Methods of Determining the Size of the Adult Liver Using 2D Ultrasound: A Systematic Review of Articles Reporting Liver Measurement Techniques

Jessie T. Childs, BHSc, Grad Dip Medical Sonography¹, Adrian J. Esterman, PhD, FACE², Maureen Phillips, DMU, MHealth, PhD¹, Kerry A. Thoirs, PhD¹, and Richard C. Turner, MBBS, BMedSc, FRACS³

Abstract
Sonography is an effective, noninvasive, safe, and inexpensive technique for measurement of the liver. Measurements of the liver using 2D ultrasound aid in diagnosing and tracking liver disease and in surgical planning. Multiple studies have developed techniques to measure the adult liver using 2D ultrasound. The aim of this systematic literature review was to determine whether a correctly developed, reliable, and reproducible technique for measuring the liver using 2D ultrasound has been reported. Analysis of studies meeting the inclusion criteria showed that none of these studies has been rigorously undertaken in terms of sample size justification, validity, reliability, and/or statistical analysis.

Keywords
ultrasound, liver, measurement

The capability to determine the size of the human liver allows sonographers and physicians improved ability to diagnose and understand underlying liver pathology and to determine the most effective treatment approach to disease.¹⁻⁴ Accurate liver measurement techniques allow for more effective surgical planning, aiding in liver transplantation and determination of adequate residual liver volume after resection. In addition, it enables superior postsurgical monitoring to reduce postoperative liver failure and patient complications.³ Simple evaluation of liver size also gives information regarding the presence and course of disease and the success or failure of treatment through serial studies over variable lengths of time.²⁻⁴⁻⁶

Sonography is a real-time, rapid, and inexpensive technique to evaluate the size of the adult liver. With the increasing availability of smaller, more portable and affordable ultrasound units, office and bedside 2D sonography performed by clinicians is becoming more attractive as a point-of-care diagnostic technique. At present, computed tomography (CT) is considered the gold standard for assessment of liver size and volume. However, there are substantial advantages to using sonography over CT in that the patient is not exposed to radiation; the size of the liver can be assessed as frequently as necessary. Currently, the determination of liver size by many clinicians is predominantly performed using the percussion technique during physical examination.³ Only when more detailed assessment is required do clinicians turn to radiographic techniques such as CT, sonography, or magnetic resonance imaging (MRI).

A common guideline in clinical ultrasound practice and ultrasound textbooks is that a liver measuring over 16 cm in the midclavicular line is enlarged.¹ However, different protocols have been described to make this measurement. To date, no systematic review has been published investigating the accuracy, reliability, and practicality of measurements that assess the size of the adult liver using 2D ultrasound.

¹University of South Australia, Adelaide, South Australia, Australia
²University of South Australia and Centre for Research Excellence in Chronic Disease Prevention, James Cook University, Australia
³University of Tasmania School of Medicine, Tasmania, Australia

Corresponding Author:
Jessie T. Childs, University of South Australia, University of South Australia City East Campus, Office BJ1–18, Frome Rd, Adelaide 5000, Australia.
Email: Jessie.childs@unisa.edu.au
ultrasound. Given the potential clinical importance of such measurements, there is a clear need for an accurate and reliable technique of determining the size of the adult liver using sonography, one that can be undertaken simply and quickly with modern-day equipment using a defined protocol. Given that clinicians with a wide variety of backgrounds and training are increasingly performing sonography, it is paramount that this method be able to be performed with limited training. Such a method would be of additional benefit if it could determine the volume of the liver to aid in liver transplantation and resection. This review was undertaken to analyze the measurement methods reported in the literature to assess their usefulness in the development of such a measurement technique.

