


# The Acute Dialysis Orders Objective Structured Clinical Examination (OSCE)

## Fellow Performance on a Formative Assessment of Acute Kidney Replacement Therapy Competence

Lisa K. Prince, Robert Nee, and Christina M. Yuan  for the Nephrology Education Research and Development Consortium (NERDC)

### Abstract

**Background and objectives** Acute kidney replacement therapy (KRT) prescription is a critical nephrology skill. We administered a formative objective structured clinical examination (OSCE) to nephrology fellows to assess acute KRT medical knowledge, patient care, and systems-based practice competencies.

**Design, setting, participants, & measurements** Prospective cohort study of an educational test using the unified model of construct validity. We tested 117 fellows: 25 (four programs) in 2016 and 92 (15 programs) in 2017; 51 first-year and 66 second-year fellows. Using institutional protocols and order sets, fellows wrote orders and answered open-ended questions on a three-scenario OSCE, previously validated by board-certified, practicing clinical nephrologists. Outcomes were overall and scenario pass percentage and score; percent correctly answering predetermined, evidence-based questions; second-year score correlation with in-training examination score; and satisfaction survey.

**Results** A total of 76% passed scenario 1 (acute continuous RRT): 92% prescribed a  $\geq 20$  ml/kg per hour effluent dose; 63% estimated clearance as effluent volume. Forty-two percent passed scenario 2 (maintenance dialysis initiation); 75% correctly prescribed 3–4 mEq/L K<sup>+</sup> dialysate and 12% identified the two absolute, urgent indications for maintenance dialysis initiation (uremic encephalopathy and pericarditis). Six percent passed scenario 3 (acute life-threatening hyperkalemia); 20% checked for rebound hyperkalemia with two separate blood draws. Eighty-three percent correctly withheld intravenous sodium bicarbonate for acute hyperkalemia in a nonacidotic, volume-overloaded patient on maintenance dialysis, and 32% passed overall. Second-year versus first-year fellow overall score was  $44.4 \pm 4$  versus  $42.7 \pm 5$  (one-tailed  $P=0.02$ ), with 39% versus 24% passing ( $P=0.08$ ). Second-year in-training examination and OSCE scores were not significantly correlated ( $r=0.15$ ;  $P=0.26$ ). Seventy-seven percent of fellows agreed the OSCE was useful in assessing “proficiency in ordering” acute KRT. Limitations include lack of a validated criterion test, and unfamiliarity with open-ended question format.

**Conclusions** The OSCE can provide quantitative data for formative Accreditation Council for Graduate Medical Education competency assessments and identify opportunities for dialysis curriculum development.

CJASN 14: 1346–1354, 2019. doi: <https://doi.org/10.2215/CJN.02900319>

### Introduction

We previously developed an acute dialysis objective structured clinical examination (OSCE) to formatively assess fellow competence in managing three commonly encountered situations requiring acute kidney replacement therapy (KRT): (1) acute continuous RRT (CRRT), (2) maintenance hemodialysis (HD) initiation in moderate uremia, and (3) acute HD for life-threatening hyperkalemia and volume overload in ESKD (1). Examinees use institutional procedures and order sets to write scenario-specific dialysis orders, and then answer related, open-ended questions. Two questions per scenario address evidence-based concepts. The OSCE does not require sophisticated simulation

techniques, takes <2 hours, is easy to administer, and is freely available.

Acute HD and CRRT are critical nephrology skills that are difficult to quantitatively and longitudinally assess in high-stakes summative examinations using multiple choice questions, the format of the nephrology certifying and in-training examinations. The 2018 American Board of Internal Medicine nephrology certification examination blueprint indicates that 11.5% of questions pertain to ESKD (HD, peritoneal dialysis and their complications; home HD; ESKD complications; and dialysis medical director topics), and 4% to acute KRT (2). Thus, few questions on the certifying or in-training examination (which parallels the certifying

Nephrology Service,  
Department of  
Medicine, Walter  
Reed National Military  
Medical Center,  
Bethesda, Maryland

**Correspondence:**  
Dr. Christina M. Yuan,  
Nephrology Service,  
Department of  
Medicine, Walter Reed  
National Military  
Medical Center, 4494  
North Palmer Road,  
Bethesda, MD 20889.  
Email: christina.m.  
yuan.civ@mail.mil

examination) directly assess acute KRT prescribing ability (3). The Accreditation Council for Graduate Medical Education (ACGME) milestones framework requires program directors ensure fellows demonstrate skill in performing acute and maintenance dialysis, a Patient Care sub-competency (4).

We prospectively administered the OSCE to fellows at 15 training programs in 2016 and 2017, determining performance overall, on each scenario, and on clinically relevant, evidence-based questions. We compared first- and second-year fellow scores and evidence-based question performance. We also assessed fellow and program director satisfaction with the OSCE as a formative assessment.

## Materials and Methods

### Test Development and Initial Validation

As previously described, the test assesses medical knowledge, patient care, and systems-based practice competencies in three common, critically necessary, acute KRT skills (1). They are as follows: scenario 1, acute CRRT in a septic, hypotensive oncology patient; scenario 2, maintenance HD initiation in a moderately uremic patient with CKD, with congestive heart failure and volume overload; and scenario 3, acute HD in a patient on maintenance dialysis with severe, life-threatening hyperkalemia and volume overload. The blueprint, questions, and rubric (Supplemental Material, Appendices 1,2, and 3) were developed by the principal investigators (L.K.P. and C.M.Y.) and refined by a nine-member test committee of board-certified, practicing clinical nephrologists. Examinees write dialysis orders for each scenario (history, physical examination,

radiology and laboratory data), and answer open-ended clinical questions. Institutional order sets/protocols are used at the program director's discretion.

