Int J Clin Exp Med 2020;13(10):7684-7690
www.ijcem.com /ISSN:1940-5901/IJCEM0115539

Original Article

An analysis of hepatic steatosis based on radiomics and deep learning

Zheng Zhang[1], Renbin Ge[2], Lei Zhang[2], Chi Zhang[1], Xinghui Li[1], Zhongze Gu[1]

[1]*School of Biological Sciences and Medical Engineering, Southeast University, Nanjing, Jiangsu, China;* [2]*Department of Radiology, Shanghai General Hospital, Shanghai Jiaotong University, Shanghai, China*

Received June 3, 2020; Accepted July 19, 2020; Epub October 15, 2020; Published October 30, 2020

Abstract: Objective: This study explores hepatic steatosis grading through imaging and deep learning in order to achieve an automatic classification of hepatic steatosis, therefore providing an intelligent and non-invasive method of hepatic steatosis classification. Methods: A retrospective study was performed on adult patients who underwent MR imaging of the upper abdomen with the mDixon sequence at Shanghai First People's Hospital from June to August 2018. The MRI data and materials from the regions of interest were collected. Cases of mild to moderate hepatic steatosis were graded using radiomics and deep learning. A Spearman correlation analysis was performed comparing the clinical information and the liver steatosis grades. Results: The highest AUC obtained on the validation cohort was 0.81, and the AUC of the deep learning model was 0.70. Conclusion: Radiomics and deep learning can aid in automatic grading for hepatic steatosis. Correlation analysis also provides useful ideas for hepatic steatosis grading.

Keywords: Radiomics, deep learning, hepatic steatosis

Introduction

With the development of China's economy, people's living standards are elevated, and their dietary habits have gradually changed, the common metabolic abnormalities associated with obesity. The long-term intake of high-sugar and high-fat foods causes an excessive accumulation of lipid substances and fat deposits in certain organs and tissues, such as the liver and pancreas, causing ectopic lipid deposition [1] and fatty degeneration. As the central organ of energy metabolism, the liver plays an important role in metabolism. Hepatic steatosis affects the physiological function of the liver and further aggravates the abnormal lipid metabolism. Varying degrees of steatosis can cause many diseases, such as fatty liver, liver fibrosis, cirrhosis, and even liver cancer [2]. Often these are not irreversible processes, so it is of great clinical significance to explore mild to moderate steatosis.

The histological examination is the gold standard for diagnosing the degree of liver steatosis, but it is invasive, and the potential risks of infection and bleeding at the puncture site cannot be ignored, and there are shortcomings such as sampling and interpretation errors and poor repeatability [3, 4]. At present, ultrasound imaging is mainly used in clinical examinations, but it has a maximum sensitivity of 55% when the fat content is less than 20%, and it provides an insufficient penetration depth for obese patients. CT examinations determine liver fatty degeneration mainly by measuring the liver attenuation. Owing to the limited accuracy in the diagnosis of mild hepatic steatosis [5], a more accurate method should be developed. CT is also radioactive. On basis of the difference in the precession frequency of hydrogen protons in fat and water molecules under different magnetic field strengths, MRT can more accurately measure the volume fraction of fat in the liver, so that it can be more reliable for mild liver fatty degeneration.

According to the results of a national survey [6], the overall prevalence of dyslipidemia in Chinese adults is as high as 40.40%, a signifi-

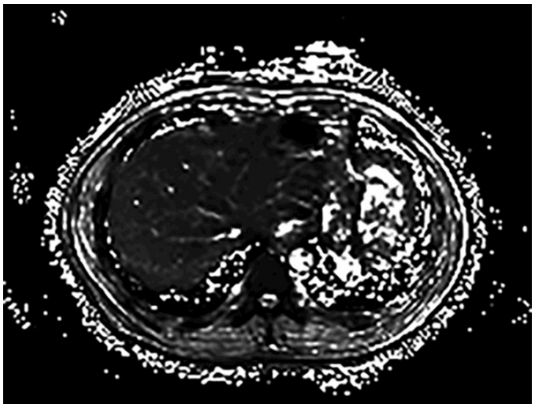

Figure 1. MR upper abdomen mDixon imaging.

cant increase from 2002. A study [7] shows that the increase of serum cholesterol will lead to an increase of about 9.2 million cardiovascular patients in China during the period 2010-2030.

