Published in final edited form as:

Acad Radiol. 2013 June ; 20(6): 675–684. doi:10.1016/j.acra.2013.01.011.

Assessing Splenomegaly: Automated Volumetric Analysis of the Spleen

Marius George Linguraru, DPhil^{1,2}, Jesse K. Sandberg², Elizabeth C. Jones, MD³, and Ronald M. Summers, MD PhD²

¹Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Medical Center, 111 Michigan Avenue NW, Washington DC 20010

²Imaging Biomarkers and Computer-aided Detection Laboratory, Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Building 10 Room 1C224D MSC 1182, Bethesda, MD 20892

³Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Building 10 Room 1C224D MSC 1182, Bethesda, MD 20892

Abstract

Rationale and Objectives—To define systematic volumetric thresholds to identify and grade splenomegaly, and retrospectively evaluate the performance of radiologists to assess splenomegaly in computed tomography image data.

Materials and Methods—A clinical tool was developed to segment spleens from 172 contrast-enhanced clinical CT studies. There were 45 normal and 127 splenomegaly cases confirmed by radiological reports. Spleen volumes were compared to manual measurements using overlap/error. Volumetric thresholds for mild/massive splenomegaly were defined at 1/2.5 standard deviations above the average splenic volume of the healthy population. The thresholds were validated against consensus reports. The performance of radiologists in assessing splenomegaly was retrospectively evaluated.

Results—The automated segmentation of spleens was robust with volume overlap/error of 95.2/3.3%. There were no significant differences ($p > 0.2$) between manual and automated segmentations for either normal/splenomegaly subgroups. Comparable correlations between interobserver and manual-automated measurements were found ($R = 0.99$ for all). The average volume of normal spleens was 236.89 ± 77.58 ml. For splenomegaly, average volume was 1004.75 ± 644.27 ml. Volumetric thresholds of 314.47/430.84 ml were used to define mild/massive splenomegaly (± 18.86 ml 95% CI). Radiologists disagreed in 23.25% ($n = 40$) of the diagnosed cases. The area under the ROC curve of the volumetric criterion for splenomegaly detection was 0.96. Using the volumetric thresholds as the reference standard, the sensitivity of radiologists in detecting all/mild/massive splenomegaly was 95.0/66.6/99.0% at 78.0% specificity, respectively.

Conclusion—Thresholds for the identification and grading of splenomegaly from automatic volumetric spleen assessment were introduced. The volumetric thresholds match well with clinical interpretations for splenomegaly and may improve splenomegaly detection compared with splenic

Correspondence to: Marius George Linguraru.

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.

cephalocaudal height measurements or visual inspection commonly used in current clinical practice.

Keywords

spleen; splenomegaly; volume analysis; volumetric threshold; segmentation

Introduction

The enlargement of the spleen, or splenomegaly, is the most common condition associated with diseases of the spleen [1]. Splenomegaly is a nonspecific finding which is diagnostically challenging to assess because the manifestation is nearly always secondary to another primary disorder and diseases associated with the spleen are numerous [1–8]. Pozo et al. grouped them into six categories: infection, hematological, congestive, inflammatory, neoplastic and infiltrative miscellaneous diseases [7]. Hematological disorders were found to be the most common (up to 66%) with lymphoma being the most prevalent disease within the category.

The assessment of spleen size, defined by its volume, is of importance in the diagnosis of and determination of the severity of many of these diseases and in the selection and monitoring of therapies. Moreover, given the spleen's irregular shape, volume is the best summary indicator of spleen changes over time. It is common clinical practice for physicians to estimate splenic size and assess for an enlarged spleen using palpation [9]. In palpation, the spleen is considered normal in size when it is not palpable below the left costal edge [11]; if the spleen is palpable, it is enlarged. But 16% of palpable spleens were found to be of normal size on radiological assessments [10].

The advent of cross-sectional imaging, such as ultrasound and computed tomography (CT), enabled the non-invasive visualization of the spleen size and shape. Currently, the main method for diagnosing splenomegaly is through cross-sectional radiological assessment [7]. A popular method routinely used by clinicians to estimate splenic size is manually measuring the organ's cephalocaudal (CC) height from radiological images [1, 5, 7, 19–25]; this measure is only two-dimensional (2D) and does not factor in interpatient anatomical variability [1]. Even though Rosenberg et al. [26] found that splenic height assessed spleen size with good accuracy, most radiological studies have emphasized the importance of volumetric spleen measurements in assessing spleen disorders and size [19, 21, 23, 27–31]. Thus, some studies have adopted a method for measuring splenic volume by calculating the “splenic index”, which is the product of the length, depth and width at certain anatomical locations [1, 21, 29, 32, 33]. As a surrogate of splenic size, the splenic index accounts better for anatomical variability than a single linear measurement.

Presently, most radiologists do not rely on volume measurements of the spleen for assessing splenomegaly due to the lack of robust and accessible methods that could be easily adapted to current image viewing software (such as PACS). Consequently, the current radiological assessments of the spleen consist of subjective assignments into qualitative size categories (e.g. small, normal, mild, marked splenomegaly) and or measurement of linear assessment (single longest dimension) for quantification of spleen size.[1, 33, 35]. The development of computer-based image processing techniques now allows for rapid and accurate assessment of spleen volumes on radiologic data [36] and their inclusion in the radiologist's interpretation and report. While several papers have investigated spleen size nomograms using the volume [27,28,29,30,43,44,50], our study proposes and evaluates the systematic definition of a volumetric threshold to detect and grade splenomegaly.

