

Online Supplementary file 2 – Search strategy

Keyword searches conducted in the databases *PubMed*, *EMBASE* and *Web of Science* as part of the scoping review.

1) PubMed Query

Search: (((("Machine learning" OR "statistical learning" OR "supervised learning" OR "unsupervised learning")) OR ("Machine Learning"[Mesh]) AND ("Biomarkers"[Mesh]) OR "Precision Medicine"[Mesh]) AND (((Omic* OR "omic based" OR "multi omic" OR "multi omics" OR genomic* OR transcriptomic* OR proteomic* OR metabolomic* OR lipidomic* OR epigenomic* OR microarray OR "RNA seq" OR "mass spectrometry")) OR ("Genomics"[Mesh] OR "Metabolomics"[Mesh] OR "Epigenomics"[Mesh] OR "Microarray Analysis"[Mesh] OR "Mass Spectrometry"[Mesh])) AND ((validation OR validity OR validated OR "cross validation" "cross validated" OR "clinical utility*" OR accuracy OR robustness OR reliability* OR sensitivity OR specificity OR benchmark* OR bias OR "cross study" OR "cross studies")) AND ("2000/01/01"[Date - Entry]: "2021/07/20"[Date - Entry])) Filters: English, French, Italian, Spanish

2) Embase Query

#25: #24 AND [embase]/lim NOT [medline]/lim
#24: #23 AND [2000-2021]/py
#23: #20 AND #21 AND ([english]/lim OR [french]/lim OR [italian]/lim OR [spanish]/lim)
#22: #20 AND #21
#21: omic*:ti,ab OR 'machine learning':ti,ab OR 'personalized medicine':ti,ab OR 'personalised medicine':ti,ab
#20: #4 AND #10 AND #16 AND #19
#19: #17 OR #18
#18: validation:ti,ab OR validity:ti,ab OR validated:ti,ab OR 'cross validation':ti,ab OR 'cross validated':ti,ab OR test*:ti,ab OR 'clinical utility*':ti,ab OR accuracy:ti,ab OR robustness:ti,ab OR reliability*:ti,ab OR sensitivity:ti,ab OR specificity:ti,ab OR benchmark*:ti,ab OR bias:ti,ab OR 'cross study:ti,ab' OR 'cross studies':ti,ab
#17: 'validation study'/exp OR 'reliability'/exp OR 'sensitivity and specificity'/exp OR 'benchmarking'/exp
#16: #14 OR #15
#15: omic*:ti,ab OR 'omic based':ti,ab OR 'multi omic*':ti,ab OR genomic*:ti,ab OR transcriptomic*:ti,ab OR proteomic*:ti,ab OR metabolomic*:ti,ab OR lipidomic*:ti,ab OR epigenomic*:ti,ab OR microarray:ti,ab OR 'rna seq':ti,ab OR 'mass spectrometr*':ti,ab
#14: #11 OR #12 OR #13
#13: 'mass spectrometry'/exp
#12: 'microarray analysis'/exp
#11: 'omics'/exp OR 'genomics'/exp OR 'epigenetics'/exp
#10: #5 OR #6 OR #7 OR #8 OR #9
#9: 'individualized medicine':ti,ab OR 'individualised medicine':ti,ab OR 'individualized therapy':ti,ab OR 'individualised therapy':ti,ab
#8: 'personalised medicine':ti,ab
#7: 'personalized medicine':ti,ab
#6: 'stratified medicine':ti,ab OR cluster*:ti,ab OR 'sub group*':ti,ab OR subgroup*:ti,ab OR biomarker*:ti,ab OR diagnos*:ti,ab OR prognos*:ti,ab OR 'precision medicine':ti,ab
#5: 'biological marker'/exp OR 'personalized medicine'/exp
#4: #1 OR #2 OR #3
#3: 'machine learning'/exp
#2: 'statistical learning'/exp

#1: 'machine learning':ti,ab OR 'statistical learning':ti,ab OR 'supervised learning':ti,ab OR 'unsupervised learning':ti,ab

3) Web of Science Query

(((((#1) AND #1) AND #2) AND #3) AND #4) AND #5

5: (ALL=((validation OR validity OR validated OR "cross validation" "cross validated" OR "clinical utility*" OR accuracy OR robustness OR reliability* OR sensitivity OR specificity OR benchmark*

OR bias OR "cross study" OR "cross studies")))) AND ALL=((omic* OR "machine learning" OR

"personalized medicine" OR "personalised Medicine")))

4: ALL=((validation OR validity OR validated OR "cross validation" "cross validated" OR "clinical utility*" OR accuracy OR robustness OR reliability* OR sensitivity OR specificity OR benchmark*

OR bias OR "cross study" OR "cross studies"))))

3: ALL=(TOPIC: ((Omic* OR "omic based" OR "multi omic*" OR genomic* OR transcriptomic* OR proteomic* OR metabolomic* OR lipidomic* OR epigenomic* OR microarray OR "RNA seq"

OR "mass spectrometr*")))

2: (ALL=(TOPIC: ("Machine learning" OR "statistical learning" OR "supervised learning" OR "unsupervised learning")))) AND ALL=(TOPIC: ("stratified medicine" OR cluster* OR "sub group*" OR Subgroup* OR biomarker* OR diagnos* OR prognos* OR "precision medicine" OR

"personalized medicine" OR "personalised medicine" OR "individualized Medicine" OR "individualised Medicine" OR "individualized therapy" OR "individualised therapy")))

1: ALL=(TOPIC: ("Machine learning" OR "statistical learning" OR "supervised learning" OR "unsupervised learning"))))