Point-of-Care Ultrasonography in Assessing Fluid Responsiveness in Sepsis Patients: Sonographer Characteristics, Noninferential Statistics, and Study Design

Answers to the September 2012 Journal Club Questions

Alan Chiem, MD, MPH

Editor’s Note: You are reading the 29th installment of Annals of Emergency Medicine Journal Club. This Journal Club refers to the article by Haydar et al that was published in the September 2012 Annals of Emergency Medicine issue.

Information about Journal Club can be found at http://www.annemergmed.com/content/journalclub.

Readers should recognize that these are suggested answers. We hope they are accurate; we know that they are not comprehensive. There are many other points that could be made about these questions or about the article in general. Questions are rated “novice,” (NOV) “intermediate,” (INT) and “advanced” (ADV) so that individuals planning a journal club can assign the right question to the right student. The “novice” rating does not imply that a novice should be able to spontaneously answer the question. “Novice” means we expect that someone with little background should be able to do a bit of reading, formulate an answer, and teach the material to others. Intermediate and advanced questions also will likely require some reading and research, and that reading will be sufficiently difficult that some background in clinical epidemiology will be helpful in understanding the reading and concepts.

We are interested in receiving feedback about this feature. Please e-mail journalclub@acep.org with your comments.

DISCUSSION POINTS

1. Sepsis is a complex cascade of immunologic mechanisms related to systemic infection. There are many causes of sepsis, and many of its hemodynamic and metabolic derangements can occur concomitantly or independently. To counter toxin-mediated vasodilatation, early goal-directed therapy (EGDT) advocates aggressive fluid resuscitation. However, several criticisms have been made about the use of central venous pressure (CVP) in assessing fluid responsiveness and guiding EGDT. The Figure below highlights the major mechanisms in sepsis, and how fluid responsiveness is assessed with CVP and inferior vena cava (IVC) measurement. Areas of intersection (A, B, C) indicate the variance shared by these factors and the degree to which CVP or IVC measurement is able to predict fluid responsiveness.

Figure. IVC measurement. A = Change in IVC that is attributable to fluid responsiveness B = Change in CVP that is attributable to fluid responsiveness C = Change in IVC that is attributable to change in CVP.

NOV A. Several articles have assessed probe positioning in relation to IVC measurement and its correlation with CVP. Describe potential patient-specific or operator-specific issues that might confound the accurate measurement of the inferior vena cava in septic patients.

NOV B. Discuss the evidence for using bedside ultrasonographic measurements of IVC variability to estimate fluid responsiveness. Are ultrasonographic measurements a more accurate measure of fluid responsiveness than CVP, urine output, jugular venous distention, or continuous pulse rate and blood pressure monitoring? What is the criterion standard for measuring the success of fluid resuscitation in patients with septic shock?

NOV C. In the Haydar et al study, only 54% of patients in the septic shock group had a CVP below 10 mm Hg compared with more than 60% in the severe sepsis and sepsis groups. The measured IVC diameter variation exceeded 50% in half of the septic shock patients, 56% in the sepsis group, and more than 70% in the severe sepsis group. Do these trends make sense according to the clinical spectrum of sepsis? What threats to the reliability and validity of bedside ultrasonography are
suggested by this discrepancy? What are possible explanations?

2. The point-of-care ultrasonography in this study differs significantly from how it is used in most EDs. In other words, the ultrasounds were performed over 2 hours after patient arrival in the ED, and treating physicians—who did not perform the ultrasounds—did not have access to the images until 20 minutes after they were performed.

- A. If many patients had treatment initiated as part of their sepsis bundle, how would this have influenced the ultrasonographic measures on the outcomes of interest?
- B. Most emergency department–based ultrasonographic research studies are performed by highly experienced physicians who have completed an ultrasound fellowship or equivalent training. Discuss the limitations associated with generalizing these study results to emergency physicians without such advanced ultrasonographic training.
- C. Might this study’s results represent the Hawthorne effect or expectation bias? Did this investigation truly measure ultrasonography’s utility or the effect of an explicit cueing of the sepsis management guidelines?
- D. How would you design a study that fully incorporates the use of ultrasonography in managing sepsis patients? What is the unit of analysis for an emergency department–based ultrasonographic research investigation? How does one calculate the required sample size for such an investigation?