This paper assesses the accuracy and advantages of splenic volume over subjective qualitative and simpler quantitative methods in the assessment of splenomegaly. First, we use a fully-automated computer-aided segmentation method [34] to measure the volume and CC height of spleens from 172 contrast-enhanced CT images of normal and splenomegaly populations. The method is evaluated against manual segmentation of splenic volumes (sum of discs method). Our technique is designed to help radiologists differentiate between normal and enlarged spleens through quantitative, robust and repeatable measurements. Then, we define volumetric thresholds for assessing splenomegaly and retrospectively evaluate the performance of radiologists to assess splenomegaly.

Materials and Methods

Study Patients

This retrospective study follows the HIPAA-compliance standard and was IRB approved. The need for informed consent was waived by the IRB. Inclusion criteria were that CT scans had to be acquired using intravenous contrast enhancement at the portal venous phase without visible artifacts or focal masses in the spleen. Cases with motion or imaging artifacts were excluded from the study.

From January 2005 to March 2009, 127 consecutive subjects with splenomegaly diagnosed radiologically on CT (97 males [mean age 47; range 18–76], 30 females [mean age 50; range 19–74]) met the inclusion criteria for our study. The patient cases were collected based upon search parameters defined within the Radiology Information Systems (RIS - Cerner Millennium, Cerner Corporation). One of the following keywords had to be included in the radiological report of the CT scan: splenomegaly, hepatosplenomegaly, enlarged spleen, or spleen enlarged. Diagnoses in the clinical reports were established by one of 14 radiologists without following any single criterion and including visual inspection and/or measurements of the spleen height. All cases were reviewed to avoid the inclusion of cases with misrepresentative keywords in their radiological reports. See Appendix I for a review of clinical diagnoses in the splenomegaly data set. Appendix II presents the STARD chart of the selection process of splenomegaly image data.

The control population of normal spleens was selected from kidney donors enrolled from January 2001 to August 2010. Control cases represent a random sample of the general normal population who did not exhibit pathologies in the spleen. In all, 45 subjects with normal spleens (18 males [mean age 44, range 17–76] and 27 females [mean age 45, range 18–72]) were selected.

The 172 cases (127 splenomegaly patients and 45 controls) were reevaluated by two experienced radiologists (XX and YY, with 17 and 20 years of post-fellowship experience reading CT scans). 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.

Contrast-enhanced CT images were acquired during the portal venous phase of enhancement during a single breath using fixed delays (65–70 s depending on the scanner) or bolus-tracking [37] after patients were administered 130 ml Isovue-300. Data were collected on LightSpeed Ultra and 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 among the patient images ranged from 0.52 to 0.93 mm in the axial view with a slice thickness from 1 to 5 mm. Spleens were manually segmented from 20 random cases (10 normal and 10 splenomegaly) by two observers (post-

graduate research fellows: XX and YY) under the supervision of a board-certified radiologist and an image processing scientist and their volumes were recorded using the sum of discs method. A subsample of data was used to obtain volumetric segmentations of the spleen, due to the time-consuming process of manually segmenting volumetric CT data. The spleen CC heights were manually measured in all data by the two observers for additional evaluation.

Segmentation

The automated segmentation method measured the spleen volumes on all subjects and is outlined in detail in [34]. The method involves a combination of appearance, shape and location statistics to segment the spleen. For the coarse estimation, mean models from an atlas of the spleen were aligned to the patient's contrast-enhanced CT images via rigid, affine and non-rigid registration. The registration was based on normalized mutual information and B-splines. The estimation was improved by a geodesic active contour, a three-dimensional deformable front that adapts to the appearance and shape of the spleen. The contour followed by an adaptive convolution to take into account patient specific contrast-enhancement characteristics. The contrast enhancement was estimated from the geodesic active contour. Using the convolution, only homogenous tissue areas that satisfy the enhancement constraints are labeled as spleen. Lastly, shape and location information from the normalized probabilistic atlas are utilized to provide an accurate representation of each spleen's morphology.

Definition of Splenomegaly and Volumetric Thresholds

Spleen volumes from cases with splenomegaly are outliers from the average spleen size in the healthy population. Although the study of the distribution of normal and abnormal splenic volume has been studied in literature, to date, there are no established volumetric thresholds for the assessment of splenomegaly. Nevertheless, standards have been defined for the detection, for example, of osteoporosis [47,48] and hepatomegaly [37]. 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 [47,48]. Similarly, the H-score for the detection of hepatomegaly was defined at a liver volume (normalized by body surface area - BSA) above one SD from the average healthy population with massive hepatomegaly above 2.5 SD from the average [37].

We define the volumetric thresholds for splenomegaly detection from the splenic volume. The average spleen volume and its SD are computed from our healthy population in the consensus reports. Only cases that were found normal by all radiologists were used in the computation of the volumetric thresholds. Following the approach used to determining osteoporosis and hepatomegaly, the threshold for mild splenomegaly is defined at one SD above the average. The massive splenomegaly threshold is defined at 2.5 SD above the average.