Continuous improvements in new hardware and standardized protocols in the imaging field have advanced medical big data technology, and radiomics [8, 9] and deep learning have been increasingly integrated into medical imaging. Through the in-depth analysis of patient images, more potential and effective information can be tapped.

In this study, radiomics and deep learning methods were used to study the region of interest from patients' abdominal MRI images, and their liver steatosis was graded, and then the differences between the results of two methods were compared to provide more reliable means and ideas for the study of hepatic steatosis.

Materials and methods

Patient data

This study collaborated with the First People's Hospital of Shanghai Jiaotong University to collect the medical records and data of patients who underwent MR imaging of the upper abdomen with the mDixon sequence from June 2018 to August 2018 (**Figure 1**). Inclusion criteria: Patients over 18 years old; patients able to provide a complete personal history and complete relevant imaging and laboratory examinations. Exclusion criteria: Patients with other liver diseases who are undergoing chemical and radiotherapy for malignant tumors. A total of 50 patients were included in this study, including 38 patients with mild hepatic steatosis and 12 patients with moderate hepatic steatosis.

MR equipment parameters

The Philips Ingenia 1.5T MRI system was adopted with the following sequence acquisition parameters: the number of echoes = 6; TE = n × 1.15 ms, N = 1, ..., 6; flip angle = 5°; TR = 10.3 ms; FOV = 400 × 350 × 210 mm^3; acquisition matrix = 320 × 256 × 70, the scan time was 16 s; breath holding. The receiving coils were 32-channel surface coils covering the entire abdomen.

Data preprocessing

The data preprocessing mainly includes ROI (region of interest) extraction and data amplification.

ROI extraction. 0 points for the percentage of fat within the hepatocytes ranging 0 to 5%; 1 point for 5% to 33%; 2 points for 34% to 66%, and 3 points for 66% or more. The distribution of the liver steatosis grades in the 50 patients includes 38 cases of 0 points (24 males and 14 females) and 12 cases of 1 point (6 males and 6 females).

3D slicer software was used to outline the ROI (region of interest) from each patient's image. The ROI was positioned at the hepatic parenchyma, avoiding large blood vessels, focal lesions, and significant artifacts. 6 ROIs were selected for each patient, of which 4 were located in the right lobe parenchyma (segments IV, VI, VII, and VIII), and 2 were located in the left lobe parenchyma (segments II and III). The size of each ROI was 16 × 16 pixels.

Data amplification. According to ROI selection rules, a total of 300 ROIs were selected from the 50 patients. The small sample size in this study may cause problems such as over-fitting in the deep learning models, so the number of samples was increased through data amplification. In this study, the data amplification methods include translation, mirroring, rotation, adding random noise, etc., to amplify the data by 8 times.