Passing score was determined by the test committee (1,5,6), and validated by ten board-certified, practicing volunteers (median 3.5 years from graduation), none of whom were on the test committee. There were 49 test items (58 points). As previously described (Table 1), pass threshold was 46 out of 58 points (15 out of 20 for scenario 1; 17 out of 21 for scenario 2, and 14 out of 17 for scenario 3) (1). No item had a median relevance less than "important," and 95% of positive point items were rated easy to medium difficulty. Content validity index (a measure of how well test items represent essential dialysis skills) was 0.91 (95% confidence interval [95% CI], 0.85 to 0.95) (1,7,8). There were two evidence-based questions per scenario (9–14). Validator tests were graded using the rubric (L.K.P. and C.M.Y., blinded to the other's scoring). Inter-rater reliability was good ( $\kappa=0.68$ ; 95% CI, 0.59 to 0.77). Median test time was 75 minutes. Mean validator score was  $49\pm 3$  (95% CI, 46 to 51). Cronbach  $\alpha$  (a measure of test item internal consistency) was 0.84 for validators and 0.76 for fellows (1,15,16).

### Fellow Testing

The testing protocol was approved by the Walter Reed National Military Medical Center (WRNMMC) Department of Research Programs as exempt from institutional review board review per 32CFR219.101(b) (1,2).

Four ACGME-accredited programs (including WRNMMC) administered the OSCE in May–July 2016. One did not test first-year fellows. Results are reported in the initial validation study (1). Fifteen ACGME-accredited programs

**Table 1. Acute dialysis orders objective structured clinical examination description**

| Question Scenario and Topic (1)   | Total Points | Passing Score (%) | Evidence-Based/Standard-of-Care Questions   |
|---|--------------|-------------------|---|
| 1. Order acute CRRT in a septic, acidemic, hypoxic, coagulopathic, hypotensive, oncology patient.                             | 20           | 15 (75)           | A. Correct for hypalbuminemia when calculating anion gap.<br>B. Obtain at least 20 ml/kg per hour effluent.<br>C. (2017 administration only) Estimate clearance using effluent rate.  |
| 2. Order maintenance HD initiation for a uremic patient with volume overload and an AV fistula <sup>a</sup>                   | 21           | 17 (81)           | A. Avoid low K dialysate (<3 meq/L) in those with normal serum K, unless only a low K dialysate is available.<br>B. Identify uremic encephalopathy (mild to severe) and serositis (pleural, pericardial) as urgent/absolute indications for dialysis. |
| 3. Manage acute, life-threatening hyperkalemia and volume overload in anuric patient with ESKD on maintenance HD <sup>b</sup> | 17           | 14 (82)           | A. Bicarbonate therapy not indicated in acute hyperkalemia in volume-overloaded patient with ESKD without acidosis. Negligible effect.<br>B. Repeat serum K at 2–4 h and at 6 h after dialysis, due to rebound.                                       |
| Overall   | 58           | 46 (79)           | NA  |

CRRT, continuous RRT; HD, hemodialysis; AV, arteriovenous; K, potassium; NA, not applicable.

<sup>a</sup>One item could yield one bonus point (use of smaller gauge dialysis needles in a new AV fistula in scenario 2).

<sup>b</sup>Points could be lost on this question if intravenous sodium bicarbonate was administered (–1 point), if epinephrine was administered (–1 point), or if intravenous Lasix was administered in this anuric patient with ESKD (–1 point).

(including those from 2016) tested fellows in May–August 2017.

Each program received a randomly generated numeric identifier series equal to the number of fellows scheduled. Fellows were assigned an identifier by the program, with fellow-identifier association known only to the program. Fellows from the four programs that tested both years were assigned new identifiers in the second year, not linked to those used previously. Examinees were told beforehand when the OSCE would be given, knew the general topic, but were not encouraged to study for it. They had 2 hours to complete the test, and indicated on the answer sheet their training year (second-year fellows testing in July and August were scored as first-year fellows), time to take the test, and in 2017, whether they had taken the test before. After testing, they received a link for an optional, online, anonymous satisfaction survey. Program directors also received a link for their own anonymous survey (Supplemental Material, Appendices 4 and 5).

Program directors graded the test and shared results with fellows. Using the identifier, they submitted the total score, each scenario score, and in-training examination score for the same training year. Graded tests (anonymous identifier only) were returned to WRNMMC for rescoring (L.K.P. or C.M.Y.) and evidence-based question scoring.

Evidence-based questions 2B (uremic encephalopathy and serositis/pericarditis as urgent/absolute indications for initiation of maintenance dialysis) (12) and 3B (repeat serum potassium (K) at 2–4 and 6 hours after dialysis for acute hyperkalemia to check for rebound) (14) were answered incorrectly by >50% of fellows during initial validation (1). In 2017, we expanded analysis of these questions. We also recorded whether fellows correctly estimated CRRT clearance as effluent volume (scenario 1) (17) Objectives were as follows:

- (1) to determine median time to take the OSCE;
- (2) to determine interrater scoring agreement between programs and WRNMMC investigators;
- (3) to determine overall and scenario pass percentages and mean scores, hypothesizing that second-year fellows (third-years analyzed as second-years) would perform better than first-year fellows, and fellows in programs testing in 2016 would improve in 2017, and perform better than those at programs administering the test for the first time;
- (4) to identify evidence-based questions incorrectly answered by >50% of second-year fellows;
- (5) to determine fellow satisfaction with the OSCE as a formative evaluation tool;
- (6) to determine whether OSCE score correlated with in-training examination score for second-year fellows (3,18).

### Statistical Analyses

Percentages, medians (ranges), means (SD and 95% CI), and counts reported as appropriate. *t* Test, Fisher exact test,  $\kappa$  statistic, and Pearson *r* statistic were used as appropriate. Significance thresholds were  $P < 0.05$ . One-tailed *P* values used for comparisons between second- and first-year fellows, using the hypothesis that second-year fellows would perform significantly better than first-year fellows. All other comparisons were two-tailed.

