

The Interrater Reliability of Inferior Vena Cava Ultrasound by Bedside Clinician Sonographers in Emergency Department Patients

J. Matthew Fields, MD, Paul A. Lee, MD, Katherine Y. Jenq, MD, Dustin G. Mark, MD, Nova L. Panebianco, MD, MPH, and Anthony J. Dean, MD

Abstract

Objectives: Inferior vena cava ultrasound (IVC-US) is a noninvasive bedside tool to assess intravascular volume status. This study set out to investigate the interrater reliability of IVC-US by bedside clinician sonographers and determine whether alternative methods of IVC-US such as B-mode and visual estimation are equally reliable to traditional M-mode.

Methods: A convenience sample of adult emergency department (ED) patients was prospectively enrolled. Each patient underwent IVC-US by two different emergency physicians (EPs), each of whom first performed visual estimation of IVC percent collapse and of volume status, followed by caliper measurements in M-mode and B-mode. EPs were blinded to patient data and to the other sonographer's results. For each technique, interrater reliability was determined between the two EPs' assessments using intraclass correlation coefficients (ICC) for continuous data and Cohen's weighted kappa for categorical data. In addition, analysis was performed on M-mode diameter measurements to determine the relationship between sonographer and patient characteristics on interrater reliability.

Results: Five EPs performed 92 US exams on 46 patients. Using M-mode, the ICC for maximum IVC diameter was 0.81 (95% confidence interval [CI] = 0.67 to 0.89), and for minimum diameter was 0.77 (95% CI = 0.62 to 0.87). There were no statistically significant differences between the caliper methods used for IVC measurements (M-mode diameter, B-mode diameter, or B-mode area). Agreement for visually estimated IVC collapse (0.60, 95% CI = 0.36 to 0.76) was similar to agreement for calculated M-mode IVC collapse index (0.52, 95% CI = 0.27 to 0.71). Cohen's weighted kappa for volume status based on visual estimation of IVC filling (size, shape, and collapse) was 0.64 (95% CI = 0.53 to 0.73). ICC values for M-mode diameter measurements were significantly higher in studies involving patients who were noneuvolemic and studies in which sonographers had each performed at least five prior IVC-US.

Conclusions: Emergency physicians' US measurements of IVC diameter have a high degree of interrater reliability. IVC percent collapse by visual estimation or based on caliper measurements have lower, but still moderate to good reliability. The use of the visual estimation technique should be considered by clinicians who have learned to obtain measured parameters of IVC filling because it is equally reliable to traditional M-mode and can be performed more rapidly.

ACADEMIC EMERGENCY MEDICINE 2011; 18:98-101 © 2011 by the Society for Academic Emergency Medicine

Inferior vena cava ultrasound (IVC-US) provides a rapidly deployable, noninvasive, and repeatable measure of intravascular volume status.¹⁻⁴ IVC-US

From the Department of Emergency Medicine, University of Pennsylvania, Philadelphia, PA.

Received March 16, 2010; revision received May 5, 2010; accepted May 6, 2010.

The authors have no disclosures or conflicts of interest to report.

Supervising Editor: Alan Jones, MD.

Address for correspondence and reprints: Anthony J. Dean, MD; e-mail: Anthony.Dean@uphs.upenn.edu.

is increasingly used by clinician sonographers in the emergency department (ED) and intensive care unit to guide fluid management and resuscitation.⁴⁻¹⁰ Despite increased use, the interrater reliability of IVC-US among clinician sonographers is unknown.

With IVC measurement, several anatomic and physiologic factors exist that may affect interrater reliability. These include extrinsic structures compressing the IVC, variations in shape and orientation of the IVC, alterations in the respiratory rate and cycle, and degree of diaphragmatic excursion. Technical factors may also potentially lead to differences between measurements including sonographer experience, patient body habi-

tus, and the method used to perform measurements (visual estimation vs. M-mode vs. B-mode). This study set out to determine the interrater reliability of IVC-US in the hands of clinician sonographers using several different techniques of IVC assessment.

METHODS

Study Design

This was a prospective observational study examining interrater reliability between emergency physician (EP) sonographers in a convenience sample of ED patients. The institutional review board approved the study protocol and all patients provided informed consent.

Study Setting and Population

This study was conducted during April 2008 at an urban ED with an annual census of 55,000. The enrolling EPs consisted of two Postgraduate Year (PGY)-3 residents, two PGY-4 residents, and one US fellow. All EPs had met the training requirements for emergency ultrasonography as delineated in the 2001 guidelines of the American College of Emergency Physicians, including at least 16 hours of didactic lectures and over 150 technically adequate scans.¹¹ In addition, the study EPs underwent a standardized 1-hour training session (didactic, video/image review, and hands-on practice) by a single instructor (AJD) in IVC-US followed by 10 proctored and reviewed exams. In other respects, the study residents had not received US training beyond that routinely provided in our residency program, and study residents had little to no prior exposure to IVC-US, as it is not commonly performed in our ED.

