Supplementary Table 1: additional information on MRI scanners used in the included studies. NR = not reported

Author	Journal	Year	MRI Field Strength (Tesla)	Manufacturer	Sequence description	Coil description
Eom	AJR	2016	1.5	Siemens	The standard dynamic protocol used for the evaluation of supraclavicular and axillary lymph nodes included the following sequences and parameters: an axial 2D T2-weighted STIR turbo spin- echo pulse sequence (TR/TE/time interval, 6700/74/150; FOV, 300 × 300 mm2; matrix, 448 × 448; and slice thickness, 5 mm), axial unenhanced and contrastenhanced fat-saturated 3D T1-weighted FLASH volume-interpolated breathhold examination pulse sequences (TR/TE-5/2/41-FOV/240 mr2; matrix, 240 m	dedicated bilateral phase array coil
Andrad	Cancer	2017	NR	NR	NR	NR
Ramsh orst, M.S.	Breast Cancer Res Treat	2017	1.5 and 3	1.5T Siemens 3T Phil	First, an unenhanced coronal 3D fast field echo (FFE) (thrive) sense T1- weighted sequence was performed. Subsequently, contrast containing gadolinium (0.1 mmol/kg) was administered intravenously followed by dynamic imaging in five consecutive series at 90-s intervals with voxel size 1.21 9 1.21 9 1.69 mm3 (1.5 T) or 1.1 9 1.1 9 1.2 mm3 (3.0 T).	dedicated bilateral phased array breast coil, philips: dedicated 7- element sense breast coil was
Iwase, M.	Breast Cancer	2018	3.0	GE healthcare Signa	pre-contrast and dynamic contrastenhanced imaging. Dynamic study was performed before and 45, 180, and 315 s after starting intravenous gadolinium injection.	breast array coils
Marin Alcala M.	Ann. Oncol.	2018	NR	NR	NR	NR
Murphy , C.	Intern Med J	2018	3T (except for 3 studies at 1.5)	Siemens	a fat-suppressed sequence, a non-fat-suppressed T2 sequence, the dynamic rapid sequence fat-suppressed T1 sequence and a high- resolution fat-suppressed T1 sequence	16-channel AL- dedicated breast coil
Namur a, M.	Clin Breast Cancer	2018	3.0	GE Health Care, Siemens	high-resolution precontrast imaging and dynamic contrast-enhanced imaging. It included a localizing sequence followed by unilateral T2- weighted axial imaging (repetition time [TR]/echo time [TE]: 4748/86, slice thickness ¼ 5 mm, field of view [FoV] ¼ 350 mm, matrix ¼ 512 256). After a scout scan, sagittal, unilateral T1-weighted, precontrast images (TR/TE: 5.2/2.2 slice thickness ¼ 1 mm, slice gap ¼ 0 mm, FoV ¼ 350 mm, matrix ¼ 384 330) were acquired. A dynamic study in the axial plane was performed before and 60, 180, 300, and 420 seconds after starting intravenous injection of 1.5 mL/kg of meglumine gadopentetate at the beginning of the fifth acquisition, followed by 10 mL saline for flushing.	dedicated 8- channel (GE Health Care) and 32- channel (Siemens Healthineers) breast array coils
Gasol Cudos A.	Ann. Oncol.	2019	1.5	NR	NR	NR
Gampe nrieder, S.P,	Breast Cancer Res	2019	3.0 or 1.5, 11 scans on <1.5T	Philips	axial T2-weighted fat-suppressed images (echo time (TE)/repetition time (TR)/ inversion recovery (IR) of 60/9065/230 msec; with a slice thickness of 3mm and a field of view of 30–40 cm) and axial diffusion- weighted images (TE IR/TR of 59/8157 msec; b-values up to 600; slice thickness 3 mm). Finally, four axial T1-weighted fat-suppressed dynamic acquisitions (TE/TR, 2.3/4.1msec with a slice thickness of 1mm) were registered over the duration of 5 min after intravenous contrast medium injection. As gadolinium contrast agent we used either 15 mL gadoteric acid (DOTAREM®, Guerbet, Cedex, France) or 7mL gadobutrol (GADOVIST®, Bayer Schering Pharma, Berlin, Germany followed by a saline bolus flash. The remaining 11 breast MR exams were performed in external institutes using scanners of less than 1.5 Tesla.	dedicated phased- array breast coil
Negrao	Radiol Bras	2019	1.5	Signa GE or Achieva Philips	dedicated breast coil	
Zhang, X.	Quant Imagin g Med Surg	2020	1.5	GE Healthcare	First, sagittal fat-saturated T2-weighted, fast-spin echo images of the bilateral breasts were recorded using the following parameters: repetition time/echo time, 4,040 ms/81 ms; echo train length, 19; slice thickness, 5 mm; slice gap, 1 mm; field of view, 220 mm; matrix size, 320×224; number of excitations, 2; imaging time, 1 min 41 s. Then, unenhanced and enhanced sequences of axial fat-saturated gradient recalled echo T1-weighted images were acquired with the following parameters: repetition time/echo time/inversion time, 6.1 ms/2.9 ms/13 ms; flip angle, 10°; slice thickness, 3.2 mm; slice gap, 0 mm;	resolute 8-channel high-definition breast coil

