Physicians’ Clinical Behavior During Fluid Evaluation Encounters

OBJECTIVES: We sought to identify factors affecting physicians’ cognition and clinical behavior when evaluating patients that may need fluid therapy.

BACKGROUND: Proponents of dynamic fluid responsiveness testing advocate measuring cardiac output or stroke volume after a maneuver to prove that further fluids will increase cardiac output. However, surveys suggest that fluid therapy in clinical practice is often given without prior responsiveness testing.

DESIGN: Thematic analysis of face-to-face structured interviews.

SETTING: ICUs and medical-surgical wards in acute care hospitals.

SUBJECTS: Intensivists and hospitalist physicians.

INTERVENTIONS: None.

MEASUREMENTS AND MAIN RESULTS: We conducted 43 interviews with experienced physicians in 19 hospitals. Hospitalized patients with hypotension, tachycardia, oliguria, or elevated serum lactate are commonly seen by physicians who weigh the risks and benefits of more fluid therapy. Encounters are often with unfamiliar patients and evaluation and decisions are completed quickly without involving other physicians. Dynamic testing for fluid responsiveness is used much less often than static methods and fluid boluses are often ordered with no testing at all. This approach is rationalized by factors that discourage dynamic testing: unavailability of equipment, time to obtain test results, or lack of expertise in obtaining valid data. Two mental calculations are particularly influential: physicians’ estimate of the base rate of fluid responsiveness (determined by physical examination, chart review, and previous responses to fluid boluses) and physicians’ perception of patient harm if 500 or 1,000 mL fluid boluses are ordered. When the perception of harm is low, physicians use heuristics that rationalize skipping dynamic testing.

LIMITATIONS: Geographic limitation to hospitals in Minnesota, United States.

CONCLUSIONS: If dynamic responsiveness testing is to be used more often in routine clinical practice, physicians must be more convinced of the benefits of dynamic testing, that they can obtain valid results quickly and believe that even small fluid boluses harm their patients.

KEY WORDS: fluid therapy; resuscitation; sepsis; shock; ultrasound

IV crystalloid fluids are indicated in hypovolemia, severe sepsis, and some other shock states. Recommendations surrounding fluid therapy (FT) emphasize frequent reassessment in the early stages of septic shock (1). The condition(s) that prompted initial FT (e.g., arterial hypotension, tachycardia, rising creatinine, elevated lactic acid) often persist or reappear hours or days later. This leads to therapeutic uncertainty: was the initial fluid order physiologically sound but now needs redosing? Or should alternatives such as catecholamines be used? The Surviving Sepsis Campaign recommends (1–3) evaluation of fluid responsiveness (FR) after initial resuscitation using dynamic measures. However,
there are no guidelines on the timing of a fluid needs evaluation (FNE); which dynamic measure to use, or whether these strategies apply to nonsepsis conditions.

As a diagnostic test, demonstration of FR, defined as a change in cardiac output (CO) or stroke volume (SV) after volume loading, is problematic. Even when the test is positive, FT does not always improve the patient's vital signs or laboratory abnormality (4). Meanwhile, clinicians are cautioned to avoid volume overload and consider fluid removal or de-resuscitation (5).

Numerous technologies and maneuvers have been tested to improve clinicians’ decision-making process during what we term a FNE episode (6). Most of these studies are from academic hospitals describing a measurement from a device before and after an intervention (7). Others describe the short-term, and occasionally long-term, effects on the patient’s course. The promotion of commercially available devices has been coincident with publications demonstrating the association between iatrogenic hypervolemia and adverse outcomes (8–13).

The Fluid challenges in intensive care study study (14) described physicians’ practice of fluid challenges (FCs) in 2,213 ICU patients. It showed significant variability in fluid type and volume ordered and that dynamic maneuvers to determine FR were used at a low frequency (22%). More importantly, a positive, negative, or ambiguous response to a FC did not change the amount of fluids subsequently prescribed. In other words, physicians’ fluid-ordering behavior was paradoxically unaffected by the test result (15). To better understand clinical behavior around FNE and explain the paradox in the FENICE results, we designed a qualitative study to understand the extent to which physicians’ beliefs and behaviors during a FNE are related to their factual knowledge, training experience and access to devices.

