Online supplementary appendices

Appendix 1 Search strategies

Table 1. Summary of electronic database searches and dates

Database	Platform	Searched on date	Date range of search	Update search
MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print	Ovid SP	09 September 2020	2010 to Present	17 May 2021
Embase	Ovid SP	09 September 2020	2010 to 2020 Week 36	17 May 2021
The Cochrane Library, including: - **Cochrane Database of Systematic Reviews (CDSR)** - **Cochrane Central Register of Controlled Trials (CENTRAL)** - **Database of Abstracts of Reviews of Effects (DARE)**	Wiley Online	09 September 2020	January 2010 to September 2020	17 May 2021
Web of Science	Ovid SP	09 September 2020	2010 to 2020	17 May 2021

Database: Ovid MEDLINE(R) ALL <1946 to September 08, 2020>
Search Strategy:
----
1 exp Breast Neoplasms/ (293428)
2 (breast adj5 (cancer* or neoplasm* or tumor* or tumour*or carcino* or malignan*or disease*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (394921)
3 1 or 2 (395368)
4 exp artificial intelligence/ or exp machine learning/ or exp deep learning/ or exp supervised machine learning/ or exp support vector machine/ or exp unsupervised machine learning/ (99304)
5 ai.mp. (28888)
6 ((artificial or machine or deep) adj5 (intelligence or learning or reasoning)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (70647)
7 exp Neural Networks, Computer/ or exp Algorithms/ or neural network*.mp. (354996)
8 exp Diagnosis, Computer-Assisted/ (83632)
9 4 or 5 or 6 or 7 or 8 (452038)
10 3 and 9 (10492)
11 exp Mammography/ (30025)
12 mammogra*.mp. (41086)
13 screen*.mp. or exp Mass Screening/ (844672)

14 exp "Early Detection of Cancer"/ or early detect*.mp. (86182)
15 11 or 12 or 13 or 14 (921015)
16 10 and 15 (3324)
17 exp "Sensitivity and Specificity"/ or sensitivity.mp. or specificity.mp. (1898406)
18 exp "Predictive Value of Tests"/ (203774)
19 exp roc curve/ or roc.mp. or receiver operating characteristic*.mp. (119948)
20 exp Area Under Curve/ or auc.mp. (96772)
21 exp False Positive Reactions/ (27763)
22 exp False Negative Reactions/ (17783)
23 exp Observer Variation/ (42540)
24 exp Diagnostic Errors/ (116740)
25 (false adj4 (negativ* or positiv*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (100096)
26 (true adj4 (positiv* or negativ*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (10701)
27 likelihood ratio*.mp. (15918)
28 ((predict* or test*) adj1 (value* or accura* or error*)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (342222)
29 exp Reproducibility of results/ (403133)
30 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 (2397952)
31 Randomized controlled trials as Topic/ (135939)
32 Randomized controlled trial/ (512638)
33 Random allocation/ (103549)
34 Double blind method/ (159672)
35 Single blind method/ (28987)
36 Clinical trial/ (524613)
37 exp Clinical Trials as Topic/ (345470)
38 (clinic$ adj trial$1).tw. (373312)
39 ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw. (174345)
40 Randomly allocated.tw. (29185)
41 (allocated adj2 random).tw. (802)
42 (test-treat trial* or test treat trial*).mp. (1)
43 or/31-42 (1423959)
44 30 or 43 (3676443)
45 3 and 9 and 15 and 44 (2179)
46 Case report.tw. (316143)
47 Letter/ (1098543)
48 Historical article/ (359991)
49 Review of reported cases.pt. (0)
50 Review, multicase.pt. (0)
51 or/46-50 (1758549)
52 45 not 51 (2164)
53 30 or 52 (2398016)
54 3 and 9 and 15 and 44 (2179)
55 54 not 51 (2164)
56 limit 55 to (english language and yr="2010 -Current") (1228)

Database: Embase <1980 to 2020 Week 36>
Search Strategy:
----
1 exp breast tumor/ (522987)
2 exp breast cancer/ (459643)
3 (breast adj5 (neoplasm* or cancer* or tumor* or tumour* or malignanc* or carcino* or disease*)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (609633)
4 or/1-3 (617527)
5 exp artificial intelligence/ (41063)
6 exp machine learning/ (215028)
7 exp deep learning/ (9250)
8 exp supervised machine learning/ (1511)
9 exp support vector machine/ (22089)
10 exp unsupervised machine learning/ (745)
11 ai.mp. (37967)
12 ((artificial or machine or deep) adj5 (intelliegence or learning or reasoning)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (64514)
13 exp artificial neural network/ or neural network*.mp. (76598)
14 exp algorithm/ (381894)
15 exp computer assisted diagnosis/ (1123074)
16 or/5-15 (1645988)
17 exp mammography/ or mammogra*.mp. (62455)
18 screen*.mp. (1308438)
19 exp mass screening/ or exp screening/ (661930)
20 exp early cancer diagnosis/ or early detect*.mp. (97077)
21 or/17-20 (1419229)
22 exp "sensitivity and specificity"/ or sensitivity.mp. or specificity.mp. (1803827)
23 exp reproducibility/ (217747)
24 exp receiver operating characteristic/ or exp roc curve/ or roc.mp. (163201)
25 exp predictive value/ or ((predict* or test*) adj1 (value* or error* or accura*)).mp. (420596)
26 auc.mp. or exp area under the curve/ (211311)
27 exp false positive result/ (30242)
28 exp false negative result/ (18705)
29 exp observer variation/ (19992)
30 exp diagnostic error/ (97269)
31 (false adj4 (negativ* or positiv*)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (120409)
32 (true adj4 (positiv* or negativ*)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (15519)
33 likelihood ratio.mp. (16617)
34 or/22-33 (2487428)
35 clinical trial/ (971925)
36 Randomized controlled trial/ (614311)
37 Randomization/ (87593)
38 Single blind procedure/ (40000)
39 Double blind procedure/ (172538)
40 Crossover procedure/ (64123)

41 Randomi?ed controlled trial$.tw. (236002)
42 Rct.tw. (38207)
43 Random allocation.tw. (2050)
44 Randomly allocated.tw. (35853)
45 Allocated randomly.tw. (2566)
46 (allocated adj2 random).tw. (822)
47 Single blind$.tw. (25171)
48 Double blind$.tw. (204835)
49 ((treble or triple) adj blind$).tw. (1182)
50 Prospective study/ (622435)
51 (test-treat trial* or test treat trial*).mp. (2)
52 or/35-51 (2076152)
53 34 or 52 (4332240)
54 Case study/ (71546)
55 Case report.tw. (410369)
56 Abstract report/ or letter/ (1114503)
57 or/54-56 (1585513)
58 53 not 57 (4229367)
59 4 and 16 and 21 and 58 (5034)
60 limit 59 to (english language and yr="2010 -Current") (3562)
61 limit 60 to (article or article in press or "review") (2808)

Database: Web of Science (Ovid SP)

# 5	881	#4 AND #3 AND #2 AND #1 *Indexes=SCI-EXPANDED, SSCI Timespan=2010-2020*
# 4	1,576,497	TS=("sensitivity and specificity" or sensitivity or specificity or ((predict* or test*) NEAR/1 (value* or error* or accura*)) or roc or "receiver operating characteristic*" or auc or "area under curve" or "observer variation" or "diagnostic error*") OR TS=(false NEAR/4 (negativ* or positiv*)) OR TS=(true NEAR/4 (negativ* or positiv*)) OR TS=("liklihood ratio*" or reproducibility) OR TS=(rct* or "randomi?ed controlled trial*" or "random allocat*" or "double blind*" or "single blind*" or "clinical trial*" or "test treat trial*" or "test-treat trial*") OR TS=((singl* or doubl* or treb* or tripl*) NEAR/1 (blind* or mask*)) OR TS=((random*) Near/2 (allocat*)) *Indexes=SCI-EXPANDED, SSCI Timespan=2010-2020*
# 3	538,555	**TOPIC:** (mammogra* or screen* or "early detect*") *Indexes=SCI-EXPANDED, SSCI Timespan=2010-2020*
# 2	895,801	**TOPIC:** ("artificial intelligence" or "machine learning" or "deep learning" or "support vector machine*" or ai) *OR* **TOPIC:** ((artificial or machine or deep) Near/5 (intelligence or learning or reasoning)) *OR* **TOPIC:** ("neural network*" or algorithm*) *OR* **TOPIC:** (diagnosis NEAR/3 computer*) *Indexes=SCI-EXPANDED, SSCI Timespan=2010-2020*
# 1	306,059	TS=((breast) NEAR/5 (neoplasm* or cancer* or tumor* or tumour* or malignan* or carcino* or disease*)) *Indexes=SCI-EXPANDED, SSCI Timespan=2010-2020*

Database: Cochrane Library (CENTRAL) (Wiley online)

Search Name: AI and Breast Cancer
Last Saved: 09/09/2020 14:10:48
Comment: Numbers for individual search lines are not captured by the saved search strategy.

ID	Search
#1	MeSH descriptor: [Breast Neoplasms] explode all trees
#2	((breast NEAR/5 (cancer* or neoplasm* or carcino* or malignan* or tumor* or tumour* or disease*))):ti,ab,kw
#3	#1 or #2
#4	MeSH descriptor: [Artificial Intelligence] explode all trees
#5	MeSH descriptor: [Machine Learning] explode all trees
#6	MeSH descriptor: [Deep Learning] explode all trees
#7	MeSH descriptor: [Supervised Machine Learning] explode all trees
#8	MeSH descriptor: [Support Vector Machine] explode all trees
#9	MeSH descriptor: [Unsupervised Machine Learning] explode all trees
#10	(ai):ti,ab,kw
#11	((artificial or machine or deep) NEAR/5 (intelligence or learning or reasoning)):ti,ab,kw
#12	MeSH descriptor: [Neural Networks, Computer] explode all trees
#13	MeSH descriptor: [Algorithms] explode all trees
#14	(neural network*):ti,ab,kw
#15	MeSH descriptor: [Diagnosis, Computer-Assisted] explode all trees
#16	#4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15
#17	MeSH descriptor: [Mammography] explode all trees
#18	(mammogra*):ti,ab,kw
#19	MeSH descriptor: [Mass Screening] explode all trees
#20	(screen*):ti,ab,kw
#21	MeSH descriptor: [Early Detection of Cancer] explode all trees
#22	(early detect*):ti,ab,kw
#23	#17 or #18 or #19 or #20 or #21 or #22
#24	#3 and #16 and #23
#25	MeSH descriptor: [Sensitivity and Specificity] explode all trees
#26	(sensitivity or specificity):ti,ab,kw
#27	MeSH descriptor: [Predictive Value of Tests] explode all trees
#28	MeSH descriptor: [ROC Curve] explode all trees
#29	(roc or "receiver operating characteristic"):ti,ab,kw
#30	MeSH descriptor: [Area Under Curve] explode all trees

#31 (auc):ti,ab,kw
#32 MeSH descriptor: [False Positive Reactions] explode all trees
#33 MeSH descriptor: [False Negative Reactions] explode all trees
#34 MeSH descriptor: [Observer Variation] explode all trees
#35 MeSH descriptor: [Diagnostic Errors] explode all trees
#36 (false NEAR/4 (negativ* or positiv*)):ti,ab,kw
#37 (true NEAR/4 (positiv* or negativ*)):ti,ab,kw
#38 (liklihood ratio*):ti,ab,kw
#39 ((predict* or test*) NEAR/1 (value* or accura* or error*)):ti,ab,kw
#40 MeSH descriptor: [Reproducibility of Results] explode all trees
#41 #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40
#42 #24 and #41

Appendix 2 QUADAS-2

Item	Response
PARTICIPANT SELECTION - A. RISK OF BIAS	
Was a consecutive or random sample of patients enrolled?	**Yes** - RCTs and cohort studies (prospective or retrospective) with unenriched (consecutive or random) sampling **Unclear** - If not stated **No** - other studies
Did the study avoid inappropriate exclusions?	**Yes** – If inappropriate exclusions were avoided **Unclear** – if not clearly reported **No** - Exclusion of more than 10% of the samples for any reason, for example retrospective studies with missing data **No** - Systematic exclusion of types of women / images (e.g. of dense breasts) **No** - Exclusion based on outcomes (e.g. exclusion of cancer types, exclusion of interval cancers, exclusion/inclusion based on recall decision)
Were the women and mammograms included in the study independent of those used to train the AI algorithm?	For test set studies, this translates as has the test set been clearly described as an external (geographically) validation set? **No** - Any internal validation (e.g. split sample, cross-validation) or temporal validation **Unclear** - No details stated about the training set and tuning set **Yes** - External geographical validation (Test set was sample from a different centre; can be in another country or the same country) For prospective applied studies in a clinical context: **Yes** - If the study is located at different centre(s) to those who provided mammograms used to train and tune the AI algorithm **Unclear** - If not stated **No** - If there is any overlap
PARTICIPANT SELECTION - B. CONCERNS REGARDING APPLICABILITY	
Is there concern that the included patients do not match the review question?	**High** - If 'yes' for any of the following statements **Unclear** - If no details are provided **Low** - If 'no' for all the following statements • Not a consecutive or random sample of women attending screening; • Enriched sample / cancer prevalence doesn't match screening context (>3%); • Mammograms not from full-field digital mammography; • Mammograms not from screening (e.g. diagnostic or symptomatic) or only subset such as recalled cases or false-negatives included

	(cancer might be easier or more difficult to detect); • Women/women's mammograms not representative of UK population (ethnicity, age)
INDEX TESTS – A. RISK OF BIAS	
Were the index test results interpreted without knowledge of the results of the reference standard?	For index tests where a human is involved (either human read comparator, AI as reader aid, or included otherwise on the AI testing pathway, e.g. arbitration): **Yes** - Require clear statement of blinding, or clear temporal relationships where the human read occurred before the reference standard **No** - Otherwise For index test where AI is used without any human element: **Yes** - AI system has not previously been trained on these mammograms or learned from these mammograms or other mammograms from the same women **No** - If any repeat use of the same cases then (unless explicit that the AI algorithm was pre-set and did not change upon repeat use, and the study did not select one of several AI systems based on use with the same cases) **Unclear** - If not explicit that there has been no repeat within same or previous studies
Were the index test results interpreted without knowledge of the results of any other index tests?	**No** - If human readers were not blinded to AI (unless that AI is specifically part of the same index test) **No** - If AI systems are trained or calibrated using decisions from human readers in same cases **Yes** - Otherwise
If a threshold was used, was it pre-specified?	**Yes** - If using a commercially available AI system which gives a yes/no result, or threshold clearly pre-specified in methods **Yes** - For systems giving a risk score and study explicitly states the pre-specified threshold **No** - Using sensitivity / specificity of the reader as benchmark using the same dataset **No** - Setting the threshold with the validation set without temporal evidence (e.g. published protocol) that threshold was truly pre-specified **NA** - Human readers or human/AI combinations
Where human readers are part of the test, were their decisions made in a clinical practice context? (i.e. avoidance of the laboratory effect)	**Yes** - If the readers made decisions in the clinical context, and those decisions were used to decide whether to recall women (either prospectively as part of a trial or test accuracy study or retrospective studies using the original decision) **No** - If readers examined a test set (of any prevalence) outside clinical practice, or any other context likely to result in the laboratory effect[1]

INDEX TESTS - B. CONCERNS REGARDING APPLICABILITY	
Is there concern that the index test(s) or comparator, its conduct, or interpretation differ from the review question?	**High** - If 'yes' for any of the following **Unclear** – If no details are provided **Low** - If 'no' for all of the following • AI system not yet commercially available, e.g. in house systems; • Study did not use a pre-specified threshold for AI system; • Not a complete testing pathway applicable to clinical practice (for example AI accuracy for single read, but not integrated into screening centre decisions, e.g. arbitration); • Human comparator not a complete testing pathway applicable to clinical practice (human double reading with arbitration at clinical threshold); • AI system / reader had no access to prior mammograms / not 4 views available
REFERENCE STANDARD – A. RISK OF BIAS	
Is the reference standard likely to correctly classify the target condition?	**Yes** - If the reference standard is histopathology results from biopsy (cancer present or absent) with at least 2 years follow up to interval cancers **No** - If the reference standard is histopathology results from biopsy (cancer present or absent) with no follow up
Were the reference standard results interpreted without knowledge of the results of the index test?	**Yes** - Retrospective studies where readers read mammograms prospectively (enriched test sets) **No** - For retrospective studies (if we include the human reader comparator as an index test) **No** - For prospective studies if the investigators did not blind the clinicians undertaking the follow up tests to which index test examined the mammograms, for example by putting location marks in the same format for AI and human readers
REFERENCE STANDARD - B. CONCERNS REGARDING APPLICABILITY	
Is there concern that the target condition as defined by the reference standard does not match the review question?	**High** - If 'yes' for any of the following **Unclear** - If no details are provided **Low** - If 'no' for all of the following • Length of screening rounds <2 years for follow-up / definition of interval cancers; • Classification not by biopsy/follow-up
FLOW AND TIMING – A. RISK OF BIAS	
Did all patients receive a reference standard?	**No** - If there was significant (>10%) loss to follow up for reference standards of interval cancers or subsequent screening results **No** - If any women who should have received a biopsy or follow-up tests after index test positive results did not receive one or results were unavailable **Yes** - otherwise