3. The point-of-care assessment of cardiac contractility is a powerful example of the role of ultrasonography in clinical care.

- A. What sonographic measures of cardiac contractility provide the best estimation of ejection fraction?
- B. How well do noncardiologists fare in estimating cardiac contractility?
- C. How does the clinician apply various findings on ultrasonography, such as IVC variation greater than 50% and normal ejection fraction or IVC variation less than 50% and low ejection fraction, in septic shock patients?

4. Haydar et al incorporated detailed graphics (Figures 2 and 3 in their original article) to present the underlying distribution of their data.

- A. Why are the graphics included in the Haydar et al article more informative than more commonly reported summary statistics? What additional information does Figure 2 provide that might be missed if the authors had included only the box plots?
- B. How do you interpret Figure 3? When pre-point-of-care ultrasonography certainty is high for patients, are they expected to change more or less after point-of-care ultrasonography than patients with low pre-point-of-care ultrasonography certainty? Identify patients for whom point-of-care ultrasonography decreased the physicians’ certainty. Do they share any clinical characteristics?

- C. What are the advantages and disadvantages to analyzing effect size, such as through Cohen’s d statistic, as opposed to traditional significance testing, such as a paired samples t test?

ANSWER 1

Q1. Sepsis is a complex cascade of immunologic mechanisms related to systemic infection. There are many causes of sepsis, and many of its hemodynamic and metabolic derangements can occur concomitantly or independently. To counter toxin-mediated vasodilatation, early goal-directed therapy (EGDT) advocates aggressive fluid resuscitation. However, several criticisms have been made about the use of central venous pressure (CVP) in assessing fluid responsiveness and guiding EGDT. The Figure below highlights the major mechanisms in sepsis, and how fluid responsiveness is assessed with CVP and inferior vena cava (IVC) measurement. Areas of intersection (A, B, C) indicate the variance shared by these factors and the degree to which CVP or IVC measurement is able to predict fluid responsiveness.

Q1.a Several articles have assessed probe positioning in relation to IVC measurement and its correlation with CVP. Describe potential patient-specific or operator-specific issues that might confound the accurate measurement of the inferior vena cava in septic patients.

Patient-specific factors that may influence IVC measurement include body mass index, presence of pulmonary hypertension or tricuspid valve disease, pulse, respiratory rate and tidal volume, chest versus abdominal breathing, and whether the patient is experiencing positive-pressure ventilation. In a morbidly obese patient, IVC measurement may be difficult to obtain because of the depth needed to visualize the IVC. Previous studies have reported that IVC measurements are not possible in 10% to 15% of patients mostly because of obesity, excessive intrathoracic air, or large amounts of intra-abdominal bowel gas. The amount of compression needed to obtain an adequate view may also either compress the anterior-posterior dimensions of the IVC or dampen IVC variability with respiration. Patients who have pulmonary hypertension or tricuspid regurgitation may have less variability in IVC collapse or increased IVC diameter because of increased right-sided atrial pressures. This is also true of patients who are tachycardic. IVC respiratory variation may also decrease with lower tidal volumes because there is reduced negative intrathoracic pressure and therefore decreased venous return. Kimura et al demonstrated that both accessory muscle use and mechanical ventilation are associated with increased venous return and more respiratory variation than normal diaphragmatic or abdominal breathing. This “chest breathing” allows increased negative intrathoracic pressure while minimizing any increase in intraabdominal pressure associated with inferior diaphragm excursion.

