European Journal of Radiology 121 (2019) 108736

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Review

Exploring breast cancer response prediction to neoadjuvant systemic therapy using MRI-based radiomics: A systematic review

R.W.Y. Granzier[a,b,*], T.J.A. van Nijnatten[c], H.C. Woodruff[c,b,d], M.L. Smidt[a,b], M.B.I. Lobbes[b,c,e]

[a] Department of Surgery, Maastricht University Medical Centre+, Maastricht, The Netherlands
[b] GROW – School for Oncology and Developmental Biology, Maastricht University Medical Centre+, Maastricht, The Netherlands
[c] Department of Radiology and Nuclear Medicine, Maastricht University Medical Centre+, Maastricht, The Netherlands
[d] The D-Lab, Department of Precision Medicine, Maastricht University, Maastricht, The Netherlands
[e] Department of Radiology, Zuyderland Medical Center, Sittard-Geleen, The Netherlands

ARTICLE INFO

Keywords:
Breast neoplasms
Tumor response prediction
Neoadjuvant therapy
Magnetic resonance imaging
Radiomics

ABSTRACT

Purpose: MRI-based tumor response prediction to neoadjuvant systemic therapy (NST) in breast cancer patients is increasingly being studied using radiomics with outcomes that appear to be promising.

The aim of this study is to systematically review the current literature and reflect on its quality.

Methods: PubMed and EMBASE databases were systematically searched for studies investigating MRI-based radiomics for tumor response prediction. Abstracts were screened by two reviewers independently. The quality of the radiomics workflow of eligible studies was assessed using the Radiomics Quality Score (RQS). An overview of the methodologies used in steps of the radiomics workflow and current results are presented.

Results: Sixteen studies were included with cohort sizes ranging from 35 to 414 patients. The RQS scores varied from 0 % to 41.2 %. Methodologies in the radiomics workflow varied greatly, especially region of interest segmentation, features selection, and model development with heterogeneous outcomes as a result. Seven studies applied univariate analysis and nine studies applied multivariate analysis. Most studies performed their analysis on the pretreatment dynamic contrast-enhanced T1-weighted sequence. Entropy was the best performing individual feature with AUC values ranging from 0.83 to 0.85. The best performing multivariate prediction model, based on logistic regression analysis, scored a validation AUC of 0.94.

Conclusion: This systematic review revealed large methodological heterogeneity for each step of the MRI-based radiomics workflow, consequently, the (overall promising) results are difficult to compare. Consensus for standardization of MRI-based radiomics workflow for tumor response prediction to NST in breast cancer patients is needed to further improve research.

1. Introduction

Neoadjuvant systemic therapy (NST) is increasingly used for breast cancer treatment [1–4]. Compared to adjuvant chemotherapy, NST bears the advantages of observing *in vivo* tumor response, decreasing tumor size (*i.e.,* enabling breast-conserving therapy where initially mastectomy was indicated), and the possibility of achieving pathologic complete response (pCR) [5,6]. Tumor pCR is clinically relevant as it correlates with better prognosis expressed in improved disease-free survival and overall survival [7–9]. Unfortunately, not all patients respond well to NST, with treatment response ranging from pCR to non-response, and even to disease progression.

Different imaging modalities can be used to assess and predict tumor response to NST in breast cancer patients. Breast magnetic resonance imaging (MRI) is considered the most accurate imaging modality for both tumor assessment and response prediction [10,11]. However, its accuracy in assessing and predicting tumor response to NST remains insufficient to adapt treatment in clinical practice [12–14]. In addition, it is not possible to predict tumor response to NST based solely on the pretreatment MRI. Therefore, there is a continuous need to further improve breast MRI accuracy for this purpose.

Recent developments in tumor response prediction to NST show promising results in objectively interpreting MR images (usually from pre- and mid-treatment exams) using quantitative imaging analysis

Abbreviations: RQS, Radiomics Quality Score; ROI, Region of interest; NST, Neoadjuvant systemic therapy; pCR, pathologic complete response; DCE-MRI, Dynamic contrast-enhanced magnetic resonance imaging; RECIST, Response evaluation criteria in solid tumors; DWI, diffusion weighted images

* Corresponding author at: Department of Surgery Maastricht University Medical Center+, P.O. Box 5800, 6202 AZ, Maastricht, The Netherlands.
E-mail address: r.granzier@maastrichtuniversity.nl (R.W.Y. Granzier).

https://doi.org/10.1016/j.ejrad.2019.108736
Received 11 July 2019; Received in revised form 26 September 2019; Accepted 31 October 2019

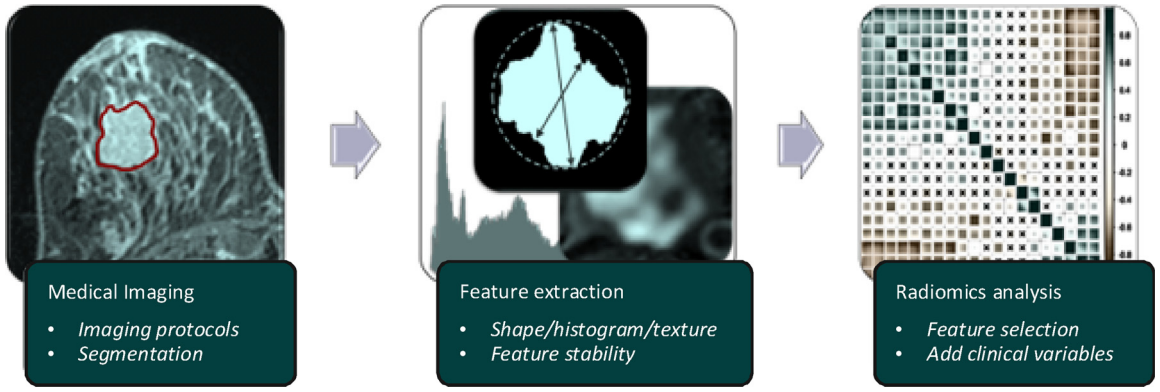

Fig. 1. The radiomics workflow.

(QIA), such as radiomics. Radiomics is a high-throughput QIA method which extracts a large number of features from standard-of-care medical images [15] (Fig. 1). It is hypothesized that these features capture underlying information in which the underlying structural heterogeneity of the entire breast tumor can be detected and quantified [16,17] at a far higher level of detail than by the radiologist's visual assessment. Correlations between imaging features and clinical or biological endpoints are found using machine learning tools which are able to combine multiple variables from differing sources of information into complex, non-linear models designed to classify (*e.g.,* tumor vs. healthy) or predict (*e.g.,* overall survival) outcomes.

The radiomics workflow consists of several consecutive steps. The methodology employed in all steps determines the quality of the final model. Different methodologies can lead to heterogeneous results which are difficult to compare. To the best of our knowledge, the variation in methodologies used in MRI-based radiomics for tumor response prediction to NST in breast cancer had not been evaluated before. This systematic review aims to assess the quality of the radiomics workflow using the Radiomics Quality Score (RQS) [18] and to report on the different methodologies used in all consecutive steps of the radiomics workflow. The RQS assigns points to all steps in the radiomics workflow (*e.g.,* ROI segmentation, feature reduction, the use of validation cohorts) were the maximum number of points per item depends on the relevance to quality. Furthermore, we summarize and briefly evaluate the current results reported on the topic of MRI-based radiomics for tumor response prediction to NST in breast cancer patients.

2. Methods

2.1. Literature search

For this systematic review, a structured search using the PubMed and EMBASE (OVID) databases was conducted. The search was performed until May 8, 2019. Two researchers with two and five years of experience in breast imaging (R.G. and T.v.N.) independently performed the literature search to select potential studies. Full search strategies are described in Appendix 1.

