



## ORIGINAL ARTICLE

# The dialysis orders objective structured clinical examination (OSCE): a formative assessment for nephrology fellows

Lisa K. Prince<sup>1</sup>, Ruth C. Campbell<sup>2</sup>, Sam W. Gao<sup>3</sup>, Jessica Kendrick<sup>4</sup>, Christopher J. Lebrun<sup>5</sup>, Dustin J. Little<sup>1</sup>, David L. Mahoney<sup>6</sup>, Laura A. Maursetter<sup>7</sup>, Robert Nee<sup>1</sup>, Mark Saddler<sup>8</sup>, Maura A. Watson<sup>1</sup> and Christina M. Yuan<sup>1</sup> for the Nephrology Education Research & Development Consortium

<sup>1</sup>Nephrology Division, Walter Reed National Military Medical Center, Bethesda, MD, USA, <sup>2</sup>Nephrology Division, Medical University of South Carolina, Charleston, SC, USA, <sup>3</sup>Nephrology Division, Portsmouth Naval Medical Center, Portsmouth, VA, USA, <sup>4</sup>Department of Renal Diseases and Hypertension, University of Colorado, Denver, CO, USA, <sup>5</sup>Department of Internal Medicine, Baptist Memorial Hospital, Golden Triangle, Columbus, MS, USA, <sup>6</sup>Nephrology Associates, Fairfax, VA, USA, <sup>7</sup>Nephrology Division, University of Wisconsin, Madison, WI, USA and <sup>8</sup>Nephrology Associates, Mercy Regional Medical Center, Durango, CO, USA

Correspondence and offprint requests to: Christina M. Yuan; E-mail: christina.m.yuan.civ@mail.mil

## Abstract

**Background:** Few quantitative nephrology-specific simulations assess fellow competency. We describe the development and initial validation of a formative objective structured clinical examination (OSCE) assessing fellow competence in ordering acute dialysis.

**Methods:** The three test scenarios were acute continuous renal replacement therapy, chronic dialysis initiation in moderate uremia and acute dialysis in end-stage renal disease-associated hyperkalemia. The test committee included five academic nephrologists and four clinically practicing nephrologists outside of academia. There were 49 test items (58 points). A passing score was 46/58 points. No item had median relevance less than 'important'. The content validity index was 0.91. Ninety-five percent of positive-point items were easy–medium difficulty. Preliminary validation was by 10 board-certified volunteers, not test committee members, a median of 3.5 years from graduation. The mean score was 49 [95% confidence interval (CI) 46–51],  $\kappa = 0.68$  (95% CI 0.59–0.77), Cronbach's  $\alpha = 0.84$ .

**Results:** We subsequently administered the test to 25 fellows. The mean score was 44 (95% CI 43–45); 36% passed the test. Fellows scored significantly less than validators ( $P < 0.001$ ). Of evidence-based questions, 72% were answered correctly by validators and 54% by fellows ( $P = 0.018$ ). Fellows and validators scored least well on the acute hyperkalemia question. In self-assessing proficiency, 71% of fellows surveyed agreed or strongly agreed that the OSCE was useful.

Received: May 7, 2017. Editorial decision: June 21, 2017

Published by Oxford University Press on behalf of ERA EDTA, 2017.

This work is written by US Government employees and is in the public domain in the US.

**Conclusions:** The OSCE may be used to formatively assess fellow proficiency in three common areas of acute dialysis practice. Further validation studies are in progress.

**Key words:** dialysis, nephrology, fellowship, education, testing, objective structured clinical examination

## Introduction

Few quantitative, validated nephrology-specific simulation tools exist to assess Accreditation Council for Graduate Medical Education (ACGME) competency performance of nephrology fellows [1–3]. Prescription of acute hemodialysis (HD) and continuous renal replacement therapy (CRRT) are critical skills that are difficult to test in the multiple choice format used in the nephrology certifying examination. The 2016 American Board of Internal Medicine (ABIM) nephrology certification examination blueprint indicates that 10.5% of questions pertain to end-stage renal disease (ESRD) (HD, peritoneal dialysis and their complications home HD; ESRD complications and dialysis medical director topics) and 4% to acute renal replacement therapy (RRT) [4]. Thus there are few questions on the nephrology certifying examination or in-training examination (whose blueprint parallels the certifying examination) that directly assess ability to prescribe acute RRT [5]. The ACGME subspecialty curricular milestones framework requires that program directors ensure that nephrology fellows demonstrate skill in performing acute and chronic RRT, a patient care subcompetency (PC4a) [6]. This vital clinical skill should be quantitatively and longitudinally assessed and fellows receive feedback regarding their progress.