### Results

Fifteen programs participated in 2017. Four were repeat programs (2016 and 2017). Figure 1 shows the testing flow diagram. A total of 117 fellows took the test (51 first- and 66 second-year fellows, including three third-years), and 114 tests were rescored by WRNMMC. Of these, 105 were program-scored. Repeat programs tested 25 fellows in 2016 (seven first- and 18 second-year) and 22 fellows in 2017 (11 first- and 11 second-year). Eight fellows (all second-year) self-identified as having tested in 2016 and 2017.

In total, 105 out of 117 tests were graded by the program and at WRNMMC. Inter-rater agreement for passing overall between programs and WRNMMC was moderate ( $\kappa = 0.56$ ; 95% CI, 0.40 to 0.72). Programs passed 44 out of 105 (42%) overall, whereas WRNMMC scorers passed 36 out of 105 (34%) ( $P = 0.32$ ).

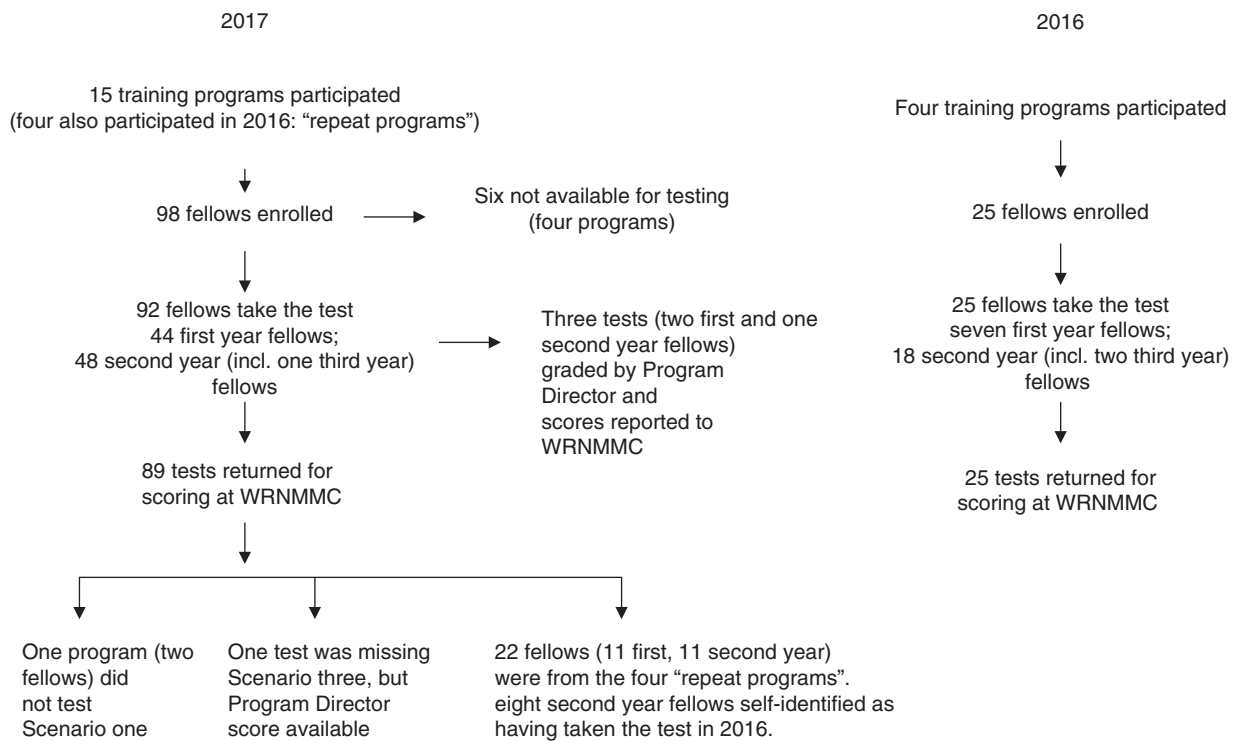
Median testing time reported by first-time takers ( $n = 86$ ) was 65 minutes (range 35–120 minutes). Correlation between testing time and overall score was weak (Pearson  $r = 0.22$ ;  $P = 0.05$ ).

Overall and scenario scores are shown in Table 2. Fellows performed best on scenario 1 (acute CRRT), with 76% passing, and intermediately on scenario 2 (urgent initiation of maintenance HD), with 43% passing. They performed least well on scenario 3 (management of severe hyperkalemia in ESKD), with 6% passing. Second-years were no more likely to pass overall or for a given scenario than first-years, although their mean scores were higher for the overall test and for scenario 2.

Figure 2 shows performance on evidence-based questions. Overall, second-years performed no better than first-years (56% versus 58% correct;  $P = 0.36$ ). Ninety-two percent correctly prescribed  $\geq 20$  ml/kg per hour effluent volume CRRT dose in scenario 1 (Q1B). Sixty-three percent calculated CRRT clearance as effluent volume (Q1C), a question considered “hard” by the test committee. Seventy-five percent correctly prescribed a 3–4 mEq/L K dialysate for maintenance HD initiation of a patient with a normal serum K (Q2A), and 83% recognized that intravenous bicarbonate was not indicated in a volume-overloaded, hyperkalemic patient on maintenance HD without acidosis (Q3A).

Two evidence-based questions were answered correctly by <50% of fellows. Only 12% correctly identified the two urgent/absolute indications for maintenance dialysis initiation—uremic encephalopathy and serositis/pericarditis (Q2B). We investigated this further in 2017, and 61% identified pericarditis and 39% identified encephalopathy as urgent/absolute indications for maintenance dialysis initiation. Eighty-nine percent indicated “uremia” (without qualifiers) as an indication. Only 3% made no mention of any uremic symptom or sign as an urgent indication. The other question (Q3B) required that K levels be checked for rebound at 2–4 and 6 hours after HD for acute hyperkalemia. Only 20% answered correctly, but 93% (82 out of 88) did check K at least once between 2 and 6 hours after dialysis.