2.2. Eligibility criteria & study selection

Studies were included if they met the following criteria: 1) The cohort consists of breast cancer patients who underwent at least a pretreatment breast MRI; 2) Patients were treated with NST; 3) Breast MRI radiomics analysis was used to predict tumor response to NST. 4) Studies are reported in English with institutional full-text availability. We excluded reviews, technical reports, letters to editors, comments to published studies, conference proceedings, as well as duplicate studies. When identical datasets were selected, the study reporting the most radiomics workflow details was chosen. The reviewers read all titles and abstracts independently and rejected studies that did not meet the aforementioned criteria. The full text of the remaining studies was determined for further eligibility by the same reviewers. In the case of discrepancies, a third reviewer (H.W.) was consulted to reach final consensus. For each included study, reference lists were searched manually for additional eligible studies.

2.3. Data extraction

The following data were extracted (if available): first author, year of publication, study design, analysis strategy, sample size, breast cancer subtype, NST regimen, MRI specifications (manufacturer, field strength, coil specifications), MR imaging parameters, MRI sequence, image preprocessing, feature extraction software, number of features, ROI segmentation methods (including profession and years of those who performed segmentations), response definition, and results (including feature selection method, chosen classifier and diagnostic performance in terms of Area Under the receiver operating characteristic Curve (AUC)).

2.4. Data analysis

To quantify the radiomics workflow quality and its reporting, the included studies were evaluated in consensus by the two initial reviewers using the RQS. This checklist is comprised of 16 components in the radiomics workflow, with 36 points (100 %) representing a maximum score [18]. Most items can receive 0, 1 or 2 point(s). To reflect the importance of prospective studies, seven points are scored when these are performed and published. Another important step, the external validation on multiple datasets, yields 5 more points, but if no external validation is performed at all, 5 points are deducted. The item 'feature reduction or adjustment for multiple testing' can also score negatively, making a negative total score possible. Negative total scores were set at 0 %. Not all items were applicable to studies performing univariate analysis, for these items n/a was entered. Explanation of all RQS items including the attainable points is presented as supplemental information (Appendix 2). Descriptive analysis of the methodologies used in the included studies was performed according to the consecutive steps of the radiomics workflow: 1) image acquisition and preprocessing, 2) region of interest (ROI) segmentation, 3) feature extraction and selection 4) feature analysis and modeling, and 5) performance evaluation [15,19]. The response prediction results will be summarized in steps four and five of the radiomics workflow in the result section.

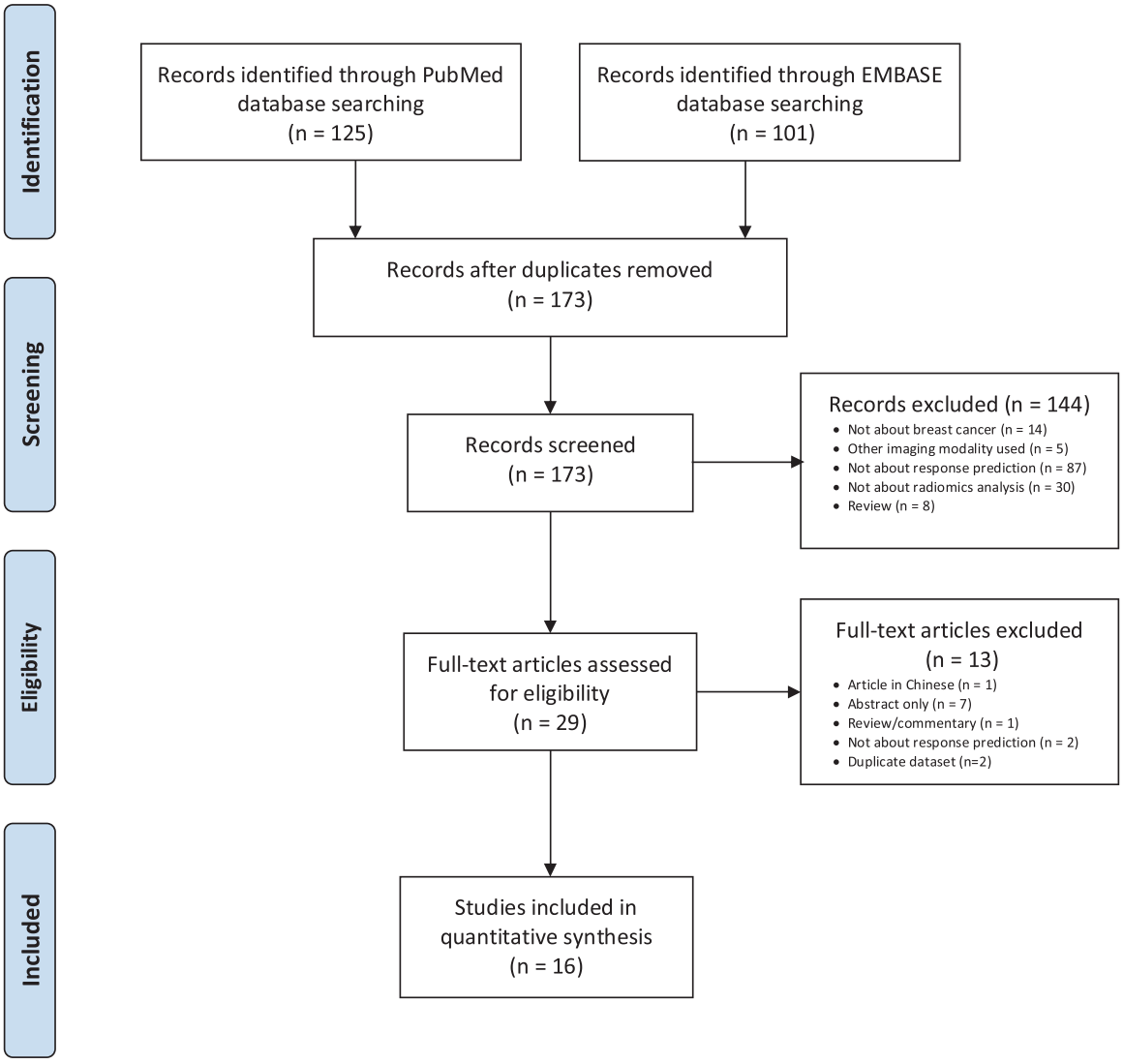

Fig. 2. Flow diagram of preferred reporting items for systematic reviews and meta-analyses (PRISMA).

3. Results

3.1. Study selection

A total of 208 records were identified through the searches in PubMed (113 records) and EMBASE (95 records). After removing duplicates, the titles and abstracts of 173 records were screened, resulting in twenty-nine records eligible for inclusion. The full text of these studies was read and the selection criteria were applied, yielding a total of sixteen studies to be included in this systematic review. Fig. 2 details the Preferred Reporting Item for Systematic reviews and Meta-Analyses (PRISMA) flow diagram [20], including the different screening phases.

3.2. General characteristics of included studies

A total of 1636 patients were retrospectively included in the sixteen studies presented in this review [21–36] (Table 1). Cohort sizes ranged from 35 to 414 patients, with a largest validation cohort of 137 patients. All breast cancer subtypes were included in all studies with the exception of Banerjee et al. [25], which included triple-negative (TN) breast tumors only. To assess tumor response to NST, radiologic (RECIST) and pathologic methods were used. The studies that based their tumor response on pathology used varying definitions of response. While the majority of studies only used patients with no residual cancer burden as 'responders', four studies added patients with invasive residual cancer burden to that group [23,29,35,36].