Nagdev et al reported that in their study of patients requiring invasive hemodynamic measurement for hypotension, there was a greater percentage of intubated patients in the higher CVP category (26% versus 4%). The authors detailed the different techniques used in intubated and spontaneously breathing patients, saying,
“Measurements were taken during a normal respiratory cycle for patients who were not intubated. For intubated patients, inspiratory inferior vena cava was the maximal diameter during forced inspiration, whereas expiratory inferior vena cava was the minimal diameter at the end of forced inspiration.” There was no mention of how these technical differences might affect IVC measurement, especially in critically ill patients who may be taking variable tidal-volume breaths because of tachypnea or decreased sensorium. In spontaneous breathing patients, Kircher et al8 showed that the observed IVC diameter decrease (defined as expiratory diameter–inspiratory diameter) correlated well with right-sided atrial pressure. However, this index measurement is no longer valid in patients undergoing positive-pressure ventilation because the IVC diameter is maximal at inspiration and minimal at expiration. Furthermore, Barbier et al9 reported “for a given plasma volume in a given patient, the increase in tidal volume and the application of PEEP [positive end-expiratory pressure] can alter IVC diameter.”10 Their article also noted that IVC diameter depended partly on intra-abdominal pressures, and they hypothesized, according to their data, that the accuracy of IVC diameter measurements would not accurately predict fluid responsiveness in individuals with high intra-abdominal pressures. Additional high-quality studies are needed to investigate how positive-pressure ventilation affects point-of-care ultrasonographic IVC measurements accuracy to estimate low CVP status and to measure response to volume therapy.

Operator-specific factors include the segment of IVC being imaged, as well as the level of training in IVC measurement. Saul et al10 showed that the diaphragm/coronal view of the IVC (with probe placed on the midaxillary line) had the greatest interrater reliability among emergency physicians.10 Another study compared 3 sites of IVC measurement, observing that among healthy volunteers, respiratory variation was most similar at the hepatic vein inlet and left renal vein inlet and least similar at the right atrial junction.11 Both studies have limitations, including the use of healthy volunteers and the use of a small number (ie, 1 to 3) of ultrasonograph–experienced emergency physicians to perform all sonograms. De Lorenzo et al12 investigated the ease of image acquisition and found that adequate subxiphoid/hepatic vein inlet views were acquired by novice sonographers in 90% of patients. Novice sonographers had much lower success rates for other views, with 43% for suprahepatic and 68% for midabdominal views.12

In a recent meta-analysis of 5 emergency department (ED)-based studies of hypotensive patients, end expiratory IVC diameter was 6 mm smaller than in normotensive controls.13 The studies were limited, however, in that the location of IVC measurement varied and neither caval index nor respiratory variation was reported. Weekes et al,14 in a panel design in which serial IVC and left ventricular systolic function estimations were made between fluid boluses, showed good correlation between a simplified scale and quantitative measures. There was excellent interrater reliability of IVC assessment between study physicians and a blinded reviewer. However, all sonographers were emergency ultrasonography faculty. An unknown number of patients also were not included in the study because of inadequate views of IVC or left ventricular function.

Haydar et al3 assessed IVC diameter and variability and left ventricular systolic function. The study did not assess for operator- or patient-specific factors that could influence IVC measurement. No data were reported on body mass index, presence of right-sided heart dysfunction, or presence of pulmonary disease. No patients were receiving positive-pressure ventilation at enrollment, so the effect of abdominal versus chest breathing on IVC measurement was not assessed. Regarding operator factors, only 1 view was obtained of the IVC, so there was no comparison with midaxillary or left renal vein segment views of the IVC. Because pulse rate can affect IVC measurement, it was expected that IVC measurements on tachycardic patients were perceived to be less useful.

Q1.b Discuss the evidence for using bedside ultrasonographic measurements of IVC variability to estimate fluid responsiveness. Are ultrasonographic measurements a more accurate measure of fluid responsiveness than CVP, urine output, jugular venous distention, or continuous pulse rate and blood pressure monitoring? What is the criterion standard for measuring the success of fluid resuscitation in patients with septic shock?

Fluid responsiveness, defined as an increase in cardiac output or stroke volume by more than 10% to 15% after a fluid challenge, is an important concept in the resuscitation of patients with sepsis. From a physiologic standpoint, the rapid increase in cardiac filling leads to an increase in myocardial stretch and contractility. The patient’s Starling curve will shift rightward, with a resultant increase in cardiac output until, after successive fluid boluses, there is no further cardiac output increase and the patient is preload optimized.15 In septic shock patients, fluid responsiveness is confounded by the high prevalence of vasodilatation, azotemia, and myocardial depression. Fluid challenges in these patients may lead to increased pulmonary edema and hypoxemia, worsening cardiac dysfunction, and, consequently, prolonged mechanical ventilation. This motivated the search for predictors of fluid responsiveness before fluid administration, a search that has yielded several competing options but no criterion standard.16

In terms of validity, CVP has been traditionally used as a proxy for preload and fluid responsiveness, especially after significant mortality reductions were observed with the advent of EGDT. Studies that assess IVC measurement as a predictor of fluid responsiveness have used CVP as the comparison standard. Wallace et al13 demonstrated in healthy volunteers that IVC measurement in both suprahepatic and infrahepatic views were equivalent in terms of correlation with CVP. Nagdev et al17 reported that an IVC variation greater than 50% corresponded to a CVP of 8 mm Hg or less in an ED patient population. In hypotensive patients, Weekes et al14 demonstrated fair correlation with the infrahepatic view for IVC and response to fluid challenge.