**Background**

**Hepatomegaly**

The liver is the largest gland in the human body and is involved in many processes. The major functions of the liver include metabolism of proteins, carbohydrates, fats, vitamins, and other substances; synthesis of cholesterol and bile acids; bilirubin transport and metabolism; synthesis of factors of blood coagulation; regulation of blood volume; destruction and metabolism of erythrocytes; metabolism and transport of medications; and removal of waste products from the blood.8

The increase in the size of the liver beyond its normal dimensions is termed hepatomegaly. Hepatomegaly is a sign of a liver in distress and undergoing a disease process.8 The causes of hepatomegaly are varied and include pathophysiological causes such as dilated hepatic sinusoids, high venous pressure causing congestion, conditions causing the hepatocytes to malfunction and enlarge, fatty infiltration of the parenchymal cells causing fibrous tissue formation, disease processes causing distension of liver cells with glycogen or amyloid, or conditions resulting in space-occupying lesions within the liver tissue.3,4,8 Depending on initial cause, hepatomegaly can potentially be reversible if treated correctly.9

**Percussion Technique**

Liver span is determined by physicians using a percussion technique. The liver/lung edge and lower liver edge in the mideclavicular line are localized by means of differentiating a resonant percussion note from dullness.7 Normal values are 9 ± 2 cm in women and 11 ± 2 cm in men.7 Values superior to this would be considered to represent hepatomegaly. In general, estimation of the size of the liver using palpation or percussion is not accurate and may be misleading.9

**Sonography**

Ultrasound is an imaging modality that uses high-frequency sound waves inaudible to the human ear to create a 2D gray-scale image of structures within the body.10 Diagnostic ultrasound was developed in the 1940s and has since evolved into a mainstay of the medical imaging field.11 Ultrasound allows rapid assessment of the liver and its related vasculature. The normal sonographic appearance of the liver is that of a homogeneous light gray structure containing dark vessels and bright ligaments.

Recent advances in ultrasound have seen the liver able to be scanned using a 3D ultrasound technique. This technique requires more expensive, less portable equipment and greater training time, which limits its accessibility to many departments and clinicians. Consequently, articles describing methods of determining the size of the liver using 3D ultrasound techniques were not included in this review.

For clinicians and sonographers to convey information regarding the size of the liver, a uniform approach needs to be taken with regard to its measurement. This measurement technique needs to be simple and fast to perform on readily available equipment. The literature suggests that there are multiple different techniques to measure the adult liver using ultrasound. This systematic review was undertaken based on the guidelines of the 2009 PRISMA statement.12 The rigor of studies reporting methods of measuring the adult liver using 2D ultrasound was assessed in terms of their methods, equipment, sample size, statistical analysis, reliability, and validity. The review was undertaken to assist in the development of a reliable, simple, and accurate measurement technique for determining the size of the adult liver.

**Method**

**Search Terms**

The electronic databases of Ovid, Medline, PubMed, Cinhall, Scopus, and Wiley Interscience were searched multiple times by one of the authors (J.T.C.) for articles reporting methods of measuring the size of the liver using ultrasound. The last search was run on April 18, 2013. Search terms included for each database were ultrasound, sonographic, ultrasonic, ultrasonography, sonographic measurement, hepatomegaly, liver enlargement, liver size, and liver measurement. An outline of the search strategy used for each database search can be found in Table 1.

**Selection Criteria**

Any duplicate articles were removed from the search results. Based on independent scrutiny of the title and
abstract of each article by two of the authors (JTC and KAT) using defined selection criteria (Table 2), the full text of studies identified as being potentially relevant were retrieved. The full text of all retrieved studies was evaluated for suitability for inclusion into the review by both authors (J.T.C. and K.A.T.) using the defined selection criteria. Further analysis of the full text by a third investigator (A.J.E.) focused on excluding any articles that did not aim primarily to develop or evaluate a technique to measure the liver using sonography. On occasions where there was disagreement between the investigators on the inclusion or exclusion of an article, a consensus was reached following discussion.

Table 1. Database Search Strategy.