Study Protocol

Sonographers were assigned days in a rotating manner to ensure each sonographer was paired at least once with each of the other four sonographers. A convenience sample of ED patients was enrolled. Inclusion criteria were age > 18 years, nonpregnant, and the ability to tolerate supine position for US examination.

Separate IVC studies were performed on each patient (one by each EP) less than 15 minutes apart. Each EP was blinded to the findings of the other, as well as to patient clinical data. Each sonographer first recorded a visual estimation of inspiratory IVC collapse (0%–100%). The sonographer then recorded a “gestalt” estimation of volume status by combining visually estimated collapse with the visual appearance of IVC size and shape on a four-point scale: 1 = hypovolemic, 2 = lower range of normal, 3 = higher range of normal, and 4 = hypervolemic. Sonographers then performed caliper measurements of the IVC using M-mode and B-mode and obtained video clips and still images for expert analysis. Upon completion of both studies, demographic and clinical information was obtained.

All IVC studies were performed with the patient in supine position, and all visual estimations and caliper measurements were made at approximately 2 cm below the entry of the hepatic veins using a Sonosite Micromaxx P17 phased array probe (Sonosite Inc., Bothwell, WA). Collapse index (also known as the caval index) was determined during normal passive

inspiration. The measured IVC collapse index (IVC-CI) was obtained with the usual formula: $IVC-CI = (IVC-max - IVC-min)/IVC-max$, where IVC-max is the maximal IVC diameter or area and IVC-min is the minimal IVC diameter or area.^{5,6,8} All studies were reviewed by a bedside sonographer with more than 10 years of experience (AJD) to determine correct vessel identification and caliper placement. The reviewer was blinded to sonographer visual estimations and patient data. Inadequate examinations (incorrect vessel identification or inappropriate caliper placements) were excluded from analysis.

Data Analysis

Summary statistics were performed on patient demographic and clinical characteristics. Interrater reliability was calculated for each US method using one-way random effects model intraclass correlation coefficients (ICCs) for continuous variables and Cohen's linear weighted kappa (κ) for categorical variables. In addition, the effects of sonographer and patient characteristics on ICC values were analyzed. To achieve a 95% confidence interval (CI) width of 0.4 for ICC values, 45 patients were required utilizing sampling theory of linear correlation. The enrollment goal of 50 patients allowed for a 10% rate of unobtainable IVC views, based on previous studies.^{5,6,8,9} Statistical analyses were performed using Stata release 11 (StataCorp, College Station, TX).

RESULTS

Fifty patients were enrolled. Four patients were excluded from analysis due to inability of one or both sonographers to obtain adequate images (three patients were reported as technically difficult, and one patient refused the second IVC-US exam). The 46 remaining patients were each evaluated by two of the five study EPs. The five study sonographers performed an average of 18 studies each (range = 15–22) and at least three studies with each of the other four study EPs. Expert review found no studies inadequate.

The median age was 41 years (interquartile range = 29–56). The mean (\pm SD) heart rate was 86 (\pm 16) beats/min, with 17% of patients (8/46) having a heart rate > 100 beats/min. The mean (\pm SD) arterial pressure was 92 (\pm 15) mm Hg. Fourteen percent of patients (6/46) had a respiratory rate > 20 breaths/min, and 26% (12/46) met body mass index criteria for obesity. Clinicians caring for the patients estimated that 39% (18/46) were hypovolemic, 50% (23/46) were euvolemic, and 11% (5/46) were hypervolemic.

There were no statistically significant differences in IVC measurements or collapse between the four methods (Table 1). Interrater agreement for gestalt assessment of volume status using visual estimation was substantial ($\kappa = 0.64$, 95% CI = 0.46 to 0.78). Analysis of M-mode diameter measurements revealed that sonographer IVC-US experience and the patient's volume status were significantly associated with interrater reliability (Table 2). Interrater reliability was not significantly associated with sonographer training level, patient's respiratory rate, or body mass index (Table 2).

