

# Metodologias Ativas em Fisioterapia: Estudo de Confiabilidade Interexaminador do Método Osce

## Active Methods in Physical Therapy: Interreliability Study of the Osce Method

Roberta de Oliveira Cacho<sup>I</sup>

Marina Pegoraro Baroni<sup>II</sup>

João Afonso Ruaro<sup>II</sup>

Johnnatas Mikael Lopes<sup>III</sup>

Heloisa Maria Jácome de Sousa Britto<sup>I</sup>

Thaiana Barbosa Ferreira<sup>I</sup>

Marcello Barbosa Otoni Gonçalves Guedes<sup>I</sup>

Débora Carvalho de Oliveira<sup>I</sup>

### PALAVRAS-CHAVE

- Aprendizagem;
- Avaliação Educacional;
- Fisioterapia;
- Educação Médica.

### KEYWORDS

- Learning;
- Educational Measurement;
- Physiotherapy;
- Medical Education.

Recebido em: 20/06/2012

Reencaminhado em: 20/06/2014

Aprovado em: 21/12/2015

### RESUMO

O presente estudo relata a experiência da aplicação do Exame Clínico Objetivo Estruturado (do inglês Objective Structured Clinical Examination – Osce) em um curso de Fisioterapia e demonstrar a confiabilidade interexaminador dos instrumentos de avaliação da simulação de atendimento fisioterapêutico. Participaram do estudo 29 estudantes regularmente matriculados no sétimo período do curso de Fisioterapia da Faculdade de Ciências da Saúde do Trairi, da Universidade Federal do Rio Grande do Norte (Facisa/UFRN). Foram elaboradas quatro estações de casos clínicos referentes à área de aparelho locomotor; para cada estação, havia um instrumento de avaliação da simulação do atendimento fisioterapêutico com as opções de resposta “sim”, “não” e “insuficiente”, com avaliação realizada por dois docentes em cada estação. O aluno teve sete minutos em cada estação para realizar a tarefa clínica, sendo que quatro alunos foram avaliados simultaneamente. Houve diferença significativa na distribuição de notas entre as estações 1 e 2 ( $p < 0,001$ ) e 1 e 3 ( $p = 0,001$ ). Verificou-se confiabilidade interexaminador excelente nas estações 1 (CCI = 0,89), 2 (CCI = 0,99) e 3 (CCI = 0,99), enquanto na estação 4 a confiabilidade interexaminador foi satisfatória (CCI = 0,73). Dessa forma, os achados indicam que o Osce na prática fisioterapêutica possui confiabilidade interexaminador de satisfatória a excelente, independentemente do perfil de ensino-aprendizagem e do instrumento de avaliação adotado, podendo ser um método útil para o processo avaliativo da formação profissional em saúde.

### ABSTRACT

This study aimed to report the experience with Objective Structured Clinical Examination in a Physiotherapy undergraduate course and demonstrate the interrater reliability of the developed checklists. The study sample consisted of twenty-nine students in the seventh semester of the Physiotherapy course at the Rio Grande do Norte Federal University Trairi Health Sciences School. Four different stations, each simulating a clinical case of the locomotor system were set up and a checklist contained options to answer “yes”, “no” or “insufficient” to assess the physiotherapeutic care, with two teachers at each station. Four students were evaluated simultaneously and had 7 minutes at each station. There was a significant difference in the distribution of grades between stations 1 and 2 ( $p < 0.001$ ), and 1 and 3 ( $p = 0.001$ ). An excellent level of reliability was found at station 1 (ICC = 0.89), 2 (ICC=0.99) and 3 (ICC = 0.99), whereas at station 4 a this level was satisfactory (ICC = 0.73). Thus, these findings indicate that OSCE in physical therapy practice presents satisfactory to excellent interrater reliability results regardless of the teaching-learning profile and the adopted evaluation, and may be a useful method for professional health care training.

<sup>I</sup> Faculdade de Ciências da Saúde do Trairi, Santa Cruz, RN, Brasil.

<sup>II</sup> Universidade Estadual do Centro-Oeste, Guarapuava, PR, Brasil.

<sup>III</sup> Universidade Federal do Rio Grande do Norte, Santa Cruz, RN, Brasil.

## INTRODUÇÃO

Uma avaliação apropriada na área da saúde deve envolver métodos que foquem o aprendizado e o desenvolvimento de habilidades e comportamentos. As provas tradicionais (múltipla escolha ou prova oral) avaliam a habilidade de memorizar e reconhecer fatos, mas não a de interpretar informações e aplicá-las no atendimento do paciente<sup>1</sup>. Como forma de avaliar estas habilidades não testadas em provas tradicionais, Harden e colaboradores<sup>2</sup>, em 1975, criaram o Exame Clínico Objetivo e Estruturado (do inglês *Objective Structured Clinical Examination – Osce*). Atualmente, o Osce vem sendo considerado um dos métodos mais confiáveis e eficazes para avaliação de habilidades de acadêmicos, principalmente na área da saúde<sup>3,4</sup>.

O Osce permite, no âmbito acadêmico, aproximar a relação cíclica existente entre os conceitos de avaliação e aprendizagem<sup>5</sup>. Além disso, estimula o reconhecimento de uma possível insuficiência no conhecimento em determinada área abordada no exame, proporcionando ao indivíduo avaliado a oportunidade de melhorar suas aptidões e diminuir dificuldades<sup>6</sup>. Desta forma, em longo prazo, o Osce pode contribuir com a prevenção de falhas clínicas na rotina profissional desses indivíduos.

Tipicamente, este exame consiste numa série de estações independentes, cada uma envolvendo um cenário com caso clínico e respectivo paciente simulado, onde o aluno deverá executar determinada tarefa clínica<sup>7</sup>. O desempenho é mensurado por meio de um instrumento de avaliação da simulação do atendimento, elaborado previamente pelo avaliador<sup>8</sup>.

Nos casos clínicos contidos em cada estação, diversas temáticas podem ser exploradas, como executar um procedimento (anotações breves sobre história da doença atual, achados no exame físico, conduta terapêutica, por exemplo), dar ao paciente uma notícia ruim, observar o comportamento perante diferenças culturais, postura ética e até a linguagem no atendimento do paciente<sup>1,9</sup>.