We developed and initially validated a formative objective structured clinical examination (OSCE) to ascertain fellows' ability to write appropriate orders in three commonly encountered acute RRT scenarios. The test is easy to implement and freely available, using institutionally available protocols and order sets.

## Methods

### OSCE development

The test assesses medical knowledge and patient care competency in three areas, representing common, necessary acute RRT skills (Figure 1). These are (i) acute CRRT in a septic, hypotensive oncology patient; (ii) chronic HD initiation in a moderately uremic patient with volume overload and (iii) acute HD in a chronic dialysis patient with life-threatening hyperkalemia and volume overload. The test blueprint (Supplementary Material 1), test questions (Supplementary Material 2) and rubric (Supplementary Material 3) were developed by the principal investigators (L.K.P. and C.M.Y.) and refined by the nine-member test committee. Five of these members were academic nephrologists from a single training program and four were nonacademic clinical nephrologists in rural (two) or suburban/urban (two) practice. All were board certified in nephrology.

There were no multiple choice or true/false questions. Examinees were required to write acute dialysis orders after reading each question scenario (pertinent history, physical examination, radiology and laboratory data) and answer pertinent clinical questions. Standard order sets and protocols are used at the program director's discretion. The final test contained 49 items, with 58 possible points and two evidence-based/standard-of-care items per question (Table 1) [7–12].

### Passing threshold and validity of content

The test committee set the pass threshold using Ebel's method, rating difficulty and item relevance using individual ballots [13, 14]. The difficulty scale was 1 = easy, 2 = medium and 3 = hard. The relevance scale was 1 = essential, 2 = important, 3 = acceptable and 4 = questionable. Each estimated the percentage of borderline second-year fellows likely to answer each item correctly. Passing threshold was determined by adding the products of the median threshold percentage for each item and the number of points per item (Supplementary Material 4: Example applied to Question Scenario 1).

The median relevance for all items yielding positive points ( $n = 44$ ) was either 'important' (23%) or 'essential' (77%). The median content validity ratio (CVR) ( $n = 44$ ) was 1 (range 0.56–1.0) with a content validity index (CVI) of 0.91 [95% confidence interval (CI) 0.85–0.95] [15, 16]. The median difficulty was 'easy' or 'medium' for 42/44 items (95%). Twenty-two items were rated 'easy', 20 'medium' and 2 'hard'. The 'hard' items, 1 point each (Question Scenario 1), required calculation of CRRT urea clearance using effluent volume and recognition that CRRT drug dosing is based on clearance and sieving coefficients. One test committee member rated the urea clearance calculation of 'questionable' relevance.

The passing score was 46 of 58 points (79%). Passing scores for each scenario are summarized in Table 1.

### Initial test validation

Validators were 10 volunteers who were board-certified, clinically active nephrologists, a median of 3.5 years (1–11 years) from fellowship graduation. None were test committee members.

Each test was graded using the rubric (L.K.P. and C.M.Y., blinded to the other's scoring). Interrater reliability was calculated using kappa (<http://www.graphpad.com/quickcalcs/kappa1/>). The number correctly answering each of the six

Draft test blueprint, 3 question scenarios and rubric (LKP, CMY).

Test committee ( $n = 9$ ) critique and refine blueprint (Supplement 1), question scenarios (Supplement 2) and rubric (Supplement 3), creating final draft.

Test committee individually rates difficulty and relevance of each item in each question scenario (Supplement 4; 49 items, 58 points).

Members individually estimate % of borderline second-year fellows likely to answer each item correctly. Median % determined for each item and pass threshold determined for each question scenario and entire test (Supplement 4).

Validators ( $n = 10$ ) take the test, indicate time to complete and offer recommendations for improving individual items.

Validation tests scored by L.K.P. and C.M.Y. Median completion time,  $\kappa$ , Cronbach's  $\alpha$ , and % passing entire test (and each question scenario) determined.

Final test draft completed.

Fig. 1. Flow diagram of OSCE development.