3.3. RQS

The RQS heat map summarized the overall RQS scores which were considered as 'poor' with a mean score of 11.8 % (range 0 %–41.2 %). The two most recently published articles [21,22] showed the highest scores, which is mainly due to the points achieved on the item 'validation' (Fig. 3). Seven articles scored n/a for the items 'features reduction or adjustment multiple testing' and 'multivariate analysis', due to the fact that these articles performed univariate analysis only. Eight articles scored 0 %, which could also mean that the article had a negative total score.

3.4. Radiomics workflow

3.4.1. Image acquisition and preprocessing

The authors of all studies, except Wu et al. [32], performed their analysis on the pretreatment MRI. Two studies [28,34] analyzed the differences between MRI exams at two-time points. Both 1.5 and 3.0 T MR scanners, with dedicated breast receiver coils, were used. The majority of the studies used the dynamic contrast-enhanced T1 weighted (DCE T1W) sequence [21,22,24,25,27,29–33,35,36], three

Table 1
Overview of included studies in this review.

Author	Year	Study design	Sample Size	Analysis strategy	Software	MR Specification			Sequence	Phase DCE T1W used	Image pre-processing	Number of features extracted	Feature type (number)	Regimen neoadjuvant systemic therapy (n)	ROI segmentation method	Definition of responders
						Manufacturer	Field strength	Channels of breast Coil								
Xiong [21]	2019	R	125	M	–	Philips, GE	1.5 T, 3.0T	–	DCE T1W, T2WI, DWI	2	No*	647 (per sequence)	Geo(8), 1[e](17), GLCM(22), GLRLM(14), NGTDM(5), GLSZM(13), wavelet	T (110), AC-T (15)*	3D, M	Non-responders[a]
Liu [22]	2019	R	414	M	Built in-house Matlab toolbox	Philips	1.5 T, 3.0T	–	DCE T1W, T2WI, DWI	2	N	4650 (per sequence)	Geo(8), 1[e](17), T(99), wavelet (4535)	T (193), AC (63), T-AC (158)	3D, M	Complete response[b]
Yoon [23]	2018	R	83	U	CGITA 1.3	Philips	3.0T	7	DWI	n/a	n/a n/a	46	IH(5), CM(6), VAM(11), NIDM (5), ISZM(11), TSM(2), NGLCM (6), NGLFM(5)	A-D (13), AC (60), H/Tr (10)	2D, M	Complete with partial response[d]
Hope Cain [24]	2018	R	288	M	–	Siemens, GE	1.5 T, 3.0T	–	DCE T1W	1	No	529	Geo(15), T-E(30), FGT-E(82), T-E-T(135), FGT-E-T(135), T-E-V (35), FGT-E-V (34), SVTH(4), TFGT-E(18)	NS	3D, SA	Complete response[c]
Banerjee [25]	2018	R	53	M	–	25 different sites	1.5 T, 3.0T	–	DCE T1W, T2W	NS	N	538	GLCM(24), Riesz (72), Geo(20), IH (24), 1[e](398)	G-C-I (53)	2D, M	Complete response[d]
Chamming's [26]	2018	R	85	U	TexRAD	GE	1.5 T	8	T2W	1	No	6	IH(6)	AC-P (72), AC-D (4), EC-P (2), FEC-P (4), Ca-P (3)	2D, M	Complete response[b]
Giannini [27]	2017	R	44	M	Built in-house	GE	1.5 T	8	DCE T1W	1	N	27	GLCM (17),GLRLM(10)	A-P (28), A-P-Tr (14)	3D, A	Complete response[b]
Henderson [28]	2017	R	88	U	MaZda 4.7	Siemens	3.0T	7	T2W	n/a	N	1	E	FEC (18), FEC-D (41), FEC-D, Tr (26), D,P,TDM-1 (3)	2D, M	Complete response[e]
Fan [29]	2017	R	103	M	–	Siemens	1.5 T, 3.0T	8	DCE T1W	Pre, 1, 2	No	158	IH(11), T(33), 1[e](33), Dyn(84)	NS	3D, SA	Complete with partial response[f]
Braman [30]	2017	R	117	M	Built in-house	Philips*, Siemens*	1.5 T, 3.0T	–	DCE T1W	1	No*	1980	T(99), PK (3), 1[e](5)	A-T (89), D-Tr (28)	3D*, M	Complete response[g]
Thibault [31]	2017	R	38	U	–	Siemens	3.0T	4	DCE T1W	2	No	1043	GLCM, GLRLM, GLSZM, LBP, PS, IH	AC-T (31), ISPY-2 (7)	3D, M	Complete response[e]
Wu [32]	2016	R	35	M	–	Philips	3.0T	–	DCE T1W	3	N	4	GLCM(4)	–	3D, SA	Complete response[d]
Michoux [33]	2015	R	69	M	Open-source Matlab codes	Philips	1.5 T	–	DCE T1W	2	No	25	K(3), Geo(2), GLCM(9), GLRLM (11)	A-T (43), A-T-Tr (26)	2D, M,A	Non-responders

(*continued on next page*)

Table 1 (*continued*)

Author	Year	Study design	Sample Size	Analysis strategy	Software	MR Specification			Sequence	Phase DCE T1W used	Image pre-processing	Number of features extracted	Feature type (number)	Regimen neoadjuvant systemic therapy (n)	ROI segmentation method	Definition of responders
						Manufacturer	Field strength	Channels of breast Coil								
Parikh [34]	2014	R	36	U	TexRAD	Siemens	1.5 T	4 or 16	T2W	n/a	No	2	E, U	FEC-D (7), EC-D (29), EC-D, Tr (6)	2D, M	Complete response[b]
Teruel [35]	2014	R	58	U	In-house build	Siemens	3.0T	4	DCE T1W	2	N	16	GLCM(16)	FEC (28), FEC-T (30)	3D*, SA	Complete with partial response[h]
Ahmed [36]	2013	R	100	U	In-house build	GE	3.0T	8	DCE T1W	1,2,3,4 and 5*	N	16	GLCM(16)	EC-D (57), NS (38)	3D*, SA	Complete with partial response[k]