O Osce é parte integrante do processo avaliativo de muitos programas de residência médica no exterior e de exames de certificação profissional<sup>10,11</sup>. Embora muito se discuta na literatura a respeito desta forma de avaliação, pouco se sabe sobre sua aplicabilidade entre acadêmicos de Fisioterapia, sendo as evidências de reprodutibilidade interexaminador ainda incipientes<sup>12,13</sup>. Desta forma, o objetivo deste trabalho é relatar a experiência da aplicação do Osce nas disciplinas da área de atenção fisioterapêutica do aparelho locomotor, do curso de Fisioterapia da Faculdade de Ciências da Saúde do Trairí, da Universidade Federal do Rio Grande do Norte (Facisa/UFRN), e demonstrar a confiabilidade interexaminador dos instrumentos de avaliação da simulação da prática fisioterapêutica.

## MATERIAIS E MÉTODOS

### Disciplinas envolvidas

O Osce foi aplicado no término do primeiro semestre de 2013 para a turma matriculada no sétimo período do curso de Fisioterapia da Facisa/UFRN, a fim de avaliar as competências técnicas e de comunicação das disciplinas relacionadas à área de atenção fisioterapêutica do aparelho locomotor (Atenção Fisioterapêutica em Ortopedia e Traumatologia, Atenção Fisioterapêutica em Reumatologia e Atenção Fisioterapêutica em Neurologia). No momento da realização do Osce, 29 estudantes estavam regularmente matriculados nas disciplinas e todos participaram do exame. O resultado obtido no Osce representava peso 3 em cada uma destas disciplinas.

### Construção das estações e processo de avaliação

O processo de construção desta forma de avaliação passou por três etapas: (i) planejamento da execução, elaboração de casos clínicos e proposição dos instrumentos de avaliação; (ii) preparação dos cenários, dos pacientes simulados e dos auxiliares; (iii) avaliação dos estudantes pelo instrumento desenvolvido.

#### Etapa 1

Na primeira etapa, os professores responsáveis pelas três disciplinas envolvidas no processo de elaboração do Osce reuniram-se periodicamente três meses antes da data agendada para a execução do exame. Decidiu-se elaborar quatro estações referentes a conteúdos estudados no semestre com o objetivo de simular os campos de futura prática profissional pelos quais os estudantes passaram:

- Estação 1: caso clínico de paciente neurológico inserido na Atenção Básica em saúde;
- Estação 2: avaliação de paciente ortopédico/reumatológico;
- Estação 3: caso clínico de paciente ortopédico/reumatológico inserido no serviço de saúde de média complexidade;
- Estação 4: caso clínico de paciente neurológico inserido no serviço de saúde de média complexidade.

As estações foram planejadas para ocorrerem em ambientes isolados, porém dentro de um espaço comum, possibilitando a avaliação simultânea de quatro alunos. O tempo estabelecido para cada estação foi de sete minutos<sup>10,14</sup>.

Para cada estação foram elaboradas tarefas clínicas (procedimento, comunicação ou atitude a serem executados pelos acadêmicos) específicas para o caso que estava sendo apresentado. Para a estação 1 foram elaboradas três tarefas clínicas; para a estação 2, cinco tarefas; e para as estações 3 e 4, duas tarefas. Concomitantemente à execução dessas tarefas, dois avaliadores presentes em cada estação preenchiam no instrumen-

**QUADRO 1.**  
**Exemplo de estruturação da estação 4**

<b>Caso clínico</b>				
Dona Maria sofreu um Acidente Vascular Encefálico Isquêmico há três meses que deixou seqüela de hemiparesia proporcionada à esquerda. Dona Maria não deambula (faz uso de cadeira de rodas), mas assume ortostatismo com auxílio do terapeuta. Ela está em sua oitava sessão de fisioterapia na clínica-escola da Facisa.				
<b>Comandos</b>				
1. Execute um exercício de equilíbrio no tablado; 2. Execute um exercício de equilíbrio em pé.				
<b>Checklist</b>				
Item	Sim	Não	Insuf.	Valor
Chama o paciente pelo nome				0,5
Freia a cadeira de rodas antes de iniciar a transferência?				0,5
Executa a transferência da CR para o tablado de forma adequada? (técnica e segurança do paciente)				1,0
<b>COMANDO 1</b>				
O exercício escolhido no comando 1 está treinando o equilíbrio?				1,0
Realiza o posicionamento das mãos e do próprio corpo para oferecer direcionamento e suporte necessário ao paciente durante a execução dos exercícios?				1,0
Preocupou-se em posicionar o lado acometido durante o exercício?				0,5
Explica e demonstra o exercício ao paciente antes de iniciá-lo?				0,5
Oferece <i>feedback</i> adequado ao paciente no comando 1? (linguagem, clareza)				0,5
Usa equipamentos adequados ao exercício proposto no comando 1 (bolas, rolos, objetos)?				0,5
<b>COMANDO 2</b>				
O exercício escolhido no comando 2 está treinando o equilíbrio?				1,0
Realiza o posicionamento das mãos e do próprio corpo para oferecer direcionamento e suporte necessário ao paciente durante a execução dos exercícios?				1,0
Preocupou-se em posicionar o lado acometido durante o exercício?				0,5
Explica e demonstra o exercício ao paciente antes de iniciá-lo?				0,5
Oferece <i>feedback</i> adequado ao paciente no comando 2 (linguagem, clareza)?				0,5
Usa equipamentos adequados ao exercício proposto no comando 2 (bolas, rolos, objetos)?				0,5
<b>Total</b>				<b>10,0</b>

to de avaliação a execução da tarefa. Cada avaliador possuía um instrumento de avaliação do atendimento fisioterapêutico com as expectativas de resposta para cada tópico avaliado, divididas em “sim” e “não” para as estações 2 e 3, e em “sim”, “não” e “insuficiente” para as estações 1 e 4.

Era assinalado o conceito “sim” quando o aluno executava a tarefa conforme esperado e pontuava integralmente, e “não” quando o aluno não a executava conforme esperado e, portanto, não pontuava no item. A opção “insuficiente” era assinalada quando o aluno não executava a ação exatamente conforme esperado, mas também não errava por completo, e a pontuação correspondia a 50% do valor do item. A soma final de todos os itens era de 10 pontos.

O Quadro 1 explicita como exemplo as tarefas clínicas e o instrumento de avaliação do atendimento fisioterapêutico da estação 4.

Como esta metodologia de avaliação não havia sido aplicada ainda na turma, todo o processo foi previamente explicado aos discentes, e as dúvidas foram esclarecidas.