**Table 1.** Acute dialysis orders OSCE test description

Question scenario and topic	Total points <sup>a</sup>	Total items <sup>a</sup>	Passing score (%)	Evidence-based/standard-of-care questions
Order acute CRRT in a septic, acidemic, hypoxic, coagulopathic, hypotensive oncology patient	20	17	15 (75)	Hypoalbuminemia correction when calculating an anion gap [7]
Order initiation of chronic HD in a moderately uremic patient with volume overload and an AV fistula	21	14	17 (81)	Obtain at least 20 mL/kg/h effluent [8] Avoid low K dialysate (<3 mEq/L) in a patient with normal serum K, unless a low-K dialysate is the only one available [9] Must identify uremic encephalopathy (mild to severe) and serositis (pleural, pericardial) as urgent/absolute indications for dialysis [10]
Manage acute, life-threatening hyperkalemia and volume overload in an anuric ESRD patient on chronic HD	17	18	14 (82)	Bicarbonate therapy not indicated in acute hyperkalemia in an ESRD patient without acidosis and with volume overload, as there is negligible effect on serum potassium [11] Must repeat serum K at 2–4 h and at 6 h after dialysis, due to rebound [12]
Overall	58	49	46 (79)	NA

<sup>a</sup>Some items were worth >1 point. Five items could yield either a 0 or 1 negative point per item (use of heparin in Question Scenario 1, incorrect use of mannitol in Question Scenario 2 and use of intravenous bicarbonate, epinephrine or furosemide in Question Scenario 3). One item could yield 1 bonus point (use of smaller gauge dialysis needles in a new AV fistula in Question Scenario 2).

evidence-based/standard-of-care items was recorded. Each reported test completion time. Two did not do Question Scenario 1 because they were no longer performing CRRT. Some made suggestions to better clarify item wording.

The median test time was 75 min. The mean score was 49 ± 3 (95% CI 46–51; n = 8). Interrater agreement was good: κ = 0.68 (95% CI 0.59–0.77). Validator results are shown in Table 2 and evidence-based question performance in Figure 1. Cronbach's α (n = 8) was 0.84.

**Preliminary fellow testing**

Eight ACGME-accredited programs expressed interest in administering the OSCE (after presentation at the American Society of Nephrology Program Directors Annual Meeting, Chicago, IL, USA, May 2016). Four withdrew because of scheduling constraints. Four [including Walter Reed National Military Medical Center (WRNMMC)] administered the test in May–July 2016 (training year 2015–2016). One program did not give the test to first-year fellows. Fellows were informed several weeks to a month beforehand that the OSCE was scheduled. They were given the general topic, but were not encouraged to prepare. Program directors received the tests 1 week before test administration, with a test administration checklist (including necessary standard RRT order sets). Fellows had 2 h to complete the test, which included detailed instructions. Program directors graded the test using the rubric and scores were shared with the fellows. Each fellow was assigned an anonymous identifier. The program director, using this identifier, reported year in training, score on each question scenario and in-training exam (ITE) score for training year 2015–2016 to investigators at WRNMMC. Graded tests (with anonymous identifier only) were returned to WRNMMC for rescoring (L.K.P. and C.M.Y.) and recording of answers to evidence-based questions.

Fellows entered the time to take the test on their answer sheet and were invited to take an anonymous satisfaction survey (SurveyMonkey) immediately after the test (Supplementary Material 5).

**Table 2.** Validator results on acute dialysis orders OSCE

Self-reported time to take test, median (range)	75 min (60–180)
Overall score, mean ± SD (95% CI)	49 ± 3 (46–51)
Those reaching passing score threshold of 46/58 points (n = 8)	88% (7/8)
Question Scenario 1 (acute CRRT) score, mean ± SD (95% CI)	17 ± 1 (17–18)
Those reaching passing score threshold of 15/20 points (n = 8)	100% (8/8)
Question Scenario 2 (initiation of chronic HD) score, mean ± SD (95% CI)	18 ± 2 (17–19)
Those reaching passing score threshold of 17/21 points (n = 10)	90% (9/10)
Question Scenario 3 (management of acute hyperkalemia in ESRD) score, mean ± SD (95% CI)	12 ± 2 (11–14)
Those reaching passing score threshold of 14/17 points (n = 10)	50% (5/10)

The following objectives and hypotheses were tested:

- Determine the median time to take the OSCE.
- Determine the overall and scenario mean scores, hypothesizing that second-year fellows would score higher than first-years and initially estimating the score difference between the two groups.
- Identify evidence-based questions incorrectly answered by > 50% of second-year fellows.
- Determine fellow satisfaction with the OSCE as a formative evaluation tool.
- Determine whether the OSCE score correlated with the 2015–2016 ITE score.