Abbreviations: n/a = not applicable, - = unknown, NS = not specified.**Study design:** P = prospective, R = retrospective.
Analysis strategy: U = Univariate feature analysis, M = multivariate prediction models. **Sample size:** *92 new patients compared to Braman ([30]).
MR Manufacturer: * three different Philips scanners, * three different Siemens scanners.
Field strength: T = Tesla.
Sequence: DCE-T1W = dynamic contrast-enhanced T1-weighted image, T2W = T2-weighted image, DWI = diffusion weighted image.
Phase DCE T1W used: pre = pre-contrast, 1 = 1 min post-contrast, 2 = 2 min post-contrast, 3 = 3 min post-contrast, 4 = 4 min post-contrast, 5 = 5 min post-contrast, *phases were used as separate models.
Image pre-processing: N = normalization, No = not performed, * = features values were normalized.E = Entropy, U = Uniformity, IH = intensity histogram, CM = co-occurrence matrix, VAM = voxel-alignment, GLCM = grey-level co-occurrence matrices, NIDM = neighborhood intensity difference matrix, ISZM = intensity size-zone matrix, NGLCM = normalized gray-level co-occurrence matrix, NGLDM = neighborhood gray-level dependence matrix, TSM = texture spectrum matrix, GLRLM = gray-level run length matrix, GLSZM = gray-level size zone matrix, LBP = local binary pattern, PS = pattern spectrum, Geo = geometric, T = texture, T-E-(T)= tumor enhancement (texture), FGT-E-(T) = fibroglandular tissue enhancement (texture), T-E-V = tumor enhancement variation, FGT-E-V = FGT-E-variation, SVTH = spatial variation of tumor heterogeneity, TFGT-E = combination of tumor and FGT enhancement, 1[e] = first order statistics, Dyn = dynamic, K = kinetic, PK = pharmacokinetic.
Regimen neoadjuvant systemic therapy: D = docetaxel, C = carboplatin, Tr = trastuzumab, PM = pertuzumab, A = doxorubicin, T = taxane, P = paclitaxel, AC = anthracycline, cyclophosphamide, H/Tr = hormone therapy or trastuzumab, G = gemcitabine, I = iniparib, EC = epirubicin- cyclophosphamide, FEC = fluorouracil, epirubicin- cyclophosphamide, Ca = carboplatin, TDM-1 = trastuzumab, emtansine, ISPY -2 = chemotherapy trial, *53 patients received trastuzumab.
Delineation: M = manual, SA = semi-automatic, A = automatic, *three adjacent representative slices with largest tumor area.
Definition of responders: [a]pCR defined as the absence of invasive cancer in the breast surgical specimen (ypT0/is), [b]Miller-Payne grade 1 and 2, [c]pCR defined as ypT0/isN0, [d]complete and partial response via Sataloff classification, [e]pCR defined as absence of invasive or in situ disease in the breast and/or lymph nodes (ypT0N0), [f]pCR defined via RCB scoring system; RCB = 0, [g]RECIST criteria, complete responder (CR) and partial responder (PR), [h] absence of any residual invasive cancer OR residual cancer burden below 1 cm, [k]absence of any residual invasive cancer OR a decrease in longest tumor diameter of greater than 50 %.

RQS component / Study	Image protocol quality (2)	Multiple segmentation (1)	phantom study (1)	Test-retest (imaging at multiple time points) (1)	Feature reduction or adjustment multiple testing (3)	Multivariate analysis with non radiomics features (1)	Biological correlates (1)	Cut-off analyses (1)	Discrimination statistics (2)	Calibration statistics (2)	Prospective study (7)	Validation (5)	Comparison to 'gold standard' (2)	Potential clinical utility (2)	Cost-effectiveness (1)	Open science and data (4)	Total score
Xiong (22)	1	0	0	0	3	1	0	1	2	1	0	2	2	0	0	0	36.1%
Liu (23)	1	1	0	0	3	1	0	0	2	0	0	5	2	0	0	0	41.2%
Yoon (24)	1	0	0	0	n/a	n/a	0	0	0	0	0	-5	2	0	0	0	0.0%
Hope Cain (25)	0	0	0	0	3	0	0	0	1	0	0	2	2	0	0	0	19.4%
Banerjee (26)	0	0	0	0	3	1	0	0	2	0	0	-5	2	0	0	0	8.3%
Chamming's (27)	1	0	0	0	n/a	n/a	0	0	1	0	0	-5	2	0	0	0	0.0%
Giannini (28)	1	0	0	0	3	0	0	0	2	0	0	-5	2	1	0	0	11.1%
Henderson (29)	1	0	0	0	n/a	n/a	0	0	1	0	0	-5	2	1	0	0	0.0%
Fan (30)	1	0	0	0	3	0	0	0	2	0	0	2	2	1	0	0	30.6%
Braman (31)	1	0	0	0	3	0	0	0	2	0	0	2	2	1	0	0	30.6%
Thibault (32)	1	0	0	1	n/a	n/a	0	0	1	1	0	-5	2	0	0	0	0.0%
Wu (33)	1	0	0	0	3	1	0	0	2	0	0	-5	2	0	0	0	11.1%
Michoux (34)	1	0	0	0	-3	1	0	0	2	0	0	-5	2	0	0	0	0.0%
Parikh (35)	1	0	0	0	n/a	n/a	0	0	1	0	0	-5	2	0	0	0	0.0%
Teruel (36)	1	0	0	0	n/a	n/a	0	0	1	0	0	-5	2	1	0	0	0,0%
Ahmed (37)	1	0	0	0	n/a	n/a	0	0	1	0	0	-5	2	0	0	0	0.0%

Fig. 3. Heat map of radiomics quality score (RQS) per component of all included studies, n/a = not applicable.

studies T2W-sequences [26,28,34], and one study diffusion-weighted imaging sequence (DWI, using b-values of 0 and 1000s/mm^2) [23]. Two studies applied multiparametric MRI for their analysis, consisting of DWI, DCE T1W, and T2W [21,22]. To exclude inter-scan differences one study used healthy contralateral breast parenchyma [28]. Wu et al. [32] performed intratumor partitioning and Fan et al. [29] added breast background parenchyma to their analysis region. Half of the studies reported on image pre-processing performed by image discretization with histogram normalization [22,25,27,28,32,35,36]. Two studies performed only feature normalization [21,30].

3.4.2. Region of interest segmentation

Six studies used 2D segmentations, five of which were performed manually [23,25,26,28,34], and one fully automatically [33]. Three studies used three adjacent representative slices with the maximum tumor diameter as their ROI, one performed manually [30], and two semi-automatically [35,36]. Seven studies performed 3D segmentation, three performed manually [21,22,31], three semi-automatically [24,29,32], and one automatically [27]. Different approaches were used for semi-automatic segmentation: 1) Point-and-click 3D segmentation, where segmentation started from a seed location determined by a mouse click [29]. 2) Using a threshold to exclude necrotic, healthy, and/or fatty tissue followed by manual adjustments [32]. 3) 3D box placing around the tumor after which an algorithm (fuzzy C-means) generated the tumor mask [24].

Segmentations were mostly performed by (breast) radiologists with 1–23 years of experience. Three studies did not report on this topic [23,28,36]. It remained unclear whether a radiologist performed the segmentations in the study of Wu et al. [32] since they only reported on years of experience without naming a profession.

3.4.3. Feature extraction and selection

The number of features extracted ranged from 1 to 4650, mostly extracted with in-house built software. All studies, with the exception of four [26,28,32,34], extracted at least all Haralick textural features [37]. Feature selection only applied to the studies developing multivariate prediction models, where these studies all opted for a different approach, with regression methods chosen most often. (Table 2). The study of Michoux et al. [33] did not report on their feature selection method, and the study of Wu [32] did not perform feature selection at all. The number of features included in the models varied between two and twenty-four with a mean of six features.

3.4.4. Feature analysis and modeling

Over half of all studies presented individual features significant for tumor response prediction to NST (Table 2). Nine studies developed multivariate prediction, where most of these studies used a logistic regression model for the discriminant analysis [21,24,27,29,32,33]. Four studies used a support vector machine model (SVM). In addition to intratumoral feature extraction, the study of Braman et al. [30] added a peritumoral region to the feature extraction region. The combination of peritumoral and intratumoral features yielded higher clustering accuracies for pCR prediction when compared to the feature groups individually (88 % vs 71 % vs 79 %). Three studies [22,24,30] performed subgroup analyses for patients with different breast cancer subtypes, all showing better results in the models developed for the specific subtypes (Table 2). Henderson et al. [28] showed excellent sensitivity and specificity for pCR prediction based on individual features within each subtype: ER+: 100 %/ 100 %; HER2+: 83.3 % /95.7 %, TN: 87.5 % /80.0 %. Furthermore, two studies explored individual features to distinguish TN breast cancer from all other subtypes [26,36]. Xiong et al. [21] combined a clinical model based on HER2 and KI67 status,

Table 2
Results of all included studies.

