Radiology

Automated Liver Fat Quantification at Nonenhanced Abdominal CT for Population-based Steatosis Assessment

Peter M. Graffy, BA, MPH • Veit Sandfort, MD • Ronald M. Summers, MD, PhD • Perry J. Pickhardt, MD

From the Department of Radiology, University of Wisconsin School of Medicine and Public Health, E3/311 Clinical Science Center, 600 Highland Ave, Madison, Wis 53792-3252 (P.M.G., P.J.P.); and Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Department of Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Bethesda, Md (V.S., R.M.S.). Received March 5, 2019; revision requested April 8; final revision received July 27; accepted August 6. **Address correspondence to** P.J.P. (e-mail: *ppickhardt2@uwhealth.org*).

This research was supported in part by the Intramural Research Program of the National Institutes of Health (NIH) Clinical Center and made use of the high-performance computing capabilities of the NIH Biowulf system.

Conflicts of interest are listed at the end of this article.

Radiology 2019; 293:334–342 • https://doi.org/10.1148/radiol.2019190512 • Content codes: CT GI

Background: Nonalcoholic fatty liver disease and its consequences are a growing public health concern requiring cross-sectional imaging for noninvasive diagnosis and quantification of liver fat.

Purpose: To investigate a deep learning–based automated liver fat quantification tool at nonenhanced CT for establishing the prevalence of steatosis in a large screening cohort.

Materials and Methods: In this retrospective study, a fully automated liver segmentation algorithm was applied to noncontrast abdominal CT examinations from consecutive asymptomatic adults by using three-dimensional convolutional neural networks, including a subcohort with follow-up scans. Automated volume-based liver attenuation was analyzed, including conversion to CT fat fraction, and compared with manual measurement in a large subset of scans.

Results: A total of 11 669 CT scans in 9552 adults (mean age ± standard deviation, 57.2 years ± 7.9; 5314 women and 4238 men; median body mass index [BMI], 27.8 kg/m^2) were evaluated, including 2117 follow-up scans in 1862 adults (mean age, 59.2 years; 971 women and 891 men; mean interval, 5.5 years). Algorithm failure occurred in seven scans. Mean CT liver attenuation was 55 HU ± 10, corresponding to CT fat fraction of 6.4% (slightly fattier in men than in women [7.4% ± 6.0 vs 5.8% ± 5.7%; $P < .001$]). Mean liver Hounsfield unit varied little by age (<4 HU difference among all age groups) and only weak correlation was seen with BMI ($r^2 = 0.14$). By category, 47.9% (5584 of 11 669) had negligible or no liver fat (CT fat fraction <5%), 42.4% (4948 of 11 669) had mild steatosis (CT fat fraction of 5%–14%), 8.8% (1025 of 11 669) had moderate steatosis (CT fat fraction of 14%–28%), and 1% (112 of 11 669) had severe steatosis (CT fat fraction >28%). Excellent agreement was seen between automated and manual measurements, with a mean difference of 2.7 HU (median, 3 HU) and r^2 of 0.92. Among the subcohort with longitudinal follow-up, mean change was only −3 HU ± 9, but 43.3% (806 of 1861) of patients changed steatosis category between first and last scans.

Conclusion: This fully automated CT-based liver fat quantification tool allows for population-based assessment of hepatic steatosis and nonalcoholic fatty liver disease, with objective data that match well with manual measurement. The prevalence of at least mild steatosis was greater than 50% in this asymptomatic screening cohort.

The increasing prevalence of hepatic steatosis and nonalcoholic fatty liver disease (NAFLD) as a component of the obesity and metabolic syndrome epidemics is emerging as a major public health concern in the United States and throughout the developed world. Over 60% of American adults are considered to be obese, along with nearly 20% of children, while approximately half of American adults may have some degree of hepatic steatosis (1–5). The prevalence of NAFLD within most populations with obesity is thought to be between 70%–90% (5–7), which carries important implications for comorbidities including cardiovascular disease, metabolic syndrome, and (less often) progression to nonalcoholic steatohepatitis and cirrhosis (8–12). However, despite the increasing trend of obesity and fatty liver disease, there is no reliable nonimaging clinical method at present, to our knowledge, for accurately capturing the degree of hepatic steatosis at the population level, either prospectively or retrospectively.

Among the cross-sectional imaging techniques to quantify liver fat, MRI proton density fat fraction, or PDFF, is emerging as the noninvasive reference standard of choice, given its accuracy and whole-liver assessment (13–15). Recent studies have established a linear correlation between liver fat measurement at MRI PDFF and nonenhanced CT, allowing for the latter to also accurately quantify liver fat (16). Going forward, both MRI and CT now represent superior reference standards to histopathologic analysis for liver fat quantification, and measurements are obtained noninvasively. Furthermore, because abdominal CT is much more frequently performed than MRI in clinical practice for a wide variety of indications (17,18), the potential now exists for population-based opportunistic screening for hepatic steatosis and NAFLD. Although liver Hounsfield unit measurement at nonenhanced CT is typically achieved by using a manual region-of-interest (ROI) approach, we have developed a fully

Abbreviations

BMI = body mass index, CI = confidence interval, NAFLD = nonalcoholic fatty liver disease, ROI = region of interest

Summary

Population-based quantification of liver fat at nonenhanced CT in a large asymptomatic adult outpatient cohort by using a fully automated deep learning CT-based method may help to identify individuals at risk for nonalcoholic fatty liver disease and metabolic syndrome, while also allowing for large-scale population-based studies.

Key Results

- The prevalence of hepatic steatosis was greater than 50% in the asymptomatic screening cohort according to an automated deep learning CT-based tool, including 10% with moderate or severe steatosis.
- This automated deep learning liver segmentation and fat quantification algorithm used at nonenhanced CT produces results that correlated well with manual liver Hounsfield unit measurement (mean difference <3 HU).
- Correlation of hepatic steatosis with body mass index, age, and sex were all weak, indicating that clinical demographic and morphometric assessment is ineffective for predicting the likelihood of underlying nonalcoholic fatty liver disease.

automated algorithm for CT-based liver segmentation and attenuation assessment. The purpose of our study was, therefore, to apply this recently trained and tested deep learning–based automated liver algorithm to nonenhanced CT in a large, longitudinal, asymptomatic adult screening cohort, including correlation with manual measurement, to assess the prevalence of hepatic steatosis.