## Etapa 2

A segunda etapa refere-se à execução do Osce. Os cenários de cada estação foram preparados no ginásio terapêutico da Clínica Escola de Fisioterapia da Facisa, conferindo um aspecto o mais realista possível ao desenvolvimento das práticas.

No dia anterior à aplicação do Osce, foi feito um piloto com a participação de docentes de outras áreas do curso, que não sabiam do formato da avaliação, para corrigir eventuais falhas (na apresentação escrita do caso, nas tarefas clínicas, no tempo planejado de execução e no instrumento de avaliação da simulação do atendimento). Nesse momento, detectou-se que havia excesso de tarefas clínicas em uma estação e que, portanto, o tempo predeterminado seria insuficiente. Como solução, o número de tarefas foi reduzido, e o instrumento de avaliação foi adequado à nova realidade.

Ao chegar a cada estação, o aluno encontrava um ambiente simulado de atendimento, com paciente simulado posicionado e as respectivas tarefas clínicas que deveria executar. Cada cenário deixava à disposição do acadêmico materiais

que ele poderia utilizar ou não para atender às solicitações das tarefas.

Os casos clínicos eram apresentados aos acadêmicos por meio de uma folha fixada na parede, em fonte arial com tamanho mínimo de 20, a fim de facilitar a leitura. Primeiramente, era apresentado o caso (breve resumo do diagnóstico funcional e quadro clínico) e, em seguida, as tarefas clínicas que ele deveria executar (Quadro 1).

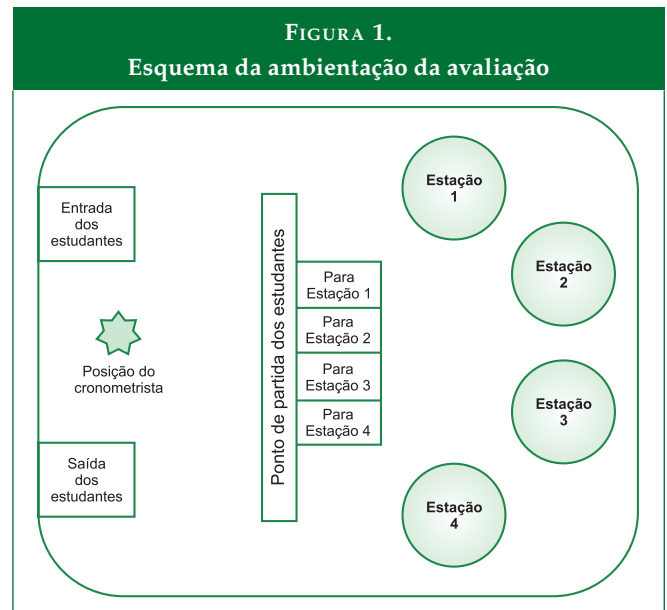
Como pacientes simulados, foram convidados estudantes do primeiro ano do curso de Fisioterapia, escolhidos por terem menor conhecimento acerca da prática fisioterapêutica aplicada ao aparelho locomotor, e, assim, não influenciar na avaliação dos colegas. Eles receberam um texto com informações acerca do caso clínico que simulariam, além de orientações de como se comportar de maneira imparcial. Não foi solicitado que os pacientes simulados utilizassem estratégias dramáticas ou de conotação emocional nas tarefas solicitadas nas estações. Os pacientes simulados tiveram conhecimento do caso clínico 30 minutos antes da execução da avaliação, para não haver risco de divulgação das informações da prova. Desta forma, 30 minutos antes da atividade avaliativa, os pacientes simulados se dirigiam à respectiva estação para uma reunião prévia com os professores responsáveis pela mesma, tempo adequado para conhecer o caso clínico, receber instruções e sanar eventuais dúvidas. Isso permitiu que eles agissem de forma coerente, conforme esperado e planejado.

Como as avaliações nas quatro estações ocorriam simultaneamente, foi necessário recrutar dois auxiliares (servidores administrativos da universidade), um responsável pela manutenção dos estudantes numa sala de espera (para que os que estivessem esperando não tivessem contato com os estudantes que já haviam sido avaliados) e outro responsável por cronometrar o tempo de duração de cada estação.

Aos estudantes que ficavam na sala de espera era permitida a consulta a qualquer material impresso e a discussão sobre os conteúdos da avaliação. Entretanto, foi proibido o uso de eletrônicos (celulares, *tablets*, *notebooks*, entre outros). Além disso, foi recomendado previamente que eles levassem alimentos e bebidas, e, caso houvesse necessidade de ir ao banheiro, os estudantes que esperavam pela avaliação eram acompanhados por um dos auxiliares. Tais medidas foram tomadas a fim de evitar contato com os estudantes que já haviam sido avaliados.

Para iniciar o Osce, quatro estudantes entravam no local da avaliação e se dispunham no ponto de partida de cada estação (Figura 1). Como o tempo determinado foi de sete minutos, o cronometrista passava três comandos aos estudantes: “início”, quando o cronômetro era acionado; “faltam dois mi-

nutos”, quando o tempo estava em cinco minutos (para informar ao aluno que o final estava próximo); e “final”, quando o cronômetro marcava sete minutos.



Dentro do tempo disponível, o aluno lia o caso clínico, as orientações e executava as tarefas clínicas solicitadas. Finalizado o tempo, avançava até o ponto de partida de uma nova estação e esperava um novo sinal para iniciar outra avaliação. Assim, quando os quatro estudantes passavam pelas quatro estações, eles deixavam o ambiente do Osce, e outros quatro estudantes entravam para dar início à avaliação, processo que durou aproximadamente 30 minutos para cada grupo de alunos. Este ciclo se repetiu até que todos os 29 estudantes fossem avaliados.

### Etapa 3

Como eram sete docentes responsáveis pelas disciplinas e foram realizadas quatro estações, um oitavo docente (de outra área do conhecimento da Fisioterapia) foi convidado a integrar a equipe. Desta forma, cada estação contou com dois docentes avaliadores, sendo que o oitavo membro auxiliou na avaliação da estação 4.

Os docentes avaliadores estavam presentes nos ambientes em posições estratégicas para melhor visualizar as tarefas clínicas executadas e anotá-las no instrumento de avaliação. Todos se portaram com total discrição, não responderam a nenhuma pergunta e somente observaram as ações dos estudantes.

Ao final deste processo, a pontuação foi somada e gerou a nota do aluno para cada professor. A pontuação dada ao aluno em cada estação foi considerada a maior nota registrada entre

os dois professores. Como cada estação gerou uma nota, foi realizada a média aritmética entre as quatro estações para se chegar à nota final do Osce.