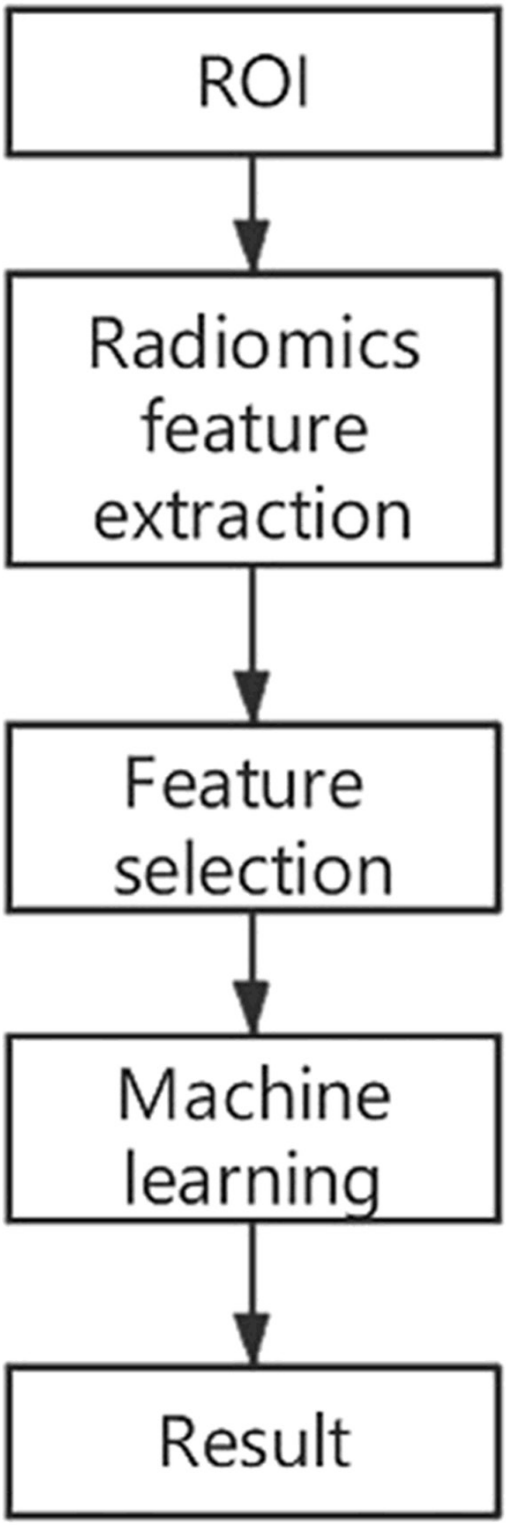

Figure 2. Radiomics process.

Radiomics methods

The radiomics process used in this study (**Figure 2**) mainly includes: (1) The extraction of radiomics features (2) Feature processing (3) Machine learning modeling and classification.

Feature extraction

Radiomics features, including intensity features and texture features, were extracted from 6 ROIs of each patient. Intensity features describe the first-order characteristic values of the voxel intensity in the ROI area, including the average, maximum, and minimum values. The texture features include the gray level co-occurrence matrix, the gray scale run length matrix, the gray level size zone matrix, the gray tone difference matrix, and the gray correlation matrix. A total of 126 features were extracted using Python, including 18 intensity features and 108 texture features.

Feature processing

The extracted features vary in their dimension and value ranges, so all the features need to be processed in advance, including feature selection [10] and normalized processing.

The purpose of feature selection is to reduce the amount of useless calculations in the modeling process by eliminating overly redundant features. There is no modification of the feature values in the feature selection, and more emphasis is placed on finding a few features that greatly improve the performance of the model. Since two-dimensional image data is explored in this study, some feature vectors describing three-dimensional features can be eliminated, such as flatness and so on. After the feature selection, 27 features were removed, and 93 features remained.

The normalization process mainly deals with the difference between the maximum and minimum values of the different features. If the gap in a feature value is too large, it will induce unnecessary feature dimension calculations during the training of the classification model. After the feature processing, the data set was prepared for the training of the machine learning classification model.

Machine learning classification

After the features are processed, a machine learning model needs to be established to classify them. This study uses integrative learning to guide the machine learning. Compared with the traditional machine learning classification

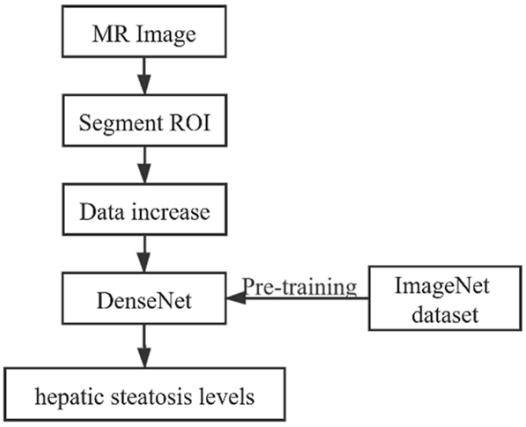

Figure 3. Deep learning process.

methods, integrative learning generates multiple learners through specific rules and then combines these learners to obtain the final result. Multiple learners in integrated learning are generally homogeneous "weak learners". Multiple learners are generated through a disturbance sample set, input disturbance, output disturbance, algorithm parameter disturbance, etc., from weak learners. After integration, a "strong learner" with better accuracy is obtained, which is the final integrative learning model.