Materials and Methods

Study Cohort

This was an institutional review board–approved, Health Insurance Portability and Accountability Act–compliant, retrospective cohort study of 9552 consecutive patients who underwent initial nonenhanced abdominal CT at a single academic medical center for the purpose of colonography screening between April 2004 and December 2016. The need for signed informed consent was waived for this retrospective study. Including CT follow-up in a subcohort of 1862 asymptomatic individuals undergoing 2117 additional CT colonography scans, a total of 11 682 noncontrast CT examinations in asymptomatic adults were eligible for inclusion (Fig 1), allowing for assessment of longitudinal changes over time. Causes for exclusion were failure of the algorithm, liver not included on scan series, and severe artifact resulting in unusable output. All patient information was anonymized before the automated measurements were performed. In separate studies, we have investigated the use of other automated CT tools with this patient cohort, including algorithms for quantifying fat, bone mineral density, muscle, and aortic calcification (19–22).

CT Scanning Protocol

Specifics relating to CT colonography technique, including bowel preparation and colonic distention, have been previously described and are not of relevance to our liver-focused study (23). Breath-hold CT acquisition of the abdomen and pelvis without intravenous contrast enhancement was obtained in both supine and prone positions, but only the former was used for liver assessment. All CT scans were obtained with 8–64-section multidetector–row scanners (GE Healthcare, Waukesha, Wis) by using 120 kVp with low-dose milliampere settings (typically modulated between 30–300 mA). All CT scanners were calibrated daily for quality control of Hounsfield unit measurements. Of note, although varying the kilovolt setting affects mean Hounsfield unit measurements, varying milliampere will only affect image noise (ie, the standard deviation of Hounsfield unit measurements), but should not meaningfully affect mean Hounsfield unit values. The kilovolt settings were kept constant for this reason. Overlapping image reconstruction included both thin (1.25 mm) and thick (5 mm) sections for colonic and extracolonic assessment, respectively.

Automated Algorithm for Liver Segmentation and Liver Fat Quantification

The detailed description of the segmentation method can be found in a separate technical publication (24). Briefly, a modified three-dimensional U-Net was used for segmentation (25,26). To enable processing of large three-dimensional volumes with limited graphics processing unit, or GPU, memory, an initial strided convolution (step size 2) and a complementing final transposed convolution were added. Training data were obtained from the Medical Segmentation Decathlon project (27). As liver training data were available at contrast material–enhanced CT only, while our population data are at noncontrast CT, we performed domain transfer of contrast CT to noncontrast CT by using CycleGAN (28). Data augmentation was performed by using three-dimensional rotation, crop, elastic deformation, CycleGAN noncontrast image, and random flips. Model training and inference took place on the National Institutes of Health Biowulf high-performance computing cluster by using 4 CPU threads and up to 48 GPU nodes (NVIDIA K80 or P100; Santa Clara, Calif; 12 GB or 16 GB GPU memory), respectively. Batch size was 4, resolution was $256 \times 256 \times 192$, and initial filters were 32. The initial learning rate was 0.0001, and training was performed for 10 000 iterations. All voxels designated as liver by the segmentation algorithm were analyzed, and the mean and median Hounsfield unit were computed.

Manual ROI Method for Liver Attenuation Measurement

In a subset of 5562 CT scans, manual mean liver attenuation was measured (in Hounsfield units) by using a simple and previously validated technique that consists of placement of an ROI over a representative parenchymal portion of the right hepatic lobe (Fig 2) (29–33). Typical size of the rounded ROI was 500–1500 mm^2 and care was taken to avoid vessels, focal nodules or masses, or other causes of heterogeneity. Manual ROI measurements were placed independently by several trained individuals in large consecutive subsamples by using the approach described above, representing the full spectrum

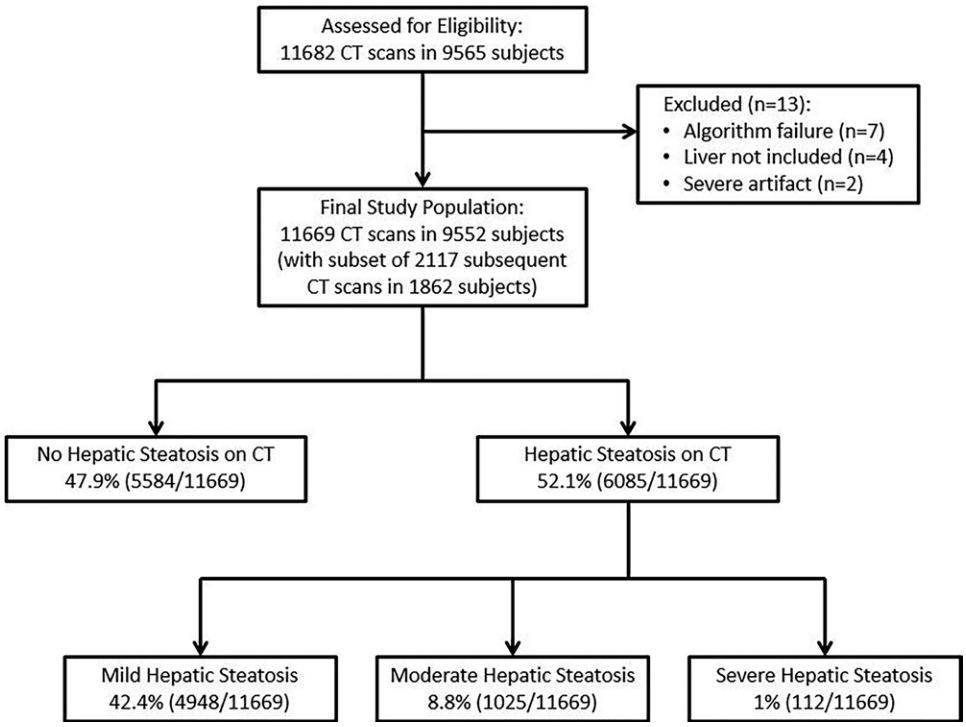

Figure 1: Flowchart shows study cohort, including final categories of liver fat.