Ao saírem do ambiente de avaliação, os alunos eram direcionados a uma nova sala, onde tinham acesso à visualização das tarefas clínicas, do instrumento de avaliação da simulação do atendimento fisioterapêutico e da respectiva pontuação de cada item no instrumento. O resultado final de cada estação foi repassado aos alunos no dia seguinte à aplicação da avaliação, haja vista que o processo todo levou quase seis horas de execução.

### Análise estatística

As notas emitidas pelos avaliadores foram descritas na forma de média e desvio padrão (DP). Utilizou-se o teste de Shapiro-Wilk para análise de distribuição de normalidade, tendo em vista o tamanho da amostra, inferior a 50 observações.

Para verificar a diferença de notas entre as estações, optou-se pela Análise de Variância (Anova) de medidas repetidas. A confiabilidade interexaminador da avaliação em cada estação foi determinada por estimativa relativa por meio do Coeficiente de Correlação Intraclasse (CCI) com medida simples e modelagem *two-way random Anova*, onde valores de CCI < 0,40 eram considerados de confiabilidade pobre,  $0,40 \leq \text{CCI} < 0,75$  como satisfatória, e  $\text{CCI} \geq 0,75$  uma confiabilidade excelente<sup>15</sup>.

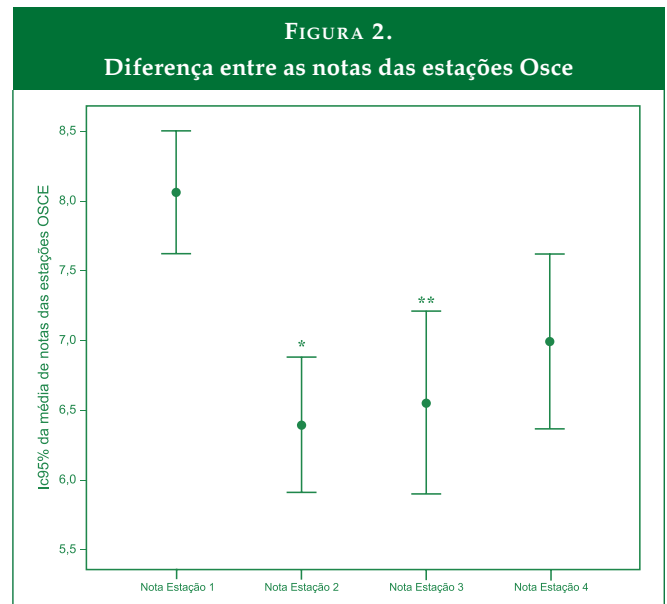
Estimou-se também a confiabilidade absoluta nas estações com a elaboração de gráficos de Altman-Bland, determinando os limites de concordância (LC) de 95% das diferenças entre as notas pela seguinte fórmula:  $\text{LC} = (\text{média das notas}) \pm 1,96 \text{ DP}$ , admitindo-se o pressuposto de distribuição de normalidade da diferença das notas entre os avaliadores. Também foi executado teste de regressão linear simples com a média e diferença de notas entre os avaliadores a fim de detectar viés proporcional. Os dados foram analisados no SPSS®, versão 20.0, da IBM®, adotando-se um  $\alpha \leq 0,05$  como meio de minimizar o erro tipo I.

## RESULTADOS

Todos os professores envolvidos com a atividade concordaram em que o Osce proporciona um meio eficaz de mensurar o conhecimento do aluno, por ser um método que utiliza a simulação do atendimento fisioterapêutico real. O uso do instrumento de avaliação da simulação do atendimento fisioterapêutico visou padronizar os critérios de avaliação e minimizar a discrepância entre os avaliadores.

Após a análise da amostra de 29 discentes avaliados, observou-se a seguinte distribuição de notas: estação 1 com nota média 8,04 (1,16); estação 2 com 6,37 ( $\pm 1,27$ ); estação 3 com 6,53 ( $\pm 1,72$ ); estação 4 com 6,97 ( $\pm 1,64$ ). Houve diferença

estatística significativa entre as estações 1 e 2 ( $p < 0,001$ ) e 1 e 3 ( $p = 0,001$ ) (Figura 2).



(\*  $p < 0,05$  entre as estações 1 e 2; \*\*  $p < 0,05$  entre as estações 1 e 3)

A Tabela 1 revela a confiabilidade relativa entre os avaliadores em cada estação. Foi verificada confiabilidade excelente nas estações 1 (CCI = 0,89), 2 (CCI = 0,99) e 3 (CCI = 0,99), enquanto na estação 4 observou-se apenas confiabilidade satisfatória (CCI = 0,73).

