Published in final edited form as:

Acad Radiol. 2012 May ; 19(5): 588–598. doi:10.1016/j.acra.2012.01.015.

Assessing Hepatomegaly: Automated Volumetric Analysis of the Liver

Marius George Linguraru, DPhil^{1,2,*}, Jesse K. Sandberg¹, Elizabeth C. Jones, MD¹, Nicholas Petrick, PhD³, and Ronald M. Summers, MD, PhD¹

¹Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, 10 Center Drive, Bethesda, MD 20892

²The Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Medical Center, 111 Michigan Avenue NW, Washington DC 20010

³Center for Devices and Radiological Health, U.S. Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD 20993

Abstract

Rationale and Objectives—To define volumetric nomograms for identifying hepatomegaly and retrospectively evaluate the performance of radiologists in assessing hepatomegaly.

Materials and Methods—Livers were automatically segmented from 148 abdominal contrast-enhanced CTs: 77 normal and 71 hepatomegaly cases (diagnosed by visual inspection and/or linear liver height by radiologists). Quantified liver volumes were compared to manual measurements using overlap/error (VO/VE). Liver volumes were normalized by patient body-surface-area (BSA) from which hepatomegaly nomograms were defined (H-score) by analyzing the distribution of liver sizes in the healthy population. The H-score was validated against consensus reports. The performance of radiologists in diagnosing hepatomegaly was retrospectively evaluated.

Results—The automated segmentation of livers was robust with VO/VE of 96.2/2.2%. There were no significant differences ($p > 0.1$) between manual and automated segmentations for either normal/hepatomegaly subgroups. The average volume of normal/enlarged livers was $1.51 \pm 0.25 / 2.32 \pm 0.75$ liter. One-way ANOVA found BSA ($p = 0.004$) and gender ($p = 0.02$), but not age significantly affected normal liver volume. No significant effects were observed for two-/three-way interactions among the three variables ($p > 0.18$). The H-score cutoffs of 0.92/1.08 l/m^2 were used to define mild/massive hepatomegaly ($\pm 0.02 \text{ l/m}^2$ 95% CI). Using the H-score as the reference standard, the sensitivity of radiologists in detecting all/mild/massive hepatomegaly was 84.4/56.7/100.0% at 90.1% specificity, respectively. Radiologists disagreed in 20.9% ($n = 31$) of the diagnosed cases. **The area under the ROC curve of the H-score criterion for hepatomegaly detection was 0.98.**

Conclusion—Nomograms for the identification and grading of hepatomegaly from automatic volumetric liver assessment normalized by body-surface-area (H-score) were introduced. The H-score matches well with clinical interpretations for hepatomegaly and may improve hepatomegaly

*corresponding author: 1-202-476-3059, mlingura@cnmc.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

detection compared with height measurements or visual inspection commonly used in current clinical practice.

Keywords

hepatomegaly; volumetric analysis; liver; segmentation; nomogram

Introduction

Hepatomegaly is an abnormal enlargement of liver size and is inherently defined by a volumetric change. Patient's size has been reported to correlate with the volume of the liver [1–4]. To date, there are no defined volumetric liver nomograms to detect hepatomegaly.

Hepatic size has been an important biomarker for assessing disorders [4–13] and surgical planning [8, 11, 14–18]. Predictably, hepatic size estimates by physicians using palpation and percussion are approximate [19, 20] and adequate for diagnosing only massive hepatomegaly cases [21–24]. Blendis et al. [21] found that only half of enlarged livers detected by plain radiography were also identified by physical examination, while approximately half of the normal livers were diagnosed as enlarged.

Radiological imaging modalities such as magnetic resonance imaging, ultrasound and computed tomography (CT) are better able to assess liver size [25–29]. To detect hepatomegaly in radiological scans, radiologists commonly focus on landmark-based visual evaluations of the hemidiaphragm, displacement of the stomach, duodenum, hepatic flexure of colon, right kidney and the lower costal cartilage [30]. Nevertheless, the great diversity of normal liver shapes within the population can make landmark-based evaluation of liver volume unreliable [31]. In the case of Riedel's lobe, a normal variant, the elongated right liver lobe extends past the lower costal cartilage [32, 33], but these livers have normal volumes. Without easily accessible volume measurements, 3 to 31% of the population with this normal variation could potentially be misdiagnosed [32, 34, 35]. Additionally, differentiating between cases of mild and massive hepatomegaly, the latter category of patients may benefit from medical interventions, is unreliable without systematic quantitative measures of liver size [36, 37].

Another popular method routinely used by clinicians is manually measuring the liver's height at the mid-hepatic line (MHL) from radiological scans [7, 38]. MHL is the half-point distance between the mid-point of the spine and the outermost point on the liver surface (right side) in axial planes. This two-dimensional (2D) measure does not fully characterize the liver morphology and Riedel's lobe livers may exhibit large MHL heights. Although correlations between liver size and patient size were reported [2–4, 28, 39, 40], liver height measurements were not normalized by patient size in previous studies to detect hepatomegaly [7, 38].

CT studies have shown that manual volumetric liver measurements are important for diagnoses [1, 41]. Manual segmentations suffer from two major drawbacks: they are time consuming due to the high and expensive human interaction; and if the same organ is segmented twice (by the same or different people), the result will likely differ. Currently, most radiologists do not rely on liver volume or even height measurements for diagnosing hepatomegaly due to the lack of robust and accessible methods adapted to the image viewing software. Automated volumetric measurements of liver volume have the potential to assist clinicians in the systematic and accurate analysis of hepatic size.

