10.1148/ryai.210268

# Appendix E1

## Multidetector CT Technique

As the datasets consist of CT scans collected for over a decade, they contain scanners and software of various manufactures. Dataset 1 includes 16-or 64-multidetector CT scanners from GE Medical systems (LightSpeed and subsequent models), and Dataset 2 includes 4-to 192-multidetector CT scanners from GE (LightSpeed, HiSpeed, and other models), Siemens (Definition, SOMATOM, and other models), and Philips (Brilliance, Mx8000 and other models) with various versions of soft tissue kernels, such as 'Soft', 'Br40', and 'B'. The parameters were 100–120 kVp, using a 2.5-mm to 5-mm section thickness (Dataset 1, 8/406 scans were 2.5-mm and 398/406 were 5-mm; Dataset 2, all scans were 5-mm), and patient-specific tube current settings. We selected portal venous phase scans for all measurements, which was scanned at approximately 70 seconds from the start of the injection, based on a time/density graph or 45–55 seconds after aortic threshold enhancement. The typical contrast media protocol for Dataset 1 included 50–200 mL (Omni 300) or 60–100 mL (Omni 350) of Omnipaque (GE, iohexol injection) at a rate of 3 mL/sec. Dataset 2 used 120–130 mL of Oxilan-300 (Guerbet, ioxilan) or Isovue-300 (Bracco, iopamidol injection) at a rate of 2 mL/sec.

## Deep Learning Model Development

### Training Data Used for DL Model Development

The training data used to develop the model included both public and in-house data, which contained cases of normal liver and spleen, liver tumors, liver cirrhosis, and splenomegaly. For training the first stage liver-only model we used 131 ground truth segmentations from the liver tumors challenge of the Medical Data Decathlon (32,33) and 443 ground truth segmentations for Data Decathlon hepatic vessels CT provided by Tian et al (20). These were supplemented with 24 ascites cases and 10 splenomegaly cases that were downloaded from our institute's (NIH) PACS and labeled in-house by a board-certified radiologist with 12 years of experience. A large number of cases with large liver tumors were removed. In total, we used 268 cases for training, 7 for validation, and 14 for testing. The spleen segmentations for this dataset were generated using a previously developed spleen model (34). For training the second stage liver Couinaud model we used 193 ground truth Couinaud segmentations which have been made available by Tian et al (20). The ground truth labels are associated with 193 CT scans from the hepatic vessels challenge of the Medical Data Decathlon (32,33). They were supplemented by 7 splenomegaly cases and 10 ascites cases. Altogether we used 188 images for training, 7 for validation, and 15 for testing.

For training the spleen deep learning (DL) model we used 40 noncontrast CT colonography scans from (University of Wisconsin-Madison) with segmentations generated in-house for a prior work (34), 41 CECT from the Data Decathlon spleen challenge (32,33), 50 contrast-enhanced CT (CECT) from the "Beyond the Cranial Vault" challenge (35) with labels provided by Gibson et al (36), and 45 CECT from the Cancer Imaging Archive Pancreas-CT dataset (37) from the Cancer Imaging Archive (38) with labels provided by Gibson et al (36).

The CECT were supplemented with synthetic noncontrast versions generated using the UNIT image translation technique described previously (39,40). This dataset was supplemented with 7 cases with ascites and 13 splenomegaly cases downloaded from our institute's PACS and labeled by a radiologist with 12 years of experience. Altogether we used 8 cases for validation and 20 for testing.

## Details on DL Model Pipeline

The liver DL model consists of two stages-the first stage which segments the liver and spleen and a second stage which does the Couinaud segmentation using a box tightly cropped around the liver segmentation from the first stage, with nonliver pixels set to -1000 HU (See Fig E1 for illustration). The reason for segmenting both the liver and spleen in the first stage was to help the model not confuse the two organs, which have symmetrical shapes and appearance, especially in splenomegaly cases. Both stages used the same 3D U-Net (41) architecture with dropout and skip connections. The spleen DL model is a one-step 3D U-Net (See Fig E2 for illustration). The spleen DL model achieved an average Dice of 0.94 ± 0.02 on the hold-out test set of 20 cases, with a relative average volume difference (RAVD) of 0.056 ± 0.04. In a separate test set of 10 challenging cases (where the primitive model failed to segment the liver and spleen due to abnormal conditions such as ascites and splenomegaly) from the hold-out test set, it achieved a Dice of 0.95 ± 0.01 and RAVD of 0.03 ± 0.03. The runtime for inference was an average of 12 seconds per CT for the spleen segmentation and 33 seconds per CT for the Couinaud segmentation on a P100 GPU (16 Gb of memory). The standard deviations of the DL models performed on the same subject 20 times are close to zero ($< 1 \times 10^{-13}$).

The CT is first put into canonical orientation and then preprocessed by clipping the HU to a soft tissue window of -150–240 HU and rescaled by subtracting the mean and dividing by the standard deviation of the HU. After the liver is segmented by the first stage 3D U-Net, the largest connected component is extracted and the height of the liver is measured and the scan is cropped in the Z-direction from 15% of the liver height below to 15% above. The cropped-z box is then fed back into the liver segmentation model to get a refined output. This helps the model work for scans with a wide variety of fields of view. A connected components analysis is then done and the largest connected component in the segmentation is used as the liver. A tightly cropped box (with ± 1% border in the x, y, and z directions) is taken around the liver segmentation and pixels outside the segmentation are set to -1000. The image is then reclipped to -150–240 HU and renormalized and fed into the second stage U-Net. The second U-Net outputs the Couinaud segmentations. Any stray segmentation lying outside the liver segmentation from the first stage is removed.

For the spleen model, the architecture and training procedure are the same as the first stage of the liver model, but with only one class (the spleen) as output. We did not use a tightly cropped box for the spleen. The images were resampled to 192 × 192 × 128 and a "temporal ensemble" of two models (saved at 14000 and 25000 iterations of training) was used.