<table>
<thead>
<tr>
<th>Search Term Plus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ultrasound</td>
</tr>
<tr>
<td>Ultrasound</td>
</tr>
<tr>
<td>Liver size</td>
</tr>
<tr>
<td>Liver measurement</td>
</tr>
<tr>
<td>Hepatomegaly</td>
</tr>
<tr>
<td>Liver enlargement</td>
</tr>
<tr>
<td>Liver volume</td>
</tr>
<tr>
<td>Sonographic</td>
</tr>
<tr>
<td>Liver size</td>
</tr>
<tr>
<td>Liver measurement</td>
</tr>
<tr>
<td>Hepatomegaly</td>
</tr>
<tr>
<td>Liver enlargement</td>
</tr>
<tr>
<td>Liver volume</td>
</tr>
<tr>
<td>Ultrasonic</td>
</tr>
<tr>
<td>Liver size</td>
</tr>
<tr>
<td>Liver measurement</td>
</tr>
<tr>
<td>Hepatomegaly</td>
</tr>
<tr>
<td>Liver enlargement</td>
</tr>
<tr>
<td>Liver volume</td>
</tr>
<tr>
<td>Ultrasonography</td>
</tr>
<tr>
<td>Liver size</td>
</tr>
<tr>
<td>Liver measurement</td>
</tr>
<tr>
<td>Hepatomegaly</td>
</tr>
<tr>
<td>Liver enlargement</td>
</tr>
<tr>
<td>Liver volume</td>
</tr>
<tr>
<td>Sonography</td>
</tr>
<tr>
<td>Liver size</td>
</tr>
<tr>
<td>Liver measurement</td>
</tr>
<tr>
<td>Hepatomegaly</td>
</tr>
<tr>
<td>Liver enlargement</td>
</tr>
<tr>
<td>Liver volume</td>
</tr>
</tbody>
</table>

Table 2. Criteria on Which Studies Were Selected for Systematic Review.

<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>The study aimed to develop, evaluate, or compare against another measurement technique a method for determining liver size with 2D ultrasound.</td>
</tr>
<tr>
<td>Exclusion Criteria</td>
</tr>
<tr>
<td>Study not in English and no English translation was available.</td>
</tr>
<tr>
<td>Subjects were animals.</td>
</tr>
<tr>
<td>All subjects were younger than 18 years.</td>
</tr>
<tr>
<td>The report was a case study or review.</td>
</tr>
</tbody>
</table>

170 studies identified from database search

24 studies retrieved following review of abstracts

27 studies retrieved following assessment of reference lists of the 24 studies above

51 studies assessed in detail by two investigators

28 studies chosen as not having any of the exclusion criteria and assessed by a third investigator

18 studies which met all the inclusion and exclusion criteria were included in this review

Search Results

The search of all databases identified 170 studies potentially reporting techniques of measuring liver size using sonography. Following review of the abstracts against the selection criteria, 24 studies were deemed potentially relevant and the full text retrieved. The reference list of these 24 studies was perused for more potentially relevant studies. From this, the full text of a further 27 studies was retrieved, resulting in 51 studies overall. The full text of the 51 retrieved studies was assessed in detail, and those found to have any of the exclusion criteria were removed. As a result, 28 studies were chosen for further analysis by the third investigator. Following this analysis and discussion by the authors, a further 10 studies were excluded as they did not primarily aim to develop or evaluate a method to measure the liver using sonography. This left the 18 studies included in this review. All included studies were observational and provided descriptions of a technique to measure the adult liver using 2D sonography. Figure 1 shows a flowchart with the results of the literature search.

Measurement Techniques

Four groups of measurement techniques were evident within the included articles. These were volume measurements, body ratio measurements, representative measurements, and volumetric measurements. A volume
measurement was considered any reported measurement technique that resulted in an estimate of the overall volume of the liver. Body ratio measurements were defined as those that measured the size of the liver as a ratio to body size. Representative measurements were those that did not reflect the true volume or size of the liver in more than one dimension. Volumetric measurements were those that used simple measurement techniques to develop a hepatic volumetric index (HVI). Some of these studies then went on to use a mathematical formula to translate this into a true liver volume.