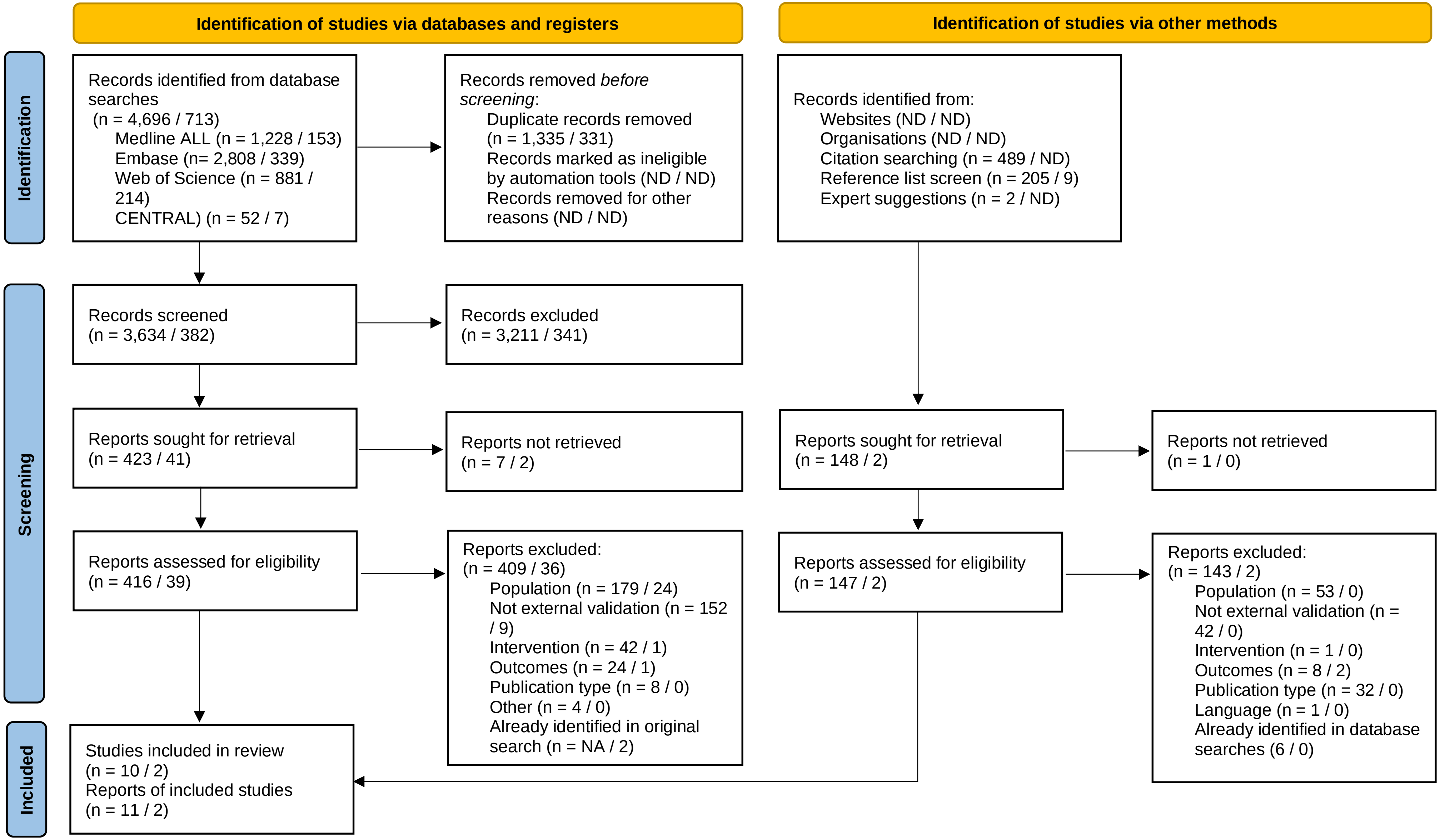
Did the study avoid choosing which reference standard based on results of just one of the index tests? (All studies will necessarily have differential verification, because not all women can or should be biopsied. Here we are measuring whether deciding which reference standard is received based on results of just one of the index tests is avoided.)	**Yes** - For test-treat RCTs randomizing to different test strategies and their associated recall decisions **Yes** - If women testing positive in any of the included index tests (AI pathways or comparator human pathways) all receive follow up tests/biopsy in a prospective study **No** - If women were recalled for further tests on the basis of one of the index tests, and not other(s) then this will cause bias because cancer, when present, is more likely to be found if the person receives follow-up tests after recall from screening **No** - In retrospective studies, the decision whether to recall for follow-up tests/biopsy was made on the basis of the human readers' decision. We do not know whether AI positive, human reader negative women are false positive or true positive, and what type of true positive. Follow-up to development of interval cancers will detect some, but not all of these cancers, so reduces, but does not eliminate this bias **No** - For prospective studies where decision to recall is informed by one index test but not all, or is more influenced by one index test than others **Unclear** - Retrospective reader studies (enriched test set studies) in which readers prospectively read retrospective data, the reference standard is not based on any index test but the reference standard is based on the original human reader decision. The reviewers are unclear about the risk of bias.
Were all patients included in the analysis?	**Yes** - If there were any exclusions after the point of selecting the cohort, for example intermediate or indeterminate results **No** - Otherwise

References for Appendix 4

1. Gur D, Bandos AI, Cohen CS, et al. The "laboratory" effect: comparing radiologists' performance and variability during prospective clinical and laboratory mammography interpretations. *Radiology* 2008;249(1):47-53. doi: 10.1148/radiol.2491072025 [published Online First: 2008/08/07]

Supplementary Figure 1. PRISMA diagram. Summary of publications included and excluded at each stage of the review (original searches / update searches)

FFDM, Full field digital mammogram; NA, Not applicable; ND, Not done



Appendix 3 Publications and sub-studies excluded after review of full-text articles

Key to reasons for exclusions and justifications

Population – Image type: Studies evaluating AI on image types other than full field digital mammography (FFDM) (mainly digitised film images). Screening mammography is typically undertaken using FFDM in women attending breast screening. Other imaging techniques not classed as mammography or not digital are not relevant. Results from imaging other than FFDM may not be applicable to FFDM in screening programmes.

Population – Mammography type not reported: Studies evaluating AI on image types which were not clearly identified as FFDM. Screening mammography is typically undertaken using full field digital mammography (FFDM) in women attending breast screening. Other imaging techniques not classed as mammography or not digital are not relevant. Results from imaging other than FFDM may not be applicable to FFDM in screening programmes.

Population – Incomplete images: Studies evaluating AI using regions of interest (ROIs) of images. Images of part of a mammogram do not represent the use case in the screening context, which requires recall or not decisions to be made on women's (craniocaudal and mediolateral oblique) screening mammograms for both breasts.

Population – Subpopulation: Studies only including images with cancer. These are not sufficient to estimate test accuracy of AI for screening mammograms, as it excludes specificity, and the trade-off between sensitivity and specificity. Studies on images of subpopulations by screening risk or screening outcome. They do not represent the screening population and no inference on the performance of the AI system in a screening population can be drawn (however, subpopulations by ethnicity or socioeconomic status are included as the impact of any change on equity is important). This does not apply to populations which represent a group of women at any stage within the screening pathway (e.g. recalled women without selection on final diagnosis) on the assumption that AI could be incorporated for this subgroup only.

Population – <90% screening mammograms or unclear proportion: Studies on images of diagnostic mammograms, with >10% diagnostic mammograms or an unclear proportion of diagnostic mammograms. These do not represent the use case in the screening context.

Internal validation – Cross validation, Leave-one out, Split sample: Studies using internal validation whereby the validation dataset used to assess a model uses data which were used to develop that model. Cross validation and leave-one out are resampling techniques which use the original data to assist with preliminary assessment and fine-tuning of the model during validation. The issue of using data on which an algorithm was trained with is that models can be prone to overfitting; whereby the model fits the trained data extremely well, but to the detriment of the model's ability to perform when presented with new data, which is known as poor generalization. The split-sample approach is generally an inefficient form of internal validation because it does not accurately reflect a model's generalisability.

Intervention – Detecting subtypes: Studies using AI to detect cancer subtypes such as microcalcifications or architectural distortions only. The detection of cancer subtypes does not present the complete picture of cancer detection (e.g. microcalcifications are associated mainly with DCIS and not with cancer; detecting microcalcifications only will miss some types of cancers). On their own, these AI systems do not provide the information on cancer present/ not present to inform a decision whether to recall or not recall. Systems reporting single features could be combined to provide a more complete picture, however, studies would need to report an overall outcome of test accuracy of the combination of systems.

Intervention – No detection/classification: Studies using AI for lesion enhancement or segmentation of pectoral muscle regions. These studies do not report on the classification of images into recall or no recall and are uninformative in terms of test performance of AI.

Intervention – Not AI: Studies using traditional computer aided detection without machine learning features for the detection of lesions in images. In AI the layers of features are not designed by human expertise; instead they are learnt from the underlying data. Therefore, studies reporting the test accuracy and effectiveness of traditional CAD systems were considered significantly different from studies assessing machine learning AI systems.

Intervention – Prediction of cancer: Studies using AI for the prediction of future cancer risk including the detection of breast density and parenchymal patterns as risk factors. These studies did not consider the detection of cancer present on screening mammograms.

No relevant outcomes: Studies reporting accuracy without outcomes characterising the trade-off between false positive and false negative results including global measures such as the area under the curve (AUC). The trade-off between false positive and false negative results is critical to test accuracy.

Study type – Systematic reviews with no relevant outcomes: Studies reporting systematic reviews that were off topic and did not provide additional references for the review.

Document Supply cancelled request: no location found: Studies unavailable following internet searches, contacting authors and pursuing interlibrary requests.

Publications excluded with reason – Original database searches

Reference		Main reason for exclusion
Population – Image type (e.g. digitised film images; not FFDM images) (n=150)		
1.	Abbas Q, Fondo'n I, Celebi E. A Computerized System for Detection of Spiculated Margins based on Mammography. International Arab Journal of Information Technology. 2015;12(6):582-8.	Population – Image type
2.	Agnes SA, Anitha J, Pandian SIA, Peter JD. Classification of Mammogram Images Using Multiscale all Convolutional Neural Network (MA-CNN). J Med Syst. 2019;44(1):30.	Population – Image type
3.	Anitha J, Dinesh Peter J, Immanuel Alex Pandian S. A dual stage adaptive thresholding (DuSAT) for automatic mass detection in mammograms. Computer Methods and Programs in Biomedicine. 2017;138:93-104.	Population – Image type
4.	Bartolotta TV, Orlando A, Cantisani V, Matranga D, Ienzi R, Cirino A, et al. Focal breast lesion characterization according to the BI-RADS US lexicon: role of a computer-aided decision-making support. La Radiologia medica. 2018;123(7):498-506.	Population – Image type
5.	Beheshti SM, AhmadiNoubari H, Fatemizadeh E, Khalili M. An efficient fractal method for detection and diagnosis of breast masses in mammograms. J Digit Imaging. 2014;27(5):661-9.	Population – Image type
6.	Chakraborty J, Midya A, Mukhopadhyay S, Rangayyan RM, Sadhu A, Singla V, et al. Computer-Aided Detection of Mammographic Masses Using Hybrid Region Growing Controlled by Multilevel Thresholding. Journal of Medical and Biological Engineering. 2019;39(3):352-66.	Population – Image type
7.	Chithra Devi M, Audithan S. Analysis of different types of entropy measures for breast cancer diagnosis using ensemble classification. Biomedical Research (India). 2017;28(7):3182-6.	Population – Image type
8.	Choi JY, Ro YM. Multiresolution local binary pattern texture analysis combined with variable selection for application to false-positive reduction in computer-aided detection of breast masses on mammograms. Phys Med Biol. 2012;57(21):7029-52.	Population – Image type
9.	Chougrad H, Zouaki H, Alheyane O. Deep Convolutional Neural Networks for breast cancer screening. Comput Methods Programs Biomed. 2018;157:19-30.	Population – Image type

Reference		Main reason for exclusion
10.	Chowdhary CL, Mittal M, P K, Pattanaik PA, Marszalek Z. An Efficient Segmentation and Classification System in Medical Images Using Intuitionist Possibilistic Fuzzy C-Mean Clustering and Fuzzy SVM Algorithm. Sensors (Basel). 2020;20(14):13.	Population – Image type
11.	Cunningham CA, Drew T, Wolfe JM. Analog Computer-Aided Detection (CAD) information can be more effective than binary marks. Atten Percept Psychophys. 2017;79(2):679-90.	Population – Image type
12.	de Oliveira Silva LC, Barros AK, Lopes MV. Detecting masses in dense breast using independent component analysis. Artif Intell Med. 2017;80:29-38.	Population – Image type
13.	Dheeba J, Albert Singh N, Tamil Selvi S. Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach. J Biomed Inform. 2014;49:45-52.	Population – Image type
14.	Dheeba J, Jaya T, Singh NA. Breast cancer risk assessment and diagnosis model using fuzzy support vector machine based expert system. Journal of Experimental & Theoretical Artificial Intelligence. 2017;29(5):1011-21.	Population – Image type
15.	Dheeba J, Tamil Selvi S. An improved decision support system for detection of lesions in mammograms using Differential Evolution Optimized Wavelet Neural Network. J Med Syst. 2012;36(5):3223-32.	Population – Image type
16.	Dubey AK, Gupta U, Jain S. Analysis of k-means clustering approach on the breast cancer Wisconsin dataset. Int. 2016;11(11):2033-47.	Population – Image type
17.	El Fahssi K, Elmoufidi A, Abenaou A, Jai-Andaloussi S, Sekkaki A. Novel approach to classification of Abnormalities in the mammogram image. International Journal of Biology and Biomedical Engineering. 2016;10:72-9.	Population – Image type
18.	El-Shazli AMA, Youssef SM, Elshennawy M. Computer-aided model for breast cancer detection in mammograms. International Journal of Pharmacy and Pharmaceutical Sciences. 2016;8(Supplement 2):31-4.	Population – Image type
19.	Elmoufidi A, El Fahssi K, Jai-andaloussi S, Sekkaki A, Gwenole Q, Lamard M. Anomaly classification in digital mammography based on multiple-instance learning. Iet Image Processing. 2018;12(3):320-8.	Population – Image type
20.	Eltoukhy MM, Faye I, Samir BB. Breast cancer diagnosis in digital mammogram using multiscale curvelet transform. Comput Med Imaging Graph. 2010;34(4):269-76.	Population – Image type

Reference		Main reason for exclusion
21.	Ericeira DR, Silva AC, de Paiva AC, Gattass M. Detection of masses based on asymmetric regions of digital bilateral mammograms using spatial description with variogram and cross-variogram functions. Comput Biol Med. 2013;43(8):987-99.	Population – Image type
22.	Ganesan K, Acharya RU, Chua CK, Min LC, Mathew B, Thomas AK. Decision support system for breast cancer detection using mammograms. Proc Inst Mech Eng [H]. 2013;227(7):721-32.	Population – Image type
23.	Garma FB, Hassan MA. Classification of breast tissue as normal or abnormal based on texture analysis of digital mammogram. Journal of Medical Imaging and Health Informatics. 2014;4(5):647-53.	Population – Image type
24.	Gedik N. Breast cancer diagnosis system via contourlet transform with sharp frequency localization and least squares support vector machines. Journal of Medical Imaging and Health Informatics. 2015;5(3):497-505.	Population – Image type
25.	Gedik N, Atasoy A. Performance evaluation of the wave atom algorithm to classify mammographic images. Turkish Journal of Electrical Engineering and Computer Sciences. 2014;22(4):957-69.	Population – Image type
26.	Gedik N, Atasoy A, Sevim Y. Investigation of wave atom transform by using the classification of mammograms. Applied Soft Computing. 2016;43:546-52.	Population – Image type
27.	Gorgel P, Sertbas A, Ucan ON. Mammographical mass detection and classification using Local Seed Region Growing-Spherical Wavelet Transform (LSRG-SWT) hybrid scheme. Computers in Biology and Medicine. 2013;43(6):765-74.	Population – Image type
28.	Guan JS, Lin LY, Ji GL, Lin CM, Le TL, Rudas IJ. Breast Tumor Computer-aided Diagnosis using Self-Validating Cerebellar Model Neural Networks. Acta Polytechnica Hungarica. 2016;13(4):39-52.	Population – Image type
29.	Hu K, Gao XP, Li F. Detection of Suspicious Lesions by Adaptive Thresholding Based on Multiresolution Analysis in Mammograms. Ieee Transactions on Instrumentation and Measurement. 2011;60(2):462-72.	Population – Image type
30.	James JJ, Gilbert FJ, Wallis MG, Gillan MG, Astley SM, Boggis CR, et al. Mammographic features of breast cancers at single reading with computer-aided detection and at double reading in a large multicenter prospective trial of computer-aided detection: CADET II. Radiology. 2010;256(2):379-86.	Population – Image type

Reference		Main reason for exclusion
31.	Jebamony J, Jacob D. Classification of Benign and Malignant Breast Masses on Mammograms for Large Datasets using Core Vector Machines. Curr Med Imaging. 2020;16(6):703-10.	Population – Image type
32.	Kadam VJ, Jadhav SM, Vijayakumar K. Breast Cancer Diagnosis Using Feature Ensemble Learning Based on Stacked Sparse Autoencoders and Softmax Regression. J Med Syst. 2019;43(8):263.	Population – Image type
33.	Kanadam KP, Chereddy SR. Mammogram classification using sparse-ROI: A novel representation to arbitrary shaped masses. Expert Systems with Applications. 2016;57:204-13.	Population – Image type
34.	Kanchana M, Varalakshmi P. Computer aided system for breast cancer in digitized mammogram using shearlet band features with LS-SVM classifier. International Journal of Wavelets Multiresolution and Information Processing. 2016;14(3).	Population – Image type
35.	Kanchanamani M, Perumal V. Performance evaluation and comparative analysis of various machine learning techniques for diagnosis of breast cancer. Biomedical Research (India). 2016;27(3):623-31.	Population – Image type
36.	Kashyap KL, Bajpai MK, Khanna P. Globally supported radial basis function based collocation method for evolution of level set in mass segmentation using mammograms. Comput Biol Med. 2017;87:22-37.	Population – Image type
37.	Kashyap KL, Bajpai MK, Khanna P. An efficient algorithm for mass detection and shape analysis of different masses present in digital mammograms. Multimedia Tools and Applications. 2018;77(8):9249-69.	Population – Image type
38.	Kashyap KL, Bajpai MK, Khanna P, Giakos G. Mesh-free based variational level set evolution for breast region segmentation and abnormality detection using mammograms. Int j numer method biomed eng. 2018;34(1):01.	Population – Image type
39.	Kayode AA, Akande NO, Adegun AA, Adebiyi MO. An automated mammogram classification system using modified support vector machine. Medical Devices: Evidence and Research. 2019;12:275-84.	Population – Image type
40.	Kelder A, Zigel Y, Lederman D, Zheng B. A new computer-aided detection scheme based on assessment of local bilateral mammographic feature asymmetry - a preliminary evaluation. Conf Proc IEEE Eng Med Biol Soc. 2015;2015:6394-7.	Population – Image type

Reference		Main reason for exclusion
41.	Khan AA, Arora AS. Computer aided diagnosis of breast cancer based on level set segmentation of masses and classification using ensemble classifiers. Biomedical Research (India). 2018;29(19):3610-5.	Population – Image type
42.	Khan HN, Shahid AR, Raza B, Dar AH, Alquhayz H. Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network. Ieee Access. 2019;7.	Population – Image type
43.	Krishnan MMR, Banerjee S, Chakraborty C, Chakraborty C, Ray AK. Statistical analysis of mammographic features and its classification using support vector machine. Expert Systems with Applications. 2010;37(1):470-8.	Population – Image type
44.	Li JB, Wang YH, Tang LL. Mammogram-based discriminant fusion analysis for breast cancer diagnosis. Clin Imaging. 2012;36(6):710-6.	Population – Image type
45.	Li P, Bi T, Huang J, Li S. Breast cancer early diagnosis based on hybrid strategy. Biomed Mater Eng. 2014;24(6):3397-404.	Population – Image type
46.	Li Z, Sun J, Zhang J, Hu D, Wang Q, Peng K. Quantification of acoustic radiation force impulse in differentiating between malignant and benign breast lesions. Ultrasound Med Biol. 2014;40(2):287-92.	Population – Image type
47.	Lo CM, Moon WK, Huang CS, Chen JH, Yang MC, Chang RF. INTENSITY-INVARIANT TEXTURE ANALYSIS FOR CLASSIFICATION OF BI-RADS CATEGORY 3 BREAST MASSES. Ultrasound in Medicine and Biology. 2015;41(7):2039-48.	Population – Image type
48.	Onan A. A stochastic gradient descent based SVM with fuzzy-rough feature selection and instance selection for breast cancer diagnosis. Journal of Medical Imaging and Health Informatics. 2015;5(6):1233-9.	Population – Image type
49.	Singh WJ, Nagarajan B. Automatic diagnosis of mammographic abnormalities based on hybrid features with learning classifier. Comput Methods Biomech Biomed Engin. 2013;16(7):758-67.	Population – Image type
50.	Valarmathi P, Robinson S. An Improved Neural Network for Mammogram Classification Using Genetic Optimization. Journal of Medical Imaging and Health Informatics. 2016;6(7):1631-5.	Population – Image type
51.	Liu X, Zeng Z. A new automatic mass detection method for breast cancer with false positive reduction. Neurocomputing. 2015;152:388-402.	Population – Image type