In this paper, nomograms are defined to identify and grade hepatomegaly from automated liver volumetry in CT imaging. The automated liver volumetry software was previously developed by our group [Anonymous] and allows quantifying healthy and diseased livers. From readily available volumes of normal livers normalized by the patient's size (body surface area), nomograms to detect hepatomegaly are defined to assist radiologists in routine clinical assessments. The nomograms are validated against consensus reports. Finally, we evaluate retrospectively the performance of radiologists to assess hepatomegaly.

Materials and Methods

Study Patients and CT Imaging

This retrospective study follows HIPAA-compliance standard, was IRB approved and informed consent was waived. Data were acquired at the anonymized institution using contrast-enhanced CT at the portal venous phase without imaging or motion artifacts or large pathologies in the liver.

From January 2001 to March 2009, 71 consecutive subjects with clinically diagnosed hepatomegaly (48 males [mean age 43, range 18–76], 23 females [mean age 41, range 19–66]) met the inclusion criteria shown in Appendix 1. The criteria include adult patients with CT scans acquired with intravenous Isovue contrast agent at fully enhanced portal venous phase with available patient data. Additionally, the data should not have imaging or motion artifacts and missing slices in the liver. Diagnoses in the clinical reports were established by one of 11 radiologists without following any single criterion and including visual inspection and/or measurements of liver height. Cases were selected using search parameters in Radiology Information System (RIS - Cerner Corporation) and Clinical Research Information System (CRIS - Eclipsys Corporation). Radiological reports included one of the keywords: hepatomegaly, hepatosplenomegaly, enlarged liver or liver enlarged; see Appendix 1 for STARD chart. Appendix 2 presents the clinical diagnoses of the hepatomegaly data.

The control population of normal livers was formed from consecutive kidney donors enrolled at the anonymized institution from January 2001 to August 2010. Liver function parameters were not available. 77 subjects with healthy livers (33 males [mean age 43, range 17–76] and 44 females [mean age 44, range 18–72]) were selected.

The 148 cases (71 hepatomegaly patients and 77 controls) were reevaluated by two experienced radiologists working by visual inspection and occasional linear measurements of the liver size in the cranial-caudal direction. Cases were presented in random order and the two radiologists were blinded to the clinical reports (the first evaluation of data). The reassessment of data was used to create consensus reports between three radiologists: two who reevaluated the cases and one from the clinical report.

Additionally, 23 consecutive cases of partial hepatectomy fulfilled the selection criteria between January 2004 and March 2009 at the anonymized institution (15 males [mean age 49, range 33–70], 8 females [mean age 54, range 28–71]). These cases were not included in the definition or evaluation of nomograms. Instead, they were used as an independent set to test the automated segmentation method.

Contrast-enhanced CT images were acquired at portal venous phase during a single breath using fixed delays (65–70 s depending on the scanner) or bolus-tracking [42] after patients were administered 130 ml Isovue-300. Data were collected on LightSpeed Ultra/QX/I (GE Healthcare), Brilliance64 (Philips Healthcare), Definition (SIEMENS Healthcare), and Aquilion ONE (Toshiba Medical Systems) scanners with 100–240 mAs and 120 kVp. Image

resolutions ranged between 0.52–0.93 mm in axial view with 1–5 mm slice thickness. Livers were manually segmented from 20 cases by two observers supervised by an experienced radiologist. The liver heights were manually measured at MHL in all data, excepting the hepatectomy cases, by the two observers.

Lastly, 25 random pairs of supine/prone non-contrast CT datasets (1 mm slice thickness) from Walter Reed Army Medical Center in Washington, DC (14 males [mean age 58, range 51–73] and 11 females [mean age 55, range 49–68], data courtesy of J. Richard Choi, ScD MD) were analyzed for intra-observer intra-patient variability in MHL height measurements under the supervision of an experienced radiologist. Cases were acquired on LightSpeed/LightSpeed Ultra machines (GE Healthcare) [43] and declared exempt from IRB review by the anonymized institution's Office of Human Subjects Research.

Segmentation

The automated segmentation method quantified liver volumes from contrast-enhanced CT; it was presented in [Anonymous]. The software was validated on normal, hepatomegaly and partial hepatectomy cases and involved appearance, shape and location statistics of livers. A liver atlas was aligned to the patient's CT data via non-linear registration. The liver estimation was improved by geodesic active contour and adaptive convolution using patient specific contrast-enhancement. Lastly, shape and location statistics were used to optimize liver segmentation.

Liver volumes were computed automatically by summing voxel volumes in the rendered 3D segmentations. Liver heights were estimated at MHL by computer-aided diagnosis (CAD) along the sagittal plane [7].

Definition of Hepatomegaly and Nomograms

Liver volumes exhibiting hepatomegaly are outliers from the average liver size in the healthy population. To account for the relation between liver volume and patient size, volumes of livers were normalized by patients' body surface area (BSA).

To date, there are no established nomograms for volumetric assessment of hepatomegaly. Nevertheless, standards have been defined for the detection, for example, of osteoporosis [44, 45]. According to the World Health Organization, osteopenia is diagnosed if the T-score of bone mineral density is below one standard deviation (SD) from the average of healthy population. Osteoporosis is defined below 2.5 SD from the average.

