Saadallah Azor Fakhouri Filho

Instrução formal de médicos residentes em técnicas de ensino

Tese apresentada à Faculdade de Medicina
da Universidade de São Paulo para
obtenção do título de Doutor em Ciências
Programa de Ciências Médicas
Área de Concentração: Educação e Saúde

Orientadora: Profa. Dra. Maria do Patrocínio
Tenório Nunes

(Versão corrigida. Resolução CoPGr 6018/11, de 13 de outubro de 2011. A versão original está disponível na Biblioteca da FMUSP)

São Paulo
2018

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Fakhouri Filho, Saadallah Azor
Instrução formal de médicos residentes em
técnicas de ensino / Saadallah Azor Fakhouri Filho.
-- São Paulo, 2019.
Tese(doutorado)--Faculdade de Medicina da
Universidade de São Paulo.
Programa de Ciências Médicas.Área de
Concentração: Educação e Saúde.
Orientadora: Maria do Patrocínio Tenório Nunes.

Descritores: 1.Residentes como professor
2.Educação médica 3.Modelos educacionais
4.Treinamento por simulação 5.Avaliação educacional
6.Avaliação objetiva de técnicas de ensino
7.Internato e residência

USP/FM/DBD-011/19

Responsável: Erinalva da Conceição Batista, CRB-8 6755

DEDICATÓRIA

Aos meus pais, Saadallah e Rawak

Aos meus irmãos, Amir e Fauze

Aos CCF

AGRADECIMENTOS

À minha orientadora, Profa. Dra. Maria do Patrocínio Tenório Nunes (Patrô), pelo exemplo de profissional e educadora.

Aos alunos do curso de teatro da Universidade Federal de Uberlândia que atuaram como alunos do internato nas avaliações simuladas da pesquisa.

Aos alunos do internato da Universidade Federal de Uberlândia que auxiliaram nas filmagens das avaliações simuladas.

Aos residentes de geriatria do Hospital de Clínicas da Universidade Federal de Uberlândia pela ajuda na organização das atividades simuladas

Aos residentes do Hospital de Clínicas da Universidade Federal de Uberlândia que concordaram em participar voluntariamente da pesquisa

Aos avaliadores dos vídeos da pesquisa que trabalhavam incansavelmente para que os resultados fossem produzidos

Aos professores das disciplinas de pós-graduação, pelo aperfeiçoamento deste estudo e pelo crescimento pessoal.

Aos membros da Banca de Qualificação do doutorado, Profs. Drs. Aécio Flávio Teixeira de Góis, Alfredo Almeida Pina de Oliveira, Itamar de Souza Santos, pelas colaborações no aprimoramento do trabalho

*"Se enxerguei mais longe é porque me
apoiei em ombros de gigantes"
(Isaac Newton)*

NORMALIZAÇÃO ADOTADA

Esta tese está de acordo com as seguintes normas, em vigor no momento desta publicação:

Referências: adaptado de *International Committee of Medical Journals Editors* (Vancouver).

Universidade de São Paulo. Faculdade de Medicina. Divisão de Biblioteca e Documentação. *Guia de apresentação de dissertações, teses e monografias*. Elaborado por Anneliese Carneiro da Cunha, Maria Julia de A. L. Freddi, Maria F. Crestana, Marinalva de Souza Aragão, Suely Campos Cardoso, Valéria Vilhena. 3a ed. São Paulo: Divisão de Biblioteca e Documentação; 2011.

Abreviaturas dos títulos dos periódicos de acordo com *List of Journals Indexed in Index Medicus*.

SUMÁRIO

1-INTRODUÇÃO:	19
2-OBJETIVOS:	23
2.1-Objetivo Geral ou Primário:	23
2.2-Objetivos Específicos ou Secundários:	23
3-REVISÃO DA LITERATURA:	25
3.1-História da medicina	25
3.2-Ensino Médico nos Estados Unidos:	26
3.3-Os Programas de Residência Médica:	27
3.4-A Relação Residente-Aluno:	27
3.5-Intervenção Educacional - Preceptor em um Minuto (PEM) – One Minute Preceptor – (OMP)	29
3.6-A avaliação de atividades simuladas:	29
4-MÉTODOS:	32
4.1-Tipo de Estudo:	32
4.2-Comitê de ética em pesquisa:	32
4.3- Critérios de Inclusão e Exclusão:	32
4.4- Amostra:	32
4.4.1- Estações de Avaliação:	33
4.4.2- Estudantes Padronizados:	33
4.4.3- Treinamento dos Avaliadores:	34
4.4.4- Instrumento de Avaliação:	35
4.4.5- Intervenções:	36
4.4.6- Etapas do Estudo:	37
4.4.7- Perdas de amostra:	38
4.4.8- Registro em vídeo das atividades:	38
5- ANÁLISE ESTATÍSTICA:	41
6- RESULTADOS:	44
6.1- Dados Sociodemográficos:	44
6.2- Estação Enfermaria:	45
6.2.1- Perdas na amostra:	45

6.2.2- Imputação de dados:.....	47
6.2.3- Confiabilidade dos dados:	48
6.2.4- Teste de hipóteses e comparação por pares:	48
6.3- Estação Procedimento:	53
6.3.1- Perdas na amostra:.....	53
6.3.2- Imputação de dados:.....	56
6.3.3- Confiabilidade dos dados:	56
6.3.4- Teste de hipóteses e comparação por pares:	57
6.4- Estação Emergência:	60
6.4.1- Perdas na amostra:.....	60
6.4.2- Imputação de dados:.....	62
6.4.3- Confiabilidade dos dados:	63
6.4.4- Teste de hipóteses e comparação por pares:	64
6.4.4.1- Comparação dos escores entre os avaliadores nas 3 etapas do estudo para a estação emergência:	64
6.4.4.2-Comparação dos escores entre os residentes nas 3 etapas do estudo para a estação emergência:	68
6.5- Estação Ambulatório:	72
6.5.1- Perdas na amostra:.....	72
6.5.2- Imputação de dados:.....	75
6.5.3- Confiabilidade dos dados:	75
6.5.4- Teste de hipóteses e comparação por pares:	76
6.5.4.1- Comparação dos escores entre os avaliadores nas 3 etapas do estudo para a estação ambulatório:	76
6.5.4.2- Comparação dos escores entre os residentes nas 3 etapas do estudo para a estação ambulatório:	80
7- DISCUSSÃO:.....	85
8- CONCLUSÃO:.....	93
REFERÊNCIAS.....	95

LISTA DE ABREVIATURAS, SÍMBOLOS E SIGLAS

ACGME - American College of Graduate Medical Education

CMN - Comunicação de Más Notícias

OMP - One Minute Preceptor

OSTE - Objective Structured Teaching Examination

OSCEs - Objective Structured Clinical Examination

PEM - Preceptor em um Minuto

PFC - Proporção de Dados Falsamente Classificados

RaTs - Residents as Teachers

SFDP26 - Stanford Faculty Development Program

UFU - Universidade Federal de Uberlândia

MAR - missing atrandom

MCAR - missing completely atrandom

MR – Médicos Residentes

LISTA DE FIGURAS

Figura 1- Seqüências de avaliações e intervenções, datas das atividades, grupos randomizados, amostra em cada etapa/intervenção, principais motivos para perda de amostra e amostragem final após todas as etapas e intervenções.....	39
Figura 2- porcentagem de dados faltantes em cada variável analisada da estação enfermaria.....	46
Figura 3- Distribuição dos dados faltantes no banco de dados da estação enfermaria.	47
Figura 4 - Representação em <i>box plot</i> da Escala Likert: Somatória dos avaliadores para a Estação Enfermaria (sem imputações).	50
Figura 5 - Escala Likert: Somatório dos Avaliadores para a estação enfermaria (com imputações) em <i>boxplot</i>	51
Figura 6 - Escala Global: Somatória dos avaliadores para a estação enfermaria representada em <i>boxplot</i>	52
Figura 7 - Porcentagem de dados faltantes em cada variável analisada da estação procedimento	53
Figura 8- Distribuição dos dados faltantes no banco de dados da estação procedimento.....	55
Figura 9- Escala Likert: Somatória dos Avaliadores para a estação procedimento (sem imputações) representada em <i>boxplot</i>	58
Figura 10- Escala Likert: Somatória dos Avaliadores para a estação procedimento (com imputações) através de <i>boxplot</i>	59
Figura 11- Escala Global: Somatória dos Avaliadores para a estação procedimento através de <i>boxplot</i>	60
Figura 12-porcentagem de dados faltantes em cada variável analisada da estação emergência;	61
Figura 13- Distribuição dos dados faltantes no banco de dados da estação emergência.....	62
Figura 14- Comparação, em <i>boxplot</i> , dos escores produzidos por cada avaliador na primeira etapa da pesquisa (D1), na estação emergência.....	65
Figura 15- Comparação, em <i>boxplot</i> , dos escores produzidos por cada avaliador na segunda etapa da pesquisa (D2), na estação emergência.....	66
Figura 16- Comparação, em <i>boxplot</i> , dos escores produzidos por cada avaliador na terceira etapa da pesquisa (D3), na estação emergência.....	67

Figura 17- representa, graficamente em <i>boxplot</i> , o desempenho dos MR nos dias de avaliação.....	70
Figura 18- Análise dos dados da amostra na estação procedimento, após imputação dos dados faltantes na amostra original. Não houve diferença estatística entre os momentos avaliados.	71
Figura 19. Demonstração da evolução dos escores na avaliação global dos médicos residentes nos diferentes dias de avaliação (D1, D2 e D3) na estação emergência. Não houve diferença estatística entre os dias avaliados.	72
Figura 20 - Porcentagem de dados faltantes em cada variável analisada da estação ambulatorio	73
Figura 21- Distribuição dos dados faltantes no banco de dados da estação ambulatorio.....	74
Figura 22- Comparação, em <i>boxplot</i> , dos escores produzidos por cada avaliador na primeira etapa da pesquisa (D1), na estação ambulatorio.....	77
Figura 23- Comparação, em <i>boxplot</i> , dos escores produzidos por cada avaliador na segunda etapa da pesquisa (D2), na estação ambulatorio.....	78
Figura 24- Comparação, em <i>boxplot</i> , dos escores produzidos por cada avaliador na terceira etapa da pesquisa (D3), na estação ambulatorio.....	79
Figura 25 - Análise dos dados da amostra original na estação ambulatorio, sem a imputação de dados. Houve diferença significativa dos primeiros dias de avaliação em relação à D3 e representativa perda de dados.	81
Figura 26- Análise dos dados da amostra na estação ambulatorio, após imputação dos dados faltantes na amostra original, com diferença significativa de desempenho em D3.	82
Figura 27- Demonstração da evolução dos escores na avaliação global dos médicos residentes nos diferentes dias de avaliação (D1, D2 e D3) na estação ambulatorio.....	83

LISTA DE TABELAS

Tabela1*- Características dos avaliadores das estações OSTE	35
Tabela 2*- Dados sociodemográficos dos médicos residentes que participaram do estudo nas 03 etapas (N=24)	44
Tabela 3*- Coeficientes de confiabilidade para a estação enfermaria.....	48
Tabela 4*- Teste de Hipóteses da escala Likert [†] : somatório dos avaliadores para a estação enfermaria (planilha sem imputação de dados).....	49
Tabela 5*-Teste de Hipóteses da escala Likert [†] : somatória dos avaliadores para a estação enfermaria (planilha com imputação de dados)	51
Tabela 6*-Teste de Hipóteses da escala Global [†] por avaliadores para a estação enfermaria	52
Tabela 7*- Coeficientes de confiabilidade para a estação procedimento	56
Tabela 8*- Teste de Hipóteses da escala Likert [†] : somatória dos avaliadores para a estação procedimento (planilha sem imputação de dados)	57
Tabela 9*-Teste de Hipóteses da escala Likert [†] : somatória dos avaliadores para a estação procedimento (planilha com imputação de dados)	58
Tabela 10*-Teste de Hipóteses da escala Global [†] por avaliadores para a estação procedimento	59
Tabela 11*- Coeficientes de confiabilidade para a estação emergência	63
Tabela 12*-Teste de Hipóteses da escala Likert [†] para comparação dos escores produzidos pelos avaliadores no primeiro dia de avaliações (D1) da estação emergência.....	65
Tabela 13*- Teste de Hipóteses da escala Likert [†] para comparação dos escores produzidos pelos avaliadores no segundo dia de avaliações (D2) da estação emergência	66
Tabela 14*- Teste de Hipóteses da escala Likert [†] para comparação dos escores produzidos pelos avaliadores no terceiro dia de avaliações (D3) da estação emergência	67
Tabela 15*: Teste de Hipóteses da escala Likert [†] : somatória dos avaliadores para a estação emergência (planilha sem imputação de dados)	69
Tabela 16*- Teste de Hipóteses da escala Likert [†] : somatória dos avaliadores para a estação emergência (planilha com imputação de dados)	70
Tabela 17*- Teste de Hipóteses da escala Global [†] por avaliadores para a estação emergência	71
Tabela 18*- Coeficientes de confiabilidade para a estação ambulatório	75

Tabela 19*- Teste de Hipóteses da escala Likert [†] para comparação dos escores produzidos pelos avaliadores no primeiro dia de avaliações (D1) da estação ambulatorio	77
Tabela 20*- Teste de Hipóteses da escala Likert [†] para comparação dos escores produzidos pelos avaliadores no segundo dia de avaliações (D2) da estação ambulatorio	78
Tabela 21*- Teste de Hipóteses da escala Likert [†] para comparação dos escores produzidos pelos avaliadores no terceiro dia de avaliações (D3) da estação ambulatorio	79
Tabela 22*- Teste de Hipóteses da escala Likert [†] : somatória dos avaliadores para a estação ambulatorio (planilha sem imputação de dados)	81
Tabela 23*- Teste de Hipóteses da escala Likert [†] : somatória dos avaliadores para a estação ambulatorio (planilha com imputação de dados)	82
Tabela 24*- Teste de Hipóteses da escala Global [†] por avaliadores para a estação ambulatorio	83

Resumo

Fakhouri Filho SA. *Instrução formal de médicos residentes em técnicas de ensino* [tese]. São Paulo: Faculdade Medicina, Universidade de São Paulo; 2019.