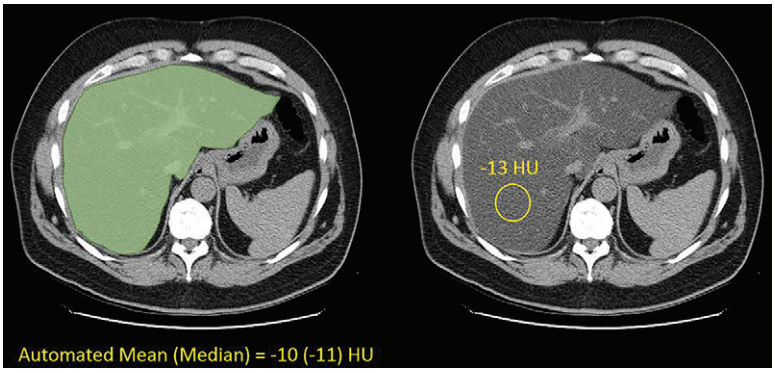

Figure 2: Images show automated and manual methods for measuring liver attenuation at nonenhanced CT in an asymptomatic 60-year-old man. Transverse (axial) CT image showing schematic depiction of automated (left) and manual (right) techniques. Automatically segmented liver is in green and manual region-of-interest (ROI) placement is in yellow. Automated Hounsfield units are based on volumetric segmentation of entire imaged liver, whereas manual Hounsfield units are based on areal mean within single ROI. Nonetheless, two values match well.

of liver fat fraction. Manual measures were performed without any knowledge of automated results or steatosis category. Reader experience for this simple ROI task varied from medical students (with <1 year of experience) to abdominal radiologists (with >10 years of experience). Manual measurements for study scans were included from routine clinical practice recorded in our clinical database, as well as others from previously conducted steatosis trials (29,30).

The automated and manual methods for assessing liver Hounsfield units have methodologic differences that could impact direct comparison. The automated method segments the entire available liver (including hepatic blood pool) and derives a mean volumetric liver attenuation, whereas the manual ROI method samples only a portion of the right hepatic lobe on a single section, deriving the mean Hounsfield unit only within that restricted area.

Statistical Analysis

Summary statistics were collected for the patient population based on age and sex. Conversion from the nonenhanced CT Hounsfield unit numbers to estimated CT fat fraction was calculated from the MRI PDFF equivalent by using the recently published (16) formula: CT fat fraction (%) = (CT HU) × (−0.58) + 38.2. This equation is specific to 120-kV scanning. For apparently negative CT fat-fraction numbers (ie, when CT attenuation is 65.9 HU or greater), CT fat fraction is necessarily constrained to 0%. Steatosis categories are herein defined as normal for fat fraction less than 5% (>57 HU), mild for 5%–14% (42–57 HU), moderate for 14%–28% (18–42 HU), and severe for greater than 28% (<18 HU) (16).

For comparison of automated versus manual Hounsfield unit measures, a 4 × 4 confusion matrix was constructed to compare steatosis categorization. In addition, for each category threshold, sensitivity and specificity (with 95% confidence intervals [CIs]) were calculated for automated Hounsfield units by using the manual Hounsfield unit measures as the reference standard.

In patients with follow-up CT scans, changes in liver attenuation values (and CT fat fraction) were recorded. Relative percent change in liver attenuation was calculated by taking the difference in Hounsfield unit between the two scans, dividing this by the initial result, and multiplying by 100. Annual percent change over time was calculated by further dividing this by the interval years between scans (in years). Sankey diagram analysis was performed to illustrate changes in steatosis categorization (ie, normal, mild, moderate, or severe) over time. *P* values were derived by using two-sided *t* tests and validated by using nonparametric testing (Wilcoxon rank sum test) in all cases; $P < .05$ was considered to indicate statistical significance. A linear regression model was used to compare body mass index (BMI) and CT fat fraction. In addition, a multivariable linear regression analysis was performed to control for the covariates of BMI, sex, and age by using the ordinary least-squares method. Data

processing and analysis was performed by using base R and ggplot2 packages (R Core Team, version 3.4.2; R Foundation for Statistical Computing, Vienna, Austria).

Results

Study Cohort

The mean age ± standard deviation of the cohort of 9552 asymptomatic adults was 57.2 years ± 7.9, including 4238 men and 5314 women. Mean BMI was 28.9 kg/m^2 ± 6.5 (median, 27.8 kg/m^2; interquartile range, 24.6–31.8 kg/m^2). There were 1862 patients (mean age, 59.2 years; 891 men and 971 women) with 2117 subsequent follow-up CT scans (mean time interval between scans was 5.5 years); 245 patients underwent more than one follow-up scan and 10 patients underwent three follow-up scans. The mean age of the 4439-patient subset with manual measurement in 5265 scans was 57.7 years, including 2069 men and 2343 women. Thirteen of 11 682 total CT scans were excluded for algorithm failure (n = 7), liver not included (n = 4), and severe artifact precluding either manual or automated Hounsfield unit (n = 2), yielding a final cohort of 11 669 total CT examinations (Fig 1). Failure rate of the algorithm was 0.06% (seven of 11 676).

Automated Tool Results

The mean automated volumetric CT liver attenuation was 55 HU ± 10, corresponding to a CT fat fraction of 6.4%. Overall, 47.9% (5584 of 11 669) had no or negligible steatosis (CT fat fraction <5%), 42.4% (4948 of 11 669) had mild steatosis (CT fat fraction of 5%–14%), 8.8% (1025 of 11 669) had moderate steatosis (CT fat fraction of 14%–28%), and 1% (112 of 11 669) had severe steatosis (CT fat fraction >28%). The population distribution of automated liver attenuation and CT fat fraction is shown in Figure 3. Men had a slightly higher fat fraction than did women on average (7.4% ± 6.0 vs 5.8% ± 5.7; $P < .001$). Little change was seen in mean liver Hounsfield unit according to age in this cohort (Table 1), with mean values differing by less than 4 HU among all age groups spanning from younger than 40 years to older than 80 years. The standard deviations of the mean also varied by only 4 HU or less. A slight decrease in mean Hounsfield unit with increasing age was also seen across age categories by sex (Table 2).

There was only a weak positive correlation between CT fat fraction and BMI (r^2 = 0.14) (Fig 4). Multiple linear regression modeling was also performed to assess the collective impact of age, sex, and BMI on liver CT fat fraction. The model produced the equation CT fat fraction (%) = $78.6 - 0.1_{age} - 2.4_{sex} - 0.6_{BMI}$, demonstrating that sex had the greatest influence on liver attenuation in the presence of age and BMI. The impact of age and BMI was minimal. However, CT fat fraction was higher in patients with a BMI greater than or equal to 30 kg/m^2 compared with those with a BMI less than 30 kg/m^2 (9.3% ± 7.1 vs 5.0% ± 4.5; $P < .001$). Furthermore, CT scans in patients with obesity (BMI ≥30 kg/m^2) showed steatosis in 69.6% (2772 of 3983), including moderate or severe steatosis in 20.5% (817 of 3983).