### METHODS

A structured interview guide was designed (M.K.H.S., C.W.) and tested by completing nine pilot interviews with physicians (Supplemental Digital Appendix Section 1, http://links.lww.com/CCX/B208). We then conducted face-to-face interviews with hospital-based physicians using the modified guide. The sampling frame included at least one physician at hospitals within 150 miles of Minneapolis, Minnesota including the 10 largest hospitals in the state and smaller hospitals with at least a six-bed ICU. We included physicians with the goal of interviewing a diverse group of physicians who regularly do FNE for hospitalized patients both in the ICU and wards. We planned a sample size of 40 interviews in 20 hospitals. We collaborated with the Minnesota Hospital Association to approach physicians for participation. Initial contact was usually via email followed by telephone to schedule an interview at the physician’s workplace. We created a list of FNE methods or devices to assist physician’s recall if needed (Supplemental Digital Appendix Section 2, http://links.lww.com/CCX/B208). We designed the interview to characterize physicians’ actual practice during a FNE and their self-reported clinical reasoning. We encouraged discussion of FNE not just for obvious shock states but for other scenarios where fluids may be beneficial such as oliguria, mild arterial hypotension, and tachycardia or elevated lactate. To minimize response bias, participants were informed that their responses were anonymous and confidential. Although we used an interview guide, physicians were encouraged to speak at length about FT issues important to their practice. Part of the interview guide asked participants how their practice group decided to purchase...
devices used in FNE episodes, but those comments are not the subject of this report.

The author (M.K.H.S.) completed all interviews and took contemporaneous field notes. Interviews were recorded with an iPhone and microphone. The TranscribeMe service transcribed the interviews and the authors (M.K.H.S., C.W.) made final edits against the recording. We used NVIVO 12 Pro (Lumivero, Denver, CO) to assist in content analysis.

The authors (M.K.H.S., C.W.) read the 43 transcripts and independently derived themes they felt were mentioned frequently by participants or were relevant to the FENICE paradox: how do physicians approach an FNE, decide to assess for FR and then order therapies? Preliminary themes were compared in research meetings and consolidated into nine themes. To improve thematic credibility, we asked two physicians not involved in the interviews (K.P., J.P.) to independently confirm the presence of the themes. They were given anonymized transcripts and rated each of the themes across all interviews as “prominently present,” “moderately present,” or “minimally present.” We took an extra step by adding two decoy themes (plausible themes about FNE but not substantially present in the transcript) to the nine themes without telling the raters which ones they were. We then calculated the kappa statistic for agreement in classifying themes between the blinded raters. After unblinding, the raters re-read the transcripts to offer their opinions on whether additional themes were present.

The University of Minnesota Institutional Review Board waived oversight of the study as it did not meet their definition of human subject research.

RESULTS

We conducted 43 interviews (24 intensivists and 19 hospitalists) over 7 months in 19 hospitals in 2019. Almost all who agreed to participate completed their interview, and all but one was conducted in-person. Participants (Table 1) were experienced (average of 10 yr of practice) in hospital-based medicine and worked in hospitals with substantial ICU capacity (average 22 beds). Most hospitals had open units indicating that a hospitalist could manage or co-manage an ICU patient.

The unweighted kappa statistic for rater agreement was 0.59 (95% CI, 0.19–0.99) for two independent raters, three levels of classification, and 11 themes. Both raters classified the two decoy themes as “minimally present” and, after unblinding, concluded no other themes were present and there was saturation of concepts in the sample. As expected, there was overlap between themes and for this report; we consolidated the analysis from nine to eight themes accompanied by representative quotations (Supplemental Digital Appendix Section 3, http://links.lww.com/CCX/B208).

The mean interview duration was 30.2 minutes. Total interview time was 21.7 hours and the transcript was 376 pages with 184,302 words (Supplemental Digital Appendix Section 4, http://links.lww.com/CCX/B208).

Themes

**Personal Process of Fluid Needs Evaluation and the Pre-Test Probability.** Physicians recognized the lack of guidelines for FT after initial resuscitation. They used their own clinical process for FNE, which they had developed over the course of their training and practice. Most had not modified their FNE style since their training or early practice years although they acknowledged contemporary issues such as the harms of severe hypervolemia.