Introdução: Médicos residentes (MR) representam uma grande força de trabalho na maioria dos hospitais universitários brasileiros. O contato destes com alunos da graduação e do internato acrescenta atividades didáticas ao seu repertório diário de atividades assistenciais. Estudos norte-americanos e canadenses demonstram que até 25% do tempo de trabalho de médicos residentes é exercido como atividade de ensino. Pesquisadores apontaram que alunos de medicina creditam aos MR30-85% de toda a teoria clínica adquirida na graduação. Nos Estados Unidos criaram-se programas de capacitação didática, conhecidos como “*Residents as Teachers*” (*RaTs*) – residentes como professores, atualmente presentes em 80% daquelas instituições. Em nosso meio, todo esse contato se dá com pouco ou nenhum treinamento formal dos residentes em técnicas de ensino. No presente estudo investigou-se a hipótese de que o treinamento formal de MR em técnicas de ensino poderia aumentar o desempenho em habilidades de ensino clínico. **Métodos:** Estudo prospectivo, onde escores pré e pós-intervenção foram aferidos por meio de atividades simuladas desenvolvidas, especificamente, para este fim. Quarenta e nove médicos residentes dos programas de clínica médica, cirurgia geral, dermatologia e infectologia da Universidade Federal de Uberlândia participaram do estudo. Foram elaboradas 4 estações simuladas para avaliação de habilidades didáticas em cenários de ambulatório, enfermaria, unidade de emergências e durante a orientação de um procedimento. Todos os MR tiveram seus desempenhos gravados em vídeo para avaliação posterior, por 3 investigadores independentes, que atribuíram escores padronizados para cada cenário. Os MR foram avaliados em 3 etapas, com 1 semana de intervalo entre elas: basal (D1), após primeira intervenção (D2), após segunda intervenção (D3). A intervenção principal baseou-se no método de ensino clínico “Preceptoria em um Minuto - PEM” e uma intervenção controle de comunicação de más notícias. Os residentes foram randomizados eletronicamente quanto à ordem de recebimento de cada uma das intervenções após D1. Todos os residentes foram submetidos às intervenções, principal e controle, com uma semana de intervalo, de acordo com a randomização. **Resultados:** Houve aumento estatisticamente significativo nos escores das avaliações de desempenho dos MR quando comparadas avaliações pré e pós-intervenções nas estações enfermaria e procedimento. As estações emergência e ambulatório produziram baixas confiabilidades entre avaliadores. Na estação enfermaria, foram observadas medianas de escores de 109,5 para a D1; 142,5 para D2 e 137,5 para D3, com diferença significativa entre as etapas de avaliação ($p=0,009$). Análise *pós-hoc* para comparação de pares demonstrou diferença significativa entre D1 e D2 ($p=0,011$). Análise de confiabilidade da estação enfermaria pela teoria da generalização (teoria G) demonstrou

coeficientes de 0,545 para D1; 0,702 para D2 e 0,860 para D3. Os coeficientes de consistência interna e confiabilidade entre avaliadores foram respectivamente 0,913 e 0,595. Na estação procedimentos, foram observadas medianas de escores de 99 para D1; 119 D2; e 119,5 D3. Houve diferença significativa entre as etapas de avaliação com $p < 0,001$. Análise *pós-hoc* para comparação de pares demonstrou diferença significativa entre D1 e D2 ($p = 0,001$) e entre D1 e D3 ($p = 0,004$). Coeficientes G para as etapas de avaliação da estação procedimentos foram calculados em 0,683 para D1; 0,681 para D2 e 0,670 para D3. Consistência interna resultou um coeficiente de 0,726 e a confiabilidade entre avaliadores 0,772. **Conclusões:** O presente estudo demonstrou ganho na aquisição de habilidades de ensino clínico, em MR brasileiros após intervenção específica com método PEM. Observou-se aumento progressivo da confiabilidade geral para a estação enfermagem, com um coeficiente condizente com avaliações de alto padrão obtido em D3. A estação enfermagem também produziu uma alta consistência interna e confiabilidade moderada entre avaliadores. A estação procedimentos produziu coeficientes de confiabilidade bons, porém estáveis através das etapas do estudo. As estações emergência e ambulatório produziram coeficientes de confiabilidade entre observadores ruins, o que prejudicou a análise final dos resultados.

Descritores: residentes como professor; educação médica; modelos educacionais; treinamento por simulação; avaliação educacional; avaliação objetiva de técnicas de ensino; internato e residência.

Abstract

Fakhouri Filho SA. Formal instruction of medical residents in teaching techniques [thesis]. São Paulo: "Faculdade Medicina, Universidade de São Paulo"; 2019.

Introduction: Medical residents (MR) represent a powerful work force in most of Brazilian teaching hospitals. Their contact with undergraduate students adds to their daily work teaching activities. Studies from the United States (US) and Canada indicate that 25% of a resident's working time is spent in teaching duties. Other researchers point to the fact that medical students attribute to MR 30-85% of all clinical theory acquired during graduation. In the US, programs developed specifically to foster MR's teaching skills, known as "Residents as Teachers" (RaTs), are now present at nearly 80% of American institutions. In Brazil, the close contact between MR and interns is conceived under little or no formal training in clinical teaching skills. The present study investigated if delivering teaching skills to MR would enhance their ability to teach simulated interns. **Methods:** We developed a prospective study where scores pre and post intervention were obtained through simulated activities, specifically designed for this purpose. Forty-nine medical residents from internal medicine, general surgery, dermatology and infectious diseases programs in Universidade Federal de Uberlandia took part in the activities. Four simulated stations were developed to assess clinical teaching skills in different scenarios: a clinical ward, in the emergency department, during ambulatory consultation and while teaching a simple procedure. All participants had their performances recorded in video for further evaluation. Three independent raters used a standardized questionnaire to attribute scores for every MR's recorded performance. All MRs were evaluated in 3 distinct occasions, one week apart: pre- intervention (D1), after first intervention (D2) and after second intervention (D3). The main intervention was based on a lecture about a teaching skill's method, the "One Minute Preceptor". Control intervention was based on a lecture about how to deliver bad news. All residents were electronically randomized to receive the main or control interventions after D1. All residents received both interventions, one week apart, according to randomization. **Results:** There was a statistically significant increase in the performance's scores of MRs when we compared pre and post intervention evaluations for stations clinical ward and teaching a procedure. For the clinical ward station, median scores were 109.5 for D1, 142.5 for D2 and 137.5 for D3, with a significant difference between 3 occasions ($p=0.009$). Post-hoc analysis was carried out and evidenced significant differences between D1 and D2 ($p=0.011$). Reliability estimates were calculated using generalizability theory (G theory). The overall coefficient was 0.545 for D1, 0.702 for D2 and 0.860 for D3. Internal consistency resulted in a 0.913 coefficient and interrater reliability 0.595. In teaching a procedure station, we observed median scores of 99 for D1, 119 for D2 and 119.5 for D3. The difference across all occasions was statistically different ($p<0.001$). Post-hoc analysis evidenced significant differences between D1 and D2 ($p=0.001$); D1 and D3 ($p=0.004$). G coefficients for the occasions were 0.683 for D1, 0.681 for

D2 and 0.670 for D3. Internal consistency resulted in a 0.726 coefficient and interrater reliability was 0.772. **Conclusions:** The present study evidenced a gain in clinical teaching skills after specific training session based on the “One Minute Preceptor” method, in studied Brazilian MRs. The reliability analysis demonstrated progressive G coefficients across occasions in the clinical ward station, reaching levels expected for high stakes evaluations in E3. The clinical ward station also showed a high internal consistency and moderate interrater reliability. The procedure station resulted in good reliability coefficients across occasions, but they were stable in all of them. Clinical ward and emergency stations produced poor interrater reliability coefficients, which impaired interpretation of final results.

Descriptors: residents as teachers; education, medical; teaching, models; simulation training; educational measurement; objective structured teaching examination; internship and residency.

1-INTRODUÇÃO:

A importância dos médicos residentes é inquestionável na atual conjuntura da saúde pública no Brasil. Os maiores hospitais universitários brasileiros têm, nesses profissionais, importante força de trabalho. Durante as últimas décadas o número de residentes cresceu, assim como o escopo de suas responsabilidades. Muitos residentes brasileiros desenvolvem seu trabalho em cenários onde há, também, alunos de graduação em treinamento. Pouco se sabe a respeito do grau de interação pedagógica entre esses diferentes entes. Desta maneira, faz-se necessário um melhor entendimento da relação entre médicos residentes e alunos do internato. Um dos pontos de interesse mais estudados dessa relação, em publicações internacionais, é a capacidade didática que médicos residentes demonstram ao orientar alunos durante suas atividades diárias.

Entendendo a relação acima descrita entre residente e aluno do internato como uma atividade de preceptoria acreditamos ser importante definir e avaliar essa atividade acadêmica. Santos e cols¹ descreveram a atividade do facilitador como um médico mais experiente e que por esse motivo auxilia no desenvolvimento profissional de alunos da graduação, médicos residentes e seus pares. Esses mesmos autores se referem ao trabalho de Vygotsky² para quem o processo de aprendizado é oriundo de fontes externas e concebido através de interações individuais com o ambiente. Conclui-se que o papel principal do facilitador, ou preceptor, é propiciar a aquisição de conhecimento teórico e habilidades ao estimular seus alunos a fazerem suas próprias descobertas.

Muitos autores reforçam que a palavra doutor deriva do latim *docere*³ que significa professor. Esta associação deixa clara a importância do ensino para todos os profissionais médicos, mesmo aqueles em treinamento. Fica explícita também a noção de que o ato de ensinar não é uma escolha para os médicos, mas sim uma imposição da profissão. Estamos envolvidos em processos pedagógicos quando discutimos um caso com colegas, ao estudar em grupo e até mesmo quando orientamos um paciente a respeito da prescrição de uma medicação.

Especialistas em educação reforçam a teoria de que a capacidade de ensino seja uma habilidade a ser aprendida e não um traço de personalidade restrito a alguns indivíduos. Para ensinar, deve-se antes aprender inúmeras habilidades e técnicas que estão muito além de, simplesmente, ter um bom conhecimento sobre a disciplina a ser ensinada. Pesquisadores norte-americanos estimam que os residentes gastam cerca de um quarto do programa de residência em atividades de ensino, apesar de terem recebido pouco ou nenhum treinamento a esse respeito⁴. Grande parte do “currículo informal”, incluindo valores profissionais, é transmitida aos alunos pelos residentes⁵.

Estudos pioneiros provenientes dos Estados Unidos e Canadá, além de demonstrar que entre 20 e 25% do tempo total de trabalho dos residentes é gasto com atividades didáticas^{4,6,7}, observaram que os estudantes de medicina atribuem aos mesmos residentes 30-85% de toda a teoria clínica adquirida durante a graduação^{8,9}. A maior parte desta troca de informações acontece durante longos períodos de trabalho e exaustivas jornadas de plantão.

A percepção de que os médicos residentes têm, de fato, trabalhado ao lado de alunos da graduação como preceptores levou à criação de programas extracurriculares especificamente desenhados para melhorar as habilidades em ensino clínico dos médicos em treinamento. Conhecidos como “Residents as Teachers” (RaTs), esses programas se popularizaram na maioria das instituições norte-americanas e atualmente estão presentes em 80% delas¹⁰. Os programas RaTs não estão restritos apenas a residências essencialmente clínicas, sendo descritos também na psiquiatria, anestesiologia, radiologia, dermatologia e outras especialidades cirúrgicas como a oftalmologia^{4,7,11-13}.

O autor desta tese, em companhia de outros colaboradores, submeteu um artigo de revisão narrativa para publicação versando especificamente sobre a experiência dos programas RaTs em outros países. O trabalho foi aceito para publicação na revista São Paulo Medical Journal após processo de revisão por pares (ANEXO A). Nesta revisão, identificamos apenas uma publicação em língua portuguesa a respeito da necessidade de se desenvolver habilidades em técnicas de ensino para médicos residentes. O trabalho que encontramos é uma dissertação de mestrado a respeito da formação docente na residência em

medicina de família e comunidade da Universidade Federal de Pernambuco¹⁴. O autor concluiu através de dados qualitativos que é possível desenvolver competências pedagógicas em médicos residentes em conformidade com as novas políticas de saúde e educação.

Devido à grande difusão dos programas RaT principalmente nos Estados Unidos e no Canadá, tornou-se necessária a criação de instrumentos válidos e confiáveis para avaliação dessas atividades. Em uma revisão crítica a respeito dos currículos em RaT, Post e cols¹⁵ descreveram as avaliações simuladas em técnicas de ensino, como o método ideal para avaliar esses programas. Conhecido como “Objective Structured Teaching Examination” (OSTE), o método utiliza o mesmo princípio das atividades simuladas para avaliação de habilidades clínicas, ou “Objective Structured Clinical Examination” (OSCEs).

Os OSCEs são muito utilizados em instituições de ensino brasileiras. Porém, não há relatos da utilização de atividades simuladas para avaliação de habilidades em ensino clínico em nosso meio. Desta forma, o autor desta tese escreveu uma nova revisão narrativa com foco nos OSTEs com a finalidade de delimitar e disseminar o tema através de uma publicação brasileira. Após processo de revisão por pares tivemos a revisão aceita para publicação na revista São Paulo Medical Journal (ANEXO B). Nesta revisão, não encontramos nenhum trabalho em língua portuguesa que utilizou o OSTE como método para avaliação de habilidades em técnicas de ensino clínico.

2OBJETIVOS

2-OBJETIVOS:

2.1-Objetivo Geral ou Primário:

Avaliar se a instrução formal de médicos residentes em técnicas de ensino, utilizando uma intervenção de curta duração, é capaz de melhorar o desempenho destes em atividades simuladas, desenvolvidas para a avaliação de habilidades em ensino clínico.

2.2-Objetivos Específicos ou Secundários:

1. Avaliar a validade e a confiabilidade do questionário utilizado para medir a aquisição de habilidades em técnicas de ensino durante as avaliações;

2. Avaliar a confiabilidade global do questionário para cada estação simulada;

3. Avaliar a confiabilidade entre os avaliadores e consistência interna dos elementos do questionário;

4. Desempenho de habilidades de ensino clínico de médicos residentes antes e após exposição a uma técnica de discussão de caso

3-REVISÃO DA LITERATURA:

3.1-História da medicina

O ensino médico está intimamente relacionado ao desenvolvimento da medicina ao longo da história. As práticas médicas iniciais datam da pré-história, 8 mil anos antes de Cristo. Conseguimos identificar os primórdios da medicina através de estudos arqueológicos que estudaram pinturas em cavernas e artefatos simbólicos. Crânios com sinais claros de trepanação, realizadas no período neolítico, sugerem que rituais médicos ocorriam desde o início da história da humanidade¹⁶.

A partir do advento da escrita, os antigos egípcios passaram a descrever procedimentos, sintomas de doenças e tratamentos instituídos para as mais diversas enfermidades. Estava em desenvolvimento, 2 mil anos antes de Cristo, o prontuário médico e o princípio da publicação em medicina. À mesma época, a anatomia humana se iniciava através dos rituais de mumificação.

Hipócrates, considerado o pai da medicina ocidental, desenvolveu teorias relacionadas aos processos de doença. Essas teorias foram usadas e estudadas em toda a Grécia antiga e permaneceram vigentes durante o Império Romano. Com a queda de Roma e o início da idade média, a medicina passou a ser monopólio das instituições religiosas. Houve, nesse período, o surgimento das universidades, porém os hospitais eram instalados em monastérios ou outras instituições religiosas.

A prática médica, durante a idade média, desenvolveu-se no mundo árabe com o refinamento das teorias de Hipócrates. Em 931 depois de Cristo, grandes hospitais estavam envolvidos no processo de treinamento e licenciatura de médicos e farmacêuticos. Seria, talvez, o embrião dos programas de residência médica caracterizado pelo treinamento em serviço.

O renascimento retomou a valorização da educação na Europa. O advento da literatura impressa iniciou uma revolução educacional em todo o mundo ocidental. Durante os séculos XVIII e XIX, com o advento da revolução industrial, houve disseminação do conhecimento científico e as universidades europeias adquiriram autonomia progressiva.

Todos esses marcos históricos nos ajudam a compreender como a medicina e principalmente a disseminação do conhecimento médico, possibilitaram o desenvolvimento humano nos dias atuais.

3.2-Ensino Médico nos Estados Unidos:

O fato de os Estados Unidos liderarem a produção científica mundial não se deu por acaso. O desenvolvimento das escolas médicas naquele país se deu de forma planejada e foi aperfeiçoado durante décadas.

Até 1910, havia 166 escolas médicas em todo o território norte-americano. A maioria delas tinha compromisso com a geração de lucro. O aprendizado se dava através da transmissão da tradição médica. Isso resultava em um número grande de profissionais sem capacitação suficiente para a prática do dia-a-dia e também para a geração de conhecimento.