Performance of Radiologists Relative to the Splenomegaly Thresholds

The performance of radiologists to diagnose splenomegaly was retrospectively evaluated using the volumetric thresholds for splenomegaly to determine the ground truth for splenomegaly. For comparison with previous criteria to assess splenomegaly based on spleen height, the diagnostic performance using the spleen CC height with 11 and 12 cm cutoffs was compared to the reference standard from volumetric thresholds.

Statistical Analysis

Prospective power analysis was performed using a one-sample test with 0.05 significance level to determine the control sample size to detect a 10% effect size on the normal splenic volume. The power analysis was based on historical data reported in [29, 30]. Additionally, 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=132 cases after consensus) to detect a significant difference between normal and splenomegaly cases.

Manual and automated volumetric segmentations were compared by volume overlap (twice the volume of intersection between the manual and automatic volumes over the union of the two volumes) and volume error (absolute volume difference between the manual and automatic volumes relative to the manual volume). Intra and interobserver variability and error analysis for measuring spleen height were performed following the Bland-Altman method [38]. Bland-Altman plots allow to investigate the existence of any systematic difference between the automatic and manual measurements and compute the estimated bias (mean difference). The Mann-Whitney U test assessed significance between inter/intraobserver, and observer-CAD agreements. Spearman nonparametric correlation coefficients and associated p-values (95% confidence level) were calculated between spleen size and patients' BSA and age for comparisons with the same metrics reported in literature. Also, 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 confounding factors on normal spleen volumes.

Fisher's exact test assessed the significance between the sensitivity of radiologists versus the spleen CC height criteria to detect splenomegaly with the volumetric thresholds as the reference standard. The Spearman correlation between the splenic volume and CC height was also analyzed. For the analysis of performance to assess splenomegaly, receiver operating characteristic (ROC) curves were calculated to find the sensitivity and specificity of the volumetric criterion and the area under the curve (AUC) was recorded; the consensus reports were used as the reference standard for splenomegaly. Statistical analysis was performed using the STATA Data Analysis and Statistical Software (StataCorp LP).

Results

Automatic Segmentation

The evaluation of the automated spleen segmentation tool [34] showed an average volume overlap of 95.2% and volume error of 3.3% between automatically and manually segmented spleens. The volume overlap was 95.1% and 95.3%, and volume error was 3.7% and 2.8% for the control and splenomegaly cases, respectively. There was no statistical significant difference ($p>0.2$) for either volume overlap or error between automated and manual segmentations on normal or abnormal cases. The automatic volumetric assessment of the spleen is 100% reproducible, as it does not require any interaction with the human operator.

The Bland-Altman height measurement agreement plots between two different observers and between each observer and the automatic measurement are shown in Figures 1a–c. Note that values are quantized due to 5 mm slice resolution used for the majority of cases. The interobserver variability was 0.02 ± 0.74 cm and the average bias between the CAD method and each observer was 0.08 ± 1.41 cm at 95% limits of agreement. Significant correlations ($R=0.99$, $p<0.001$) were found between each observer and the automatic measurement, comparable to interobserver measurements correlation ($R=0.99$, $p<0.001$). Outliers in Figure 1 corresponded to unusually shaped spleens (thin and tall), which increased the variability of CC height measurements.

Consensus Report

The consensus reports agreed on $n=132$ cases: 44 normal and 88 splenomegaly cases. The radiologists disagreed on $n=40$ (23.25%) of cases; $n=1$ (2.22%) were normal and $n=39$ (30.70%) were abnormal cases, according to the clinical reports. The volumes and CC heights for the automatically segmented normal and enlarged spleens are displayed in Table 1. Significant differences were found between normal and enlarged spleen volumes ($p<0.001$) and between normal and enlarged spleen heights ($p<0.001$). Organ volumes normalized by BSA are also presented in Table 1.

Relations between Splenic Volume and Patient Size

Table 2 shows Spearman correlation coefficients (R) between spleen volumes and patient's BSA and age. Significant but moderate correlations were noted between spleen volumes and BSA in normal ($R=0.28$, $p=0.02$) and splenomegaly ($R=0.32$, $p=0.002$) cases, and between the spleen CC heights and BSA in normal ($R=0.35$, $p=0.004$) and splenomegaly ($R=0.27$, $p=0.008$) cases. No significant correlations were observed between the spleen volumes and age of patients. No significant effects on the normal splenic volumes were observed for three-way interactions among the BSA, age and gender of patients ($p>0.2$).

Volumetric Thresholds for Splenomegaly

A minimum of 34 control cases are required to detect an effect size of 10% on the normal splenic volume with 95% power. Our control data sample consists of 44 cases, after consensus. The volumetric threshold of 314.47 ml was used to identify mild splenomegaly (± 18.86 ml 95% CI). The volumetric threshold for massive splenomegaly was 430.84 ml. The statistical power of our dataset ($n=132$) for the observed volumetric threshold for splenomegaly was 100%. From ROC analysis, the operating point of the volumetric thresholds was found with a sensitivity of 93.18% and specificity of 88.63%. The area under the ROC curve of the volumetric criterion for splenomegaly detection was 0.96, when the radiological consensus was used as reference standard. The area under the ROC curve of the criterion based on spleen volume normalized by BSA was also 0.96. The lack of significant difference between the two criteria ($p=0.9$) and the reduced correlation between spleen volumes and BSA ($R=0.28$) indicate that the BSA normalization does not improve the detection of splenomegaly.

