractures	17	159	LR			0.8690			

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
Lee, S	2008	Train	F	Multiple fractures	47	94	SVM				0.8500	0.4900	
Jacobs, J. W. G	2010	Train	M	Vertebral fracture	58	109	LR			0.5100			
Jacobs, J. W. G	2010	Train	F	Vertebral fracture	98	205	LR			0.7400	0.6700	0.7100	
Eller-Vainicher, C	2011	Train	F	Vertebral fracture	33	372	LR			0.8230	0.3730	0.9030	0.6380
Eller-Vainicher, C	2011	Train	F	Vertebral fracture	33	372	ANN			0.6990	0.7480	0.8780	0.8130
Zhong, B. Y	2017	Train	M+F	Vertebral fracture	33	256	SM			0.7800			
Zhong, B. Y	2017	Test	M+F	Vertebral fracture	23	165	SM			0.7200			
Xiao, X	2021	Train	F	Hip fracture	25	699	SM			0.8040			
Du,J	2022	Train	M+F	Femur fracture		96	SVM					0.6250	
Du,J	2022	Train	M+F	Femur fracture		96	RF					0.5000	
Du,J	2022	Train	M+F	Femur fracture		96	DT					0.5833	
Du,J	2022	Train	M+F	Femur fracture		96	Boosted tree					0.5000	
Du,J	2022	Train	M+F	Femur fracture		96	ANN					0.5833	
Du,J	2022	Train	M+F	Femur fracture		96	Boosted tree					0.5417	
Du,J	2022	Test	M+F	Femur fracture		24	SVM						0.9167

Author	Year	Data set	Gender	Fracture site	Events	Sample size	Model type	Verification method	Missing value processing	C-index	Sensitivity	Specificity	Accuracy
Du,J	2022	Test	M+F	Femur fracture		24	RF						0.8333
Du,J	2022	Test	M+F	Femur fracture		24	DT						0.9167
Du,J	2022	Test	M+F	Femur fracture		24	Boosted tree						0.8750
Du,J	2022	Test	M+F	Femur fracture		24	ANN						0.9583
Du,J	2022	Test	M+F	Femur fracture		24	Boosted tree						0.9167
Wang,M	2022	Train	M+F	Vertebral fracture	72	7906	SM	10-fold cross validation		0.820			
Dong,Q	2022	Train	M+F	Vertebral fractur		3413	Other DL			0.990	0.5980	0.9990	0.9950
Dong,Q	2022	Test	M+F	Vertebral fractur		379	Other DL			0.820	0.9770	0.9510	0.9510
Wen,Z	2022	Train	M+F	Vertebral fractur	208	220	LR			0.854	0.7310	0.8460	
Wen,Z	2022	Test	M+F	Vertebral fractur	50	50	LR			0.979	0.8942	0.9545	
Pluskiewicz,W	2023	Train	F	Multiple fractures	129	640	LR			0.660			
Kong,X	2022	Train	M+F	Multiple fractures	109	1730	SM	Bootstrapping	Mean interpolation	0.803			
Agarwal,A	2023	Test	F	Multiple fractures	264	9716	SM			0.710			

*M:Male;F:Female;LR:Logistic Regression;ANN:artificial neural network;SVM:support-vector machine; CNN:convolutional ANN; kNN:k-nearest neighbors; RF:random forests;DT:decision tree;BT:Boosted tree;SM:Survival model;NB:Naive Bayes;DL:deep learnimg model.

.

Table S5 Risk of bias assessment grading of the machine learning predictive modelling studies of osteoporosis populations as per the PROBAST criteria

Study	Participants bias	Predictors bias	Outcome bias	Analysis bias	Overall bias rating
Wu, Q	low	low	low	high	high
Wu, Q	low	low	low	high	high
Wu, Q	low	low	low	high	high
Wu, Q	low	low	low	high	high
Villamor, E	high	unclear	unclear	high	high
Villamor, E	high	unclear	unclear	high	high
Villamor, E	high	unclear	unclear	high	high
Villamor, E	high	low	unclear	high	high
Van Geel, Tacm	low	low	low	high	high
Ulivieri, F. M	low	low	low	high	high
Yoda, T	low	low	low	high	high
Jiang, X. Z	low	low	low	high	high
Schousboe, J. T	high	low	low	unclear	high
Sandhu, S. K	high	unclear	unclear	high	high
Rubin, K. H	low	low	low	unclear	unclear
Pluskiewicz, W	high	low	low	unclear	high
Jang, E. J	low	low	low	high	high
Barret A. Monchka	high	low	low	unclear	high
Mehta, S. D	high	unclear	unclear	high	high
Langsetmo, L	low	low	low	unclear	unclear
Ioannidis, G	high	low	low	unclear	high
K. K. Nishiyama	low	low	low	high	high