O ponto de inflexão nesta trajetória que levava a resultados pífios foi a publicação do relatório Flexner¹⁷. Abraham Flexner foi um professor escolar a quem a fundação Carnegie deu a missão de avaliar a qualidade das escolas médicas dos Estados Unidos e do Canadá. Em seu relatório, publicado em 1910, Flexner fez as seguintes recomendações: (1) o curso de medicina da Universidade Johns Hopkins deveria ser o modelo a ser seguido, (2) deveria haver educação universitária prévia à admissão na faculdade de medicina, (3) as escolas médicas deveriam ser acreditadas por entidades de classe, (4) deveria haver vínculo das escolas médicas com universidades, (5) o modelo curricular deveria ser dividido entre ciências básicas e prática clínica, (6) o ensino deveria ser baseado no conhecimento científico.

Os principais resultados dessas recomendações foram: (1) encerramento das atividades de 1/3 das escolas consideradas deficitárias à época, (2) expansão da produção científica, (3) incremento da base científica na formação dos alunos de medicina, (4) início da era do cientista clínico. Considera-se que esse trabalho reconstituiu o profissionalismo médico nos Estados Unidos e no Canadá, transformando a natureza e o processo da educação médica nesses países, havendo como se sabe, enormes críticas sobre o tema¹⁷.

3.3-Os Programas de Residência Médica:

Os programas de residência médica, criados como uma oportunidade de treinamento avançado para egressos do curso de medicina em especialidades clínicas e cirúrgicas, começaram a tomar forma a partir do final do século XIX. Eram considerados como um treinamento adicional em áreas de interesse especial. Foram desenvolvidos de forma concomitante à transformação do ensino em medicina norte-americano, no mesmo movimento produzido pelo relatório Flexner. Os primeiros programas formais de residência médica foram criados por William Halsted e William Osler no Hospital Johns Hopkins.

A partir do século XX os programas formais de treinamento em serviço para médicos recém-formados se disseminaram por todos os estados norte-americanos e o modelo foi adotado por outros países, incluindo o Brasil. Atualmente, ter participado de um programa de residência confere às instituições e aos aprendizes maior qualificação, além de maiores rendimentos, admiração de seus pares e alunos da graduação para os médicos residentes.

A residência médica no Brasil iniciou-se em 1945, no Hospital das Clínicas da Universidade de São Paulo e no Hospital dos Servidores do Rio de Janeiro¹⁸. Em 1977, foi legalmente reconhecida através de decreto presidencial. Assim como nos Estados Unidos, os programas se disseminaram por todo o país. Atualmente o número de vagas cresceu, porém ainda há um déficit que se reflete em concursos muito concorridos.

3.4-A Relação Residente-Aluno:

Observações da íntima relação pedagógica entre médicos residentes e alunos da graduação remontam à década de 70^{6,19}. Desde então, essa relação passou a ser estudada de maneira mais profunda. As evidências são cada vez mais crescentes a demonstrar que os residentes têm o desejo de ensinar e que os mesmos respondem de forma positiva a programas formais para treinamento em técnicas de ensino²⁰.

A necessidade de se treinar residentes em técnicas de ensino clínico se dá por dois motivos. O volume de conhecimento que os alunos, principalmente do internato, adquirem através de seus colegas residentes é grande. O tempo que os residentes dispõem em atividades de ensino em suas grades horárias

é igualmente substancial. Os números que corroboram essas afirmativas foram citados na introdução dessa tese. Contudo, tanto os médicos residentes como os alunos de medicina recebem pouca ou nenhuma instrução a respeito do ato de ensinar^{8,21,22}.

Com essa finalidade foram criados os programas formais em técnicas de ensino para médicos residentes e alunos de medicina. Reconheceu-se de forma objetiva que existe uma necessidade que pode ser satisfeita de maneira efetiva. Mais uma vez, temos as instituições norte-americanas e canadenses na vanguarda do desenvolvimento da educação médica mundial. E, mais uma vez, devemos reconhecer que esse sucesso não se deu por acaso. A vasta literatura publicada a respeito dos programas RaTs torna explícito o trabalho árduo sobre o tema.

As recomendações atuais sobre a implementação de programas de treinamento são provenientes de várias instituições sérias. Tanto o *American College of Graduate Medical Education* (ACGME)²³ quanto o *Liaison Committee on Medical Education*²⁴ recomendam a instrução de médicos residentes em técnicas de ensino. Alguns autores consideram essas intervenções prioridades nacionais e até mesmo obrigações pessoais dos educadores em medicina³.

Podemos citar alguns motivos pelos quais esses programas conquistaram tamanho sucesso: (1) alunos do internato gostam de trabalhar com residentes, (2) existe supervisão muito próxima dos residentes em relação aos alunos, (3) residentes são fontes de influência positiva e exemplos de profissionalismo²⁵⁻²⁷. Outros motivos incluem o longo tempo que alunos e residentes passam juntos em enfermarias ou unidades de emergência, a proximidade em atividades práticas, faixas etárias similares, níveis de desenvolvimento profissional também próximos e o fato de que os residentes se sentem mais satisfeitos com o trabalho quando desenvolvem responsabilidades didáticas^{28,29}.

3.5. Intervenção Educacional - Preceptor em um Minuto (PEM) – One Minute Preceptor – (OMP)

O One Minute Preceptor (OMP) é um modelo de desenvolvimento educacional amplamente utilizado com a finalidade de melhorar técnicas de ensino. Originalmente projetado para utilização em prática ambulatorial de serviços com sobrecarga de atendimentos, o método facilita a interação entre preceptor e alunos através do uso de 5 micro habilidades. Estas habilidades incluem: (1) obter compromisso do aluno em articular seu próprio plano terapêutico e diagnóstico, (2) avaliar o conhecimento atual do aluno, (3) apresentar conceitos gerais que poderão ser usados em outros casos, (4) oferecer feedback positivo, (5) oferecer feedback construtivo. O OMP pode ser disseminado de maneira rápida e eficiente e usado com sucesso em ambientes com forte pressão de tempo, exatamente como a maioria dos serviços hospitalares brasileiros se apresentam ao residente.³⁰

3.6-A avaliação de atividades simuladas:

Atividades simuladas são utilizadas, com frequência crescente, como forte ferramenta para avaliações formativas e somativas. Desde que o neurologista Howard Barrows³¹ empregou atores treinados para avaliar as habilidades clínicas de seus residentes em um exame clínico na década de 60, o método passou a ser bastante estudado e aperfeiçoado.

Atualmente, atividades simuladas estão solidificadas, inclusive no Brasil, como instrumentos de formação e avaliação. O exame clínico, objetivo e estruturado (OSCE- “*Objective Structured Clinical Examination*”), faz parte do arsenal de métodos de avaliação em escolas médicas de todo o país. Foi construído de forma a simular uma situação real e com isso possibilitar a avaliação de habilidades clínicas de alunos do curso de medicina ou de médicos residentes. Pode ser utilizado também em concursos, na forma de provas práticas.

Partindo-se do mesmo princípio de que habilidades clínicas podem ser avaliadas por meio de simulações, foram desenvolvidas atividades, também simuladas, com a finalidade de avaliar técnicas de ensino clínico. O exame objetivo e estruturado em técnicas de ensino clínico (OSTE - “*Objective*

Structured Teaching Examination ou *Exercise* ou *Encounter*”) é uma modalidade de avaliação simulada que começou a ser desenvolvida a partir da década de 90.

Os OSTEs consistem de atividades simuladas, baseadas em ambientes de ensino, que utilizam alunos padronizados com o objetivo de fornecer feedback objetivo e imediato ao professor ou docente que participa da atividade³¹. Os cenários têm duração entre 10 a 15 minutos e podem ser elaborados para avaliar múltiplas situações que posteriormente serão abordadas em grandes grupos de discussão³³.

Durante os OSTEs, o docente, ou médico residente, tem a oportunidade de demonstrar suas habilidades de ensino e interação com estudantes padronizados, em um ambiente controlado e protegido. O OSTE é atualmente considerado o método padrão ouro para avaliação das habilidades de ensino de médicos residentes³⁴ e o instrumento com melhor acurácia para avaliar programas de desenvolvimento docente³⁵.

4-MÉTODOS:

4.1-Tipo de Estudo:

Desenvolvemos um estudo experimental, prospectivo, do tipo ensaio clínico com avaliações pré e pós intervenção, longitudinal, randomizado, “cego” quanto às datas de avaliações para os avaliadores.

4.2-Comitê de ética em pesquisa:

Aprovação da Comissão de Ética em Pesquisas da Universidade Federal de Uberlândia foi obtida antes que qualquer coleta de dados fosse iniciada (parecer número 1.894.307) (ANEXO C). Obtivemos também a aprovação da Comissão de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina da USP como instituição coparticipante da pesquisa (ANEXO D).

4.3- Critérios de Inclusão e Exclusão:

Os critérios de inclusão utilizados para recrutamento da amostra foram: médicos residentes cursando o primeiro ou segundo anos dos programas de clínica médica, cirurgia geral, dermatologia, infectologia e neurologia do Hospital de Clínicas da Universidade Federal de Uberlândia-UFU ou aqueles, cursando os mesmos programas nas unidades de origem, que estavam em estágio eletivo nesta mesma instituição à época da coleta dos dados.

Os critérios de exclusão foram: a) período de férias concomitante ou próximo ao período de coleta dos dados; b) rodízios em estágios que dificultariam a presença dos residentes nas atividades práticas; c) restrição dos médicos residentes quanto à gravação em vídeo das atividades simuladas.

4.4- Amostra:

Quarenta e nove residentes fizeram parte de uma amostra de conveniência que participou das atividades. Escolhemos tais programas de residência médica devido ao grande envolvimento de tais programas com

alunos do internato. Termo de consentimento livre e esclarecido foi obtido de todos os participantes (ANEXO E).

4.4.1- Estações de Avaliação:

A criação de enredos (*scripts*) para as estações foi o passo inicial para início das atividades. Procuramos contemplar as situações mais corriqueiras da interação aluno (interno) - residente, assim como possíveis armadilhas no tratamento de alguns temas.

Todos os residentes foram avaliados em quatro estações simuladas, nas diversas etapas do estudo. Cada estação teve duração de no máximo seis minutos. Usamos como modelo de avaliação simulada em técnicas de ensino clínico o método OSTE. Três etapas foram desenvolvidas ao longo de três semanas consecutivas. As estações simuladas consistiam de situações clínicas vivenciadas, diariamente, pelos residentes ao orientar alunos do internato na realização de procedimentos e na discussão de casos em enfermaria, ambulatório e setor de emergência (ANEXO F), garantindo desta forma os requisitos de validade. Contamos com a contribuição de alunos do curso de teatro da mesma instituição, UFU, como estudantes (internos) simulados. Todas as simulações foram gravadas em vídeo para posterior análise independente.

Alunos simulados padronizados e avaliadores independentes e treinados aumentam a fidedignidade, assim como o fato de todos os residentes serem submetidos às mesmas condições de avaliação³⁶.

4.4.2- Estudantes Padronizados:

Uma semana antes do início das atividades o autor se reuniu com a equipe de atores para esclarecer dúvidas concernentes a termos técnicos, nomenclatura de doenças e medicações e comportamentos esperados nas diferentes situações. Esta reunião teve duração de cerca de 90 minutos. Cada ator recebeu um formulário com orientações úteis (roteiro) para guiar seu comportamento de acordo com as respostas dos residentes em avaliação. As orientações foram específicas para cada uma das estações (ANEXO G).

O recrutamento dos alunos padronizados foi feito por meio de contato com a coordenação do curso de teatro da Universidade Federal de Uberlândia. Havia um grupo de alunos que tinha experiência prévia com atividades simuladas em medicina. Eles/elas participaram, como pacientes simulados, de OSCEs confeccionados para a avaliação de técnicas semiológicas de alunos da graduação.

Selecionamos atores com idade entre 21 e 35 anos e com experiência prévia em atividades simuladas. Todos passaram por treinamento prévio com duração média de 20 horas, entre preparação de figurinos, estudo dos roteiros e ensaios.

Todos os atores que participaram das atividades foram remunerados pelo trabalho, assim como receberam auxílio financeiro para alimentação e transporte para o local das simulações. As despesas foram totalmente custeadas pelo autor da pesquisa.

4.4.3- Treinamento dos Avaliadores:

As estações foram avaliadas por três indivíduos independentes, “cegados” quanto ao dia da avaliação, a sequência de exposição dos residentes às intervenções, a especialidade dos residentes e seus nomes.

Todos os três avaliadores possuíam experiência prévia em avaliação e participaram de treinamento específico antes do início da atividade desenvolvida no presente estudo. Antes de iniciar os procedimentos, os avaliadores se reuniram 2 vezes, por videoconferência, para discutir os métodos de avaliação e uniformizar a forma de pontuação dos itens da escala que seria usada. Essas duas primeiras reuniões tiveram duração de cerca de 90 minutos. Após as reuniões iniciais, os avaliadores se reuniram mais 4 vezes com duração também de 90 minutos. Tais reuniões, também por videoconferência, antecederam o início das avaliações de cada uma das 4 estações. Para padronização e treinamento dos avaliadores foram usados vídeos de residentes, os quais foram excluídos da análise final. Foram analisados 03 vídeos por estação, num total de 12 vídeos de treinamento, totalizando 9 horas de treinamento dos avaliadores.

Nenhum dos avaliadores recebeu financiamento para as atividades. A tabela 1 descreve as principais características dos indivíduos que avaliaram as estações.

Tabela1- Características dos avaliadores das estações OSTE

Avaliador	Profissão	Idade (anos)	Tempo de Formado (anos)	Experiência em ensino na Graduação e Residência Médica (anos)
Avaliador 1	Médico	43	20	12
Avaliador 2	Médico	30	07	4
Avaliador 3	Médico	39	15	10

Tabela 1. Perfil dos médicos avaliadores independentes, das estações de OSTE. Dois da região nordeste e um da região sudeste

4.4.4- Instrumento de Avaliação:

Para a avaliação foi empregada a versão validada para o português, por nossa equipe, do "*Stanford Faculty Development Program*" – SFDP26 (ANEXO H)³⁷. O instrumento, desenvolvido originalmente nos Estados Unidos, avalia as seguintes habilidades em técnicas de ensino: (1) capacidade de estabelecer ambiente propício para o ensino; (2) controle da sessão de ensino; (3) comunicação dos objetivos de aprendizado; (4) capacidade de facilitar a compreensão e retenção de conteúdo; (5) avaliação do conhecimento prévio do aluno; (6) promover aprendizado auto direcionado; (7) capacidade de oferecer "feedback". Utilizamos os subitens da escala que refletiam, com melhor acurácia, as situações abordadas em cada estação. O instrumento SFDP26 utiliza escala Likert de pontuação para cada competência avaliada, sendo que as pontuações variam de 1 a 5 (1= descordo totalmente, 2= discordo, 3= indiferente ou não é possível avaliar, 4= concordo, mas com alguma ressalva, 5= concordo totalmente).

O SFDP26 foi traduzido para o português pelo primeiro autor da pesquisa. A seguir um professor de língua inglesa com cidadania americana retraduziu para o inglês e as inconformidades (poucas) foram revistas. O instrumento final, em português, foi apresentado a um painel de especialistas

brasileiros (médicos com amplo e profundo conhecimento de residência médica, das 5 regiões do país). Tais especialistas, 15 indivíduos, analisaram a pertinência de cada tópico, também em escala Likert. Nenhuma questão recebeu avaliação igual ou inferior a indiferente, por mais de 50% dos painelistas.

Todas as avaliações de cada estação, nas diferentes etapas da pesquisa, receberam também um escore de 0 a 100 que foi anotado ao final do questionário contendo o instrumento SFDP26. Esse escore global atribuía uma nota final ao desempenho do MR naquela estação. Escores mais próximos de 0 denotavam a performance de um professor limitado, enquanto aqueles próximos de 100, um professor ideal.