Correlation with Manual Measurement

Excellent agreement was seen for the subset of 5265 scans with both automated and manual liver Hounsfield unit measurement, despite the fundamental differences in the automated volumetric versus limited manual areal ROI approach. The overall mean difference between automated and manual measurement was 2.7 HU (median, 3 HU), with an r^2 value of 0.93 (Fig 5). Mean population-based automated and manual measurements were 56 HU ± 10 and 59 HU ± 12, respectively. Bland-Altman analysis revealed a proportional bias of 2.7 HU toward manual measurements, which increased slightly at higher Hounsfield units, presumably related to the inclusion of hepatic blood pool in the automated technique. This would slightly lower mean Hounsfield units when blood pool attenuation was less than liver attenuation, but would eventually increase automated Hounsfield units in advanced steatosis, when blood pool attenuation is greater than liver attenuation.

Comparison of steatosis categorization by using automated versus manual measures is demonstrated by the 4 × 4 confusion matrix (Table 3). By using the manual Hounsfield unit measure as the reference standard, the sensitivity and specificity of the automated technique for categorizing a patient as healthy (no steatosis) was 82.6% (3137 of 3800; 95% CI: 81.3%, 83.8%) and 96.3% (1411 of 1465; 95% CI: 95.2%, 97.2%), respectively. Similarly for mild, moderate, and severe steatosis, the automated performance was 88.7% (892 of 1006; 95% CI: 86.6%, 90.6%) and 83.5% (3556 of 4259; 95% CI: 82.3%, 84.6%), 87.5% (343 of 392; 95% CI: 83.8%, 90.6%) and 98.3% (4792 of 4873; 95% CI: 97.9%, 98.7%), and 74.6% (50 of 67; 95% CI: 62.5%, 84.5%) and 99.9% (5193 of 5198; 95% CI: 99.9%, 99.9%), respectively.

Longitudinal Subcohort

Among the subcohort of 1862 individuals with longitudinal follow-up (mean time interval of 5.5 years), the overall average change in mean liver attenuation was minimal at −3 HU ± 9, reflecting a decrease of 6% in Hounsfield unit. A total of 42.0% (782 of 1861) of patients decreased in liver attenuation by at least 5 HU between their initial CT and first follow-up, whereas 15.3% (284 of 1861) increased in attenuation by at least 5 HU, and 42.7% (795 of 1861) remained within 5 HU (Fig 6). A total of 43.3% (806 of 1861) of individuals changed steatosis category (normal, mild, moderate, severe) between their first and last scans, more often toward greater degrees of steatosis, which is better illustrated by the Sankey diagram in Figure 7.

Discussion

We demonstrate the feasibility of using an automated deep learning tool for population-based screening for hepatic steatosis. Given our relatively unique noncontrast CT cohort of asymptomatic and generally healthy outpatient adults, our study may provide further insights into the "normal" population range of liver fat, including prevalent degrees of nonalcoholic fatty liver disease (NAFLD). Of note, the prevalence of at least mild hepatic steatosis was greater than 50% in our asymptomatic adult screen-

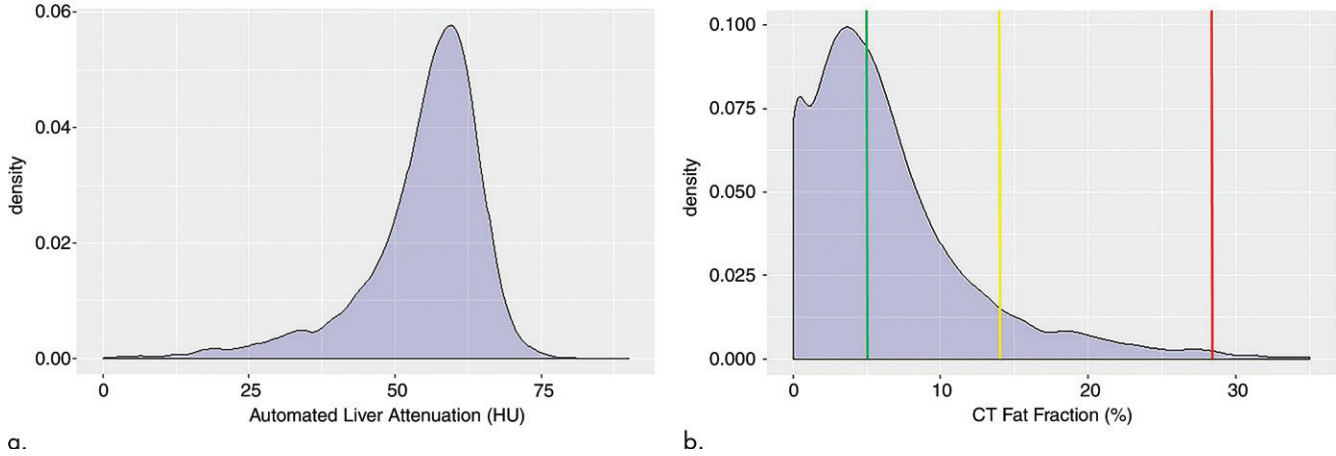

Figure 3: Density plots show population-based distribution of mean liver attenuation (in Hounsfield units) at **(a)** unenhanced CT and **(b)** corresponding estimated CT fat fraction. In **b**, all patients to right of green line had at least mild steatosis (fat fraction ≥5%; prevalence of 52.1%), whereas those patients to right of yellow line had at least moderate steatosis (fat fraction ≥14%; prevalence of 9.7%), and those patients to right of red line had severe steatosis (fat fraction ≥28%; prevalence of 1.0%).

Table 1: Population-based Values of Automated CT Liver Attenuation and Estimated Fat Fraction

Age (y)	No.	Mean Liver Attenuation (HU)	Median Liver Attenuation (HU)	Interquartile Range (HU)	Mean CT Fat Fraction (%)
<40	59	57 ± 8	58	55–63	5.1 ± 4.2
40–44	88	56 ± 11	59	54–62	5.8 ± 6.1
45–49	251	57 ± 9	58	54–62	5.3 ± 4.8
50–54	3944	56 ± 10	58	52–62	6.0 ± 5.7
55–59	3042	55 ± 11	57	51–62	6.4 ± 6.0
60–64	2294	54 ± 12	56	50–61	7.2 ± 6.5
65–69	981	54 ± 10	56	50–60	7.1 ± 5.6
70–74	538	54 ± 10	55	50–60	7.1 ± 5.6
75–79	261	54 ± 10	56	51–60	7.0 ± 5.8
>80	211	55 ± 8	55	50–59	6.8 ± 4.1

Note.—Unless otherwise specified, data are means ± standard deviation.