The fellow testing protocol was reviewed and approved by the WRNMMC Department of Research Programs as exempt from institutional review board review per 32 CFR 219.101 (b)(1)–(2).

Table 3. Results of fellow testing

Result	All fellows	First year	Second year	P-value
Number of fellows	25	7	16	NA
Self-reported time to take test, min, median (range)	60 (35–120)	60 (40–120)	65 (35–120)	NA
Overall score, mean $\pm$ SD (95% CI)	44 $\pm$ 3 (43–45)	43 $\pm$ 3 (41–45)	45 $\pm$ 3 (43–46)	0.30
Those reaching passing score threshold of 46/58 points	36% (9/25)	29% (2/7)	44% (7/16)	0.66
Question Scenario 1 (acute CRRT) score, mean $\pm$ SD (95% CI)	17 $\pm$ 2 (16–17)	17 $\pm$ 1 (16–18)	17 $\pm$ 2 (16–18)	0.90
Those reaching passing score threshold of 15/20 points	84% (21/25)	100% (7/7)	88% (14/16)	1.00
Question Scenario 2 (initiation of chronic HD) score, mean $\pm$ SD (95% CI)	16 $\pm$ 2 (16–17)	16 $\pm$ 1 (15–17)	17 $\pm$ 2 (16–17)	0.31
Those reaching passing score threshold of 17/21 points	48% (12/25)	14% (1/7)	56% (9/16)	0.09
Question Scenario 3 (management of acute hyperkalemia in ESRD) score, mean $\pm$ SD (95% CI)	11 $\pm$ 2 (10–12)	11 $\pm$ 2 (9–12)	11 $\pm$ 2 (10–12)	0.54
Those reaching passing score threshold of 14/17 points	8% (2/25)	14% (1/7)	6% (1/16)	0.53

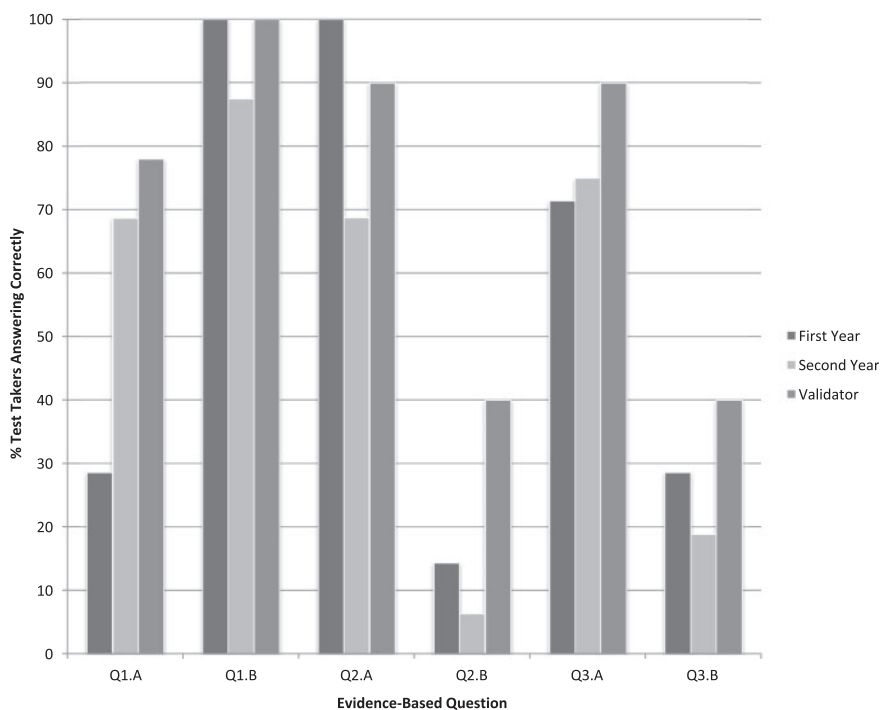


Fig. 2. Performance on evidence-based/standard-of-care question by validators and fellows. Q1.A: Perform hypoalbuminemia correction when calculating an anion gap [7]. Q1.B: Obtain at least 20 mL/kg/h effluent during CRRT [8]. Q2.A: Avoid low-K dialysate (<3 mEq/L) in a patient with normal serum K unless a low-K dialysate is the only one available [9]. Q2.B: Must identify uremic encephalopathy (mild to severe) and serositis (pleural, pericardial) as urgent/absolute indications for dialysis [10]. Q3.A: Bicarbonate therapy not indicated in acute hyperkalemia in ESRD patient without acidosis and with volume overload, as there is a negligible effect on serum K [11]. Q3.B: Must repeat serum K at 2–4 h and at 6 h after dialysis, due to rebound [12].