Table 1
ICCs for IVC Maximum, Minimum, and CI Using Various
Ultrasound Methods

	IVC Maximum (95% CI)	IVC Minimum (95% CI)	IVC-CI (95% CI)
M-mode diameter	0.81 (0.67–0.89)	0.77 (0.62–0.87)	0.52 (0.27–0.71)
B-mode diameter	0.77 (0.62–0.87)	0.81 (0.69–0.89)	0.58 (0.34–0.74)
B-mode area	0.79 (0.65–0.88)	0.85 (0.74–0.91)	0.56 (0.32–0.73)
Visual estimation	N/A	N/A	0.60 (0.36–0.76)

ICCs = intraclass correlation coefficients; IVC = inferior vena cava; IVC-CI = IVC collapse index.

DISCUSSION

To our knowledge, the interrater reliability of IVC-US between clinician sonographers has never been studied. One study by Randazzo et al.⁷ measured agreement between EPs' and cardiologists' determination of central venous pressure by visual estimation of IVC collapse; however, caliper measurements were not performed. The higher agreement found in our study ($\kappa = 0.64$ vs. 0.40) may be due to greater similarity between technique, training, and experience of sonographers in our study or may be due to temporal proximity of examinations in our study (<15 minutes vs. < 4 hours) reducing the likelihood of interval changes in intravascular volume.

The current study demonstrates strong agreement between clinician sonographers for M-mode IVC diameter measurements, while agreement for M-mode IVC-CI was moderate. This finding may appear anomalous, but is due to the fact that the calculation of IVC-CI

causes the differences in diameter measurement to be augmented multiplicatively, so that the correlation coefficients between measurements of IVC-CI will always be lower than those between the measurements of its constituent components. The study also evaluated two other caliper techniques for IVC measurement (diameter and area measured from B-mode images) as well as visual estimation of IVC collapse. Our findings suggest that both measured and visually estimated IVC assessments are similar and repeatable among different providers.

While there was only moderate agreement for IVC-CI, there was moderate to substantial agreement for volume status using a four-point visually estimated scale of IVC filling. This improvement may be explained by the ability of the sonographer to rapidly assimilate multiple data points using visual estimation (IVC collapse averaged over multiple respiratory cycles, IVC shape and size, extrinsic causes of IVC collapse such as liver masses, and anatomic strictures) compared to isolated M-mode diameter measurements. Accurate use of visual estimation in ultrasonography has previously been described for assessment of left ventricular ejection fraction, which has been validated for both cardiologists and EPs.^{6,7}

Analysis of the data showed a significant improvement in interrater reliability when assessing hyper- and hypovolemic patients and when sonographers had performed at least five previous exams. The former finding reflects the fact that extreme states (e.g., 100% or 0% collapse) are visually less ambiguous than midrange states (e.g., distinguishing 40% vs. 60% collapse). This finding supports a previous study that suggested that IVC-US was reliable in critically ill patients whose central venous pressures are often at the extremes.⁸ Since each sonographer had performed 10 prior examinations, the latter finding, that interrater reliability improves after 5 studies, suggests that 15 studies may

Table 2
ICCs of M-mode IVC Maximum and Minimum Diameter Measurements by Various Sonographer and Patient Characteristics

	Maximum Diameter (95% CI)	p-value	Minimum Diameter (95% CI)	p-value
Sonographers				
All (<i>n</i> = 46)	0.80 (0.67 to 0.89)		0.76 (0.61 to 0.86)	
Sonographer 1 (<i>n</i> = 19)	0.69 (0.34 to 0.87)		0.67 (0.31 to 0.86)	
Sonographer 2 (<i>n</i> = 22)	0.74 (0.46 to 0.88)		0.76 (0.50 to 0.90)	
Sonographer 3 (<i>n</i> = 15)	0.88 (0.68 to 0.96)		0.75 (0.38 to 0.91)	
Sonographer 4 (<i>n</i> = 18)	0.85 (0.63 to 0.94)		0.84 (0.61 to 0.94)	
Sonographer 5 (<i>n</i> = 18)	0.84 (0.61 to 0.94)		0.77 (0.48 to 0.91)	
IVC-US studies performed				
≤5 studies for both (<i>n</i> = 13)	0.51 (0.18 to 0.71)	0.03	0.49 (0.18 to 0.70)	0.05
>5 studies for either (<i>n</i> = 33)	0.85 (0.72 to 0.92)		0.81 (0.70 to 0.90)	
Body mass index				
Nonobese (<i>n</i> = 34)	0.76 (0.56 to 0.87)	0.42	0.56 (0.27 to 0.76)	0.29
Obese (<i>n</i> = 12)	0.93 (0.76 to 0.98)		0.40 (−0.23 to 0.80)	
Respiratory rate (breaths/min)				
Normal (<i>n</i> = 40)	0.76 (0.58 to 0.86)	0.14	0.75 (0.57 to 0.86)	0.42
Elevated (<i>n</i> = 6)	0.93 (0.49 to 0.99)		0.80 (−0.05 to 0.98)	
Estimated volume status				
Euvolemic (<i>n</i> = 25)	0.59 (0.27 to 0.78)	0.01	0.62 (0.30 to 0.82)	0.07
Hypo- or hypervolemic (<i>n</i> = 21)	0.92 (0.80 to 0.97)		0.83 (0.63 to 0.93)	

ICCs = intraclass correlation coefficients; IVC = inferior vena cava; US = ultrasound.

be a reasonable number for training in this sonographic skill.