**TABELA 1.**  
**Concordância observada entre as notas dos avaliadores nas estações Osce**

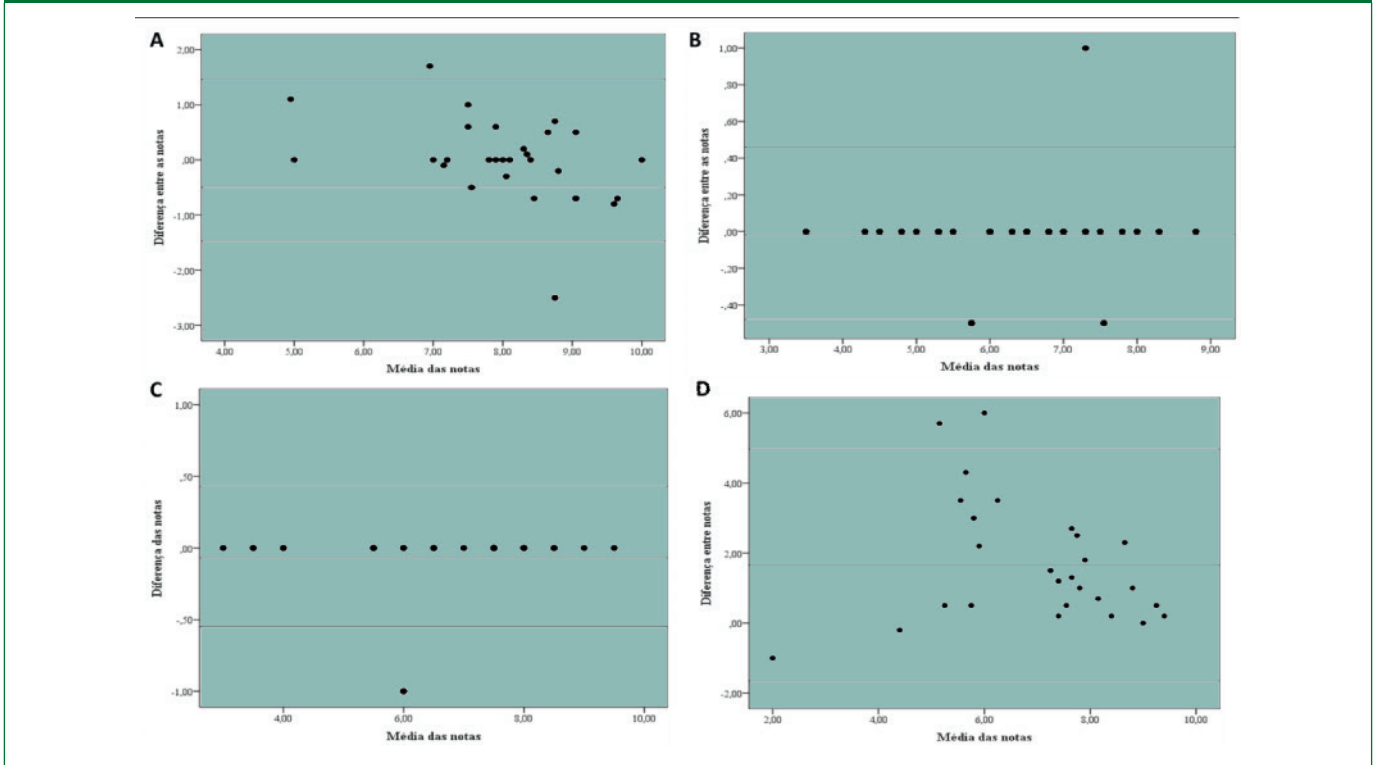
	Avaliador A Média (EP) (IC95%)	Avaliador B Média (EP) (IC95%)	CCI (IC95%)	Interpretação
Estação 1	8,04 (0,19) (5,0-10,0)	8,05 (0,25) (4,4-10,0)	0,89 (0,43-0,86)	Excelente
Estação 2	6,37 (0,24) (3,5-8,8)	6,38 (0,23) (3,5-8,8)	0,99 (0,98-0,99)	Excelente
Estação 3	6,5 (0,32) (3,0-9,5)	6,56 (0,32) (3,0-9,5)	0,99 (0,98-0,99)	Excelente
Estação 4	7,8 (0,32) (1,5-9,8)	6,1 (0,37) (2,3-9,3)	0,73 (0,43-0,88)	Satisfatória

Examinando-se a confiabilidade absoluta entre os avaliadores, na estação 1 detectou-se que os limites de concordância das notas estavam dentro do intervalo de -1,47 a 1,46 ponto, mostrando que a máxima diferença poderia chegar a pouco



FIGURA 3.

Limites de concordância das notas interexaminador (A: estação 1; B: estação 2; C: estação 3; D: estação 4). Observa-se uma confiabilidade mais estreita nas estações 2 e 3.



menos de um ponto e meio. No entanto, estes limites apresentaram um viés proporcional moderado ( $r = -0,45$ ;  $p = 0,02$ ) (Figura 3, A).

Confiabilidade mais estreita foi obtida nas estações 2 e 3, onde os limites absolutos estiveram entre  $-0,48$  e  $0,46$  ponto na estação 2 e  $-0,55$  e  $0,43$  ponto na estação 3, podendo a diferença máxima nessas estações atingir aproximadamente meio ponto. Nestas estações, não foi observado qualquer viés proporcional,  $r = 0,10$  ( $p = 0,57$ ) e  $r = 0,08$  ( $p = 0,65$ ), respectivamente (Figura 3, B e C), porém as diferenças de notas entre os avaliadores não apresentaram distribuição normal ( $p < 0,05$ ), diferentemente das demais estações.

Já na estação 4, os limites de concordância estiveram entre  $-1,62$  e  $4,97$  pontos (Figura 3, D), com a diferença máxima podendo atingir quase 5 pontos no valor da nota do aluno, sem apresentar viés proporcional ( $r = 0,17$ ;  $p = 0,36$ ).

## DISCUSSÃO

Inúmeros métodos têm sido desenvolvidos e utilizados na avaliação dos alunos em diversas áreas da saúde, e a maioria

deles ainda tende a focar o “saber” ou “saber como”, habilidades da esfera cognitiva que são essenciais, mas insuficientes para compor a competência da prática clínica de forma integral, abordando competências como a tomada de decisão. Esses níveis cognitivos podem ser avaliados por testes de múltipla escolha, exame oral, dissertações ou trabalhos escritos<sup>16,17</sup>.

Avaliar a “ação” ou a tomada de decisão significa avaliar o que ocorre na prática clínica, e o Osce se mostrou confiável para medir o “saber fazer” e não somente o “saber como”, ao se aproximar ao máximo de uma situação real do dia a dia da prática profissional<sup>11,18</sup>, mesmo que realizada por pacientes simulados diferentes no processo de avaliação nas suas quatro estações.

A literatura tem apontado que deficiências de padronização por treinamento insuficiente dos pacientes simulados têm sido um dos fatores que contribui para a baixa confiabilidade do método Osce<sup>19,20</sup>. Baig e colaboradores<sup>20</sup> observaram que houve boa consistência interna no comando verbal fornecido pelo paciente simulado, mas que a expressão facial e linguagem corporal diferiram nas quatro estações Osce que conti-

nham o mesmo caso clínico. Já tem se falado em Osce virtual, no qual o ambiente de um simulador permite que o estudante interaja com um paciente virtual, efetue ações e diagnostique casos clínicos virtuais em um ambiente controlado, a fim de minimizar os erros durante a aplicação deste exame<sup>19</sup>.

No entanto, em nosso estudo observou-se que não houve prejuízo ou deficiência de padronização ao se usar um estudante do primeiro ano do curso como paciente simulado. Os bons resultados de confiabilidade obtidos neste estudo podem indicar que os pacientes simulados aderiram às orientações fornecidas na reunião que precedeu o início do Osce e executaram a simulação de forma adequada. No entanto, é necessário realizar outros estudos para confirmar tal inferência. Além disto, em nosso estudo, não foi preciso desenvolver mais que uma estação com o mesmo caso clínico, não havendo necessidade de vários pacientes simulados para o mesmo caso, o que foi apontado como outro fator de baixa confiabilidade do exame<sup>20</sup>. Corroborando nosso estudo, Sadeghi e colaboradores<sup>21</sup> observaram que houve concordância na avaliação dos pacientes padronizados em sete das oito estações realizadas, sugerindo acurácia dos pacientes padronizados para atuação nos casos clínicos propostos.