For training both U-Net models, we utilized the Generalized Dice loss (42), which has been shown to outperform standard Dice loss for multiclass segmentation. We trained the model using the Rectified Adam optimizer (43) with a batch size of 2. For data augmentation, we used random rotations between ± 10 degrees around one of the XYZ axes and random elastic deformations using a B spline. The ascites and splenomegaly cases were reweighted in the training sampler to comprise about 50% of the training iterations. Temporal ensembling was

used in both the first and second stages where 3–5 models taken from between 20,000–30,000 iterations of training are utilized for prediction and their results averaged.

## Multivariable Models in Predicting Cirrhosis and Advanced Fibrosis

### Materials and Methods

Using combinations of the automated measurements (whole liver volume, spleen volume, LSVR, volume proportions, attenuations), the following multivariable models were built using multivariable logistic regression.

Multivariate model using the [S]pleen volume and [L]SVR will be marked as the 'S+L model', [W]hole liver volume, [S]pleen, and [L]SVR as the 'W+S+L model'. Model using the [V]olume proportions (the volume of each Couinaud segment divided by the entire liver volume) of all the liver Couinaud segments will be marked as the 'V model', model using the [M]edian HU attenuation of all the liver Couinaud segments as the 'M model', model using the standard [D]eviation of the attenuation in all the liver Couinaud segments as the 'D model', model using the [S]pleen volume and [V]olume proportions (compared with the entire liver) of all the liver Couinaud segments as the 'S+V model', model using the [S]pleen, [L]SVR, and standard [D]eviation of the attenuation in all the liver Couinaud segments as the 'S+L+D model', and model using the [S]pleen, [L]SVR, [V]olume proportions (compared with the entire liver) of all the liver Couinaud segments, and standard [D]eviation of the attenuation in all the liver Couinaud segments as the 'S+L+V+D model'.

Before building the multivariable models, Dataset 1 was first randomly split into a ratio of 80:20. The multivariable models were built on the 80% (325/406) split of Dataset 1, then tested on the remaining 20% (81/406) of Dataset 1. The same models were also tested on Dataset 2 (See Fig 1 for illustration). AUCs between the multivariable models were also compared for noninferiority.

Since Dataset 1 consisted solely of patients with HCV, while Dataset 2 had multiple etiologies including viral hepatitis and steatohepatitis, we divided Dataset 2 into patients with HCV-only ($n$ = 79, patients with HCV in Dataset 2) and patients with non-HCV ($n$ = 128, all other etiologies) and calculated the performance of the automated measurements in predicting advanced fibrosis and cirrhosis.

### Results

In predicting cirrhosis in Dataset 1, the automated S+L model was comparable to the manual S+L model (AUCs, 0.90; CI: 0.84, 0.97 versus 0.93; CI: 0.88, 0.98 for automated versus manual S+L model, significantly noninferior with $P < .001$) (Table E6). The V model had an AUC of 0.79 (CI: 0.68, 0.89). The S+V model had similar performance with the S+L model (AUCs, 0.93; CI: 0.88, 0.98 versus 0.90; CI: 0.84, 0.97 for S+V model versus S+L model, significantly noninferior with $P < .001$). The M model and D model had AUCs lower than 0.75. However, the S+L+V+D model had an AUC of 0.94 (CI: 0.89, 0.99) that was significantly noninferior ($P < .001$) to the best performance of the manual model (manual S+L model: AUC, 0.93; CI: 0.88, 0.98). A similar pattern was observed in the prediction of advanced fibrosis.

In Dataset 2, the multivariable models had a generally lower performance compared with Dataset 1, but had a similar pattern, with S+L+V+D as the best performing model (AUCs, 0.79;

CI: 0.71, 0.87 in the Ishak staging system and 0.78; CI: 0.70, 0.86 in the Knodell HAI system for predicting cirrhosis, Table E6). However, the HCV-only subset of Dataset 2 had a generally higher performance than the whole dataset (AUCs, 0.82; CI: 0.72, 0.91 versus 0.79; CI: 0.71, 0.87 for S+L+V+D in predicting cirrhosis), and the non-HCV subset of Dataset 2 had a generally lower performance than the whole dataset (AUCs, 0.73; CI: 0.57, 0.90 versus 0.79; CI: 0.71, 0.87 for S+L+V+D in predicting cirrhosis) in predicting advanced fibrosis and cirrhosis in all multivariable models (Table E7). Receiver operating characteristic curves of the multivariable models in Dataset 1 and 2 can be found in Figure E10.

## References


32. Medical Segmentation Decathlon. http://medicaldecathlon.com/. Accessed May 29, 2021.

33. Simpson AL, Antonelli M, Bakas S, et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. arXiv preprint arXiv:1902.09063. https://arxiv.org/abs/1902.09063. Posted February 25, 2019. Accessed July 14, 2021.

34. Sandfort V, Yan K, Pickhardt PJ, Summers RM. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. Sci Rep 2019;9(1):16884.

35. Landman B, Xu Z, Iglesias J, Styner M, Langerak T, Klein A. 2015 MICCAI Multi-Atlas Labeling Beyond the Cranial Vault - Workshop and Challenge. https://doi.org/10.7303/syn3193805 Published February 13, 2015. Accessed July 14, 2021.

36. Gibson E, Giganti F, Hu Y, et al. Automatic multi-organ segmentation on abdominal CT with dense v-networks. IEEE Trans Med Imaging 2018;37(8):1822–1834.

37. Roth HR, Lu L, Farag A, et al. DeepOrgan: Multi-level Deep Convolutional Networks for Automated Pancreas Segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9349. Cham, Switzerland: Springer, 2015; 556–564.

38. Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 2013;26(6):1045–1057.

39. Zhu Y, Elton DC, Lee S, Pickhardt PJ, Summers RM. Image translation by latent union of subspaces for cross-domain plaque detection. arXiv preprint arXiv:2005.11384. arXiv preprint arXiv:2005.11384. https://arxiv.org/abs/2005.11384. Posted May 22, 2020. Accessed July 14, 2021.