The integrative learning methods used in this study include: AdaBoost [11] (Adaptive Boosting), GBDT [12] (Gradient Boosting Decision Tree), and XGBoost [13] (Extreme Gradient Boosting).

Deep learning methods

The DenseNet model: The deep learning process in this study is shown in **Figure 3**, where the deep learning model is DenseNet (Dense Convolutional Network). DenseNet won the top award at CVPR2017. It is different from the method of deepening the number of network layers, such as ResNet, and the method of widening the network structure is represented by Inception to improve the performance. DenseNet facilitates feature reuse in each layer, which not only greatly reduces the amount of network parameters but also alleviates the problem of gradient disappearance.

DenseNet is mainly composed of two parts: a dense block and a transition layer, and each layer of dense block is associated with all the previous layers [14]. In the traditional convolutional neural network, the L-layer network has L connections, but in DenseNet, there are L (L+)/2 connections. The input of each layer comes from the output of all previous layers, namely feature reuse. A transition layer was set up between the different dense blocks to achieve downsampling. The transition layer generally consists of a batch normalization layer, a convolution layer, and an average pooling layer (**Figure 4**).

Statistics

The statistical analyses were performed using SPSS 25.0. $P < 0.05$ indicated statistical significance, and all the tests were two sided. Continuous variables were expressed as the average ± standard deviation, and categorical variables were presented as *n* (%). The correlation analysis adopts the Pearson correlation analysis method.

Results

Classification performance

This study used two indicators to evaluate the classification results of the radiomics and deep learning methods: AUC (area under the ROC curve) and ACC (accuracy). AUC is an evaluation indicator to measure the pros and cons of the binary classification model. It represents the probability that the predicted positive example will be ranked before the negative example.

$$ACC = \frac{TN + TP}{TN + TP + FN + FP}$$

TP (true positive) represents the number of positive samples predicted by the model as positive, FP (false positive) represents the number of negative samples predicted as positive by the model, and FN (false negative) represents the number of positive samples predicted by the model as negative, TN (true negative) represents the number of negative samples predicted to be negative by the model.

Classification evaluation

The classification results of the two methods were used to calculate the performance index (**Table 1**).

Correlation analysis

SPSS 24.0 statistical software was used to analyze the baseline data. Spearman correla-

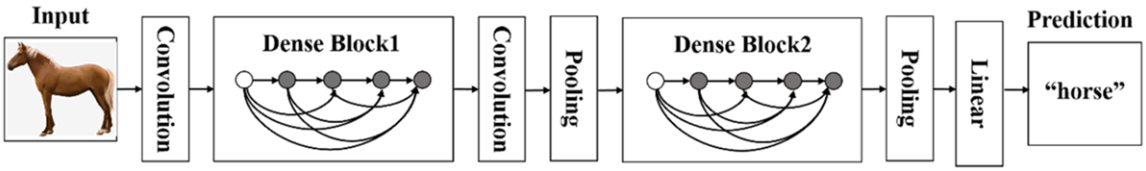

Figure 4. DenseNet.

Table 1. Performance index of the classification using the two methods

		AUC (95% CI)		ACC (95% CI)	
		Training set	Test set	Training set	Test set
Radiomics	AdaBoost	0.79	0.72	78.60	71.33
	GBDT	0.73	0.70	76.80	67.12
	XGBoost	0.84	0.81	81.94	78.26
Deep learning	DenseNet	0.76	0.70	64.77	66.27

Table 2. The correlation coefficients between the baseline data and the liver steatosis grades

Baseline data	Correlation coefficient	P
Age	0.003	0.959
Gender	0.035	0.570
Pancreatic steatosis grading	0.702	0.000
Metabolic syndrome	0.890	0.147

tion analyses were performed on the liver steatosis grading and baseline data including age, gender, pancreatic steatosis grading, and metabolic syndrome. The same method was used to correlate the clinical information of the patients with the grading of hepatic steatosis. $P < 0.05$ indicated that a difference was statistically significant.

The correlations between the clinical information and the hepatic steatosis grades are shown in **Table 2**. There is a high correlation between the liver steatosis grades and the pancreatic steatosis grades. We found that the liver steatosis grades in the female patients had a significant correlation with age ($P = 0.002$) and metabolic syndrome ($P = 0.007$).