Reference		Main reason for exclusion
52.	Liu XM, Tang JS. Mass Classification in Mammograms Using Selected Geometry and Texture Features, and a New SVM-Based Feature Selection Method. Ieee Systems Journal. 2014;8(3):910-20.	Population – Image type
53.	Liu XM, Zhai LL, Zhu T, Liu J, Zhang K, Hu W. Multiple TBSVM-RFE for the detection of architectural distortion in mammographic images. Multimedia Tools and Applications. 2018;77(12):15773-802.	Population – Image type
54.	Mahersia H, Boulehmi H, Hamrouni K. Development of intelligent systems based on Bayesian regularization network and neuro-fuzzy models for mass detection in mammograms: A comparative analysis. Comput Methods Programs Biomed. 2016;126:46-62.	Population – Image type
55.	Mencattini A, Salmeri M. Breast masses detection using phase portrait analysis and fuzzy inference systems. Int. 2012;7(4):573-83.	Population – Image type
56.	Meselhy Eltoukhy M, Faye I, Belhaouari Samir B. A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation. Comput Biol Med. 2012;42(1):123-8.	Population – Image type
57.	Midya A, Rabidas R, Sadhu A, Chakraborty J. Edge Weighted Local Texture Features for the Categorization of Mammographic Masses. Journal of Medical and Biological Engineering. 2018;38(3):457-68.	Population – Image type
58.	Milosevic M, Jovanovic Z, Jankovic D. A comparison of methods for three-class mammograms classification. Technol Health Care. 2017;25(4):657-70.	Population – Image type
59.	Mohammadi-Sardo S, Labibi F, Shafiei SA. A new approach for detecting abnormalities in mammograms using a computer-aided windowing system based on Otsu's method. Radiol Phys Technol. 2019;12(2):178-84.	Population – Image type
60.	Mohammed SHA, Yousuf SEK. A computer-aided diagnosis system for the detection and classification of breast cancer. Journal of Clinical Engineering. 2016;41(2):97-100.	Population – Image type
61.	Mohammed SHAA, Mustafa ZA. Breast Tumors Classification Using Adaptive Neuro-Fuzzy Inference System. Journal of Clinical Engineering. 2017;42(2):68-72.	Population – Image type
62.	Mohanty F, Rup S, Dash B. Automated diagnosis of breast cancer using parameter optimized kernel extreme learning machine. Biomedical Signal Processing and Control. 2020;62 (no pagination).	Population – Image type

Reference		Main reason for exclusion
63.	Mohanty F, Rup S, Dash B, Majhi B, Swamy MNS. Mammogram classification using contourlet features with forest optimization-based feature selection approach. Multimedia Tools and Applications. 2019;78(10):12805-34.	Population – Image type
64.	Mohanty F, Rup S, Dash B, Majhi B, Swamy MNS. A computer-aided diagnosis system using Tchebichef features and improved grey wolf optimized extreme learning machine. Applied Intelligence. 2019;49(3):983-1001.	Population – Image type
65.	Mohanty F, Rup S, Dash B, Majhi B, Swamy MNS. An improved scheme for digital mammogram classification using weighted chaotic salp swarm algorithm-based kernel extreme learning machine. Applied Soft Computing. 2020;91.	Population – Image type
66.	Mohanty F, Rup S, Dash B, Majhi B, Swamy MNS. Digital mammogram classification using 2D-BDWT and GLCM features with FOA-based feature selection approach. Neural Computing & Applications. 2020;32(11):7029-43.	Population – Image type
67.	Moslemi H, Kazerouni IA, Hourali F. Breast cancer diagnosis in mammogram images using coordinate logic filters. Biomedical Research (India). 2017;28(22):10108-11.	Population – Image type
68.	Muduli D, Dash R, Majhi B. Automated breast cancer detection in digital mammograms: A moth flame optimization based ELM approach. Biomedical Signal Processing and Control. 2020;59 (no pagination).	Population – Image type
69.	Mughal B, Muhammad N, Sharif M. Adaptive hysteresis thresholding segmentation technique for localizing the breast masses in the curve stitching domain. Int J Med Inf. 2019;126:26-34.	Population – Image type
70.	Nagarajan V, Britto EC, Veeraputhiran SM. Feature extraction based on empirical mode decomposition for automatic mass classification of mammogram images. Medicine in Novel Technology and Devices. 2019;1 (no pagination).	Population – Image type
71.	Nagthane DK, Rajurkar AM. An improved diagnosis technique for breast cancer using LCFS and TreeHiCARe classifier model. Sensor Review. 2019;39(1):107-20.	Population – Image type
72.	Nanni L, Brahnam S, Lumini A. A very high performing system to discriminate tissues in mammograms as benign and malignant. Expert Systems with Applications. 2012;39(2):1968-71.	Population – Image type
73.	Narvaez F, Alvarez J, Garcia-Arteaga JD, Tarquino J, Romero E. Characterizing Architectural Distortion in Mammograms by Linear Saliency. J Med Syst. 2017;41(2):26.	Population – Image type

Reference		Main reason for exclusion
74.	Naseem MT, Sulong GZB, Jaffar MA. MRT letter: Quantum noise removal and classification of breast mammogram images. Microscopy Research and Technique. 2012;75(12):1609-12.	Population – Image type
75.	Naveed N, Jaffar MA, Choi TS. MRT Letter: Segmentation and Texture-Based Classification of Breast Mammogram Images. Microscopy Research and Technique. 2011;74(11):985-7.	Population – Image type
76.	Neto OPS, Silva AC, Paiva AC, Gattass M. Automatic mass detection in mammography images using particle swarm optimization and functional diversity indexes. Multimedia Tools and Applications. 2017;76(18):19263-89.	Population – Image type
77.	Nishikawa RM, Schmidt RA, Linver MN, Edwards AV, Papaioannou J, Stull MA. Clinically missed cancer: how effectively can radiologists use computer-aided detection? AJR Am J Roentgenol. 2012;198(3):708-16.	Population – Image type
78.	Nugroho HA, Fajrin HR, Soesanti I, Budiani RL. Analysis of texture for classification of breast cancer on mammogram images. International Journal of Medical Engineering and Informatics. 2018;10(4):382-91.	Population – Image type
79.	P S, R T. Aiding the Digital Mammogram for Detecting the Breast Cancer Using Shearlet Transform and Neural Network. Asian Pac J Cancer Prev. 2018;19(9):2665-71.	Population – Image type
80.	Pak F, Kanan HR, Alikhassi A. Breast cancer detection and classification in digital mammography based on Non-Subsampled Contourlet Transform (NSCT) and Super Resolution. Comput Methods Programs Biomed. 2015;122(2):89-107.	Population – Image type
81.	Panetta K, Zhou Y, Agaian S, Jia H. Nonlinear unsharp masking for mammogram enhancement. IEEE Trans Inf Technol Biomed. 2011;15(6):918-28.	Population – Image type
82.	Paquerault S, Hardy PT, Wersto N, Chen J, Smith RC. Investigation of optimal use of computer-aided detection systems: the role of the "machine" in decision making process. Acad Radiol. 2010;17(9):1112-21.	Population – Image type
83.	Paramkusham S, Rao KMM, Prabhakar Rao BVVSN, Sharma S. Application of TAR signature for breast mass analysis. Biomedical Research (India). 2018;29(10):2030-4.	Population – Image type
84.	Parmeggiani D, Avenia N, Sanguinetti A, Ruggiero R, Docimo G, Siciliano M, et al. Artificial intelligence against breast cancer (A.N.N.E.S-B.C.-Project). Ann Ital Chir. 2012;83(1):1-5.	Population – Image type

Reference		Main reason for exclusion
85.	Patel BC, Sinha GR, Soni D. Detection of masses in mammographic breast cancer images using modified histogram based adaptive thresholding (MHAT) method. International Journal of Biomedical Engineering and Technology. 2019;29(2):134-54.	Population – Image type
86.	Pawar MM, Talbar SN, Dudhane A. Local Binary Patterns Descriptor Based on Sparse Curvelet Coefficients for False-Positive Reduction in Mammograms. J. 2018;2018:5940436.	Population – Image type
87.	Perre AC, Alexandre LA, Freire LC. Lesion classification in mammograms using convolutional neural networks and transfer learning. Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization. 2019;7(5-6):550-6.	Population – Image type
88.	Pezeshki H, Rastgarpour M, Sharifi A, Yazdani S. Extraction of spiculated parts of mammogram tumors to improve accuracy of classification. Multimedia Tools and Applications. 2019;78(14):19979-20003.	Population – Image type
89.	Povyakalo AA, Alberdi E, Strigini L, Ayton P. How to discriminate between computer-aided and computer-hindered decisions: a case study in mammography. Med Decis Making. 2013;33(1):98-107.	Population – Image type
90.	Qasim KR, Ouda AJ. An accurate breast cancer detection system based on deep learning cnn. Medico Legal Update. 2020;20(1).	Population – Image type
91.	Quellec G, Lamard M, Cozic M, Coatrieux G, Cazuguel G. Multiple-Instance Learning for Anomaly Detection in Digital Mammography. IEEE Trans Med Imaging. 2016;35(7):1604-14.	Population – Image type
92.	Rabidas R, Arif W. Characterization of mammographic masses based on local photometric attributes. Multimedia Tools and Applications. 2020;79(29-30):21967-85.	Population – Image type
93.	Rabidas R, Chakraborty J, Midya A. Analysis of 2D singularities for mammographic mass classification. Iet Computer Vision. 2017;11(1):22-32.	Population – Image type
94.	Rabidas R, Midya A, Chakraborty J. Neighborhood Structural Similarity Mapping for the Classification of Masses in Mammograms. IEEE j. 2018;22(3):826-34.	Population – Image type
95.	Rabottino G, Mencattini A, Salmeri M, Caselli F, Lojacono R. Performance evaluation of a region growing procedure for mammographic breast lesion identification. Computer Standards & Interfaces. 2011;33(2):128-35.	Population – Image type
96.	Ragab DA, Sharkas M, Marshall S, Ren J. Breast cancer detection using deep convolutional neural networks and support vector machines. Peerj. 2019;7:e6201.	Population – Image type

Reference		Main reason for exclusion
97.	Raghavendra U, Acharya UR, Fujita H, Gudigar A, Tan JH, Chokkadi S. Application of Gabor wavelet and Locality Sensitive Discriminant Analysis for automated identification of breast cancer using digitized mammogram images. Applied Soft Computing. 2016;46:151-61.	Population – Image type
98.	Rangayyan RM, Banik S, Chakraborty J, Mukhopadhyay S, Desautels JE. Measures of divergence of oriented patterns for the detection of architectural distortion in prior mammograms. Int. 2013;8(4):527-45.	Population – Image type
99.	Rangayyan RM, Banik S, Desautels JE. Computer-aided detection of architectural distortion in prior mammograms of interval cancer. J Digit Imaging. 2010;23(5):611-31.	Population – Image type
100.	Rangayyan RM, Banik S, Desautels JE. Detection of architectural distortion in prior mammograms via analysis of oriented patterns. J. 2013;78:30.	Population – Image type
101.	Rangayyan RM, Oloumi F. Fractal analysis and classification of breast masses using the power spectra of signatures of contours. Journal of Electronic Imaging. 2012;21(2).	Population – Image type
102.	Rangayyan RM, Oloumi F, Nguyen TM. Fractal analysis of contours of breast masses in mammograms via the power spectra of their signatures. Conf Proc IEEE Eng Med Biol Soc. 2010;2010:6737-40.	Population – Image type
103.	Rauch T, Rieger J, Pelzer G, Horn F, Erber R, Wunderle M, et al. Discrimination analysis of breast calcifications using x-ray dark-field radiography. Med Phys. 2020;47(4):1813-26.	Population – Image type
104.	Reyad YA, Berbar MA, Hussain M. Comparison of statistical, LBP, and multi-resolution analysis features for breast mass classification. J Med Syst. 2014;38(9):100.	Population – Image type
105.	Roberts T, Newell M, Auffermann W, Vidakovic B. Wavelet-based scaling indices for breast cancer diagnostics. Stat Med. 2017;36(12):1989-2000.	Population – Image type
106.	Roseline R, Manikandan S. Determination of Breast Cancer Using KNN Cluster Technique. Indian Journal of Public Health Research and Development. 2018;9(2):418-23.	Population – Image type
107.	Rouhi R, Jafari M. Classification of benign and malignant breast tumors based on hybrid level set segmentation. Expert Systems with Applications. 2016;46:45-59.	Population – Image type
108.	Rouhi R, Jafari M, Kasaei S, Keshavarzian P. Benign and malignant breast tumors classification based on region growing and CNN segmentation. Expert Systems with Applications. 2015;42(3):990-1002.	Population – Image type

Reference		Main reason for exclusion
109.	S RS, Rajaguru H. Comparison Analysis of Linear Discriminant Analysis and Cuckoo-Search Algorithm in the Classification of Breast Cancer from Digital Mammograms. Asian Pac J Cancer Prev. 2019;20(8):2333-7.	Population – Image type
110.	Safdar Gardezi SJ, Faye I. Mammogram classification based on morphological component analysis (MCA) and Curvelet decomposition. Neuroscience and Biomedical Engineering. 2015;3(1):27-33.	Population – Image type
111.	5Saki F, Tahmasbi A, Soltanian-Zadeh H, Shokouhi SB. Fast opposite weight learning rules with application in breast cancer diagnosis. Comput Biol Med. 2013;43(1):32-41.	Population – Image type
112.	Salazar-Licea LA, Pedraza-Ortega JC, Pastrana-Palma A, Aceves-Fernandez MA. Location of mammograms ROI's and reduction of false-positive. Comput Methods Programs Biomed. 2017;143:97-111.	Population – Image type
113.	Sampaio WB, Diniz EM, Silva AC, de Paiva AC, Gattass M. Detection of masses in mammogram images using CNN, geostatistic functions and SVM. Comput Biol Med. 2011;41(8):653-64.	Population – Image type
114.	Samulski M, Karssemeijer N. Optimizing Case-based detection performance in a multiview CAD system for mammography. IEEE Trans Med Imaging. 2011;30(4):1001-9.	Population – Image type
115.	Saranyaraj D, Manikandan M, Maheswari S. A deep convolutional neural network for the early detection of breast carcinoma with respect to hyper- parameter tuning. Multimedia Tools and Applications. 2020;79(15-16):11013-38.	Population – Image type
116.	Saraswathi D, Srinivasan E. A CAD system to analyse mammogram images using fully complex-valued relaxation neural network ensembled classifier. J Med Eng Technol. 2014;38(7):359-66.	Population – Image type
117.	Saybani MR, Teh YW, Aghabozorgi SR, Shamshirband S, Kiah MLM, Balas VE. Diagnosing breast cancer with an improved artificial immune recognition system. Soft Computing. 2016;20(10):4069-84.	Population – Image type
118.	Schonmeyer R, Athelogou M, Sittek H, Ellenberg P, Feehan O, Schmidt G, et al. Cognition Network Technology prototype of a CAD system for mammography to assist radiologists by finding similar cases in a reference database. Int. 2011;6(1):127-34.	Population – Image type
119.	Selvamani I, Arasu GT. Computer aided system for detection and classification of breast cancer. Current Medical Imaging Reviews. 2015;11(2):77-84.	Population – Image type

Reference		Main reason for exclusion
120.	Selvi C, Suganthi M. A Novel Enhanced Gray Scale Adaptive Method for Prediction of Breast Cancer. J Med Syst. 2018;42(11).	Population – Image type
121.	Senthilkumar B, Umamaheswari G. Combination of novel enhancement technique and fuzzy C means clustering technique in breast cancer detection. Biomedical Research (India). 2013;24(2):252-6.	Population – Image type
122.	Sha ZJ, Hu L, Rouyendegh BD. Deep learning and optimization algorithms for automatic breast cancer detection. International Journal of Imaging Systems and Technology. 2020;30(2):495-506.	Population – Image type
123.	Shahin OR, Alruily M, Alsmarah M, Alruwaill M. Breast cancer detection using modified hough transform. Biomedical Research (India). 2018;29(16):3188-91.	Population – Image type
124.	Sharma S, Khanna P. Computer-aided diagnosis of malignant mammograms using Zernike moments and SVM. J Digit Imaging. 2015;28(1):77-90.	Population – Image type
125.	Shen LZ, He MF, Shen N, Yousefi N, Wang C, Liu GQ. Optimal breast tumor diagnosis using discrete wavelet transform and deep belief network based on improved sunflower optimization method. Biomedical Signal Processing and Control. 2020;60.	Population – Image type
126.	Shirazinodeh A, Noubari HA, Rabbani H, Dehnavi AM. Detection and classification of breast cancer in wavelet sub-bands of fractal segmented cancerous zones. Journal of Medical Signals and Sensors. 2015;5(3):162-70.	Population – Image type
127.	Shobha Rani N, Rao CS. Exploration and evaluation of efficient pre-processing and segmentation technique for breast cancer diagnosis based on mammograms. International Journal of Research in Pharmaceutical Sciences. 2019;10(3):2071-81.	Population – Image type
128.	Singh B, Kaur M. An approach for classification of malignant and benign microcalcification clusters. Sadhana-Academy Proceedings in Engineering Sciences. 2018;43(3).	Population – Image type
129.	Singh L, Jaffery ZA. Computer-aided diagnosis of breast cancer in digital mammograms. International Journal of Biomedical Engineering and Technology. 2018;27(3):233-46.	Population – Image type
130.	Singh SP, Urooj S. An Improved CAD System for Breast Cancer Diagnosis Based on Generalized Pseudo-Zernike Moment and Ada-DEWNN Classifier. J Med Syst. 2016;40(4):105.	Population – Image type
131.	Singh SP, Urooj S, Lay-Ekuakille A. Breast Cancer Detection Using PCPCET and ADEWNN: A Geometric Invariant Approach to Medical X-Ray Image Sensors. Ieee Sensors Journal. 2016;16(12):4847-55.	Population – Image type