40. Zhu Y, Tang Y, Tang Y, et al. Cross-domain Medical Image Translation by Shared Latent Gaussian Mixture Model. In: Martel AL, Abolmaesumi P, Stoyanov D, et al, eds. Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. MICCAI 2020. Lecture Notes in Computer Science, vol 12262. Cham, Switzerland: Springer, 2020; 379–389.

41. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In: Ourselin S, Joskowicz L, Sabuncu M, Unal G, Wells W, eds. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. MICCAI 2016. Lecture Notes in Computer Science, vol 9901. Cham, Switzerland: Springer, 2016; 424–432.

42. Sudre CH, Li W, Vercauteren T, Ourselin S, Cardoso MJ. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. In: Cardoso MJ, Arbel T, Carneiro G, et al, eds. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA ML-CDS 2017 2017. Lecture Notes in Computer Science, vol 10553. Cham, Switzerland: Springer, 2017; 240–248.

43. Liu L, Jiang H, He P, et al. On the Variance of the Adaptive Learning Rate and Beyond. arXiv preprint arXiv:1908.03265. arXiv preprint arXiv:1908.03265. https://arxiv.org/abs/1908.03265. Posted August 8, 2019. Accessed July 14, 2021.

## Table E1

### Automatically Measured Parameter Values of Each Fibrosis Stage (Metavir System) in Dataset 1

| | | Dataset 1 | | | | | |
|---|---|---|---|---|---|---|---|
| | | Metavir System | | | | | |
| | | F0 (*n* = 47) | F1 (*n* = 62) | F2 (*n* = 90) | F3 (*n* = 59) | F4 (*n* = 148) | Kruskal-Wallis *P* Value |
| Whole liver volume (mL) | | 1722 (1501–1955) | 1682 (1548–1942) | 1772 (1566–2010) | 1989 (1752–2247) | 1803 (1294–2141) | < 0.001 |
| Spleen volume (mL) | | 249 (210–344) | 268 (213–355) | 301 (231–406) | 380 (274–623) | 736 (500–1049) | < 0.001 |
| LSVR | | 0.32 (0.28–0.37) | 0.31 (0.27–0.34) | 0.33 (0.28–0.41) | 0.37 (0.3–0.45) | 0.48 (0.39–0.56) | < 0.001 |
| Volume proportions of segments compared with entire liver (%)* | I | 4.83 (4.13–5.49) | 4.54 (4.03–5.09) | 4.39 (3.96–4.88) | 4.61 (4.02–5.26) | 5.12 (4.12–6.44) | < 0.001 |
| | II | 12.71 (11.73–14.63) | 13.16 (11.18–14.63) | 13.68 (12.12–15.10) | 14.34 (11.14–16.56) | 16.24 (13.67–19.16) | < 0.001 |
| | III | 5.82 (3.91–8.09) | 5.84 (4.27–7.46) | 6.56 (4.26–9.57) | 8.14 (5.56–10.6) | 9.49 (6.57–12.61) | < 0.001 |
| | IV | 11.32 (9.61–12.56) | 10.85 (9.18–12.24) | 10.84 (9.27–12.65) | 9.79 (8.31–11.23) | 9.70 (8.03–11.99) | 0.003 |
| | V | 13.89 (10.36–15.7) | 13.80 (11.14–16.06) | 12.95 (11.25–14.95) | 13.69 (11.79–16.00) | 10.59 (8.58–13.88) | < 0.001 |
| | VI | 11.30 (10.03–14.17) | 12.09 (10.25–14.61) | 12.39 (10.31–15.5) | 12.21 (10.13–14.21) | 10.88 (7.76–13.66) | 0.002 |
| | VII | 18.37 (15.85–21.16) | 17.95 (15.61–21.23) | 17.67 (14.86–19.80) | 17.67 (15.65–21.26) | 17.52 (15.07–19.64) | 0.56 |
| | VIII | 20.29 (18.80–23.23) | 20.69 (18.00–22.98) | 19.74 (16.90–21.95) | 18.97 (16.38–20.53) | 17.84 (15.53–19.91) | < 0.001 |
| Attenuation of segments (Median HU) | I | 27.47 (23–32.13) | 30.05 (25.75–36.12) | 29.49 (26.15–34.13) | 26.64 (24.4–30.11) | 27.83 (23.81–33.34) | 0.01 |
| | II | 33.81 (28.84–39.52) | 36.71 (30.83–41.7) | 35.83 (31.14–40.8) | 33.7 (30.42–39.12) | 29.72 (24.76–36.58) | < 0.001 |
| | III | 37.2 (32.34–44.18) | 38.91 (32.35–45.54) | 40.63 (35.18–45.12) | 38.25 (32.68–42.55) | 31.94 (26.15–38.1) | < 0.001 |
| | IV | 32.89 (28.58–40.54) | 35.55 (30.78–43.96) | 36.67 (30.62–42.6) | 36.22 (31.42–43) | 33.79 (27.56–43.32) | 0.19 |
| | V | 29.18 (25.58–33.84) | 29.63 (26.14–36.05) | 31.56 (26.1–36.53) | 28.91 (25.25–35.07) | 28.34 (23.07–34.18) | 0.19 |
| | VI | 23.29 (20.08–27.97) | 24.17 (21.08–27.43) | 25.01 (20.65–30.52) | 23.18 (20.73–26.01) | 21.7 (18.04–25.67) | < 0.001 |
| | VII | 30 (25.19–37.59) | 31.17 (26.58–43.57) | 32.89 (28.97–40.67) | 35.24 (27.58–44.64) | 28.64 (21.94–40.01) | 0.002 |
| | VIII | 39.05 (33.71–50.58) | 42.56 (37.13–56.78) | 44.99 (38.48–54.3) | 47.68 (38.02–56.9) | 40.59 (28.7–50.89) | 0.003 |
| Attenuation of segments (Standard deviation) | I | 112 (97.5–124) | 115.5 (101–127) | 110 (99.25–123) | 107 (93.5–115) | 92 (80–110) | < 0.001 |
| | II | 118 (103.5–129) | 117 (101.25–131.75) | 113.5 (105.25–127) | 107 (94–119.5) | 93.5 (82–108.25) | < 0.001 |
| | III | 113 (102.5–129) | 114.5 (98.25–130) | 112.5 (102–125.75) | 105 (92.5–119) | 94 (82.75–109.25) | < 0.001 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | IV | 118 (104–133) | 118.5 (103–132.75) | 115 (105.25–129.75) | 108 (95–119) | 93 (80.75–110) | < 0.001 |
| | SV | 116 (106–135) | 123.5 (105.5–135.75) | 118 (106–129.25) | 112 (96.5–124) | 95.5 (82.75–111) | < 0.001 |
| | VI | 118 (108–137) | 122.5 (105.5–136.75) | 118 (107–131.5) | 113 (95.5–123.5) | 99 (84–113) | < 0.001 |
| | VII | 117 (102.5–133) | 119.5 (100.5–133.25) | 115 (103–126.75) | 107 (91–118) | 94.5 (80–111.25) | < 0.001 |
| | VIII | 118 (103.5–136) | 121 (104.5–135.75) | 116 (104.25–128.25) | 109 (92–121) | 94 (82.75–110) | < 0.001 |