4.4.5- Intervenções:

O objetivo era que tivéssemos avaliações pré e pós intervenções para que fosse possível medir o efeito da intervenção principal sobre o comportamento dos médicos residentes nas atividades práticas.

A intervenção principal baseou-se na apresentação da técnica de ensino “preceptoria em um minuto” (PEM)³⁷. A intervenção controle baseou-se na apresentação de técnicas para comunicação de más notícias (CMN). Ambas intervenções foram oferecidas a todos os médicos residentes de acordo com processo de randomização eletrônica realizado após a primeira atividade prática realizada para coleta de dados.

Cada intervenção durou em média 50 minutos e foi ofertada mais de uma vez, no mesmo dia, para que todos pudessem participar. As intervenções consistiram de uma aula expositiva utilizando recursos visuais de *Power Point* e a reprodução de uma animação descrevendo os principais tópicos das intervenções. A animação correspondente à intervenção CMN foi retirada do site da *Irish Hospice Foundation* (<https://hospicefoundation.ie/education-training/video-wall/communication/>) e à correspondente ao PEM foi retirada de um canal do *youtube* da Universidade de Columbia nos EUA - The Virginia Apgar Academy of Medical Educators (<https://www.youtube.com/watch?v=hmKvei3thwQ>). Durante as intervenções os

residentes eram encorajados, ao longo das discussões, a compartilhar suas experiências relacionadas aos temas tratados.

4.4.6- Etapas do Estudo:

A primeira etapa do estudo consistiu na avaliação de todos os médicos residentes (MR), que passaram pelas quatro estações simuladas. Antes de iniciar cada avaliação, os residentes podiam ler um breve resumo (ANEXO I) a respeito das tarefas a serem realizadas e eram informados do tempo que teriam para desenvolver as discussões e orientações aos alunos simulados. Evidentemente que os residentes desconheciam serem os alunos, simulados. Compareceram quarenta e oito residentes, totalizando 192 avaliações. Esta etapa foi considerada o tempo zero, sem intervenção – **primeira etapa**, ou etapa pré-intervenção. O desempenho dos residentes, durante as orientações simuladas, foi gravado em vídeo, com o devido consentimento de cada participante.

Após a primeira etapa do estudo, os residentes foram eletronicamente randomizados para a ordem em que receberiam as duas intervenções. Um grupo foi inicialmente submetido a um treinamento baseado no método de ensino PEM, enquanto que o outro grupo recebeu intervenção baseada no treinamento em comunicação de más notícias - CMN. A instrução PEM foi considerada como a intervenção principal, enquanto a instrução CMN foi considerada como o controle. As intervenções ocorreram após a primeira e a segunda etapa de avaliações, na mesma semana, porém em dias diferentes.

Após as intervenções os residentes voluntários da pesquisa foram novamente expostos às situações (estações), denominada **segunda etapa**. O intervalo entre a primeira e a segunda etapa foi de 7 dias. As estações foram as mesmas, assim como a distribuição dos atores. Na segunda etapa compareceram 41 residentes e produzimos 164 vídeos para avaliação. Na mesma semana da segunda etapa os residentes foram submetidos às intervenções CMN e PEM. Aqueles submetidos à intervenção PEM na semana anterior foram desta vez submetidos à intervenção CMN e todos os que foram submetidos à intervenção CMN na semana anterior foram submetidos à intervenção principal (cruzamento dos treinamentos).

Sete dias depois iniciamos a **terceira etapa** de coleta dos dados, com o mesmo rigor das etapas anteriores. Compareceram 35 residentes com a produção de 140 vídeos. Com esta etapa encerramos as avaliações (Figura 1).

4.4.7- Perdas de amostra:

Houve perda de indivíduos tanto nas etapas quanto nas intervenções. Todas as perdas de amostra foram devidas a conflito de agenda com as atividades da residência médica ou com atividades fora da instituição. Não houve relato de absenteísmo por doença ou recusa em participar das atividades. Em todo o curso do estudo não houve pré-requisito de assiduidade para participação nas etapas subsequentes de avaliações ou intervenções. Ao final da coleta de dados contabilizamos 24 indivíduos que participaram de todas as 3 etapas e das 2 intervenções propostas.

4.4.8- Registro em vídeo das atividades:

Todas as três etapas das avaliações foram registradas em vídeo usando *smartphones*, tablets e câmeras fotográficas. Os arquivos em vídeo foram mantidos gravados em um disco rígido externo e na memória de um computador. Foram selecionadas 444 gravações que foram codificadas (apenas o autor da tese conhece os códigos) para posterior avaliação. Excluimos aquelas com duração inferior a 1 minuto ou com erros de reprodução que inviabilizariam sua avaliação.

As etapas do estudo e as intervenções, assim como o número de participantes em cada uma delas estão ilustradas na **figura 1**.

Figura 1- Seqüências de avaliações e intervenções, datas das atividades, grupos randomizados, amostra em cada etapa/intervenção, principais motivos para perda de amostra e amostragem final após todas as etapas e intervenções.

5- ANÁLISE ESTATÍSTICA:

Os dados sociodemográficos foram avaliados por meio de análise de frequência. Incluímos como dados sociodemográficos a idade dos médicos residentes, o tempo em meses desde a formatura até a atividade desenvolvida, a instituição de origem de cada participante e o programa de residência que cada um deles cursa no momento. A análise comparativa da idade e do tempo da graduação até o início das atividades entre indivíduos do sexo masculino e feminino foi feita através do teste t de *student*.

As análises de confiabilidade foram integralmente feitas pela teoria da generalização (teoria G). Seguimos a teoria de Brennan³⁹ e nos orientamos pela publicação de Bloch e cols⁴⁰ para os cálculos da confiabilidade geral de cada estação, confiabilidade interobservador e consistência interna das escalas de avaliação. Os cálculos foram feitos com o programa “G_String” que pode ser baixado de forma gratuita no endereço http://fhspcrd.mcmaster.ca/g_string/index.html.

Devido à necessidade de bancos de dados completos nos cálculos de confiabilidade, usamos um método de imputação para lidar com informações perdidas nas avaliações das diferentes estações. Utilizamos o ambiente estatístico R, versão 3.3.0 (*The R Foundation for Statistical Computing Platform*), com o pacote MissForest. Orientamo-nos através pela publicação de Stekhoven e cols⁴¹ para as análises que utilizaram imputações de dados. Descrevemos também, para cada estação, a porcentagem de dados falsamente classificados durante os procedimentos de imputação.

Consideramos todos os valores atribuídos a cada residente pelos avaliadores como variáveis categóricas ordinais. Para fins de comparações, utilizamos a somatória das pontuações para cada item do instrumento de avaliação. Foram realizadas comparações de desempenho entre os dias de avaliação para cada avaliador, individualmente e, para a somatória das avaliações de todos os avaliadores. Assumimos como hipótese nula que os desempenhos dos residentes, nos três dias de avaliação, eram iguais.

A análise de variância por postos de Friedman com dois fatores foi utilizada para as comparações de desempenho dos médicos residentes, nas diferentes estações, entre os 3 dias de avaliações. Para os cálculos do teste de Friedman também utilizamos o ambiente estatístico R. Análise *pós hoc* de Nemenyi foi utilizada para comparação entre os seguintes pares de grupos: a) Dia 1 (D1) e Dia 2 (D2); b) Dia 1 e Dia 3 (D3); c) Dia 2 e Dia 3. Para a análise *pós hoc* utilizamos o pacote PMCMR do ambiente R.

Utilizamos o programa Microsoft Excel 2016 para confecção dos bancos de dados e o ambiente R para análise dos dados e geração dos gráficos. Foram considerados significativos valores de $p < 0,05$.

6 RESULTADOS

6- RESULTADOS:

6.1- Dados Sociodemográficos:

Os dados sociodemográficos da amostra de MR que participaram da análise final da pesquisa está ilustrada na tabela 2. Não foram encontradas diferenças estatisticamente significativas quando comparamos residentes do sexo masculino e feminino quanto à idade ($p=0,52$) e quanto ao tempo de formatura até o início das atividades de avaliação ($p=0,59$).

Tabela 2- Dados sociodemográficos dos médicos residentes que participaram do estudo nas 03 etapas (N=24)

Sexo	Masculino	Feminino	Valor de p
N (%)	11 (45,83)	13 (54,17)	-
Idade*	28,36 (2)	27,62 (3,5)	0,52
Média ($\pm$ DP)			
Tempo de formado (meses)†	29,91 (12,8)	34,77 (28,7)	0,59
Média ($\pm$ DP)			
Instituição de origem**(n)	UFU (6)	UFU (8)	
	UNIPAC (1)	ICSMOC (1)	
	UNIUBE (1)	UFAC (1)	
	UFJF (2)	UFGD (1)	-
	UFTM (1)	UFT (1)	
		UNAERP (1)	
Programa de Residência (n)	Clínica Médica R1 (5)	Clínica Médica R1 (5)	
	Clínica Médica R2 (4)	Clínica Médica R2 (4)	
	Infectologia R1 (2)	Clínica Cirúrgica R2	-
	Neurologia R1 (1)	(2)	
	Infectologia R1 (1)	Dermatologia R1(1)	
		Dermatologia R2 (1)	

Tabela 2: *medida em anos†medida em meses da data de formatura até o início das avaliações. **Instituições: UFU: Universidade Federal de Uberlândia;UFTM: Universidade Federal do Triângulo Mineiro;UFJF: Universidade Federal de Juiz de Fora; UNIUBE: Universidade de Uberaba; UNIPAC: Universidade Presidente Antônio Carlos; Instituto de Ciências da Saúde de Montes Claros; UNAERP: Universidade de Ribeirão Preto; UFT: Universidade Federal de Tocantins; UFGD: Universidade Federal da Grande Dourados; UFAC: Universidade Federal do Acre

6.2- Estação Enfermaria:

6.2.1- Perdas na amostra:

Houve perdas na amostra da estação enfermaria. Essa perda foi calculada em 9,23% de todos os dados da estação. A figura 2 ilustra a porcentagem de dados perdidos segundo cada variável e a figura 3 demonstra a distribuição das perdas no banco de dados.

Figura 2- porcentagem de dados faltantes em cada variável analisada da estação enfermaria; no eixo y observamos a porcentagem de dados faltantes para cada variável (onde 0,2=20%), no eixo x observamos cada variável analisada na estação enfermaria (onde AE= ambiente de ensino, CR= promoção de compreensão e retenção, AD=promoção de aprendizado auto direcionado, CO= comunicação dos objetivos e FB=feedback). Os números 1,2 e 3 antes das abreviações se referem aos dias de avaliação, os números após as abreviações se referem aos itens da mesma habilidade de ensino avaliada e as letras F, G e K no final das variáveis se referem à inicial do nome do avaliador). Total de variáveis da estação: 171. A escala para a estação enfermaria possuía 19 itens divididos em 5 habilidades que foram examinados por 3 avaliadores independentes e em 3 ocasiões distintas.

Figura 3- Distribuição dos dados faltantes no banco de dados da estação enfermaria. Os dados faltantes são representados pelas barras verticais em amarelo e os dados completos em azul. As colunas representam as variáveis da estação enfermaria (onde AE=ambiente de ensino, CR=promoção de compreensão e retenção, AD=promoção de aprendizado auto direcionado, CO=comunicação dos objetivos e FB=feedback). Os números 1,2 e 3 antes das abreviações se referem aos dias de avaliação. Os números após as abreviações se referem aos itens da mesma habilidade de ensino avaliada e as letras F, G e K no final das variáveis se referem à inicial do nome do avaliador). As linhas representam a porcentagem de dados faltantes conforme padrões específicos de perda das variáveis (por exemplo, a última linha demonstra que houve dados faltantes em 3 variáveis específicas em 12,5% dos casos).

6.2.2- Imputação de dados:

A imputação de dados para análise estatística da estação enfermaria resultou em uma proporção de dados falsamente classificados (PFC) calculada em 0,35 pela estimativa de erro do método MissForest. O método produz bons resultados com valores próximos de 0 e resultados ruins com valores em torno de 1 (hum).

6.2.3- Confiabilidade dos dados:

Os cálculos de confiabilidade, por meio da teoria G, para a estação enfermaria resultaram em um coeficiente geral para toda a estação de 0,425. A consistência interna foi calculada em 0,913 e a confiabilidade entre os avaliadores resultou em um coeficiente de 0,595. Foram também calculados os mesmos coeficientes para os diferentes dias de avaliação, como se fossem etapas independentes. Para **o primeiro dia de avaliações** (basal), os coeficientes gerais, consistência interna e confiabilidade entre os avaliadores foram de, respectivamente, 0,545; 0,775; 0,392. Para **o segundo dia de avaliações** (primeira intervenção) os mesmos coeficientes foram calculados em: 0,702; 0,890; 0,495. **No terceiro dia** (segunda intervenção – cruzamento) os coeficientes foram calculados em: 0,860; 0,921; 0,698. A tabela 3 resume todos os cálculos de coeficientes de confiabilidade para a estação enfermaria.

Tabela 3*- Coeficientes de confiabilidade para a estação enfermaria

Avaliação	Coeficiente G**	Consistência Interna**	Confiabilidade entre avaliadores –score total†
Enfermaria total	0,425	0,913	0,595
Enfermaria D1	0,545	0,775	0,392
Enfermaria D2	0,702	0,890	0,495
Enfermaria D3	0,860	0,921	0,698

Tabela 3. *Resumo dos cálculos de coeficientes de confiabilidade para a estação enfermaria total e para os três diferentes dias de avaliação (D1, D2 e D3). **coeficiente relativo. †coeficiente absoluto

6.2.4- Teste de hipóteses e comparação por pares:

Considerando a somatória dos escores de todos os avaliadores foram realizados testes de hipóteses utilizando o banco de dados com perdas de amostras assim, como o banco de dados com imputações.

As avaliações globais de cada residente com escores de desempenho entre 0 e 100 não foram submetidas a imputações de dados.

Considerando o banco de dados sem imputações foi observada diferença estatisticamente significativa entre os 3 grupos de avaliação ($p=0,009$). Análise pós-hoc de Nemenyi para comparação de pares evidenciou diferença entre D1 e D2 apenas ($p= 0,011$). A tabela 4 ilustra um resumo das medianas dos escores obtidos nos diferentes dias de avaliações, percentis 25 e 75 (P25-P75) assim como os valores de p obtidos. A figura 4 representa, graficamente, por *boxplot*, o desempenho dos MR nos dias de avaliação.

Tabela 4*: Teste de Hipóteses da escala Likert[†]: somatório dos avaliadores para a estação enfermagem (planilha sem imputação de dados)

Dia	Mediana	P25-P75	Valor de p^{**} (análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	109,5	96,75-131	0,009404
D2	142,5	121,5-172,5	D1-D2=0,011
D3	137,5	103-188,2	D1-D3=0,055 D2-D3=0,832

Tabela 4. * Comparação das medianas e P25-P75 dos escores de desempenho dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/D3- pós intervenção). † pontuação mínima: 57; pontuação máxima: 285. **Análise de variância por postos de Friedman com 2 fatores

Figura 4- Representação em *box plot* da Escala Likert: Somatória dos avaliadores para a Estação Enfermaria (sem imputações). A figura demonstra a análise estatística com a amostra original da pesquisa, sem a imputação de dados

Figura 4. Representação gráfica, em *boxplot*, da comparação das medianas de escores de desempenho dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2, D3-pós intervenção). O *box-plot* é formado pelo primeiro (linha inferior da caixa) e terceiro quartil (linha superior da caixa) e pela mediana ou segundo quartil (linha média da caixa). As hastes inferiores e superiores se estendem, respectivamente, do quartil inferior até o menor valor não inferior ao limite inferior e do quartil superior até o maior valor não superior ao limite superior. O ponto (aberto) em D1 corresponde a um ponto fora da curva (*outlier*)

Considerando o banco de dados com imputações foi observada diferença estatisticamente significativa entre os 3 grupos de avaliação ($p=0,0015$). Análise *pós-hoc* de Nemenyi para comparação de pares evidenciou diferença entre D1 e D2 ($p= 0,0015$) e entre D1 e D3 ($p= 0,025$). A tabela 5 ilustra um resumo das medianas dos escores obtidos nos diferentes dias de avaliações assim como os valores de p obtidos. A figura 5 representa, graficamente, por *boxplot*, o desempenho dos MR nos dias de avaliação, com imputação de dados.