Their FNE incorporated history of present illness, comorbid conditions like heart or renal failure, estimation of fluid loss, and the volume of recent fluids

---

**TABLE 1.**

Demographics of Participants and Hospitals

<table>
<thead>
<tr>
<th></th>
<th>Mean (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interviews</td>
<td>43</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>43 (37–47)</td>
</tr>
<tr>
<td>Years of practice post-training</td>
<td>10.5 (5–15)</td>
</tr>
<tr>
<td>Gender</td>
<td>27 Men, 16 Women</td>
</tr>
<tr>
<td>Participants’ proportion of work in clinical settings</td>
<td>79% (70–100%)</td>
</tr>
<tr>
<td>Participants’ proportion of their clinical work that is in the hospital</td>
<td>87% (80–100%)</td>
</tr>
<tr>
<td>Hospitalist/intensivist group size</td>
<td>21 Physicians (8–30)</td>
</tr>
<tr>
<td>ICU bed capacity at their main facility</td>
<td>22 Beds (14–20)</td>
</tr>
<tr>
<td>Open ICU style units</td>
<td>90%</td>
</tr>
</tbody>
</table>

IQR = interquartile range.
given with special attention paid to the patient’s response to prior FCs. No single data point was determinative. Sometimes the FNE occurred in patients they knew well but frequently it was a cross-cover situation that required medical record review. If that review created enough certainty, then FNE decisions were made without examining the patient. Clinicians were strongly influenced by prior results of FNE episodes and heuristics such as “if it worked before, I’ll do it again” were used. Physicians’ pretest probability that FT would improve the condition they were addressing (e.g., hypotension, oliguria, elevated lactate) influenced their interpretation of FNE test results. Physicians’ clinical judgment could override results from any test.

When the pretest probability was ambiguous, the FNE process led to a FC to detect improving vital signs or laboratory tests (e.g., lactate). Even though most clinicians were aware that dynamic monitoring meant a change in a physiologic measure such as SV or CO, their priority was to correct the abnormality they were called to evaluate which was almost never a SV or CO measurement. If they strongly believed that a volume challenge, (by IV fluids or leg raising) would result in a positive response in heart rate, blood pressure, or urinary output, they felt justified in skipping the intermediate physiology step, which required a device, expertise, and their clinical time.

**Most Used Modality/Process.** To supplement the history and physical examination, physicians employed other modalities. Portable ultrasound was widely available and 67% of participants obtained and interpreted their own images. Physicians preferred ultrasound because of its availability, familiarity (they have used it for other reasons such as line placement), its noninvasive nature and that ultrasound gives visual information on pathophysiology such as myocardial contractility. The FloTrac system (pulse wave contour analysis device) was less available but even when it was, physicians were reluctant to place an arterial line just to use this device for FR testing. Central venous pressure was not used in decision-making unless at very high or low values. A passive leg raise (PLR) test with CO/SV monitoring was rarely used. There was one hospital where a noninvasive cardiac output monitoring (NICOM) device (also termed Cheetah devices by participants) operated by nurses was available if requested. Another physician did PLR routinely during a FNE but looked for changes in vital signs, not device-measured CO. Clinicians reported that the PLR often required a helper who was not always available, and the full test took too much time and was sometimes not feasible because of body habitus or patient’s postoperative status.

In non-ICU settings, FNE was mostly chart review supplemented by a bedside physical examination without the use of a device or dynamic measures. Some participants reported making FNE decisions based solely on chart review, explaining it provides them with enough information to decide.

**FNE: A Rapid, Independently Made Decision.** FNE episodes were frequent and familiar work for hospital-based physicians and they consider a FNE within their scope of practice expertise. FNE were almost always unscheduled events in response to an abnormality. Clinicians were called by nurses about abnormal vital signs or oliguria or physicians noticed that a follow-up laboratory value was not improving as expected. The turnaround time from awareness of the abnormality to the FNE was minutes to 2 hours depending on the perceived seriousness and the physician’s workload. Although FNE decisions were acknowledged as “complex,” “difficult,” or “challenging,” in most cases, the physician did the entire FNE and intervention independently.

**Crude Measurements Not Precise Calculations.** Although physicians knew what objective changes in parameters (e.g., inferior vena cava size and collapsibility or arterial pulse pressure variation) were predictive of FR, few exactly measured device output. “Eyeballing it” was usually deemed adequate.