We define the hepatomegaly score (H-score) as the measure of liver volume normalized by patient's BSA. The average H-score and its SD are computed from our healthy population according to the consensus reports. Only cases that were found normal by all radiologists were used in the computation of the nomograms. Following the approach used to determining osteoporosis, mild hepatomegaly is defined for H-scores above one SD from the average. Massive hepatomegaly is defined for H-scores above 2.5 SD from the average.

Performance of Radiologists Relative to the Hepatomegaly Nomogram

The performance of radiologists to diagnose hepatomegaly was retrospectively evaluated using the H-score as the reference standard. For comparison with previous nomograms based on liver height [7, 38], the diagnosis performance using the liver MHL height with a 15.5 cm cutoff was compared to both the consensus reports and H-score reference standards.

Statistical Analysis

Retrospective power analysis was performed using a two-sided test with 0.05 significance level and a binomial distribution to determine the power of our sample size ($n=148$) to detect a significant difference between normal and hepatomegaly cases. The variation of the H-score with the increase in the number of consecutive samples was analyzed. We also analyzed the receiver operating characteristic (ROC) curve to identify the operating point with the highest accuracy of the H-score defined as the sum of sensitivity and specificity; the consensus reports were used as the reference standard for hepatomegaly.

Manual and automated volumetric segmentations were compared by volume overlap (VO - twice the volume intersection over the union) and volume error (VE - volume difference relative to the manual volume). Intra- and inter-observer variability and error analysis for measuring liver height were performed following the Bland-Altman method [46]. The Mann-Whitney U test assessed significance between inter-/intra- observer, and observer-CAD agreements.

One, two and three-way analyses of variance (ANOVA) with full interaction were performed for combinations of patients' BSA, age and gender to determine the impact of these factors on normal liver volumes. Pearson correlation coefficients were calculated between liver size and patients' height, weight, BSA and age for comparisons with the same metrics reported in the literature. Student's paired t-test was used to assess the significance of correlations after testing the normal distribution of data.

Fisher's exact test assessed the significance between the sensitivity of radiologists versus the MHL height criterion to detect hepatomegaly with the H-score as the reference standard. We assessed if the sample size ($n=148$) had 0.90 power to detect a 10% (0.08) change of sensitivity between the two criteria. The Pearson correlation between the H-score and MHL height was also analyzed.

Results

The evaluation of the CAD tool [Anonymous] showed a VO of 96.2% and VE of 2.2% between automatically and manually segmented livers. There was no statistical significant difference ($p>0.1$) for either VO or VE between automated and manual segmentations on normal or abnormal cases.

The consensus reports agreed on $n=117$ cases: 74 normal and 43 hepatomegaly cases. The radiologists disagreed on $n=31$ (20.9%) of cases; $n=3$ (3.8%) were normal and $n=28$ (39.4%) were abnormal cases, according to the clinical reports. Table 1 presents the volumes, volumes normalized by BSA and MHL heights for the automatically segmented normal, enlarged and partial hepatectomy livers. There was no significant difference between the volumes of normal and partial hepatectomy livers ($p=0.7$). Significant differences were found between normal and enlarged liver volumes/heights ($p<0.001$).

Table 2 shows Pearson correlation coefficients (R) between liver measurements and patient's BSA and age. The highest correlation ($R=0.60$, $p<0.001$) was noted between liver volumes and BSA in normal cases. Moderate correlations were also observed between the enlarged liver volumes and BSA and age of patients, while a negative moderate correlation was found between the liver MHL heights and ages of controls. No significant correlations existed between liver height and patient's BSA.

One-way ANOVA found that BSA ($p=0.004$) and gender ($p=0.01$), but not age ($p=0.5$) significantly affected normal liver volume. No significant effects were observed for two- and three-way interactions among the three variables ($p>0.1$).

Bland-Altman agreement plots for height measurements between two observers and between each observer and CAD are shown in Figures 1a–c. The inter-observer variability was 0.12 ± 1.28 cm and the bias between the CAD method and each observer was 0.06 ± 1.5 cm at 95% limits of agreement. Significant correlations ($R=0.98$, $p<0.001$) were found between each observer's and CAD measurements, comparable to the inter-observer measurements correlation ($R=0.98$, $p<0.001$). Outliers in Figure 1 corresponded to unusually shaped livers, which increased the variability of MHL height measurements.

Figure 1d presents Bland-Altman agreements for consecutive intra-observer MHL height measurements on pairs of supine/prone non-contrast CT scans of normal livers ($n=25$). The variability in measurement was 1.6 ± 2.8 cm at 95% limits of agreement. These intra-observer errors were significantly larger ($p<0.001$) than the errors between CAD and manual measurements on supine scans (Figures 1b–c).

The H-score cutoff of 0.92 l/m^2 was used to identify mild hepatomegaly ($\pm 0.02 \text{ l/m}^2$ 95% CI). The H-score cutoff for massive hepatomegaly was 1.08 l/m^2 . Figure 2 shows the variation of the H-score with the number of consecutive samples used to compute it. The H-score becomes stable after approximately 50 cases. From ROC analysis, the operating point of the highest accuracy in the H-score using the consensus reports as reference standard was found at 0.93 l/m^2 , which is within the 95% CI. The sensitivity and specificity at this operating point were 97.6% and 86.4% respectively. The area under the ROC curve of the H-score criterion for hepatomegaly detection was 0.98. The statistical power of our dataset ($n=148$) for the observed $\Delta H\text{-score}=0.81$ is 0.99.

