Clinical practice guideline for body composition assessment based on upper abdominal magnetic resonance images annotated using artificial intelligence

Han Lv[1], Mengyi Li[2], Zhenchang Wang[1], Dawei Yang[1], Hui Xu[1], Juan Li[1], Yang Liu[2], Di Cao[1], Yawen Liu[3], Xinru Wu[1], He Jin[1], Peng Zhang[2], Liqin Zhao[4], Rixing Bai[5], Yunlong Yue[6], Bin Li[6], Nengwei Zhang[7], Mingzhu Zou[8], Jinghai Song[9], Weibin Yu[10], Pin Zhang[11], Weijun Tang[12], Qiyuan Yao[13], Liheng Liu[14], Hui Yang[15], Zhenghan Yang[1], Zhongtao Zhang[2], National Clinical Research Center for Digestive Diseases in Beijing Friendship Hospital, Capital Medical University; Beijing Quality Control and Improvement Center of Medical Imaging; National Health Commission Capacity Building and Continuing Education Center, Chinese College of Metabolic and Bariatric Surgeons; Beijing Quality Control and Improvement Center of Metabolic and Bariatric Surgery

[1]Department of Radiology, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China;

[2]Department of General Surgery, Beijing Friendship Hospital, Capital Medical University & National Clinical Research Center for Digestive Diseases, Beijing 100050, China;

[3]School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China;

[4]Department of Radiology, Beijing Tiantan Hospital, Capital Medical University, Beijing 100070, China;

[5]Department of General Surgery, Beijing Tiantan Hospital, Capital Medical University, Beijing 100070, China;

[6]Department of MR, Beijing Shijitan Hospital, Capital Medical University/Peking University, Ninth Clinical Medical College, Beijing 100038, China;

[7]Department of General Surgery, Beijing Shijitan Hospital, Capital Medical University/Peking University, Ninth Clinical Medical College, Beijing 100038, China;

[8]Department of Radiology, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing 100730, China;

[9]Department of General Surgery, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing 100730, China;

[10]Department of Radiology, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai 200233, China;

[11]Department of Bariatric and Metabolic Surgery, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai 200233, China;

[12]Department of Radiology, Huashan Hospital, Fudan University, Shanghai 200040, China;

[13]Center for Obesity and Metabolic Surgery, Huashan Hospital, Fudan University, Shanghai 200040, China;

[14]Department of Radiology, Zhongshan Hospital, Fudan University, Shanghai 200032, China;

[15]Department of Radiology, The Second Affiliated Hospital of Shandong First Medical University, Tai'an, Shandong 271000, China.

Introduction

Upper abdominal magnetic resonance (MR) imaging is appropriate for body composition analysis.[1] Especially for individuals with obesity, it is of great value to quantify the hepatic proton density fat fraction (PDFF) and the amount of abdominal adipose tissue during clinical evaluation and for research on obesity-related risks. Analytical results may be used to determine the optimal choice of surgical procedure and evaluate treatment outcomes. Multiple artificial intelligence (AI) algorithms and systems have been developed for the automated measurement of body composition. The basis of AI development and application is to have uniform standards for clinical data acquisition and management. The uneven quality of MR images is one of the major obstacles to AI system development and analytical

results. A standardized process of MR scanning and clinical data management is urgently needed.

Purpose and Target Audience

This guideline aims to standardize data acquisition, utilization, and storage for AI systems that target the automatic quantification of body composition. This guide is recommended for surgeons, clinical researchers, and radiologists who focus on body composition analysis and obesity-related topics, for example, type 2 diabetes, metabolic syndrome, and bariatric surgery.

Han Lv and Mengyi Li contributed equally to this work.

Correspondence to: Dr. Zhongtao Zhang, Department of General Surgery, Beijing Friendship Hospital, Capital Medical University & National Clinical Research Center for Digestive Diseases, No. 95, Yong An Road, Xicheng District, Beijing 100050, China
E-Mail: zhangzht@ccmu.edu.cn;
Dr. Zhenchang Wang, Department of Radiology, Beijing Friendship Hospital, Capital Medical University, No.95, Yong An Road, Xicheng District, Beijing 100050, China
E-Mail: cjr.wzhch@vip.163.com

Data Acquisition and Evaluation

The standardized data acquisition process guarantees high-quality MR images for AI analysis. An image used for AI labeling and clinical diagnosis must follow the basic process mentioned in the following sections.

Subject preparation

Patients should have an empty stomach before MR scanning. Metal articles should be removed. For overweight subjects with a large waist circumference, wide-bore MR equipment is preferred. According to clinical practice experience, if the patient is >125 kg, doctors or technicians should carefully evaluate the feasibility of upper abdominal MR examination. One challenge is that the bore may not be sufficiently large to accommodate the subject's abdomen. Furthermore, it may be difficult for the patient to hold their breath during MR examination, leading to significant motion artifacts.