Table 2: Automated Liver Results according to Sex and Age Categories

	Automated Mean Liver Attenuation (HU)		
Age (y)	Men (n = 5262)	Women (n = 6407)	P Value
<40	57 ± 8	57 ± 7	.84*
40–49	56 ± 10	57 ± 9	.207*
50–59	54 ± 10	57 ± 10	<.001
60–69	52 ± 11	55 ± 11	<.001
70–79	53 ± 10	55 ± 9	<.001
≥80	53 ± 7	56 ± 9	.032*
All	53 ± 11	56 ± 10	<.001

Note.— Unless otherwise specified, data are means ± standard deviation.

* Corresponding P values from Wilcoxon rank sum test were P = .740, P = .653, and P = .018, respectively.

ing cohort. Although we are aware of other studies that have used automated CT techniques for such biomarkers in smaller cohorts (of approximately 100 individuals) (34), our study also provided correlation with manual measures, converted results to MRI proton density fat fraction, or PDFF, equivalent values, and assessed for longitudinal changes over time. With our data and automated CT tool, clinicians and researchers can potentially evaluate even larger-scale populations for hepatic steatosis by using nonenhanced CT scans obtained for any reason that includes the liver, either prospectively or retrospectively.

NAFLD and the associated metabolic syndrome represent a growing and potentially major public health issue, especially if left unchecked. Hepatic steatosis is largely an asymptomatic condition but—whether a cause or effect—is related to diabetes, obesity, hyperlipidemia, and metabolic syndrome. Some have argued that even mild steatosis, present in nearly half of our otherwise healthy screening cohort, carries potentially serious health risks. Assuming the degree of steatosis is relevant, we found a 10% prevalence of moderate or severe hepatic steatosis within our patient population. Furthermore, although individuals with

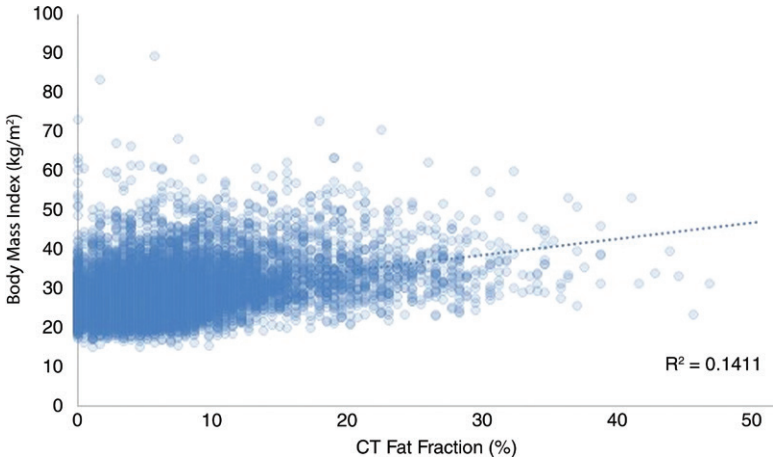

Figure 4: Graph shows relationship between CT fat fraction and body mass index (BMI). Surprisingly little correlation is seen, with wide range of fat fractions at any given BMI value. As such, an individual's BMI would appear to be poor predictor of hepatic steatosis.

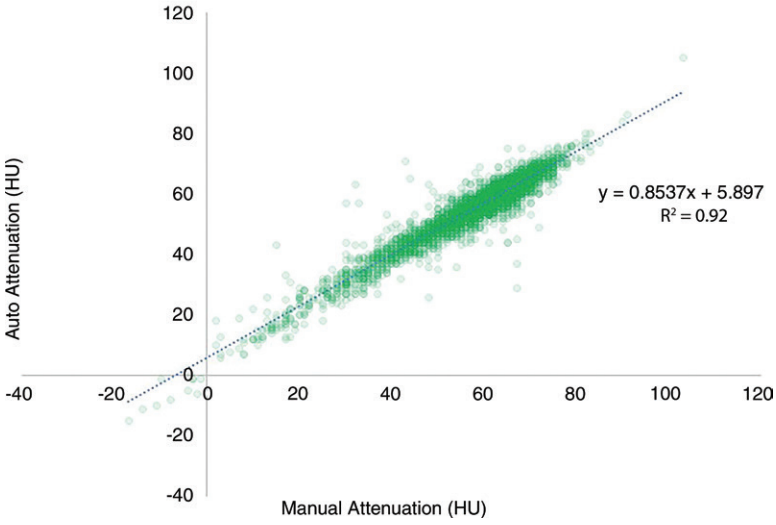

Figure 5: Scatter plot shows automated versus manual measurement of liver attenuation at nonenhanced CT in subcohort of 5265 scans with excellent agreement (r^2 = 0.934). Slight bias (2.7 HU) was observed toward manual values, which increases slightly at higher Hounsfield units, presumably related to inclusion of hepatic blood pool with automated technique.

Table 3: 4 × 4 Confusion Matrix for Steatosis Categorization by Using Automated versus Manual Measures (*n* = 5265)

	Automated Steatosis Categorization			
Manual Steatosis Categorization	Normal	Mild	Moderate	Severe
Normal	3137	661	2	0
Mild	51	892	63	0
Moderate	3	41	343	5
Severe	0	1	16	50

Note.—Data are CT scans of patients.

obesity (body mass index [BMI] ≥30 kg/m^2) did have a slightly fattier liver on average than did those with a BMI less than 30 kg/m^2, we found there was surprisingly little correlation between hepatic steatosis and BMI. This is relevant because the presence or absence of hepatic steatosis can therefore not be reliably inferred from a patient's body habitus. Although neither hepatic nor visceral fat are defining features of metabolic syndrome, they may provide unique information that is lacking from BMI alone (32). These two measures are readily available from nonenhanced CT, and further investigation is warranted in terms of their predictive value.

Similarly, although men had slightly more liver fat compared with women on average, and mean liver fat slightly increased with patient age, no meaningful correlation could otherwise be derived between age and steatosis. As such, simple clinical assessment of basic demographic and morphometric features cannot reliably predict the presence of underlying hepatic steatosis. Mild elevation in liver enzymes is relatively common in the setting of steatosis but is also highly nonspecific, and requires cross-sectional imaging for confirmation.