Performance of Radiologists to Detect Splenomegaly

Using the volumetric thresholds for splenomegaly as the new reference standard, the performance of radiologists to detect splenomegaly was retrospectively analyzed using the clinical reports. Table 3 presents the sensitivity and specificity of radiologists to detect splenomegaly. The performance of diagnosis by spleen CC height is also shown using two different thresholds based on previous studies (see Table 4): CC11 and CC12 to detect splenomegaly at a CC height larger than 11 and 12 cm, respectively. Radiologist detected splenomegaly with both higher sensitivity and specificity than CC11 and CC12 criteria. However, the performance in diagnosis was significant only between the radiologists and the CC12 criterion for mild ($p=0.05$) and all ($p=0.03$) splenomegaly cases. There were no significant differences between the performances of the CC11 and CC12 criteria ($p>0.07$ for all).

Relation between Splenic Volume and Height

There was a significant correlation between the volumes and CC heights of cases ($n=172$, $R=0.85$, $p<0.001$), as shown in Figure 2. To show the agreement between the quantitative scores and the radiology reports, an 11 cm cutoff at CC height [7, 38] detected splenomegaly with 96.85% sensitivity at 84.44 % specificity against the radiological reports. A similar 12

cm threshold at CC height [5] detected splenomegaly with 92.91% sensitivity at 95.55 % specificity using the radiological reports as reference standard. In comparison, using a volumetric threshold of 314.47ml for splenomegaly showed a sensitivity of 91.33 % and specificity of 86.66% against the radiological reports.

Figure 3 shows examples of spleens with normal CC height and abnormal volumes, which affected the sensitivity of the CC height criteria to detect mild splenomegaly. For instance, in the example in Figure 3.a, the CC height of the spleen was close to the average height of the normal spleens; however, the splenic volume was two standard deviations above the average normal volume.

Discussion

We presented a volumetric method to assess splenomegaly. The method was presented in conjunction with an automated technique to segment the spleen from CT images. Please note that the software for spleen analysis in this study is not commercially available and was used as a research tool. The technique was accurate and able to successfully segment both normal and abnormal spleens [34]. Bland-Altman agreements showed that the automatic measurement errors are comparable to interobserver variability. Hammon et al. [51] also reported 0.99 correlations between CAD and manual segmentations of the spleen in a study based on 15 lymphoma patients. In our application, segmentation errors were observed when enlarged spleens extended into the small intestines because both splenic and bowel tissue are of similar intensity in CT images. In addition, intensity variations in the form of partial volume effects induced errors at the superior and inferior ends of the spleen.

The average volumes of the normal spleens in our study (Table 1) were comparable to those found in other publications (Table 4). As seen in Table 4, the variability of the enlarged spleen volumes tended to be much larger between publications. Our study found high and significant correlations between the splenic volume and CC height of our populations; Lamb et al. [35] also observed this relationship ($R=0.86$, $p<0.001$). In addition, Prassopoulos et al. [29] observed that splenic volume correlated well with all the linear and the maximal cross-sectional area measurements of the spleen.

It is of general agreement that there is a negative correlation between the splenic size, expressed as volume or height, and a patient's age [6, 7, 27, 28, 31, 39, 40, 41]. We observed such negative correlation only for the normal splenic height, though not significant. In particular, there was no correlation between the splenic volume and age of the control cases. This observation was also confirmed by the study in [29], which included data from 140 normal patients, one of the largest volumetric studies of the spleen in literature. What is not of general agreement is a correlation between splenic size and a patient's BSA. In particular, Prassopoulos et al. [29] reports that there is no relation between splenic volume and a patient's height, weight or BMI. In addition, a study focused on enlarged spleens [42] found no trend between the splenic size and a patient's BSA or weight. As seen in Table 2, our study found moderate but significant correlations between patient's BSA and both splenic volume and height for both normal and splenomegaly cases, which made us investigate the role of BSA in the assessment of splenomegaly. However, we did not detect a significant effect in the detection of splenomegaly by using BSA normalization.

As shown in Table 4, several papers have discussed the definition of spleen size nomograms using the splenic CC height, and occasionally its volume. In terms of an upper limit of normal height; thresholds ranging from 9.76–14 cm have been suggested [1, 5, 7, 19–22]. Most of the CC height thresholds are in the 11–12 cm range, the values we evaluated in our study. In particular, Peddu et al. [6] presented a review of pathologies and defined the

normal spleen height to up to 11 cm (95% CI). Publications have also shown the upper limit for normal spleen volume to range between 250–335 ml [19, 21, 28, 29]. Our study proposed a volumetric threshold of 314.47 ml for mild splenomegaly and 430.84 ml for massive splenomegaly. Using volumetric measurements to detect splenomegaly, an area under the ROC curve of 0.96 was obtained against the consensus report of radiologists. The sensitivity and specificity of the volumetric thresholds against the radiological consensus were 93.18% and 88.63%, respectively.