**Volume Measurement Techniques**

Nine studies were identified that calculated liver volumes. These studies used five different variations of technique:

- **Technique 1:** Serial longitudinal scans of the liver were taken, typically at 1-cm intervals. The liver contour on each slice was displayed via computer software and delineated with a light pen, which allowed tracings to be made on the screen. An area-measuring program calculated the area of each slice. The overall volume was the sum of the areas of each slice. 1, 13, 14

- **Technique 2:** Serial longitudinal scans of the liver were taken at 1-cm intervals extending from the lateral most visible point of the liver to the left extremity of the liver. The outline of the liver on each image was traced and the area calculated using a tabletop engineering drafting instrument for measuring areas of any shape called a polar planimeter. Liver volume was determined by summing the areas of each slice. The volume of the remaining right lateral-most tip was determined using a formula that was based on the area of the first longitudinal section and the height of liver. Overall liver volume was determined by adding these two volumes together. 1, 13, 14

- **Technique 3:** Serial longitudinal scans of the liver were taken at 1-cm intervals extending along the right lower chest from left to right. The outline of the liver on each image was traced and the area calculated using a polar planimeter as in technique 2. The portion of the liver to the left of the first slice was calculated using the following formula: 1/3 × the area of the last left slice × the distance from this slice to the tip of the left lobe of the liver. The portion of the liver to the right of the last slice was calculated using the following formula: 2/3 × the area of the last right slice × the distance from this slice to the right extremity of the liver. Total liver volume was determined as a sum of the areas of each slice and the volume of the left and right tips. 17

- **Technique 4:** Serial longitudinal scans of the liver were taken with an interval of 10 degrees through a common axis of rotation perpendicular to the transverse plane. A photograph of each image was taken and a line indicating the transverse scanning section was drawn. On this line, the distance was measured from the skin surface to the axis of rotation. These images were then used for digitizing and subsequent computer calculation of liver volume. 18, 19

- **Technique 5:** Serial scans of the left lobe of the liver were taken at 1.4-cm intervals from the umbilical portion of the left portal vein to the medial edge of the left segment of the liver. The area of liver on each scan was measured with proprietary built-in software and the volume of the left lobe of the liver calculated as the sum of all areas multiplied by the thickness of each slice (1.4 cm). 20 Clinically, the size of the entire liver needs to be able to be measured, and as this study measured only the volume of the left lobe, it is of little value and was excluded from further evaluation.

**Reliability and Validity of Volume Measurement Techniques**

The reliability and validity of the volume estimation techniques were well reported in the literature. The validity of technique 1 was established by Raeth et al. 13 who reported close agreement (coefficient of variation 8%) between liver volumes measured using this technique on cadavers compared with the volume of the same livers measured using water displacement after removal. Van Theil et al. 14 and Fritschi et al. 4 compared their measurements against water displacement and CT, respectively. While neither study reported statistical analysis, intrinsic validity can be assumed. Van Thiel et al. 14 was the only study using technique 1, which reported interobserver reliability. Their interobserver error for measuring a single slice or for measuring the entire liver volume was reported to be <5%. Raeth et al. 13 indicated in their article that they performed test-retest reliability, but they did not report these results.

The validity of technique 2 was not addressed by either study using this technique. Both studies, however, reported reliability, with Carr et al. 15 reporting intrarater reliability discrepancies averaging 9.4% of the mean. Baddeley et al. 16 used an analysis of variance to test intrarater reliability, reporting an F value of 13.7 with 33 and 12 degrees of freedom (P < .001). Using technique 3, Leung et al. 17 did not assess validity but reported high reliability with an intraobserver error of 15% to 20% and an interobserver error of 20% to 25%.