The Haydar et al3 investigation of IVC measurement’s effect on physician management of septic patients is based on the
The aforementioned studies that compared IVC measures to invasive CVP measures. Like several of these studies, Haydar et al. did not directly compare IVC measurement with CVP to determine effect on clinically important outcomes (eg, duration of vasoactive medication use, intensive care unit [ICU] stay, ventilator-free days, and mortality).

Since Marik et al.17 strongly questioned the utility of CVP in assessing fluid responsiveness in their recent meta-analysis, there has been substantial discussion and research to evaluate other modalities for assessing it, including pulse contour analyses and bioelectance, along with low-tech modalities such as the passive leg raise or a small monitored fluid challenge of approximately 200 to 300 mL. Each purports to predict fluid responsiveness much more accurately than CVP. However, 2 important points need to be made in regard to CVP measurement and fluid responsiveness. First, serial measurements with or without actual fluid challenge will be a much more accurate gauge of fluid responsiveness than a single measurement. A single value taken from these new modalities will likely have the same limitations as a single CVP measurement. Therefore, a more nuanced interpretation of the meta-analysis by Marik et al.17 and the studies included, would be that although a single CVP measurement alone cannot predict fluid responsiveness, serial CVP measurements may be able to do so. This is borne out by several studies—including those previously mentioned in this article—that both CVP and IVC do increase with fluid bolus. Magder et al.18 were able to show that CVP variation, much like IVC variation, with respiration predicted fluid responsiveness. The second important point is that fluid responsiveness is determined not only by left ventricular filling but also by cardiac function, including right ventricular filling. It is commonly known that left ventricular dysfunction occurs in at least 50% of patients with severe sepsis, so although these patients may be intravascularly depleted, they may not be fluid responsive because of cardiac dysfunction. Because CVP is equivalent to right-sided atrial pressures, a fluid challenge must overcome CVP or right-sided atrial pressure to fill the right ventricle to affect its Starling curve and increase right-sided output. Without an increase in the right ventricle (RV) output, there is no increase in left ventricle (LV) filling to increase left-sided or cardiac output. Because the vena cava drains into the right atrium, IVC measurements are prone to the same pitfalls as CVP measurements, in addition to those highlighted above. Key advantages for IVC measurement are that it is noninvasive, is less labor intensive, and can be measured serially.

Data suggest that serial IVC measurements may be a better indicator of fluid responsiveness than other modalities for measuring it. Machare-Delgado et al.19 compared IVC change to pulse contour analysis in an ICU population after a 500-mL bolus and demonstrated much better correlation between stroke volume improvement and IVC change than with stroke volume variation in pulse contour analysis. Other promising echocardiographic methods include measuring aortic blood flow through a 5-chamber apical view with the use of pulse wave Doppler.20 This method allows visualization of respiratory variation in the velocity time integral, which can be multiplied by left ventricular outflow tract cross-sectional area to calculate stroke volume. A change in approximately 20% in velocity time integral is highly predictive of fluid responsiveness.21

Q1.c Only 54% of patients in the septic shock group had a CVP below 10 mm Hg compared with more than 60% in the severe sepsis and sepsis groups. The measured IVC diameter variation exceeded 50% in half of the septic shock patients, 56% in the sepsis group, and more than 70% in the severe sepsis group. Do these trends make sense according to the clinical spectrum of sepsis? What threats to the reliability and validity of bedside ultrasonography are suggested by this discrepancy? What are possible explanations?