An arguable drawback in our study is the relatively small sample of normal population (n=44). Please note that the only three image-based volumetric studies in Table 4 that included larger data samples are those by Kaneko et al. [27], Prassopoulos et al. [29], and Srisajjakul et al. [52]. The studies in [27, 52] included Japanese and Thai subjects only and their conclusions cannot be directly related to our study. Moreover, the ranges of spleen volumes and CC heights in [52] are unusually wide. More importantly, Prassopoulos et al. [29] concluded that the upper volumetric bound of the normal spleen is 314.50 ml in the most comprehensive prior volumetric study of the spleen, which included 140 CT scans of normal spleens (both male and female). Note that the systematic computation of a volumetric threshold for splenomegaly in our study also found that a splenic volume of 314.47 ml is the correct upper bound of the normal spleen. This agreement can be interpreted as a confirmation that our normal data sample is sufficiently large and representative of the general normal population. The advantages that our method brings over the study in [29] are the automation and repeatability of the volumetric measurements, and the systematic computation of the volumetric threshold instead of choosing the maximum splenic volume in the healthy population.

The agreement between the quantitative scores and the radiology reports was also evaluated. A CC height threshold of 11 cm resulted in 96.85% sensitivity at 84.44 % specificity and a CC height threshold of 12 cm obtained 92.91% sensitivity at 95.55 % specificity. For comparison, Bezerra et al. [21] reported a sensitivity of 90.6% and specificity of 90.3% to detect splenomegaly using spleen CC heights. These results suggest that the radiologists in our study are in closer agreement with height measurements than volumetric assessments of the spleen. This is not surprising, given that radiologists could measure the CC height in their assessment, but did not have access to volumes. Moreover, the splenic height is simple to observe on CT images, while the volume, although the inherent definition of size requires a more complex 3D assessment. Please note that our study recognized a large disagreement between individual radiological reports (23.25% of cases).

Ultimately, our study introduced volumetric thresholds to detect splenomegaly, as spleen size is inherently a volume. The evaluations of radiological report and CC height criteria to detect splenomegaly were presented in Table 3. All three criteria showed moderate specificity in detecting splenomegaly with variable sensitivity. In particular height criteria showed very low sensitivity to detect mild splenomegaly (33.3%). Radiologists detected abnormal spleens significantly better than CC height thresholds and with high sensitivity, whether assessing mild or massive cases of splenomegaly. The superiority of sensitivity of the radiologists' interpretation of splenomegaly over determination of splenomegaly based on CC is likely due to the radiologists' appreciation of increase in the spleen's cross-sectional area in addition to CC dimension. Using the volumetric thresholds, systematic quantitative thresholds of splenic size, as reference standard, radiologists detected all, mild and massive cases of splenomegaly with 95.0%, 66.6% and 99.0% sensitivity at 78.0% specificity, respectively.

The volumetric thresholds may be useful as systematic indications of splenomegaly, as CAD can offer robust and reproducible spleen volume measurements to support routine radiologic

image analysis. If CAD were to be adopted in the clinical work flow, the diagnosis of splenomegaly could become a seamless automated process. Also, differentiating between cases of mild and massive splenomegaly is unreliable without systematic measures of spleen size. But importantly, the introduction of spleen volume thresholds, with or without CAD, promises to reduce variability and error in the radiologic interpretation of abdominal data.

The study has certain limitations. The distribution of the patient population with splenomegaly may not be representative; however, patients with splenomegaly were not admitted to our hospital on the basis of splenomegaly diagnosis, which was a secondary radiologic finding. Additionally, the reader variability in the clinical determination of splenomegaly is potentially large, but representative for the clinical environment of a medium sized to large hospital. Finally, gender and race differences [27] were not accounted in our study.

In conclusion, it has been shown that for accurate detection of splenomegaly, radiologists and other health professionals should be observant of the volume of the spleen when evaluating for splenomegaly. Thresholds for the identification and grading of splenomegaly from automatic volumetric spleen assessment were introduced. The volumetric thresholds matched well with clinical interpretations for splenomegaly and may improve splenomegaly detection compared with height measurements or visual inspection, commonly used in current clinical practice.

Acknowledgments

This work was supported in part by the Intramural Research Programs of the National Institutes of Health, Clinical Center. The authors thank John Pura and Jianfei Liu, PhD 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. Robertson F, Leander P, Ekberg O. Radiology of the spleen. *Eur Radiol.* 2001; 11(1):80–95. [PubMed: 11194923]
2. O'Reilly RA. Splenomegaly in 2,505 patients at a large university medical center from 1913 to 1995. 1963 to 1995: 449 patients. *West J Med.* 1998; 169(2):88–97. [PubMed: 9735689]
3. Elsayes KM, et al. MR imaging of the spleen: spectrum of abnormalities. *Radiographics.* 2005; 25(4):967–982. [PubMed: 16009818]
4. Hermann RE, De Haven KE, Hawk WA. Splenectomy for the diagnosis fo splenomegaly. *Ann Surg.* 1968; 168(5):896–900. [PubMed: 5684193]
5. O'Reilly RA. Splenomegaly at a United States County Hospital: diagnostic evaluation of 170 patients. *Am J Med Sci.* 1996; 312(4):160–165. [PubMed: 8853064]
6. Peddu P, Shah M, Sidhu PS. Splenic abnormalities: a comparative review of ultrasound, microbubble-enhanced ultrasound and computed tomography. *Clin Radiol.* 2004; 59(9):777–792. [PubMed: 15351243]
7. Pozo AL, Godfrey EM, Bowles KM. Splenomegaly: investigation, diagnosis and management. *Blood Rev.* 2009; 23(3):105–111. [PubMed: 19062140]
8. Swaroop J, O'Reilly RA. Splenomegaly at a university hospital compared to a nearby county hospital in 317 patients. *Acta Haematol.* 1999; 102(2):83–88. [PubMed: 10529511]
9. Schloesser LL. The diagnostic significance of splenomegaly. *Am J Med Sci.* 1963; 245:84–90. [PubMed: 13991866]
10. Arkles LB, Gill GD, Molan MP. A palpable spleen is not necessarily enlarged or pathological. *Med J Aust.* 1986; 145(1):15–17. [PubMed: 3724622]