Fellows at the four repeating programs did not have higher pass percentages or scores in 2017 versus 2016. In 2016, 36% (nine out of 25) passed versus 45% (ten out of 22) in 2017 ( $P = 0.56$ , Fisher exact test, two-tailed). Overall scores were not significantly different:  $44.1 \pm 3.3$  (2016) versus  $44.9 \pm 5.9$  (2017) ( $P = 0.56$ ; *t* test, two-tailed). In 2017, fellows



**Figure 1.** | Flow diagram of fellow testing in 2016 and 2017.

from the four repeating programs ( $n=22$ ) did not have significantly higher pass percentages than those from the 11 first-time programs ( $n=64$ ): 45% (ten out of 22) versus 25% (16 out of 64) ( $P=0.11$ , Fisher exact test, two-tailed). However, in 2017, 64% of second-year fellows from repeating programs ( $n=11$ ) passed overall (Figure 3), significantly greater than the overall pass percentage for first-year fellows at programs initially giving the OSCE in 2016 and 2017 ( $n=38$ , 24%;  $P=0.03$ , Fisher exact test, two-tailed), but not significantly greater versus first-years ( $n=11$ ) at the four repeating programs, or second-years in 2016 and 2017 at programs initially giving the test ( $n=51$ ).

There was no significant correlation between in-training examination scores and overall OSCE score for second-year fellows ( $n=57$ ; Pearson  $r=0.15$ ;  $P=0.26$ ).

The fellow satisfaction survey (2016 and 2017) had a 56% response (65 out of 117; first-years 51%, second-years 58%). Over 80% strongly agreed/agreed that each scenario “permitted me to assess my proficiency” in ordering KRT (Figure 4). Seventy-seven percent strongly agreed/agreed the OSCE overall was “useful to me in assessing my proficiency in ordering” acute KRT.

In 2017, the program director satisfaction survey response was 80% (12 out of 15). Seventy-five percent strongly agreed/agreed that overall the OSCE was “useful to fellows in assessing their proficiency in ordering” acute KRT. Program director feedback indicated that many institutions’ CRRT orders (especially hemodynamic monitoring and citrate anticoagulation) are protocolized templates, and fellows may not be able to reproduce orders without referring to them. Another criticism was that some questions confused fellows, who

were unsure of the detail needed or the type of answer being sought.

## Discussion

The acute dialysis orders OSCE is a formative assessment (1,19,20), testing commonly used, critically important KRT skills. Locally graded, it permits timely, personalized feedback. Program directors may identify specific fellow or curriculum deficiencies, and adjust accordingly. Fellows have the opportunity for self-assessment in a low-stakes setting.

The OSCE evaluates the ACGME Patient Care and Systems-Based Practice competencies, asking fellows to translate medical knowledge into clinical practice (4). The simulation is simple, inexpensive, and freely available. Standardized patients and sophisticated equipment are not required. Scenarios may be given individually, if desired.

We addressed OSCE construct validity as a unified model, assessing content, response process, internal structure, relation to other variables, and consequences (21). We previously focused on content and internal structure (1). The OSCE was developed and initially validated by clinically active, board-certified nephrologists, who knew the “performance domain” of acute dialysis (7). The content validity index indicated test items were highly representative of the construct, with relevance “essential” or “important” for all test items. Test items appear to have internal consistency (*i.e.*, item performance correlates with the overall test outcome), assessed by Cronbach  $\alpha$  (15,16). Inter-rater reliability was good for investigator graders (1), and moderate for programs versus investigators (programs tended to



**Table 2. Results of fellow testing**

| Test Overall <sup>a</sup>   | All Fellows             | First Year              | Second Year             | P Value (First versus Second Year) <sup>b</sup> |
|---|-------------------------|-------------------------|-------------------------|---|
| No. of fellows  | 111                     | 49                      | 62                      | NA  |
| Overall score, mean±SD (95% CI)   | 43.6±4.6 (42.7 to 44.5) | 42.7±5.0 (41.3 to 44.1) | 44.4±4.0 (43.4 to 45.4) | P=0.02  |
| Proportion reaching pass threshold (46/58 points)                         | 32% (36/111)            | 24% (12/49)             | 39% (24/62)             | P=0.08  |
| Scenario 1  |                         |                         |                         |   |
| No. of fellows  | 112                     | 49                      | 63                      | NA  |
| Overall score, mean±SD (95% CI)   | 16.1±2.0 (15.7 to 16.5) | 16.0±1.9 (15.5 to 16.5) | 16.3±2.1 (15.8 to 16.8) | P=0.2   |
| Proportion reaching pass threshold (15/20 points)                         | 76% (85/112)            | 69% (34/49)             | 81% (51/63)             | P=0.1   |
| Scenario 2  |                         |                         |                         |   |
| No. of fellows  | 114                     | 49                      | 65                      | NA  |
| Overall score, mean±SD (95% CI)   | 16.5±2.3 (16.1 to 16.9) | 16.0±2.5 (15.3 to 16.7) | 16.8±2.1 (16.3 to 17.3) | P=0.03  |
| Proportion reaching pass threshold (17/21 points; 1 bonus point possible) | 43% (49/114)            | 35% (17/49)             | 49% (32/65)             | P=0.09  |
| Scenario 3  |                         |                         |                         |   |
| No. of fellows  | 113                     | 49                      | 64                      | NA  |
| Overall score, mean±SD (95% CI)   | 11.0±1.8 (10.7 to 11.3) | 10.7±2.0 (10.1 to 11.3) | 11.2±1.7 (10.8 to 11.6) | P=0.08  |
| Proportion reaching pass threshold (14/17 points)                         | 6% (7/113)              | 6% (3/49)               | 6% (4/64)               | P=0.65  |

NA, not applicable; 95% CI, 95% confidence interval.  
<sup>a</sup>Scored by Walter Reed National Military Medical Center investigators (L.K.P., C.M.Y.).  
<sup>b</sup>On the basis of the hypothesis that second-year fellow performance would be better than that of first-year fellows, *P* values are for one-tailed tests. Fisher exact test used for pass threshold comparisons. Unpaired *t* test used for score comparisons.