Kruse, C	low	low	low	unclear	unclear
Kruse, C	low	low	low	unclear	unclear
Kolanu, N	high	low	low	unclear	high
Kim, H. Y	low	low	low	unclear	unclear
Hsieh, C. I	low	low	low	unclear	unclear
Hong, N	low	low	low	unclear	unclear
Hong, N	low	low	low	unclear	unclear
Hong, N	low	low	low	unclear	unclear
Hong, N	low	low	low	unclear	unclear
Hong, N	low	low	low	unclear	unclear
Ho-Le, T. P	low	low	low	high	high
Ho-Le, T. P	low	low	low	high	high
Ho-Le, T. P	low	low	low	high	high
Ho-Le, T. P	low	low	low	high	high
Henry, M. J	low	low	low	unclear	unclear
Galassi, A	low	low	low	high	high
Galassi, A	low	low	low	high	high
Galassi, A	low	low	low	high	high
Galassi, A	low	low	low	high	high
FitzGerald, G	low	low	low	unclear	unclear
Ferizi, U	high	unclear	unclear	high	high
Ferizi, U	high	unclear	unclear	high	high
Ferizi, U	high	unclear	unclear	high	high
Ferizi, U	high	unclear	unclear	high	high
Ferizi, U	high	unclear	unclear	high	high
Ferizi, U	high	unclear	unclear	high	high

Enns-Bray, W. S	high	low	low	high	high
Engels, A	low	low	low	unclear	unclear
Engels, A	low	low	low	unclear	unclear
Engels, A	low	low	low	unclear	unclear
Engels, A	low	low	low	unclear	unclear
Engels, A	low	low	low	unclear	unclear
Engels, A	low	low	low	unclear	unclear
de Vries, B. C. S	high	low	low	unclear	high
de Vries, B. C. S	high	low	low	unclear	high
de Vries, B. C. S	high	low	low	unclear	high
Cheung, E. Y	low	low	low	unclear	unclear
Chanplakorn, P	high	low	low	unclear	high
Bredbenner, T. L	high	unclear	unclear	high	high
Beyaz, S	high	low	low	unclear	high
Berry, S. D	low	low	low	unclear	unclear
Beaudoin, C	high	low	low	unclear	high
Baleanu, F	low	low	low	unclear	unclear
Baleanu, F	low	low	low	unclear	unclear
Almog, Y. A	high	low	low	unclear	unclear
Zagorski, P	low	low	low	high	high
Diez-Perez, A	low	low	low	unclear	unclear
Lix, L. M	low	low	low	unclear	unclear
Li, Q. J	high	low	low	high	high
Lee, S	high	unclear	low	high	high
Jacobs, J. W. G	low	low	low	unclear	unclear
Eller-Vainicher, C	low	low	low	high	high

Eller-Vainicher, C	low	low	low	high	high
Zhong, B. Y	high	low	low	high	high
Xiao, X	low	low	low	high	high
Du,J	low	low	low	high	high
Du,J	low	low	low	high	high
Du,J	low	low	low	high	high
Du,J	low	low	low	high	high
Du,J	low	low	low	high	high
Du,J	low	low	low	high	high
Wang,M	high	low	low	high	high
Dong,Q	low	low	low	unclear	unclear
Wen,Z	high	low	low	high	high
Pluskiewicz,W	low	low	low	high	high
Kong,X	high	low	low	high	high
Agarwal,A	low	low	low	unclear	unclear

*When a single study included multiple models, risk of bias concerns were assessed for each model.