					field of view, 360 mm; matrix size, 350×350; number of excitations, 0.8; imaging time, 58 s per phase; 8 phases of enhancement. Gadopentetate dimeglumine was administered at a dose of 15 mL and a rate of 2.5 mL/s, followed by a 20-mL saline flush given with an automatic injector.	
De Los Santos, J.	Clin Breast Cancer	2011	1.5	GE	T1 axial, T2 fat-sat axial, pre- and post-contrast sagittal dynamic sequence with four repeats at 90-second intervals, 2- to 3-mm slice thickness, and delayed post-contrast T1 fat-sat axial, color map analysis, and maximum intensity projection	8-channel In Vivo dedicated breast coil
De Los Santos, J. F.	Cancer	2013	NR	NR	NR	NR
Hayashi , Y.	Oncol Lett	2013	1.5	Philips	Transverse images were obtained by diffusion-weighted (DW) imaging. Coronal images were obtained by contrast-enhanced dynamic imaging. Sagittal images were obtained by contrast- enhanced late-phase imaging. Additionally, depending on the case, sagittal images were obtained by T2-weighted fat-suppressed imaging prior to infusion of contrast material.	SENSE body coil
Sabade II M.D.	Eur. J. Cancer	2014	NR	NR	NR	NR
Kim, M. J.	Acta Radiol	2015	3.0	Trio Tim Siemens	T1-weighted (T1W) non-fat-suppressed axial sequence (repetition time/echo time, 280/2.6 ms; flip angle, 65; bandwidth, 543 Hz/pixel; matrix size, 512512; voxel size, 0.70.73.0 mm), T1W non-fat-suppressed precontrast and 2D dynamic contrast-enhanced (DCE) axial (repetition time/echo time, 280/2.6 ms; flip angle, 65; bandwidth 540 Hz/pixel; matrix size, 512343; voxel size, 1.00.73.0 mm) with intravenous injection of 0.2 cc/kg gadolinium-diethylenetriaminepenta acetic acid (Gd-DTPA, Magnevist; Berlex Laboratories Inc., Montville, NJ, USA) in order to obtain one precontrast set and six postcontrast sets, the temporal resolution was 73 s for each frame, and finally T2-weighted (T2W) turbo spin echo axial images (repetition time/ echo time, 4360/82 ms; flip angle, 150; bandwidth, 305 Hz/pixel; matrix size, 512512; voxel size, 0.70.73.0 mm). The slice thickness was 3mm and field of view was 32–34cm for all of the MRI sequences.	edicated fourchannel breast array coil.
Fukuda, T.	Springe rplus	2016	1.5	GE	Our imaging protocol included a localizing sequence followed by unilateral fast spin-echo T2-weighted coronal imaging (TR/TE, 4800/85 ms; echo train length 16, and matrix 384 × 224) with fat suppression by chemical shiftselective imaging sequences. Other parameters were as follows: field of view, 260 mm; section thickness, 3 mm and interslice gap, 0 mm. This examination was followed by combined dynamic contrast-enhanced unilateral coronal breast imaging. An enhanced T1-weighted examination 3D gradient echo sequence with fat suppression by spectral inversion recovery was performed before and after contrast material injection. The image parameters were as follows: TR/TE/FA, 3.6 ms/1.0 ms/15°; FOV, 26 × 26 cm; matrix, 320 × 240; section thickness, 3.0 mm; interslice gap, 0 mm and acquisition time, 60 s. A dynamic study in the coronal plane was performed before and 60, 120, 180 and 240 s after starting intravenous injection of 0.2 mmol/kg of gadodiamide hydrate (Omniscan®, Daiichi-Sankyo, Tokyo, Japan) at a rate of 3 mL/s, followed by a 20 mL saline flush at the rate of 3 mL/s.	dedicated four- channel breast array coil.
Schaef gen, B.	Ann Surg Oncol	2016	1.5	Siemens	NR	bilateral breast coil (Seven-Channel Biopsy Breast Array
Bufi	Eur J Radiol	2014	1.5	GE	STIR, DWI, 3D fast spolied graideint echo	dedicated breast coil
Santam aria G.	Europe an Journal of Radiolo gy	2019	1.5	GE + Siemens	T2 FSE, 3D t1 weighted, DWI (EPI)	edicated bilateral fourchannel breast surface coil
Zhang	Cancer manag ement and researc h	2020	1.5	Aurora	precontrast axial T2-weighted fat-suppressed sequence (TR 6,680 ms, TE 29 ms, thickness 3 mm) and axial T1-weighted fat-suppressed sequences (TR 4.8 ms, TE 29 ms, thickness 1.1 mm, FOV 360 mm, matrix 360×360×128) before and after a bolus of a gadolinium-based contrast agent (gadopentetate dimeglumine, 0.1 mmol/kg) was injected at a rate of 2 mL/s. Postcontrast images were obtained at 90, 180, 270, and 360 seconds after the injection	breast coil
Pasque ro	In vivo	2020	0.4 and 1.5	Multiple	Open MRI Hitachi 0.4 Tesla, MRI Philips Ingenia 1.5 Tesla or MRI Philips Achieva D-Stream 1.5 Tesla.	NR