Tabela 5*- Teste de Hipóteses da escala Likert[†]: somatória dos avaliadores para a estação enfermaria (planilha com imputação de dados)

Dia	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	110,5	101,5-132,2	0,001503
D2	147	130,2-174,5	D1-D2=0,0015
D3	142	113-189,2	D1-D3=0,0254 D2-D3=0,6618

Tabela 5. Comparação das medianas e P25-P75 dos escores de desempenho dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/D3- pós intervenção) † pontuação mínima: 57; pontuação máxima: 285 ** Análise de variância por postos de Friedman com 2 fatores

Figura 5- Escala Likert: Somatório dos Avaliadores para a estação enfermaria (com imputações) em *boxplot*

Figura 5. Representação gráfica, em *boxplot*, após imputação dos dados faltantes na amostra original dos escores de desempenho dos médicos residentes, nos três diferentes momentos de avaliação (D1 – basal, D2/D3-pós intervenção). Houve diferença significativa ao se comparar D1 com D2 e D3.

A avaliação estatística do desempenho dos MR por meio da avaliação global revelou diferença significativa entre os dias de avaliação ($p < 0,001$). Análise *pós-hoc* de Nemenyi para comparação de pares evidenciou diferença entre D1 e D2 ($p < 0,001$) e entre D1 e D3 ($p = 0,017$). A tabela 6 apresenta resumo das medianas dos escores obtidos nos diferentes dias de avaliações, assim como os valores de p obtidos. A figura 6 representa, em *box plot*, o desempenho dos MR nos dias de avaliação.

Tabela 6*- Teste de Hipóteses da escala Global[†] por avaliadores para a estação enfermaria

Dia	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	82,5	72,5-107	<0,001 D1-D2<0,001 D1-D3=0,017 D2-D3=0,319)
D2	129,5	108,2-170,5	
D3	127	93,5-169,2	

Tabela 6. * Comparação das medianas e P25-P75 dos escores de desempenho global dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/D3-pós intervenção) † pontuação mínima: 0; pontuação máxima: 300 ** Análise de variância por postos de Friedman com 2 fatores

Figura 6- Escala Global: Somatória dos avaliadores para a estação enfermaria representada em *boxplot*

Figura 6. Demonstração da evolução dos escores na avaliação global dos médicos residentes nos diferentes dias de avaliação (D1, D2 e D3), com diferença significativa entre D1 e D2/D3.

A análise estatística através dos itens da escala SFDP26 para a estação enfermaria revelou diferença estatisticamente significativa entre os dias de avaliação para as seguintes competências: ambiente de ensino ($p < 0,001$), aprendizado auto direcionado ($p < 0,001$) e feedback ($p = 0,003$). Não foram observadas diferenças para as competências de ensino: promoção da

compreensão e retenção ($p = 0,62$), comunicação dos objetivos ($p = 0,1143$). Para a análise de significância estatística dos itens da escala de avaliação foi utilizado apenas o banco de dados sem imputações.

6.3- Estação Procedimento:

6.3.1- Perdas na amostra:

Houve perdas na amostra da estação procedimento. Essa perda foi calculada em 13% de todos os dados da estação. A figura 7 ilustra a porcentagem de dados perdidos segundo cada variável e a figura 8 demonstra a distribuição das perdas no banco de dados.

Figura 7- Porcentagem de dados faltantes em cada variável analisada da estação procedimento

Figura 7 -porcentagem de dados faltantes em cada variável analisada da estação procedimento; no eixo y observamos a porcentagem de dados faltantes para cada variável (onde 0,2=20%), no eixo x observamos cada variável analisada na estação procedimento (onde AE= ambiente de ensino, CR= promoção de compreensão e retenção, CO= comunicação dos objetivos). Os números 1,2 e 3 antes das abreviações se referem aos dias de avaliação. Os números após as abreviações se referem aos itens da mesma habilidade de ensino avaliada e as letras F, G e K no final das variáveis se referem à inicial do nome do avaliador). Total de variáveis da estação: 90 (a escala para a estação enfermagem possuía 10 itens divididos em 3 habilidades que foram examinados por 3 avaliadores independentes e em 3 ocasiões distintas)

Figura 8- Distribuição dos dados faltantes no banco de dados da estação procedimento

Figura 8- Distribuição dos dados faltantes no banco de dados da estação procedimento. Os dados faltantes são representados pelas barras verticais em amarelo e os dados completos em azul. As colunas representam as variáveis da estação enfermária (onde AE= ambiente de ensino, CR= promoção de compreensão e retenção, CO= comunicação dos objetivos). Os números 1,2 e 3 antes das abreviações se referem aos dias de avaliação. Os números após as abreviações se referem aos itens da mesma habilidade de ensino avaliada e as letras F, G e K no final das variáveis se referem à inicial do nome do avaliador). As linhas representam a porcentagem de dados faltantes conforme padrões específicos de perda das variáveis (por exemplo, a última linha demonstra que em 41,7% das variáveis não houve dados faltantes).

6.3.2- Imputação de dados:

A imputação de dados para análise estatística da estação procedimento resultou em uma PFC calculada em 0,35 através da estimativa de erro do método MissForest.

6.3.3- Confiabilidade dos dados:

Os cálculos de confiabilidade, através de teoria G, para a estação procedimento resultaram em um coeficiente geral para toda a estação de 0,398. A consistência interna foi calculada em 0,726 e a confiabilidade entre os avaliadores resultou em um coeficiente de 0,772. Foram também calculados os mesmos coeficientes para os diferentes dias de avaliação, como se fossem etapas independentes. Para o primeiro dia de avaliações, os coeficientes gerais, consistência interna e confiabilidade entre os avaliadores foram de, respectivamente, 0,683; 0,599; 0,662. Para o segundo dia de avaliações os mesmos coeficientes foram calculados em: 0,681; 0,703; 0,599. No terceiro dia os coeficientes foram calculados em: 0,670; 0,629; 0,352. A tabela 7 resume todos os cálculos de coeficientes de confiabilidade para a estação procedimento.

Tabela 7*- Coeficientes de confiabilidade para a estação procedimento

Avaliação	Coeficiente G**	Consistência Interna**	Confiabilidade entre avaliadores –escore total†
Procedimento total	0,398	0,726	0,772
Procedimento D1	0,683	0,599	0,662
Procedimento D2	0,681	0,703	0,599
Procedimento D3	0,670	0,629	0,352

Tabela 7. * Resumo dos cálculos de coeficientes de confiabilidade para a estação procedimento total e para os três diferentes dias de avaliação (D1, D2 e D3)*coeficiente relativo. †coeficiente absoluto

6.3.4- Teste de hipóteses e comparação por pares:

Considerando a somatória dos escores de todos os avaliadores foram realizados testes de hipóteses utilizando o banco de dados com perdas de amostras assim como o banco de dados com imputações.

As avaliações globais de cada residente com escores de desempenho entre 0 e 100 não foram submetidas a imputações de dados.

Considerando o banco de dados sem imputações foi observada diferença estatisticamente significativa entre os 3 grupos de avaliação ($p < 0,001$). Análise *pós-hoc* de Nemenyi para comparação de pares evidenciou diferença entre D1 e D2 ($p = 0,0012$) e entre D1 e D3 ($p = 0,0043$). A tabela 8 ilustra um resumo das medianas dos escores obtidos nos diferentes dias de avaliações, P25-P75 assim como os valores de p obtidos. A figura 9 representa, graficamente em *boxplot*, o desempenho dos MR nos dias de avaliação.

Tabela 8*: Teste de Hipóteses da escala Likert[†]: somatória dos avaliadores para a estação procedimento (planilha sem imputação de dados)

Dia	Mediana	P25-P75	Valor de p^{**} (análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	99	85-111,2	<0,001 D1-D2=0,0012 D1-D3=0,0043 D2-D3=0,9307
D2	119	106,2-131,2	
D3	119,5	102,8-127,2	

Tabela 8. * Comparação das medianas e P25 - P75 dos escores de desempenho dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/D3- pós intervenção). † pontuação mínima: 30; pontuação máxima: 150 ** Análise de variância por postos de Friedman com 2 fatores.

Figura 9- Escala Likert: Somatória dos Avaliadores para a estação procedimento (sem imputações) representada em *boxplot*.

Figura 9- Análise dos dados da amostra original na estação procedimento, sem a imputação de dados.

Considerando o banco de dados com imputações foi observada diferença estatisticamente significativa entre os 3 grupos de avaliação ($p < 0,001$). Análise *pós-hoc* de Nemenyi para comparação de pares evidenciou diferença entre D1 e D2 ($p = 0,0043$) e entre D1 e D3 ($p < 0,001$). A tabela 9 ilustra o resumo das medianas dos escores obtidos nos diferentes dias de avaliações assim como os valores de p obtidos. A figura 10 representa, graficamente, em *boxplot*, o desempenho dos MR nos dias de avaliação.

Tabela 9*-Teste de Hipóteses da escala Likert[†]: somatória dos avaliadores para a estação procedimento (planilha com imputação de dados)

Dia	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	100,5	91,5-111,2	<0,001 D1-D2=0,0043 D1-D3<0,001 D2-D3=0,3186
D2	119	110,2-132	
D3	123,5	117-134,5	

Tabela 9. * Comparação das medianas e P25-P75 dos escores de desempenho dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/D3- pós intervenção) [†] pontuação mínima: 30; pontuação máxima: 150 ** Análise de variância por postos de Friedman com 2 fatores

Figura 10- Escala Likert: Somatória dos Avaliadores para a estação procedimento (com imputações) através de *boxplot*

Figura 10 - Análise dos dados da amostra na estação procedimento, após imputação dos dados faltantes na amostra original

A avaliação estatística do desempenho dos MR por meio da avaliação global revelou diferença estatisticamente significativa entre os dias de avaliação ($p=0,009$). Análise *pós-hoc* de Nemenyi para comparação de pares evidenciou diferença entre D1 e D2 ($p=0,031$) e entre D1 e D3 ($p=0,021$). A tabela 10 contém o resumo das medianas dos escores obtidos nos diferentes dias de avaliações, assim como os valores de p obtidos. A figura 11 representa, graficamente, em *boxplot*, o desempenho dos MR nos dias de avaliação.

Tabela 10*- Teste de Hipóteses da escala Global[†] por avaliadores para a estação procedimento

Dia	Mediana	P25-P75	Valor de p^{**} (análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	120	93,25-149,2	0,009
D2	181,5	139-235,8	D1-D2=0,031
D3	195	75-238,5	D1-D3=0,021 D2-D3=0,989

Tabela 10. * Comparação das medianas e P25-P75 dos escores de desempenho global dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/ D3-pós intervenção) † pontuação mínima: 0; pontuação máxima: 300** Análise de variância por postos de Friedman com 2 fatores

Figura 11- Escala Global: Somatória dos Avaliadores para a estação procedimento através de *boxplot*

Figura 11. Demonstração da evolução dos escores na avaliação global dos médicos residentes nos diferentes dias de avaliação (D1, D2 e D3) na estação procedimento, com diferença significativa após as intervenções.

A análise estatística através dos itens da escala SFDP26 para a estação procedimento revelou diferença estatisticamente significativa entre os dias de avaliação para as seguintes competências: promoção da compreensão e retenção ($p < 0,001$), controle da sessão ($p = 0,005$), comunicação dos objetivos ($p < 0,001$). Não foram observadas diferenças para competência ambiente de ensino ($p = 0,3679$). Para a análise de significância estatística dos itens da escala de avaliação foi utilizado apenas o banco de dados sem imputações.

6.4- Estação Emergência:

6.4.1- Perdas na amostra:

Houve perdas na amostra da estação emergência. Essa perda foi calculada em 17,53% de todos os dados da estação. A figura 12 ilustra a porcentagem de dados perdidos segundo cada variável e a figura 13 a distribuição das perdas no banco de dados.

Figura 12- Porcentagem de dados faltantes em cada variável analisada da estação emergência; no eixo y observamos a porcentagem de dados faltantes para cada variável (onde 0,2 = 20%), no eixo x observamos cada variável analisada na estação procedimento (onde AE = ambiente de ensino, CR = promoção de compreensão e retenção, CO = comunicação dos objetivos). Os números 1, 2 e 3 antes das abreviações se referem aos dias de avaliação. Os números após as abreviações se referem aos itens da mesma habilidade de ensino avaliada e as letras F, G e K no final das variáveis se referem à inicial do nome do avaliador). Total de variáveis da estação: 153(a escala para a estação emergência contém 17 itens, divididos em 4 habilidades, que foram examinados por 3 avaliadores independentes e em 3 ocasiões distintas)

6.4.2- Imputação de dados:

A imputação de dados para análise estatística da estação emergência resultou em uma PFC calculada em 0,39 através da estimativa de erro do método MissForest.

6.4.3- Confiabilidade dos dados:

Os cálculos de confiabilidade, através de teoria G, para a estação emergência resultaram em um coeficiente geral para toda a estação de 0,459. A consistência interna foi calculada em 0,926 e a confiabilidade entre os avaliadores resultou em um coeficiente de 0,189. Foram também calculados os mesmos coeficientes para os diferentes dias de avaliação, como se fossem etapas independentes. Para o primeiro dia de avaliações, os coeficientes gerais, consistência interna e confiabilidade entre os avaliadores foram de, respectivamente, 0,656; 0,903; 0,264. Para o segundo dia de avaliações os mesmos coeficientes foram calculados em: 0,839; 0,920; 0,404. No terceiro dia os coeficientes foram calculados em: 0,566; 0,854; 0,076. A tabela 11 resume todos os cálculos de coeficientes de confiabilidade para a estação emergência.

Tabela 11*- Coeficientes de confiabilidade para a estação emergência

Avaliação	Coeficiente G**	Consistência Interna**	Confiabilidade entre avaliadores –score total†
Emergência total	0,459	0,926	0,189
Emergência D1	0,656	0,903	0,264
Emergência D2	0,839	0,920	0,404
Emergência D3	0,566	0,854	0,076

Tabela 11. * Resumo dos cálculos de coeficientes de confiabilidade para a estação emergência total e para os três diferentes dias de avaliação (D1, D2 e D3) *coeficiente relativo †coeficiente absoluto

6.4.4- Teste de hipóteses e comparação por pares:

6.4.4.1- Comparação dos escores entre os avaliadores nas 3 etapas do estudo para a estação emergência:

Observou-se que um dos avaliadores da estação emergência produziu escores sistematicamente diferentes dos outros dois. Essa observação é corroborada pelos baixos coeficientes de correlação entre avaliadores observados na tabela 11. Foi realizada uma comparação entre os escores de cada avaliador nas diferentes etapas da pesquisa com a finalidade de identificar se um avaliador que produziu respostas diferentes dos demais.