From H-score nomograms (new reference standard), the performance of radiologists was retrospectively analyzed using the clinical reports. Table 3 presents the sensitivity and specificity of radiologists to detect hepatomegaly. The performance of diagnosis by liver MHL height is also shown. The statistical power of our dataset ($n=148$) to detect an effect size of 0.08 on the sensitivity between diagnosis criteria is 0.93. There was no significant difference in the sensitivity of the detection of mild hepatomegaly between radiologists (56.7%, $p=0.09$) and MHL height. The sensitivity of radiologists to detect massive hepatomegaly was significantly higher (100.0%, $p=0.006$) than using MHL height. Radiologists also detected all cases of hepatomegaly with significantly higher sensitivity (84.4%, $p=0.007$) than the MHL height criterion.

The linear regression model in Figure 3.a showed a significant correlation between the H-scores and MHL heights of normal cases in the consensus reports ($R=0.34$, $p=0.02$). A significant correlation was also found between the H-scores and MHL heights of enlarged livers (Figure 3.b, $R=0.58$, $p<0.001$). A 15.5 cm cutoff at MHL height [7, 38] detected hepatomegaly with 83.7% sensitivity at 90.5% specificity using the consensus reports as reference standard, and 67.5% sensitivity at 88.7% specificity with the H-score as reference standard.

Discussion

This retrospective study presented the evaluation of an automated method to segment livers from contrast-enhanced CT images [Anonymous]. The automated technique was accurate and consistent with manual segmentations for liver volume and height for cohorts of normal and enlarged livers. Table 1 illustrates the average normal liver volumes from our data, which are compared to reports in the literature in Table 5. Our average diseased liver

volumes were generally larger than in the literature. The liver volumes of cases that underwent partial hepatectomy were comparable to those of normal cases. Although their average volumes were similar to those of the normal cases, we preferred to compute hepatomegaly nomograms from data from patients that did not have liver surgery.

It is generally agreed that the liver volume and patient's BSA/weight correlate (Table 4). This observation, and the statistical analysis in this paper, supports the normalization of the liver volume by BSA in the definition of the H-score. A significant correlation between normal liver volume and age was indicated in [3, 47–49]; this was confirmed in [1] or our study (Table 2). Reports also did not agree on the correlations between liver volume and patient's height [1, 2, 28, 39, 47] or liver volume and patient's age [1, 3, 47–49]. The sensitivity and specificity of the H-score to detect hepatomegaly using consensus reports as reference standard were 97.6% and 86.4% respectively with 0.98 area under the ROC curve.

Despite an abundance of papers that analyzed normal and diseased livers' volumes (Table 5), we are not aware of another effort to define nomograms to classify hepatomegaly using a volumetric measurement such as the H-score. More commonly, hepatomegaly is classified using liver MHL height. Gosink and Leymaster [7] found that 93% of normal livers had MHL heights less than 13 cm, while 75% of cases with MHL heights greater than 15.5 cm had hepatomegaly. Similarly, Rosenfield et al. [38] defined an upper limit of 15.5 cm MHL height for normal livers. Niederau et al. [50] defined the upper normal limit of liver height at the mid-clavicular line to be about 12.7 cm. The standards in [7] and [38] were based on 36 and 33 cases, respectively. Our study found that a 15.5 cm cutoff in MHL height detected hepatomegaly with 67.5% sensitivity at 88.7% specificity using the H-score as reference standard.

Four cases clinically misclassified with hepatomegaly had confirmed Riedel's lobe. Figure 4 exemplifies the segmentation of a case with Riedel's lobe with an abnormally high liver MHL height, but normal H-score.

Radiologists use experience and visual landmarks to diagnose the liver. Unsurprisingly, Table 3 showed that radiologists detected hepatomegaly with significantly higher sensitivity than the liver MHL height criterion. The H-score, a systematic quantitative approach to diagnose hepatomegaly, should be preferable to visual inspection for consistency and generalization. A uniform quantitative assessment of hepatomegaly has additional advantages; the criterion can be subdivided into categories, such as mild and massive hepatomegaly, which is less reliable using visual inspection. This finer categorization has the potential to customize treatments and monitor stages of diseases, such as cirrhosis or drug-related hepatitis. Selective patients with massive hepatomegaly can also benefit from operative intervention [36, 37, 51].

The large intra-observer errors noted for consecutive liver height measurements on supine/prone CT pairs (Figure 1d) indicate that MHL height measurement is sensitive to patient position in the scan and unreliable for detecting hepatomegaly. However, Leung et al. [2] also found substantial diurnal variations in liver volume with a minimum value between 12 am and 2 pm (17% fall average, range 9–31%). Our study did not include time and position variables.

This study has certain limitations. The distribution of the patient population with hepatomegaly may not be representative. Nevertheless, hepatomegaly patients were not admitted to the anonymized institution based on hepatomegaly diagnosis, which was a secondary radiological finding. The definition of the H-score would also benefit from a larger data sample. Finally, the reader variability in the clinical determination of

hepatomegaly is potentially very large, as radiologists disagreed in 20.9% (n=31) of cases, however representative for the clinical environment of a medium to large hospital.