Graeser	Breast	2021	1.5 and	NR	The standardized imaging protocol of 1.5-T and 3.0-T systems with a	dedicated breast
	cancer	-	3.0		dedicated breast multichannel surface coil consisted of an axial	multichannel
	researc				bilateral twodimensional multi-section gradient-echo dynamic series	surface coil
	h : BCR				(repetition time 250 ms; echo time 4.6 ms (1.5 T) or 2.3 ms (3 T); flip	
					angle 90°) with a section thickness of 3 mm and full 512×512	
					acquisition matrix. The voxel size of all scans was kept constant along	
					all exams with a maximum of 0.8. To ensure a high spatial resolution.	
					the field of view was adapted to the individual breast size of the	
					patient with a minimum of 280 mm and a maximum of 350 mm	
					keeping the voxel size within the intended range. The dynamic	
					sequence was performed prior to and four times after bolus injection	
					of macrocyclic gadolinium agent, gadobutrol (Gadovist®/Gadavist®,	
					Bayer AG, Leverkusen, Germany; 0.1 mmol/kg body weight), followed	
					by a saline flush. Depending on the site preference, fat suppression or	
					image subtraction was used for visualization of enhancement in the	
					T1 gradient-echo sequence. A standard center of k-space between 60	
					and 90 s was used in all exams. In addition, an axial T2-weighted fast	
					spin-echo sequence without fat suppression and with identical	
					anatomic parameters as the T1-gradient echo sequence was	
					performed.	
Nakash	Breast	2021	3.0	Philips	T1-weighted high-resolution isotropic volume examination sequences	dedicated 16-
ima	cancer				were performed using a 3D gradient-echo technique (TR/TE, 3.4/1.8	channel surface
	(Tokyo,				ms; flip angle, 10°; FOV, 33 × 33 cm; matrix, 400 × 320; section	breast coi
	Japan)				thickness, 1.8 mm interpolated to 0.9 mm) with active fat suppression.	
					Images in the axial plane were acquired before and 20, 85, and 280 s	
					after the injection of contrast material, either 0.1 mmol/kg	
					gadopentetate dimeglumine (Magnevist; Bayer, Osaka, Japan) or	
					gadobutrol (Gadovist; Bayer, Osaka, Japan). In addition, contrast-	
					enhanced unilateral sagittal images with high spatial resolution were	
					obtained for the breast with cancer and the contralateral breast using	
					a 3D gradient-echo sequence (TR/TE, 3.9/1.9 ms; flip angle, 10°; FOV,	
					20 × 20 cm; matrix, 256 × 230; section thickness, 1.0 mm interpolated	
					to 0.5 mm) 160 and 360 s, respectively, after the injection of contrast	
					material.	
Palshof	Annals	2021	1.5	GE	one axial I2W tast spin echo (FSE), one axial diffusionweighted	Liberty 9000 8
	ot				Imaging (DWI), and one TIW sequence before infusion of Multihance	Breast coi
	surgical				0.2 ml/kg at an infusion rate 1.5 ml/ s. After admission of contrast, five	
	oncolo				I LW sequences (multiphase), one I LW sagittal, and one TW (with	
	gy	2024			phase AP) including subtraction recordings were performed.	
Winder	ANZ	2021	3.0	GE	pre-contrast axial T1, axial ideal and dynamic axial T1 (one phase pre	Dedicated breast
	Journal				and five phases post), three-dimensional fatsaturated post-contrast.	COIL
	ot					
	curgon	1	1	1		