Reference		Main reason for exclusion
132.	Singh VP, Srivastava S, Srivastava R. Effective mammogram classification based on center symmetric-LBP features in wavelet domain using random forests. Technology and Health Care. 2017;25(4):709-27.	Population – Image type
133.	Singh WJ, Nagarajan B. Automatic diagnosis of mammographic abnormalities based on hybrid features with learning classifier. Comput Methods Biomech Biomed Engin. 2013;16(7):758-67.	Population – Image type
134.	Soulami KB, Saidi MN, Honnit B, Anibou C, Tamtaoui A. Detection of breast abnormalities in digital mammograms using the electromagnetism-like algorithm. Multimedia Tools and Applications. 2019;78(10):12835-63.	Population – Image type
135.	Sriramkumar D, Malmathanraj R, Mohan R, Umamaheswari S. Mammogram tumour classification using modified segmentation techniques. International Journal of Biomedical Engineering and Technology. 2013;13(3):218-39.	Population – Image type
136.	Suganthi M, Madheswaran M. An improved medical decision support system to identify the breast cancer using mammogram. J Med Syst. 2012;36(1):79-91.	Population – Image type
137.	Suvetha K, Sultana M. Analysis of breast cancer using tetrolet transform. Indian Journal of Public Health Research and Development. 2017;8(3 Supplement):19-21.	Population – Image type
138.	Tahmasbi A, Saki F, Shokouhi SB. Classification of benign and malignant masses based on Zernike moments. Comput Biol Med. 2011;41(8):726-35.	Population – Image type
139.	Tai SC, Chen ZS, Tsai WT. An automatic mass detection system in mammograms based on complex texture features. IEEE j. 2014;18(2):618-27.	Population – Image type
140.	Tan M, Pu J, Zheng B. Optimization of breast mass classification using sequential forward floating selection (SFFS) and a support vector machine (SVM) model. Int. 2014;9(6):1005-20.	Population – Image type
141.	Thivya KS, Sakthivel P, Sai PMV. Analysis of framelets for breast cancer diagnosis. Technology and Health Care. 2016;24(1):21-9.	Population – Image type
142.	Timp S, Varela C, Karssemeijer N. Computer-aided diagnosis with temporal analysis to improve radiologists' interpretation of mammographic mass lesions. IEEE Trans Inf Technol Biomed. 2010;14(3):803-8.	Population – Image type
143.	Ting FF, Tan YJ, Sim KS. Convolutional neural network improvement for breast cancer classification. Expert Systems with Applications. 2019;120:103-15.	Population – Image type

Reference		Main reason for exclusion
144.	Uppal MTN. Classification of mammograms for breast cancer detection using fusion of discrete cosine transform and discrete wavelet transform features. Biomedical Research (India). 2016;27(2):322-7.	Population – Image type
145.	Vaijayanthi N, Caroline BE, Murugan VS. Automatic detection of masses in mammograms using bi-dimensional empirical mode decomposition. Journal of Medical Imaging and Health Informatics. 2018;8(7):1326-41.	Population – Image type
146.	Velikova M, Lucas PJ, Karssemeijerb N. Using local context information to improve automatic mammographic mass detection. Stud Health Technol Inform. 2010;160(Pt 2):1291-5.	Population – Image type
147.	Vikhe PS, Thool VR. Mass Detection in Mammographic Images Using Wavelet Processing and Adaptive Threshold Technique. J Med Syst. 2016;40(4):82.	Population – Image type
148.	Wang H, Feng J, Bu Q, Liu F, Zhang M, Ren Y, et al. Breast Mass Detection in Digital Mammogram Based on Gestalt Psychology. J. 2018;2018:4015613.	Population – Image type
149.	Wang X, Li L, Liu W, Xu W, Lederman D, Zheng B. An interactive system for computer-aided diagnosis of breast masses. J Digit Imaging. 2012;25(5):570-9.	Population – Image type
150.	Wei J, Chan HP, Zhou C, Wu YT, Sahiner B, Hadjiiski LM, et al. Computer-aided detection of breast masses: four-view strategy for screening mammography. Med Phys. 2011;38(4):1867-76.	Population – Image type
Population – Mammography type not reported (n=8)		
151.	Li Y, Chen H, Yang Y, Cheng L, Cao L. A bilateral analysis scheme for false positive reduction in mammogram mass detection. Comput Biol Med. 2015;57:84-95.	Population – Mammography type not reported
152.	Moin P, Deshpande R, Sayre J, Messer E, Gupte S, Romsdahl H, et al. An observer study for a computer-aided reading protocol (CARP) in the screening environment for digital mammography. Acad Radiol. 2011;18(11):1420-9.	Population – Mammography type not reported
153.	Mutasa S, Chang P, Nemer J, Van Sant EP, Sun M, McIlvride A, et al. Prospective Analysis Using a Novel CNN Algorithm to Distinguish Atypical Ductal Hyperplasia From Ductal Carcinoma in Situ in Breast. Clinical Breast Cancer. 2020.	Population – Mammography type not reported
154.	Padmavathy TV, Vimalkumar MN, Bhargava DS. Adaptive clustering based breast cancer detection with ANFIS classifier using mammographic images. Cluster Computing-the Journal of Networks Software Tools and Applications. 2019;22:13975-84.	Population – Mammography type not reported

Reference		Main reason for exclusion
155.	Ribli D, Horvath A, Unger Z, Pollner P, Csabai I. Detecting and classifying lesions in mammograms with Deep Learning. Sci. 2018;8(1):4165.	Population – Mammography type not reported
156.	Sasikala S, Bharathi M, Ezhilarasi M, Reddy MR, Arunkumar S. Fusion of MLO and CC view binary patterns to improve the performance of breast cancer diagnosis. Current Medical Imaging Reviews. 2018;14(4):651-8.	Population – Mammography type not reported
157.	Sasikala S, Ezhilarasi M. Comparative analysis of serial and parallel fusion on texture features for improved breast cancer diagnosis. Current Medical Imaging Reviews. 2018;14(6):957-68.	Population – Mammography type not reported
158.	Vimalkumar MN, Helenprabha K. Adaptive neuro-fuzzy inference system for classification of mammographic image using electromagnetism-like optimisation. International Journal of Biomedical Engineering and Technology. 2018;26(3-4):376-84.	Population – Mammography type not reported
Population – Incomplete images (e.g. regions of interest) (n=8)		
159.	Sun W, Tseng TL, Zhang J, Qian W. Computerized breast cancer analysis system using three stage semi-supervised learning method. Comput Methods Programs Biomed. 2016;135:77-88.	Population – Incomplete images
160.	Tan M, Pu J, Zheng B. A new and fast image feature selection method for developing an optimal mammographic mass detection scheme. Med Phys. 2014;41(8):081906.	Population – Incomplete images
161.	Wang XH, Park SC, Zheng B. Assessment of performance and reliability of computer-aided detection scheme using content-based image retrieval approach and limited reference database. J Digit Imaging. 2011;24(2):352-9.	Population – Incomplete images
162.	Yu X, Kang C, Guttery DS, Kadry S, Chen Y, Zhang YD. ResNet-SCDA-50 for breast abnormality classification. IEEE/ACM transactions on computational biology and bioinformatics. 2020;13.	Population – Incomplete images
163.	Zhang Y, Tomuro N, Furst J, Raicu DS. Building an ensemble system for diagnosing masses in mammograms. Int. 2012;7(2):323-9.	Population – Incomplete images
164.	Zyout I, Togneri R. Empirical mode decomposition of digital mammograms for the statistical based characterization of architectural distortion. Conf Proc IEEE Eng Med Biol Soc. 2015;2015:109-12.	Population – Incomplete images
165.	Zyout I, Togneri R. A new approach for the detection of architectural distortions using textural analysis of surrounding tissue. Conf Proc IEEE Eng Med Biol Soc. 2016;2016:3965-8.	Population – Incomplete images

Reference	Main reason for exclusion
166. Zyout I, Togneri R. A computer-aided detection of the architectural distortion in digital mammograms using the fractal dimension measurements of BEMD. Comput Med Imaging Graph. 2018;70:173-84.	Population – Incomplete images
Population – Subpolulation (e.g. only cancer cases) (n=3)	
167. Bolivar AV, Gomez SS, Merino P, Alonso-Bartolome P, Garcia EO, Cacho PM, et al. Computer-aided detection system applied to full-field digital mammograms. Acta Radiol. 2010;51(10):1086-92.	Population - Subpopulation
168. Cho KR, Seo BK, Woo OH, Song SE, Choi J, Whang SY, et al. Breast Cancer Detection in a Screening Population: Comparison of Digital Mammography, Computer-Aided Detection Applied to Digital Mammography and Breast Ultrasound. Journal of Breast Cancer. 2016;19(3):316-23.	Population - Subpopulation
169. Hamza AO, El-Sanosi MD, Habbani AK, Mustafa NA, Khider MO. Computer-aided detection of benign tumors of the female breast. Journal of Clinical Engineering. 2013;38(1):32-7.	Population - Subpopulation
Population – <90% screening mammograms or unclear proportion (n=10)	
170. Al-Najdawi N, Biltawi M, Tedmori S. Mammogram image visual enhancement, mass segmentation and classification. Applied Soft Computing. 2015;35:175-85.	Population – <90% screening mammograms or unclear proportion
171. Angayarkanni N, Kumar D, Arunachalam G. The application of image processing techniques for detection and classification of cancerous tissue in digital mammograms. Journal of Pharmaceutical Sciences and Research. 2016;8(10):1179-83.	Population – <90% screening mammograms or unclear proportion
172. Cascio D, Fauci F, Iacomi M, Raso G, Magro R, Castrogiovanni D, et al. Computer-aided diagnosis in digital mammography: Comparison of two commercial systems. Imaging in Medicine. 2014;6(1):13-20.	Population – <90% screening mammograms or unclear proportion
173. Diz J, Marreiros G, Freitas A. Applying Data Mining Techniques to Improve Breast Cancer Diagnosis. J Med Syst. 2016;40(9).	Population – <90% screening mammograms or unclear proportion
174. Langarizadeh M, Mahmud R, Bagherzadeh R. Detection of masses and microcalcifications in digitalmammogram images using fuzzy logic. Asian Biomedicine. 2016;10(4):345-50.	Population – <90% screening mammograms or unclear proportion
175. Mutasa S, Chang P, Van Sant EP, Nemer J, Liu M, Karcich J, et al. Potential Role of Convolutional Neural Network Based Algorithm in Patient Selection for DCIS Observation Trials Using a Mammogram Dataset. Acad Radiol. 2020;27(6):774-9.	Population – <90% screening mammograms or unclear proportion

Reference		Main reason for exclusion
176.	Sasaki M, Tozaki M, Rodriguez-Ruiz A, Yotsumoto D, Ichiki Y, Terawaki A, et al. Artificial intelligence for breast cancer detection in mammography: experience of use of the ScreenPoint Medical Transpara system in 310 Japanese women. Breast Cancer. 2020;27(4):642-51.	Population – <90% screening mammograms or unclear proportion
177.	Rodriguez-Ruiz A, Lang K, Gubern-Merida A, Teuwen J, Broeders M, Gennaro G, et al. Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. Eur Radiol. 2019;29(9):4825-32.	Population – <90% screening mammograms or unclear proportion
178.	Soulami KB, Kaabouch N, Saidi MN, Tamtaoui A. An evaluation and ranking of evolutionary algorithms in segmenting abnormal masses in digital mammograms. Multimedia Tools and Applications. 2020;79(27-28):18941-79.	Population – <90% screening mammograms or unclear proportion
179.	Zheng J, Lin DA, Gao ZJ, Wang S, He MJ, Fan JP. Deep Learning Assisted Efficient AdaBoost Algorithm for Breast Cancer Detection and Early Diagnosis. Ieee Access. 2020;8:96946-54.	Population – <90% screening mammograms or unclear proportion
Internal validation – Cross validation (n=91)		
180.	Abdar M, Zomorodi-Moghadam M, Zhou XJ, Gururajan R, Tao XH, Barua PD, et al. A new nested ensemble technique for automated diagnosis of breast cancer. Pattern Recognition Letters. 2020;132:123-31.	Internal validation – Cross validation
181.	Agarwal R, Diaz O, Yap MH, Llado X, Marti R. Deep learning for mass detection in Full Field Digital Mammograms. Comput Biol Med. 2020;121:103774.	Internal validation – Cross validation
182.	Ahmadi A, Afshar P. Intelligent breast cancer recognition using particle swarm optimization and support vector machines. Journal of Experimental & Theoretical Artificial Intelligence. 2016;28(6):1021-34.	Internal validation – Cross validation
183.	Al-Antari MA, Al-Masni MA, Choi MT, Han SM, Kim TS. A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. Int J Med Inf. 2018;117:44-54.	Internal validation – Cross validation
184.	Al-Masni MA, Al-Antari MA, Park JM, Gi G, Kim TY, Rivera P, et al. Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. Comput Methods Programs Biomed. 2018;157:85-94.	Internal validation – Cross validation
185.	Aminikhanghahi S, Shin S, Wang W, Jeon SI, Son SH. A new fuzzy Gaussian mixture model (FGMM) based algorithm for mammography tumor image classification. Multimedia Tools and Applications. 2017;76(7):10191-205.	Internal validation – Cross validation

Reference		Main reason for exclusion
186.	Arzehgar A, Khalilzadeh MM, Varshoei F. Assessment and classification of mass lesions based on expert knowledge using mammographic analysis. Current Medical Imaging Reviews. 2019;15(2):199-208.	Internal validation – Cross validation
187.	Ayer T, Alagoz O, Chhatwal J, Shavlik JW, Kahn CE, Jr., Burnside ES. Breast cancer risk estimation with artificial neural networks revisited: discrimination and calibration. Cancer. 2010;116(14):3310-21.	Internal validation – Cross validation
188.	Azar AT, El-Metwally SM. Decision tree classifiers for automated medical diagnosis. Neural Computing & Applications. 2013;23(7-8):2387-403.	Internal validation – Cross validation
189.	Azar AT, El-Said SA. Probabilistic neural network for breast cancer classification. Neural Computing & Applications. 2013;23(6):1737-51.	Internal validation – Cross validation
190.	Azar AT, El-Said SA. Performance analysis of support vector machines classifiers in breast cancer mammography recognition. Neural Computing & Applications. 2014;24(5):1163-77.	Internal validation – Cross validation
191.	Beura S, Majhi B, Dash R, Roy S. Classification of mammogram using two-dimensional discrete orthonormal S-transform for breast cancer detection. Healthc. 2015;2(2):46-51.	Internal validation – Cross validation
192.	Bouyer A. Breast cancer diagnosis using data mining methods, cumulative histogram features, and gary level co-occurrence matrix. Current Medical Imaging Reviews. 2017;13(4):460-70.	Internal validation – Cross validation
193.	Cao P, Liu X, Bao H, Yang J, Zhao D. Restricted Boltzmann machines based oversampling and semi-supervised learning for false positive reduction in breast CAD. Bio-Medical Materials and Engineering. 2015;26(Supplement 1):S1541-S7.	Internal validation – Cross validation
194.	Carneiro G, Nascimento J, Bradley AP. Automated Analysis of Unregistered Multi-View Mammograms With Deep Learning. IEEE Trans Med Imaging. 2017;36(11):2355-65.	Internal validation – Cross validation
195.	Casti P, Mencattini A, Salmeri M, Ancona A, Lorusso M, Pepe ML, et al. Towards localization of malignant sites of asymmetry across bilateral mammograms. Computer Methods and Programs in Biomedicine. 2017;140:11-8.	Internal validation – Cross validation
196.	Casti P, Mencattini A, Salmeri M, Ancona A, Mangeri F, Pepe ML, et al. Contour-independent detection and classification of mammographic lesions. Biomedical Signal Processing and Control. 2016;25:165-77.	Internal validation – Cross validation

Reference		Main reason for exclusion
197.	Casti P, Mencattini A, Salmeri M, Ancona A, Mangieri F, Rangayyan RM. Development and validation of a fully automated system for detection and diagnosis of mammographic lesions. Conf Proc IEEE Eng Med Biol Soc. 2014;2014:4667-70.	Internal validation – Cross validation
198.	Celaya-Padilla J, Martinez-Torteya A, Rodriguez-Rojas J, Galvan-Tejada J, Trevino V, Tamez-Pena J. Bilateral Image Subtraction and Multivariate Models for the Automated Triaging of Screening Mammograms. Biomed Res Int. 2015;2015:231656.	Internal validation – Cross validation
199.	Celaya-Padilla JM, Guzman-Valdivia CH, Galvan-Tejada CE, Galvan-Tejada JI, Gamboa-Rosales H, Garza-Veloz I, et al. Contralateral asymmetry for breast cancer detection: A CADx approach. Biocybernetics and Biomedical Engineering. 2018;38(1):115-25.	Internal validation – Cross validation
200.	Chakraborty J, Midya A, Rabidas R. Computer-aided detection and diagnosis of mammographic masses using multi-resolution analysis of oriented tissue patterns. Expert Systems with Applications. 2018;99:168-79.	Internal validation – Cross validation
201.	Chen HL, Yang B, Liu J, Liu DY. A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. Expert Systems with Applications. 2011;38(7):9014-22.23.	Internal validation – Cross validation
202.	Chen HL, Yang B, Wang G, Wang SJ, Liu J, Liu DY. Support vector machine based diagnostic system for breast cancer using swarm intelligence. J Med Syst. 2012;36(4):2505-19.	Internal validation – Cross validation
203.	Chen X, Zargari A, Hollingsworth AB, Liu H, Zheng B, Qiu Y. Applying a new quantitative image analysis scheme based on global mammographic features to assist diagnosis of breast cancer. Comput Methods Programs Biomed. 2019;179:104995.	Internal validation – Cross validation
204.	Choi JY. A generalized multiple classifier system for improving computer-aided classification of breast masses in mammography. Biomedical Engineering Letters. 2015;5(4):251-62.	Internal validation – Cross validation
205.	Choi JY, Kim DH, Plataniotis KN, Ro YM. Combining multiple feature representations and AdaBoost ensemble learning for reducing false-positive detections in computer-aided detection of masses on mammograms. Conf Proc IEEE Eng Med Biol Soc. 2012;2012:4394-7.	Internal validation – Cross validation
206.	Choi JY, Kim DH, Plataniotis KN, Ro YM. Computer-aided detection (CAD) of breast masses in mammography: combined detection and ensemble classification. Phys Med Biol. 2014;59(14):3697-719.	Internal validation – Cross validation