Embora alguns Osces pareçam ser mais confiáveis do que outros, por motivos que não estão totalmente compreendidos<sup>22</sup>, os autores concordam em que o treinamento dos pacientes simulados deve ser fortalecido e constantemente monitorado durante o exame, a fim de assegurar que o resultado da avaliação seja reflexo real da competência do avaliado, desprovido de erros do Osce<sup>20</sup>.

Os índices de confiabilidade interexaminador alcançados neste estudo concordam com os de outros autores<sup>1,23,24</sup> e reafirmam que o Osce é um método com reprodutibilidade de satisfatória a excelente para a avaliação em situações práticas. Sakurai e colaboradores<sup>13</sup> examinaram a concordância interexaminador do Osce ao avaliarem estudantes de graduação e profissionais de Fisioterapia e Terapia Ocupacional no Japão, e revelaram alta reprodutibilidade na avaliação dos aspectos comportamentais e técnicos dos examinados.

Observou-se uma discrepância entre a nota final da estação 1 e as demais estações. Tal fato não permite inferir que os conhecimentos de Atenção Primária abordados na estação 1 tenham sido mais bem ensinados ou que os itens de avaliação eram mais facilmente executados pelos alunos que nas outras estações, ou, ainda, que o critério dos docentes tenha sido diferente. Para tanto, seria necessária uma análise mais contundente da metodologia de ensino e avaliação dos docentes, fato que não foi objetivo desta pesquisa. Os achados apenas permitem afirmar que, mesmo em situações de variabilidade no processo

de ensino-aprendizagem e avaliação, o Osce mantém alta reprodutibilidade das suas observações interexaminadores.

No Osce, é possível envolver diferentes situações-problema da prática fisioterapêutica. Desta forma, parece ser convincente a alegação de que a falta de sistematização de prática profissional e avaliadores com especialidades distintas levam a avaliações menos concordantes, como ocorreu na estação 4. Harden e colaboradores<sup>2</sup> perceberam que esse processo resulta em considerável subjetividade na avaliação, com os riscos de o viés e o acaso serem elementos-chave que afetam negativamente a validade e confiabilidade do exame. Apesar do exposto, o estudo em questão demonstrou bom CCI nas estações propostas. Ademais, o teste de regressão linear simples entre a média e diferença de notas entre os avaliadores não detectou viés proporcional entre os avaliadores das estações 2, 3 e 4, havendo um viés proporcional moderado somente na estação 1. Isso implica afirmar que não houve grandes interferências de variáveis externas na maior parte das estações avaliadas.

Dentro do instrumento de avaliação da simulação do atendimento fisioterapêutico em cada uma das estações, como opção de resposta para pontuação do aluno, as estações 1 e 4 apresentaram itens que contemplaram uma terceira possibilidade de pontuação, denominada “insuficiente”. Embora as duas estações se assemelhem nas opções de respostas executadas pelo aluno e na possibilidade de pontuação pelo docente, a estação 4 obteve um CCI mais baixo em relação às demais estações, ainda assim, considerado satisfatório. Mesmo com um critério de pontuação a mais, a estação 1 obteve índices de CCI compatíveis com as estações 2 e 3.

Devemos ressaltar que as estações 2 e 3 apresentaram uma variação de aproximadamente 0,5 ponto em nossa análise estatística, fato que pode ser respaldado pelo instrumento de avaliação apresentar apenas duas chaves de resposta, “sim” ou “não”, o que deixa os avaliadores com menos possibilidades de pontuação e, portanto, com menor variabilidade final.

Já nas estações 1 e 4, o instrumento de avaliação possibilitou três chaves de respostas, “sim”, “não” e “insuficiente”, permitindo maior variabilidade na pontuação do avaliador. No entanto, a estação 1, mesmo com as três opções de resposta, manteve um CCI excelente. Partindo do pressuposto de que em um processo de avaliação é esperado que o estudante execute brilhantemente uma ação, se ele a executou de forma correta, mas com alguma falha, ele o fez de forma insuficiente. Apesar das três possibilidades de resposta, a excelente CCI nesta estação (1, sobre a Atenção Básica) pode ter sido oriunda do fato de os professores planejarem suas atividades em conjunto, de modo a executarem atividades semelhantes durante todo o período de prática supervisionada.

No estudo de Wilkinson e colaboradores<sup>25</sup>, fica claro que o instrumento de avaliação da simulação do atendimento deve conter critérios bem definidos, pois a inclusão de muitos itens está associada negativamente com a confiabilidade e validade. Por outro lado, o envolvimento do examinador na elaboração do Osce foi positivamente associado com a confiabilidade interexaminador.

No entanto, o processo de avaliação realizado pelos docentes não pode se restringir apenas a uma estimativa global de reprodutibilidade expressa pelo CCI. Ela deve também apresentar aspectos de avaliação que individualizam o aluno e permitam ao professor emitir pareceres singulares para o aperfeiçoamento de cada um dos envolvidos. É nesse atributo (Figura 3) que detectamos variabilidade considerável nas estações 1 e 4, e similitudes nas estações 2 e 3.

Nas estações 1 e 4, estão evidenciadas características pertinentes aos avaliadores e ao instrumento de avaliação, que sobressaem como explicação para as diferenças máximas encontradas de aproximadamente um ponto e meio na estação 1 e de quase 5 pontos na estação 4, o que poderia aprovar ou reprovar um aluno, dependendo do docente. A primeira hipótese levantada, e concernente à estação 1, é a defasagem de literatura específica para a atuação profissional do fisioterapeuta no nível da atenção à saúde primária, mais especificamente na Estratégia de Saúde da Família<sup>26</sup>, o que ocasiona arbitrariedades em diversas situações da prática profissional e justificaria de maneira plausível o viés proporcional aqui encontrado, mesmo os docentes sendo especializados no tema foco da estação. Para a estação 4, levanta-se a hipótese da formação especializada distinta entre os avaliadores, pois um pertence à área tema da estação (Neurologia), e o outro é da área cardiopulmonar. Apesar de Wilkinson e colaboradores<sup>25</sup> não terem observado associação da experiência do examinador com a área da avaliação, e sim com a falta de envolvimento do avaliador na elaboração do Osce, este fato ocorreu parcialmente em nosso estudo.