Supplementary Table 2: Search string used for obtaining relevant articles and abstracts on Pubmed and Embase.

Blocks	Pubmed	Embase
combined with AND Domain (1)	((((breast cancer[Title/Abstract]) OR (breast neoplasm*[Title/Abstract])) OR (Breast carcinoma[Title/Abstract])) OR (Breast carcinomas[Title/Abstract])) OR (Breast Neoplasms[MeSH])	'breast cancer':ab,ti,kw OR 'breast neoplasm*':ab,ti,kw OR 'breast carcinoma':ab,ti,kw OR 'breast carcinomas':ab,ti,kw OR 'breast cancer'/exp OR 'breast carcinoma'/exp
Domain (2)	(((("Neoadjuvant Therapy"[Mesh]) OR (Neoadjuvant[Title/Abstract])) OR (Neo adjuvant[Title/Abstract])) OR (Preoperative[Title/Abstract])) OR (Pre operative[Title/Abstract])	'neoadjuvant therapy'/exp OR 'neoadjuvant chemotherapy'/exp OR 'neoadjuvant':ab,ti,kw OR 'neo adjuvant':ab,ti,kw OR 'preoperative':ab,ti,kw OR 'pre operative':ab,ti,kw
Determinant	((((Magnetic Resonance Imaging[Mesh]) OR (Magnetic resonance imaging[Title/Abstract])) OR (MRI[Title/Abstract])) OR (Radiologic* response[Title/Abstract])))	'nuclear magnetic resonance imaging'/exp OR 'magnetic resonance imaging':ab,ti,kw OR 'MRI':ab,ti,kw OR 'radiologic* response':ab,ti,kw
Outcome	((((((((((((((((((((((((((((((((((((((<pre>'pathological response*':ab,ti,kw OR 'pathologic response*':ab,ti,kw OR 'pathological complete response*':ab,ti,kw OR 'pathologic complete response*':ab,ti,kw OR 'pathologic complete response'/exp OR 'pCR':ab,ti,kw OR 'residual cancer burden':ab,ti,kw OR 'residual cancer burden':ab,ti,kw OR 'residual cancer 'therapeutic effect':ab,ti,kw OR 'therapeutic effect':ab,ti,kw OR 'therapy effect'/exp OR 'therapeutic effects':ab,ti,kw OR 'response to treatment':ab,ti,kw OR 'response to treatments':ab,ti,kw OR 'therapeutic response':ab,ti,kw OR 'therapeutic response':ab,ti,kw OR 'therapeutic response':ab,ti,kw OR 'therapeutic response':ab,ti,kw OR 'treatment effect':ab,ti,kw OR 'treatment effect':ab,ti,kw OR 'residual disease'/exp OR 'treatment outcome'/exp OR 'treatment response':ab,ti,kw OR 'complete response':ab,ti,kw OR 'complete response':ab,ti,kw OR 'complete response'/exp</pre>

Supplementary Figure 1: Risk of bias assessment of included studies based on QUADAS-2 criteria

(a) Summary plot total risk of bias per category (b) traffic light plot: risk of bias per category in each individual study