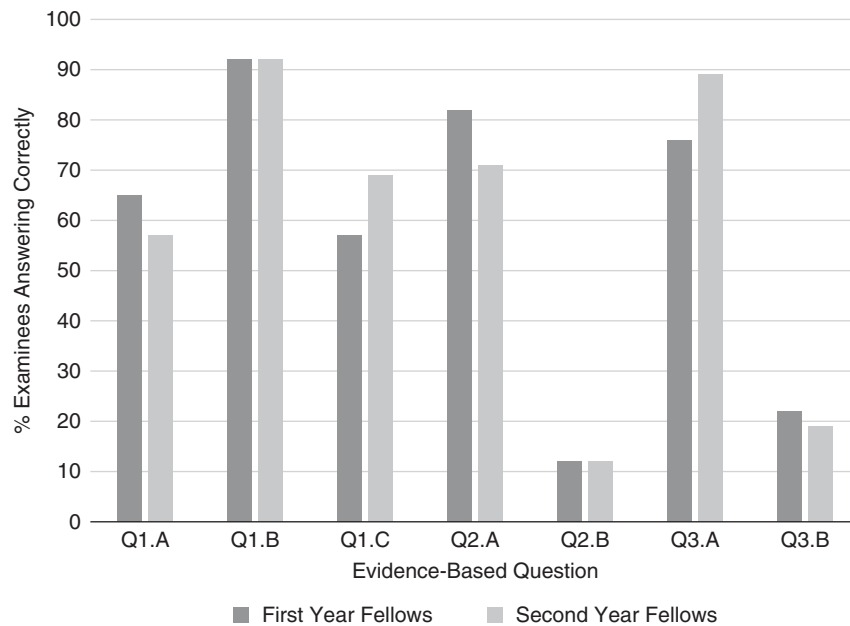
upgrade), suggesting the rubric is sufficiently clear and detailed. The final rubric (<https://nerdc.org>) was modified for greater clarity on the basis of program director surveys. All reflect structural validity (21). Seventy-seven percent of fellows responding to the post-test survey agreed the OSCE was “useful in assessing proficiency in ordering” acute KRT, suggesting examinees understood the test construct in the same way as the test committee (a response process validity measure).

Because there is no validated criterion test of KRT competency, we addressed the relationship of OSCE performance to other variables by prospectively evaluating performance of validators versus fellows, first-versus second-year fellows, and repeating versus first-time programs. Validators had significantly higher pass percentages, scores, and evidence-based question performance, as predicted (1). Although overall scores were greater for second-versus first-years, this did not result in higher overall/individual scenario pass percentages, or evidence-based question performance. This suggests that fellows learn most KRT skills in the first year. Second-years at repeating programs had a significantly higher pass percentage (64%) than did first-years at first-time programs, suggesting curriculum changes and/or individual formative-feedback led to improvement as a consequence of testing. Although second-year fellow in-training examination scores did not correlate with overall score, the in-training

examination addresses the whole spectrum of nephrology medical knowledge. Acute and maintenance dialysis comprise only a small part.

Fellow performance differed substantially between scenarios. They performed well on scenario 1 (acute CRRT), with almost all correctly prescribing a  $\geq 20$  ml/kg per hour effluent rate (Q1B) (10). However, many programs had order sets that prevented effluent prescriptions  $< 20$ –25 ml/kg per hour using constrained pick lists. Over 60% estimated urea clearance using effluent rate (Q1C), suggesting fairly sophisticated knowledge of CRRT clearance (17). The majority of first-year rotations are inpatient, often intensive care unit-based, and fellows appear well prepared to manage CRRT (22). At many programs, cardiovascular monitoring and responses to decompensation during CRRT are not routinely managed by nephrologists, but by intensivists. Monitoring was not included in some standard order sets, and fellows often did not address these issues in the orders, although specifically asked to do so. Constrained picklists and standard order sets (often within the electronic medical record) may have biased fellow performance, an example of an unintended educational consequence of electronic order entry (23,24).

Fellows performed less well on scenario 2 (initiation of maintenance dialysis). As in scenario 1, some lost points because of failure to order monitoring. Over 70% ordered an appropriate K dialysate for a patient with a normal



**Figure 2. | Fellow performance on evidence-based questions.** Scenario 1: Order acute CRRT in septic, hypotensive patient. Q1A: Correct for hypalbuminemia when calculating anion gap (9). Q1B: Obtain at least 20 ml/kg per hour effluent (10). Q1C: (2017 administration only) Estimate clearance using effluent rate (17). Scenario 2: Order maintenance HD initiation for moderately-uremic, volume-overloaded patient. Q2A: Avoid low K dialysate (<3 meq/L) in those with normal serum K (11). Q2B: Identify uremic encephalopathy and serositis (pleural, pericardial) as urgent/absolute indications for dialysis (12). Scenario 3: Manage acute hyperkalemia and volume overload in an anuric patient with ESKD. Q3A: Bicarbonate therapy not indicated in volume-overloaded ESKD patient without acidosis (13). Q3B: Repeat serum K at 2–4 hours and at 6 hours after dialysis, due to rebound (14).