**Physicians’ Desired FNE Process Is Often Not Possible.** Physicians were sometimes unable to deploy their preferred FNE process due to constraining factors: too busy with other duties, lack of immediately available technology, being off-ward or managing the patient by telephone. Some had a preferred technology that was unavailable at their hospital. Physicians acknowledged that their approach to FNE was often based on training and skills acquired years ago. Some had read literature about new technologies and expressed a desire to obtain new skills. But only a few did acquire new skills after their formal training had ended.

**Perceived Risk of Under-Resuscitation Versus Over-Resuscitation.** At the time of a FNE, physicians were more concerned with the risk of giving insufficient fluids compared with too much fluid although...
the presence of cardiac and renal disease made clinicians cautious. The perceived accuracy of the risk-benefit balance changed over the hospital course, as more data were available. For instance, significant weight gain, positive fluid balance, or knowledge of prior FCs influenced decision-making. Adverse effects of fluid overload were acknowledged but felt to be benign or reversible. Physicians believed that fluid overload could be detected early by lung rales, lung ultrasound, chest radiograph, or increased oxygen needs. Then, fluids would be stopped and diuretics given to reverse the process. Individual boluses of 500 mL were almost never perceived as harmful. Fluid overload harm was conceptualized as a hypothetical, future-oriented problem that was less important than fixing the current issue in front of them.

**FNE in Non-ICU Settings/Vasopressors.** In non-ICU settings, vasopressors cannot be easily started for hypotension. The process to transfer a patient to the ICU and obtain central access involves time-consuming steps. Therefore, physicians were more likely to use fluid boluses to reverse hypotension or oliguria as it is the only readily available option, even if they believed there was a lower probability of favorable response. In other words, if the barriers to ordering vasopressors were high but the perceived harm of ordering fluid boluses and ease of ordering fluids were both low, then patients were given more fluids.

**Properties of an Ideal FNE Tool/Protocol.** Physicians identified deficiencies in current FNE practice and methods. When a FNE requires presence at the bedside for more than a brief period, they will often choose other methods or will order a fluid bolus as a therapeutic test. If the patient improves with fluid, the test was positive. Devices that produce images (e.g., ultrasound or variation in arterial waveforms) are considered more influential than devices that produce a number (e.g., FloTrac, NICOM). However, most ultrasound users estimated that they could not obtain good images in a third to a half of their attempts. An ideal FNE tool or protocol would be credible, feasible, and efficient with device images that can be easily seen and understood.

**DISCUSSION**

We asked experienced physicians to describe their clinical reasoning and behavior during a FNE, a distinct encounter that acute care physicians experience routinely: does this patient need more IV fluids? We characterized clinicians’ reasoning and their behavior in their own words in both ICU and non-ICU settings and in academic and community hospitals.

Current research emphasizes a fluid strategy that gives boluses only to patients who demonstrate increases in cardiac parameters after a FC. Presumably, this will improve oxygen delivery to fluid responsive patients while reducing volume overload-related harm to nonresponsive patients. Our results suggest that this reasoning is not how most experienced clinicians approach a FNE.

Only a minority of physicians in our study routinely employ PLR or use commercial devices for dynamic testing. Similarly low rates were reported in a study of physician practice in 34 hospitals (16). Many more obtain static echocardiographic images and interpret them qualitatively (e.g., IVC is “plump,” “collapsing,” or “IV looks underfilled”) but without quantitative before-and-after dynamic imaging. Physicians with recent residency or fellowship training were more likely to have obtained bedside echocardiography skills but routine use of even static echocardiography during a FNE remained low for most. More commonly though, if physicians believed that a FC would be positive, they felt justified in skipping the intermediate physiologic step which required a device, expertise, and their clinical time. When the pretest probability was more ambiguous, the FNE process still usually led to a FC trial with the test result being an improvement in the abnormal vital signs or laboratory tests they were evaluating.

Interestingly, before physicians decide to order the FT, they have already rapidly obtained and interpreted a large amount of clinical data including physical examination findings, weight trends, fluid balance, prior response to FT, laboratories, and presence of relevant comorbid conditions. They use this information to make an informal calculation of the pre-test probability of FR. In short, their cognitive process seems influenced by Bayesian reasoning (although no physician mentioned that term). If they believe that the pretest probability is so high or low that no test result will substantially change the post-test probability, they feel justified in giving the fluid (or withholding it) without obtaining the test or even seeing the patient. If the pre-test probability is intermediate, then the decision to do a test is positively affected (i.e., more likely to perform) by their expectation of a minimal time commitment,
device availability, perceived self-expertise in consistently obtaining test results, and perception of additional fluid causing harm to the patient. The latter is particularly influential. If physicians believe that their test bolus of 500 or 1,000 mL is harmless, then most will proceed with the heuristic of “Don’t know, might help, won’t hurt, do it and see what happens.”