			Risk o	of bias dor	nains		
		D1	D2	D3	D4	Overall	
	Eom, 2017	+	?	+	+	-	
	Andrade, 2017	X	?	?	?	×	
	Ramshorst, 2017	+	+	+	+	+	
	lwase, 2018	×	+	+	+	-	
	Marin Alcala, 2018	?	?	+	+	-	
	Murphy, 2018	+	+	+	+	+	
	Namura, 2018	+	+	+	?	-	
	Gasol Cudos, 2019	?	+	?	+	-	
	Gampenrieder, 2019	+	+	+	+	+	
	Negrao, 2019	X	+	+	+	-	
	Zhang, X., 2020	?	+	+	+	-	
	De Los Santos, 2011	+	?	+	+	-	
ĥ	De Los Santos, 2013	+	×	+	+	-	
วี	Hayashi, 2013	+	+	+	+	+	
	Sabadell, 2014	?	?	+	+	-	
	Kim, 2015	+	+	+	+	+	
	Fukuda, 2016	+	+	+	+	+	
	Schaefgen, 2017	X	×	+	+	×	
	Bufi, 2014	+	+	+	+	+	
	Santamaria, 2019	?	?	+	+	-	
	Zhang, K., 2020	+	+	+	+	+	
	Pasquero, 2020	+	×	X	+	×	
	Graeser, 2020	+	+	X	+	-	
	Nakashima, 2021	?	+	+	+	-	
	Palshof, 2021	+	+	+	+	+	
	Winder, 2021	+	+	+	+	+	
	(b)	Domains: D1: Patient	selection.	Judgement			
		D2: Index te D3: Referer	est. nce standard.		High		
		D4: Flow &	timing.			Low	
					?	No information	

Study

Sensitivity of MR	I for detecting pCR in HR·	+HER2+ subtype	Specificity of MRI for detecting pCR in HR+HER2+ subtype			
Ramshorst, 2017	⊨∎⊣	0.84 [0.74, 0.91]	Ramshorst, 2017	⊨∎⊣	0.48 [0.37, 0.58]	
Iwase, 2018	⊢ −−−−	0.70 [0.30, 0.93]	lwase, 2018	⊢−■	0.95 [0.80, 0.99]	
Numura, 2018	⊢ _	0.95 [0.80, 0.99]	Numura, 2018	⊢■⊣	0.80 [0.70, 0.88]	
Gasol Cudos, 2019	⊨−■−−1	0.64 [0.47, 0.78]	Gasol Cudos, 2019	⊢∎→	0.20 [0.12, 0.32]	
Gampenrieder, 2019	⊢	0.71 [0.43, 0.89]	Gampenrieder, 2019	⊢ ■ 1	0.54 [0.36, 0.71]	
Negrao, 2019	┝──■─┤	0.77 [0.57, 0.89]	Negrao, 2019	⊢ −■−+	0.71 [0.54, 0.84]	
De Los Santos, 2013	⊢	0.49 [0.34, 0.64]	De Los Santos, 2013	⊢∎⊣	0.77 [0.69, 0.84]	
Hayashi, 2013	⊢− ■−−−−1	0.31 [0.14, 0.54]	Hayashi, 2013	⊢∎ ⊣	0.83 [0.68, 0.92]	
Fukada, 2016	■	0.17 [0.02, 0.69]	Fukada, 2016	⊢	0.76 [0.56, 0.89]	
Bufi, 2014	⊢	0.71 [0.43, 0.89]	Bufi, 2014	⊢∎	0.58 [0.36, 0.78]	
Santamaria, 2019	⊢−−−− 1	0.75 [0.20, 0.97]	Santamaria, 2019	⊢	0.75 [0.44, 0.92]	
Zhang K., 2020	⊢∎	0.40 [0.29, 0.51]	Zhang K., 2020	⊢∎I	0.92 [0.85, 0.95]	
Graeser, 2021	⊢■	0.56 [0.40, 0.71]	Graeser, 2021	⊢ ∎+	0.91 [0.82, 0.96]	
Nakashima, 2021	⊢−−−− 1	0.58 [0.24, 0.86]	Nakashima, 2021		0.77 [0.57, 0.89]	
Palshof, 2021	⊢	0.43 [0.22, 0.67]	Palshof, 2021	⊢−■⊣	0.87 [0.72, 0.94]	
Winder, 2021	⊢	0.44 [0.17, 0.74]	Winder, 2021		0.69 [0.49, 0.84]	
Pooled sensitivity	-	0.60 [0.50, 0.70]	Pooled specificity		0.74 [0.63, 0.83]	
(4)						
(a)	0 0.25 0.5 0.75 1			0 0.25 0.5 0.75 1		