MR parameter setting

A 3.0-T or 1.5-T MR device is preferred for data acquisition. The standard parameters for MR examination for AI analysis are listed in Supplementary Table 1, http://links. lww.com/CM9/A951. PDFF is a reliable measure that can be used to accurately evaluate hepatic steatosis.[2] The Dixon image is 3-dimensional with high resolution. The fat image at the axial level of lumbar 1 to lumbar 2 (L1–L2) intervertebral disc on the Dixon image is considered the best choice for quantifying adipose tissue.[1] If a 1.5-T MR device cannot perform 3-dimensional Dixon imaging, dual-echo scanning is also available.

Availability of the MR image for AI quantification

All acquired images should be saved in the Digital Imaging and Communications in Medicine (DICOM) format, as the slice thickness and other important information can be stored. The image quality required by the AI analysis is similar to that required for clinical diagnosis. Overall, images with significant artifacts are judged not acceptable. The other requirements for different applications are listed in the following sections.

The image quality required for PDFF quantification is usually high.[1] For AI analysis of hepatic PDFF, the first step is to recognize the margin of the liver parenchyma. For patients with hepatic steatosis over grade 2 (PDFF >17.4%),[2] the signal intensity of the liver parenchyma is significantly higher than that of the vessels and adjacent organs. As a result, the margin of the liver is relatively easy for the AI system to recognize. However, for patients with hepatic steatosis grade 1 (6.4% < PDFF ≤ 17.4%),[2] the margin of the liver parenchyma is difficult to recognize. Therefore, AI annotation may not be precise. For subjects without hepatic steatosis (PDFF ≤6.4%),[2] AI annotation may fail if only based on PDFF images. The principles and examples of different degrees of hepatic steatosis are listed in Supplementary Table 2, http://links.lww.com/CM9/A951.

According to the recently developed appropriateness criteria, a Dixon sequence image of fat at the axial level of the L1–L2 intervertebral disc can be used for abdominal adipose tissue quantification.[1] It is essential to cover the skin of the abdomen; otherwise, subcutaneous adipose tissue (SAT) cannot be measured. Owing to an insufficient field of view, parallel acquisition, or an uneven magnetic field, the quality of the image may vary. The principles and examples of different image quality degrees are shown in Supplementary Table 3, http://links.lww.com/CM9/A951.

Comprehensive evaluation of MR image quality

Quality evaluation of both the PDFF and fat images should be considered. These principles are listed in Supplementary Table 4, http://links.lww.com/CM9/A951.

Annotation standards

Image for PDFF quantification

The whole liver parenchyma should be included. Large vessels, local lesions, regions beyond the margin of the liver, and imaging artifacts should be avoided [Figure 1A].

The average PDFF value can be calculated by averaging the values of all voxels included in the region of interest. Since the PDFF value and its change after bariatric surgery vary in different parts of the liver,[3] AI systems are being developed to record the values of different liver segments [Figure 1B].

Image for abdominal adipose tissue quantification

Visceral adipose tissue (VAT) and SAT can be recognized and labeled. Different regions of interest, for example, muscle, can also be defined for analysis [Figure 1C and 1D].
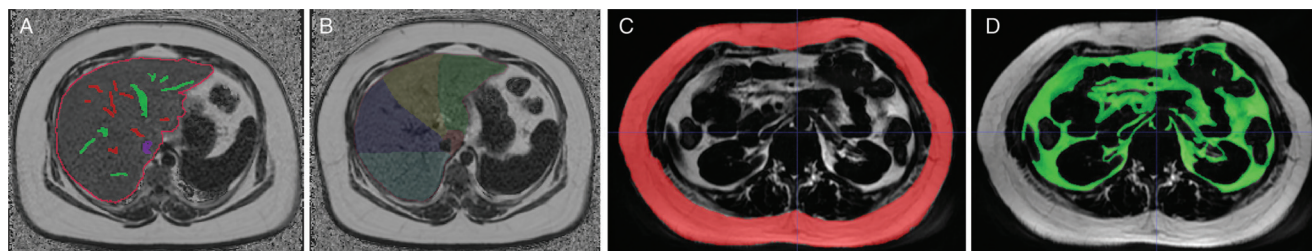

Figure 1: Examples of annotation. (A and B) Annotated areas on a PDFF image. The whole parenchyma of the liver is Included. Large vessels, local lesions, regions beyond the margin of the liver, and imaging artifacts are avoided. Different segments of the liver are also annotated. (C and D) Annotated areas on a fat image of the Dixon sequence at the axial level of the L1–L2 intervertebral disc. The red color represents SAT (19,287 mm^2), while the green color represents VAT (9718 mm^2). L1–L2: Lumbar 1 to lumbar 2; PDFF: Proton density fat fraction; SAT: Subcutaneous adipose tissue; VAT: Visceral adipose tissue.