We observed excellent overall agreement between the automated liver attenuation values derived from the deep learning algorithm and the manual ROI-based measurements, despite the underlying difference in methodology (volumetric vs single-section ROI). This suggests that the effect of marked regional heterogeneity in liver attenuation, which exists but is relatively uncommon in practice, does not meaningfully impact overall population assessment. Nearly all livers in our asymptomatic outpatient cohort were relatively homogeneous in appearance at CT. The slight difference in mean attenuation presumably relates to inclusion of the intrahepatic vasculature with the volumetric automated approach, which would have a negative impact when liver Hounsfield units exceed blood pool Hounsfield unit, and vice-versa for advanced steatosis when liver parenchyma drops below blood pool attenuation. Regardless, the automated method can serve as a natural extension of the existing CT literature based on manual ROI measurement, which was previously shown to be a representative measure despite the undersampling (29–32,35,36). Furthermore, the linear relationship between noncontrast CT Hounsfield units and MRI PDFF percentages allows for direct correlation between these imaging modalities and reporting of fat fractions (16).

Abdominal CT is a commonly performed examination among middle-age and older adults in the United States (18), which provides an opportunity to screen for multiple conditions beyond the initial indication for imaging. In addition to hepatic steatosis (29–31), we found that CT can also be used to retrospectively assess for hepatic fibrosis (37–41). Beyond NAFLD and liver disease, we are also investigating the use of abdominal CT for opportunistic screening of a wide variety of additional entities, including osteoporosis, abdominal aortic aneurysm and calcification, visceral fat, sarcopenia, hemochromatosis, cancer, and the metabolic syndrome (32,42–48). Furthermore, nearly all these tasks

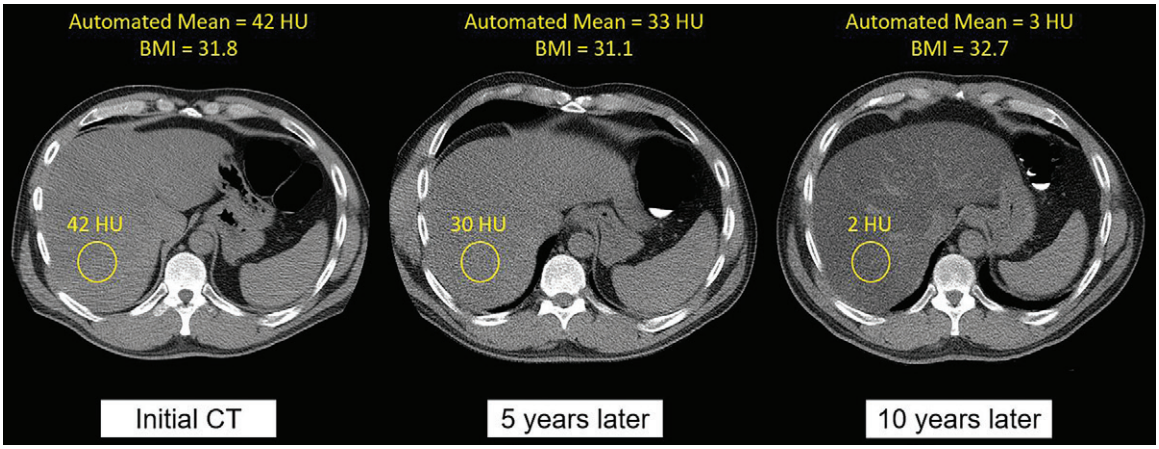

Figure 6: Images show screening nonenhanced transverse (axial) CT scans over 1 decade in an asymptomatic man who was 51 years old at time of initial imaging. There is substantial variation in liver attenuation over time (corresponding to fat fraction change of >20%), with only minor changes in body mass index (BMI). There is excellent agreement between automated and manual measures of liver attenuation.

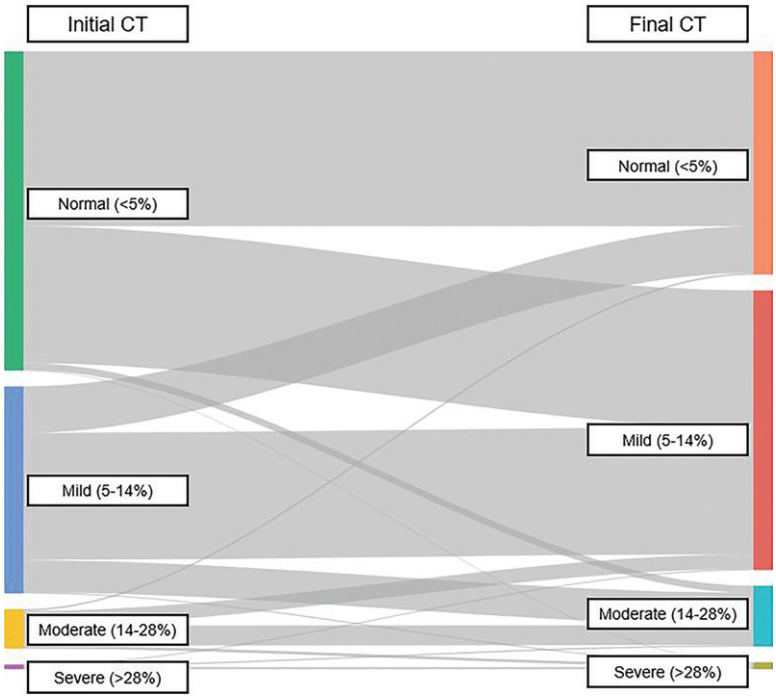

Figure 7: Sankey diagram illustrates change in steatosis categorization between initial and final CT scans in subcohort of 1861 individuals. Of those patients initially classified as normal (fat fraction <5%), 42.8% (446 of 1041) increased to mild (fat fraction 5%–14%) and 2.4% (25 of 1041) increased to moderate (fat fraction 14%–28%) or severe (fat fraction >28%) steatosis. Of those starting at mild, 22.5% (152 of 675) reverted to normal, and 16.1% (109 of 675) increased to moderate or severe steatosis.

can be fully automated through means of artificial intelligence, presumably eliminating any variability related to subjective interobserver measurements. To date, prior to liver evaluation, we have investigated automated CT-based bone mineral density, visceral fat, muscle, and aortic calcification (19–21). These additional objective measures can be added prospectively or retrospectively without any additional time or dose to the patient. If utilization of CT colonography for colorectal cancer screening were to substantially increase in the future, then all the additional opportunities described above could be included without any additional scanning.