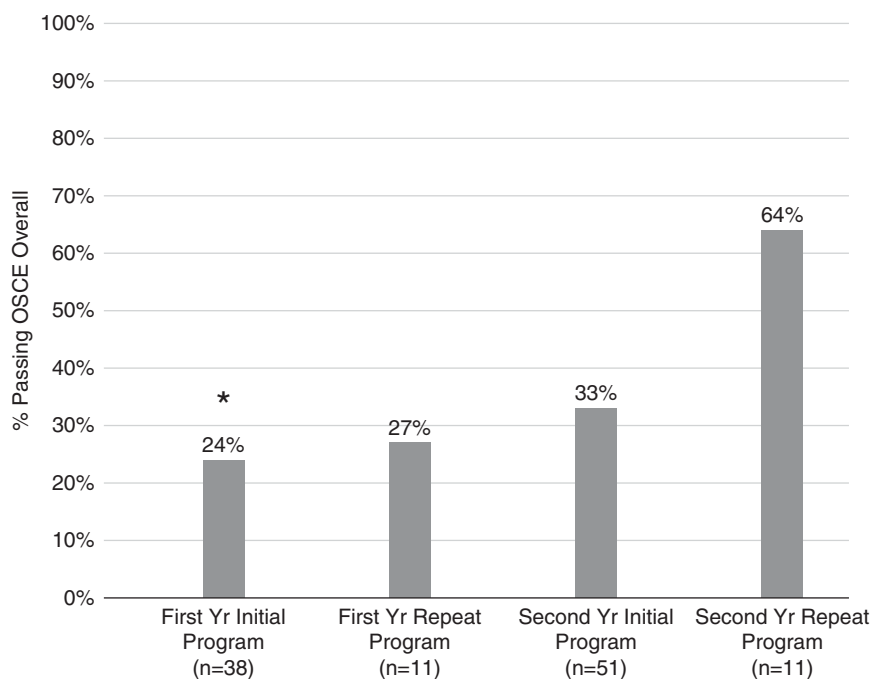
serum K and congestive heart failure (11). Only 12% answered that the two absolute/urgent indications for maintenance dialysis initiation were uremic serositis/pericarditis and encephalopathy (12,25). Some were unsure whether the question referred to the patient in the scenario, although the text indicated that the two were “not necessarily (present) in this patient.” The patient was described as having “mild asterixis,” but only 39% answered “uremic encephalopathy,” including in the follow-up question asking for four other indications for maintenance dialysis initiation. 89% did answer that “uremia” was an indication. Only 3% made no mention of uremic signs or symptoms at all. Initiation thresholds for maintenance dialysis are subjective (26). Patients are initiated earlier than formerly. Fellows may be less aware of uremic encephalopathy and serositis as absolute indications for initiating maintenance dialysis, as they are now rarely seen (25).

Fellows performed least well on scenario 3 (acute hyperkalemia in ESKD). The first question asked examinees to provide “orders, monitoring, treatments, and dispositions” to an ER intern with an anuric patient with ESKD with a K of 7.9 meq/L, weakness, dyspnea, and marked volume overload after obvious dietary indiscretion. Many lost points because of insufficient detail, especially for dosing, sequence, and frequency of intravenous calcium, insulin, and glucose. However, 75% recognized that intravenous sodium bicarbonate was not indicated for hyperkalemia in a nonacidotic patient with ESKD, in acute congestive heart failure (13). Although over 90% checked for rebound hyperkalemia after dialysis, only 20% checked twice (14),

and many did not check at the conclusion of dialysis, trusting that hyperkalemia had resolved.

Acute hyperkalemia treatment is controversial, and medical and dialytic standard of care may differ between institutions. Several program directors commented that examinees were unsure how much detail to include, and lost points, although further questioning revealed they knew the material. The final test version (<https://nerdc.org>) was modified to encourage examinees to include detailed answers in the first part of scenario 3. This is one of the benefits of a formative OSCE: fellows can practice skills in a low-risk environment and program directors can interact with fellows to determine if knowledge deficits do indeed exist, and adjust scores accordingly. It is also important that fellows not approach the treatment and monitoring of acute, symptomatic hyperkalemia casually.

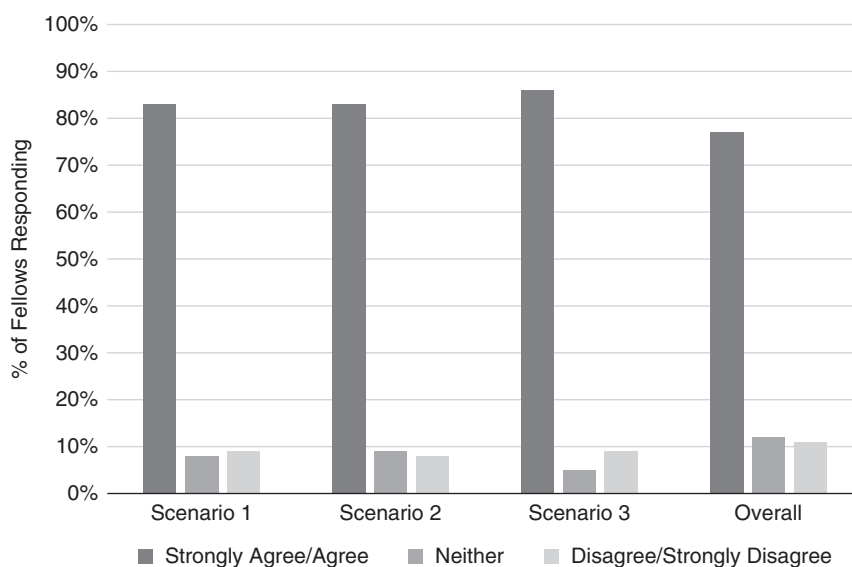
The differences in scenario performance is an outcome (consequence) of testing that can be addressed at a curriculum level. At our own program, we intensified the dialysis didactic curriculum, introduced material earlier in the training year, and focused more on hyperkalemia management. Another participating program also reported making curriculum changes. An advantage of the formative OSCE is that grading is at program level. Allowances can be made for local protocols/procedures, while ensuring fellows know the data underlying constrained pick lists and protocols. Fellows might have been more comfortable with multiple choice testing, where at least one answer is correct and need only be recognized (24). Because fellows receive specific, detailed feedback, and can refer directly to the



**Figure 3.** | Pass performance of first- and second-year fellows from initial testing programs (2016 and 2017,  $n=15$ ) and repeat testing programs (2017,  $n=4$ ). \* $P=0.03$ , Fisher exact test, two-tailed versus second-year fellows at four programs that repeated administration of the test in 2017.

rubric, they may develop a more nuanced self-directed study plan than with a centralized, infrequent, general test of clinical nephrology knowledge, such as the in-training examination. Individual scenarios, which only take about 30 minutes, could be administered frequently at relevant time points (e.g., after specific rotations, at the end of the first year) or diagnostically for a struggling fellow.