Fig.S1 Forest plots for subgroup analysis of C-index statistics by fracture site and machine learning type in training set

Subgroup	Number of models	Cindex(95%CI)
Fracture site		
Vertebral body	15	0.80(0.74,0.87)
Hip	20	0.76(0.72,0.81)
Multiple site	31	0.70(0.67,0.72)
Model type		
LR	26	0.75(0.72,0.78)
ANN	4	0.73(0.64,0.82)
CNN	2	0.95(0.94,0.96)
RF	3	0.70(0.68,0.72)
SVM	5	0.72(0.60,0.85)
DT	2	0.78(0.56,0.99)
kNN	1	0.51(0.46,0.55)
NB	2	0.74(0.39,1.00)
Survival model	13	0.70(0.69,0.72)
Boosted tree	5	0.71(0.68,0.74)
Enseble learning	1	0.72(0.71,0.73)
other DL	2	0.97(0.96,0.97)
Overall	66	0.75(0.72,0.78)

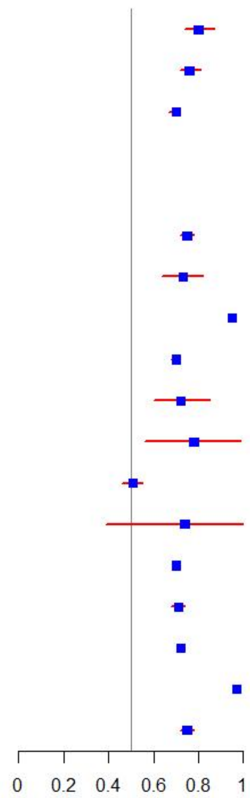

Fig.S2 Forest plots for subgroup analysis of C-index statistics by fracture site and machine learning type in validation set