We acknowledge certain limitations to our study. All scans were derived from a single medical center by using scanners from a single CT vendor and noncontrast technique. Further external validation of the tool using a variety of different patient care settings and CT techniques is warranted. These liver Hounsfield unit results apply only to nonenhanced CT. We plan to study a cohort of patients with both nonenhanced and contrast-enhanced CT imaging of the liver to determine if the impact of hepatic enhancement on fat quantification is predictable. As noted, there were differences in measurement technique between the automated and manual liver Hounsfield unit approaches. Nonetheless, these measures correlated strongly. The relatively smaller number of cases with moderate (n = 1025) and severe (n = 112) steatosis limit our conclusions somewhat, especially toward the severe end of the fatty liver spectrum. However, performance data from confusion matrix analysis shows that performance is relatively preserved across the Hounsfield unit spectrum. Lastly, our study did not attempt to correlate liver attenuation values with downstream adverse clinical outcomes. If we are successful in demonstrating clinical utility of this automated tool in risk profiling, then the next logical step would be widespread implementation as a prospective clinical tool.

In conclusion, we provide validation for a deep learning–based automated liver segmentation tool at abdominal CT for quantifying liver fat. This fully automated CT tool provides both rapid and objective assessment that allows for application to large retrospective cohorts for future research. The relative lack of correlation between liver fat and clinical data including body mass index, age, and sex further demonstrate the need for quantitative cross-sectional imaging in nonalcoholic fatty liver disease (NAFLD). We found the overall prevalence of steatosis to be greater than 50% in our asymptomatic adult cohort, including 10% with moderate

or severe steatosis. We have also demonstrated how this tool can be used for longitudinal assessment of change in liver fat fraction. If hepatic steatosis is shown to be an independent risk factor for future adverse events, then this automated tool could also be potentially used for opportunistic NAFLD screening with any nonenhanced CT that includes the liver (abdominal or thoracic), regardless of the clinical indication for imaging.

Author contributions: Guarantor of integrity of entire study, P.J.P.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, P.M.G., P.J.P.; clinical studies, all authors; statistical analysis, P.M.G., V.S.; and manuscript editing, P.M.G., V.S., P.J.P.

Disclosures of Conflicts of Interest: P.M.G. disclosed no relevant relationships. **V.S.** disclosed no relevant relationships. **R.M.S.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: receives royalties from iCAD, Philips, PingAn, and ScanMed; receives research support from PingAn and NVIDIA. Other relationships: received donation of GPU cards from NVIDIA; institution has a cooperative research and development agreement with PingAn. **P.J.P.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: is a consultant for Bracco; is a shareholder in Cellectar, Elucent, and SHINE. Other relationships: disclosed no relevant relationships.

References

1. Wang Y, Beydoun MA. The obesity epidemic in the United States: gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. Epidemiol Rev 2007;29(1):6–28.
2. Parikh NI, Pencina MJ, Wang TJ, et al. Increasing trends in incidence of overweight and obesity over 5 decades. Am J Med 2007;120(3):242–250.
3. Wang Y, Beydoun MA, Liang L, Caballero B, Kumanyika SK. Will all Americans become overweight or obese? Estimating the progression and cost of the US obesity epidemic. Obesity (Silver Spring) 2008;16(10):2323–2330.
4. Bedogni G, Nobili V, Tiribelli C. Epidemiology of fatty liver: an update. World J Gastroenterol 2014;20(27):9050–9054.
5. Bellentani S. The epidemiology of non-alcoholic fatty liver disease. Liver Int 2017;37(Suppl 1):81–84.
6. Bellentani S, Scaglioni F, Marino M, Bedogni G. Epidemiology of non-alcoholic fatty liver disease. Dig Dis 2010;28(1):155–161.
7. Perumpail BJ, Khan MA, Yoo ER, Cholankeril G, Kim D, Ahmed A. Clinical epidemiology and disease burden of nonalcoholic fatty liver disease. World J Gastroenterol 2017;23(47):8263–8276.
8. Anstee QM, Targher G, Day CP. Progression of NAFLD to diabetes mellitus, cardiovascular disease or cirrhosis. Nat Rev Gastroenterol Hepatol 2013;10(6):330–344.
9. Lonardo A, Ballestri S, Marchesini G, Angulo P, Loria P. Nonalcoholic fatty liver disease: a precursor of the metabolic syndrome. Dig Liver Dis 2015;47(3):181–190.
10. Targher G, Byrne CD, Lonardo A, Zoppini G, Barbui C. Non-alcoholic fatty liver disease and risk of incident cardiovascular disease: A meta-analysis. J Hepatol 2016;65(3):589–600.
11. Targher G, Marchesini G, Byrne CD. Risk of type 2 diabetes in patients with non-alcoholic fatty liver disease: Causal association or epiphenomenon? Diabetes Metab 2016;42(3):142–156.
12. Wree A, Broderick L, Canbay A, Hoffman HM, Feldstein AE. From NAFLD to NASH to cirrhosis-new insights into disease mechanisms. Nat Rev Gastroenterol Hepatol 2013;10(11):627–636.
13. Liu CY, McKenzie CA, Yu H, Brittain JH, Reeder SB. Fat quantification with IDEAL gradient echo imaging: correction of bias from T(1) and noise. Magn Reson Med 2007;58(2):354–364.
14. Reeder SB, Cruite I, Hamilton G, Sirlin CB. Quantitative assessment of liver fat with magnetic resonance imaging and spectroscopy. J Magn Reson Imaging 2011;34(4):729–749.
15. Reeder SB, Hu HH, Sirlin CB. Proton density fat-fraction: a standardized MR-based biomarker of tissue fat concentration. J Magn Reson Imaging 2012;36(5):1011–1014.
16. Pickhardt PJ, Graffy PM, Reeder SB, Hernando D, Li K. Quantification of Liver Fat Content With Unenhanced MDCT: Phantom and Clinical Correlation With MRI Proton Density Fat Fraction. AJR Am J Roentgenol 2018;211(3):W151–W157.
17. Agarwal R, Bergey M, Sonnad S, Butowsky H, Bhargavan M, Bleshman MH. Inpatient CT and MRI utilization: trends in the academic hospital setting. J Am Coll Radiol 2010;7(12):949–955.
18. Moreno CC, Hemingway J, Johnson AC, Hughes DR, Mittal PK, Duszak R Jr. Changing Abdominal Imaging Utilization Patterns: Perspectives From Medicare Beneficiaries Over Two Decades. J Am Coll Radiol 2016;13(8):894–903.
19. Lee SJ, Liu J, Yao J, Kanarek A, Summers RM, Pickhardt PJ. Fully automated segmentation and quantification of visceral and subcutaneous fat at abdominal CT: application to a longitudinal adult screening cohort. Br J Radiol 2018;91(1089):20170968.
20. Pickhardt PJ, Lee SJ, Liu J, et al. Population-based opportunistic osteoporosis screening: Validation of a fully automated CT tool for assessing longitudinal BMD changes. Br J Radiol 2019;92(1094):20180726.
21. Graffy PM, Liu J, O'Connor S, Summers RM, Pickhardt PJ. Automated segmentation and quantification of aortic calcification at abdominal CT: application of a deep learning-based algorithm to a longitudinal screening cohort. Abdom Radiol (NY) 2019;44(8):2921–2928.
22. Graffy PM, Liu J, Pickhardt PJ, Burns JE, Yao J, Summers RM. Deep learning-based muscle segmentation and quantification at abdominal CT: application to a longitudinal adult screening cohort for sarcopenia assessment. Br J Radiol 2019;92(1100):20190327.
23. Pickhardt PJ. Imaging and Screening for Colorectal Cancer with CT Colonography. Radiol Clin North Am 2017;55(6):1183–1196.
24. Sandfort V, Pickhardt PJ, Summers RM. Data augmentation using generative adversarial networks to improve generalizability in CT segmentation tasks. submitted.
25. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. Cham, Switzerland: Springer International, 2016; 424–432.
26. Kayalibay B, Jensen G, van der Smagt P. CNN-based segmentation of medical imaging data. arXiv 2017; 1701.03056 [preprint]. https://arxiv.org/abs/1701.03056. Posted July 25, 2017.
27. Medical Segmentation Decathlon: Generalisable 3D Semantic Segmentation. In.
28. Zhu JY, Park T, Isola P, Efros AA. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. 2017 IEEE International Conference on Computer Vision (ICCV), 2017; 2242-2251.
29. Boyce CJ, Pickhardt PJ, Kim DH, et al. Hepatic steatosis (fatty liver disease) in asymptomatic adults identified by unenhanced low-dose CT. AJR Am J Roentgenol 2010;194(3):623–628.
30. Hahn L, Reeder SB, Muñoz del Rio A, Pickhardt PJ. Longitudinal Changes in Liver Fat Content in Asymptomatic Adults: Hepatic Attenuation on Unenhanced CT as an Imaging Biomarker for Steatosis. AJR Am J Roentgenol 2015;205(6):1167–1172.
31. Pickhardt PJ, Hahn L, Muñoz del Rio A, Park SH, Reeder SB, Said A. Natural history of hepatic steatosis: observed outcomes for subsequent liver and cardiovascular complications. AJR Am J Roentgenol 2014;202(4):752–758.
32. Pickhardt PJ, Jee Y, O'Connor SD, del Rio AM. Visceral adiposity and hepatic steatosis at abdominal CT: association with the metabolic syndrome. AJR Am J Roentgenol 2012;198(5):1100–1107.
33. Speliotes EK, Massaro JM, Hoffmann U, et al. Liver fat is reproducibly measured using computed tomography in the Framingham Heart Study. J Gastroenterol Hepatol 2008;23(6):894–899.
34. Kullberg J, Hedström A, Brandberg J, et al. Automated analysis of liver fat, muscle and adipose tissue distribution from CT suitable for large-scale studies. Sci Rep 2017;7(1):10425.
35. Graffy PM, Pickhardt PJ. Quantification of hepatic and visceral fat by CT and MR imaging: relevance to the obesity epidemic, metabolic syndrome and NAFLD. Br J Radiol 2016;89(1062):20151024.
36. Pickhardt PJ, Park SH, Hahn L, Lee SG, Bae KT, Yu ES. Specificity of unenhanced CT for non-invasive diagnosis of hepatic steatosis: implications for the investigation of the natural history of incidental steatosis. Eur Radiol 2012;22(5):1075–1082.
37. Pickhardt PJ, Malecki K, Kloke J, Lubner MG. Accuracy of Liver Surface Nodularity Quantification on MDCT as a Noninvasive Biomarker for Staging Hepatic Fibrosis. AJR Am J Roentgenol 2016;207(6):1194–1199.
38. Pickhardt PJ, Malecki K, Hunt OF, et al. Hepatosplenic volumetric assessment at MDCT for staging liver fibrosis. Eur Radiol 2017;27(7):3060–3068.
39. Lubner MG, Jones D, Kloke J, Said A, Pickhardt PJ. CT texture analysis of the liver for assessing hepatic fibrosis in patients with hepatitis C virus. Br J