LIMITATIONS

Potential sources of error in our study include selection bias arising from convenience sampling. Study EPs may have enrolled patients they felt more confident about scanning. Data from four patients were excluded due to inability to obtain measurements, which was anticipated in the study design, and coheres with the 5% to 16% rate reported in other studies.^{5,6,8,9} US interest among the study physicians may limit generalizability due to a higher level of performance than would be obtained by typical EPs. However, it should be noted that prior to the study, the four resident sonographers had received the same US training as their peers, with the exception of a 1-hour review session and 10 proctored exams. To reflect typical ED conditions, inspiratory effort was not standardized; however, any resulting bias could have only have served to lessen interrater reliability.

Measurement bias may have occurred if EPs tried to match measurements to their previously determined visual estimates. Since sonographers were aware that their measurements were being studied, there may have been bias secondary to a Hawthorne effect. The expert review process was designed in an attempt to limit such bias by reviewing for inappropriate measurements. No such mismeasurements were identified by the reviewer.

CONCLUSIONS

Emergency physicians' ultrasound measurement of inferior vena cava diameter has a high degree of interrater reliability. Measurement of inferior vena cava collapse index based on these measurements has lower, but still moderate to good, reliability due to the multiplicative nature of this technique. These test characteristics make ultrasound assessment of the inferior vena cava by emergency bedside sonographers a potentially useful noninvasive and repeatable modality for intravascular volume assessment. The use of the visual estimation technique should be considered by clinicians who have learned to obtain measured parameters of inferior vena cava filling, since it is equally reliable and can be performed very rapidly.

References

1. Feissel M, Michard F, Faller JP, Teboul JL. The respiratory variation in inferior vena cava diameter as a guide to fluid therapy. *Intensive Care Med.* 2004; 30:1834–7.
2. Lyon M, Blaivas M, Brannam L. Sonographic measurement of the inferior vena cava as a marker of blood loss. *Am J Emerg Med.* 2005; 23:45–50.
3. Moreno FL, Hagan AD, Holmen JR, Pryor TA, Strickland RD, Castle CH. Evaluation of size and dynamics of the inferior vena cava as an index of right-sided cardiac function. *Am J Cardiol.* 1984; 53:579–85.
4. Nagdev AD, Merchant RC, Tirado-Gonzalez A, Sisson CA, Murphy MC. Emergency department bedside ultrasonographic measurement of the caval index for noninvasive determination of low central venous pressure. *Ann Emerg Med.* 2010; 55:290–5.
5. Carr BG, Dean AJ, Everett WW, et al. Intensivist bedside ultrasound (INBU) for volume assessment in the intensive care unit: a pilot study. *J Trauma.* 2007; 63:495–500.
6. Mark DG, Hayden GE, Ky B, et al. Hand-carried echocardiography for assessment of left ventricular filling and ejection fraction in the surgical intensive care unit. *J Crit Care.* 2009; 24:470.e1–7.
7. Randazzo MR, Snoey ER, Levitt MA, Binder K. Accuracy of emergency physician assessment of left ventricular ejection fraction and central venous pressure using echocardiography. *Acad Emerg Med.* 2003; 10:973–7.
8. Stawicki SP, Braslow BM, Panebianco NL, et al. Intensivist use of hand-carried ultrasonography to measure IVC collapsibility in estimating intravascular volume status: correlations with CVP. *J Am Coll Surg.* 2009; 209:55–61.
9. Wallace DJ, Allison M, Stone MB. Inferior vena cava percentage collapse during respiration is affected by the sampling location: an ultrasound study in healthy volunteers. *Acad Emerg Med.* 2010; 17:96–9.
10. Yanagawa Y, Sakamoto T, Okada Y. Hypovolemic shock evaluated by sonographic measurement of the inferior vena cava during resuscitation in trauma patients. *J Trauma.* 2007; 63:1245–8.
11. American College of Emergency Physicians. ACEP emergency ultrasound guidelines–2001. *Ann Emerg Med.* 2001; 38:470–81.