Two studies used technique 4, and both addressed the validity of this method. Rasmussen (1972) 18 and
Rasmussen (1978)\(^1\) both compared ultrasound volumes on cadavers with those obtained by water displacement following autopsy. In his first study, Rasmussen\(^1\) reported the differences between both measurement techniques to correspond to a standard deviation of 168 mL and a 95% confidence interval for a single determination of ±375 mL. In the second study, Rasmussen\(^1\) reported a coefficient of variation of 13.8% with a significant correlation, \(r = 0.85, P < .001\). The earlier study by Rasmussen was the only study to address the reliability of this method, with intrarater reliability tested on 30 subjects. Differences between observers varied from 0 to 180 mL, with 95% confidence intervals for a single determination of ±142 mL for females and ±88 mL for males\(^1\).

### Ease of Use of Volume Measurement Techniques

For a liver measurement technique to be used in clinical practice, it must be a rapid assessment that is easy to perform as well as reliable and valid. The primary limiting factor of volume techniques to measure the liver is that those that have been reported are time-consuming and tedious, with measurements done offline and not in real time. The methods of measurement reported in the studies evaluated involved dividing the liver into multiple sections and summing the value of the volumes calculated from each segment. Another limiting factor of these techniques is that the equipment used in these studies is now out of date and not commercially available. Volume measurements may be more easily attainable with newer, more modern ultrasound equipment, but no studies have yet evaluated this. While reliability and validity were well reported, the results of these tests as reported in the literature were not ideal. For these measurement techniques to be applied clinically in a routine fashion, methods with better reliability and validity are necessary.

### Body Ratio Measurement Techniques

One study was identified that used a body ratio measurement to determine the size of the liver\(^2\).

Technique: An image was taken at the level of the maximum transverse diameter of the liver. The cross-sectional area of the liver on this scan was measured by tracing the outline with a planimeter on a recorded (Polaroid) image. The cross-sectional area of the abdomen at this level was similarly determined and the two compared.

The reliability and validity of this technique were not addressed in the article. In addition, this technique, while simple, was undertaken on equipment no longer commercially available. An adaptation of this technique could be produced with modern-day equipment; however, it would be laborious and time-consuming. The use of a single measurement at a specific level to predict liver size also has a high potential for significant error since it may not always be reflective of what is happening over the entire liver. A technique that results in a true measured volume of the liver would be of more use clinically.

### Representative Measurement Techniques

Five studies were found in the literature search that used a representative measurement of the liver to estimate volume, but none of these dealt primarily with determining a technique to measure the liver but rather focused on tracking changes in size over time or developing normal reference ranges for a single measurement of the liver\(^2\). Two studies described the techniques to determine a representative measurement of the liver, each using a different technique\(^2\). These studies provided a semi-quantitative estimate of liver size by measuring the liver in only one or two dimensions.

Technique 1: Measurements of the craniocaudal and anteroposterior diameters of the liver in the midclavicular line and midline were taken (Figure 2). The anteroposterior diameter was measured at the midpoint of the longitudinal diameters\(^2\).

Technique 2: The craniocaudal diameter only of the liver was measured in the midclavicular line, as shown in Figure 3\(^2\).

### Reliability and Validity of Representative Measurement Techniques

The validity and reliability of the representative measurement techniques were not reported by either study. Review of the current literature shows that validity and reliability of this measurement technique are still widely unreported, with no validity being reported for any study found in our search using representative techniques. The only mention of reliability in the literature was found in the study by Gosink and Leymaster\(^2\), who stated that two observers made an independent determination of the ultrasound measurements from the same film, but they did not report the results for interobserver or intraobserver agreement.

### Ease of Use of Representative Measurement Techniques

Representative techniques are quickly and easily performed in the clinical setting and give a rapid semiquantitative estimate of liver size. These techniques, however,
have the same problem as the body ratio technique in that they rely on the assumption that a single plane is representative of the entire liver, which reduces their usefulness. Niderau and Sonnenberg reported taking four measurements in their study but in only two planes. Representative measurements are easily performed on modern-day equipment, and it should be noted that many imaging departments currently use these techniques or variations of these techniques to determine liver size. A common clinical guideline is that a liver measuring over 16 cm in the midclavicular line is enlarged. This protocol is based on the study by Kratzer et al., who reported validity but failed to report reliability. This study also used suboptimal statistical methods of data analysis. Many imaging departments using this protocol also measure in the anteroposterior plane rather than in the cranio-caudal plane as described in the original studies.