### Statistical analysis

Percentages, medians (ranges), means (SD and 95% CI), and counts were reported as appropriate. Two-tailed t-test, paired t-test, Fisher's exact test and Pearson's *r* were used as appropriate. Cronbach's  $\alpha$  was calculated using unstandardized scores for 44 items, permitting negative signs (STATA 12.1, StataCorp, College Station, TX, USA) [17, 18]. The CVR and CVI were calculated defining as 'essential' relevancy scores of 1 (essential) and 2 (important), 'useful but not essential' as 3 (acceptable) and 'not necessary' as 4 (questionable) [13, 15]. The significance threshold was  $P < 0.05$ .

### Results

Twenty-five fellows took the OSCE: 7 first-year, 16 second-year and 2 third-year. The median test time was 60 min (range 35–

120). The mean overall score (C.M.Y. and L.K.P.) was  $44 \pm 3$  (95% CI 43–45), not significantly different than that of program directors:  $45 \pm 5$  (95% CI 42–47) ( $P = 0.44$ , paired t-test). Validators performed better than fellows ( $P = 0.0004$ , unpaired t-test). Nine of 25 fellows passed (36%), Significantly fewer fellows passed [9/25 (36%)] than validators [7/8 (88%);  $P = 0.017$ ].

Table 3 shows overall fellow performance and compares first- and second-year fellows. There was no significant difference in performance overall or on any individual question scenario between first- and second-year fellows. Fellows did best on Question Scenario 1 (84% passed) and least well on Question Scenario 3 (8% passed).

Performance on evidence-based/standard-of-care questions is shown in Figure 2. Validators were significantly more likely overall to answer correctly than fellows (72% versus 54%;  $P = 0.018$ ). Second-year fellows were no more likely to answer correctly overall than first-years (54% versus 57%;  $P = 0.85$ ).

Table 4. Validation matrix and sources of evidence for the acute dialysis orders OSCE [21, 22]

Construct validity [21, 23–25]	Definition	Sources of evidence to establish construct validity
Source of evidence	Component	Content (Do test items represent the construct?) Response (Do test takers engage in the performance being measured and understand the construct being measured in the same way as the test developers?) Structural (Is test reliable? Are predicted differences confirmed? Is scoring reproducible?) Relationships with external variables (Is there correlation with scores from another instrument?) Consequences (Are intended outcomes achieved?) Measurement(s) performed in this study (or future studies) <sup>a</sup>
Content [15]	The degree to which the OSCE is representative of the knowledge being measured—the ‘job performance domain’ (acute RRT management)	Test committee (board-certified, clinically active nephrologists) who know the ‘job performance domain’ first-hand agreed on blueprint and determined pass threshold using accepted methods CVI high Median item relevance deemed essential or important for all items
Response	The degree to which test construct is understood and demonstrated by those taking the test	Performance and feedback from validators. 88% passed the OSCE—all were board-certified, credentialed, clinically active nephrologists 71% of fellows surveyed agreed that the OSCE was useful in assessing proficiency in ordering acute RRT
Structural	Internal consistency  Inter-rater reliability Confirmation of predicted differences	Cronbach’s $\alpha$ to measure internal consistency was acceptable for both validators and fellows [18] Good. $\kappa = 0.68$ Board-certified, clinically active validators had high overall pass rate and significantly higher scores and pass rates than fellows
Relationships with external variables (predictive and concurrent validity)	The degree to which the OSCE correlates with or predicts performance on an independent (criterion) measure of the same attribute, i.e. fellow competence in management of acute RRT. (An independent measure may not exist. Both the ITE and Nephrology Board examination are general measures of medical knowledge and are not specific to RRT)	88% of validators passed the test—all were board-certified, credentialed, clinically active nephrologists. (concurrent validity) Correlation with fellow ITE scores not demonstrated (concurrent validity) <i>Hypothesize that OSCE performance predicts passing the ABIM nephrology examination (predictive validity)</i>
Consequences	Are intended outcomes of this formative OSCE achieved? Are there unintended outcomes?	<i>Hypothesize that: Second-year fellows will have higher pass rates/scores than first-year fellows First-year fellows who previously took the test will do better as second-year fellows, after formative feedback First-year fellows who previously took the test will do better as second-year fellows, after formative feedback</i>

<sup>a</sup>Potential future studies to further demonstrate construct validity are shown in italics.