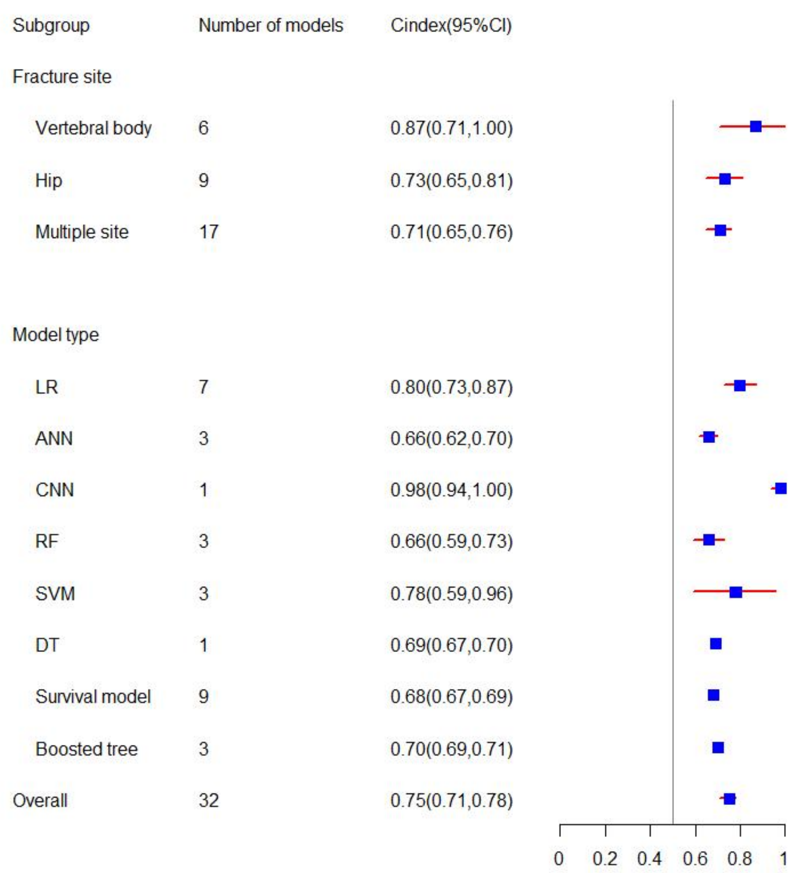

Fig.S3 Sensitivity analysis of multiple fracture model in training set

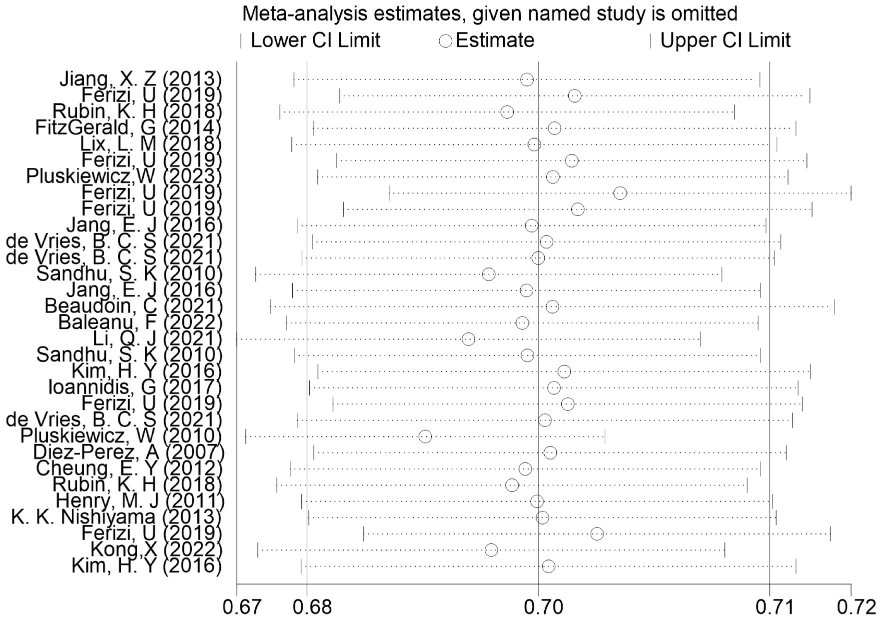

Fig.S4 Sensitivity analysis of vertebral fracture model in training set

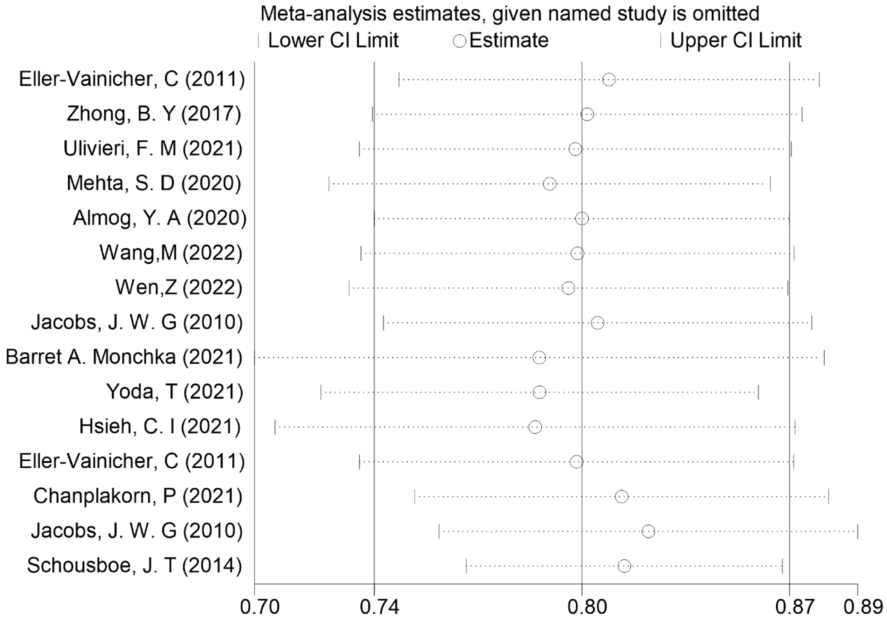

Fig.S5 Sensitivity analysis of hip fracture model in training set