Foi utilizado o banco de dados sem imputações para comparação dos escores produzidos pelos avaliadores, nos diferentes dias. No primeiro dia de avaliações foi observado que os avaliadores produziram resultados diferentes ($p < 0,001$). Análise *pós-hoc* de Nemenyi evidenciou diferenças entre os avaliadores 1 e 3 ($p < 0,001$) e entre os avaliadores 2 e 3 ($p = 0,005$). No segundo dia de avaliações observou-se a ocorrência de diferença entre os avaliadores ($p < 0,001$). Novamente as diferenças foram encontradas nas comparações entre avaliadores 1 e 3 ($p = 0,0087$) e entre os avaliadores 2 e 3 ($p < 0,001$). No terceiro dia de avaliações os escores também foram diferentes ($p < 0,001$), porém observou-se diferenças apenas entre os avaliadores 2 e 3 ($p < 0,001$). As tabelas 12, 13 e 14 ilustram as medianas dos escores de cada avaliador nas 3 etapas da pesquisa, assim como P25-P75, o valor de p para a comparação total e os valores de p para a análise *pós-hoc* para comparação por pares. As figuras 14, 15 e 16 representam, graficamente, em *boxplot*, a comparação dos escores produzidos por cada avaliador nas 3 etapas da pesquisa.

Tabela 12*- Teste de Hipóteses da escala Likert[†] para comparação dos escores produzidos pelos avaliadores no primeiro dia de avaliações (D1) da estação emergência

Avaliador	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
Avaliador 1 (Aval 1)	40,5	34,5-50,75	Aval 1-Aval 2= 0,62 Aval 1-Aval 3<0,001 Aval 2-Aval 3=0,005
Avaliador 2 (Aval 2)	44	34-57,25	
Avaliador 3 (Aval 3)	71,5	40-82	

Tabela 12. * Comparação das medianas e P25-P75 dos escores produzidos por cada avaliador na primeira etapa da pesquisa (D1) †pontuação mínima: 17; pontuação máxima: 85** Análise de variância por postos de Friedman com 2 fatores. Houve diferença significativa entre o avaliador 3 e os outros dois.

Figura 14- Comparação, em *boxplot*, dos escores produzidos por cada avaliador na primeira etapa da pesquisa (D1), na estação emergência

Figura 14. AVAL1= avaliador 1; AVAL2= avaliador 2; AVAL3= avaliador 3. Houve diferença significativa entre o avaliador 3 e os outros dois.

Tabela 13*- Teste de Hipóteses da escala Likert[†] para comparação dos escores produzidos pelos avaliadores no segundo dia de avaliações (D2) da estação emergência

Avaliador	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
Avaliador 1 (Aval 1)	58	45,75-73	<0,001 Aval 1-Aval 2= 0,14 Aval 1-Aval 3=0,009 Aval 2-Aval 3<0,001
Avaliador 2 (Aval 2)	48	34,75-60,5	
Avaliador 3 (Aval 3)	76,5	69,75-85	

Tabela 13. * Comparação das medianas e P25-P75 dos escores produzidos por cada avaliador na segunda etapa da pesquisa (D2) †pontuação mínima: 17; pontuação máxima: 85** Análise de variância por postos de Friedman com 2 fatores. Houve diferença significativa entre o avaliador 3 e os outros dois.

Figura 15- Comparação, em *boxplot*, dos escores produzidos por cada avaliador na segunda etapa da pesquisa (D2), na estação emergência.

Figura 15. AVAL1= avaliador 1; AVAL2= avaliador 2; AVAL3= avaliador 3. Houve diferença significativa entre o avaliador 3 e os outros dois.

Tabela 14*- Teste de Hipóteses da escala Likert[†] para comparação dos escores produzidos pelos avaliadores no terceiro dia de avaliações (D3) da estação emergência

Avaliador	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
Avaliador 1 (Aval 1)	50	0-63,5	<0,001 Aval 1-Aval 2= 0,28 Aval 1-Aval 3=0,07 Aval 2-Aval 3<0,001
Avaliador 2 (Aval 2)	37	0-46,25	
Avaliador 3 (Aval 3)	73	0-82	

Tabela 14. * Comparação das medianas e P25-P75 dos escores produzidos por cada avaliador na terceira etapa da pesquisa (D3) †pontuação mínima: 17; pontuação máxima: 85 (a presença dos valores mínimos representados por 0 se deve à presença de dados faltantes) ** Análise de variância por postos de Friedman com 2 fatores. Houve diferença significativa entre o avaliador 3 e os outros dois.

Figura 16- Comparação, em *boxplot*, dos escores produzidos por cada avaliador na terceira etapa da pesquisa (D3), na estação emergência

Figura 16. AVAL1= avaliador 1; AVAL2= avaliador 2; AVAL3= avaliador 3. Houve diferença significativa entre o avaliador 3 e os outros dois.

Quanto utilizado o banco de dados com imputações de valores observou-se diferenças entre os escores produzidos por cada avaliador nas 3 etapas da pesquisa. Os valores de *p* para D1, D2 e D3 foram igualmente

calculados e menores que 0,001. Em D1 as diferenças foram observadas entre os avaliadores 1 e 3 ($p < 0,001$) e entre 2 e 3 ($p < 0,001$). Em D2 as diferenças foram observadas entre os avaliadores 1 e 2 ($p = 0,04$), entre os avaliadores 1 e 3 ($p = 0,02$) e entre os avaliadores 2 e 3 ($p < 0,001$). Em D3 as diferenças foram observadas entre os avaliadores 1 e 2 ($p = 0,007$), entre os avaliadores 1 e 3 ($p = 0,0015$) e entre os avaliadores 2 e 3 ($p < 0,001$).

As pontuações na escala de avaliação global (com escores entre 0 e 100) não foram diferentes na comparação entre os avaliadores em D1 ($p = 0,256$), porém observou-se diferenças em D2 e D3 ($p < 0,001$ para ambas as comparações). Em D2 as diferenças foram observadas entre os avaliadores 1 e 2 ($p < 0,001$) e entre os avaliadores 2 e 3 ($p < 0,001$). Em D3 as diferenças foram observadas entre os avaliadores 1 e 3 ($p < 0,001$) e entre os avaliadores 2 e 3 ($p < 0,001$).

Não houve imputação de dados para a escala global.

6.4.4.2-Comparação dos escores entre os residentes nas 3 etapas do estudo para a estação emergência:

Considerando a somatória dos escores de todos os avaliadores foram realizados testes de hipóteses utilizando o banco de dados com perdas de amostras assim como o banco de dados com imputações.

As avaliações globais de cada residente com escores de desempenho entre 0 e 100 não foram submetidas a imputações de dados.

Considerando o banco de dados sem imputações foi observada diferença estatisticamente significativa entre os 3 grupos de avaliação ($p = 0,03$). Análise *pós-hoc* de Nemenyi para comparação de pares evidenciou diferenças limítrofes entre os dias de avaliação. Os valores de p para a diferença entre D1 e D2 foram de 0,055, entre D1 e D3 resultaram em 0,997 e entre D2 e D3 0,065. A tabela 15 ilustra um resumo das medianas dos escores obtidos nos diferentes dias de avaliações, P25-P75 assim como os valores de p obtidos. A figura 17 representa, graficamente em *boxplot*, o desempenho dos MR nos dias de avaliação.

Tabela 15*: Teste de Hipóteses da escala Likert[†]: somatória dos avaliadores para a estação emergência (planilha sem imputação de dados)

Dia	Mediana	P25-P75	Valor de p**(análise pós-hoc de Nemenyi para comparação de pares)
D1	162	119,2-184,8	0,03
D2	180,5	152-215,5	D1-D2=0,055
D3	163,5	0-187	D1-D3=0,997 D2-D3=0,065

Tabela 15. * Comparação das medianas e P25 - P75 dos escores de desempenho dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/D3 pós-intervenção) † pontuação mínima: 51; pontuação máxima: 255 ** Análise de variância por postos de Friedman com 2 fatores. Não houve diferença estatística entre os momentos de avaliação sem ocorrência de diferença estatística entre os 3 momentos de avaliação.

Figura 17- Representa, graficamente em *boxplot*, o desempenho dos MR nos dias de avaliação.

Tabela 16*- Teste de Hipóteses da escala Likert[†]: somatória dos avaliadores para a estação emergência (planilha com imputação de dados)

Dia	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	171,5	143,2-190	0,065
D2	184	156,2-217,2	
D3	185	174,8-190	

Tabela 16. * Comparação das medianas e P25-P75 dos escores de desempenho dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/D3- pós intervenção) † pontuação mínima: 51; pontuação máxima: 255 ** Análise de variância por postos de Friedman com 2 fatores. Não houve diferença estatística entre os momentos de avaliação

Figura 18- Análise dos dados da amostra na estação procedimento, após imputação dos dados faltantes na amostra original. Não houve diferença estatística entre os momentos avaliados.

A avaliação estatística do desempenho dos MR por meio da avaliação global não revelou diferença estatisticamente significativa entre os dias de avaliação ($p = 0,43$). A tabela 17 contém o resumo das medianas dos escores obtidos nos diferentes dias de avaliações, assim como os valores de p obtidos. A figura 19 representa, graficamente, em *boxplot*, o desempenho dos MR nos dias de avaliação.

Tabela 17*- Teste de Hipóteses da escala Global[†] por avaliadores para a estação emergência

Dia	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	188,5	107,5-205	0,43
D2	184	156,2-217,2	
D3	185	174,8-190	

Tabela 17. * Comparação das medianas e P25-P75 dos escores de desempenho global dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/ D3-pós intervenção) † pontuação mínima: 0; pontuação máxima: 300 ** Análise de variância por postos de Friedman com 2 fatores. Não houve diferença estatística

Figura 19. Demonstração da evolução dos escores na avaliação global dos médicos residentes nos diferentes dias de avaliação (D1, D2 e D3) na estação emergência. Não houve diferença estatística entre os dias avaliados.

6.5- Estação Ambulatório:

6.5.1- Perdas na amostra:

Houve perdas na amostra da estação ambulatório. Essa perda foi calculada em 32,47% de todos os dados da estação. A figura 20 ilustra a porcentagem de dados perdidos segundo cada variável e a figura 21 mostra a distribuição das perdas no banco de dados.

Figura 20- Porcentagem de dados faltantes em cada variável analisada da estação ambulatório; no eixo y observamos a porcentagem de dados faltantes para cada variável (onde 0,2 = 20%), no eixo x observamos cada variável analisada na estação procedimento (onde AE = ambiente de ensino, CR = promoção de compreensão e retenção, CO = comunicação dos objetivos). Os números 1, 2 e 3 antes das abreviações se referem aos dias de avaliação. Os números após as abreviações se referem aos itens da mesma habilidade de ensino avaliada e as letras F, G e K no final das variáveis se referem à inicial do nome do avaliador). Total de variáveis da estação: 189 (a escala para a estação emergência possuía 21 itens divididos em 4 habilidades que foram examinados por 3 avaliadores independentes e em 3 ocasiões distintas)

Figura 21- Distribuição dos dados faltantes no banco de dados da estação ambulatório. Os dados faltantes são representados pelas barras verticais em amarelo e os dados completos em azul. As colunas representam as variáveis da estação enfermária (onde AE=ambiente de ensino, CR=promoção de compreensão e retenção, AD=promoção de aprendizado auto direcionado, CO=comunicação dos objetivos e FB=feedback). Os números 1, 2 e 3 antes das abreviações se referem aos dias de avaliação. Os números após as abreviações se referem aos itens da mesma habilidade de ensino avaliada e as letras F, G e K no final das variáveis se referem à inicial do nome do avaliador). As linhas representam a porcentagem de dados faltantes conforme padrões específicos de perda das variáveis (por exemplo, a última linha demonstra que em 8,3% das variáveis houve dados faltantes com aquele padrão específico de distribuição); os dados faltantes são representados pelas barras verticais em amarelo e os dados completos em azul.

6.5.2- Imputação de dados:

A imputação de dados para análise estatística da estação ambulatorio resultou em uma PFC calculada em 0,28 através da estimativa de erro do método MissForest.

6.5.3- Confiabilidade dos dados:

Os cálculos de confiabilidade, através de teoria G, para a estação ambulatorio resultaram em um coeficiente geral para toda a estação de 0,683. A consistência interna foi calculada em 0,884 e a confiabilidade entre os avaliadores resultou em um coeficiente de 0,131. Foram também calculados os mesmos coeficientes para os diferentes dias de avaliação, como se fossem etapas independentes. Para o primeiro dia de avaliações, os coeficientes gerais, consistência interna e confiabilidade entre os avaliadores foram de, respectivamente, 0,645; 0,827; 0,081. Para o segundo dia de avaliações os mesmos coeficientes foram calculados em: 0,720; 0,799; 0,155. No terceiro dia os coeficientes foram calculados em: 0,699; 0,838; 0,245. A tabela 18 resume todos os cálculos de coeficientes de confiabilidade para a estação ambulatorio.

Tabela 18*- Coeficientes de confiabilidade para a estação ambulatorio

Avaliação	Coeficiente G**	Consistência Interna**	Confiabilidade entre avaliadores –score total†
Ambulatório total	0,683	0,884	0,131
Ambulatório D1	0,645	0,827	0,081
Ambulatório D2	0,720	0,799	0,155
Ambulatório D3	0,699	0,838	0,245

Tabela 18. * Resumo dos cálculos de coeficientes de confiabilidade para a estação ambulatorio total e para os três diferentes dias de avaliação (D1, D2 e D3) *coeficiente relativo †coeficiente absoluto.

6.5.4- Teste de hipóteses e comparação por pares:

6.5.4.1- Comparação dos escores entre os avaliadores nas 3 etapas do estudo para a estação ambulatório:

Observou-se que um dos avaliadores da estação ambulatório produziu escores sistematicamente diferentes dos outros dois avaliadores. Essa observação é corroborada pelos baixos coeficientes de correlação entre avaliadores observados na tabela 19. Foi realizada uma comparação entre os escores de cada avaliador nas diferentes etapas da pesquisa com a finalidade de identificar a presença de um avaliador que produziu respostas diferentes dos demais.

Foi utilizado o banco de dados sem imputações para comparação dos escores produzidos pelos avaliadores nos diferentes dias. No primeiro dia de avaliações foi observado que os avaliadores produziram resultados diferentes ($p < 0,001$). Análise *pós-hoc* de Nemenyi evidenciou diferenças entre os avaliadores 1 e 3 ($p < 0,001$) e entre os avaliadores 2 e 3 ($p = 0,009$). No segundo dia de avaliações também foi observado que os avaliadores produziram resultados diferentes ($p < 0,001$). Novamente as diferenças foram encontradas nas comparações entre avaliadores 1 e 3 ($p = 0,09$) e entre os avaliadores 2 e 3 ($p < 0,065$). No terceiro dia de avaliações os escores também foram diferentes ($p < 0,001$), com diferenças entre os avaliadores 1 e 3 ($p < 0,001$) assim como entre os avaliadores 2 e 3 ($p < 0,001$). As tabelas 19, 20 e 21 ilustram as medianas dos escores de cada avaliador nas 3 etapas da pesquisa, assim como P25-P75, o valor de p para a comparação total e os valores de p para a análise *pós-hoc* para comparação por pares. As figuras 22, 23 e 24 ilustram, através de *boxplot*, a comparação dos escores produzidos por cada avaliador nas 3 etapas da pesquisa.

Tabela 19*- Teste de Hipóteses da escala Likert[†] para comparação dos escores produzidos pelos avaliadores no primeiro dia de avaliações (D1) da estação ambulatorio

Avaliador	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
Avaliador 1 (Aval 1)	32,5	15-58	<0,001 Aval 1-Aval 2= 0,57 Aval 1-Aval 3<0,001 Aval 2-Aval 3=0,009
Avaliador 2 (Aval 2)	42,5	32-54,5	
Avaliador 3 (Aval 3)	54	37,5-79,5	

Tabela 19. * Comparação das medianas e P25-P75 dos escores produzidos por cada avaliador na primeira etapa da pesquisa (D1) † pontuação mínima: 21; pontuação máxima: 105** Análise de variância por postos de Friedman com 2 fatores. Houve diferença estatística entre Aval 3 e os outros dois.