The clinical spectrum of sepsis would suggest that a larger proportion of patients in septic shock should have CVPs lower than 8 mm Hg compared with those in sepsis or severe sepsis. However, the sepsis literature shows that only about 50% of patients are actually fluid responsive. This is because vasodilatation is only one manifestation of sepsis, and other factors such as endothelial dysfunction and myocardial suppression may contribute to the hemodynamic derangements that occur. This may be one reason why CVP trends did not correspond with the severity of sepsis in the study population and why IVC diameter may not as well. Another possible explanation is that patients were enrolled after they were identified as possibly septic while in triage, and a sepsis bundle—which included a fluid bolus—was initiated. Although there is no specific mention about the mean time elapsed between study enrollment and IVC measurement, study subjects had a mean time elapsed of 138 minutes between ED arrival and IVC ultrasonography. Thus, patients may have already responded to the fluid bolus before IVC ultrasonography, and those who remained hypotensive may not have been fluid responsive.

ANSWER 2

Q2. The point-of-care ultrasonography in this study differs significantly from how it is used in most EDs. In other words, ultrasounds were performed over 2 hours after patient arrival in the ED, and treating physicians—who did not perform the ultrasounds—did not have access to the images until 20 minutes after they were performed.

Q2.a If many patients had treatment initiated as part of their sepsis bundle, how would this have influenced the ultrasonographic measures on the outcomes of interest?

There would be a temporal bias because the volume status of patients would have changed between the ultrasonographic measurement, fluid administration, and clinician integration of IVC data. This may have caused some confusion among the clinicians because the IVC measurements would no longer be contemporaneous with the patient’s current volume status. Therefore, a decrease in confidence may result, or this may lead to overresuscitation of patients. Clinical ultrasonographic studies should be designed to reproduce the “real-time” availability of point-of-care ultrasonography studies. This
immediate data awareness is crucial to accurately understand ultrasonography’s effect on managing critically ill patients.

Q2.b Most emergency department–based ultrasonographic research studies are performed by highly experienced physicians who have completed an ultrasonographic fellowship or equivalent training. Discuss the limitations associated with generalizing these study results to emergency physicians without such advanced ultrasonographic training.

This is a common limitation in ultrasonographic studies because the test characteristics published are often superior to those of similar studies that use nonultrasonographic fellowship-trained physicians. Generalizing these study results to a typical ED should be done with caution for this reason. Ways to improve the generalizability of these studies is to explicitly state the number and training of personnel performing the ultrasonography, as well as to incorporate a larger number of non–fellowship-trained emergency physicians. Haydar et al used study physicians who received 3 hours of didactics and had to perform 25 cardiac and IVC ultrasonographic procedures before participation in the study. Although the limitations section suggests that the focused training of the study “mirrors the reality of many” ED-based ultrasonographic studies, there was no mention of the level of training or the number of study physicians who participated.

Q2.c Might this study’s results represent the Hawthorne effect or expectation bias? Did this investigation truly measure ultrasonography’s utility or the effect of an explicit cueing of the sepsis management guidelines?

It is unclear how much influence participating in this study had on the practice of the study clinicians compared with the true influence of the IVC data on clinical management. Even in the original EGDT literature by Rivers et al, there was a high likelihood that the study itself cued physicians to be more aggressive in their management of sepsis because the control group mortality was much lower than the baseline sepsis mortality. The study reports a small to medium improvement in physician certainty about the cause of clinical findings overall, with a greater change in physician confidence when managing patients without tachycardia.

Q2.d How would you design a study that fully incorporates the use of ultrasonography in managing sepsis patients? What is the unit of analysis for an emergency department–based ultrasonographic research investigation? How does one calculate the required sample size for such an investigation?

Haydar et al were able to show that IVC and cardiac ultrasonography was associated with increased physician certainty in the cause of vital sign abnormalities and their treatment plans in septic patients. Study physicians also were given essentially the minimum level of IVC and cardiac training and practice as required by most emergency ultrasonography training programs. However, a major limitation of the study is that clinically relevant outcomes were not measured.

Only a randomized controlled trial that uses IVC measurement in the study arm and conventional or invasive estimation of fluid responsiveness in the control arm, with primarily clinical outcomes such as mortality, would help to answer the question of the clinical utility of IVC variation in septic patients. As mortality from sepsis continues to decline to below 20% to 25%, any sample size estimation would have to account for likely further reductions in baseline mortality if the study is to be adequately powered. The investigation would also have to address the aforementioned limitations, including requiring a large number of physician sonographers with diverse levels of ultrasonographic expertise, the need for real-time ultrasonographic data availability to the treating clinicians, and a proper randomization scheme to account for many of the patient-specific and operator-specific confounders mentioned in Answer 1a.