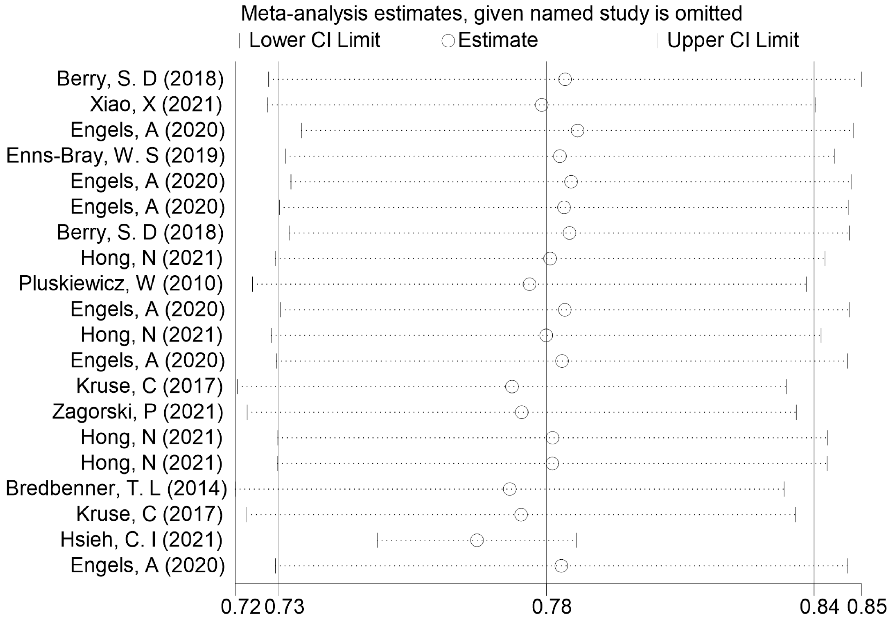

Fig.S6 Sensitivity analysis of multiple fracture model in validation set

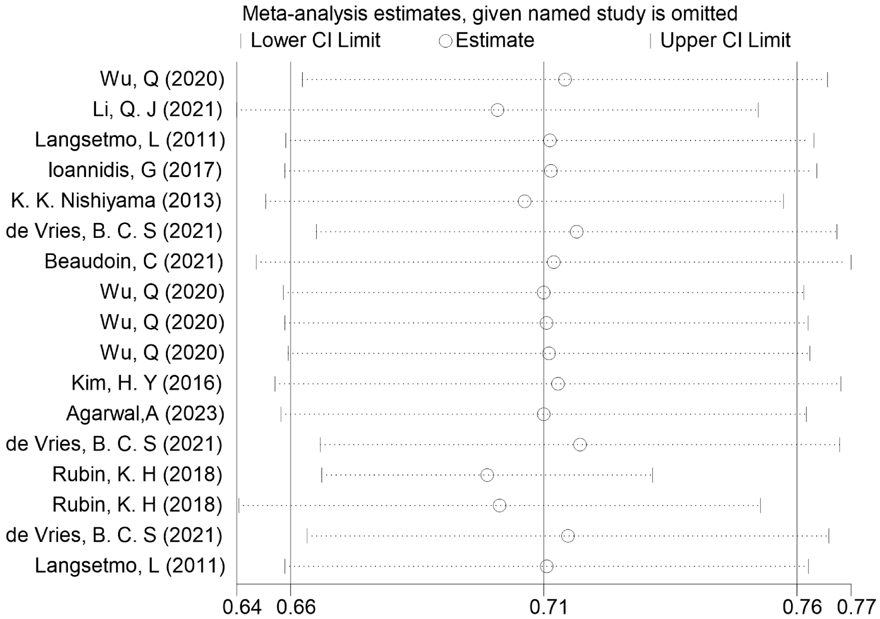

Fig.S7 Sensitivity analysis of vertebral fracture model in validation set

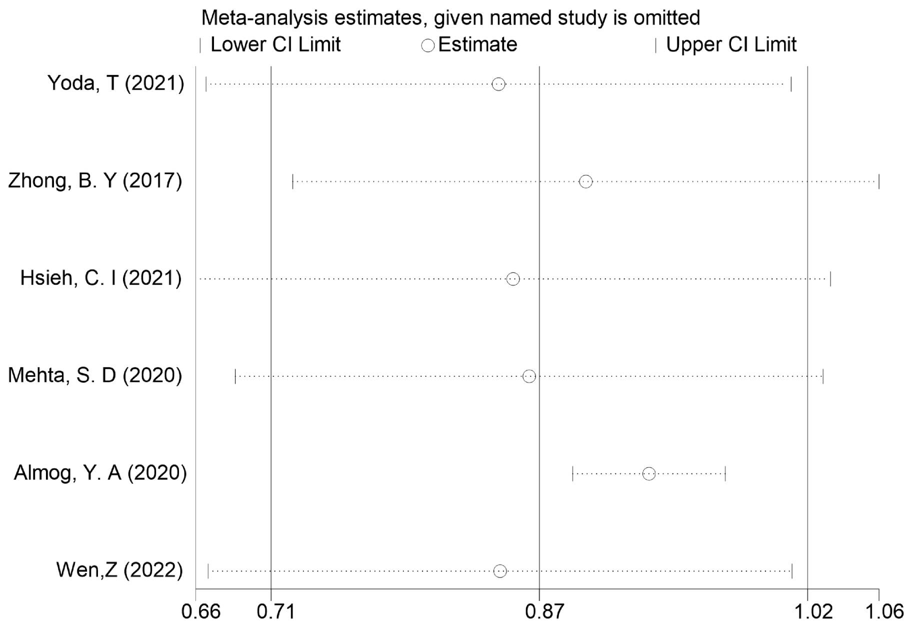

Fig.S8 Sensitivity analysis of hip fracture model in validation set