We proposed to use liver size nomograms for detecting hepatomegaly. Liver size is inherently a volume and correlated to patient's BSA. Computer-aided diagnosis can offer robust and reproducible liver volume measurements in a fully automated manner to support routine radiological image analysis. In this light, the H-score defined from liver volume normalized by patient's BSA may be useful as a systematic indication of hepatomegaly.

With the introduction of volumetric measurements in routine radiological assessment, whether organ or tumor volumes, nomograms have the potential to offer systematic and reproducible tools for diagnosis. Whether supported or not by CAD, the introduction of liver volume nomograms promises to reduce variability and error in the radiological interpretation of abdominal data. Our study found 20.9% disagreement between radiologists in diagnosing hepatomegaly. If liver CAD would be adopted in the clinical flow in combination with volumetric nomograms, the diagnosis of hepatomegaly could become a seamless automated process. Finally, differentiating between cases of mild and massive hepatomegaly is unreliable without methodical quantitative measures of liver size.

Automatic volumetric liver assessment normalized by body-surface-area may improve the identification of hepatomegaly detection compared with height measurements or visual inspection commonly used in clinical practice. Using the H-score, a systematic quantitative measure of liver size, as the reference standard, radiologists detected all, mild and massive cases of hepatomegaly with 84.4%, 56.7% and 100.0% sensitivity at 90.1% specificity, respectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by the Intramural Research Programs of the National Institutes of Health, Clinical Center and the U.S. Food and Drug Administration. The authors would like to thank Zhixi Li, Furhawn Shah and Visal Desai for assistance with data analysis, and Andrew J. Dwyer, MD for critical review. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services.

References

1. Andersen V, et al. The volume of the liver in patients correlates to body weight and alcohol consumption. *Alcohol Alcohol*. 2000; 35(5):531–2. [PubMed: 11022029]
2. Leung NW, Farrant P, Peters TJ. Liver volume measurement by ultrasonography in normal subjects and alcoholic patients. *J Hepatol*. 1986; 2(2):157–64. [PubMed: 3514743]
3. Wynne HA, et al. The effect of age upon liver volume and apparent liver blood flow in healthy man. *Hepatology*. 1989; 9(2):297–301. [PubMed: 2643548]
4. Zoli M, et al. Prognostic indicators in compensated cirrhosis. *Am J Gastroenterol*. 1991; 86(10): 1508–13. [PubMed: 1928047]
5. Ellert J, Kreel L. The role of computed tomography in the initial staging and subsequent management of the lymphomas. *J Comput Assist Tomogr*. 1980; 4(3):368–91. [PubMed: 7372869]
6. Elstein D, et al. Accuracy of ultrasonography in assessing spleen and liver size in patients with Gaucher disease: comparison to computed tomographic measurements. *J Ultrasound Med*. 1997; 16(3):209–11. [PubMed: 9166820]
7. Gosink BB, Leymaster CE. Ultrasonic determination of hepatomegaly. *J Clin Ultrasound*. 1981; 9(1):37–44. [PubMed: 6792230]

8. Kawasaki S, et al. Preoperative measurement of segmental liver volume of donors for living related liver transplantation. *Hepatology*. 1993; 18(5):1115–20. [PubMed: 8225216]
9. Okamoto E, et al. Prediction of the safe limits of hepatectomy by combined volumetric and functional measurements in patients with impaired hepatic function. *Surgery*. 1984; 95(5):586–92. [PubMed: 6324403]
10. Sekiyama K, et al. Prognostic value of hepatic volumetry in fulminant hepatic failure. *Dig Dis Sci*. 1994; 39(2):240–4. [PubMed: 8313804]
11. Soyer P, et al. Hepatic metastases from colorectal cancer: influence of hepatic volumetric analysis on surgical decision making. *Radiology*. 1992; 184(3):695–7. [PubMed: 1509051]
12. Tsushima Y, Endo K. Spleen enlargement in patients with nonalcoholic fatty liver: correlation between degree of fatty infiltration in liver and size of spleen. *Dig Dis Sci*. 2000; 45(1):196–200. [PubMed: 10695635]
13. Van Thiel DH, et al. In vivo hepatic volume determination using sonography and computed tomography. Validation and a comparison of the two techniques. *Gastroenterology*. 1985; 88(6):1812–7. [PubMed: 3888769]
14. Broering DC, Sterneck M, Rogiers X. Living donor liver transplantation. *J Hepatol*. 2003; 38(Suppl 1):S119–35. [PubMed: 12591190]
15. Chen YS, et al. Evaluation of living liver donors. *Transplantation*. 2003; 75(3 Suppl):S16–9. [PubMed: 12589132]
16. Hermoye L, et al. Liver segmentation in living liver transplant donors: comparison of semiautomatic and manual methods. *Radiology*. 2005; 234(1):171–8. [PubMed: 15564393]
17. Kasai H, et al. Intraoperative color Doppler ultrasonography for partial-liver transplantation from the living donor in pediatric patients. *Transplantation*. 1992; 54(1):173–5. [PubMed: 1631931]
18. Sugarbaker PH. Surgical decision making for large bowel cancer metastatic to the liver. *Radiology*. 1990; 174(3 Pt 1):621–6. [PubMed: 2406776]
19. Delp, MH.; Manning, RT. Examination of the abdomen and Gastrointestinal and Hepatobiliary Disorders. *Major's Physical Diagnosis* 9th; 1981; Philadelphia.
20. Sherlock, S. Disease of the Liver and Biliary System. Vol. 1. Boston: Blackwell Scientific; 1981.
21. Blendis LM, et al. Observer variation in the clinical and radiological assessment of hepatosplenomegaly. *Br Med J*. 1970; 1(5698):727–30. [PubMed: 5440548]
22. Castell DO, et al. Estimation of liver size by percussion in normal individuals. *Ann Intern Med*. 1969; 70(6):1183–9. [PubMed: 5795743]
23. Meyhoff HH, Roder O, Andersen B. Palpatory estimation of liver size. Within- and between-observer variation. *Acta Chir Scand*. 1979; 145(7):479–81. [PubMed: 539331]
24. Riemenschneider PA, Whalen JP. Relative Accuracy of Estimation of Enlargement of Liver and Spleen by Radiologic and Clinical Methods. *American Journal of Roentgenology Radium Therapy and Nuclear Medicine*. 1965; 94(2):462–468.
25. Cheng YF, et al. Single imaging modality evaluation of living donors in liver transplantation: magnetic resonance imaging. *Transplantation*. 2001; 72(9):1527–33. [PubMed: 11707741]
26. Schroeder T, et al. Potential living liver donors: evaluation with an all-in-one protocol with multi-detector row CT. *Radiology*. 2002; 224(2):586–91. [PubMed: 12147860]
27. Heymsfield SB, et al. Accurate measurement of liver, kidney, and spleen volume and mass by computerized axial tomography. *Ann Intern Med*. 1979; 90(2):185–7. [PubMed: 443650]
28. Raeth U, Johnson PJ, Williams R. Ultrasound determination of liver size and assessment of patients with malignant liver disease. *Liver*. 1984; 4(5):287–93. [PubMed: 6209520]
29. Rasmussen SN. Liver volume determination by ultrasonic scanning. *Br J Radiol*. 1972; 45(536):579–85. [PubMed: 5045968]
30. Mazonakis M, et al. Comparison of two volumetric techniques for estimating liver volume using magnetic resonance imaging. *J Magn Reson Imaging*. 2002; 15(5):557–63. [PubMed: 11997897]
31. McAfee JG, Ause RG, Wagner HN Jr. Diagnostic Value of Scintillation Scanning of the Liver. *Arch Intern Med*. 1965; 116:95–110. [PubMed: 14338958]
32. Gillard JH, et al. Riedel's lobe of the liver: fact or fiction? *Clin Anat*. 1998; 11(1):47–9. [PubMed: 9445097]