Ainda uma última hipótese pode servir para as duas estações e reside na existência da opção “insuficiente” como resposta no instrumento de avaliação da estação, que inclusive pode ser potencializada pela escassez de sistematização da prática profissional avaliada (estação 1) e pela diferença na formação/engajamento na elaboração do Osce (estação 4). No entanto, estudos recentes têm utilizado o *global rating scale* em substituição aos *checklists* e têm observado maior confiabilidade interexaminador<sup>27</sup>. Aquele possui, geralmente, mais de quatro opções de resposta no instrumento de avaliação, diferentemente dos *checklists*, que são instrumentos dicotomizados. Embora nosso estudo tenha usado em duas estações um

instrumento de avaliação no modelo de *checklists* (estações 2 e 3) e similar nas outras duas estações, houve boa confiabilidade interexaminador.

Quanto à experiência do avaliador na área, Regher e colaboradores<sup>24</sup> relataram haver menor risco de viés se respeitada a área de atuação docente. Ainda assim, estudos apontaram que mesmo graduandos dos cursos avaliados podem se mostrar eficientes para aplicação do Osce, desde que previamente treinados<sup>8,28</sup>.

O estudo de Silva e colaboradores<sup>12</sup> demonstrou elevada consistência interna no Osce entre estudantes de pós-graduação de Fisioterapia, atribuindo estes resultados à harmonia existente com relação às habilidades educacionais avaliadas em cada estação e ao preparo cuidadoso dos itens do instrumento de avaliação. Os autores destacam a etapa 1 (planejamento e estruturação) como a mais importante na implantação de um exame do tipo Osce. Cada detalhe deve ser pensado e discutido para que não aconteçam falhas no momento da aplicação. O piloto executado anteriormente para analisar o andamento de cada estação foi de extrema importância para moldar o Osce final, pois fez surgir dúvidas e equívocos não percebidos enquanto a ação estava somente no planejamento.

Outro momento fundamental do processo de ensino-aprendizagem é o *feedback* do professor sobre o desempenho do aluno após o término da prova, pois, dessa forma, o aluno tem a oportunidade de aprender com seus eventuais erros<sup>29</sup>. Recomenda-se efetuar esse momento de correção no mesmo ambiente da aplicação da avaliação, pois isto permite demonstrar as ações esperadas. No presente estudo, a correção das ações do aluno foi realizada após a conclusão das quatro estações, no dia posterior à execução do Osce, sem a divulgação dos conceitos (notas), visto que experiências prévias com a aplicação do exame demonstraram que, quando são divulgadas as notas, os alunos tendem a não se concentrar nas correções posteriores.

Destacam-se como aspectos benéficos da avaliação: possibilidade de interposição de disciplinas que são de áreas comuns, mas nem sempre são vistas de forma correlacionada (ex. integrar o conhecimento de todas as disciplinas que compõem o estágio em aparelho locomotor: Ortopedia e Traumatologia, Reumatologia e Neurologia), o que respeita as Diretrizes Curriculares Nacionais dos cursos de saúde, que buscam práticas educativas de integralidade do currículo, tanto de componentes curriculares, como de níveis de atenção<sup>30</sup>; avaliação da prática clínica padronizada para todos os acadêmicos; e avaliação de atitudes que muitas vezes passam despercebidas no cotidiano da prática.



Entre as dificuldades ou limitações encontradas na execução da atividade proposta, destaca-se a necessidade e disponibilidade de tempo para planejamento e execução, bem como amplo espaço físico e uma numerosa equipe para execução da atividade avaliativa conforme idealizada. Quanto ao delineamento do Osce, apesar de não estimada a fidedignidade do paciente simulado nas estações<sup>20,22,31</sup> e de se saber que o elevado número de estações e examinadores melhora a confiabilidade do exame<sup>22</sup>, percebe-se que estas parecem não ter sido facetas comprometedoras do resultado final, como evidenciado nas estimativas apresentadas. Talvez sua aplicação de forma repetida aprimore as mensurações.

### CONSIDERAÇÕES FINAIS

A metodologia Osce aplicada à avaliação do “saber fazer” na prática fisioterapêutica nos três níveis de atenção à saúde possui confiabilidade interexaminador de satisfatória a excelente, podendo ser um método útil para o processo avaliativo da formação profissional em saúde.

Além disso, existem aspectos que devem ser ressaltados ao se elaborar este tipo de avaliação: os avaliadores precisam ter formação afim com o tema escolhido para o teste; é preferível utilizar um instrumento de avaliação da simulação do atendimento com possibilidade de resposta estreita, tipo correto ou incorreto; e áreas da Fisioterapia com pouca sistematização profissional, como, por exemplo, a Atenção Primária, devem ser mais bem delineadas antes do exame, para evitar arbitrariedades.