Reference		Main reason for exclusion
207.	Costa DD, Campos LF, Barros AK. Classification of breast tissue in mammograms using efficient coding. Biomed. 2011;10:55.	Internal validation – Cross validation
208.	Dhungel N, Carneiro G, Bradley AP. A deep learning approach for the analysis of masses in mammograms with minimal user intervention. Med Image Anal. 2017;37:114-28.	Internal validation – Cross validation
209.	do Nascimento MZ, Martins AS, Neves LA, Ramos RP, Flores EL, Carrijo GA. Classification of masses in mammographic image using wavelet domain features and polynomial classifier. Expert Systems with Applications. 2013;40(15):6213-21.	Internal validation – Cross validation
210.	Dong M, Lu X, Ma Y, Guo Y, Ma Y, Wang K. An Efficient Approach for Automated Mass Segmentation and Classification in Mammograms. J Digit Imaging. 2015;28(5):613-25.	Internal validation – Cross validation
211.	Drukker K, Giger ML, Joe BN, Kerlikowske K, Greenwood H, Drukteinis JS, et al. Combined Benefit of Quantitative Three-Compartment Breast Image Analysis and Mammography Radiomics in the Classification of Breast Masses in a Clinical Data Set. Radiology. 2019;290(3):621-8.	Internal validation – Cross validation
212.	Eltrass AS, Salama MS. Fully automated scheme for computer-aided detection and breast cancer diagnosis using digitised mammograms. Iet Image Processing. 2020;14(3):495-505.	Internal validation – Cross validation
213.	Esmaeili M, Ayyoubzadeh SM, Ahmadinejad N, Ghazisaeedi M, Nahvijou A, Maghooli K. A decision support system for mammography reports interpretation. Health Inf Sci Syst. 2020;8(1):17.	Internal validation – Cross validation
214.	Fanizzi A, Basile TMA, Losurdo L, Bellotti R, Bottigli U, Dentamaro R, et al. A machine learning approach on multiscale texture analysis for breast microcalcification diagnosis. BMC Bioinformatics. 2020;21(Suppl 2):91.	Internal validation – Cross validation
215.	Ganesan K, Acharya UR, Chua CK, Lim CM, Abraham KT. One-Class Classification of Mammograms Using Trace Transform Functionals. Ieee Transactions on Instrumentation and Measurement. 2014;63(2):304-11.	Internal validation – Cross validation
216.	Ganesan K, Acharya UR, Chua CK, Min LC, Abraham TK. Automated diagnosis of mammogram images of breast cancer using discrete wavelet transform and spherical wavelet transform features: a comparative study. Technol Cancer Res Treat. 2014;13(6):605-15.	Internal validation – Cross validation
217.	Gao F, Wu T, Li J, Zheng B, Ruan L, Shang D, et al. SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis. Comput Med Imaging Graph. 2018;70:53-62.	Internal validation – Cross validation

Reference		Main reason for exclusion
218.	Garcia-Manso A, Garcia-Orellana CJ, Gonzalez-Velasco H, Gallardo-Caballero R, Macias MM. Consistent performance measurement of a system to detect masses in mammograms based on blind feature extraction. Biomed. 2013;12:2.	Internal validation – Cross validation
219.	Ghasemzadeh A, Azad SS, Esmaeili E. Breast cancer detection based on Gabor-wavelet transform and machine learning methods. International Journal of Machine Learning and Cybernetics. 2019;10(7):1603-12.	Internal validation – Cross validation
220.	Ghosh A. Artificial Intelligence Using Open Source BI-RADS Data Exemplifying Potential Future Use. J. 2019;16(1):64-72.	Internal validation – Cross validation
221.	Gomez-Flores W, Hernandez-Lopez J. Assessment of the invariance and discriminant power of morphological features under geometric transformations for breast tumor classification. Computer Methods and Programs in Biomedicine. 2020;185.	Internal validation – Cross validation
222.	Ha R, Chang P, Karcich J, Mutasa S, Pascual Van Sant E, Liu MZ, et al. Convolutional Neural Network Based Breast Cancer Risk Stratification Using a Mammographic Dataset. Acad Radiol. 2019;26(4):544-9.	Internal validation – Cross validation
223.	Hai J, Tan H, Chen J, Wu M, Qiao K, Xu J, et al. Multi-level features combined end-to-end learning for automated pathological grading of breast cancer on digital mammograms. Computerized Medical Imaging and Graphics. 2019;71:58-66.	Internal validation – Cross validation
224.	Heidari M, Mirniaharikandehei S, Liu W, Hollingsworth AB, Liu H, Zheng B. Development and Assessment of a New Global Mammographic Image Feature Analysis Scheme to Predict Likelihood of Malignant Cases. IEEE Trans Med Imaging. 2020;39(4):1235-44.	Internal validation – Cross validation
225.	Huynh BQ, Li H, Giger ML. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Med Imaging (Bellingham). 2016;3(3).	Internal validation – Cross validation
226.	Jacomini RS, Nascimento MZ, Dantas RD, Ramos RP. Classification of mass in two views mammograms: Use of analysis of variance (ANOVA) for reduction of the features. Recent Patents on Medical Imaging. 2013;3(1):80-8.	Internal validation – Cross validation
227.	Keles A, Keles A. Extracting fuzzy rules for the diagnosis of breast cancer. Turkish Journal of Electrical Engineering and Computer Sciences. 2013;21(5):1495-503.	Internal validation – Cross validation
228.	Khan S, Hussain M, Aboalsamh H, Mathkour H, Bebis G, Zakariah M. Optimized Gabor features for mass classification in mammography. Applied Soft Computing. 2016;44:267-80.	Internal validation – Cross validation

Reference		Main reason for exclusion
229.	Khan S, Khan A, Maqsood M, Aadil F, Ghazanfar MA. Optimized Gabor Feature Extraction for Mass Classification Using Cuckoo Search for Big Data E-Healthcare. Journal of Grid Computing. 2019;17(2):239-54.	Internal validation – Cross validation
230.	Kilic N, Gorgel P, Ucan ON, Sertbas A. Mammographic mass detection using wavelets as input to neural networks. J Med Syst. 2010;34(6):1083-8.	Internal validation – Cross validation
231.	Kim DH, Choi JY, Ro YM. Region based stellate features combined with variable selection using AdaBoost learning in mammographic computer-aided detection. Computers in Biology and Medicine. 2015;63:238-50.	Internal validation – Cross validation
232.	Kim DH, Lee SH, Ro YM. Mass type-specific sparse representation for mass classification in computer-aided detection on mammograms. Biomed. 2013;12 Suppl 1:S3.	Internal validation – Cross validation
233.	Kim S. Margin-maximised redundancy-minimised SVM-RFE for diagnostic classification of mammograms. Int J Data Min Bioinform. 2014;10(4):374-90.	Internal validation – Cross validation
234.	Kooi T, van Ginneken B, Karssemeijer N, den Heeten A. Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural network. Med Phys. 2017;44(3):1017-27.	Internal validation – Cross validation
235.	Kozegar E, Soryani M. A cost-sensitive Bayesian combiner for reducing false positives in mammographic mass detection. Biomed Tech (Berl). 2019;64(1):39-52.	Internal validation – Cross validation
236.	Kozegar E, Soryani M, Minaei B, Domingues I. Assessment of a novel mass detection algorithm in mammograms. J Cancer Res Ther. 2013;9(4):592-600.	Internal validation – Cross validation
237.	Kyono T, Gilbert FJ, van der Schaar M. Improving Workflow Efficiency for Mammography Using Machine Learning. J. 2020;17(1 Pt A):56-63.	Internal validation – Cross validation (UK Tommy dataset)
238.	Lakshmanan R, Shiji TP, Jacob SM, Pratab T, Thomas C, Thomas V. Detection of architectural distortion in mammograms using geometrical properties of thinned edge structures. Intelligent Automation and Soft Computing. 2017;23(1):183-97.	Internal validation – Cross validation
239.	Lee J, Nishikawa RM. Detecting mammographically occult cancer in women with dense breasts using deep convolutional neural network and Radon Cumulative Distribution Transform. J Med Imaging (Bellingham). 2019;6(4):044502.	Internal validation – Cross validation
240.	Li H, Zhuang S, Li DA, Zhao J, Ma Y. Benign and malignant classification of mammogram images based on deep learning. Biomedical Signal Processing and Control. 2019;51:347-54.	Internal validation – Cross validation

Reference		Main reason for exclusion
241.	Shan LH, Faust O, Yu W. Data mining framework for breast cancer detection in mammograms: A hybrid feature extraction paradigm. Journal of Medical Imaging and Health Informatics. 2014;4(5):756-65.	Internal validation – Cross validation
242.	Liu N, Qi ES, Xu M, Gao B, Liu GQ. A novel intelligent classification model for breast cancer diagnosis. Information Processing & Management. 2019;56(3):609-23.	Internal validation – Cross validation
243.	Luo ST, Cheng BW. Diagnosing breast masses in digital mammography using feature selection and ensemble methods. J Med Syst. 2012;36(2):569-77.	Internal validation – Cross validation
244.	Mednikov Y, Nehemia S, Zheng B, Benzaquen O, Lederman D. Transfer Representation Learning using Inception-V3 for the Detection of Masses in Mammography. Conf Proc IEEE Eng Med Biol Soc. 2018;2018:2587-90.	Internal validation – Cross validation
245.	Melendez J, Sanchez CI, van Ginneken B, Karssemeijer N. Improving mass candidate detection in mammograms via feature maxima propagation and local feature selection. Med Phys. 2014;41(8):081904.	Internal validation – Cross validation
246.	Milosevic M, Jankovic D, Peulic A. Comparative analysis of breast cancer detection in mammograms and thermograms. Biomed Tech (Berl). 2015;60(1):49-56.	Internal validation – Cross validation
247.	Min H, Chandra SS, Crozier S, Bradley AP. Multi-scale sifting for mammographic mass detection and segmentation. Biomedical Physics and Engineering Express. 2019;5(2).	Internal validation – Cross validation
248.	Naghibi S, Teshnehlab M, Shoorehdeli MA. Breast cancer classification based on advanced multi dimensional fuzzy neural network. J Med Syst. 2012;36(5):2713-20.	Internal validation – Cross validation
249.	Nassif H, Wu Y, Page D, Burnside E. Logical Differential Prediction Bayes Net, improving breast cancer diagnosis for older women. AMIA Annu Symp Proc. 2012;2012:1330-9.	Internal validation – Cross validation
250.	Nilashi M, Ibrahim O, Ahmadi H, Shahmoradi L. A knowledge-based system for breast cancer classification using fuzzy logic method. Telematics and Informatics. 2017;34(4):133-44.	Internal validation – Cross validation
251.	Oliver A, Freixenet J, Marti J, Perez E, Pont J, Denton ER, et al. A review of automatic mass detection and segmentation in mammographic images. Med Image Anal. 2010;14(2):87-110.	Internal validation – Cross validation
252.	Peng J, Bao C, Hu C, Wang X, Jian W, Liu W. Automated mammographic mass detection using deformable convolution and multiscale features. Medical and Biological Engineering and Computing. 2020;58(7):1405-17.	Internal validation – Cross validation

Reference		Main reason for exclusion
253.	Perez NP, Guevara Lopez MA, Silva A, Ramos I. Improving the Mann-Whitney statistical test for feature selection: an approach in breast cancer diagnosis on mammography. Artif Intell Med. 2015;63(1):19-31.	Internal validation – Cross validation
254.	Qiu Y, Yan S, Gundreddy RR, Wang Y, Cheng S, Liu H, et al. A new approach to develop computer-aided diagnosis scheme of breast mass classification using deep learning technology. Journal of X-Ray Science and Technology. 2017;25(5):751-63.	Internal validation – Cross validation
255.	Ragab DA, Sharkas M, Attallah O. Breast cancer diagnosis using an efficient CAD system based on multiple classifiers. Diagnostics (Basel). 2019;9(4).	Internal validation – Cross validation
256.	Sapate S, Talbar S, Mahajan A, Sable N, Desai S, Thakur M. Breast cancer diagnosis using abnormalities on ipsilateral views of digital mammograms. Biocybernetics and Biomedical Engineering. 2020;40(1):290-305.	Internal validation – Cross validation
257.	Sapate SG, Mahajan A, Talbar SN, Sable N, Desai S, Thakur M. Radiomics based detection and characterization of suspicious lesions on full field digital mammograms. Comput Methods Programs Biomed. 2018;163:1-20.	Internal validation – Cross validation
258.	Suresh A, Udendhran R, Balamurgan M, Varatharajan R. A Novel Internet of Things Framework Integrated with Real Time Monitoring for Intelligent Healthcare Environment. J Med Syst. 2019;43(6):165.	Internal validation – Cross validation
259.	Tan M, Aghaei F, Wang Y, Zheng B. Developing a new case based computer-aided detection scheme and an adaptive cueing method to improve performance in detecting mammographic lesions. Phys Med Biol. 2017;62(2):358-76.	Internal validation – Cross validation
260.	Tan M, Pu JT, Zheng B. Reduction of false-positive recalls using a computerized mammographic image feature analysis scheme. Physics in Medicine and Biology. 2014;59(15):4357-73.	Internal validation – Cross validation
261.	Torabi M, Razavian SM, Vaziri R, Vosoughi-Vahdat B. A Wavelet-packet-based approach for breast cancer classification. Conf Proc IEEE Eng Med Biol Soc. 2011;2011:5100-3.	Internal validation – Cross validation
262.	Velikova M, Lucas PJ, Samulski M, Karssemeijer N. A probabilistic framework for image information fusion with an application to mammographic analysis. Med Image Anal. 2012;16(4):865-75.	Internal validation – Cross validation
263.	Velikova M, Lucas PJ, Samulski M, Karssemeijer N. On the interplay of machine learning and background knowledge in image interpretation by Bayesian networks. Artif Intell Med. 2013;57(1):73-86.	Internal validation – Cross validation

Reference	Main reason for exclusion
264. Wang Z, Huang Y, Li M, Zhang H, Li C, Xin J, et al. Breast mass detection and diagnosis using fused features with density. Journal of X-Ray Science and Technology. 2019;27(2):321-42.	Internal validation – Cross validation
265. Wang Z, Yu G, Kang Y, Zhao Y, Qu Q. Breast tumor detection in digital mammography based on extreme learning machine. Neurocomputing. 2014;128:175-84.	Internal validation – Cross validation
266. Xie W, Li Y, Ma Y. Breast mass classification in digital mammography based on extreme learning machine. Neurocomputing. 2016;Part 3. 173:930-41.	Internal validation – Cross validation
267. Yang LY, Xu ZS. Feature extraction by PCA and diagnosis of breast tumors using SVM with DE-based parameter tuning. International Journal of Machine Learning and Cybernetics. 2019;10(3):591-601.	Internal validation – Cross validation
268. Zadeh HG, Seryasat OR, Haddadnia J. Assessment of a novel computer aided mass diagnosis system in mammograms. Biomedical Research (India). 2017;28(7):3129-35.	Internal validation – Cross validation
269. Zeng JM, Gimenez F, Burnside ES, Rubin DL, Shachter R. A Probabilistic Model to Support Radiologists' Classification Decisions in Mammography Practice. Med Decis Making. 2019;39(3):208-16.	Internal validation – Cross validation
270. Zhang C, Zhao J, Niu J, Li D. New convolutional neural network model for screening and diagnosis of mammograms. PLoS ONE. 2020;15(8):e0237674.	Internal validation – Cross validation
Internal validation – Leave-one out (n=12)	
271. Casti P, Mencattini A, Salmeri M, Rangayyan RM. Analysis of structural similarity in mammograms for detection of bilateral asymmetry. IEEE Trans Med Imaging. 2015;34(2):662-71.	Internal validation – Leave-one out
272. Dhahbi S, Barhoumi W, Zagrouba E. Breast cancer diagnosis in digitized mammograms using curvelet moments. Comput Biol Med. 2015;64:79-90.	Internal validation – Leave-one out
273. Drukker K, Duewer F, Giger ML, Malkov S, Flowers CI, Joe B, et al. Mammographic quantitative image analysis and biologic image composition for breast lesion characterization and classification. Med Phys. 2014;41(3):031915.	Internal validation – Leave-one out
274. Kelder A, Lederman D, Zheng B, Zigel Y. A new computer-aided detection approach based on analysis of local and global mammographic feature asymmetry. Med Phys. 2018;45(4):1459-70.	Internal validation – Leave-one out
275. Kendall EJ, Barnett MG, Chytyk-Praznik K. Automatic detection of anomalies in screening mammograms. BMC med. 2013;13:43.	Internal validation – Leave-one out

Reference		Main reason for exclusion
276.	Kendall EJ, Flynn MT. Automated breast image classification using features from its discrete cosine transform. PLoS ONE. 2014;9(3):e91015.	Internal validation – Leave-one out
277.	Liang C, Bian Z, Lv W, Chen S, Zeng D, Ma J. A computer-aided diagnosis scheme of breast lesion classification using GLGLM and shape features: Combined-view and multi-classifiers. Phys Med. 2018;55:61-72.	Internal validation – Leave-one out
278.	Muramatsu C, Hara T, Endo T, Fujita H. Breast mass classification on mammograms using radial local ternary patterns. Comput Biol Med. 2016;72:43-53.	Internal validation – Leave-one out
279.	Ramos-Pollan R, Guevara-Lopez MA, Suarez-Ortega C, Diaz-Herrero G, Franco-Valiente JM, Rubio-Del-Solar M, et al. Discovering mammography-based machine learning classifiers for breast cancer diagnosis. J Med Syst. 2012;36(4):2259-69.	Internal validation – Leave-one out
280.	Rangayyan RM, Nguyen TM, Ayres FJ, Nandi AK. Effect of pixel resolution on texture features of breast masses in mammograms. J Digit Imaging. 2010;23(5):547-53.	Internal validation – Leave-one out
281.	Wang X, Lederman D, Tan J, Wang XH, Zheng B. Computerized detection of breast tissue asymmetry depicted on bilateral mammograms: a preliminary study of breast risk stratification. Acad Radiol. 2010;17(10):1234-41.	Internal validation – Leave-one out
282.	Wang Y, Aghaei F, Zarafshani A, Qiu Y, Qian W, Zheng B. Computer-aided classification of mammographic masses using visually sensitive image features. Journal of X-Ray Science and Technology. 2017;25(1):171-86.	Internal validation – Leave-one out
Internal validation – Split sample (n=49)		
283.	Aboutalib SS, Mohamed AA, Berg WA, Zuley ML, Sumkin JH, Wu S. Deep Learning to Distinguish Recalled but Benign Mammography Images in Breast Cancer Screening. Clin Cancer Res. 2018;24(23):5902-9.	Internal validation – Split sample
284.	Akselrod-Ballin A, Chorev M, Shoshan Y, Spiro A, Hazan A, Melamed R, et al. Predicting Breast Cancer by Applying Deep Learning to Linked Health Records and Mammograms. Radiology. 2019;292(2):331-42.	Internal validation – Split sample
285.	Akselrod-Ballin A, Karlinsky L, Alpert S, Hashoul S, Ben-Ari R, Barkan E. A CNN based method for automatic mass detection and classification in mammograms. Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization. 2019;7(3):242-9.	Internal validation – Split sample