33. Kudo M. Riedel's lobe of the liver and its clinical implication. *Internal Medicine*. 2000; 39(2):87–88. [PubMed: 10732821]
34. Baum S, Locko RC, Davignon MB. Functional-Anatomy and Radionuclide Imaging - Riedels Lobe of the Liver. *Anatomia Clinica*. 1982; 4(2):121–123.
35. Sham R, Sain A, Silver L. Hypertrophic Riedel's lobe of the liver. *Clin Nucl Med*. 1978; 3(3):79–81. [PubMed: 657671]
36. Aussilhou B, et al. Extended liver resection for polycystic liver disease can challenge transplantation. *Annals of Surgery*. 2010; 252(5):735–743. [PubMed: 21037428]
37. Schnellendorfer T, et al. Polycystic liver disease. *Annals of Surgery*. 2009; 250(1):112–118. [PubMed: 19561475]
38. Rosenfield AT, Schneider PB. Rapid evaluation of hepatic size on radioisotope scan. *J Nucl Med*. 1974; 15(4):237–40. [PubMed: 4818293]
39. Hausken T, et al. Estimation of the human liver volume and configuration using three-dimensional ultrasonography: effect of a high-caloric liquid meal. *Ultrasound Med Biol*. 1998; 24(9):1357–67. [PubMed: 10385959]
40. Urata K, et al. Calculation of child and adult standard liver volume for liver transplantation. *Hepatology*. 1995; 21(5):1317–21. [PubMed: 7737637]
41. Sandrasegaran K, et al. Measurement of liver volume using spiral CT and the curved line and cubic spline algorithms: reproducibility and interobserver variation. *Abdom Imaging*. 1999; 24(1):61–5. [PubMed: 9933675]
42. Goshima S, et al. Multi-detector row CT of the kidney: optimizing scan delays for bolus tracking techniques of arterial, corticomedullary, and nephrographic phases. *Eur J Radiol*. 2007; 63(3):420–6. [PubMed: 17367973]
43. Pickhardt PJ, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *N Engl J Med*. 2003; 349(23):2191–200. [PubMed: 14657426]
44. Link TM, et al. In vitro and in vivo spiral CT to determine bone mineral density: initial experience in patients at risk for osteoporosis. *Radiology*. 2004; 231(3):805–11. [PubMed: 15105454]
45. Engelke K, et al. Clinical Use of Quantitative Computed Tomography and Peripheral Quantitative Computed Tomography in the Management of Osteoporosis in Adults: the 2007 ISCD Official Positions. *Journal of Clinical Densitometry: Assessment of Skeletal Health*. 2008; 11(1):123–162.
46. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986 February 8; 1(8476):307–310. [PubMed: 2868172]
47. Kratzer W, et al. Factors affecting liver size: a sonographic survey of 2080 subjects. *J Ultrasound Med*. 2003; 22(11):1155–61. [PubMed: 14620885]
48. Boyd E. Normal variability in weight of the adult human liver and spleen. *Archives of Pathology*. 1933; 16(3):350–372.
49. Calloway NO, Foley CF, Lagerbloom P. Uncertainties in Geriatric Data. Ii. Organ Size. *J Am Geriatr Soc*. 1965; 13:20–8. [PubMed: 14256217]
50. Niederau C, Sonnenberg A. Liver size evaluated by ultrasound: ROC curves for hepatitis and alcoholism. *Radiology*. 1984; 153(2):503–5. [PubMed: 6385110]
51. Kornasiewicz O, et al. Choice of transplantation techniques and indications for liver transplantation in polycystic liver disease in patients with no signs of end-stage liver disease. *Transplantation Proceedings*. 2008; 40:1536–1538. [PubMed: 18589146]
52. Kardel T, et al. Ultrasonic determination of liver and spleen volumes. *Scand J Clin Lab Invest*. 1971; 27(2):123–8. [PubMed: 5572884]
53. McNeal GR, et al. Liver volume measurements and three-dimensional display from MR images. *Radiology*. 1988; 169(3):851–4. [PubMed: 3187015]
54. Kwo PY, et al. Gender differences in alcohol metabolism: relationship to liver volume and effect of adjusting for body mass. *Gastroenterology*. 1998; 115(6):1552–7. [PubMed: 9834284]
55. Farraher SW, et al. Liver and spleen volumetry with quantitative MR imaging and dual-space clustering segmentation. *Radiology*. 2005; 237(1):322–8. [PubMed: 16126927]