Determining unit of analysis is a crucial step in study design, but it is often overlooked. In ultrasonographic studies, the patient is often not the study subject. The physician sonographer is the primary focus of the study. Because the physicians under study may not perform identically, data should be presented for each physician because the physician is the unit of analysis. A common mistake is to ignore the physician and pretend the patient is the unit of analysis, which is equivalent to assuming that all physician sonographers perform equally, an assumption that is quite likely untrue.

The sample size calculation for an ultrasonographic study must use methods that account for the nesting of patients within sonographers. Such calculations require the investigator to make assumptions about the variance among ultrasonographers as it compares to the variance among patients (within ultrasonographers). This ratio can be expressed as the intraclass correlation coefficient. A simple but useful way to think about this is that if the sonographers perform identically, then the effective N of the study is the number of patients, but if each sonographer is very consistent (all of the patients scanned by one physician receive equivalent quality scans/decisions) but the sonographers differ greatly from one another in performance, then the effective N is the number of sonographers. Typically, the effective N is somewhere between these 2 extremes. The lower the variance between sonographers, the closer the N gets to the number of patients.

ANSWER 3

Q3. The point-of-care assessment of cardiac contractility is a powerful example of the role of ultrasonography in clinical care.

Q3.a What sonographic measures of cardiac contractility provide the best estimation of ejection fraction?

Ejection fraction is widely used by cardiologists and noncardiologists as a proxy for left ventricular function and contractility. Several imaging modalities are used to assess left ventricular function, including magnetic resonance imaging (MRI), multislice computed tomography (CT), and radionuclide ventriculography. In a recent meta-analysis by Asfeng et al, accuracy of transthoracic echocardiography was similar to that of MRI and multislice CT. The main advantages of transthoracic echocardiography are that it has been well studied, is widely available, is low cost, and does not expose patients to radiation. The primary disadvantages include lower
interrater reliably and operator dependence. Qualitative “eyeball” methods for left ventricular ejection fraction estimation also have been shown to be just as accurate as quantitative methods such as the Simpson method.²³

Q.3.b How well do noncardiologists fare in estimating cardiac contractility?

Several studies in the critical care and emergency medicine literature suggest that noncardiologists’ interpretation of cardiac contractility is very reliable. In a study of 50 hypotensive patients presenting to the ED, emergency physicians with 6 hours of didactics and 10 hours of hands-on training were able to estimate ejection fraction with accuracy similar to that of the study cardiologist. Specifically, the Pearson’s correlation coefficient was 0.86, which compares favorably with that of studies that examine correlation of ejection fraction (EF) estimations between cardiologists (R=0.77 to 0.90).²⁴ Weekes et al²⁴ were able to demonstrate in a recent study that emergency physicians’ gross estimation of cardiac contractility and IVC variation had a high correlation with traditional quantitative measures, such as fractional shortening and the caval index. Secco et al²⁵ demonstrated that emergency residents could use mitral valve E-point septal separation on M-mode to accurately estimate EF compared with 2 cardiologists. Randazzo et al²⁶ showed that emergency physicians given about 5 hours of didactics and hands-on practice were able to accurately assess gross left ventricular function (LVF) with 86% agreement compared with a formal transthoracic echocardiography.

Q.3.c How does the clinician apply various findings on ultrasonography, such as IVC variation greater than 50% and normal ejection fraction or IVC variation less than 50% and low ejection fraction, in septic shock patients?

Utility in point-of-care ultrasonography is maximized when the hypotensive patient is managed. From the Undifferentiated Hypotensive Patient to Rapid Ultrasound in Shock protocols, the concept is to elucidate the shock cause, differentiating patients with obstructive, hypovolemic, and cardiogenic causes.²⁷,²⁸ Hypotensive patients who have less than 50% IVC variation and depressed cardiac contractility are exhibiting a predominantly cardiogenic shock presentation. These patients would likely benefit more from pressors than additional volume resuscitation. Hypotensive patients who have greater than 50% IVC variation or a flat IVC, along with a normal or hyperdynamic heart, are exhibiting a hypovolemic or distributive shock presentation. Patients with hypovolemic shock should be adequately fluid resuscitated before initiation of pressors. Many protocols also evaluate for obstructive shock physiology, whether from tension pneumothorax or cardiac tamponade.