Note.—The values in the cell represent the median and 1st–3rd interquartile values of each parameter.

* Volume proportion was calculated by the volume of each Couinaud segment divided by the entire liver volume.

## Table E2

## Automatically Measured Parameter Values of Each Fibrosis Stage (Knodell HAI System) in Dataset 2

| | | Dataset 2 | | | | |
|---|---|---|---|---|---|---|
| | | Knodell HAI | | | | |
| | | 0 (*n* = 58) | 1 (*n* = 56) | 3 (*n* = 52) | 4 (*n* = 41) | Kruskal-Wallis *P* Value |
| Whole liver volume (mL) | | 1970 (1627–2328) | 1643 (1438–2118) | 2032 (1402–2429) | 1787 (1331–2109) | 0.045 |
| Spleen volume (mL) | | 284 (205–385) | 290 (194–409) | 441 (247–625) | 360 (290–717) | 0.002 |
| LSVR | | 0.33 (0.28–0.36) | 0.32 (0.28–0.39) | 0.33 (0.24–0.45) | 0.45 (0.36–0.6) | < 0.001 |
| Volume proportions of segments compared with entire liver (%)* | I | 4.10 (3.52–4.63) | 4.31 (3.77–5.04) | 4.26 (3.78–5.05) | 4.55 (3.84–5.96) | 0.06 |
| | II | 13.39 (11.46–15.28) | 13.23 (10.87–15.05) | 12.76 (11.06–15.14) | 15.78 (13.32–18.42) | < 0.001 |
| | III | 6.50 (4.37–9.45) | 7.07 (4.10–9.20) | 6.26 (3.77–10.05) | 10.43 (7.73–13.79) | < 0.001 |
| | IV | 11.55 (10.49–13.10) | 10.62 (9.09–12.95) | 9.16 (7.19–11.36) | 10.19 (8.14–11.92) | < 0.001 |
| | V | 13.49 (11.42–15.55) | 14.53 (12.10–16.21) | 13.58 (10.90–16.41) | 11.57 (8.41–14.54) | 0.006 |
| | VI | 12.13 (9.85–14.19) | 12.55 (8.58–15.54) | 14.14 (10.00–17.18) | 10.88 (8.30–13.60) | 0.12 |
| | VII | 17.05 (15.23–18.81) | 17.75 (14.79–20.39) | 17.67 (14.68–21.67) | 16.63 (14.4–18.91) | 0.36 |
| | VIII | 20.60 (18.57–23.80) | 19.88 (18.10–22.26) | 18.71 (16.29–20.38) | 17.67 (15.70–19.45) | < 0.001 |
| Attenuation of segments (Median HU) | I | 28.09 (23.28–35.14) | 26.8 (24.11–31.82) | 27.13 (24.82–32.21) | 29.48 (23.65–37.47) | 0.87 |
| | II | 33.4 (30.12–41) | 33.29 (26.76–38.2) | 32.23 (30–38.8) | 29.41 (24.73–38.06) | 0.29 |
| | III | 36.19 (31.52–40.33) | 34.16 (27.88–39.31) | 35.72 (30.39–39.1) | 32.42 (27.41–38.94) | 0.33 |
| | IV | 33.86 (28.17–37.92) | 32.37 (27.87–40.19) | 34.14 (28.9–46.46) | 35.46 (29.11–42.85) | 0.7 |
| | V | 27.82 (23.04–33.79) | 27.56 (23.48–31.96) | 30.06 (22.97–33.34) | 27.52 (23.06–34.84) | 0.92 |
| | VI | 23.47 (20.26–27.45) | 21.79 (17.49–27.42) | 22.07 (18.57–25.14) | 21.37 (16.42–25.22) | 0.31 |
| | VII | 27.52 (24.75–52.24) | 28.05 (23.73–82.03) | 27.12 (24.43–38.13) | 28.21 (21.89–46.53) | 0.88 |
| | VIII | 36.94 (31.45–79.9) | 36.69 (32.98–95.68) | 37.55 (33.13–50.14) | 39.12 (29.65–61.58) | 0.97 |
| Attenuation of segments | I | 100 (83.75–108.5) | 100 (83.75–109) | 88.5 (77.5–103) | 91 (79–103) | 0.11 |
| | II | 95.5 | 101.5 | 88 | 91 | 0.21 |

| | | | | | | |
|---|---|---|---|---|---|---|
| (Standard deviation) | | (83.25–110.75) | (84–113) | (75.75–108.25) | (81–104) | |
| | III | 95.5 (82.25–108.5) | 100 (82.75–114.25) | 87 (74.75–104) | 94 (81–104) | 0.14 |
| | IV | 96 (82.25–110) | 99.5 (81.5–114) | 87 (73–104.75) | 92 (80–104) | 0.15 |
| | V | 96.5 (82.25–111.75) | 102 (84.25–116) | 88.5 (76.75–111.5) | 92 (83–106) | 0.14 |
| | VI | 99.5 (83.25–111) | 100.5 (83.75–118.25) | 90 (77.75–113.25) | 94 (85–107) | 0.19 |
| | VII | 93 (80–107) | 96 (83.25–110.25) | 84 (74.5–108.25) | 90 (80–102) | 0.23 |
| | VIII | 95.5 (79.5–109) | 99.5 (83.25–112.5) | 87 (75.75–110) | 91 (83–105) | 0.28 |