### REFERÊNCIAS

1. Casey PM, Goepfert AR, Espey EL, Hammoud MM, Kaczmarczyk JM, Katz NT, et al. To the point: reviews in medical education - the Objective Structured Clinical Examination. *American Journal Obstetrics Gynecology*. 2009;200(1):25-34.
2. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Brit Med J*. 1975;5955(1):447-51.
3. Troncon LEA. Clinical skills assessment: limitations to the introduction of an “OSCE” (Objective Structured Clinical Examination) in a traditional Brazilian medical school. *São Paulo Medical Journal*. 2004; 22(1):12-17.
4. Rushforth HE. Objective structured clinical examination (OSCE): review of literature and implications for nursing education. *Nurse Education Today*. 2007;27(5):481-490.
5. Gamboa-Salcedo T, Martinez-Viniegra N, Peña-Alonso YR, Pacheco-Ríos A, García-Durán R, Sánchez-Medina J. Objective Structured Clinical Examination as an instrument for evaluation of clinical competence in pediatrics. A pilot study. *Bol Med Hosp Infantil Mex*. 2011; 68(3):169-176.
6. Galato D, Alano GM, Trauthman SC, França TF. Pharmacy practice simulations: performance of senior pharmacy students at a University in southern Brazil. *Pharm Pract (Granada)*. 2011; 9(3):136-140.
7. Khan KZ, Ramachandran S, Gaunt K, Pushkar P. The objective structured clinical examination (OSCE): AMEE Guide No. 81. Part I: An historical and theoretical perspective. *Med Teach*. 2013;35:1437-1446.
8. Hodges B, Regehr G, McNaughton N, Tiberius R, Hanson M. OSCE checklists do not capture increasing levels of expertise. *Academic Medicine*. 1999; 74(10):1129-1134.
9. Mazor KM, Zanetti ML, Alper EJ, Hatem D, Barrett SV, Meterko V, et al. Assessing professionalism in the context of an objective structured clinical examination: an in-depth study of the rating process. *Med Educ*. 2007;41:331-40.
10. Sloan DA, Donnelly MB, Schwartz RW, Felts JL, Blue AV, Strodel WE. The Use of the Objective Structured Clinical Examination (OSCE) for Evaluation and Instruction in Graduate Medical Education. *Journal of Surgical Research*. 1996; 63(1):225-230.
11. Amaral FTV, Troncon LEA. Participação de estudantes de medicina como avaliadores em exame estruturado de habilidades clínicas (OSCE). *Rev Bras Educ Med*. 2007;31(1):81-89.
12. Silva CCBM, Lunardi AC, Mendes FAR, Souza FFP, Carvalho CRF. Objective structured clinical evaluation as an assessment method for undergraduate chest physical therapy students: a cross-sectional study. *Rev Bras Fisioter*. 2011;15(6):481-486.
13. Sakurai H, Kanada YI, Sugiura YI, Motoya I, Wada Y, Yamada M, Tomida M, Tanabe S, Teranish T, Tsujimura T, Sawa S, Okanish T. Reliability of the OSCE for Physical and Occupational Therapists. *J Phys Ther Sci*. 2014; 26:1147-1152.
14. Galato D, Alano GM, França TF, Vieira AC. Exame Clínico Objetivo Estruturado (ECO): uma experiência de ensino por meio da simulação do atendimento farmacêutico. *Interface: Comunicação, Saúde, Educação*. 2011;15(36):309-319.
15. Bloch KV, Medronho RA. *Epidemiologia*. 2 ed. Atheneu: Rio de Janeiro; 2008.
16. Craddock D, Mathias H. Assessment option education. *Assessment and evaluation in higher Education*. 2009; 34: 127-140.
17. Pereira DR, Flores MA. Percepções dos estudantes universitários sobre a avaliação das aprendizagens: um estudo exploratório. *Avaliação*. 2012; 17:529-556.

18. Mitchell ML, Henderson A, Groves M, Dalton M, Nulty D. The objective structured clinical examination (OSCE): Optimising its value in the undergraduate nursing curriculum. *Nurse Education Today*. 2009;29(4):398-404.
19. Andrade AF, Madeira CAG, Aires SF. OSCE virtual: simulação de avaliação de casos clínicos. II Congresso Brasileiro de Informática na Educação. 2013:454-464.
20. Baig LA, Beran TN, Vallevand A, Baig ZA, Monroy-Cuadros M. Accuracy of portrayal by standardized patients: results from four OSCE stations conducted for high stakes examinations. *BMC Medical Education*. 2014; 14:97-104.
21. Sadeghi M, Taghva A, Mirsepassi G, Hassanzadeh M. How do examiners and examinees think about role-playing of standardized patients in an OSCE setting? *Acad Psychiatry*. 2007; 31(5):358-362.
22. Brannick MT, Erol-Korkmaz TE, Prewett M. A systematic review of the reliability of objective structured clinical examination scores. *Medical Education*. 2011; 45:1181-1189.
23. Sandoval GE, Valenzuela PM, Monge MM, Toso PA, Triviño XC, Wright AC, et al. Analysis of a learning assessment system for pediatric internship based upon objective structured clinical examination, clinical practice observation and written examination. *J Pediatr*. 2010;86(2):131-136.
24. Regher G, Machae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE. *Journal of the Association of American Medical Colleges*. 1998;73(9):993-997.
25. Wilkinson TJ, Frampton CM, Thompson-Fawcett M, Regan T. Objectivity in objective structured clinical examinations: checklists are no substitute for examiner commitment. *Journal of the Association of American Medical Colleges*. 2003; 78(2):219-223.
26. Bispo Júnior JP. Fisioterapia e saúde coletiva: desafios e novas responsabilidades profissionais. *Ciência & Saúde Coletiva*. 2010; 15(1):1627-1636.
27. Malau-Aduli BS, Mulcahy S, Warnecke E, Otahal P, Teague PA, Turner R, Vleuten CV. Inter-rater reliability: comparison of checklist and global scoring for OSCEs. *Creative Education*. 2012; 3:937-942.
28. Amaral E, Domingues RCL, Bicudo-Zeferino AM. Avaliando competência clínica: o método de avaliação estruturada observacional. *Revista Brasileira de Educação Médica*. 2007; 31(3):287-290.
29. Moineau G, Power B, Pion AMJ, Wood TJ, Humphrey-Murto S. Comparison of student examiner to faculty examiner scoring and feedback in an OSCE. *Medical Education*. 2011; 45:183-191.
30. Brasil. Diretrizes Curriculares Nacionais do Curso de Graduação em Fisioterapia. Resolução CNE/CES, 4 de 19 de fevereiro de 2002.
31. Chesser A, Cameron H, Evans P, Cleland J, Boursicot K, Mires G. Sources of variation in performance on a shared OSCE station across four UK medical schools. *Medical Education*. 2009; 43:526-532.

### CONFLITO DE INTERESSES

Os autores declaram não haver conflito de interesses.

### CONTRIBUIÇÃO DOS AUTORES

Os autores Roberta de Oliveira Cacho, João Afonso Ruaro e Marina Pegoraro Baroni foram responsáveis pela concepção e delineamento do estudo, bem como pela aplicação da metodologia proposta e redação do conteúdo intelectual do artigo. O autor Johnnatas Mikael Lopes foi responsável pela aplicação da metodologia e pela análise estatística e interpretação dos dados. Os autores Heloísa Maria Jácome de Sousa Britto, Thaiana Barbosa Ferreira, Marcello Barbosa Otoni Gonçalves Guedes e Débora Carvalho de Oliveira foram responsáveis pela aplicação da metodologia e revisão crítica relevante do material intelectual do manuscrito.

### ENDEREÇO PARA CORRESPONDÊNCIA

Roberta de Oliveira Cacho  
Universidade Federal do Rio Grande do Norte  
Faculdade de Ciências da Saúde do Trairi – Curso de Fisioterapia  
Rua Vila Trairi, s/nº  
Centro – Santa Cruz  
CEP 59200-000 – RN – Brasil  
E-mail: ro\_fisio1@hotmail.com