Eighty-eight percent of fellows provided a minimum CRRT effluent rate of 20 mL/kg/h (Q1.B) [8]. Seventy-six percent of fellows avoided a dialysate potassium <3 mEq/L in a patient with normal serum potassium (Q2.A) [9], and 76% correctly did not use intravenous sodium bicarbonate to treat acute hyperkalemia in a non-acidotic, volume-overloaded chronic dialysis patient (Q3.A) [11]. Fewer than 50% of second-year fellows identified uremic encephalopathy and pericarditis/serositis as urgent/absolute indications for chronic dialysis initiation (Q2.B) [10]. Thirty-six percent of fellows identified both as absolute/compelling indications for initiation; 70% of validators did so. Fewer than 50% of second-year fellows correctly monitored potassium for rebound after dialysis for acute hyperkalemia (Q3.B) [12]. While only 29% of first-year fellows corrected for

hypoalbuminemia when calculating anion gap (Q1.A), 69% of second-years did so [7].

Scores on Question Scenario 3 (management of acute hyperkalemia in ESRD) were lowest for validators and fellows. In addition to writing HD orders, this scenario required a detailed order set for management and monitoring of life-threatening hyperkalemia, before and after dialysis. Meeting the passing threshold (14/17 points) were 2/25 fellows (8%) and 5/10 validators (50%). Negative-point items (intravenous sodium bicarbonate [11] or furosemide) contributed, but many did not specify electrocardiogram monitoring, repeat potassium determination and correct dosing and sequence of intravenous calcium, insulin, glucose and inhaled beta-agonists [19]. Many did not order intradialytic potassium monitoring or monitor for rebound

hyperkalemia. This is reflected by performance on Q3.B (Table 1), which required that potassium be repeated at 2–4 and 6 hours after dialysis for rebound [12]. Twenty percent of fellows and 40% of validators answered Q3.B correctly.

There was no significant correlation between ITE score and OSCE score ( $r=0.104$ ,  $P=0.62$ ). Cronbach's  $\alpha$  for fellow test administration was 0.76.

Seventeen of 25 fellows responded to the satisfaction survey (68%)—first-year 57% and second-year 81%. Twelve (71%) agreed/strongly agreed that the OSCE overall was 'useful to me in assessing my proficiency in ordering acute RRT'. Two (12%) disagreed/strongly disagreed.

## Discussion

To our knowledge, there are no quantitative tests that specifically assess competence in initiation and management of acute RRT, a defining skill for nephrologists [20]. We developed, initially validated and administered a formative OSCE to assess acute RRT skills in three common and vital clinical situations: (i) acute CRRT in a septic, hypotensive oncology patient; (ii) chronic HD initiation in a moderately uremic, volume-overloaded patient and (iii) acute HD in a chronic dialysis patient with life-threatening hyperkalemia and volume overload. The OSCE is easy to administer, takes < 2 h and is simple and low cost, requiring only institutional order sets/protocols.

The OSCE emphasized writing complete, individualized RRT order sets—a simulation of RRT management knowledge put into clinical practice. The evidence used to support initial validity of the test construct (competence in delivery of acute RRT) and studies planned in the future are summarized in Table 4. Interpretations drawn from the OSCE (i.e. scores) must be justifiable and actionable [21], allowing worthwhile formative feedback. For any test, including this one, a single administration cannot establish construct validity, which builds with repeated administrations, sources of evidence and (ultimately) the prospective association of test scores with real clinical performance [23, 24].

As expected, validators had significantly higher scores and pass rates than fellows. There was no significant difference in scores between first- and second-year fellows in this small, preliminary sample, ~3% of the 863 US nephrology fellows in training year 2015–2016 [26]. Fellow ITE scores did not correlate with OSCE performance. This is not surprising since so little acute RRT is covered on the ITE (10% on acute kidney injury and intensive care unit (ICU) nephrology and 10% on chronic kidney disease, including chronic dialysis) [5].