Note.—The values in the cell represent the median and 1st–3rd interquartile values of each parameter.

* Volume proportion was calculated by the volume of each Couinaud segment divided by the entire liver volume.

## Table E3

## Automatically Measured Parameter Values of Each Fibrosis Stage (Ishak Scoring System) in Dataset 2

| | | Dataset 2 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Ishak Scoring System | | | | | | | |
| | | 0 (*n* = 58) | 1 (*n* = 32) | 2 (*n* = 24) | 3 (*n* = 29) | 4 (*n* = 22) | 5 (*n* = 10) | 6 (*n* = 32) | Kruskal-Wallis *P* Value |
| Whole liver volume (mL) | | 1922 (1627–2315) | 1724 (1382–2279) | 1603 (1475–2118) | 2021 (1427–2291) | 2065 (1308–2459) | 1658 (1399–1925) | 1883 (1329–2353) | 0.19 |
| Spleen volume (mL) | | 284 (200–385) | 292 (189–380) | 283 (204–429) | 341 (220–506) | 511 (277–628) | 379 (307–720) | 466 (294–728) | 0.006 |
| LSVR | | 0.33 (0.28–0.36) | 0.34 (0.27–0.39) | 0.31 (0.28–0.36) | 0.3 (0.23–0.44) | 0.34 (0.3–0.44) | 0.54 (0.33–0.62) | 0.45 (0.37–0.6) | < 0.001 |
| Volume proportions of segments compared with entire liver (%)* | I | 4.10 (3.52–4.62) | 4.17 (3.76–4.60) | 4.52 (4.15–5.21) | 4.13 (3.75–4.71) | 4.61 (3.90–5.50) | 4.83 (3.94–5.91) | 4.54 (3.89–5.81) | 0.08 |
| | II | 13.43 (11.46–15.37) | 13.62 (11.26–15.05) | 12.08 (10.12–14.04) | 11.86 (10.54–14.5) | 14.06 (12.53–15.74) | 17.10 (13.89–21.54) | 15.83 (13.22–18.18) | < 0.001 |
| | III | 6.50 (4.37–9.43) | 6.91 (5.09–8.96) | 7.56 (3.49–9.30) | 6.14 (3.57–10.14) | 6.40 (5.16–9.05) | 8.14 (7.72–16.96) | 10.67 (8.86–13.86) | 0.003 |
| | IV | 11.64 (10.74–13.10) | 11.05 (9.3–13.67) | 9.88 (9.09–11.42) | 8.84 (7.11–10.93) | 9.31 (7.97–11.67) | 9.59 (6.96–12.84) | 10.24 (8.11–11.92) | 0.002 |
| | V | 13.55 (11.42–15.61) | 14.47 (12.13–15.78) | 14.54 (12.10–16.05) | 14.75 (10.91–16.52) | 13.29 (10.23–15.69) | 9.81 (6.75–14.03) | 11.97 (9.7–14.67) | 0.07 |
| | VI | 12.11 (9.79–14.18) | 12.94 (8.93–14.62) | 12.50 (9.08–15.84) | 14.46 (10.00–17.44) | 14.14 (10.72–16.23) | 11.41 (8.63–15.37) | 10.87 (7.8–13.36) | 0.27 |
| | VII | 17.01 (15.22–18.81) | 16.66 (13.52–19.61) | 18.32 (16.69–21.35) | 18.03 (14.92–22.03) | 17.65 (14.37–21.41) | 16.75 (16.03–19.18) | 16.38 (14.38–18.89) | 0.15 |
| | VIII | 20.60 (18.57–23.8) | 19.71 (18.76–21.8) | 20.50 (17.37–22.39) | 18.96 (16.9–20.43) | 18.29 (15.49–20.45) | 17.73 (14.28–18.46) | 17.37 (16.12–19.46) | < 0.001 |
| Attenuation of segments (Median HU) | I | 28.09 (23.28–35.14) | 26.52 (24.14–30.38) | 28.74 (24.08–36.5) | 27.27 (24.89–30.69) | 27.23 (25.09–32.3) | 27.73 (23.82–64.71) | 28.93 (23.52–36.28) | 0.91 |
| | II | 33.4 (30.12–41) | 34 (25.9–37.08) | 32.75 (29.38–40.12) | 32.69 (30.96–38.89) | 31.37 (27.26–38.32) | 29.08 (26.38–50.01) | 29.86 (24.38–37.83) | 0.52 |
| | III | 36.19 (31.87–40.33) | 33.66 (27.88–38.92) | 34.8 (27.45–40.09) | 36.31 (31.57–39.49) | 35.62 (30.51–40.5) | 33.78 (28.5–35.9) | 29.82 (26.6–36.5) | 0.33 |
| | IV | 33.86 (28.94–37.92) | 31.48 (26.69–37.24) | 35.18 (29.43–43.26) | 40.52 (30.67–46.57) | 34.54 (29.22–43.57) | 33.77 (30.58–38.84) | 34.98 (27.58–43.27) | 0.73 |
| | V | 28.13 (23.21–33.79) | 27.56 (23.18–32.27) | 25.76 (23.42–31.53) | 30.53 (23.85–33.59) | 28.96 (23.6–33.35) | 28.68 (25.74–36.54) | 25.52 (21.56–32.21) | 0.76 |
| | VI | 23.64 (20.26–28.24) | 21.49 (17.1–27.57) | 22.14 (18.13–25.31) | 22.37 (19.18–25.4) | 21.84 (18.66–28.9) | 21.94 (17.07–29.39) | 20.51 (15.87–24.92) | 0.38 |
| | VII | 27.52 (24.75–52.24) | 27.33 (23.22–50.11) | 28.3 (25.42–86.64) | 27.23 (24.64–46.75) | 27.22 (23.27–40.75) | 28.27 (24.08–37.62) | 26.83 (21.57–52.1) | 0.81 |
| | VIII | 36.94 | 35.72 | 37.97 | 40.93 | 34.54 | 36 | 39.09 | 0.52 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | (31.45–79.9) | (32.12–57.95) | (34.28–103.86) | (34.46–76.58) | (29.25–40.02) | (33.01–52.34) | (28.66–65.62) | |
| Attenuation of segments (Standard deviation) | I | 100.5 (83.75–110.5) | 91 (80.5–103.5) | 107.5 (87.75–115.75) | 94 (83–108) | 85.5 (70–94.75) | 83 (79.25–93) | 93 (79–103.25) | 0.04 |
| | II | 97.5 (83.25–111) | 93.5 (79–104.5) | 108 (85.5–117.25) | 95 (81–109) | 79 (68.25–100.5) | 81.5 (79.5–97.5) | 92.5 (81.75–104.5) | 0.06 |
| | III | 97 (82.25–109) | 95 (80–104) | 105 (86.75–119.75) | 95 (83–107) | 80.5 (66.5–97.25) | 85 (79–97.25) | 92.5 (81–104.25) | 0.06 |
| | IV | 96 (82.25–110.75) | 92.5 (78.25–105) | 110 (84.25–117.5) | 96 (78–111) | 78.5 (65.25–92.75) | 83 (80–95.5) | 93 (80.75–104.5) | 0.04 |
| | V | 98.5 (82.25–112) | 96.5 (80.75–106) | 114 (87.75–125.25) | 96 (84–116) | 81 (69.5–93.5) | 89 (84–96.5) | 92 (81–106.25) | 0.03 |
| | VI | 101 (83.25–112.5) | 96 (80.5–103.75) | 114.5 (88–125.5) | 97 (85–116) | 84 (66.25–93.25) | 89.5 (85.5–99) | 93.5 (81.75–107.5) | 0.02 |
| | VII | 94.5 (80–107.75) | 91 (78.75–100) | 107.5 (84.75–116) | 93 (81–112) | 80.5 (61–85.75) | 84.5 (80–93.75) | 91 (80.5–104.25) | 0.03 |
| | VIII | 96 (79.5–110.5) | 93.5 (79.25–101.25) | 109.5 (84.75–118.75) | 93 (83–112) | 80.5 (65.5–87.5) | 86.5 (81.5–94.5) | 92 (83.5–105.25) | 0.04 |