Programs can use the OSCE for ongoing assessment of six of the 24 ACGME nephrology subcompetencies (patient care 1–3, medical knowledge 1–2, and systems based practice 1) (4). It provides quantitative, practical, granular data on fellow prescription of dialysis therapy, and might reveal overdependence on computerized provider order entry, particularly in the era of electronic medical record click-box dialysis orders (23,24).



**Figure 4.** | Fellow satisfaction survey results after testing. Scenario 1: “Question 1 permitted me to assess my proficiency in ordering acute CRRT in a critically ill patient with AKI.” Scenario 2: “Question 2 permitted me to assess my proficiency in ordering HD initiation in a moderately uremic patient at ESKD.” Scenario 3: “Question three permitted me to assess my proficiency in managing acute hyperkalemia and volume overload in an ESKD patient on chronic HD.” Overall: “Overall, the acute dialysis OSCE was useful to me in assessing my proficiency in ordering acute RRT.”

Our future goal is to expand the menu of available questions in the KRT performance domain, collaborating with program directors and clinical nephrologists throughout the United States. We are validating a peritoneal dialysis scenario, and previously published an OSCE simulating rare HD-specific emergencies (27). Suggested topics include vascular access assessment, home HD, water purification/monitoring, and management of dialysis-associated cardiovascular complications. Expert technical understanding and provision of KRT is a critical skill for all nephrologists, and the acute dialysis orders OSCE should prove valuable in quantitatively assessing individual KRT competence and program curriculum efficacy.

### Acknowledgments

We would like to thank the nephrology fellows who participated in the objective structured clinical examination. We would also like to thank Robert M. Perkins and Matthew A. Sparks for their critical review of the manuscript.

The views expressed in this article are those of the authors and do not reflect the official policy of the Department of Army/Navy/Air Force, Department of Defense, or the US Government.

The following Nephrology Education Research and Development Consortium members contributed to this work. Test Committee: Sam W. Gao (Portsmouth, VA), Christopher J. Lebrun (Columbus, MS), Dustin J. Little (Bethesda, MD), David L. Mahoney (Fairfax, VA), Robert Nee (Bethesda, MD), Lisa K. Prince (Bethesda, MD), Mark Saddler (Durango, CO), Maura A. Watson (Bethesda, MD), Christina M. Yuan (Bethesda, MD); Validation Committee: Jonathan A. Bolanos (Bethesda, MD), Amy J. Frankston (Bethesda, MD), Jorge I. Martinez-Osorio (El Paso, TX), Deepti S. Moon (Allentown, PA), David Owshalampur (Tacoma, WA), Bret Pasiuk (Fond du Lac, WI), Robert M. Perkins (Whippany, NJ), Ian M. Rivera (Augusta, GA), John S. Thurlow (Bethesda, MD), Sylvia C. Yoon (North Chicago, IL); Training Programs/Test Administration: Donald I. Baumstein, MD (New York Medical College Metropolitan, New York, NY), Ruth C. Campbell, MD (Medical University of South Carolina, Charleston, SC), Sarah Elfering, MD (University of Minnesota, Minneapolis, MN), Kambiz Kalantari, MD (University of Virginia, Charlottesville, VA), Jessica Kendrick, MD, MPH (University of Colorado, Denver, CO), Joshua D. King, MD (University of Virginia, Charlottesville, VA), Oliver Lenz, MD, MBA (University of Miami, Miami, FL), Yiming Zhao Lit, MD (Stanford University, Palo Alto, CA), Laura S. Maursetter, DO (University of Wisconsin, Madison, WI), Sharon E. Maynard, MD (Lehigh Valley Health Network, Allentown, PA), Michal Melamed, MD (Albert Einstein College of Medicine, Bronx, NY), David I. Ortiz-Melo, MD (Duke University, Durham, NC), Lisa K. Prince, MD (Walter Reed National Military Medical Center, Bethesda, MD), Rajeev Raghavan, MD (Baylor College of Medicine, Houston, TX), Ross J. Scalse, MD (University of Miami, Miami, FL), Matthew A. Sparks, MD (Duke University and Durham VA, Durham, NC), Amy N. Sussman, MD (University of Arizona, Tucson, AZ), Dawn F. Wolfgram, MD (Medical College of Wisconsin, Milwaukee, WI).

### Disclosures

Dr. Prince, Dr. Nee, and Dr. Yuan have nothing to disclose.

### Funding

There was no grant funding or support for this work.

### Supplemental Material

This article contains the following supplemental material online at <http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.02900319/-/DCSupplemental>.

Supplemental Appendix 1. NERDC dialysis orders OSCE blueprint.

Supplemental Appendix 2. NERDC dialysis orders OSCE test (final).

Supplemental Appendix 3. NERDC dialysis orders OSCE rubric (final).

Supplemental Appendix 4. NERDC dialysis orders OSCE fellow survey 2017.

Supplemental Appendix 5. NERDC dialysis orders OSCE program director survey 2017.