Discussion

Lipids participate in the regulation of energy conversion, material transport, signal recognition and transmission, cell development and differentiation, and cell apoptosis [15, 16]. The liver, the central organ of energy metabolism, plays an important role in regulating lipid metabolism, including lipogenesis and lipoprotein uptake and secretion. Abnormal lipid metabolism will lead to an abnormal increase in lipids in liver tissue, such as increased lipid peroxide and free fatty acid (FFA), resulting in lipid accumulation and liver fat degeneration. Hepatic steatosis results from fat accumulation. Severe hepatic steatosis causes irreversible damage to the liver. For example, hepatic stellate cell damage will develop into liver fibrosis. Therefore, the early diagnosis of hepatic steatosis has an important clinical significance.

Radiomics and deep learning methods were used to classify the patients with mild to moderate hepatic steatosis, aiming to carry out an automated, intelligent screening of mild steatosis to achieve an early and accurate diagnosis. Three different machine learning methods was used for the radiomics classification, and the highest AUC of the validation set reached 0.81. As for the deep learning, the highest AUC was 0.70. The accuracy of the deep learning methods is low, which may be caused by the small sample size. Although we performed a series of data amplifications, the small number of initial sample sets reduced the model's generalizability.

Compared with the traditional analysis that only relies on the intensity of the voxel values in the image, radiomics further explores some deep-level features in the image [17, 18] by analyzing the imaging significance of these features, it aids a more comprehensive imaging study of liver steatosis [19, 20].

At present, the imaging methods for studying liver steatosis are mainly ultrasound and CT, both of which are less sensitive to mild to mod-

erate steatosis, which could be resolved by using MRI and radiomics.

Our correlation analysis showed that the liver steatosis grade and the pancreatic steatosis grade had a significant correlation. When too much fat tissue is deposited in the pancreas parenchyma, pancreatic fatty infiltration, also known as fatty pancreas or pancreatic fat degeneration, occurs. This correlation indicates that liver fat accumulation and pancreatic fat accumulation show a certain synchronization.

Metabolic syndrome is caused by the abnormal metabolism of lipids and sugars, or by abnormal metabolic pathways [19, 21, 22]. Hepatic steatosis is associated with metabolic syndrome. For example, the risk factors for nonalcoholic fatty liver include insulin resistance, obesity, hypertension, dyslipidemia and type 2 diabetes metabolic syndrome [20, 23]. Through a correlation analysis, it can be seen that the grade of liver steatosis in female patients has a significant correlation with age and metabolic syndrome, but the liver steatosis grade in male patients has no relationship with these factors. Since this study only deals with grade 0 and 1 hepatic steatosis, the correlation between the moderate and severe hepatic steatosis grades and metabolic syndrome may be more significant.

Few studies have assessed the utility of ultrasound in quantifying steatosis, probably because of the complexity of computer-assisted image analysis techniques [11, 20, 21] and interobserver and interoperator variability [22]. There are still some shortcomings to this study. Due to the small sample size used in the early stage, the advantages of radiomics and deep learning were not fully realized. Subsequent studies will continue to collect the abdominal MRI data of patients with varying degrees of liver steatosis, further improving the classification accuracy, and mining more reliable and effective information in liver steatosis imaging data to provide a reference for imaging examination. In addition, this study only analyzed the correlation between age, gender, pancreatic steatosis grade and metabolic syndrome and liver steatosis grade. The relationships among blood lipids, blood glucose, blood pressure, and hepatic steatosis need further exploration.

In summary, liver steatosis grading based on radiomics and deep learning has the advantages of good performance and automated grading, and it can mine deeper and more effective information in medical imaging. A Spearman coefficient correlation analysis provides insights for the correlation between liver steatosis grading and the related factors such as metabolic syndrome.

Disclosure of conflict of interest

None.

Address correspondence to: Zhongze Gu, School of Biological Sciences and Medical Engineering, Southeast University, No. 2, Sipai Building, Xuanwu District, Nanjing 210000, Jiangsu, China. Tel: +86-17317629027; E-mail: guzhz027@163.com

References

[1] Després JP. Abdominal obesity and cardiovascular disease: is inflammation the missing link? Can J Cardiol 2012; 28: 642-652.