The OSCE appears to have face validity, based on the fellow satisfaction survey. But as a formative (and experimental) test, fellows may not have taken it as seriously as they might the ITE. Fellows reported less time to complete the test (60 min) than did validators (75 min). One completed the test in 35 min.

First- and second-year fellows did well on Question Scenario 1 (management of CRRT), with 91% passing. Second-year fellows did no better than first-years, suggesting this skill is learned in the first year [27]. More than 50% did not include cardiovascular monitoring and response thresholds in their order sets, representing an area for improvement in CRRT training. Program directors should consider reviewing standard order sets with fellows to ensure they understand RRT adjustment in response to blood pressure changes, vasopressor requirements and laboratory results. Fellows may be too reliant on ICU staff to manage acute patient status changes during RRT and might benefit by being more frequently called for management advice.

Both fellows and validators did least well on Question Scenario 3 (management of acute hyperkalemia in ESRD). Examinees were specifically asked to provide initial, detailed monitoring and treatment orders for acute hyperkalemia before and after dialysis. Many orders were incomplete, out of sequence or incorrect. Laboratory and cardiovascular monitoring were often absent. Perhaps examinees did not carefully read the question, defer treatment details to other providers (e.g. emergency medicine) or rely too heavily on standard dialysis order sets. Because some hyperkalemia management recommendations are empiric, there may be variation in local standard practice. The question may need further refinement, which can be explored during subsequent validation. However, we should not assume fellows complete internal medicine residency knowing how to manage hyperkalemia. Some may have learned ineffective or potentially harmful practices [11, 19].

The OSCE is formative, designed to identify gaps in knowledge and training that can (i) focus fellow learning, (ii) improve curriculum and (iii) assist program directors and clinical competency committees in quantitative assessment of milestone progress [25]. The ACGME subspecialty curricular milestones assessed are PC4a, skill in performing invasive procedures (acute and chronic RRT); PC2, develop and achieve a comprehensive management plan for each patient; Medical Knowledge 1 (MK1), possess clinical knowledge; and MK2, knowledge of diagnostic testing and procedures [6]. The OSCE should be graded and discussed with the fellow shortly after completion. Question scenarios may be given individually. Future goals include expanding validation of existing questions, and adding new questions to broaden coverage of the RRT performance domain. We invite program directors and clinical nephrologists throughout the USA to participate.

## Acknowledgements

We would like to thank the nephrology fellows who participated in the OSCE.

### NERDC Members:

**Test Committee:** Lisa K. Prince, Bethesda, MD Sam W. Gao, Portsmouth, VA Christopher J. Lebrun, Columbus, MS Dustin J. Little, Bethesda, MD David L. Mahoney, Fairfax, VA Robert Nee, Bethesda, MD Mark Saddler, Durango, CO Maura A. Watson, Bethesda, MD Christina M. Yuan, Bethesda, MD

**Validators:** Jonathan A. Bolanos, Portsmouth, VA Amy J. Frankston, Bethesda, MD Jorge I. Martinez-Osorio, Honolulu, HI Deepti S. Moon, Bethesda, MD David Owshalimpur, Tacoma, WA Bret Pasiuk, Fond du Lac, WI Robert M. Perkins, Whippany, NJ Ian M. Rivera, Augusta, GA John S. Thurlow, El Paso, TX Sylvia C. Yoon, North Chicago, IL

**Program Directors:** Lisa K. Prince, Bethesda, MD Ruth C. Campbell, Charleston, SC Jessica Kendrick, Denver, CO Laura S. Maursetter, Madison, WI

## Conflict of interest statement

D.L.M. is a consultant for DaVita. M.S. is an employee of DaVita, Durango, CO, USA. The views expressed are those of the authors and do not necessarily reflect the official policy or position of the Departments of the Army, Navy or Air Force, the Department of Defense nor the US government. The identification of specific products or scientific

instrumentation is considered an integral part of the scientific endeavor and does not constitute endorsement or implied endorsement on the part of the authors, the DoD or any component agency.