Author	Number of patients	Time point MRI exam	Analysis strategy	Univariate feature analysis			Multivariate feature analysis									
				Significant features	p-value	AUC	Feature selection	Number of features	Model/ Classifier	Model validation	AUC (training)	Accuracy training	Sens (%)	Spec (%)	AUC (Validation)	Accuracy validation
Xiong (21)	125	P	M	–	–	–	WRST, LASSO	4	MLR	TV	MRS: 0.92 (95% [0.84-1.00])	–	–	–	0.83 (95% [0.65-1.00])	–
											MCS: 0.99 (95% [0.96-1.00])	95.2% (93.4-97.1%)	–	–	0.94 (95% [0.85-1.00])	93.55% (91.44-95.69%)
Liu (22)	414	P	M	–	–	–	UFA, PCC, RFB	4	SVM	EV*	MRS: 0.79 (95% [0.71-0.87])	–	–	–	0.79 (95% [0.65-0.93])	–
											MCS: 0.86 (95% [0.80-0.92])	–	–	–	0.80 (95% [0.67-0.91])	–
											HR + /HER2- 0.81 (95% [0.69-0.93])	–	–	–	0.87 (95% [0.66-1.00])	–
											HER2 + 0.70 (95% [0.56-0.85])	–	–	–	0.79 (95% [0.59-0.99])	–
											TN 0.96 (95% [0.89-1.00])	–	–	–	0.84 (95% [0.69-1.00])	–
Yoon (23)	83	P	U	Entropy (histogram)	0.024	–	n/a	n/a	n/a	n/a	–	–	–	–	n/a	n/a
				ASM	0.033	–										
				Entropy	0.025	–										
Hope Cain (24)	288	P	M	–	–	–	MRB	2	LR	TV	–	–	–	–	0.66(95% [0.56-0.76])	–
							MRB	2	SVM	TV	–	–	–	–	0.59 (95% [0.48-0.70])	–
	151*	P	M	–	–	–	MRB	4	LR	TV	–	–	–	–	0.71 (95% [0.58-0.83])	–
							MRB		SVM	TV	–	–	–	–	0.71 (95% [0.58-0.83])	–
Banerjee (25)	53	P	M	–	–	–	–	*	LASSO	KFCV	0.69 ± 0.03	–	–	–	n/a	n/a
								24	SVM	KFCV	0.74 ± 0.01	–	–	–	n/a	n/a
Chamming's (26)	85	P	U	Kurtosis (SFF = 2)	0.015	0.67	n/a	n/a	n/a	n/a	–	–	–	–	n/a	n/a
				Kurtosis (SFF = 4)	0.044											
				Kurtosis (SFF = 6)	0.019											
Giannini (27)	44	P	M	Contrast	< 0.05	0.722	BRM	2	LR	LOOCV	0.80 (95% [0.65-0.90])		80	69	n/a	n/a
				Correlation	< 0.05	0.715										
				Sum variance	< 0.05	0.674										
				Difference variance	< 0.05	0.699										
				Difference entropy	< 0.05	0.719	FA	6	Bayesian	LOOCV	–	0.70	67	72	n/a	n/a
				LRE	< 0.05	0.676										
				LRHGE	< 0.05	0.708										
Henderson (28)	88	P, M	U	ΔEntropy fine	0.006	0.834	n/a	n/a	n/a	n/a	–	–	–	–	n/a	n/a
				ΔEntropy coarse	0.006	0.845										
Fan (29)	103	P	M	ΔBPE*	–	0.713	EAB	12	LR	TV, LOOCV	0.91 (95% [0.80-1.00])	–	90	87.2	0.71 (95% [0.54-0.89])	–

(*continued on next page*)

Table 2 (*continued*)

Author	Number of patients	Time point MRI exam	Analysis strategy	Univariate feature analysis			Multivariate feature analysis									
				Significant features	p-value	AUC	Feature selection	Number of features	Model/ Classifier	Model validation	AUC (training)	Accuracy training	Sens (%)	Spec (%)	AUC (Validation)	Accuracy validation
Braman (30)	117	P	M	–	–	–	mRMR	8	LDA	TV, KFCV	0.75 ± 0.039	0.72	–	–	0.60	0.59
							mRMR	10	DLDA	TV, KFCV	0.78 ± 0.032	0.75	–	–	0.55	0.59
							mRMR	6	QDA	TV, KFCV	0.74 ± 0.037	0.76	–	–	0.64	0.64
							mRMR	10	Naïve Bayes	TV, KFCV	0.77 ± 0.021	0.78	–	–	0.69	0.64
							mRMR	10	SVM	TV, KFCV	0.71 ± 0.076	0.71	–	–	0.47	0.64
	70**	P	M	–	–	–	mRMR	10	*DLDA	TV, KFCV	0.83 ± 0.025	0.79	–	–	–	–
	47***	P	<	–	–	–	mRMR	10	Naïve Bayes	TV, KFCV	0.93 ± 0.018	0.84	–	–	–	–
Thibault (31)	38	P, F	U	–	[c]	–	n/a	n/a	n/a	n/a	–	–	–	–	n/a	n/a
Wu (32)	35	F	M	Contrast	< 0.05	0.76	n/a	4	LR	LOOCV	0.79 (95% [0.62-0.96])		75	78	n/a	–
				Correlation	< 0.05	0.80										
				Energy	< 0.05	0.76										
				Homogeneity	< 0.05	0.76										
Michoux (33)	69	P	M	Energy	0.002	0.702	–	2	LR	LOOCV	–	0.74	74	74	n/a	–
				Entropy	0.003	0.696										
				Homogeneity	0.001	0.701										
				Inverse difference moment	0.001	0.711										
				Difference variance	0.023	0.649		4	KMC	LOOCV	–	0.68	84	62	n/a	–
				Run percentage	0.045	0.640										
				HGRE	0.038	0.644										
				Wash-in	0.008	0.685										
Parikh (34)	36	P, M	U	ΔEntropy	0.003	0.84	n/a	n/a	n/a	n/a	–	–	–	–	n/a	n/a
				ΔUniformity	0.004	0.84										
Teruel (35)	58	P	U	Sum variance	0.019	0.689	n/a	n/a	n/a	n/a	–	–	–	–	n/a	n/a
				Sum entropy	0.021	0.686										
				Entropy	0.040	–										
				Difference variance	0.040	–										
Ahmed (36)	100	P	U	Contrast	0.039[a]		n/a	n/a	n/a	n/a	–	–	–	–	n/a	n/a
				Difference variance	0.039[b]											

Abbreviations: n/a = not applicable, - = not performed. *triple negative/HER2+ subgroup analysis, **HR+/HER2- subgroup analysis, ***TN/HER2+ subgroup analysis.
Time point MRI exam: P = pretreatment, M = midtreatment, F = after first cycle of neoadjuvant systemic therapy.
Significant features: ASM = angular second moment, SFF = spatial scaling factor, Δ between pretreatment and midtreatment, LRE = long run emphasis, LRHGE = long run high grey level emphasis, HGRE = high grey level run emphasis, *average difference of background parenchymal enhancement (BPE) between the breast with a breast tumor and the contralateral normal breast (best performing individual feature).
p-value: [a]1 min post-contrast, [b]2 min post-contrast, [c] Significant individual features could not be identified.
Feature selection: WRST = Wilcoxon rank sum test, LASSO = least absolute shrinkage and selection operator, UFA = Univariate feature selection, PCC = Pearson correlation coefficient, RFB = Random forest Boruta, MRB = multilinear regression-based features selection, BRM = backward regression method, FA = filter approach, EAB = Evolutionary Algorithm-based feature selection, mRMR = minimum redundancy maximum relevance feature selection.
Number of features: *unknown number of riesz and first-order features.
Model/Classifier: MLR = multivariate logistic regression, SVM = support vector machine, LR = logistic regression, LDA = linear discriminant analysis, DLDA = diagonal linear discriminant analysis, QDA = quadratic discriminant analysis, KMC = k-means clustering, *best performing classifier.
Model validation: TV = cohort split in two; training and validation, EV = external validation, KFCV = k-fold cross validation, LOOCV = leave-one-out cross-validation.
AUC training: MRS = Multiparametric radiomics signature, MCS = Multiparametric combined signature (radiomics and clinical features), HR = hormone receptor, HER2 = Human epidermal growth factor recepter, TN = triple negative.

with a radiomics model. This combination showed improved results when compared to the radiomics model only (AUC of 0.94 vs 0.83). The best performing model developed by the study of Liu et al. [22] also consisted of both clinical and radiomics features.