Physicians are aware of medical literature that reports the association of volume overload to worse outcomes and acknowledge that high-volume resuscitations they may have performed years ago are not beneficial. But they do not equate the 500 or 1,000 mL boluses that they order with the grossly edematous patient that may result in the future.

Many participants believe that, given the complexity of FNE, it is unlikely that a single simple FNE method will improve their decision-making in all situations. Since bedside ultrasound is already familiar, greater acceptance of new FNE technology would be more likely if performed with ultrasound rather than stand-alone devices that have only one function or are available only in select hospital settings.

The strength of this study is that we interviewed physicians at a variety of academic and medium-to-large community hospitals who perform FNE regularly in ICUs and general wards. We reoriented physicians away from “quoting the literature” and toward their actual practice. Face-to-face interviews are subject to interviewer bias and subjects’ recollection of their clinical reasoning and behavior may not reflect their actual behavior. All surveys that ask physicians to self-report their practice have that limitation. However, the confidential interview method minimized incentives for participants to be untruthful about their actual practice. Thematic analysis of interview text is inherently qualitative and we have minimized categorical comparisons such as physician specialty, hospital type, or years of experience. Future surveys could provide statistical analyses on such relative proportions and could sample a broader population to confirm the generalizability of our conclusions. For the feasibility of conducting face-to-face interviews, we limited the sample frame to a drivable distance; however, that region includes about 90% of the Minnesota population. While geographic specificity is a limitation because Minnesota’s aggregate health outcomes and per-capita income rank near the top in the United States (17), it is unlikely that the low use of certain technologies is due to insufficient financial resources or training opportunities.

CONCLUSIONS

Three cognitive processes explain clinical behavior around FT and the paradox of the FENICE study: 1) Clinicians’ approach to a fluid needs encounter is oriented to correct hypotension, oliguria, or elevated lactate. 2) When the probability of FR is estimated to be high or low, then clinicians’ reason there is little need to do an extra test or maneuver that is burdensome, unavailable, or that they are not skilled at. 3) Even if the probability is intermediate, because physicians consider fluid boluses benign or easily reversible, they often proceed with the fluid bolus and see if the problem resolves without performing dynamic testing.

If these rationales and behaviors are confirmed in a larger-scale survey, proponents of dynamic fluid management who desire wider adoption have three tasks. First, they can promote FNE methods that are portable ultrasound-based and can be performed by busy clinicians on almost all patients in a few minutes. For instance, multivessel quantitative interrogations such as a Venous Excess Ultrasonography Score scan (18) are unlikely to be performed by busy clinicians. Option ally, have hospitals support technicians to obtain images and record interpretations in a timely fashion, similar to the workflow of ultrasound technicians and radiologists. Second, improve the evidence that moderate volume overload leads to meaningful harm. Then clinicians will be more cautious about ordering therapy they currently consider benign. Physicians may interpret the results of recent fluid trials such as Conservative versus Liberal Approach to Fluid Therapy of Septic Shock in Intensive Care trial and Crystalloid Liberal or Vasopressors Early Resuscitation in Sepsis trial to mean that precision FT early in the hospital course is not essential (19, 20). Third, conduct research that estimates the pre-test probability (Bayesian base rate) for FR in common clinical scenarios. Effective Bayesian decision-making depends on correctly estimating the base rate. For instance, 70% of septic shock patients remain fluid responsive after a 26 mL/kg bolus (21) and 42% were responsive during the first 48 hours in the Fluid Responsiveness Evaluation in Sepsis Hypotension and Shock Trial (7). If clinicians knew that the pre-test probability of FR was in the ambiguous range for certain scenarios, they might be more willing to perform objective testing prior to ordering additional boluses.
ACKNOWLEDGMENTS

We thank the Minnesota Hospital Association assisted the authors in identifying hospitals for subject recruitment.

REFERENCES