**Volumetric Measurement Techniques**

Six studies calculated volumetric measurements of the liver. All of these studies took simple sonographic measurements of the liver. Many studies also describe the development of an HVI, with five of these studies then converting these measurements to volumes using a derived formula. Four different techniques were used across the six studies.

Technique 1: The liver was divided into three compartments. The height (H), thickness (T), and width (W) of each compartment were measured and the volume of each compartment calculated from these measures. Compartment 1 included liver segments II and III, compartment 2 included segment IV, and compartment 3 included segments V to VIII. The volume of compartments 1 and 3 was both calculated using the formula $H \times T \times W \times \pi / 6$, while the volume of compartment 2 was calculated using formula $H \times T \times W \times \pi / 4$.

Technique 2: Images were taken at two equidistant points through the liver and the circumference of
The liver slices at these points measured. Liver volume was estimated as \((\frac{a + b}{3}) \times L\), where \(a\) and \(b\) = the circumferences at the two points and \(L\) = the transverse diameter of the liver (Figure 4). \(^{30}\)

**Technique 3:** Maximum anteroposterior (AP), craniocaudal (CC), and transverse (Trans) diameters of the liver were measured in centimeters. Liver volume was then calculated using a derived formula. \(^{2,31,32}\) Glenn et al\(^{31}\) and Elstein et al\(^{2}\) multiplied their AP, CC, and Trans measurements together to arrive at an HVI that was placed in the formula rather than the individual measurements. The formulas used by Zoli et al\(^{31}\) were Volume = 133.2 + 0.422(CC × AP × Trans) or Volume = −2364 + 89.2 × CC + 119.9 × AP + 59.6 × Trans. The HVI formula used by Glenn et al\(^{32}\) was HVI = 2.55 × liver volume − 0.12. The formula used by Elstein et al\(^{2}\) was Liver volume = 320.86 + 0.317 × HVI (Figure 5).

Technique 4: Maximum anteroposterior, craniocaudal, and transverse diameters of the liver were measured. These measurements were then multiplied together and divided by 27 to arrive at an HVI. \(^{6}\)

### Reliability and Validity of Volumetric Measurement Techniques

Technique 1, by Kitajima et al,\(^{29}\) addressed validity intrinsically by comparing liver volumes with volumes obtained using water displacement. They tested for differences by using an analysis of variance and subgroup analysis. No significant differences were found between the measures, with a correlation coefficient of \(r = 0.916\) \((P < .001)\). No intra- or interrater reliability testing was reported for this technique. Using technique 2, Gladisch et al\(^{30}\) examined validity by using linear regression to compare volumes measured using sonography with volumes measured using water displacement. They reported all values falling within the 95% confidence interval. Again, no intra- or interrater reliability testing was performed. Of the three studies using technique 3, only Zoli et al\(^{31}\) addressed reliability. Intraobserver reliability was not performed, but interobserver variation was reported being within ±8%. No study using this technique reported validity. Technique 4 was used by Boscaini and Pietri,\(^{6}\) but they did not report validity or reliability. The formulas for conversion were generated by comparing the volumetric indices with volumes calculated from CT scans,\(^{2}\) MRI scans,\(^{32}\) or direct measurement following resection.\(^{30}\) Zoli et al\(^{31}\) did not state the source of their formula, casting doubt on its usability.

### Ease of Use of Volumetric Measurement Techniques

The volumetric measurement techniques are advantageous in that they are simple and rapid to perform, and they can be reproduced easily on modern-day equipment. The ability to convert these simple measurements into a volume using an equation would be of great value in a clinical setting if their validity and reliability can be verified. Using multiple simple measurements to generate a liver volume would be superior to that of using only a single measurement, which gives a semiquantitative estimate of liver size.