11. Maia MD, et al. Evaluation of splenomegaly in the hepatosplenic form of mansonic schistosomiasis. *Acta Trop.* 2007; 101(3):183–186. [PubMed: 17336262]
12. Kumar PV, et al. Splenic lesions: FNA findings in 48 cases. *Cytopathology.* 2007; 18(3):151–156. [PubMed: 17573764]
13. Lishner M, et al. Fine needle aspiration biopsy in patients with diffusely enlarged spleens. *Acta Cytol.* 1996; 40(2):196–198. [PubMed: 8629397]
14. Solbiati L, et al. Focal lesions in the spleen: sonographic patterns and guided biopsy. *AJR Am J Roentgenol.* 1983; 140(1):59–65. [PubMed: 6600326]
15. Venkataramu NK, et al. Ultrasound guided fine needle aspiration biopsy of splenic lesions. *Br J Radiol.* 1999; 72(862):953–956. [PubMed: 10673946]
16. Zeppa P, et al. Fine needle aspiration biopsy of the spleen. A useful procedure in the diagnosis of splenomegaly. *Acta Cytol.* 1994; 38(3):299–309. [PubMed: 8191816]
17. Lieberman S, et al. Imaging-guided percutaneous splenic biopsy using a 20- or 22-gauge cutting-edge core biopsy needle for the diagnosis of malignant lymphoma. *Am J Roentgenol.* 2003; 181(4):1025–1027. [PubMed: 14500223]
18. Lopez JJ, et al. Role of ultrasound-guided core biopsy in the evaluation of spleen pathology. *Apms.* 2006; 114(7–8):492–499. [PubMed: 16907854]
19. Grover SA, Barkun AN, Sackett DL. The rational clinical examination. Does this patient have splenomegaly? *Jama.* 1993; 270(18):2218–2221. [PubMed: 8411607]
20. Koga T, Morikawa Y. Ultrasonographic determination of the splenic size and its clinical usefulness in various liver diseases. *Radiology.* 1975; 115(1):157–161. [PubMed: 1090976]
21. Bezerra AS, et al. Determination of splenomegaly by CT: is there a place for a single measurement? *AJR Am J Roentgenol.* 2005; 184(5):1510–1513. [PubMed: 15855107]
22. Larson SM, et al. Dimensions of the normal adult spleen scan and prediction of spleen weight. *J Nucl Med.* 1971; 12(3):123–126. [PubMed: 5100768]
23. Schulz HG, et al. Computerized tomography studies of normal morphology and volumetry of parenchymatous epigastric organs in humans. *Anat Anz.* 1986; 162(1):1–12. [PubMed: 3752529]
24. Frank K, et al. Sonographic determination of spleen size: normal dimensions in adults with a healthy spleen. *Ultraschall Med.* 1986; 7(3):134–137. [PubMed: 3529385]
25. Spielmann AL, DeLong DM, Kliever MA. Sonographic evaluation of spleen size in tall healthy athletes. *American Journal of Roentgenology.* 2005; 184(1):45–49. [PubMed: 15615949]
26. Rosenberg HK, et al. Normal splenic size in infants and children: sonographic measurements. *AJR Am J Roentgenol.* 1991; 157(1):119–121. [PubMed: 2048509]
27. Kaneko J, et al. Normal splenic volume in adults by computed tomography. *Hepatology.* 2002; 49(48):1726–1727. [PubMed: 12397778]
28. Hoefs JC, et al. A novel, simple method of functional spleen volume calculation by liver-spleen scan. *J Nucl Med.* 1999; 40(10):1745–1755. [PubMed: 10520718]
29. Prassopoulos P, et al. Determination of normal splenic volume on computed tomography in relation to age, gender and body habitus. *Eur Radiol.* 1997; 7(2):246–248. [PubMed: 9038125]
30. 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):525–527. [PubMed: 6974875]
31. Meier JM, et al. Assessment of age-related changes in abdominal organ structure and function with computed tomography and positron emission tomography. *Semin Nucl Med.* 2007; 37(3):154–172. [PubMed: 17418149]
32. Cools L, et al. Prediction of splenic volume by a simple CT measurement: a statistical study. *J Comput Assist Tomogr.* 1983; 7(3):426–430. [PubMed: 6841704]
33. Yetter EM, et al. Estimating splenic volume: sonographic measurements correlated with helical CT determination. *AJR Am J Roentgenol.* 2003; 181(6):1615–1620. [PubMed: 14627584]
34. Anonymous.
35. Lamb PM, et al. Spleen size: how well do linear ultrasound measurements correlate with three-dimensional CT volume assessments? *Br J Radiol.* 2002; 75(895):573–577. [PubMed: 12145129]