The images can be labeled automatically using AI[4] or manually using ITK-SNAP 3.8.0 software (http://www.itksnap.org/). Since a single slice of an MR image is volumetric, the value acquired after labeling is influenced by slice thickness. The volume is calculated as the Supplementary formulas, http://links.lww.com/CM9/A951.

Database management

For clinical follow-up and research, it is recommended to set up a database to manage clinical data and MR images.

Data registry

The clinical and radiological data should be registered in a standardized database. For example, there is a prospective national registry database named the "Greater China Metabolic and Bariatric Surgery Database" (GC-MBD®) (Clinicaltrial.gov: NCT03800160) where data of >10,000 cases have been recorded.

Data quality control

The database needs a committee to hold regular meetings to discuss issues concerning quality control. It is highly recommended that a multidisciplinary team achieve a consensus on the variables in the database. For example, according to the consensus of surgeons, radiologists, clinical researchers, and statisticians, variables that should be documented in the GC-MBD include, but are not limited to, structured demographic information, laboratory tests, PDFF values, VAT and SAT values, biological sample information, and adverse event records. Upper abdominal MR images in the DICOM format should also be uploaded. Before data entry, it is essential for the team's main participants to undergo training. The manager of the database should check and verify the authenticity, accuracy, and integrity of all information according to the source data. Data modification traces should be recorded in the system. After verification, the data should be locked.

Discussion

There are three key points in AI analysis of body composition using upper abdominal MR images: uniform data acquisition standards, imaging annotation, and database management. This guideline will promote the development and application of AI systems for the automatic quantification of PDFF and abdominal adipose tissue.

The PDFF value can significantly influence the availability of PDFF images. For patients without hepatic steatosis, the grayscale contrast between the hepatic parenchyma and vessels is insufficient to train the neural network of the AI system. New strategies may solve this problem. For example, the liver may need to be registered with other higher contrast sequences (eg, axial T1-weighted imaging[5] or portal venous phase of contrast-enhanced imaging[6]) to achieve margin recognition for the AI system. As such, additional MR sequences and related parameter standards are required.

Whole-body MR imaging can precisely quantify the volume of adipose tissue. For timesaving, scanning and analyzing a single-slice abdominal MR is preferred. Since imaging annotation for AI analysis requires only a single slice image of fat, we can acquire a single slice at the axial level of the L1–L2 intervertebral disc during the Dixon sequence acquisition to achieve further reduction of scanning time.

Clinical practice guideline registration and ethical approval

This guideline was registered on the International Practice Guideline Registry (IPGRP-2021CN177).

This work was approved by the Ethics Committees of Beijing Friendship Hospital, Capital Medical University (No. 2018-P2-022-01).

Conflicts of interest

None.

References

1. Lv H, Li M, Liu Y, Zhao L, Sun J, Cao D, et al. The clinical value and appropriateness criteria of upper abdominal magnetic resonance examinations in patients before and after bariatric surgery: a study of 837 images. Obes Surg 2020;30:3784–3791. doi: 10.1007/s11695-020-04688-w.
2. Tang A, Tan J, Sun M, Hamilton G, Bydder M, Wolfson T, et al. Nonalcoholic fatty liver disease: MR imaging of liver proton density fat fraction to assess hepatic steatosis. Radiology 2013;267:422–431. doi: 10.1148/radiol.12120896.
3. Li M, Cao D, Liu Y, Jin L, Zeng N, Wang L, et al. Alterations in the liver fat fraction features examined by magnetic resonance imaging following bariatric surgery: a self-controlled observational study. Obes Surg 2020;30:1917–1928. doi: 10.1007/s11695-020-04415-5.
4. Hui S, Zhang T, Shi L, Wang D, Ip CB, Chu W. Automated segmentation of abdominal subcutaneous adipose tissue and visceral adipose tissue in obese adolescent in MRI. Magn Reson Imaging 2018;45:97104. doi: 10.1016/j.mri.2017.09.016.
5. Wang K, Mamidipalli A, Retson T, Bahrami N, Hasenstab K, Blansit K, et al. Automated CT and MRI liver segmentation and biometry using a generalized convolutional neural network. Radiol Artif Intell 2019;1:180022. doi: 10.1148/ryai.2019180022.
6. Wang SH, Du J, Xu H, Yang D, Ye Y, Chen Y, et al. Automatic discrimination of different sequences and phases of liver MRI using a dense featurefusionneural network: a preliminarystudy. AbdomRadiol (NY) 2021;46:4576–4587. doi: 10.1007/s00261-021-03142-4.