### Critical Analysis and Qualitative Rating System

After removal of the volume study by Hatsuno et al,\(^{20}\) which assessed only the left lobe of the liver, the remaining 17 studies were collated according to their developed measurement protocol (volume, representative, body ratio, volumetric). A critical analysis table was used to group studies that used the same technique to measure the liver. The criteria for analysis were developed as an iterative process by the authors. Each study was evaluated by three independent reviewers, and after comparison of their results, each study was given a quality rating score. For each criterion, studies were awarded one point if they fulfilled the criterion and a zero score if they did not. Each included criterion was selected since it affected the integrity of the study. In those studies that developed a formula for generation of liver volume, each included...
criterion affected the usefulness of the formula. The criteria that were developed were the following:

**Source of the sample stated**: To ascertain whether a study may have introduced bias, it is important to know the source of the sample.

**Sample from a broad-based population**: This is most important in the volumetric studies, which compared their measurements with CT or MRI to determine a formula to calculate liver volume. Use of a population with all small or all large livers could result in the development of a formula that would not be accurate when used on a broader population base.

**Sample size justified**: Any study should include a statistical justification for the sample size of the study.

**Inclusion and exclusion criteria stated**: To ascertain any introduced bias, a study requires clearly defined inclusion and exclusion criteria.

**Detailed description of ultrasound measurements provided**: For a measurement to be useful clinically, clear directions need to be given on how to perform the measurement, including anatomic landmarks, diagrams, and correct measurement plane. Without clear instructions, sonographers are unable to accurately reproduce the measurements.

**Technique able to be reproduced on modern-day equipment**: The measurement technique must be able to be performed on currently available equipment; otherwise, the method is outdated and not of use.

**Measurement(s) able to be quickly and easily performed**: With modern-day time constraints on departments and clinicians, it is important that a measurement be able to be performed quickly and easily.

**Statistical analysis performed**: Studies were awarded one point if they performed any type of valid statistical analysis of their data.

**Ideal statistical analysis**: Studies were awarded one additional point if their statistical analysis method was optimal.

**Reliability**: Studies were awarded one point if they demonstrated intra- or interrater reliability. A technique needs to be reproducible by a variety of users to be widely used in the clinical environment.

**Validity**: Studies were awarded one point if they demonstrated validity of their measurement technique. There is little usefulness of a measurement that is not valid.

The highest possible overall score for this qualitative rating system was 11. For a study to be considered as being correctly undertaken and the results reliable and usable in the current clinical environment, it would need to fulfill all of these criteria (and hence require a score of 11). Table 3 shows the ratings of the studies evaluated.

### Results of the Critical Analysis

Overall, the studies performed poorly when critically evaluated, with scores ranging from 3 to 6 out of the required 11. Not a single study describing a technique to measure the liver using sonography justified the sample size, with many studies failing to state the source of their...
## Table 3. Qualitative Analysis.