56. Henderson JM, et al. Measurement of liver and spleen volume by computed tomography. Assessment of reproducibility and changes found following a selective distal splenorenal shunt. *Radiology*. 1981; 141(2):525–7. [PubMed: 6974875]
57. Stapakis J, et al. Liver volume assessment by conventional vs. helical CT. *Abdom Imaging*. 1995; 20(3):209–10. [PubMed: 7620406]

Appendix 1 - STARD Flowchart for Hepatomegaly Study

Appendix 2 - Liver Disorders

	# of Cases
Acute myeloid leukemia (AML)	2
Adult T-cell leukemia/lymphoma (ATL)	1
Alveolar soft part sarcoma (ASPS)	1
Aplastic anemia	2
Appendiceal cancer	1
Autoimmune lymphoproliferative syndrome (ALPS) type 12	1
B-cell chronic lymphocytic leukemia (CLL)	1
Chronic granulomatous disease (CGD)	7
Chronic myelogenous (or myeloid) leukemia (CML)	2
Colon Cancer with liver metastases	1
Cutaneous T cell lymphoma (CTCL)	1
Dermatomyositis lipodystrophy	1
Desmoplastic small round cell tumor	1
D-MAC/myelodysplastic syndromes (MDS)	1
Follicular lymphoma	1
Hepatitis C	1
Hepatosplenic T-cell lymphoma	1
Hereditary papillary renal carcinoma type I (HPRC)	1

	# of Cases
Histoplasmosis (Darling's disease)	1
HIV	1
HIV with Mycobacterium avium complex (MAC)	2
HIV/Hepatic Steatosis	1
HIV/Kaposi's sarcoma (KS)/Multicentric Castleman's disease	1
HIV/Primary Effusion Lymphoma/Multicentric Castleman's disease/Kaposi's sarcoma (KS)	1
HIV with Burkitt's non-Hodgkin lymphoma	1
HIV/Lymphoma	1
Hypertriglyceridemia/Pancreatitis	1
Hypertension	1
Kaposi's sarcoma	2
Large granular lymphocytic leukemia (LGL Leukemia)	2
Lymphoma	3
Mantle cell lymphoma (MCL)	1
Mastocytosis	2
Melanoma	4
Mesothelioma	1
Multicentric Castleman's disease	3
Mycobacterium avium complex (MAC)	1
Mycobacterium avium intracellulare (MAI)	1
No Diagnosis	3
Nontuberculous mycobacteria (NTM)	1
Ocular melanoma with liver metastases	3
Peritoneal mesothelioma	1
Prostate carcinoma	2
Pure red cell aplasia (PRCA)	1
Systemic lupus erythematosus (SLE)	1
Thalassemia	1
TOTAL	71

Figure 1.

Bland-Altman agreement plots for liver height measurements at the mid-hepatic line between a) two independent observers (n=148), b) automated method (CAD) and observer 1 (n=148), c) CAD and observer 2 (n=148), d) intra-observer supine and prone measurements (n=25, note scale change on axes). The mean error is shown in solid line and the 95% limits of agreement (± 1.96 SD) in dashed lines.

Figure 2.

H-score (cutoff to detect hepatomegaly) variation with the number of consecutive normal samples used for its computation. The peak near the beginning of the graph is caused by an outlier with large H-score.

Figure 3. Correlations between the H-score and liver MHL heights using a linear regression model: a) normal cases, and b) hepatomegaly cases (as defined by the H-score). The vertical red lines show the cutoff for hepatomegaly as defined in literature by an MHL height of 15.5cm [7, 38].

Figure 4.

The image data of a case with Riedel's lobe with abnormal liver MHL height of 17.6 cm and normal H-score of 0.88 l/m². The automatically segmented liver is overlaid in blue over the CT image and shown in axial slices.