(b)	0 0.20 0.0 0.70 1		0 0.2	0 0.0 0.10 1		
Sensitivity of MF	RI for detecting pCR in HR+	HER2- subtype	Specificity of MRI for detecting pCR in HR+HER2- subtype			
Iwase, 2018 Murphy, 2018 Numura, 2018 Gasol Cudos, 2019 Gampenrieder, 2019 Zhang X, 2020 De Los Santos, 2011 De Los Santos, 2013 Fukada, 2016 Schaefgen, 2010 Bufi, 2014 Santamaria, 2019 Zhang K, 2020 Nakashima, 2021		$\begin{array}{c} 0.62 \left[0.22, 0.91 \right] \\ 0.38 \left[0.09, 0.76 \right] \\ 0.48 \left[0.60, 0.95 \right] \\ 0.62 \left[0.39, 0.81 \right] \\ 0.79 \left[0.51, 0.93 \right] \\ 0.57 \left[0.39, 0.74 \right] \\ 0.50 \left[0.05, 0.95 \right] \\ 0.61 \left[0.31, 0.85 \right] \\ 0.46 \left[0.32, 0.60 \right] \\ 0.32 \left[0.14, 0.58 \right] \\ 0.48 \left[0.40, 0.99 \right] \\ 0.32 \left[0.12, 0.61 \right] \\ 0.77 \left[0.48, 0.93 \right] \\ 0.39 \left[0.15, 0.69 \right] \\ 0.39 \left[0.51, 0.69 \right] \\ 0.39 \left[0.27, 0.43 \right] \\ 0.52 \left[0.52, 0.99 \right] \\ 0.72 \left[0.40, 0.91 \right] \\ \end{array}$	Iwase, 2018 Murphy, 2018 Numura, 2018 Gaspenireder, 2019 Negrao, 2019 Zhang X, 2020 De Los Santos, 2011 De Los Santos, 2011 Hayash, 2013 Fukada, 2016 Schaefgen, 2016 Bufi, 2014 Santamaria, 2019 Zhang K, 2020 Nakashima, 2021		$\begin{array}{c} 0.89 \left[0.78, 0.95 \right] \\ 0.81 \left[0.47, 0.96 \right] \\ 0.94 \left[0.91, 0.96 \right] \\ 0.94 \left[0.91, 0.96 \right] \\ 0.81 \left[0.74, 0.87 \right] \\ 0.69 \left[0.58, 0.78 \right] \\ 0.85 \left[0.67, 0.90 \right] \\ 0.87 \left[0.69, 0.95 \right] \\ 0.87 \left[0.69, 0.95 \right] \\ 0.87 \left[0.69, 0.95 \right] \\ 0.93 \left[0.83, 0.98 \right] \\ 0.93 \left[0.83, 0.98 \right] \\ 0.81 \left[0.73, 0.87 \right] \\ 0.87 \left[0.72, 0.95 \right] \\ 0.92 \left[0.85, 0.95 \right] \\ 0.92 \left[0.85, 0.97 \right] \\ 0.92 \left[0.81, 0.97 \right] \\ 0.92 \left[0.85, 0.95 \right] \\ 0.95 \left[0.95 \right] \\ 0.95 $	
Pooled sensitivity		0.58 [0.24, 0.86] 0.55 [0.45, 0.64]	Winder, 2021 Pooled specificity		0.77 [0.62, 0.87] 0.88 [0.84, 0.91]	
(c)	0 0.25 0.5 0.75 1		0 0.25	0.5 0.75 1		