Note.—The values in the cell represent the median and 1st–3rd interquartile values of each parameter.

* Volume proportion was calculated by the volume of each Couinaud segment divided by the entire liver volume.

## Table E4

## Evaluation of Automated Segmentation Compared with Manual Segmentation (Reader 2) Between Dataset 1 and Dataset 2

| | Sample of Dataset 1 (*n* = 35) | Sample of Dataset 2 (*n* = 35) | Wilcoxon *P* Value |
|---|---|---|---|
| Dice similarity coefficient | | | |
| Whole liver volume | 0.981 (0.978–0.985) | 0.979 (0.971–0.981) | 0.02 |
| Spleen volume | 0.957 (0.952–0.962) | 0.950 (0.937–0.960) | 0.05 |
| Segments I+II+III | 0.916 (0.906–0.932) | 0.925 (0.909–0.935) | 0.54 |
| Segments IV+V+VI+VII+VIII | 0.968 (0.963–0.971) | 0.965 (0.951–0.971) | 0.07 |
| Segment I | 0.637 (0.595–0.703) | 0.656 (0.590–0.692) | 0.99 |
| Segment II | 0.910 (0.900–0.929) | 0.917 (0.867–0.938) | 0.71 |
| Segment III | 0.875 (0.823–0.901) | 0.877 (0.812–0.902) | 0.83 |
| Segment IV | 0.822 (0.803–0.862) | 0.854 (0.793–0.894) | 0.29 |
| Segment V | 0.839 (0.806–0.88) | 0.820 (0.766–0.869) | 0.41 |
| Segment VI | 0.847 (0.788–0.875) | 0.799 (0.751–0.855) | 0.12 |
| Segment VII | 0.887 (0.834–0.916) | 0.877 (0.854–0.921) | 0.58 |
| Segment VIII | 0.856 (0.805–0.872) | 0.872 (0.825–0.895) | 0.16 |
| Mean Hausdorff distance (mm) | | | |
| Whole liver volume | 1.1 (0.9–1.2) | 1.2 (0.9–1.6) | 0.23 |
| Spleen volume | 0.7 (0.5–0.9) | 0.7 (0.6–0.9) | 0.44 |
| Segments I+II+III | 1.1 (0.9–1.3) | 1.1 (0.9–1.7) | 0.27 |
| Segments IV+V+VI+VII+VIII | 1.5 (1.3–1.7) | 1.6 (1.2–1.9) | 0.29 |
| Segment I | 4.7 (3.5–6.0) | 3.9 (3.4–5.1) | 0.11 |
| Segment II | 1.3 (1.1–1.7) | 1.3 (1.0–2.1) | 0.76 |
| Segment III | 2.0 (1.3–2.5) | 1.7 (1.3–3.2) | 0.70 |
| Segment IV | 2.3 (1.8–2.7) | 1.7 (1.4–2.9) | 0.21 |
| Segment V | 3.1 (2.3–4.5) | 3.4 (2.3–4.2) | 0.92 |
| Segment VI | 3.0 (2.0–4.1) | 3.1 (2.3–4.5) | 0.46 |
| Segment VII | 2.2 (1.5–3.4) | 2.3 (1.4–3.3) | 0.99 |
| Segment VIII | 3.0 (2.4–3.9) | 2.7 (2.1–3.8) | 0.46 |

Note.—The values in the cell represent the median and 1st–3rd interquartile values of each parameter.