Reference		Main reason for exclusion
286.	Alqudah AM, Algharib AMS, Algharib HMS. Computer aided diagnosis system for automatic two stages classification of breast mass in digital mammogram images. Biomedical Engineering - Applications, Basis and Communications. 2019;31(1).	Internal validation – Split sample
287.	Andreadis, II, Spyrou GM, Nikita KS. A CADx scheme for mammography empowered with topological information from clustered microcalcifications' atlases. IEEE j. 2015;19(1):166-73.	Internal validation – Split sample
288.	Arevalo J, Gonzalez FA, Ramos-Pollan R, Oliveira JL, Guevara Lopez MA. Representation learning for mammography mass lesion classification with convolutional neural networks. Comput Methods Programs Biomed. 2016;127:248-57.	Internal validation – Split sample
289.	Bakkouri I, Afdel K. Multi-scale CNN based on region proposals for efficient breast abnormality recognition. Multimedia Tools and Applications. 2019;78(10):12939-60.	Internal validation – Split sample
290.	Banaem HY, Dehnavi AM, Shahnazi M. Ensemble Supervised Classification Method Using the Regions of Interest and Grey Level Co-Occurrence Matrices Features for Mammograms Data. Iranian Journal of Radiology. 2015;12(3).	Internal validation – Split sample
291.	Bandeira Diniz JO, Bandeira Diniz PH, Azevedo Valente TL, Correa Silva A, de Paiva AC, Gattass M. Detection of mass regions in mammograms by bilateral analysis adapted to breast density using similarity indexes and convolutional neural networks. Comput Methods Programs Biomed. 2018;156:191-207.	Internal validation – Split sample
292.	Barkana BD, Saricicek I. Classification of breast masses in mammograms using 2D homomorphic transform features and supervised classifiers. Journal of Medical Imaging and Health Informatics. 2017;7(7):1566-71.	Internal validation – Split sample
293.	Beura S, Majhi B, Dash R. Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer. Neurocomputing. 2015;154:1-14.	Internal validation – Split sample
294.	Bhardwaj A, Tiwari A. Breast cancer diagnosis using Genetically Optimized Neural Network model. Expert Systems with Applications. 2015;42(10):4611-20.	Internal validation – Split sample
295.	Boumaraf S, Liu X, Ferkous C, Ma X. A New Computer-Aided Diagnosis System with Modified Genetic Feature Selection for BI-RADS Classification of Breast Masses in Mammograms. Biomed Res Int. 2020;2020:7695207.	Internal validation – Split sample
296.	Cha KH, Petrick N, Pezeshk A, Graff CG, Sharma D, Badal A, et al. Evaluation of data augmentation via synthetic images for improved breast mass detection on mammograms using deep learning. J Med Imaging (Bellingham). 2020;7(1).	Internal validation – Split sample

Reference		Main reason for exclusion
297.	Chinnasamy VA, Shashikumar DR. Breast cancer detection in mammogram image with segmentation of tumour region. International Journal of Medical Engineering and Informatics. 2020;12(1):1-18.	Internal validation – Split sample
298.	Choi JY, Kim DH, Plataniotis KN, Ro YM. Classifier ensemble generation and selection with multiple feature representations for classification applications in computer-aided detection and diagnosis on mammography. Expert Systems with Applications. 2016;46:106-21.	Internal validation – Split sample
299.	Chu J, Min H, Liu L, Lu W. A novel computer aided breast mass detection scheme based on morphological enhancement and SLIC superpixel segmentation. Med Phys. 2015;42(7):3859-69.	Internal validation – Split sample
300.	de Nazare Silva J, de Carvalho Filho AO, Correa Silva A, Cardoso de Paiva A, Gattass M. Automatic Detection of Masses in Mammograms Using Quality Threshold Clustering, Correlogram Function, and SVM. J Digit Imaging. 2015;28(3):323-37.	Internal validation – Split sample
301.	de Sampaio WB, Silva AC, de Paiva AC, Gattass M. Detection of masses in mammograms with adaption to breast density using genetic algorithm, phylogenetic trees, LBP and SVM. Expert Systems with Applications. 2015;42(22):8911-28.	Internal validation – Split sample
302.	Dhas AS, Vijikala V. An improved CAD system for abnormal mammogram image classification using SVM with linear kernel. Biomedical Research (India). 2017;28(12):5499-505.	Internal validation – Split sample
303.	Duggento A, Aiello M, Cavaliere C, Cascella GL, Cascella D, Conte G, et al. An Ad Hoc Random Initialization Deep Neural Network Architecture for Discriminating Malignant Breast Cancer Lesions in Mammographic Images. Contrast Media Mol Imaging. 2019;2019:5982834.	Internal validation – Split sample
304.	Duraisamy S, Emperumal S. Computer-aided mammogram diagnosis system using deep learning convolutional fully complex-valued relaxation neural network classifier. Iet Computer Vision. 2017;11(8):656-62.	Internal validation – Split sample
305.	Ferreira P, Fonseca NA, Dutra I, Woods R, Burnside E. Predicting malignancy from mammography findings and image-guided core biopsies. Int J Data Min Bioinform. 2015;11(3):257-76.	Internal validation – Split sample
306.	Gao X, Wang Y, Li X, Tao D. On combining morphological component analysis and concentric morphology model for mammographic mass detection. IEEE Trans Inf Technol Biomed. 2010;14(2):266-73.	Internal validation – Split sample

Reference		Main reason for exclusion
307.	Gastounioti A, Oustimov A, Hsieh MK, Pantalone L, Conant EF, Kontos D. Using Convolutional Neural Networks for Enhanced Capture of Breast Parenchymal Complexity Patterns Associated with Breast Cancer Risk. Acad Radiol. 2018;25(8):977-84.	Internal validation – Split sample
308.	Hinton B, Ma L, Mahmoudzadeh AP, Malkov S, Fan B, Greenwood H, et al. Deep learning networks find unique mammographic differences in previous negative mammograms between interval and screen-detected cancers: a case-case study. Cancer Imaging. 2019;19(1):41.	Internal validation – Split sample
309.	Ibrahim IM, Wahed MA. Visual versus statistical features selection applied to mammography mass detection. Journal of Medical Imaging and Health Informatics. 2014;4(2):237-44.	Internal validation – Split sample
310.	Kim EK, Kim HE, Han K, Kang BJ, Sohn YM, Woo OH, et al. Applying Data-driven Imaging Biomarker in Mammography for Breast Cancer Screening: Preliminary Study. Sci. 2018;8(1):2762.	Internal validation – Split sample
311.	Kooi T, Karssemeijer N. Classifying symmetrical differences and temporal change for the detection of malignant masses in mammography using deep neural networks. J Med Imaging (Bellingham). 2017;4(4):044501.	Internal validation – Split sample
312.	Kooi T, Litjens G, van Ginneken B, Gubern-Merida A, Sanchez CI, Mann R, et al. Large scale deep learning for computer aided detection of mammographic lesions. Med Image Anal. 2017;35:303-12.	Internal validation – Split sample
313.	Lesniak JM, Hupse R, Blanc R, Karssemeijer N, Szekely G. Comparative evaluation of support vector machine classification for computer aided detection of breast masses in mammography. Phys Med Biol. 2012;57(16):5295-307.	Internal validation – Split sample
314.	Li H, Meng X, Wang T, Tang Y, Yin Y. Breast masses in mammography classification with local contour features. Biomed. 2017;16(1).	Internal validation – Split sample
315.	Liu B, Jiang Y. A multitarget training method for artificial neural network with application to computer-aided diagnosis. Med Phys. 2013;40(1):011908.	Internal validation – Split sample
316.	Mao N, Yin P, Wang Q, Liu M, Dong J, Zhang X, et al. Added Value of Radiomics on Mammography for Breast Cancer Diagnosis: A Feasibility Study. J. 2019;16(4 Pt A):485-91.	Internal validation – Split sample
317.	Memon MH, Li JP, Ul Haq A, Memon MH, Zhou W. Breast Cancer Detection in the IOT Health Environment Using Modified Recursive Feature Selection. Wireless Communications & Mobile Computing. 2019;2019.	Internal validation – Split sample

Reference		Main reason for exclusion
318.	Raj JR, Rahman SMK, Anand S. Preliminary evaluation of differentiation of benign and malignant breast tumors using non-invasive diagnostic modalities. Biomedical Research (India). 2016;27(3):596-603.	Internal validation – Split sample
319.	Ramos-Pollan R, Franco JM, Sevilla J, Guevara-Lopez MA, de Posada NG, Loureiro J, et al. Grid infrastructures for developing mammography CAD systems. Conf Proc IEEE Eng Med Biol Soc. 2010;2010:3467-70.	Internal validation – Split sample
320.	Sasikala S, Ezhilarasi M. Fusion of k-Gabor features from medio-lateral-oblique and craniocaudal view mammograms for improved breast cancer diagnosis. J Cancer Res Ther. 2018;14(5):1036-41.	Internal validation – Split sample
321.	Shankar RS, Gupta VM, Murthy KVSS, Rao CS. Breast cancer data classification using machine learning mechanisms. Indian Journal of Public Health Research and Development. 2019;10(5):214-20.	Internal validation – Split sample
322.	Shen L, Margolies LR, Rothstein JH, Fluder E, McBride R, Sieh W. Deep Learning to Improve Breast Cancer Detection on Screening Mammography. Sci. 2019;9(1):12495.	Internal validation – Split sample
323.	Teare P, Fishman M, Benzaquen O, Toledano E, Elnekave E. Malignancy Detection on Mammography Using Dual Deep Convolutional Neural Networks and Genetically Discovered False Color Input Enhancement. J Digit Imaging. 2017;30(4):499-505.	Internal validation – Split sample
324.	Wang J, Yang X, Cai H, Tan W, Jin C, Li L. Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning. Sci. 2016;6:27327.	Internal validation – Split sample
325.	Wang Y, Shi H, Ma S. A new approach to the detection of lesions in mammography using fuzzy clustering. Journal of International Medical Research. 2011;39(6):2256-63.	Internal validation – Split sample
326.	Wu N, Phang J, Park J, Shen Y, Huang Z, Zorin M, et al. Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. IEEE Trans Med Imaging. 2020;39(4):1184-94.	Internal validation – Split sample
327.	Yala A, Schuster T, Miles R, Barzilay R, Lehman C. A Deep Learning Model to Triage Screening Mammograms: A Simulation Study. Radiology. 2019;293(1):38-46.	Internal validation – Split sample
328.	Zadeh Shirazi A, Seyyed Mahdavi Chabok SJ, Mohammadi Z. A novel and reliable computational intelligence system for breast cancer detection. Med Biol Eng Comput. 2018;56(5):721-32.	Internal validation – Split sample

Reference		Main reason for exclusion
329.	Zeiser FA, da Costa CA, Zonta T, Marques NMC, Roehe AV, Moreno M, et al. Segmentation of Masses on Mammograms Using Data Augmentation and Deep Learning. J Digit Imaging. 2020;23:23.	Internal validation – Split sample
330.	Zhang YD, Pan CC, Chen XQ, Wang FB. Abnormal breast identification by nine-layer convolutional neural network with parametric rectified linear unit and rank-based stochastic pooling. Journal of Computational Science. 2018;27:57-68.	Internal validation – Split sample
331.	Zhou L, Ding M, Xu L, Zhou Y, Zhang X. Automated segmentation of malignant mass in mammography using the principal component analysis network based deep learning model. Journal of Medical Imaging and Health Informatics. 2018;8(8):1678-83.	Internal validation – Split sample
Intervention – Detecting subtypes (n=17)		
332.	Bekker AJ, Shalhon M, Greenspan H, Goldberger J. Multi-view probabilistic classification of breast microcalcifications. IEEE Trans Med Imaging. 2016;35(2):645-6536.	Intervention – Detecting subtypes
333.	Berks M, Chen Z, Astley S, Taylor C. Detecting and classifying linear structures in mammograms using random forests. Inf. 2011;22:510-24.	Intervention – Detecting subtypes
334.	Devisuganya S, Suganthe RC. A wrapper based binary shuffled frog algorithm for efficient classification of mammograms. Current Signal Transduction Therapy. 2016;11(2):105-13.	Intervention – Detecting subtypes
335.	Du GM, Dong M, Sun Y, Li SY, Mu XM, Wei HB, et al. A New Method for Detecting Architectural Distortion in Mammograms by NonSubsampled Contourlet Transform and Improved PCNN. Applied Sciences-Basel. 2019;9(22).	Intervention – Detecting subtypes
336.	Huang ML, Hung YH, Lee WM, Li RK, Wang TH. Usage of case-based reasoning, neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. J Med Syst. 2012;36(2):407-14.	Intervention – Detecting subtypes
337.	Jing H, Yang Y, Nishikawa RM. Retrieval boosted computer-aided diagnosis of clustered microcalcifications for breast cancer. Med Phys. 2012;39(2):676-85.	Intervention – Detecting subtypes
338.	Kamra A, Jain VK, Singh S, Mittal S. Characterization of Architectural Distortion in Mammograms Based on Texture Analysis Using Support Vector Machine Classifier with Clinical Evaluation. J Digit Imaging. 2016;29(1):104-14.	Intervention – Detecting subtypes
339.	Keles A, Keles A, Yavuz U. Expert system based on neuro-fuzzy rules for diagnosis breast cancer. Expert Systems with Applications. 2011;38(5):5719-26.	Intervention – Detecting subtypes
340.	Magna G, Casti P, Jayaraman SV, Salmeri M, Mencattini A, Martinelli E, et al. Identification of mammography anomalies for breast cancer detection by an ensemble of	Intervention – Detecting subtypes

Reference		Main reason for exclusion
	classification models based on artificial immune system. Knowledge-Based Systems. 2016;101:60-70.	
341.	Matsubara T, Ito A, Tsunomori A, Hara T, Muramatsu C, Endo T, et al. An automated method for detecting architectural distortions on mammograms using direction analysis of linear structures. Conf Proc IEEE Eng Med Biol Soc. 2015;2015:2661-4.	Intervention – Detecting subtypes
342.	Mordang JJ, Gubern-Merida A, Bria A, Tortorella F, den Heeten G, Karssemeijer N. Improving computer-aided detection assistance in breast cancer screening by removal of obviously false-positive findings. Med Phys. 2017;44(4):1390-401.	Intervention – Detecting subtypes
343.	Mordang JJ, Gubern-Merida A, Den Heeten G, Karssemeijer N. Reducing false positives of microcalcification detection systems by removal of breast arterial calcifications. Med Phys. 2016;43(4):1676-87.	Intervention – Detecting subtypes
344.	Scaranelo AM, Eiada R, Bukhanov K, Crystal P. Evaluation of breast amorphous calcifications by a computer-aided detection system in full-field digital mammography. Br J Radiol. 2012;85(1013):517-22.	Intervention – Detecting subtypes
345.	Shao YZ, Liu LZ, Bie MJ, Li CC, Wu YP, Xie XM, et al. Characterizing the Clustered Microcalcifications on Mammograms to Predict the Pathological Classification and Grading: A Mathematical Modeling Approach. J Digit Imaging. 2011;24(5):764-71.	Intervention – Detecting subtypes
346.	Tiedeu A, Daul C, Kentsop A, Graebling P, Wolf D. Texture-based analysis of clustered microcalcifications detected on mammograms. Digital Signal Processing. 2012;22(1):124-32.	Intervention – Detecting subtypes
347.	Wang X, Li L, Xu W, Liu W, Lederman D, Zheng B. Improving the performance of computer-aided detection of subtle breast masses using an adaptive cueing method. Phys Med Biol. 2012;57(2):561-75.	Intervention – Detecting subtypes
348.	Wang X, Li L, Xu W, Liu W, Lederman D, Zheng B. Improving performance of computer-aided detection of masses by incorporating bilateral mammographic density asymmetry: an assessment. Acad Radiol. 2012;19(3):303-10.	Intervention – Detecting subtypes
Intervention – No detection/classification (n=7)		
349.	Angayarkanni SP, Kamal NB, Thangaiya RJ. Dynamic graph cut based segmentation of mammogram. Springerplus. 2015;4.	Intervention – No detection/classification

Reference		Main reason for exclusion
350.	Dabagov AR, Gorbunov VA, Filist SA, Malyutina IA, Kondrashov DS. An Automated System for Classification of Radiographs of the Breast. Biomedical Engineering. 2020;53(6):425-8.	Intervention – No detection/classification
351.	James JJ, Giannotti E, Chen Y. Evaluation of a computer-aided detection (CAD)-enhanced 2D synthetic mammogram: comparison with standard synthetic 2D mammograms and conventional 2D digital mammography. Clin Radiol. 2018;73(10):886-92.	Intervention – No detection/classification
352.	Mayo RC, Kent D, Sen LC, Kapoor M, Leung JWT, Watanabe AT. Reduction of False-Positive Markings on Mammograms: a Retrospective Comparison Study Using an Artificial Intelligence-Based CAD. J Digit Imaging. 2019;32(4):618-24.	Intervention – No detection/classification
353.	Patel BC, Sinha GR. Abnormality detection and classification in computer-aided diagnosis (CAD) of breast cancer images. Journal of Medical Imaging and Health Informatics. 2014;4(6):881-885.	Intervention – No detection/classification
354.	Shen R, Yan K, Xiao F, Chang J, Jiang C, Zhou K. Automatic Pectoral Muscle Region Segmentation in Mammograms Using Genetic Algorithm and Morphological Selection. J Digit Imaging. 2018;31(5):680-91.	Intervention – No detection/classification
355.	Sujatha K, Shalini Punithavathani D, Mary Sowbaghya P. Model based non-rigid registration framework for high dynamic range mammography. WSEAS Transactions on Biology and Biomedicine. 2014;11(1):126-32.	Intervention – No detection/classification
Intervention – Not AI (“old” CAD) (n=16)		
356.	Bargallo X, Santamaria G, Del Amo M, Arguis P, Rios J, Grau J, et al. Single reading with computer-aided detection performed by selected radiologists in a breast cancer screening program. Eur J Radiol. 2014;83(11):2019-23.	Intervention – Not AI
357.	Bargallo X, Velasco M, Santamaria G, Del Amo M, Arguis P, Sanchez Gomez S. Role of computer-aided detection in very small screening detected invasive breast cancers. J Digit Imaging. 2013;26(3):572-7.	Intervention – Not AI
358.	Cole EB, Zhang Z, Marques HS, Hendrick RE, Yaffe MJ, Pisano ED. Impact of computer-aided detection systems on radiologist accuracy with digital mammography. American Journal of Roentgenology. 2014;203(4):909-16.	Intervention – Not AI
359.	Fenton JJ, Xing G, Elmore JG, Bang H, Chen SL, Lindfors KK, et al. Short-term outcomes of screening mammography using computer-aided detection a population-based study of medicare enrollees. Ann Intern Med. 2013;158(8):580-7.	Intervention – Not AI

Reference		Main reason for exclusion
360.	Guerriero C, Gillan MG, Cairns J, Wallis MG, Gilbert FJ. Is computer aided detection (CAD) cost effective in screening mammography? A model based on the CADET II study. BMC Health Serv Res. 2011;11:11.	Intervention – Not AI
361.	Hupse R, Samulski M, Lobbes M, den Heeten A, Imhof-Tas MW, Beijerinck D, et al. Standalone computer-aided detection compared to radiologists' performance for the detection of mammographic masses. Eur Radiol. 2013;23(1):93-100.	Intervention – Not AI
362.	Hupse R, Samulski M, Lobbes MB, Mann RM, Mus R, den Heeten GJ, et al. Computer-aided detection of masses at mammography: interactive decision support versus prompts. Radiology. 2013;266(1):123-9.	Intervention – Not AI
363.	Jung NY, Kang BJ, Kim HS, Cha ES, Lee JH, Park CS, et al. Who could benefit the most from using a computer-aided detection system in full-field digital mammography? World J Surg Oncol. 2014;12:168.	Intervention – Not AI
364.	Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL, et al. Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection. JAMA Intern Med. 2015;175(11):1828-37.	Intervention – Not AI
365.	Onega T, Aiello Bowles EJ, Miglioretti DL, Carney PA, Geller BM, Yankaskas BC, et al. Radiologists' perceptions of computer aided detection versus double reading for mammography interpretation. Acad Radiol. 2010;17(10):1217-26.	Intervention – Not AI
366.	Romero C, Varela C, Munoz E, Almenar A, Pinto JM, Botella M. Impact on breast cancer diagnosis in a multidisciplinary unit after the incorporation of mammography digitalization and computer-aided detection systems. AJR Am J Roentgenol. 2011;197(6):1492-7.	Intervention – Not AI
367.	Sato M, Kawai M, Nishino Y, Shibuya D, Ohuchi N, Ishibashi T. Cost-effectiveness analysis for breast cancer screening: Double reading versus single + CAD reading. Breast Cancer. 2014;21(5):532-41.	Intervention – Not AI
368.	Singh S, Maxwell J, Baker JA, Nicholas JL, Lo JY. Computer-aided classification of breast masses: performance and interobserver variability of expert radiologists versus residents. Radiology. 2011;258(1):73-80.	Intervention – Not AI
369.	Skaane P, Kshirsagar A, Hofvind S, Jahr G, Castellino RA. Mammography screening using independent double reading with consensus: is there a potential benefit for computer-aided detection? Acta Radiol. 2012;53(3):241-8.	Intervention – Not AI