Figura 22- Comparação, em *boxplot*, dos escores produzidos por cada avaliador na primeira etapa da pesquisa (D1), na estação ambulatorio

Figura 22. AVAL1= avaliador 1; AVAL2= avaliador 2; AVAL3= avaliador 3. Houve diferença estatística de pontuação atribuída pelo Aval 3 em comparação aos outros dois.

Tabela 20*- Teste de Hipóteses da escala Likert[†] para comparação dos escores produzidos pelos avaliadores no segundo dia de avaliações (D2) da estação ambulatorio

Avaliador	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
Avaliador 1 (Aval 1)	0	0-55,25	<0,001 Aval 1-Aval 2= 0,989 Aval 1-Aval 3= 0,09 Aval 2-Aval 3=0,06
Avaliador 2 (Aval 2)	0	0-55	
Avaliador 3 (Aval 3)	0	0-81	

Tabela 20. * Comparação das medianas e P25-P75 dos escores produzidos por cada avaliador na segunda etapa da pesquisa (D2) †pontuação mínima: 21; pontuação máxima: 105 (a presença dos valores mínimos representados por 0 se deve à ocorrência de dados faltantes) ** Análise de variância por postos de Friedman com 2 fatores

Figura 23- Comparação, em *boxplot*, dos escores produzidos por cada avaliador na segunda etapa da pesquisa (D2), na estação ambulatorio

Figura 23. AVAL1= avaliador 1; AVAL2= avaliador 2; AVAL3= avaliador 3. Não houve diferença significativa

Tabela 21*- Teste de Hipóteses da escala Likert[†] para comparação dos escores produzidos pelos avaliadores no terceiro dia de avaliações (D3) da estação ambulatorio

Avaliador	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
Avaliador 1 (Aval 1)	66,5	57-77,5	<0,001 Aval 1-Aval 2= 0,75 Aval 1-Aval 3<0,001 Aval 2-Aval 3<0,001
Avaliador 2 (Aval 2)	63,5	50,25-76,25	
Avaliador 3 (Aval 3)	91	81,75-99	

Tabela 21. * Comparação das medianas e P25-P75 dos escores produzidos por cada avaliador na terceira etapa da pesquisa (D3) †pontuação mínima: 21; pontuação máxima: 105 ** Análise de variância por postos de Friedman com 2 fatores. Houve diferença significativa entre o avaliador 3 e os outros dois.

Figura 24- Comparação, em *boxplot*, dos escores produzidos por cada avaliador na terceira etapa da pesquisa (D3), na estação ambulatorio

Figura 24. AVAL1= avaliador 1; AVAL2= avaliador 2; AVAL3= avaliador 3. O avaliador 3 se destacou significativamente dos demais.

Quando utilizado o banco de dados com imputações de valores observou-se diferenças entre os escores produzidos por cada avaliador nas 3 etapas da pesquisa. Os valores de p para D1, D2 e D3 foram igualmente

calculados em menores que 0,001. Em D1 as diferenças foram observadas entre os avaliadores 1 e 3 ($p < 0,001$) e entre 2 e 3 ($p < 0,001$). Em D2 as diferenças foram observadas entre os avaliadores 1 e 3 ($p < 0,001$) e entre os avaliadores 2 e 3 ($p < 0,001$). Em D3 as diferenças foram observadas entre os avaliadores 1 e 3 ($p < 0,001$) e entre os avaliadores 2 e 3 ($p < 0,001$).

As pontuações na escala de avaliação global (com escores entre 0 e 100) foram diferentes entre os avaliadores em todas as etapas da avaliação, D1 ($p = 0,014$), D2 ($p < 0,001$) e D3 ($p < 0,001$). Em D1 as diferenças foram observadas entre os avaliadores 1 e 3 ($p = 0,007$) e entre os avaliadores 2 e 3 ($p < 0,001$). Em D2 as diferenças foram observadas entre os avaliadores 1 e 3 ($p < 0,001$) e entre os avaliadores 2 e 3 ($p < 0,001$). Em D3 as diferenças foram observadas entre os avaliadores 1 e 3 ($p = 0,01$) e entre os avaliadores 2 e 3 ($p < 0,001$).

Não houve imputação de dados para a escala global

6.5.4.2- Comparação dos escores entre os residentes nas 3 etapas do estudo para a estação ambulatorio:

Considerando a somatória dos escores de todos os avaliadores foram realizados testes de hipóteses utilizando o banco de dados com perdas de amostras assim como o banco de dados com imputações.

As avaliações globais de cada residente com escores de desempenho entre 0 e 100 não foram submetidas a imputações de dados.

Considerando o banco de dados sem imputações foi observada diferença estatisticamente significativa entre os 3 grupos de avaliação ($p < 0,001$). Análise *pós-hoc* de Nemenyi para comparação de pares evidenciou diferenças entre os dias de avaliação. As diferenças foram encontradas entre D1 e D3 e entre D2 e D3 (ambas com $p < 0,001$). A tabela 22 ilustra um resumo das medianas dos escores obtidos nos diferentes dias de avaliações, P25-P75 assim como os valores de p obtidos. A figura 25 representa, graficamente em *boxplot*, o desempenho dos MR nos dias de avaliação.

Tabela 22*: Teste de Hipóteses da escala Likert[†]: somatória dos avaliadores para a estação ambulatorio (planilha sem imputação de dados)

Dia	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	156	84,75-183,2	<0,001 D1-D2= 0,86 D1-D3<0,001 D2-D3<0,001
D2	0	0-191	
D3	219,5	200,8-248	

Tabela 22. * Comparação das medianas e P25 - P75 dos escores de desempenho dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/D3 pós-intervenção)† pontuação mínima: 63; pontuação máxima: 315 (a presença de valores mínimos representados por 0 traduz a presença de dados faltantes)** Análise de variância por postos de Friedman com 2 fatores. ocorrendo diferença significativa de D3 em comparação à D1 e D2.

Figura 25 - Análise dos dados da amostra original na estação ambulatorio, sem a imputação de dados. Houve diferença significativa dos primeiros dias de avaliação em relação à D3 e representativa perda de dados.

Considerando o banco de dados com imputações foi observada diferença estatisticamente significativa entre os 3 grupos de avaliação ($p < 0,001$). Análise *pós-hoc* de Nemenyi para comparação de pares evidenciou

diferenças entre D1 e D3 ($p < 0,001$) e entre D2 e D3 ($p = 0,0026$). A tabela 23 ilustra o resumo das medianas dos escores obtidos nos diferentes dias de avaliações assim como os valores de p obtidos. A figura 26 representa, graficamente, em *boxplot*, o desempenho dos MR nos dias de avaliação.

Tabela 23*- Teste de Hipóteses da escala Likert[†]: somatória dos avaliadores para a estação ambulatorio (planilha com imputação de dados)

Dia	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	187,5	182,5-200,8	<0,001 D1-D2= 0,95 D1-D3< 0,001 D2-D3=0,0026
D2	197,5	184-218,2	
D3	230	204,8-256	

Tabela 23. * Comparação das medianas e P25-P75 dos escores de desempenho dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/D3- pós intervenção) † pontuação mínima: 63; pontuação máxima: 315 ** Análise de variância por postos de Friedman com 2 fatores

Figura 26- Análise dos dados da amostra na estação ambulatorio, após imputação dos dados faltantes na amostra original, com diferença significativa de desempenho em D3.

A avaliação estatística do desempenho dos MR por meio da avaliação global revelou diferença estatisticamente significativa entre os dias de

avaliação ($p < 0,001$). Análise *pós-hoc* de Nemenyi para comparação de pares evidenciou diferença entre D1 e D3 ($p < 0,001$) e entre D2 e D3 ($p < 0,001$). A tabela 24 contém o resumo das medianas dos escores obtidos nos diferentes dias de avaliações, assim como os valores de p obtidos. A figura 27 representa, graficamente, em *boxplot*, o desempenho dos MR nos dias de avaliação

Tabela 24*- Teste de Hipóteses da escala Global[†] por avaliadores para a estação ambulatorio

Dia	Mediana	P25-P75	Valor de p**(análise <i>pós-hoc</i> de Nemenyi para comparação de pares)
D1	83	33,75-156,2	D1-D2 = 0,7 D1-D3<0,001 D2-D3<0,001
D2	0	0-156,2	
D3	197,5	170-214,5	

Tabela 24. * Comparação das medianas e P25-P75 dos escores de desempenho global dos médicos residentes nos três diferentes momentos de avaliação (D1 – basal, D2/ D3-pós intervenção) † pontuação mínima: 0; pontuação máxima: 300 ** Análise de variância por postos de Friedman com 2 fatores

Figura 27- Demonstração da evolução dos escores na avaliação global dos médicos residentes nos diferentes dias de avaliação (D1, D2 e D3) na estação ambulatorio

7- DISCUSSÃO:

O presente estudo demonstrou ganho nas habilidades de ensino clínico de médicos residentes de um hospital de ensino após intervenção breve, porém específica, com a finalidade de atrair a atenção e desenvolver habilidades as demandas didáticas que o treinamento em serviço, em instituições de ensino, impõe.

Todo o entendimento teórico concernente às responsabilidades didáticas dos MR provém da literatura internacional. Bordley e cols.(28) chegam a caracterizar o residente como o coração e a alma do currículo informal (ou *hidden curriculum*, em inglês), transmitido aos alunos do curso de medicina durante a graduação. O currículo informal é ensinado aos estudantes durante plantões ou discussões em enfermarias e é complementar ao currículo formal oferecido pelos docentes ou por médicos administrativos, com alto potencial no desenvolvimento de habilidades e principalmente, atitudes. É importante ressaltar que o trabalho didático proporcionado pelos MR é complementar, e em hipótese alguma deve substituir a instrução fornecida por professores e preceptores mais experientes, sendo que parte dos MR também serão parte deste contingente.

Alguns estudos norte-americanos reconhecem a escassez de docentes médicos e investem na instrução de MR para suprir as futuras necessidades em educadores para o curso de medicina. Lin e cols⁴² desenvolveram uma oficina com 9 horas de duração denominada “docentes para o amanhã”, do inglês “*faculty for tomorrow*” (F4T). Essa oficina foi oferecida a residentes de medicina de família e 75 profissionais compareceram. Após participação na atividade a proporção de MR que consideravam altamente provável a possibilidade de seguirem carreira acadêmica, em medicina de família, aumentou de 58% para 72%.

Autores de países em desenvolvimento também reconhecem dificuldades ainda maiores na disponibilidade de docentes para o curso de medicina e dessa forma estudaram a aquisição de habilidades de ensino por MR. Owolabi e cols⁴³ avaliaram as competências de MR em uma instituição

nigeriana de ensino, sob a perspectiva de alunos da graduação. Os resultados demonstraram que o melhor desempenho se deu na criação de um ambiente de ensino favorável e o pior na promoção de compreensão e retenção. Foi utilizado o mesmo instrumento de avaliação da presente pesquisa, o SFDP26. Os autores concluem o estudo recomendando a integração de programas de desenvolvimento pedagógico às atividades habituais de formação de especialistas, que a residência médica cumpre tão bem.

Em duas estações do presente estudo – discussão de caso em enfermaria e orientação de procedimentos, a confiabilidade obtida permitiu análise robusta dos dados. Observou-se aquisição de habilidades nas competências de ensino clínico, ambiente de ensino, aprendizado auto direcionado e *feedback* para a estação enfermaria e de promoção da compreensão e retenção, controle da sessão e comunicação dos objetivos para a estação procedimento. Não foi observado ganho na competência ambiente de ensino na estação procedimento, sendo que todos os residentes, conforme os avaliadores, partiram e mantiveram escores altos nesta habilidade, em todos os dias de avaliação.

Post e cols¹⁵ publicaram em 2009 uma revisão crítica a respeito dos programas RaT. Nesta revisão os autores sugerem parâmetros metodológicos para delimitar os estudos envolvendo a avaliação de habilidades de ensino clínico de MR. Os autores sugerem a realização de avaliações simuladas pré e pós- intervenções baseadas no método OSTE, intervenções curtas com a utilização da técnica PEM, além da gravação em vídeo do desempenho dos participantes e nova avaliação de desempenho utilizando o OSTE, 6 meses após a intervenção.

O estudo descrito nesta tese teve o cuidado de seguir a maioria dos parâmetros metodológicos descritos por Post e cols¹⁵ para uma avaliação apropriada da aquisição de habilidades de ensino clínico de MR. Foram criadas 4 estações de avaliação contemplando as situações mais freqüentes, no cotidiano das atividades pedagógicas desenvolvidas no dia a dia de MR em sua relação com os estudantes da graduação, particularmente os internos.

Para a confecção dessas estações seguiu-se as recomendações de Boillat e cols³² que descrevem em 12 passos as etapas mais importantes no processo de criação de uma atividade baseada no método OSTE. A intervenção realizada para treinamento dos MR em técnicas pedagógicas foi curta, com duração de cerca de 45-50 minutos e, baseou-se integralmente, na estratégia PEM, desenvolvida por Neher³⁸ a qual propõe a discussão de casos centrada no desenvolvimento de 5 micro habilidades de ensino clínico: (1) comprometimento com o caso, (2) busca de evidências concretas que deem suporte às hipóteses, (3) transmissão de regras de conhecimento geral, (4) enfatizar o que está correto- *feedback* positivo, (5) correção do que estava incorreto- *feedback* negativo.

Não houve oportunidade de reavaliar o desempenho dos MR 6 meses após as atividades iniciais para determinar se os efeitos da intervenção principal foram duradouros nesse período.

Sob o ponto de vistas dos autores, um dos pontos fortes do presente estudo foi a utilização de teoria da generalização (teoria G) para a determinação da confiabilidade geral de cada estação, assim como da consistência interna e do coeficiente de confiabilidade entre os avaliadores. Segundo Downing⁴⁴, em uma atividade avaliativa, o coeficiente de confiabilidade entre os avaliadores determina a principal fonte de confiabilidade da avaliação em curso. O mesmo autor cita a teoria G como a forma mais estatisticamente elegante de se medir os coeficientes de confiabilidade entre avaliadores.

No presente estudo, as estimativas de confiabilidade entre avaliadores demonstraram poder suficiente para identificar estações em que pelo menos um dos avaliadores destoou dos demais ao atribuir escores mais altos para os MR nos diferentes dias de atividades. Nessas estações os resultados se mostraram inconsistentes quando comparados com aqueles em que as estimativas de confiabilidade se mostraram mais robustas. Para a classificação das estimativas de confiabilidade entre os avaliadores usamos os mesmos critérios de Landis e Koch⁴⁵ para os valores de *kappa*: (1) <0,00 precários, (2) 0,00-0,20 pequenos, (3) 0,21-0,40 razoáveis, (4) 0,41-0,60 moderados, (5) 0,61-0,80 substanciais, (6) 0,81-1,00 quase perfeitos.

Na estação enfermaria observamos moderado coeficiente total de confiabilidade entre observadores. Quando os dias de avaliação foram considerados de forma individual houve aumento progressivo dos coeficientes, com o maior deles tendo sido observado em D3, 0,698. O mesmo comportamento de aumento progressivo das estimativas de confiabilidade foi observado em relação à consistência interna e ao coeficiente geral de confiabilidade. O aumento progressivo das estimativas de confiabilidade provavelmente se deve ao efeito da intervenção, tanto no comportamento dos residentes quanto na capacidade dos avaliadores de diferenciar residentes com bons desempenhos daqueles com desempenhos ruins.