To validate the models, the majority of the studies performed leave-one-out cross-validation. Two studies divided their data into a training and a validation cohort [24,30], and two studies used an independent validation cohort from the same hospital [21,29]. Liu et al. [22] was the only study to externally validate their model, using three external validation cohorts.

3.4.5. Performance evaluation

Entropy was the most common significant denominator with a highest AUC value of 0.85. In total, 26 different features showed significance.

The AUC results of the multivariate analysis, for respective training and validation cohorts, ranged from 0.69 to 0.99 and 0.47 to 0.94. The multivariate model of the study of Xiong et al. [21] showed overall the best performing classifier with an AUC value of 0.94 in the validation cohort. According to the published results of all radiomics models, no identical features could be identified (Table 2).

4. Discussion

This systematic review provides an overview of studies published on the topic of MRI-based radiomics for tumor response prediction to NST in breast cancer. Overall, the studies showed large methodological heterogeneity in each step of the radiomics workflow. Differences mostly arise due to the use of different ROI segmentation methods (2D vs. 3D), varying response definitions, and differences in modeling strategy. Despite the radiomics workflow heterogeneities, the results seem promising. Entropy was the most statistically significant individual feature reported. Validation results for articles reporting multivariate analysis ranged from AUC of 0.47 to 0.94. No identical features were identified in the multivariate analysis, where logistic regression was the most frequently chosen model.

In this review, we applied the RQS checklist to study the quality and reporting of the radiomics workflows. The studies employing multivariate analysis reached a mean score of 20.9 % (0–41.2 %). The studies using univariate analysis scored 0 %, however, two items of the RQS (*i.e., feature reduction* and *multivariate analysis*) are not applicable to these studies, and hence no points were given. The RQS is also strongly influenced by the items *prospective research* and *validation cohort* since they account for respectively 7 and 5 points (*i.e.,* 33 % of the maximum score). Further, the RQS score only identifies whether steps are being carried out and reported and not how they are performed. However, it is of utmost importance to know how the different methodologies used in each step of the radiomics workflow have been performed and applied, since differing methodologies and parameter settings can result in heterogeneous outcomes.

The first step in the radiomics workflow after image acquisition is ROI segmentation. This is an essential step since all subsequent steps rely on the (quality of the) segmentation. Segmentations are either performed manually (which is time-consuming, labor-intensive, and prone to inter-and intra-observer variation procedure), semi-automatically (which is prone to variation due to manual contribution ([38])) or automatically (which still bears the disadvantage of needing to be checked by experienced radiologists ([39])). Recently, a study by Pandey et al. [40] showed promising results with a fast and fully automatic MRI breast tumor segmentation software that reached accuracies above 95 % comparable to manual segmentation. In this review, in ten studies segmentation was performed manually, five semi-automatically, and two automatically. In the publications using semi-automatic segmentation, four different methods were reported, which again indicates heterogeneity regarding tumor segmentation among the included studies.

Furthermore, the extent of the segmentation influences the results of the radiomics analysis. The analysis of the entire tumor, which is one of radiomics most attractive features, is only performed when 3D segmentation is applied. Previous research showed the advantages of 3D segmentation with respect to tumor heterogeneity compared to 2D segmentation [41,42]. Of the studies included in this review, eleven used 3D segmentation, although four of these studies [30,35,36] only used a few consecutive slices, not fully exploiting this advantage of radiomics analysis. Standardization of a 3D segmentation method will enhance the applicability of future research to larger cohorts.

So far, none of the segmentation approaches, either the extent of the segmentations, showed an association with higher AUC values.

The next step in the workflow which also showed great variability was feature selection and model development. Feature selection was performed differently in all studies, with Michoux et al. [33] not reporting on any feature selection details. Model overfitting seems to play a role in four studies [25,29,30,32] since the suggested need of at least ten patients for each feature in a model was exceeded [43]. It is important to exclude features that do not correlate with the outcome or highly correlate with other extracted features (*e.g.* volume) to reduce overfitting [19].

Model development was mostly performed by logistic regression. No identical features could be identified among the multivariate prediction models, to which the heterogeneity in previous steps probably contributed. A similar observation could be made for univariate feature analysis. The entropy feature demonstrated the most predictive ability for tumor response prediction to NST [23,27,28,33–35], but the lack of standardization of the preceding steps in the radiomics workflow did not allow for an overall conclusion.

The chosen outcome parameter for tumor response measurement varied among the studies. The majority of the studies used the commonly accepted gold standard: pathologic response assessment [44]. Only one study used the RECIST criteria based on imaging to quantify the response [29]. In the studies using pathologic assessment, response definitions varied, from no residual invasive cancer [22,24–28,30,31,34], to the consecutive addition of decrease in longest tumor diameter by > 50 % [36], of residual cancer burden below 1 cm [35], or of any partial response [23]. The differences in outcome parameters made it impossible to compare results.

Furthermore, additional issues need to be addressed. First, most studies lumped all breast cancer subtypes into one category. Previous research has shown that tumor response prediction to NST varied among subtypes indicating the need for specific radiomics models, or at least the need to incorporate subtype into the analysis [8,45]. Radiomics models incorporating or differentiating for breast cancer subtypes will ensure more robust and comparable results.

Second, the use of a multiparametric breast MRI approach adds additional information and yield more robust features [46]. Only the two most recently published studies [21,22] used this approach showing improved results. The majority of the studies opted for a single T1W contrast-enhanced imaging sequence [13]. However, a standard breast MRI protocol also consists of a T2W sequence and a DCE T1W sequence, sometimes combined with diffusion-weighted imaging and the derived apparent diffusion coefficient maps [47]. Future studies should further explore such multiparametric approaches, as they better reflect the clinical assessment performed by a radiologist [48,49].

Third, the lack of image pre-processing which is an important step in the radiomics workflow. Just over half of all studies applied and reported any method of image pre-processing. Image pre-processing ensures a more homogenous image quality (by interpolating all images to the same pixel size and slice spacing) [50,51]. Since feature values strongly depend on image quality it is of importance to perform this step, preferably in a standardized way.

Fourth, publication bias, a more general limitation where studies with less favorable results may not be published while these may nevertheless contribute to the radiomics analysis.

Fifth, since the field of radiomics is rapidly evolving, including the nomenclature, it might be possible that the search excluded eligible studies.

Lastly, for study inclusion, articles needed to have their full-text available through our institutional library. In our search there was one potentially eligible article of which the full-text was not available.

Twenty-six individual features showed significant results, with a mean AUC of 0.72 [range 0.65–0.85]. The mean AUC results for the multivariate models were 0.81 [range 0.70–0.99] and 0.72 [range 0.47–0.94] for the training and validation cohort, respectively. Despite encouraging results, similar to results from studies publishing on the same topic in different tumor sites [52,53], studies showed their concerns about clinical implementation. These concerns arose mostly from the methodological differences in the radiomics workflow, which also prevented us from performing a meta-analysis. Standardizing the methodologies used in the radiomics workflow will be the first step towards clinical implementation.