Reference		Main reason for exclusion
370.	Sohns C, Angic BC, Sossalla S, Konietschke F, Obenauer S. CAD in full-field digital mammography-influence of reader experience and application of CAD on interpretation of time. Clin Imaging. 2010;34(6):418-24.	Intervention – Not AI
371.	Zheng B, Sumkin JH, Zuley ML, Lederman D, Wang X, Gur D. Computer-aided detection of breast masses depicted on full-field digital mammograms: a performance assessment. Br J Radiol. 2012;85(1014):e153-61.	Intervention – Not AI
Intervention – Prediction of cancer (n=2)		
372.	Chen X, Moschidis E, Taylor C, Astley S. Breast cancer risk analysis based on a novel segmentation framework for digital mammograms. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv. 2014;17(Pt 1):536-43.	Intervention – Prediction of cancer
373.	Timmers JM, Verbeek AL, IntHout J, Pijnappel RM, Broeders MJ, den Heeten GJ. Breast cancer risk prediction model: a nomogram based on common mammographic screening findings. Eur Radiol. 2013;23(9):2413-9.	Intervention – Prediction of cancer
Outcomes – No relevant outcomes (n=24)		
374.	Antropova N, Huynh BQ, Giger ML. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Med Phys. 2017;44(10):5162-71.	No relevant outcomes
375.	Benndorf M. Conditional non-independence of radiographic image features and the derivation of post-test probabilities - A mammography BI-RADS example. Radiography. 2012;18(3):201-5.	No relevant outcomes
376.	Benndorf M, Burnside ES, Herda C, Langer M, Kotter E. External validation of a publicly available computer assisted diagnostic tool for mammographic mass lesions with two high prevalence research datasets. Med Phys. 2015;42(8):4987-96.	No relevant outcomes
377.	Clancy K, Aboutalib S, Mohamed A, Sumkin J, Wu S. Deep Learning Pre-training Strategy for Mammogram Image Classification: an Evaluation Study. J Digit Imaging. 2020;30:30.	No relevant outcomes
378.	Cole EB, Zhang Z, Marques HS, Nishikawa RM, Hendrick RE, Yaffe MJ, et al. Assessing the stand-alone sensitivity of computer-aided detection with cancer cases from the Digital Mammographic Imaging Screening Trial. AJR Am J Roentgenol. 2012;199(3):W392-401.	No relevant outcomes
379.	Li Z, Yu L, Wang X, Yu H, Gao Y, Ren Y, et al. Diagnostic Performance of Mammographic Texture Analysis in the Differential Diagnosis of Benign and Malignant Breast Tumors. Clin Breast Cancer. 2018;18(4):e621-e7.	No relevant outcomes

Reference		Main reason for exclusion
380.	Lobbes M, Smidt M, Keymeulen K, Girometti R, Zuiani C, Beets-Tan R, et al. Malignant lesions on mammography: accuracy of two different computer-aided detection systems. Clin Imaging. 2013;37(2):283-8.	No relevant outcomes
381.	Mayo RC, Leung JWT. Impact of artificial intelligence on women's imaging: Cost-benefit analysis. American Journal of Roentgenology. 2019;212(5):1172-3.	No relevant outcomes
382.	Mendel K, Li H, Sheth D, Giger M. Transfer Learning From Convolutional Neural Networks for Computer-Aided Diagnosis: A Comparison of Digital Breast Tomosynthesis and Full-Field Digital Mammography. Acad Radiol. 2019;26(6):735-43.	No relevant outcomes
383.	Murakami R, Kumita S, Tani H, Yoshida T, Sugizaki K, Kuwako T, et al. Detection of breast cancer with a computer-aided detection applied to full-field digital mammography. J Digit Imaging. 2013;26(4):768-73.	No relevant outcomes
384.	Oliver A, Llado X, Freixenet J, Marti R, Perez E, Pont J, et al. Influence of using manual or automatic breast density information in a mass detection CAD system. Acad Radiol. 2010;17(7):877-83.	No relevant outcomes
385.	Park CS, Jung NY, Kim K, Jung HS, Sohn KM, Oh SJ. Detection of breast cancer in asymptomatic and symptomatic groups using computer-aided detection with full-field digital mammography. Journal of Breast Cancer. 2013;16(3):322-8.	No relevant outcomes
386.	Punitha S, Ravi S, Devi MA, Vaishnavi J. Particle swarm optimized computer aided diagnosis system for classification of breast masses. Journal of Intelligent & Fuzzy Systems. 2017;32(4):2819-28.	No relevant outcomes
387.	Sadaf A, Crystal P, Scaranelo A, Helbich T. Performance of computer-aided detection applied to full-field digital mammography in detection of breast cancers. Eur J Radiol. 2011;77(3):457-61.	No relevant outcomes
388.	Sohns C, Angic B, Sossalla S, Konietschke F, Obenauer S. Computer-assisted diagnosis in full-field digital mammography--results in dependence of readers experiences. Breast J. 2010;16(5):490-7.	No relevant outcomes
389.	Torrents-Barrena J, Puig D, Melendez J, Valls A. Computer-aided diagnosis of breast cancer via Gabor wavelet bank and binary-class SVM in mammographic images. Journal of Experimental & Theoretical Artificial Intelligence. 2016;28(1-2):295-311.	No relevant outcomes
390.	van den Biggelaar FJ, Kessels AG, van Engelshoven JM, Boetes C, Flobbe K. Computer-aided detection in full-field digital mammography in a clinical population: performance of radiologist and technologists. Breast Cancer Res Treat. 2010;120(2):499-506.	No relevant outcomes

Reference		Main reason for exclusion
391.	Vedanarayanan V, Nandhitha NM. Advanced image segmentation techniques for accurate isolation of abnormality to enhance breast cancer detection in digital mammographs. Biomedical Research (India). 2017;28(6):2753-7.	No relevant outcomes
392.	Warren LM, Given-Wilson RM, Wallis MG, Cooke J, Halling-Brown MD, Mackenzie A, et al. The effect of image processing on the detection of cancers in digital mammography. AJR Am J Roentgenol. 2014;203(2):387-93.	No relevant outcomes
393.	Warren LM, Halling-Brown MD, Looney PT, Dance DR, Wallis MG, Given-Wilson RM, et al. Image processing can cause some malignant soft-tissue lesions to be missed in digital mammography images. Clin Radiol. 2017;72(9):799.e1-.e8.	No relevant outcomes
394.	Wu Y, Vanness DJ, Burnside ES. Using multidimensional mutual information to prioritize mammographic features for breast cancer diagnosis. AMIA Annu Symp Proc. 2013;2013:1534-43.	No relevant outcomes
395.	Yang X, Cao A, Song Q, Schaefer G, Su Y. Vicinal support vector classifier using supervised kernel-based clustering. Artif Intell Med. 2014;60(3):189-96.	No relevant outcomes
396.	Yu SD, Liu LL, Wang ZY, Dai GZ, Xie YQ. Transferring deep neural networks for the differentiation of mammographic breast lesions. Science China-Technological Sciences. 2019;62(3):441-7.	No relevant outcomes
397.	Zheng K, Harris C, Bakic P, Makrogiannis S. Spatially localized sparse representations for breast lesion characterization. Computers in Biology and Medicine. 2020;123 (no pagination).	No relevant outcomes
Study type – Systematic reviews (n=7)		
398.	Azavedo E, Zackrisson S, Mejare I, Heibert Arnlind M. Is single reading with computer-aided detection (CAD) as good as double reading in mammography screening? A systematic review. BMC med. 2012;12:22.	Study type – Systematic reviews with no relevant outcomes
399.	Eadie LH, Taylor P, Gibson AP. A systematic review of computer-assisted diagnosis in diagnostic cancer imaging. Eur J Radiol. 2012;81(1):e70-6.	Study type – Systematic reviews with no relevant outcomes
400.	Gruppo di studio G-S, Chersevani R, Ciatto S, Del Favero C, Frigerio A, Giordano L, et al. "CADEAT": considerations on the use of CAD (computer-aided diagnosis) in mammography. Radiol Med (Torino). 2010;115(4):563-70.	Study type – Systematic reviews with no relevant outcomes

Reference		Main reason for exclusion
401.	Henriksen EL, Carlsen JF, Vejborg IM, Nielsen MB, Lauridsen CA. The efficacy of using computer-aided detection (CAD) for detection of breast cancer in mammography screening: a systematic review. Acta Radiol. 2019;60(1):13-8.	Study type – Systematic reviews with no relevant outcomes
402.	Houssami N, Kirkpatrick-Jones G, Noguchi N, Lee CI. Artificial Intelligence (AI) for the early detection of breast cancer: a scoping review to assess AI's potential in breast screening practice. Expert Rev Med Devices. 2019;16(5):351-62.	Study type – Systematic reviews with no relevant outcomes
403.	Sadoughi F, Kazemy Z, Hamedan F, Owji L, Rahmanikatigari M, Azadboni TT. Artificial intelligence methods for the diagnosis of breast cancer by image processing: a review. Breast Cancer (Dove Med Press). 2018;10:219-30.	Study type – Systematic reviews with no relevant outcomes
404.	Yassin NIR, Omran S, El Houby EMF, Allam H. Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review. Comput Methods Programs Biomed. 2018;156:25-45.	Study type – Systematic reviews with no relevant outcomes
Full text not available via Document Supply (n=7)		
405.	Bhavani SR, Chilambuchelvan A, Senthilkumar J, Manjula D, Krishnamoorthy R, Kannan A. A secure cloud-based multi-agent intelligent system for mammogram image diagnosis. International Journal of Biomedical Engineering and Technology. 2018;28(2):185-202.	Document Supply cancelled request: no location found.
406.	Grout S, Dheeraj Suryaa SR, Hitesh, Venkatesan DH, Sumanth S, Vishnu Vardhan Reddy M. Anomaly detection in digital mammography using neural networks. Journal of International Pharmaceutical Research. 2019;46(3):750-4.	Document Supply cancelled request: no location found.
407.	Saraswathi D, Srinivasan E. An ensemble approach to diagnose breast cancer using fully complex-valued relaxation neural network classifier. International Journal of Biomedical Engineering and Technology. 2014;15(3):243-60.	Document Supply cancelled request: no location found.
408.	Selvan VP, Suganthi M. Clinical support system for classification of tumor in mammogram images using multiple features and neural network classifier. Journal of Pure and Applied Microbiology. 2015;9(Special Edition):253-61.	Document Supply cancelled request: no location found.
409.	Singh B, Jain VK, Singh S. Mammogram mass classification using support vector machine with texture, shape features and hierarchical centroid method. Journal of Medical Imaging and Health Informatics. 2014;4(5):687-96.	Document Supply cancelled request: no location found.
410.	Srivastava S, Sharma N, Singh SK, Srivastava R. Quantitative analysis of a general framework of a CAD tool for breast cancer detection from mammograms. Journal of Medical Imaging and Health Informatics. 2014;4(5):654-74.	Document Supply cancelled request: no location found.

Reference		Main reason for exclusion
411.	Zhou L, Ding M, Xu L, Zhou Y, Zhang X. The automatic segmentation of mammographic mass using the end-to-end convolutional network based on dense-prediction. Journal of Medical Imaging and Health Informatics. 2019;9(7):1429-34.	Document Supply cancelled request: no location found.
Other reasons (n=5)		
412.	Becker AS, Marcon M, Ghafoor S, et al. Deep Learning in Mammography: Diagnostic Accuracy of a Multipurpose Image Analysis Software in the Detection of Breast Cancer. Invest Radiol 2017;52(7):434-40. doi: https://dx.doi.org/10.1097/RLI.0000000000000358	Study 1: BCDR database; unclear proportion of screening mammograms Study 2: Temporal validation
413.	da Silva R, de Carvalho A. Automatic classification of breast lesions usingTransfer Learning. Ieee Latin America Transactions. 2019;17(12):1964-9.	Language – Not available in English
414.	Kim HE, Kim HH, Han BK, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. The Lancet Digital Health 2020;2(3):e138-e48. doi: http://dx.doi.org/10.1016/S2589-7500%2820%2930003-0	Evaluation study: unclear proportion of screening mammograms. Reader study: In parts temporal validation confirmed by corresponding author via email.
415.	Polat K. Application of Attribute Weighting Method Based on Clustering Centers to Discrimination of Linearly Non-Separable Medical Datasets. J Med Syst. 2012;36(4):2657-73.	Separation of two different mage datasets (liver and breast)
416.	Sechopoulos I, Mann RM. Stand-alone artificial intelligence - The future of breast cancer screening? Breast. 2020;49:254-60.	Narrative review

Publications excluded after review of full-text articles – Update database searches

Publications

Reference	Main reason for exclusion
Population – Image type (e.g. digitised film images; not FFDM images) (n=18)	
1. Abubacker NF, Hashem IAT, Hui LK. Mammographic Classification Using Stacked Ensemble Learning with Bagging and Boosting Techniques. Journal of Medical and Biological Engineering 2020;40(6):908-16. doi: http://dx.doi.org/10.1007/s40846-020-00567-y	Population – Image type
2. Arora R, Rai PK, Raman B. Deep feature-based automatic classification of mammograms. Medical & biological engineering & computing 2020;58(6):1199-211. doi: https://dx.doi.org/10.1007/s11517-020-02150-8	Population – Image type
3. Bakthavachalam MD, Albert Antony Raj S. A study on breast cancer analysis by using k-nearest neighbor with different distances and classification rules using machine learning. European Journal of Molecular and Clinical Medicine 2020;7(3):4842-51.	Population – Image type
4. Gautam N, Singh A, Kumar K, et al. Investigation on performance analysis of support vector machine for classification of abnormal regions in medical image. Journal of Ambient Intelligence and Humanized Computing doi: 10.1007/s12652-021-02965-9	Population – Image type
5. Graewingholt A, Duffy S. Retrospective comparison between single reading plus an artificial intelligence algorithm and two-view digital tomosynthesis with double reading in breast screening. Journal of medical screening 2021:969141320984198. doi: https://dx.doi.org/10.1177/0969141320984198	Population – Image type
6. Jahangeer GSB, Rajkumar TD. Early detection of breast cancer using hybrid of series network and VGG-16. Multimedia Tools and Applications 2021;80(5):7853-86. doi: 10.1007/s11042-020-09914-2	Population – Image type
7. Kakileti ST, Madhu HJ, Krishnan L, et al. Observational Study to Evaluate the Clinical Efficacy of Thermalytix for Detecting Breast Cancer in Symptomatic and Asymptomatic Women. JCO global oncology 2020;6:1472-80. doi: https://dx.doi.org/10.1200/GO.20.00168	Population – Image type
8. Ketabi H, Ekhlasi A, Ahmadi H. A computer-aided approach for automatic detection of breast masses in digital mammogram via spectral clustering and support vector machine. Physical and Engineering Sciences in Medicine 2021;44(1):277-90. doi: http://dx.doi.org/10.1007/s13246-021-00977-5	Population – Image type

Reference		Main reason for exclusion
9.	Khamparia A, Bharati S, Podder P, et al. Diagnosis of breast cancer based on modern mammography using hybrid transfer learning. Multidimensional systems and signal processing 2021:1-19. doi: https://dx.doi.org/10.1007/s11045-020-00756-7	Population – Image type
10.	Melekoodappattu JG, Kadan AB, Anoop V. Early detection of breast malignancy using wavelet features and optimized classifier. International Journal of Imaging Systems and Technology doi: 10.1002/ima.22537	Population – Image type
11.	Melekoodappattu JG, Subbian PS. Automated breast cancer detection using hybrid extreme learning machine classifier. Journal of Ambient Intelligence and Humanized Computing doi: 10.1007/s12652-020-02359-3	Population – Image type
12.	Melekoodappattu JG, Subbian PS, Queen MPF. Detection and classification of breast cancer from digital mammograms using hybrid extreme learning machine classifier. International Journal of Imaging Systems and Technology 2021;31(2):909-20. doi: 10.1002/ima.22484	Population – Image type
13.	Rao PMM, Singh SK, Khamparia A, et al. Multi-class Breast Cancer Classification using Ensemble of Pretrained models and Transfer Learning. Current medical imaging 2021 doi: https://dx.doi.org/10.2174/1573405617666210218101418	Population – Image type
14.	Shaikh TA, Ali R. An intelligent healthcare system for optimized breast cancer diagnosis using harmony search and simulated annealing (HS-SA) algorithm. Informatics in Medicine Unlocked 2020;21:100408. doi: http://dx.doi.org/10.1016/j.imu.2020.100408	Population – Image type
15.	Solanki YS, Chakrabarti P, Jasinski M, et al. A Hybrid Supervised Machine Learning Classifier System for Breast Cancer Prognosis Using Feature Selection and Data Imbalance Handling Approaches. Electronics 2021;10(6) doi: 10.3390/electronics10060699	Population – Image type
16.	Thawkar S. A hybrid model using teaching-learning-based optimization and Salp swarm algorithm for feature selection and classification in digital mammography. Journal of Ambient Intelligence and Humanized Computing doi: 10.1007/s12652-020-02662-z	Population – Image type
17.	Thawkar S, Ingolikar R. Classification of masses in digital mammograms using Biogeography-based optimization technique. Journal of King Saud University-Computer and Information Sciences 2020;32(10):1140-48. doi: 10.1016/j.jksuci.2018.01.004	Population – Image type
18.	Xiao M, Zhao C, Li J, et al. Diagnostic Value of Breast Lesions Between Deep Learning-Based Computer-Aided Diagnosis System and Experienced Radiologists: Comparison the	Population – Image type