## References

- Mcquillan RF, Clark E, Zahirieh A et al. Performance of temporary hemodialysis catheter insertion by nephrology fellows and attending nephrologists. *Clin J Am Soc Nephrol* 2015; 10: 1767–1772
- Prince LK, Abbott KC, Green F et al. Expanding the role of objectively structured clinical examinations in nephrology training. *Am J Kidney Dis* 2014; 63: 906–912
- Dawoud D, Lyndon W, Mrug S et al. Impact of ultrasound-guided kidney biopsy simulation of trainee confidence and biopsy outcomes. *Am J Nephrol* 2012; 36: 570–574
- American Board of Internal Medicine Nephrology Certification Examination Blueprint; January 2016. <https://www.abim.org/~media/ABIM%20Public/Files/pdf/exam-blueprints/certification/nephrology.pdf> (21 November 2016, date last accessed)
- Rosner MH, Berns JS, Parker M et al. Development, implementation, and results of the ASN in-training examination for fellows. *Clin J Am Soc Nephrol* 2010; 5: 328–334
- Internal Medicine Subspecialty Milestones Project: a joint initiative of the ACGME and the ABIM. July 2015. <http://www.acgme.org/Portals/0/PDFs/Milestones/InternalMedicineSubspecialtyMilestones.pdf> (21 November 2016, date last accessed)
- Vichot AA, Rastegar A. Use of the anion gap in the evaluation of a patient with metabolic acidosis. *Am J Kidney Dis* 2014; 64: 653–657
- VA/NIH Acute Renal Failure Trial Network. Palevsky PM, Zhang JH, O'Connor TZ et al. Intensity of renal support in critically ill patients with acute kidney injury. *N Engl J Med* 2008; 359: 7–20
- Jadoul M, Thumma J, Fuller DS et al. Modifiable practices associated with sudden death among hemodialysis patients in the dialysis outcomes and practice patterns study. *Clin J Am Soc Nephrol* 2012; 7: 765–774
- Singh A, Kari J. Management of CKD stages 4 and 5. In: Daugirdas JT, Blake PG, Ing TS (eds). *Handbook of Dialysis*, 5th edn. Philadelphia, PA: Wolters Kluwer Health, 2015, 30
- Allon M, Shanklin N. Effect of bicarbonate administration on plasma potassium in dialysis patients: Interactions with insulin and albuterol. *Am J Kidney Dis* 1996; 28: 508–514
- Blumberg A, Roser HW, Zehnder C. Plasma potassium in patients with terminal renal failure during and after hemodialysis: relationship with dialytic potassium removal and total body potassium. *Nephrol Dial Transplant* 1997; 12: 1629–1633
- Ebel RL. *Essentials of Educational Measurement*. Englewood Cliffs, NJ: Prentice-Hall, 1972
- Livingston SA, Zieky MJ. *Passing Scores: A Manual for Setting Standards of Performance on Educational and Occupational Tests*. Princeton, NJ: Educational Testing Service, 1982
- Lawshe CH. A quantitative approach to content validity. *Pers Psychol* 1975; 28: 563–575
- Wilson FR, Pan W, Schumsky DA. Recalculation of the critical values of Lawshe's content validity ratio. *Meas Eval Counsel Dev* 2012; 45: 197–210
- Bland JM, Altman DG. *Statistics notes: Cronbach's alpha*. *Br Med J* 1997; 314: 572
- Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011; 2: 53–55
- Mount DB. Treatment and prevention of hyperkalemia in adults. UpToDate (updated 9/20/29016). [https://www.uptodate.com/contents/treatment-and-prevention-of-hyperkalemia-in-adults?source=search\\_result&search=managment%20of%20hyperkalemia&selectedTitle=1~150](https://www.uptodate.com/contents/treatment-and-prevention-of-hyperkalemia-in-adults?source=search_result&search=managment%20of%20hyperkalemia&selectedTitle=1~150) (9 February 2017, date last accessed)
- Jhaveri KD, Perazella MA. Nephrologists as educators: clarifying roles, seizing opportunities. *Clin J Am Soc Nephrol* 2016; 22: 176–189
- Messick S. *Validity of test interpretation and use*. Princeton, NJ: Educational Testing Service, 1990
- Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull* 1955; 52: 281–302
- Cook DA, Brydges R, Ginsburg S et al. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ* 2015; 49: 560–575
- Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med* 2006; 119: e7–e16
- Yuan CM, Prince LK, Oliver JD 3rd et al. Implementation of nephrology subspecialty curricular milestones. *Am J Kidney Dis* 2015; 66: 15–22
- Salsberg E, Quigley L, Mehfoed N et al. *The US Nephrology Workforce 2016: Developments and Trends*. Washington, DC: American Society of Nephrology, 2016
- Liebman SC, Moore CA, Monk RD et al. What are we doing? A survey of United States nephrology program directors. *Clin J Am Soc Nephrol* 2017; 12: 518–523