### References

1. Prince LK, Campbell RC, Gao SW, Kendrick J, Lebrun CJ, Little DJ, Mahoney DL, Maursetter LA, Nee R, Saddler M, Watson MA, Yuan CM; Nephrology Education Research & Development Consortium: The dialysis orders objective structured clinical examination (OSCE): A formative assessment for nephrology fellows. *Clin Kidney J* 11: 149–155, 2018
2. American Board of Internal Medicine: Nephrology Certification Examination Blueprint, 2018. Available at: <https://www.abim.org/~media/ABIM%20Public/Files/pdf/exam-blueprints/certification/nephrology.pdf>. Accessed July 14, 2018
3. Rosner MH, Berns JS, Parker M, Tolwani A, Bailey J, DiGiovanni S, Lederer E, Norby S, Plumb TJ, Qian Q, Yeun J, Hawley JL, Owens S; ASN In-Training Examination Committee: Development, implementation, and results of the ASN in-training examination for fellows. *Clin J Am Soc Nephrol* 5: 328–334, 2010
4. Accreditation Council for Graduate Medical Education; American Board of Internal Medicine: The Internal Medicine Subspecialty Milestones Project, 2015. Available at: <http://www.acgme.org/Portals/0/PDFs/Milestones/Internal-MedicineSubspecialtyMilestones.pdf>. Accessed November 21, 2016
5. Ebel RL: *Essentials of Educational Measurement*, Englewood Cliffs, NJ, Prentice-Hall, 1972, pp 492–494
6. Livingston SA, Zieky MJ: *Passing Scores: A Manual for Setting Standards of Performance on Educational and Occupational Tests*, Princeton, NJ, Educational Testing Service, 1982, pp 26–29
7. Lawshe CH: A quantitative approach to content validity. *Person Psychol* 28: 563–575, 1975
8. Wilson FR, Pan W, Schumsky DA: Recalculation of the critical values of Lawshe's content validity ratio. *Meas Eval Couns Dev* 45: 197–210, 2012
9. Vichot AA, Rastegar A: Use of anion gap in the evaluation of a patient with metabolic acidosis. *Am J Kidney Dis* 64: 653–657, 2014
10. Palevsky PM, Zhang JH, O'Connor TZ, Chertow GM, Crowley ST, Choudhury D, Finkel K, Kellum JA, Paganini E, Schein RM, Smith MW, Swanson KM, Thompson BT, Vijayan A, Watnick S, Star RA, Peduzzi P; VA/NIH Acute Renal Failure Trial Network: Intensity of renal support in critically ill patients with acute kidney injury. *N Engl J Med* 359: 7–20, 2008
11. Jadoul M, Thumma J, Fuller DS, Tentori F, Li Y, Morgenstern H, Mendelssohn D, Tomo T, Ethier J, Port F, Robinson BM: Modifiable practices associated with sudden death among hemodialysis patients in the dialysis outcomes and practice patterns study. *Clin J Am Soc Nephrol* 7: 765–774, 2012
12. Singh A, Kari J: Management of CKD Stages 4 and 5. In: *Handbook of Dialysis, 5th Edition*, Chapter 2, edited by Daugirdas JT, Blake PG, Ing TS, Philadelphia, PA, Wolters Kluwer Health, 2015.
13. Allon M, Shanklin N: Effect of bicarbonate administration on plasma potassium in dialysis patients: Interactions with insulin and albuterol. *Am J Kidney Dis* 28: 508–514, 1996
14. Blumberg A, Roser HW, Zehnder C, Müller-Brand J: Plasma potassium in patients with terminal renal failure during and after haemodialysis; relationship with dialytic potassium removal and total body potassium. *Nephrol Dial Transplant* 12: 1629–1634, 1997



15. Bland JM, Altman DG: Cronbach's alpha. *BMJ* 314: 572, 1997
16. Tavakol M, Dennick R: Making sense of Cronbach's alpha. *Int J Med Educ* 2: 53–55, 2011
17. Claire-Del Granado R, Macedo E, Chertow GM, Soroko S, Himmelfarb J, Ikizler TA, Paganini EP, Mehta RL: Effluent volume in continuous renal replacement therapy overestimates the delivered dose of dialysis. *Clin J Am Soc Nephrol* 6: 467–475, 2011
18. Jurich D, Duhigg LM, Plumb TJ, Haist SA, Hawley JL, Lipner RS, Smith L, Norby SM: Performance on the nephrology in-training examination and ABIM nephrology certification examination outcomes. *Clin J Am Soc Nephrol* 13: 710–717, 2018
19. Bloom BS: Learning for mastery. *Evaluation Comment*. 1: 1–12, 1968
20. Boston C: *The Concept of Formative Assessment*, ERIC Digest, ERIC Clearinghouse on Assessment and Evaluation, College Park, MD
21. Cook DA, Beckman TJ: Current concepts in validity and reliability for psychometric instruments: Theory and application. *Am J Med* 119: 166.e7–166.e16, 2006
22. Liebman SE, Moore CA, Monk RD, Rizvi MS: What are we doing? A survey of United States nephrology fellowship program directors. *Clin J Am Soc Nephrol* 12: 518–523, 2017
23. Campbell EM, Sittig DF, Ash JS, Guappone KP, Dykstra RH: Types of unintended consequences related to computerized provider order entry. *J Am Med Inform Assoc* 13: 547–556, 2006
24. Tierney MJ, Pageler NM, Kahana M, Pantaleoni JL, Longhurst CA: Medical education in the electronic medical record (EMR) era: Benefits, challenges, and future directions. *Acad Med* 88: 748–752, 2013
25. Bleyer A: Indications for Initiation of Dialysis in Chronic Kidney Disease, 2018. Available at: [https://www.uptodate.com/contents/indications-for-initiation-of-dialysis-in-chronic-kidney-disease?search=uremia%20symptoms&source=search\\_result&selectedTitle=1~150&usage\\_type=default&display\\_rank=1#H1147970492](https://www.uptodate.com/contents/indications-for-initiation-of-dialysis-in-chronic-kidney-disease?search=uremia%20symptoms&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1#H1147970492). Accessed September 22, 2018
26. Rivara MB, Chen CH, Nair A, Cobb D, Himmelfarb J, Mehrotra R: Indication for dialysis initiation and mortality in patients with chronic kidney failure: A retrospective cohort study. *Am J Kidney Dis* 69: 41–50, 2017
27. Prince LK, Abbott KC, Green F, Little D, Nee R, Oliver JD 3rd, Bohlen EM, Yuan CM: Expanding the role of objectively structured clinical examinations in nephrology training. *Am J Kidney Dis* 63: 906–912, 2014

**Received:** March 9, 2019 **Accepted:** May 29, 2019

Published online ahead of print. Publication date available at [www.cjasn.org](http://www.cjasn.org).