Figure 2 in the article by Haydar et al³ shows the change in certainty of each respondent from pre-point-of-care ultrasonography findings to post-point-of-care ultrasonography. By showing the data at the level of the unit of analysis (see question 2d), the graphic provides added insight into the utility of the sonographic information because we can see the proportion of respondents who had increased and decreased certainty. When the treating physician was 90% certain about the cause of abnormal vital signs, the presentation of sonographic data often decreased certainty (Figure 2A in the article by Haydar et al³), whereas the majority of physicians who had 50% to 75% certainty had an increase in certainty after obtaining ultrasonographic data. This would not be apparent to readers if only the box plots had been reported. Take, for instance, 2 additional hypothetical studies with the same box plots showing an overall 15% increase in certainty from 75% to 90%. In study A, 50% of patients did not have any change but 50% did have a large increase in certainty. In study B, all patients had an increase in certainty, but only by 15%. With the additional information on the distribution of change, clinicians may surmise that point-of-care ultrasonography in study A is essentially worthless in half of patients, whereas in study B, it is helpful to a limited extent in all patients. It is only by studying graphics of this kind that investigators can begin to understand what is actually taking place.

Q.4.b How do you interpret Figure 3? When pre-point-of-care ultrasonography certainty is high for patients, are they expected to change more or less after point-of-care ultrasonography than patients with low pre-point-of-care ultrasonography certainty? Identify patients for whom point-of-care ultrasonography decreased the physicians’ certainty. Do they share any clinical characteristics?

The 5 graphs in Figure 3 are scatterplots that show pre-point-of-care ultrasonography certainty on the x axis and change in certainty post-point-of-care ultrasonography. These are unique in that they are organized into subgroups of patients according to volume status and left ventricular function. In patients for whom point-of-care ultrasonography shows normal cardiac contractility with low volume status, there is increased certainty when the treating physicians showed uncertainty pre-point-of-care ultrasonography, likely because they were thinking the patient was hypovolemic and ultrasonography confirmed that belief. However, Figure 3A shows a substantial number of patients for whom the finding of normal contractility and normal IVC volume decreased the physician’s certainty. Presumably, they were suspecting hypovolemia as the cause of vital signs abnormalities, and they now had to find an alternate cause or discount the validity of the point-of-care ultrasonography results. These graphics allow the reader to examine which findings tend to increase (or decrease) physician certainty.

Q.4.c What are the advantages and disadvantages to analyzing effect size, such as through Cohen’s d statistic, as opposed to traditional significance testing, such as a paired samples t test?

Cohen’s d compares the difference in means to the SD standard deviation, thereby referencing the effect size (the difference) to the variance (as expressed by the standard deviation). Measures of this
kind are particularly useful when one is trying to compare different outcome measures related to a single phenomenon. For example, similar educational interventions may have been tested in a variety of settings, each with some type of before and after test as the outcome measure. If each experiment used a different test, however, it would be meaningless to compare the differences because the scales would be different, as would the variance. For example, if one system used a test scored 0 to 100 with an standard deviation of 20 and another used one with a test scored 200 to 800 with an standard deviation of 60, comparing differences—say, 7 points on the first test and 50 on the second—would be meaningless, but comparing effect sizes such as Cohen’s $d$ would be more meaningful because the scores would have been standardized by the standard deviation.

The other advantage of statistics such as Cohen’s $d$ is that they emphasize description above hypothesis testing. In noninferential statistics or descriptive statistics, no attempt is made to determine the probability of type I or II error (ie, whether results are consistent with chance [random error] or whether they are unlikely to be due to chance, and therefore they represent a true effect or nonrandom error [bias]). Noninferential statistics are particularly appropriate in observational studies.

Section editors: Tyler W. Barrett, MD, MScI; David L. Schriger, MD, MPH

Author affiliations: Olive View–University of California, Los Angeles Medical Center, Sylmar, CA

REFERENCES