TABLE 1**AVERAGE AUTOMATED LIVER VOLUMES AND HEIGHTS**

Cases	Mean/std
Normal Volume: n = 74	1.52±0.26 l
Normal Volume/BSA: n = 74	0.81±0.11 l/m ²
Normal MHL Height: n = 74	12.86±2.21 cm
Enlarged Volume: n = 43	2.57±0.76 l
Enlarged Volume/BSA: n = 43	1.45±0.40 l/m ²
Enlarged MHL Height: n = 43	19.32±3.64 cm
Hepatectomy Volume: n = 23	1.56±0.40 l

NOTE: MHL – mid-hepatic line; BSA- body surface area. Average volumes, volumes normalized by BSA and heights at MHL for normal and enlarged livers. Only average volumes are reported for partial hepatectomy cases due to the frequent inability to calculate the MHL height and extraneous relationship to body size after surgery.

TABLE 2**CORRELATIONS BETWEEN LIVER VOLUME/HEIGHT AND PATIENT'S BSA/AGE**

Cases/Correlation factor	Correlation (p value)
Normal Volume: n = 74/Patient BSA	0.60 (<0.001)
Normal Volume: n = 74/Patient Age	0.01 (0.9)
Enlarged Volume: n = 43/Patient BSA	0.41 (0.005)
Enlarged Volume: n = 43/Patient Age	0.38 (0.01)
Normal MHL Height: n = 74/Patient BSA	-0.02 (0.8)
Normal MHL Height: n = 74/Patient Age	-0.33 (0.003)
Enlarged MHL Height: n = 43/Patient BSA	0.08 (0.5)
Enlarged MHL Height: n = 43/Patient Age	0.18 (0.2)

NOTE: MHL – mid-hepatic line; BSA- body surface area. Correlation coefficients (and associated p values) are presented between volume/MHL height measurements and patient's BSA and age for normal and enlarged livers.

TABLE 3**RETROSPECTIVE SENSITIVITY AND SPECIFICITY OF RADIOLOGISTS AND LIVER MHL HEIGHT TO DETECT HEPATOMEGALY**

Criterion (n = 148)	Sensitivity/Specificity (%)		
	Mild Hepatomegaly	Massive Hepatomegaly	All Hepatomegaly
Radiologists	56.7/90.1	<i>100.0/90.1</i>	<i>84.4/90.1</i>
MHL Height	40.0/88.7	85.1/88.7	67.5/88.7

NOTE: MHL – mid-hepatic line. The H-score was used as the reference standard with cutoffs of 0.92/1.08 l/m² to identify mild/massive. The radiological reports were used to compute the sensitivity and specificity of the radiologists. A cutoff of 15.5 cm was used to identify hepatomegaly using the liver MHL height. Significant differences (p<0.007, shown in italics) were noted between the sensitivity of the radiologist versus the sensitivity of the liver MHL height criterion for the detection of massive and all cases of hepatomegaly.

TABLE 4**CORRELATIONS BETWEEN LIVER VOLUME AND PATIENT DATA**

Study	BSA	Weight	Height	Age
This study	Yes	Yes	Yes	Yes
Andersen et al. [1]	-	Yes	No	No
Boyd et al. [48]	-	-	-	Yes
Calloway et al. [49]	-	-	-	Yes
Hauksen et al. [39]	Yes	Yes	Yes	-
Kardel et al. [52]	-	Yes	-	-
Kratzer et al. [47]	-	Yes	Yes	Yes
Leung et al. [2]	Yes	Yes	No	-
McNeal et al. [53]	-	Yes	-	-
Raeth et al. [28]	Yes	Yes	Yes	-
Urata et al. [40]	Yes	-	-	-
Wynne et al. [3]	Yes	-	-	Yes
Zoli et al. [4]	Yes	Yes	-	-

NOTE: BSA – body surface area.

TABLE 5
REVIEW OF AVERAGE VOLUMES FOR NORMAL AND DISEASED LIVERS

Study	Patients Livers	Sex	No. Cases	Mean Volume	Modality
Andersen et al. [1]	Normal	Male	7	1.60 l	CT
Kwo et al. [54]	Normal	Male	10	1.54 ± 0.08 l	CT
Sandrasegaran et al. [41]	Normal	Male	8	1.67 l	CT
Andersen et al. [1]	Normal	Female	16	1.34 l	CT
Kwo et al. [54]	Normal	Female	10	1.48 ± 0.06 l	CT
Sandrasegaran et al. [41]	Normal	Female	11	1.52 l	CT
Farragher et al. [55]	Normal	Male/Female	18	1.66 ± 0.35 l	MRI
Henderson et al. [56]	Normal	Male/Female	11	1.45 ± 0.17 l	CT
Kardel et al. [52]	Normal	Male/Female	20	1.61 ± 0.19 l	US
Mazonakis et al. [30]	Normal	Male/Female	27	1.47 ± 0.23 l	MRI
Siapakis et al. [57]	Normal	Male/Female	22	1.32 ± 0.42 l	CT
Farragher et al. [55]	Diseased	Male/Female	9	1.99 ± 0.52 l	MRI
Henderson et al. [56]	Diseased	Male/Female	12	1.64 ± 0.7 l	CT
Mazonakis et al. [30]	Diseased	Male/Female	11	1.94 ± 0.19 l	MRI
Van Thiel et al. [13]	Diseased	Male/Female	99	1.78 ± 0.09 l	CT

NOTE: Table 1 reports the average volumes of normal and diseased livers in our study. CT – computed tomography; MRI – magnetic resonance imaging; US – ultrasound.