In this review, we aimed to perform a systematic overview of the methodologies used in the radiomics workflow and reported results in the field of MRI-based tumor response prediction to NST. Though limitations (including the heterogeneous radiomics workflow) currently plaguing this research field, results are still promising. Therefore we propose several recommendations which should be considered in designing future studies on radiomics research: [1] obliging image preprocessing performance, preferably in a standardized way [2]; consensus should be reached for both ROI segmentation method and the (pathologic) response definition, or at least they should be described in detail [3]; automatic 3D segmentation of the tumor lesion to improve feature stability [4]; future models can be improved by incorporating breast cancer subtypes information, by using multiparametric MRI, and by adding peritumoral regions [5]; the application of external model validations to ensure that models are not simply reflecting localized spurious correlations between features and outcomes; [6] extensive reporting of each consecutive step in the radiomics analysis to increase transparency and reproducibility;

To conclude, studies focusing on MRI-based radiomics for tumor response prediction to NST in breast cancer patients showed promising results despite large methodological heterogeneity in each step of the radiomics workflow. This review demonstrates the requirement for more standardized methodology in the radiomics workflow since it is important to achieve robust and reproducible results in future research in order to translate the results to clinical applications.

Dr. Woodruff has (minority) shares in the company OncoRadiomics. Dr Smidt received a grant of the company Servier for microbiome research. The rest of the authors declare no competing interest.

Declaration of Competing Interest

Dr. Woodruff has (minority) shares in the company OncoRadiomics. Dr Smidt received a grant of the company Servier for microbiome research. The rest of the authors declare no competing interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrad.2019.108736.

References

[1] S. Loibl, C. Denkert, G. von Minckwitz, Neoadjuvant treatment of breast cancer–Clinical and research perspective, Breast 24 (Suppl. 2) (2015) S73–7.
[2] M. Teshome, K.K. Hunt, Neoadjuvant therapy in the treatment of breast cancer, Surg. Oncol. Clin. N. Am. 23 (3) (2014) 505–523.
[3] P.E.R. Spronk, K.M. de Ligt, A.C.M. van Bommel, S. Siesling, C.H. Smorenburg, M. Vrancken Peeters, et al., Current decisions on neoadjuvant chemotherapy for early breast cancer: experts' experiences in the Netherlands, Patient Educ. Couns. (2018).
[4] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, et al., Dermatologist-level classification of skin cancer with deep neural networks, Nature. 542 (7639) (2017) 115–118.
[5] A.M. Thompson, S.L. Moulder-Thompson, Neoadjuvant treatment of breast cancer, Ann. Oncol. 23 (Suppl. 10) (2012) x231–6.
[6] J.P. O'Connor, C.J. Rose, J.C. Waterton, R.A. Carano, G.J. Parker, A. Jackson, Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome, Clin. Cancer Res. 21 (2) (2015) 249–257.
[7] H. Bonnefoi, S. Litiere, M. Piccart, G. MacGrogan, P. Fumoleau, E. Brain, et al., Pathological complete response after neoadjuvant chemotherapy is an independent predictive factor irrespective of simplified breast cancer intrinsic subtypes: a landmark and two-step approach analyses from the EORTC 10994/BIG 1-00 phase III trial, Ann. Oncol. 25 (6) (2014) 1128–1136.
[8] P. Cortazar, L. Zhang, M. Untch, K. Mehta, J.P. Costantino, N. Wolmark, et al., Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis, Lancet. 384 (9938) (2014) 164–172.
[9] T.M. Prowell, R. Pazdur, Pathological complete response and accelerated drug approval in early breast cancer, N. Engl. J. Med. 366 (26) (2012) 2438–2441.
[10] N. Houssami, R. Turner, M. Morrow, Preoperative magnetic resonance imaging in breast cancer: meta-analysis of surgical outcomes, Ann. Surg. 257 (2) (2013) 249–255.
[11] N. Hylton, MR imaging for assessment of breast cancer response to neoadjuvant chemotherapy, Magn. Reson. Imaging Clin. N. Am. 14 (3) (2006) 383–389 vii.
[12] M.B. Lobbes, R. Prevos, M. Smidt, V.C. Tjan-Heijnen, M. van Goethem, R. Schipper, et al., The role of magnetic resonance imaging in assessing residual disease and pathologic complete response in breast cancer patients receiving neoadjuvant chemotherapy: a systematic review, Insights Imaging 4 (2) (2013) 163–175.
[13] R. Prevos, M.L. Smidt, V.C. Tjan-Heijnen, M. van Goethem, R.G. Beets-Tan, J.E. Wildberger, et al., Pre-treatment differences and early response monitoring of neoadjuvant chemotherapy in breast cancer patients using magnetic resonance imaging: a systematic review, Eur. Radiol. 22 (12) (2012) 2607–2616.
[14] K. Wasser, S.K. Klein, C. Fink, H. Junkermann, H.P. Sinn, I. Zuna, et al., Evaluation of neoadjuvant chemotherapeutic response of breast cancer using dynamic MRI with high temporal resolution, Eur. Radiol. 13 (1) (2003) 80–87.
[15] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R.G. van Stiphout, P. Granton, et al., Radiomics: extracting more information from medical images using advanced feature analysis, Eur. J. Cancer 48 (4) (2012) 441–446.
[16] H.J. Aerts, E.R. Velazquez, R.T. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, et al., Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, Nat. Commun. 5 (2014) 4006.
[17] F. Davnall, C.S. Yip, G. Ljungqvist, M. Selmi, F. Ng, B. Sanghera, et al., Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? Insights Imaging 3 (6) (2012) 573–589.
[18] P. Lambin, R.T.H. Leijenaar, T.M. Deist, J. Peerlings, E.E.C. de Jong, J. van Timmeren, et al., Radiomics: the bridge between medical imaging and personalized medicine, Nat. Rev. Clin. Oncol. 14 (12) (2017) 749–762.
[19] A. Ibrahim, M. Vallières, H. Woodruff, S. Primakov, M. Beheshti, S. Keek, et al., Radiomics analysis for clinical decision support in nuclear medicine, Semin. Nucl. Med. (2019).
[20] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, P. Group, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, Int. J. Surg. 8 (5) (2010) 336–341.
[21] Q. Xiong, X. Zhou, Z. Liu, C. Lei, C. Yang, M. Yang, et al., Multiparametric MRI-based radiomics analysis for prediction of breast cancers insensitive to neoadjuvant chemotherapy, Clin. Transl. Oncol. (2019).
[22] Z. Liu, Z. Li, J. Qu, R. Zhang, X. Zhou, L. Li, et al., Radiomics of multi-parametric MRI for pretreatment prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer: a multicenter study, Clin. Cancer Res. (2019).
[23] H.J. Yoon, Y. Kim, J. Chung, B.S.R. Kim, Predicting neo-adjuvant chemotherapy response and progression-free survival of locally advanced breast cancer using textural features of intratumoral heterogeneity on F-18 FDG PET/CT and diffusion-weighted MR imaging, Breast J. (2018).
[24] E.H. Cain, A. Saha, M.R. Harowicz, J.R. Marks, P.K. Marcom, M.A. Mazurowski, Multivariate machine learning models for prediction of pathologic response to neoadjuvant therapy in breast cancer using MRI features: a study using an independent validation set, Breast Cancer Res. Treat. (2018).
[25] I. Banerjee, S. Malladi, D. Lee, A. Depeursinge, M. Telli, J. Lipson, et al., Assessing treatment response in triple-negative breast cancer from quantitative image analysis in perfusion magnetic resonance imaging, J. Med. Imaging Bellingham (Bellingham) 5 (1) (2018) 011008.
[26] F. Chamming's, Y. Ueno, R. Ferre, E. Kao, A.S. Jannot, J. Chong, et al., Features from computerized texture analysis of breast cancers at pretreatment MR imaging are associated with response to neoadjuvant chemotherapy, Radiology 286 (2) (2018) 412–420.
[27] V. Giannini, S. Mazzetti, A. Marmo, F. Montemurro, D. Regge, L.R. Martincich, A computer-aided diagnosis (CAD) scheme for pretreatment prediction of pathological response to neoadjuvant therapy using dynamic contrast-enhanced MRI texture features, Br. J. Radiol. 90 (1077) (2017) 20170269.
[28] S. Henderson, C. Purdie, C. Michie, A. Evans, R. Lerski, M. Johnston, et al., Interim heterogeneity changes measured using entropy texture features on T2-weighted MRI at 3.0 T are associated with pathological response to neoadjuvant chemotherapy in primary breast cancer, Eur. Radiol. 27 (11) (2017) 4602–4611.
[29] M. Fan, G. Wu, H. Cheng, J. Zhang, G. Shao, L.R. Li, Radiomic analysis of DCE-MRI for prediction of response to neoadjuvant chemotherapy in breast cancer patients, Eur. J. Radiol. 94 (2017) 140–147.
[30] N. Braman, M. Etesami, P. Prasanna, C. Dubchuk, H. Gilmore, P. Tiwari, et al.,

Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI, Breast Cancer Res Journal Translated Name Breast Cancer Research. 19 (1) (2017) no pagination.
[31] G. Thibault, A. Tudorica, A. Afzal, S.Y. Chui, A. Naik, M.L. Troxell, et al., DCE-MRI texture features for early prediction of breast Cancer therapy response, Tomography 3 (1) (2017) 23–32.
[32] J. Wu, G.H. Gong, Y. Cui, R.J.R. Li, Intratumor partitioning and texture analysis of dynamic contrast-enhanced (DCE)-MRI identifies relevant tumor subregions to predict pathological response of breast cancer to neoadjuvant chemotherapy, J. Magn. Reson. Imaging 44 (5) (2016) 1107–1115.
[33] N. Michoux, S. Van den Broeck, L. Lacoste, L. Fellah, C. Galant, M. Berliere, et al., Texture analysis on MR images helps predicting non-response to NAC in breast cancer, BMC Cancer 15 (2015) 574.
[34] J. Parikh, M. Selmi, G. Charles-Edwards, J. Glendenning, B. Ganeshan, H. Verma, et al., Changes in primary breast cancer heterogeneity may augment midtreatment MR imaging assessment of response to neoadjuvant chemotherapy, Radiology. 272 (1) (2014) 100–112.
[35] J.R. Teruel, M.G. Heldahl, P.E. Goa, M. Pickles, S. Lundgren, T.F. Bathen, et al., Dynamic contrast-enhanced MRI texture analysis for pretreatment prediction of clinical and pathological response to neoadjuvant chemotherapy in patients with locally advanced breast cancer, NMR Biomed. 27 (8) (2014) 887–896.
[36] A. Ahmed, P. Gibbs, M. Pickles, L.R. Turnbull, Texture analysis in assessment and prediction of chemotherapy response in breast cancer, J. Magn. Reson. Imaging 38 (1) (2013) 89–101.
[37] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Trans. Syst. Man Cybern. SMC-3 (6) (1973) 610–621.
[38] T. Heye, E.M. Merkle, C.S. Reiner, M.S. Davenport, J.J. Horvath, S. Feuerlein, et al., Reproducibility of dynamic contrast-enhanced MR imaging. Part II. Comparison of intra- and interobserver variability with manual region of interest placement versus semiautomatic lesion segmentation and histogram analysis, Radiology 266 (3) (2013) 812–821.
[39] I.E. van Dam, J.R. van Sornsen de Koste, G.G. Hanna, R. Muirhead, B.J. Slotman, S. Senan, Improving target delineation on 4-dimensional CT scans in stage I NSCLC using a deformable registration tool, Radiother. Oncol. 96 (1) (2010) 67–72.
[40] D. Pandey, X. Yin, H. Wang, M.Y. Su, J.H. Chen, J. Wu, et al., Automatic and fast segmentation of breast region-of-interest (ROI) and density in MRIs, Heliyon. 4 (12) (2018) e01042.
[41] X. Fave, M. Cook, A. Frederick, L.F. Zhang, J.Z. Yang, D. Fried, et al., Preliminary investigation into sources of uncertainty in quantitative imaging features, Comput. Med. Imaging Graph. 44 (2015) 54–61.
[42] F. Ng, R. Kozarski, B. Ganeshan, V. Goh, Assessment of tumor heterogeneity by CT texture analysis: can the largest cross-sectional area be used as an alternative to whole tumor analysis? Eur. J. Radiol. 82 (2) (2013) 342–348.
[43] R. Gillies, P. Kinahan, H. Hricak, Radiomics: Images are more than pictures, they are data, Radiology 278 (2) (2016) 563–577.
[44] A.M. Fowler, D.A. Mankoff, B.N. Joe, Imaging neoadjuvant therapy response in breast Cancer, Radiology 285 (2) (2017) 358–375.
[45] G. von Minckwitz, M. Untch, J.U. Blohmer, S.D. Costa, H. Eidtmann, P.A. Fasching, et al., Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes, J. Clin. Oncol. 30 (15) (2012) 1796–1804.
[46] D. Truhn, S. Schrading, C. Haarburger, H. Schneider, D. Merhof, C. Kuhl, Radiomic versus convolutional neural networks analysis for classification of contrast-enhancing lesions at multiparametric breast MRI, Radiology (2018) 181352.
[47] R.M. Mann, C. Balleyguier, P.A. Baltzer, U. Bick, C. Colin, E. Cornford, et al., Breast MRI: EUSOBI recommendations for women's information, Eur. Radiol. 25 (12) (2015) 3669–3678.
[48] V.S. Parekh, M.A. Jacobs, Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI, NPJ Breast Cancer 3 (2017) 43.
[49] J. Peerlings, H.C. Woodruff, J.M. Winfield, A. Ibrahim, B.E. Van Beers, A. Heerschap, et al., Stability of radiomics features in apparent diffusion coefficient maps from a multi-centre test-retest trial, Sci. Rep. 9 (1) (2019) 4800.
[50] P. Lambin, Radiomics: The bridge between medical imaging and personalized medicine, Nat Rev Clin Oncol Journal Translated Name Nature Reviews Clinical Oncology. 14 (12) (2017) 749–762.
[51] A. Jethanandani, T.A. Lin, S. Volpe, H. Elhalawani, A.S.R. Mohamed, P. Yang, et al., Exploring applications of radiomics in magnetic resonance imaging of head and neck Cancer: a systematic review, Front. Oncol. 8 (2018) 131.
[52] N. Horvat, D.D.B. Bates, I. Petkovska, Novel imaging techniques of rectal cancer: what do radiomics and radiogenomics have to offer? A literature review, Abdom Radiol (NY). (2019).
[53] A. Chaddad, M.J. Kucharczyk, P. Daniel, S. Sabri, B.J. Jean-Claude, T. Niazi, et al., Radiomics in glioblastoma: current status and challenges facing clinical implementation, Front. Oncol. 9 (2019) 374.