Reference		Main reason for exclusion
	Performance Between Symptomatic and Asymptomatic Patients. Frontiers in Oncology 2020;10:1070. doi: http://dx.doi.org/10.3389/fonc.2020.01070	
Population – Subpolulation (e.g. only cancer cases) (n=3)		
19.	Graewingholt A, Rossi PG. Retrospective analysis of the effect on interval cancer rate of adding an artificial intelligence algorithm to the reading process for two-dimensional full-field digital mammography. Journal of medical screening 2021:969141320988049. doi: https://dx.doi.org/10.1177/0969141320988049	Population - Subpopulation
20.	Lang K, Hofvind S, Rodriguez-Ruiz A, et al. Can artificial intelligence reduce the interval cancer rate in mammography screening? European radiology 2021 doi: https://dx.doi.org/10.1007/s00330-021-07686-3	Population - Subpopulation
21.	Lee SE, Han K, Kim E-K. Application of artificial intelligence-based computer-assisted diagnosis on synthetic mammograms from breast tomosynthesis: comparison with digital mammograms. European radiology 2021 doi: https://dx.doi.org/10.1007/s00330-021-07796-y	Population - Subpopulation
Population – <90% screening mammograms or unclear proportion (n=3)		
22.	Ashiba HI. A proposed framework for diagnosis and breast cancer detection. Multimedia Tools and Applications 2021;80(6):9333-69. doi: 10.1007/s11042-020-10131-0	Population – <90% screening mammograms or unclear proportion
23.	Cui Y, Li Y, Xing D, et al. Improving the Prediction of Benign or Malignant Breast Masses Using a Combination of Image Biomarkers and Clinical Parameters. Frontiers in oncology 2021;11:629321. doi: https://dx.doi.org/10.3389/fonc.2021.629321	Population – <90% screening mammograms or unclear proportion
24.	Kalyani K. CNN analysis for mammogramdisease detection. European Journal of Molecular and Clinical Medicine 2020;7(9):1540-43.	Population – <90% screening mammograms or unclear proportion
Internal validation – Cross validation (n=5)		
25.	Africano G, Arponen O, Sassi A, et al. A Comparison of Regions of Interest in Parenchymal Analysis for Breast Cancer Risk Assessment. Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference 2020;2020:1136-39. doi: https://dx.doi.org/10.1109/EMBC44109.2020.9176200	Internal validation – Cross validation
26.	Hassan SA, Sayed MS, Abdalla MI, et al. Breast cancer masses classification using deep convolutional neural networks and transfer learning. Multimedia Tools and Applications 2020;79(41-42):30735-68. doi: 10.1007/s11042-020-09518-w	Internal validation – Cross validation

Reference	Main reason for exclusion
27. Heidari M, Lakshmivarahan S, Mirniaharikandehei S, et al. Applying a random projection algorithm to optimize machine learning model for breast lesion classification. IEEE Transactions on Biomedical Engineering 2021 doi: http://dx.doi.org/10.1109/TBME.2021.3054248	Internal validation – Cross validation
28. Hsu CH, Chen X, Lin WW, et al. Effective multiple cancer disease diagnosis frameworks for improved healthcare using machine learning. Measurement 2021;175 doi: 10.1016/j.measurement.2021.109145	Internal validation – Cross validation
29. Jiang MK, Han L, Sun H, et al. Cross-modality image feature fusion diagnosis in breast cancer. Physics in Medicine and Biology 2021;66(10) doi: 10.1088/1361-6560/abf38b	Internal validation – Cross validation
Internal validation – Split sample (n=4)	
30. Gnanasekaran VS, Joypaul S, Sundaram PM, et al. Deep learning algorithm for breast masses classification in mammograms. Iet Image Processing 2020;14(12):2860-68. doi: 10.1049/iet-ipr.2020.0070	Internal validation – Split sample
31. Guan Y, Wang X, Li H, et al. Detecting Asymmetric Patterns and Localizing Cancers on Mammograms. Patterns (New York, NY) 2020;1(7) doi: https://dx.doi.org/10.1016/j.patter.2020.100106	Internal validation – Split sample
32. Shen Y, Wu N, Phang J, et al. An interpretable classifier for high-resolution breast cancer screening images utilizing weakly supervised localization. Medical image analysis 2021;68:101908. doi: https://dx.doi.org/10.1016/j.media.2020.101908	Internal validation – Split sample
33. Suh YJ, Jung J, Cho B-J. Automated Breast Cancer Detection in Digital Mammograms of Various Densities via Deep Learning. Journal of personalized medicine 2020;10(4) doi: https://dx.doi.org/10.3390/jpm10040211	Internal validation – Split sample
Intervention – Detecting subtypes (n=1)	
34. Vedalankar AV, Gupta SS, Manthalkar RR. Addressing architectural distortion in mammogram using AlexNet and support vector machine. Informatics in Medicine Unlocked 2021;23:100551. doi: http://dx.doi.org/10.1016/j.imu.2021.100551	Intervention – Detecting subtypes
Outcomes – No relevant outcomes (n=1)	
35. Nct. Mammography Screening With Artificial Intelligence (MASAI). https://clinicaltrialsgov/show/NCT04838756 2021	Registered study protocol, no relevant outcomes
Full text not available via Document Supply (n=2)	

Reference	Main reason for exclusion
36. Sathya Priya T, Ramaprabha T. Deep learning based image segmentation with alexnet feature extraction for classification of mammogram images. International Journal of Pharmaceutical Research 2021;13(1):4995-5009. doi: http://dx.doi.org/10.31838/ijpr/2021.13.01.690	Document Supply cancelled request: no location found.
37. Yi XC, Hou J. Segmentation of Medical Image Based on Superpixel Boundary Perceptual Convolutional Network in Cancer Diagnosis and Treatment. Journal of Medical Imaging and Health Informatics 2021;11(1):254-60. doi: 10.1166/jmihi.2021.3425	Document Supply cancelled request: no location found.
Already picked up by original searches (n=2)	
38. Lindholm P, Eklund M, Dembrower K, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. The Lancet Digital Health 2020;2(9):e468-e74. doi: http://dx.doi.org/10.1016/S2589-7500%2820%2930185-0	Same article as Dembrower (2020), already picked up by original search
39. Pacile S, Lopez J, Chone P, et al. Improving Breast Cancer Detection Accuracy of Mammography with the Concurrent Use of an Artificial Intelligence Tool. Radiology Artificial intelligence 2020;2(6):e190208. doi: https://dx.doi.org/10.1148/ryai.2020190208	Already picked up by original search

Sub-studies or datasets of included studies

Reference	Excluded study / dataset and reason
Balta 2020[25]	None
Dembrower 2020[26]	None
Lang 2020[27]	None
McKinney 2020[29]	1) **Retrospective clinical comparison** with original decisions of UK and US readers, respectively, excluded due to internal validation test set (split sample). 2) **Comparison with reader study** excluded due to internal validation test set (split sample). 3) **Simulation study** excluded as it is based on test accuracy estimated obtained using internal validation test sets (split sample).
Pacilè 2020[30]	None
Rodriguez-Ruiz 2018[34]	Excluded **AI as stand-alone reader** due to lack of outcomes such as sensitivity and specificity (only AUC).
Rodriguez-Ruiz 2019[32]	Excluded **AI as stand-alone reader** due to lack of outcomes such as sensitivity and specificity (only AUC).
Rodriguez-Ruiz 2019[33]	Excluded **data sets A and D-H** as <90% screening mammograms or unclear proportion of screening mammograms. Excluded **data set B** as no relevant outcomes reported.
Salim 2020[35]	None
Schaffter 2020[36]	Excluded the **Kaiser Permanente Washington (KPW) dataset** as it was used for training and evaluation (split sample).
Watanabe 2019[37]	None
Lotter 2021[28]	The paper reports on 5 test sets; only one data set (Site D) is an external test set of screening FFDMs. Two test sets are excluded as they are also used for training, one test set uses DBT (not FFDM) images, and one test set uses diagnostic (not screening) FFDMs. The study of pre-index cancers was excluded as the analysis included a sub-population of cancers with a negative previous screening result, therefore the analysis was based on images of subpopulations by screening outcome.
Raya-Povedano 2021[31]	Simulated autonomous AI triaging strategy excluded as simulation study.

Appendix 4 Additional baseline characteristics of included studies

STUDY	Method of enrichment	Confirmed cancer (prevalence) n (%)	Cancer type n (%)	Cancer size/grade (invasive only)	Breast density n (%)
Lotter 2021[28]	Matched case-control study: Cancer cases: all patients from a single health system in Massachusetts with qualifying index (screening mammograms interpreted as suspicious and confirmed to be malignant by pathology within three months) and pre-index exams (from the same set of women as the index exams: screening exams interpreted as BI-RADS 1 or 2 12–24 months prior to the index exams) over the specified time period using a local cancer registry. Non-cancer cases: selected from a single health system in Massachusetts to have a similar distribution in patient age and breast density compared with the cancer cases using bucketing (negative exam followed by an additional BI-RADS 1 or 2 interpretation at the next screening exam 9–39 months later). BI-RADS >2: 131 (46%) BI-RADS 1 or 2: 154 (54%)	131 (46.0%)	ILC or IDC: 88 (67.2%) DCIS: 38 (29.0%) Other: 5 (3.8%) Lesion type Soft tissue 87 Calcifications 53 (adds up to 140 though)	For all 131 cancers: 0-1cm: 45 (34.35%) 1-2cm: 27 (20.6%) 2-5cm: 11 (8.4%) >5cm: 3 (2.3%) Unknown: 45 (34.35%)	Non-cancer (n=154): Non-dense (A&B) 96 (62.3%) Dense (C&D) 58 (37.7%) Cancer (n=131): Non-dense (A&B) 81 (61.8%) Dense (C&D) 50 (38.2%)

STUDY	Method of enrichment	Confirmed cancer (prevalence) n (%)	Cancer type n (%)	Cancer size/grade (invasive only)	Breast density n (%)
McKinney 2020[29]	Images from all women at one US academic medical centre who were biopsied during this time period and a random subset of women (~5%) who never underwent biopsy. BI-RADS 0, 4 or 5: 929 (30%) BI-RADS 1,2 or 3: 1,809 (58%) No Bi-RADS score: 359 (12%)	686 (22.2%)	Data for 553/686 with BI-RADS score: Invasive 364/553 (65.8%) DCIS 163/553 (29.5%) Other 26/553 (4.7%)	NR	NR
Rodriguez-Ruiz 2019[33]	Mammograms collected on cancer outcome from Dutch digital screening pilot project: 80 biopsy-proved cancer cases and 120 negative cases. Case selection: 1) cases in which the lesion was rated as obvious, cases with only microcalcifications, and cases in which not all four cranial-caudal and mediolateral oblique views of both breasts were available were excluded. 2) Prior screening mammograms in which a malignant lesion was already visible n=17 mammograms. 3) From the remaining cases, random selection of 63 screen-detected cancer cases from incident screening rounds. Negative case selection: 1) 20 false-positive cases that were verified by normal follow-up (no biopsy), 2) random selection of 100 non-referred mammograms with at least one normal follow-up screen.	79 (39.7%)	NR	NR	NR

STUDY	Method of enrichment	Confirmed cancer (prevalence) n (%)	Cancer type n (%)	Cancer size/grade (invasive only)	Breast density n (%)
Salim 2020[35]	All women with a diagnosis of breast cancer from the Swedish Cohort of screen-age women were included and a random sample of healthy women.	739 (8.4%)	Invasive 640 (86%) IDC 514/640 (80%) ILC 82/640 (13%) other 43/640 (7%) missing 1/640 (0.15%) In situ 85 (12%) Missing information 14 (2%)	Median 15mm (IQR 10-21mm)	Mammographic percent density, % Median 21.9 IQR 13.8-32.1
Schaffter 2020[36]	No enrichment: consecutive sample of screened women from 1 Swedish centre	780 (1.1%)	Invasive 681 (87.3%) DCIS 99 (12.7%)	NR	NR
Balta 2020[25]	No enrichment: consecutive sample of screened women from 1 German centre	114 (0.64%)	NR	NR	NR
Dembrower 2020[26]	All women diagnosed with breast cancer who attended two consecutive screening rounds from the Swedish Cohort of screen-age women were include. Healthy women were randomly sampled from the same cohort.	547 (7.4%)	NR	NR	NR
Lang 2020[27]	No enrichment: Consecutive sub cohort of the prospective population-based Malmö Breast Tomosynthesis Screening Trial in which every third woman who was invited to attend regular screening was invited to participate (random sample)	68 (0.71%)	IDC 33/68 (48.5%) ILC 11/68 (16.2%) ITC 10/68 (14.7%) DCIS 11/68 (16.2%) Other (e.g. papillary carcinoma, apocrine tumour) 3/68 (4.4%)	Grade 1 24/56 (42.9%) Grade 2 25/56 (44.6%) Grade 3 7/56 (12.5%)	NR

STUDY	Method of enrichment	Confirmed cancer (prevalence) n (%)	Cancer type n (%)	Cancer size/grade (invasive only)	Breast density n (%)
Rava-Povedano 2021[31]	No enrichment: Mammograms from Spanish tomosynthesis screening trial Original outcomes: Normal readings (with two-year follow-up): 14,795 (92.5%) FP recalls: 1,078 (6.7%) Screen detected cancers: 98 (0.6%) Interval cancers: 15 (0.1%)	113 (0.7%)	Cancer type: Mass 67 (59.3%) Architectural distortion 21 (18.6%) Asymmetry 4 (3.5%) Calcification 21 (18.6%) Histologic type: IDC 80 (70.8%) ILC 5 (4.4%) Other invasive 1 (0.9%) DCIS 27 (23.9%)	For all 113 cancers: Grade 1 49 (43.4%) Grade 2 40 (35.4%) Grade 3 24 (21.2%)	Category A 3,648/15,986 (22.8%) Category B 8,153/15,986 (51.0%) Category C 3,749/15,986 (23.5%) Category D 436/15,986 (2.7%)
Pacilè 2020[30]	Enrichment method not reported. The final dataset included 80 true-positive, 40 false-negative. Data underwent a quality check performed by an experienced breast radiologist to exclude examinations not meeting acquisition standards or presenting identifiable features (e.g., nipple retraction, invasive cancer larger than approximately 2.5 cm, bilateral cancer, and others to minimize recall bias), and for false-negative examinations, that malignant lesions were visible and identifiable in retrospect.	120 (50%)	Histologic type: IDC 75 (62.5%) DCIS 27 (22.5%) ILC 6 (5.0%) Other 12 (10.0%) Lesion type: Mass 64 (53.3%) Calcification 30 (25.0%) Asymmetry 13 (10.8%) Architectural distortion 13 (10.8%)	NR	Category A 15.00% (36/240) Category B 43.75% (105/240) Category C 34.58% (83/240) Category D 6.67% (16/240)

STUDY	Method of enrichment	Confirmed cancer (prevalence) n (%)	Cancer type n (%)	Cancer size/grade (invasive only)	Breast density n (%)
Rodriguez-Ruiz 2019[32]	Examinations from cancer, false-positive and normal cases were consecutively collected from one US and one European centre until predefined distribution of selection was achieved. Mammograms were reviewed by one radiologist to ensure image quality 9 were excluded (3 for poor image quality, 3 without link to case report findings and 3 with obvious signs of cancer)	100 (41.7%)	Lesion type: Mass 49 (49%) Calcifications 30 (30%) Asymmetry 10 (10%) Architectural distortion 6 (6%) Both calcifications and mass lesions 5 (5%) Histologic type: IDC 64 (64%) DCIS 13 (13%) ILC 18 (18%) Invasive tubular carcinoma 6 (6%) Other 3 (3%) (4 examinations showed 2 histologic cancer types)	Median 13 mm^2 IQR 4-22 mm^2	BI-RADS breast density A 28 (12%) B 133 (55%) C 64 (27%) D 15 (6%)
Watanabe 2019[37]	Mammograms were selected from an archive of false negative mammograms (dataset from a community healthcare facility in Southern California). mammograms were originally interpreted by community-based radiologists using the R2 ImageChecker CAD v10.0	90 (73.8%)	Mass 50 (55.6%) Microcalcifications 16 (17.8%) Mass and Microcalcifications 9 (10.0%) Architectural Distortions 5 (5.6%) Mass and Architectural Distortions 4 (4.4%) Asymmetry 3 (3.3%) Architectural Distortion and Microcalcifications 1 (1.1%) Microcalcifications and Asymmetry 1 (1.1%) Focal Asymmetry 1 (1.1%)	NR	Fatty 4 (4%) Scattered 43 (48%) Heterogeneously dense 37 (41%) Extremely dense 6 (7%)

References for Appendices 3 and 4

25. Balta C, Rodriguez-Ruiz A, Mieskes C, et al. Going from double to single reading for screening exams labeled as likely normal by AI: what is the impact? *Proceedings of SPIE* 2020;11513:115130D.
26. Dembrower K, Wahlin E, Liu Y, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. *Lancet Digital Health* 2020;2(9):E468-E74.
27. Lang K, Dustler M, Dahlblom V, et al. Identifying normal mammograms in a large screening population using artificial intelligence. *European Radiology* 2020;02:02. doi: https://dx.doi.org/10.1007/s00330-020-07165-1
28. Lotter W, Diab AR, Haslam B, et al. Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. *Nature medicine* 2021;27(2):244-49. doi: https://dx.doi.org/10.1038/s41591-020-01174-9
29. McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. *Nature* 2020;577(7788):89-94. doi: https://doi.org/10.1038/s41586-019-1799-6)
30. Pacilè S, Lopez J, Chone P, et al. Improving Breast Cancer Detection Accuracy of Mammography with the Concurrent Use of an Artificial Intelligence Tool. *Radiology: Artificial Intelligence* 2020;2(6):e190208. doi: 10.1148/ryai.2020190208
31. Raya-Povedano JL, Romero-Martin S, Elias-Cabot E, et al. AI-based Strategies to Reduce Workload in Breast Cancer Screening with Mammography and Tomosynthesis: A Retrospective Evaluation. *Radiology* 2021:203555. doi: https://dx.doi.org/10.1148/radiol.2021203555
32. Rodriguez-Ruiz A, Krupinski E, Mordang JJ, et al. Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. *Radiology* 2019;290(2):305-14. doi: https://dx.doi.org/10.1148/radiol.2018181371
33. Rodriguez-Ruiz A, Lang K, Gubern-Merida A, et al. Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. *J Natl Cancer Inst* 2019;111(9):916-22. doi: https://dx.doi.org/10.1093/jnci/djy222
34. Rodriguez-Ruiz A, Mordang JJ, Karssemeijer N, et al. Can radiologists improve their breast cancer detection in mammography when using a deep learning based computer system as decision support? *Proceedings of SPIE* 2018;10718:1071803. doi: 10.1117/12.2317937
35. Salim M, Wahlin E, Dembrower K, et al. External Evaluation of 3 Commercial Artificial Intelligence Algorithms for Independent Assessment of Screening Mammograms. *JAMA Oncol* 2020;27:27. doi: https://dx.doi.org/10.1001/jamaoncol.2020.3321
36. Schaffter T, Buist DSM, Lee CI, et al. Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. *JAMA netw* 2020;3(3):e200265. doi: https://dx.doi.org/10.1001/jamanetworkopen.2020.0265
37. Watanabe AT, Lim V, Vu HX, et al. Improved Cancer Detection Using Artificial Intelligence: a Retrospective Evaluation of Missed Cancers on Mammography. *J Digit Imaging* 2019;32(4):625-37. doi: https://dx.doi.org/10.1007/s10278-019-00192-5