## Table E5

## Automatically Measured Parameter Values Between Dataset 1 and Dataset 2

| | | Dataset 1 | | Dataset 2 | | Wilcoxon *P* Value |
|---|---|---|---|---|---|---|
| | | Median (IQR) | *n* | Median (IQR) | *n* | |
| Whole liver volume (mL) | All stages | 1788 (1528–2090) | 406 | 1855 (1482–2291) | 207 | 0.17 |
| | Cirrhosis | 1803 (1294–2141) | 148 | 1787 (1331–2109) | 41 | 0.75 |
| | Advanced fibrosis* | 1825 (1531–2144) | 207 | 1927 (1331–2291) | 93 | 0.58 |
| Spleen volume (mL) | All stages | 383 (260–668) | 406 | 328 (209–522) | 207 | < 0.001 |
| | Cirrhosis | 736 (499–1049) | 148 | 360 (290–717) | 41 | < 0.001 |
| | Advanced fibrosis* | 479 (294–794) | 207 | 422 (262–640) | 93 | 0.06 |
| LSVR | All stages | 0.37 (0.29–0.46) | 406 | 0.34 (0.28–0.45) | 207 | 0.01 |
| | Cirrhosis | 0.48 (0.39–0.56) | 148 | 0.45 (0.36–0.60) | 41 | 0.97 |
| | Advanced fibrosis* | 0.40 (0.31–0.50) | 207 | 0.38 (0.27–0.55) | 93 | 0.44 |

Note.—The values in the cell represent the median and 1st–3rd interquartile values of each parameter.

* Advanced fibrosis includes cirrhosis patients.

## Table E6

## AUC Values of Multivariable Models for Predicting Cirrhosis and Advanced Fibrosis

| | Test set (20%) of Dataset 1 | | Dataset 2 | | | |
|---|---|---|---|---|---|---|
| | Metavir System | | Knodell HAI System | | Ishak Staging System | |
| | Advanced Fibrosis (F0–1 vs F2–4) * | Cirrhosis (F0–3 vs F4) | Advanced Fibrosis (0–1 vs 3–4) * | Cirrhosis (0–3 vs 4) | Advanced Fibrosis (0–2 vs 3–6)* | Cirrhosis (0–4 vs 5–6) |
| Manual measurements | | | | | | |
| S+L | 0.85 (0.75–0.95) | 0.93 (0.88–0.98) | — | — | — | — |
| W+S+L | 0.85 (0.74–0.95) | 0.93 (0.88–0.99) | — | — | — | — |
| Automated measurements | | | | | | |
| S+L | 0.82 (0.71–0.93) | 0.9 (0.84–0.97) | 0.72 (0.65–0.8) | 0.77 (0.69–0.85) | 0.72 (0.65–0.8) | 0.79 (0.71–0.86) |
| W+S+L | 0.81 (0.7–0.92) | 0.91 (0.84–0.98) | 0.73 (0.65–0.8) | 0.75 (0.67–0.84) | 0.73 (0.65–0.8) | 0.77 (0.69–0.85) |
| V | 0.73 (0.61–0.86) | 0.79 (0.68–0.89) | 0.71 (0.63–0.78) | 0.76 (0.66–0.86) | 0.71 (0.63–0.78) | 0.76 (0.66–0.85) |
| M | 0.68 (0.55–0.81) | 0.7 (0.58–0.83) | 0.53 (0.45–0.61) | 0.59 (0.49–0.69) | 0.53 (0.45–0.61) | 0.61 (0.51–0.71) |
| D | 0.75 (0.63–0.87) | 0.74 (0.62–0.85) | 0.59 (0.51–0.67) | 0.61 (0.51–0.7) | 0.59 (0.51–0.67) | 0.6 (0.51–0.69) |
| S+V | 0.79 (0.68–0.91) | 0.93 (0.88–0.98) | 0.74 (0.67–0.81) | 0.77 (0.69–0.85) | 0.74 (0.67–0.81) | 0.78 (0.71–0.86) |
| S+L+D | 0.8 (0.69–0.91) | 0.91 (0.85–0.97) | 0.7 (0.63–0.78) | 0.75 (0.67–0.84) | 0.7 (0.63–0.78) | 0.76 (0.68–0.85) |
| S+L+V+D | 0.8 (0.69–0.91) | 0.94 (0.89–0.99) | 0.71 (0.64–0.78) | 0.78 (0.7–0.86) | 0.71 (0.64–0.78) | 0.79 (0.71–0.87) |
| *P* value between multivariable model AUCs (noninferiority†) | | | | | | |
| Manual S+L vs Automated S+L | < 0.001 | < 0.001 | — | — | — | — |
| Manual S+L vs Automated S+L+V+D | < 0.001 | < 0.001 | — | — | — | — |
| Automated S+L vs Automated W+S+L | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Automated S+L vs Automated S+V | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Automated S+L vs Automated S+L+D | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Automated S+L vs Automated S+L+V+D | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