36. Farraher SW, Jara H, Chang KJ, Hou A, Soto JA. Liver and Spleen Volumetry with Quantitative MR Imaging and Dual-Space Clustering Segmentation. *Radiology*. 2005; 237(1):322–328. [PubMed: 16126927]
37. Linguraru MG, et al. Assessing hepatomegaly: automated volumetric analysis of the liver. *Acad Rad*. 2012; 19(5):588–598.
38. 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):420–426. [PubMed: 17367973]
39. Bland JM, Altman DG, et al. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986 Feb 8; 1(8476):307–310. [PubMed: 2868172]
40. DeLand FH. Normal spleen size. *Radiology*. 1970; 97(3):589–592. [PubMed: 5491745]
41. Geraghty EM, et al. Normal organ volume assessment from abdominal CT. *Abdom Imaging*. 2004; 29(4):482–490. [PubMed: 15024516]
42. Neiman, RS.; Orazi, A. 2nd ed. Philadelphia, London: Saunder; 1999. Disorders of the spleen.
43. Sekiyama K, et al. Prognostic value of hepatic volumetry in fulminant hepatic failure. *Dig Dis Sci*. 1994; 39(2):240–244. [PubMed: 8313804]
44. Mazonakis M, et al. Comparison of two volumetric techniques for estimating liver volume using magnetic resonance imaging. *J Magn Reson Imaging*. 2002; 15(5):557–563. [PubMed: 11997897]
45. Picardi M, et al. Spleen Enlargement Following Recombinant Human Granulocyte Colony-Stimulating Factor Administration for Peripheral Blood Stem Cell Mobilization. *Hematologica*. 2003; 88(7):794–800.
46. Zhang B, Lewis SM. Use of radionuclide scanning to estimate size of spleen in vivo. *J Clin Pathol*. 1987; 40(5):508–511. [PubMed: 3294908]
47. 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):805–811. [PubMed: 15105454]
48. 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.
49. Filicori F, Stock C, Schweitzer AD, Keutgen XM, Lagratta MD, Zarnegar R, Fahey TJ 3rd. Three-Dimensional CT Volumetry Predicts Outcome of Laparoscopic Splenectomy for Splenomegaly: Retrospective Clinical Study. *World J Surg*. 2013; 37(1):52–58. [PubMed: 23079780]
50. Chowdhury AI, Sultana SZ, Mannan S. Morphometric study of splenic volume in Bangladeshi cadaver. *Mymensingh Med J*. 2012; 21(3):445–449. [PubMed: 22828541]
51. Hammon M, Dankerl P, Kramer M, Seifert S, Tsybal A, Costa MJ, Janka R, Uder M, Cavallaro A. Automated detection and volumetric segmentation of the spleen in CT scans. *Rofo*. 2012; 184(8):734–739. [PubMed: 22618479]
52. Srisajjakul S, Prapaisilp P, Laorattkul N. Normal splenic volume assessment on CT in 426 adults. *Siriraj Medical Journal*. 2012; 64(2):43–46.

APPENDIX I – Spleen Disorders in the Patient Population

Spleen Disorders	# of Cases
Adult T-cell leukemia/lymphoma (ATL)	2
AIDS/Mycobacterium avium complex (MAC)	1
Alcohol Dependent	1
Anaplastic large cell lymphoma (ALCL)	2
Acquired Hypogammaglobulinemia (CVID) (also WHIM Syndrome)	4
Autoimmune lymphoproliferative syndrome (ALPS) type 1	4
B-cell Chronic Lymphocytic Leukemia (CLL)	20
Chronic granulomatous disease (CGD)	9

Spleen Disorders	# of Cases
Chronic myelogenous (or myeloid) leukemia (CML)	5
Colon Cancer	3
Coronary artery disease (CAD)/Renal Cell Cancer	1
Cutaneous T cell lymphoma (CTCL)	1
Dermatomyositis Lipodystrophy	1
Diffuse large B-cell lymphoma (DLBCL)	1
Diffuse well-differentiated lymphocytic lymphoma (DWDL)/CLL	1
D-MAC/myelodysplastic syndromes (MDS)	1
Epstein-Barr Virus (EBV) Lymphoproliferative Disease	1
Follicular lymphoma	5
Hairy cell leukemia (HCL)	2
Hepatitis C (HCV)	1
HIV	1
HIV/Hepatitis B (HBV)	1
HIV/Kaposi's sarcoma (KS)/Multicentric Castleman's Disease	2
HIV/NHL	1
HIV/Nodular regenerative hyperplasia	1
Hodgkin's lymphoma	1
Human T-lymphotropic virus (HTLV) type I/ATL	1
Hyponatremia	1
Kaposi Sarcoma	2
Large granular lymphocytic leukemia (LGL Leukemia)	1
Liver Disease	1
Lymphoid Granulomatosis (LYG)	3
lymphoma (Unspecified)	4
Mantle Cell Lymphoma (MCL)	12
medullary thyroid cancer / cirrhosis	1
Melanoma	6
Metastatic Pancreatic Cancer	1
Multicentric Castleman's Disease	3
Mycobacterium Avium Intracellulare (MAI)	1
Neuroendocrine cancer with liver mets	1
Non-Hodgkins Lymphoma (NHL)	4
Nontuberculous mycobacteria (NTM)	1
Parasitic Infection	2
Peripheral T-cell lymphoma (PTCL)	1
Prostate Carcinoma	3
Small lymphocytic lymphoma (SLL)/CLL	1
T-cell leukemia	1
T-cell Non-Hodgkins lymphoma (NHL)	2

APPENDIX II – STARD Flowchart for Data Selection

Figure 1.