<table>
<thead>
<tr>
<th>Category</th>
<th>Study</th>
<th>Source of the Sample Stated</th>
<th>Sample From a Well-Distributed Population</th>
<th>Sample Size Justified</th>
<th>Entry and Exclusion Criteria Stated</th>
<th>Detailed Description of Ultrasound Measurement Provided</th>
<th>Technique Able to Be Reproducible on Modern-Day Equipment</th>
<th>Measurement Able to Be Quickly and Easily Performed</th>
<th>Statistical Analysis Performed</th>
<th>Ideal Statistical Analysis Performed</th>
<th>Reliability</th>
<th>Validity</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td>Baddeley et al&lt;sup&gt;16&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3/11</td>
</tr>
<tr>
<td></td>
<td>Carr et al&lt;sup&gt;15&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3/11</td>
</tr>
<tr>
<td></td>
<td>Fritschy et al&lt;sup&gt;9&lt;/sup&gt;</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4/11</td>
</tr>
<tr>
<td></td>
<td>Leung et al&lt;sup&gt;17&lt;/sup&gt;</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4/11</td>
</tr>
<tr>
<td></td>
<td>Raeth et al&lt;sup&gt;3&lt;/sup&gt;</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5/11</td>
</tr>
<tr>
<td></td>
<td>Rasmussen&lt;sup&gt;18&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>6/11</td>
</tr>
<tr>
<td></td>
<td>Rasmussen&lt;sup&gt;19&lt;/sup&gt;</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5/11</td>
</tr>
<tr>
<td></td>
<td>Van Thiel et al&lt;sup&gt;14&lt;/sup&gt;</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>6/11</td>
</tr>
<tr>
<td>Body ratio</td>
<td>Holmes et al&lt;sup&gt;21&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4/11</td>
</tr>
<tr>
<td>Representative</td>
<td>Niderau et al&lt;sup&gt;27&lt;/sup&gt;</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6/11</td>
</tr>
<tr>
<td></td>
<td>Sapira and Wilamson&lt;sup&gt;28&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4/11</td>
</tr>
<tr>
<td>Volumetric</td>
<td>Boscani and Pietri&lt;sup&gt;4&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5/11</td>
</tr>
<tr>
<td></td>
<td>Elstein et al&lt;sup&gt;3&lt;/sup&gt;</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4/11</td>
</tr>
<tr>
<td></td>
<td>Gladisch et al&lt;sup&gt;30&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5/11</td>
</tr>
<tr>
<td></td>
<td>Glenn et al&lt;sup&gt;32&lt;/sup&gt;</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5/11</td>
</tr>
<tr>
<td></td>
<td>Kitajima et al&lt;sup&gt;29&lt;/sup&gt;</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6/11</td>
</tr>
<tr>
<td></td>
<td>Zoli et al&lt;sup&gt;31&lt;/sup&gt;</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6/11</td>
</tr>
</tbody>
</table>
sample. Statistical analysis of the results in the studies was also suboptimal. In some cases, this was a reflection of the age of the study and the statistical methods available at the time of publication. The most appropriate analysis for method comparison studies involves the use of Bland-Altman plots, a formal test of zero bias, and Deming regression. Reliability is usually measured by the intraclass correlation coefficient (ICC). However, many of these methods were developed after these studies had been published.

While description of the ultrasound measurement techniques was generally good, many techniques require equipment that is no longer readily available. Despite an extensive review of the literature, the most recent study found was conducted in 2008, with a majority of studies having been done in the 1970s and 1980s. There may be more recent studies that have not been identified despite an exhaustive search, although there will be some difficulty in comparing the findings of studies 30 years apart, given the advancement in equipment, knowledge, and technique. Adaptations of these outdated techniques using modern-day equipment would be difficult. The reporting of validity and reliability among the techniques was varied, and this has already been discussed in detail. Reliability and validity should always be reported. To date, a study that has been correctly undertaken in terms of sample size justification, validity, reliability, and statistical analysis has yet to be performed.

Conclusion
The ability to determine liver volume sonographically would better aid in disease tracking and the planning and follow-up of liver transplant surgery. It would also mean volumes could be compared with those of recent CT scans for disease tracking. However, it appears that reported techniques to measure the volume of the liver are both time-consuming and outdated. Body ratio techniques and representative techniques have some merit in their ease of use but are limited in that they often only represent the liver in one dimension. It can be seen from the literature that the currently used sonographic techniques to determine whether a liver is enlarged using a single 16-cm cutoff for a representative measurement of liver size at the midclavicular line may not be accurate. Volumetric techniques hold promise as a quick and easy way to measure the liver. Some of these studies use an equation to convert such measurements into a volume, and these techniques appear potentially to be the most appropriate for accurately measuring the size of the adult liver.

Declaration of Conflicting Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

References