Note.—Dataset 1 uses the Metavir biopsy staging system and Dataset 2 uses both the Knodell histologic activity index (HAI) and Ishak staging system. The values in the cell represent the area under the receiver operating characteristic curve (AUC) and DeLong 95% confidence intervals (CI). Volume proportion was calculated by the volume of each Couinaud segment divided by the entire liver volume. LSVR = liver segmental volume ratio. Multivariable model names – [S+L]: Multivariable model using [S]pleen volume and [L]SVR, [W+S+L]; Multivariable model using [W]hole liver volume, [S]pleen volume, and [L]SVR; [V]: Multivariable model using [V]olume proportions of all the liver Couinaud segments; [M]: Multivariable model using [M]edian Hounsfield Unit attenuation of all the liver Couinaud segments; [D]: Multivariable model using Standard [D]eviation of the attenuation in all the liver Couinaud segments; [S+V]: Multivariable model using [S]pleen volume and [V]olume proportions of all the liver Couinaud segments; [S+L+D]: Multivariable model using [S]pleen volume, [L]SVR, and Standard [D]eviation of the attenuation in all the liver Couinaud segments; [S+L+V+D]: Multivariable model using [S]pleen volume, [L]SVR, [V]olume proportions, and Standard [D]eviation of the attenuation in all the liver Couinaud segments.

* Advanced fibrosis includes patients with cirrhosis.

† $P < .05$ indicates significant noninferiority between two methods.

## Table E7

## AUC Values of Dataset 2 Divided into HCV and Non-HCV Groups for Predicting Cirrhosis and Advanced Fibrosis

| | | Entire Dataset 2 (*n* = 207) | | HCV-only patients in Dataset 2 (*n* = 79) | | Non-HCV patients in Dataset 2 (*n* = 128) | |
|---|---|---|---|---|---|---|---|
| | | Advanced Fibrosis (Ishak 0–2 vs 3–6)* | Cirrhosis (Ishak 0–4 vs 5–6) | Advanced Fibrosis (Ishak 0–2 vs 3–6)* | Cirrhosis (Ishak 0–4 vs 5–6) | Advanced Fibrosis (Ishak 0–2 vs 3–6)* | Cirrhosis (Ishak 0–4 vs 5–6) |
| Univariable | Whole liver volume (mL) | 0.49 (0.41–0.57) | 0.46 (0.44–0.64) | 0.7 (0.58–0.82) | 0.61 (0.47–0.74) | 0.4 (0.49–0.72) | 0.31 (0.53–0.85) |
| | Spleen volume (mL) | 0.66 (0.58–0.73) | 0.65 (0.55–0.74) | 0.79 (0.69–0.89) | 0.68 (0.56–0.8) | 0.59 (0.48–0.71) | 0.61 (0.45–0.77) |
| | LSVR | 0.63 (0.55–0.71) | 0.75 (0.66–0.85) | 0.65 (0.53–0.77) | 0.79 (0.67–0.9) | 0.59 (0.48–0.71) | 0.69 (0.52–0.87) |
| Multivariable models | S+L | 0.72 (0.65–0.8) | 0.79 (0.71–0.86) | 0.85 (0.76–0.94) | 0.82 (0.72–0.91) | 0.64 (0.53–0.75) | 0.75 (0.6–0.9) |
| | W+S+L | 0.73 (0.65–0.8) | 0.77 (0.69–0.85) | 0.85 (0.76–0.93) | 0.76 (0.65–0.88) | 0.65 (0.54–0.76) | 0.76 (0.62–0.91) |
| | V | 0.71 (0.63–0.78) | 0.76 (0.66–0.85) | 0.67 (0.55–0.79) | 0.75 (0.63–0.87) | 0.70 (0.59–0.81) | 0.72 (0.54–0.91) |
| | M | 0.53 (0.45–0.61) | 0.61 (0.51–0.71) | 0.48 (0.33–0.62) | 0.65 (0.52–0.77) | 0.56 (0.45–0.67) | 0.56 (0.37–0.74) |
| | D | 0.59 (0.51–0.67) | 0.6 (0.51–0.7) | 0.61 (0.47–0.75) | 0.66 (0.54–0.78) | 0.59 (0.48–0.7) | 0.55 (0.39–0.72) |
| | S+V | 0.74 (0.67–0.81) | 0.78 (0.71–0.86) | 0.79 (0.69–0.89) | 0.78 (0.67–0.88) | 0.72 (0.61–0.82) | 0.77 (0.63–0.91) |
| | S+L+D | 0.7 (0.63–0.78) | 0.76 (0.68–0.85) | 0.82 (0.72–0.91) | 0.82 (0.72–0.91) | 0.65 (0.54–0.76) | 0.72 (0.55–0.88) |
| | S+L+V+D | 0.71 (0.64–0.78) | 0.79 (0.71–0.87) | 0.77 (0.67–0.88) | 0.82 (0.72–0.91) | 0.68 (0.58–0.79) | 0.73 (0.57–0.9) |

Note.—This table shows the performance of automated measurements in predicting advanced fibrosis or cirrhosis in all subjects of Dataset 2 (left column, *n* = 207, median age 50), a subset of Dataset 2 including only hepatitis C (HCV) patients (center column, *n* = 79, median age 47), and a subset of Dataset 2 including all other etiologies including other viral hepatitis and steatohepatitis (right column, *n* = 128, median age 53). Multivariable model names – [S+L]: Multivariable model using [S]pleen volume and [L]SVR; [W+S+L]: Multivariable model using [W]hole liver volume, [S]pleen volume, and [L]SVR; [V]: Multivariable model using [V]olume proportions of all the liver Couinaud segments; [M]: Multivariable model using [M]edian Hounsfield Unit attenuation of all the liver Couinaud segments; [D]: Multivariable model using Standard [D]eviation of the attenuation in all the liver Couinaud

segments; [S+V]: Multivariable model using [S]pleen volume and [V]olume proportions of all the liver Couinaud segments; [S+L+D]: Multivariable model using [S]pleen volume, [L]SVR, and Standard [D]eviation of the attenuation in all the liver Couinaud segments; [S+L+V+D]: Multivariable model using [S]pleen volume, [L]SVR, [V]olume proportions, and Standard [D]eviation of the attenuation in all the liver Couinaud segments.

* Advanced fibrosis includes cirrhosis patients.