. 0.67 [0.58, 0.74] Pooled specificity 0.85 [0.81, 0.88] Pooled sensitivity 0 0.25 0.5 0.75 0 0.25 0.5 0.75 1 (a) 1 Sensitivity of MRI for detecting pCR in HR-HER2+ subtype Specificity of MRI for detecting pCR in HR-HER2+ subtype $\begin{array}{c} 0.84 \left[0.67, 0.93 \right] \\ 0.62 \left[0.45, 0.77 \right] \\ 0.77 \left[0.59, 0.89 \right] \\ 0.93 \left[0.81, 0.98 \right] \\ 0.92 \left[0.73, 0.98 \right] \\ 0.53 \left[0.54, 0.96 \right] \\ 0.95 \left[0.54, 0.96 \right] \\ 0.95 \left[0.54, 0.96 \right] \\ 0.85 \left[0.54, 0.96 \right] \\ 0.87 \left[0.65, 0.96 \right] \\ 0.71 \left[0.55, 0.96 \right] \\ 0.75 \left[0.44, 0.92 \right] \\ 0.75 \left[0.44, 0.92 \right] \\ 0.75 \left[0.44, 0.92 \right] \\ 0.92 \left[0.52, 0.99 \right] \\ 0.89 \left[0.87, 0.99 \right] \\ 0.81 \left[0.83, 0.99 \right] \\ 0.81 \left[0.83, 0.99 \right] \\ 0.81 \left[0.83, 0.94 \right] \\ 0.92 \left[0.52, 0.99 \right] \\ 0.81 \left[0.83, 0.94 \right] \\ 0.79 \left[0.93, 0.99 \right] \\ 0.81 \left[0.83, 0.94 \right] \\ 0.79 \left[0.93, 0.99 \right] \\ 0.81 \left[0.83, 0.94 \right] \\ 0.79 \left[0.93, 0.99 \right] \\ 0.81 \left[0.83, 0.94 \right] \\ 0.79 \left[0.93, 0.99 \right] \\ 0.81 \left[0.83, 0.94 \right] \\ 0.79 \left[0.93, 0.94 \right] \\ 0.79 \left[0.93, 0.95 \right] \\ 0.81 \left[0.83, 0.94 \right] \\ 0.79 \left[0.93, 0.95 \right] \\ 0.81 \left[0.83, 0.94 \right] \\ 0.79 \left[0.73, 0.98 \right] \\ 0.75 \left[0.83, 0.94 \right] \\ 0.79 \left[0.73, 0.98 \right] \\ 0.75 \left[0.83, 0.94 \right] \\ 0.79 \left[0.73, 0.98 \right] \\ 0.75 \left[0.83, 0.94 \right] \\ 0.79 \left[0.73, 0.98 \right] \\ 0.75 \left[0.83, 0.94 \right] \\ 0.75 \left[0.83, 0.94 \right] \\ 0.75 \left[0.93, 0.95 \right] \\$ Andrade, 2017 Ramshorst, 2017 Iwase, 2018 Numura, 2018 Gasol Cudos, 2019 Gampenrieder, 2019 Negrao, 2019 Zhang X., 2020 De Los Santos 2013 0.35 [0.15, 0.61] Andrade, 2017 0.35 [0.15, 0.61] 0.82 [0.74, 0.88] 0.64 [0.41, 0.82] 0.91 [0.80, 0.96] 0.64 [0.50, 0.76] 0.62 [0.39, 0.81] Ramshorst, 2017 Iwase, 2018 Numura, 2018 Gasol Cudos, 2019 _ ---------Gampenrieder, 2019 Negrao, 2019 Zhang X., 2020 De Los Santos, 2013 0.80 [0.60, 0.92] 0.50 [0.17, 0.83] 0.47 [0.33, 0.63] 0.46 [0.33, 0.60] • De Los Santos, 2013 Hayashi, 2013 Havashi, 2013 0.77 [0.48, 0.93] 0.83 [0.51, 0.96] 0.61 [0.36, 0.81] 0.47 [0.39, 0.55] Fukada, 2016 Bufi, 2014 Santamaria, 2019 Fukada, 2016 Bufi, 2014 Santamaria, 2019 Zhang K., 2020 Graeser, 2021 Zhang K., 2020 0.94 [0.79, 0.99] 0.76 [0.58, 0.87] 0.72 [0.47, 0.88] 0.68 [0.39, 0.88] 0.82 [0.63, 0.92] Graeser, 2021 Nakashima, 2021 Palshof, 2021 Winder, 2021 0.65 [0.39, 0.85] 0.44 [0.17, 0.74] 0.64 [0.30, 0.88] Nakashima, 2021 Palshof, 2021 Winder, 2021 _ Pooled sensitivity 0.65 [0.56, 0.73] 0.81 [0.74, 0.86] Pooled specificity 0 0.25 0.5 0.75 1 Λ 0.25 0.5 0.75



Supplementary Figure 2: Forest plots for pooled sensitivity and specificity with their 95% confidence intervals of MRI for pCR in each of the breast cancer IHC subtypes (a) HR-/HER2- subtype (b) HR-/HER2+ subtype (c) HR+/HER2- (d) HR+/HER2+ subtype

Supplementary Figure 3: Deeks funnel plots per IHC subtype (a) HR+/HER2- subtype (b) HR+/HER2+ subtype (c) HR-/HER2+ subtype (d) HR-/HER2+ subtype