[2] Angrish MM, Kaiser JP, McQueen CA and Chorley BN. Tipping the balance: hepatotoxicity and the 4 apical key events of hepatic steatosis. Toxicol Sci 2016; 150: 261-268.

[3] Bonekamp S, Tang A, Mashhood A, Wolfson T, Changchien C, Middleton MS, Clark L, Gamst A, Loomba R and Sirlin CB. Spatial distribution of MRI-determined hepatic proton density fat fraction in adults with nonalcoholic fatty liver disease. J Magn Reson Imaging 2014; 39: 1525-1532.

[4] Merriman RB, Ferrell LD, Patti MG, Weston SR, Pabst MS, Aouizerat BE and Bass NM. Correlation of paired liver biopsies in morbidly obese patients with suspected nonalcoholic fatty liver disease. Hepatology 2006; 44: 874-880.

[5] Castera L, Vilgrain V and Angulo P. Noninvasive evaluation of NAFLD. Nat Rev Gastroenterol Hepatol 2013; 10: 666-675.

[6] National Center for Disease Control and Prevention NHaFPC. Nutrition and Chronic Diseases of Chinese Residents (2015). Beijing: People's Medical Publishing House; 2015.

[7] Moran A, Gu D, Zhao D, Coxson P, Wang YC, Chen CS, Liu J, Cheng J, Bibbins-Domingo K, Shen YM, He J and Goldman L. Future cardiovascular disease in china: markov model and risk factor scenario projections from the coronary heart disease policy model-china. Circ Cardiovasc Qual Outcomes 2010; 3: 243-252.

[8] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A and Aerts

HJ. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012; 48: 441-446.
[9] van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S and Aerts HJWL. Computational radiomics system to decode the radiographic phenotype. Cancer Res 2017; 77: e104-e107.
[10] Tuv E, Borisov A, Runger G and Torkkola K. Feature selection with ensembles, artificial variables, and redundancy elimination. J Mach Learn Res 2009; 10: 1341-1366.
[11] Cao Y, Miao QG, Liu JC and Gao L. Advance and prospects of AdaBoost algorithm. Zidonghua Xuebao 2013; 39: 745-758.
[12] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001; 29: 1189-1232.
[13] Chen T and Guestrin C. XGBoost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2016; 785-794.
[14] Di Paolo G and De Camilli P. Phosphoinositides in cell regulation and membrane dynamics. Nature 2006; 443: 651-657.
[15] Huang G, Liu Z, Maaten LVD and Weinberger KQ. Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017; 2261-2269.
[16] Baumruker T, Bornancin F and Billich A. The role of sphingosine and ceramide kinases in inflammatory responses. Immunol Lett 2005; 96: 175-185.
[17] Gillies RJ, Kinahan PE and Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016; 278: 563-577.
[18] Zwanenburg A, Leger S, Vallières M and Lck S; Initiative ftIBS. Image biomarker standardisation initiative. Arxiv Preprint 2016.
[19] Fan JG and Farrell GC. Epidemiology of non-alcoholic fatty liver disease in China. J Hepatol 2009; 50: 204-210.
[20] Angulo P. GI epidemiology: nonalcoholic fatty liver disease. Aliment Pharmacol Ther 2007; 25: 883-889.
[21] Molenaar MR, Vaandrager AB and Helms JB. Some lipid droplets are more equal than others: different metabolic lipid droplet pools in hepatic stellate cells. Lipid Insights 2017; 10: 1178635317747281.
[22] Singh RG, Yoon HD, Wu LM, Lu J, Plank LD and Petrov MS. Ectopic fat accumulation in the pancreas and its clinical relevance: asystematic review, meta-analysis, and meta-regression. Metabolism 2017; 69: 1-13.
[23] Kühn JP, Meffert P, Heske C, Kromrey ML, Schmidt CO, Mensel B, Völzke H, Lerch MM, Hernando D, Mayerle J and Reeder SB. Prevalence of fatty liver disease and hepatic iron overload in a northeastern german population by using quantitative MR imaging. Radiology 2017; 284: 706-716.