Assessing Hepatomegaly:

Automated Volumetric Analysis of the Liver

Marius George Linguraru, DPhil, Jesse K. Sandberg, BS, Elizabeth C. Jones, MD, Nicholas Petrick, PhD, Ronald M. Summers, MD, PhD

Rationale and Objectives: The aims of this study were to define volumetric nomograms for identifying hepatomegaly and to retrospectively evaluate the performance of radiologists in assessing hepatomegaly.

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

Results: The automated segmentation of livers was robust, with volume overlap and error of 96.2% and 2.2%, respectively. There were no significant differences ($P > .10$) between manual and automated segmentation for either the normal or the hepatomegaly subgroup. The average volumes of normal and enlarged livers were 1.51 ± 0.25 and 2.32 ± 0.75 L, respectively. One-way analysis of variance found that body surface area ($P = .004$) and gender ($P = .02$), but not age, significantly affected normal liver volume. No significant effects were observed for two-way and three-way interactions among the three variables ($P > .18$). H-score cutoffs of 0.92 and 1.08 L/m^2 were used to define mild and massive hepatomegaly (95% confidence interval, $\pm$0.02 L/m^2). Using the H score as the reference standard, the sensitivity of radiologists in detecting all, mild, and massive hepatomegaly was 84.4%, 56.7%, and 100.0% at 90.1% specificity, respectively. Radiologists disagreed on 20.9% of the diagnosed cases ($n = 31$). The area under the receiver-operating characteristic curve of the H-score criterion for hepatomegaly detection was 0.98.

Conclusions: Nomograms for the identification and grading of hepatomegaly from automatic volumetric liver assessment normalized to body surface area (H scores) are introduced. H scores match well with clinical interpretations for hepatomegaly and may improve hepatomegaly detection compared with height measurements or visual inspection, commonly used in current clinical practice.

Key Words: Hepatomegaly; volumetric analysis; liver; segmentation; nomogram.

Hepatomegaly is an abnormal enlargement of liver size and is inherently defined by a volumetric change. Patient 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 cases of massive hepatomegaly (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 normal livers were diagnosed as enlarged.

Radiologic imaging modalities such as magnetic resonance imaging, ultrasound, and computed tomographic (CT) imaging are better able to assess liver size (25–29). To detect hepatomegaly in radiologic scans, radiologists commonly focus on landmark-based visual evaluations of the hemidiaphragm, displacement of the stomach, the duodenum, hepatic flexure of the colon, the 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 mild and massive hepatomegaly (with the latter category of patients

Acad Radiol 2012; 19:588–598

From the Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, Bethesda, Maryland (M.G.L., J.K.S., E.C.J., R.M.S.); The Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Medical Center, 111 Michigan Avenue, NW, Washington, DC 20010 (M.G.L.); and the Center for Devices and Radiological Health, US Food and Drug Administration, Silver Spring, Maryland (N.P.). Received September 26, 2011; accepted January 28, 2012. This work was supported in part by the Intramural Research Programs of the National Institutes of Health, Clinical Center, and the US Food and Drug Administration. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either actual or implied endorsements of such products by the US Department of Health and Human Services. **Address correspondence to:** M.G.L. e-mail: mlingura@cnmc.org

doi:10.1016/j.acra.2012.01.015

possibly benefiting from medical intervention) 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 midhepatic line (MHL) from radiologic scans (7,38). The MHL is the half-point distance between the midpoint of the spine and the outermost point on the liver surface (right side) in axial planes. This two-dimensional measure does not fully characterize the liver's morphology, and livers with Riedel's lobe may exhibit large MHL heights. Although correlations between liver size and patient size have been 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 because of the high and expensive human interaction, and if the same organ is segmented twice (by the same or different observers), the result will likely differ. Currently, most radiologists do not rely on liver volume or even height measurements for diagnosing hepatomegaly, because of 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 report, nomograms are defined to identify and grade hepatomegaly from automated liver volumetry on CT imaging. The automated liver volumetric software was previously developed by our group (42) and allows the quantification of healthy and diseased livers. From readily available volumes of normal livers normalized to patient size (body surface area [BSA]), 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

Data were acquired at the National Institutes of Health using contrast-enhanced CT scans at the portal venous phase without imaging or motion artifacts or large pathologies in the liver. This retrospective study was in compliance with the Health Insurance Portability and Accountability Act and was institutional review board approved, and the requirement for informed consent was waived.

From January 2001 to March 2009, 71 consecutive subjects with clinically diagnosed hepatomegaly (48 men [mean age, 43 years; age range, 18–76], 23 women [mean age, 41 years; age range, 19–66 years]) met the inclusion criteria shown in Appendix A. The criteria included adult patients with CT scans acquired with intravenous Isovue contrast agent (Bracco Diagnostics, Milan, Italy) in fully enhanced portal venous phase with available patient data. Additionally, the data should not have imaging or motion artifacts or 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 the radiology information system (Cerner Corporation, Kansas City, MO) and the clinical research information system (Eclipsys Corporation, Atlanta, GA). Radiologic reports included one of the following key terms: "hepatomegaly," "hepatosplenomegaly," "enlarged liver," or "liver enlarged" (see Appendix A for the study flowchart). Appendix B presents the clinical diagnoses of the hepatomegaly data.

The control population of normal livers was formed from consecutive kidney donors enrolled at the National Institutes of Health from January 2001 to August 2010. Liver function parameters were not available. Seventy-seven subjects with healthy livers (33 men [mean age, 43 years; age range, 17–76 years], 44 women [mean age, 44 years; age range, 18–72 years]) were selected.

The 148 cases (71 patients with hepatomegaly 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 National Institutes of Health (15 men [mean age, 49 years; age range, 33–70 years], eight women [mean age, 54 years; age range, 28–71 years]). 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 in portal venous phase during a single breath using fixed delays (65–70 seconds depending on the scanner) or bolus tracking (43) after patients were administered 130 mL of Isovue-300. Data were collected on LightSpeed Ultra/QX/I (GE Healthcare, Milwaukee, WI), Brilliance 64 (Philips Medical Systems, Andover, MA), Definition (Siemens Healthcare, Erlangen, Germany), and Aquilion ONE (Toshiba Medical Systems, Tokyo, Japan) scanners at 100 to 240 mAs and 120 kVp. Image resolutions ranged between 0.52 and 0.93 mm in the axial view with slice thicknesses of 1 to 5 mm. Livers were manually segmented from 20 cases by two observers supervised by an experienced radiologist. The liver heights were manually measured at the MHL in all data, excepting the hepatectomy cases, by the two observers.

Last, 25 random pairs of supine and prone noncontrast CT data sets (slice thickness, 1 mm) from Walter Reed Army Medical Center in Washington, District of Columbia (14 men [mean age, 58 years; age range, 51–73 years], 11 women [mean age, 55 years; age range, 49–68 years]; data courtesy of J. Richard Choi, ScD, MD) were analyzed for intraobserver