Bland-Altman height agreement plots between a) two observers, b) automated method (CAD) and observer 1, c) CAD and observer 2. The mean error is shown in solid line and the 95% limits of agreement (± 1.96 SD) in dashed lines.

Figure 2. Relation between splenic volume and CC height. The normal and splenomegaly cases are marked according to the clinical reports (before consensus).

Figure 3.
Examples of spleens in two different patients (one patient per row) that have normal splenic CC height but enlarged splenic volume. Axial slices of the CT data are shown in superior to inferior order from left to right. The blue color indicates the automated segmentation.

Table 1**AVERAGE SPLEEN VOLUMES AND HEIGHTS**

Cases	Mean/std
Normal Volume: n = 44	236.89±77.58 ml
Normal Volume/BSA: n = 44	124.75±39.66 ml
Normal Height: n = 44	9.20±1.46 cm
Enlarged Volume: n = 88	1004.75±644.27 ml
Enlarged Volume/BSA: n = 88	526.53±320.84 ml
Enlarged Height: n = 88	17.31±3.5 cm

NOTE: Average volumes and cephalocaudal heights for normal and enlarged spleens. BSA represents the body surface area and was used a normalization factor.

Table 2**CORRELATION BETWEEN SPLEEN SIZE AND PATIENT'S BSA/AGE**

Cases/Correlations factor	Correlation (p-value)
Normal Volume : n=44 / Patient BSA	0.28 (p=0.02)
Normal Volume : n=44 / Patient Age:	0.00 (p=0.99)
Enlarged Volume : n=88 / Patient BSA	0.32 (p=0.002)
Enlarged Volume : n=88 / Patient Age	0.10 (p=0.32)
Normal CC Height : n=44 / Patient BSA	0.35 (p=0.004)
Normal CC Height : n=44 / Patient Age	-0.14 (p=0.26)
Enlarged CC Height : n=88 / Patient BSA	0.27 (p=0.008)
Enlarged CC Height : n=88 / Patient Age	0.01 (p=0.90)

NOTE: Correlation coefficients (R) and associated p-values are presented between volumes/cephalocaudal (CC) height measurements and patient's body surface are (BSA) and age or normal and enlarged spleens.

Table 3

RECEIVER OPERATING CHARACTERISTIC CURVE ANALYSIS

Criterion (n = 172)		Sensitivity / Specificity		
		Mild Splenomegaly (n=15)	Massive Splenomegaly (n=107)	All Splenomegaly (n=122)
Radiologists		66.66* / 78.00	99.06 / 78.00	95.08* / 78.00
CC Height	11 cm	53.33 / 68.00	99.06 / 68.00	93.44 / 68.00
CC Height	12 cm	33.33 / 76.00	96.26 / 76.00	88.52 / 76.00

NOTE: Sensitivity and specificity of radiologists and CC height criteria to detect splenomegaly were computed using the S-core as reference standard. CC height 11/12 cm refers to a threshold of 11/12 cm in CC height to detect splenomegaly.

* indicates a significant difference between the performances of radiologists and CC12.

Table 4

PREVIOUS REPORTS ON SPLEEN SIZE

Reference	Type of Data	Gender	No. Cases	Mean/Maximum Volume (ml)	Mean/Maximum Height (cm)	Imaging Modality
Kaneko et al. [27]	Normal	M	80	119±40 / -	-	CT
Spielmann et al. [25]	Normal	M	82	-	11 / -	US
Chowdhury et al. [50]	Normal	M	87	75.2±3.7/200	-	Cadaver
Kaneko et al. [27]	Normal	F	70	108±39 / -	-	CT
Spielmann et al. [25]	Normal	F	47	-	10 / -	US
Chowdhury et al. [50]	Normal	F	33	60.5±4.9/150	-	Cadaver
Henderson et al. [30]	Normal	M/F	11	219±76 / -	-	CT
Schulz et al. [23]	Normal	M/F	38	169 / -	11.1 / -	CT
Prassopoulos et al. [29]	Normal	M/F	140	214.6 / 314.5	-	CT
Bezerra et al. [21]	Normal	M/F	249	-	- / 9.76	CT
O'Reilly et al. [5]	Normal	M/F	170	-	- / 12	CT
Srisitjakul et al. [52]	Normal	M/F	426	124.1±51.8/430.8	8.5±1.4/13.9	CT
Mazonakis et al. [43]	Normal	M/F	16	204.8 / 289.8	-	MRI
Frank et al. [24]	Normal	M/F	793	-	- / 11 (95%)	US
Picardi et al. [44]	Normal	M/F	10	240 / 380	10 / 11.5	US
Hoefs et al. [28]	Normal	M/F	11	201±77 / 335	-	SPECT
Larson et al. [22]	Normal	M/F	26	-	10±1.5 / 12.9	Photoscan
Henderson et al. [30]	Diseased	M/F	12	660±336 / -	-	CT
Filicori et al. [49]	Diseased	M/F	88	1124.6±1200.1/-	-	CT
Hammon et al. [51]	Diseased	M/F	15	268.2 ±114.6/-	-	CT
Zhang et al. [45]	Diseased	M/F	14	1716 / 3585	-	US
Picardi et al. [44]	Post treatment	M/F	13	470 / 1200	-	US