Radiol 2018 Oct 11:20180153 [Epub ahead of print] https://doi.org/10.1259/bjr.20180153.
40. Lubner MG, Pickhardt PJ. Multidetector Computed Tomography for Retrospective, Noninvasive Staging of Liver Fibrosis. Gastroenterol Clin North Am 2018;47(3):569–584.
41. Pickhardt PJ, Graffy PM, Said A, et al. Multiparametric CT for Noninvasive Staging of Hepatitis C Virus-Related Liver Fibrosis: Correlation With the Histopathologic Fibrosis Score. AJR Am J Roentgenol 2019; 212(3):547–553.
42. Lee SJ, Pickhardt PJ. Opportunistic Screening for Osteoporosis Using Body CT Scans Obtained for Other Indications: the UW Experience. Clin Rev Bone Miner Metab 2017;15(3):128–137.
43. Pickhardt PJ, Pooler BD, Lauder T, del Rio AM, Bruce RJ, Binkley N. Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. Ann Intern Med 2013;158(8):588–595.
44. Pickhardt PJ, Hassan C, Laghi A, Kim DH. CT colonography to screen for colorectal cancer and aortic aneurysm in the Medicare population: cost-effectiveness analysis. AJR Am J Roentgenol 2009;192(5): 1332–1340.
45. O'Connor SD, Graffy PM, Zea R, Pickhardt PJ. Does Nonenhanced CT-based Quantification of Abdominal Aortic Calcification Outperform the Framingham Risk Score in Predicting Cardiovascular Events in Asymptomatic Adults? Radiology 2019;290(1):108–115.
46. Hassan C, Pickhardt PJ, Laghi A, et al. Computed tomographic colonography to screen for colorectal cancer, extracolonic cancer, and aortic aneurysm: model simulation with cost-effectiveness analysis. Arch Intern Med 2008;168(7):696–705.
47. Ryckman EM, Summers RM, Liu J, Munoz del Rio A, Pickhardt PJ. Visceral fat quantification in asymptomatic adults using abdominal CT: is it predictive of future cardiac events? Abdom Imaging 2015;40(1):222–226.
48. Lawrence EM, Pooler BD, Pickhardt PJ. Opportunistic Screening for Hereditary Hemochromatosis With Unenhanced CT: Determination of an Optimal Liver Attenuation Threshold. AJR Am J Roentgenol 2018;211(6):1206–1211.