Na estação procedimento, observou-se substancial coeficiente total de confiabilidade entre avaliadores mais alto entre todas as estações da pesquisa (0,772). Nesse caso observou-se um comportamento de redução progressiva da estimativa de confiabilidade entre os observadores, nos diferentes dias de atividades. É possível que um maior percentual de dados faltantes nessa estação possa ter contribuído para esse achado, já que as perdas foram maiores em D3. As estimativas de consistência interna e de confiabilidade geral nos diferentes dias de avaliação se mantiveram estáveis.

A estação emergência produziu um pequeno coeficiente total de confiabilidade entre avaliadores. Os coeficientes de confiabilidade entre os observadores foram razoáveis em D1 (0,264), moderados em D2 (0,404) e pequenos em D3 (0,076). Nesse caso, além de um maior percentual de dados faltantes observou-se que os escores de um dos avaliadores foram estatisticamente maiores quando comparados aos dos outros dois. Os coeficientes de confiabilidade geral e a consistência interna da estação nos diferentes dias de avaliação se mostraram estáveis e substancialmente mais altos que os coeficientes de confiabilidade entre os avaliadores. Em D2 observou-se um coeficiente de confiabilidade geral de 0,839, compatível com avaliações de alto padrão de discriminação.

A estação ambulatório demonstrou coeficientes de confiabilidade entre avaliadores consistentemente pequenos em praticamente todas as ocasiões, exceto em D3, quando foi calculado em 0,245 (razoável). As estimativas de confiabilidade geral e de consistência interna se mostraram estáveis e muito

superiores às aquelas de confiabilidade entre os observadores em todas as ocasiões da pesquisa. Essa estação teve o maior percentual de dados faltantes. Também nessa estação observou-se que um dos avaliadores produziu resultados estatisticamente superiores aos demais. Essas podem ser as causas mais prováveis para coeficientes de confiabilidade entre observadores consistentemente tão baixos.

As duas estações com os piores coeficientes de confiabilidade entre observadores, emergência e ambulatório, são precisamente aquelas em que houve um maior percentual de dados faltantes e em que um dos avaliadores produziu escores discrepantes dos demais nas avaliações de desempenho dos MR, nas diferentes etapas da pesquisa.

É possível que as avaliações discrepantes de um dos avaliadores tenham relação com o tempo entre as reuniões para padronização dos escores nas diferentes escalas e o momento da avaliação. Sabe-se que os dois avaliadores, com maior tempo de experiência profissional e em ensino, possuem habilidades muito próximas de avaliação. Entretanto, um deles pontuou com notas mais elevadas, destoando dos outros dois. Portanto, não parece ser ter sido o acúmulo de experiência prévia, o fator preponderante, nas diferenças entre os avaliadores.

Com relação à estação emergência, os escores de um dos avaliadores foram atribuídos aos MR mais de 6 meses depois da reunião por videoconferência para normatização dos escores. Pode haver um efeito de perda na retenção das informações quanto aos parâmetros e critérios discutidos entre os avaliadores, como de praxe fizeram, de modo independente dos demais pesquisadores. Essa é uma questão a ser respondida por estudos complementares. Tipo de treinamento e intervalo, entre as reuniões e processamento das avaliações sejam iniciadas por mais de um indivíduo. Há necessidade de estimar um coeficiente de confiabilidade entre observadores pré-determinado, utilizando-se, por exemplo, uma amostra que não seria incluída na análise final da pesquisa? Há um ponto crítico no tempo em que as informações discutidas nas reuniões para padronização dos escores sofreriam comprometimento significativo com consequente prejuízo na confiabilidade das avaliações?

Uma das limitações que podemos citar nesta pesquisa está relacionada tanto à perda na amostra de residentes quanto aos dados faltantes nos bancos de dados das diferentes estações. A amostra inicial de conveniência dos residentes era composta por 48 indivíduos. Houve perdas na amostra principalmente por conflitos nas agendas dos residentes. Algumas atividades obrigatórias dos diferentes programas de residência médica tiveram caráter imprevisível quanto ao horário de início e horário de término e por esse motivo alguns MR não puderam comparecer a uma ou mais atividades de avaliação ou intervenção. Ao final do período de coleta de dados apenas 24 MR compareceram a todas as 3 etapas do estudo e às 2 intervenções propostas. A decisão de incluir na análise final apenas os indivíduos que participaram de todas as etapas e intervenções levou a uma perda de metade da amostra. É possível que resultados mais consistentes pudessem ter sido atingidos com uma amostra maior. Porém, a inclusão de indivíduos que não tivessem participado de todas as etapas e intervenções poderia atribuir vieses importantes à análise final dos dados.

As perdas de dados nas diferentes estações de avaliação se deveram principalmente ao mal funcionamento de equipamentos usados nas gravações em vídeo dos desempenhos dos MR. Os principais problemas enfrentados sob esse ponto de vista foram o descarregamento da bateria dos equipamentos, a falta de espaço para gravação nos dispositivos e a deleção acidental de arquivos pelo autor da tese. Outra fonte de dados faltantes foi devida à falta de preenchimento, pelos avaliadores, de itens do questionário para cada estação.

O método escolhido para abordar os dados faltantes foi a imputação múltipla com finalidade de produzir bancos completos para as análises estatísticas. Considerou-se que a perda de dados era completamente ao acaso (*missing completely at random-MCAR*) ou ao acaso (*missing at random-MAR*). Nunes e cols⁴⁶ descrevem a importância da imputação múltipla quando comparada à imputação simples de dados pois a primeira leva em consideração a variabilidade das imputações na construção de modelos.

O presente estudo descreveu pela primeira vez no Brasil um experimento testado em outros países, culturas e diferentes profissões da saúde, em que MR tiveram suas habilidades de ensino clínico avaliadas por

meio de um instrumento validado antes e após uma intervenção específica, baseada no método de ensino PEM.

Os resultados demonstraram aquisição de conhecimentos pedagógicos e apontam na direção de efeitos favoráveis na interação residente-aluno de medicina. Estudos adicionais são necessários para identificar as maneiras mais eficazes para implementação de programas RaT no Brasil.

Outro tema para investigações futuras seria o tempo de duração desses programas e como eles se encaixariam na grade curricular dos MR de diferentes especialidades. Os autores acreditam firmemente no RaT como aprimoramento das competências do MR e do ensino médico em nosso meio

8 CONCLUSÕES

8- CONCLUSÕES:

O presente estudo demonstrou ganho na aquisição de habilidades de ensino clínico após intervenção baseada no método de ensino PEM (40-50 minutos de duração), em três estações de avaliação simulada (aluno simulado) desenvolvidas especificamente com essa finalidade, enfermaria, procedimento e ambulatório.

Não houve diferença estatisticamente significativa nos escores atribuídos aos MR nas diferentes etapas do estudo para a estação emergência.

Os coeficientes de confiabilidade gerais e a consistência interna, estimados através de teoria G, foram considerados satisfatórios em todas as estações. Na terceira etapa da estação enfermaria e na segunda etapa da estação emergência os coeficientes gerais atingiram valores acima de 0,8, condizentes com avaliações de alto poder discriminatório.

Os coeficientes de confiabilidade entre avaliadores, também estimados através de teoria G, foram considerados moderados para a estação enfermaria, substanciais para a estação procedimentos e pequenos para as estações emergência e ambulatório.

Os médicos residentes, com treinamento de 40-50 minutos, modificaram suas habilidades de ensino para melhor e de acordo com a literatura. Este dado pode ser comprovado nas estações em que não houve perda substancial de dados (por defeitos técnicos de gravação dos vídeos) e melhor concordância entre os avaliadores.

Embora não fosse objetivo do estudo, decidimos por utilizar método estatístico defendido na literatura, como mais apropriado ao escopo do estudo. Neste sentido, as perdas e incongruências serviram para revelar a capacidade da Teoria G em discriminar os dados avaliados.

REFERÊNCIAS

1. Santos EG, Ferreira RR, Mannarino VL, Leher EMT, Goldwasser RS, Bravo Neto GP. Assessment of preceptorship in general surgery residency in the operating room, comparison between a teaching hospital and a non teaching hospital. *Revista do Colégio Brasileiro de Cirurgiões*. 2012;39(6):547-52.
2. MARTINS JC. Vygotsky e o Papel das interações sociais na sala de aula. *Reconhecer e Desvendar o Mundo*. 1997.
3. Louie AK, Beresin EV, Coverdale J, Tait GR, Balon R, Roberts LW. Residents as teachers. *Acad Psychiatry*. 2013;37(1):1-5.
4. Burgin S, Homayounfar G, Newman LR, Sullivan A. Instruction in teaching and teaching opportunities for residents in US dermatology programs: Results of a national survey. *J Am Acad Dermatol*. 2017;76(4):703-6.
5. Stern DT. In search of the informal curriculum: when and where professional values are taught. *Acad Med*. 1998;73(10 Suppl):S28-30.
6. Brown RS. House staff attitudes toward teaching. *J Med Educ*. 1970;45(3):156-9.
7. Donovan A. Radiology resident teaching skills improvement: impact of a resident teacher training program. *Acad Radiol*. 2011;18(4):518-24.
8. Bing-You RG, Sproul MS. Medical students' perceptions of themselves and residents as teachers. *Med Teach*. 1992;14(2-3):133-8.
9. Busari JO, Arnold A. Educating doctors in the clinical workplace: unraveling the process of teaching and learning in the medical resident as teacher. *J Postgrad Med*. 2009;55(4):278-83.
10. Al Achkar M, Hanauer M, Morrison EH, Davies MK, Oh RC. Changing trends in residents-as-teachers across graduate medical education. *Adv Med Educ Pract*. 2017;8:299-306.
11. Dewey CM, Coverdale JH, Ismail NJ, Culberson JW, Thompson BM, Patton CS, et al. Residents-as-teachers programs in psychiatry: a systematic review. *Can J Psychiatry*. 2008;53(2):77-84.
12. Berger JS, Daneshpayeh N, Sherman M, Gaba N, Keller J, Perel L, et al. Anesthesiology residents-as-teachers program: a pilot study. *Journal of graduate medical education*. 2012;4(4):525-8.

13. Ryg PA, Hafler JP, Forster SH. The Efficacy of Residents as Teachers in an Ophthalmology Module. *J Surg Educ.* 2016;73(2):323-8.
14. Magalhães G. O Residente como Professor: Formação Docente no Programa de Residência em Medicina de Família e Comunidade da Universidade Federal de Pernambuco. São Paulo: Tese [Mestrado]: Universidade Federal de São Paulo-Escola Paulista de Medicina; 2012.
15. Post RE, Quattlebaum RG, Benich JJ, 3rd. Residents-as-teachers curricula: a critical review. *Acad Med.* 2009;84(3):374-80.
16. Faria MA. Neolithic trepanation decoded- A unifying hypothesis: Has the mystery as to why primitive surgeons performed cranial surgery been solved? *Surg Neurol Int.* 2015;6:72.
17. Duffy TP. The Flexner Report--100 years later. *Yale J Biol Med.* 2011;84(3):269-76.
18. Nunes MdPT. Residência médica no Brasil: situação atual e perspectivas. *Cad ABEM.* 2004;1:30-2.
19. Brown RS. Pedagogy for surgical house staff. *J Med Educ.* 1971;46(1):93-5.
20. Hill AG, Yu TC, Barrow M, Hattie J. A systematic review of resident-as-teacher programmes. *Med Educ.* 2009;43(12):1129-40.
21. Craig JL, Page G. Teaching in medicine: an elective course for third-year students. *Med Educ.* 1987;21(5):386-90.
22. Morrison EH, Hafler JP. Yesterday a learner, today a teacher too: residents as teachers in 2000. *Pediatrics.* 2000;105(1 Pt 3):238-41.
23. Education ACfGM. ACGME common program requirements. 2013. Requirement IV A. 2015;5:9.
24. Education LCoM. Functions and structure of a medical school: Standards for accreditation of medical education programs leading to the MD Degree: Liaison Committee on Medical Education; 2007.
25. Sheets KJ, Hankin FM, Schwenk TL. Preparing surgery house officers for their teaching role. *Am J Surg.* 1991;161(4):443-9.
26. Jewett LS, Greenberg LW, Goldberg RM. Teaching residents how to teach: a one-year study. *J Med Educ.* 1982;57(5):361-6.
27. Busari JO, Prince KJ, Scherpbier AJ, Van Der Vleuten CP, Essed GG. How residents perceive their teaching role in the clinical setting: a qualitative study. *Med Teach.* 2002;24(1):57-61.

28. Bordley DR, Litzelman DK. Preparing residents to become more effective teachers: a priority for internal medicine. *Am J Med.* 2000;109(8):693-6.
29. Morrison EH, Shapiro JF, Harthill M. Resident doctors' understanding of their roles as clinical teachers. *Med Educ.* 2005;39(2):137-44.
30. Furney SL. et al. Teaching the one-minute preceptor. A randomized controlled trial. *J Gen Intern Med.* 2001;16(9):620-4.
31. Barrows HS, Abrahamson S. The programmed patient: a technique for appraising student performance in clinical neurology. *J Med Educ.* 1964;39:802-5.
32. Boillat M, Bethune C, Ohle E, Razack S, Steinert Y. Twelve tips for using the objective structured teaching exercise for faculty development. *Medical teacher.* 2012;34(4):269-73.
33. Sturpe DA, Schaivone KA. A primer for objective structured teaching exercises. *American journal of pharmaceutical education.* 2014;78(5):104.
34. McAndrew M, Eidtson WH, Pierre GC, Gillespie CC. Creating an objective structured teaching examination to evaluate a dental faculty development program. *Journal of dental education.* 2012;76(4):461-71.
35. Steinert Y, Mann K, Centeno A, Dolmans D, Spencer J, Gelula M, et al. A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education: BEME Guide No. 8. *Medical teacher.* 2006;28(6):497-526.
36. Adamo G. Simulated and standardized patients in OSCEs: achievements and challenges 1992-2003. *Medical teacher.* 2003;25(3):262-70.
37. Williams BC, Litzelman DK, Babbott SF, Lubitz RM, Hofer TP. Validation of a global measure of faculty's clinical teaching performance. *Academic Medicine.* 2002;77(2):177-80.
38. Neher JO, Gordon KC, Meyer B, Stevens N. A five-step "microskills" model of clinical teaching. *J Am Board Fam Pract.* 1992;5(4):419-24.
39. Brennan R. Statistics for social science and public policy. Generalizability theory. New York, NY, US. Springer-Verlag Publishing. [http://dx. doi. org/10.1007/978-1-4757-3456-0](http://dx.doi.org/10.1007/978-1-4757-3456-0); 2001.
40. Bloch R, Norman G. Generalizability theory for the perplexed: a practical introduction and guide: AMEE Guide No. 68. *Medical teacher.* 2012;34(11):960-92.

41. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*. 2011;28(1):112-8.
42. Lin S, Gordon P. Preparing Residents for Teaching Careers: The Faculty for Tomorrow Resident Workshop. *Family medicine*. 2017;49(3):225-9.
43. Owolabi MO, Afolabi AO, Omigbodun AO. Performance of residents serving as clinical teachers: a student-based assessment. *J Grad Med Educ*. 2014;6(1):123-6.
44. Downing SM. Reliability: on the reproducibility of assessment data. *Medical education*. 2004;38(9):1006-12.
45. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *biometrics*. 1977:159-74.
46. Nunes LN, Klück MM, Fachel JMG. Comparação de métodos de imputação única e múltipla usando como exemplo um modelo de risco para mortalidade cirúrgica. *Revista Brasileira de Epidemiologia*. 2010;13:596-606.

Teaching skills for medical residents: are these important? A narrative review of the literature

Saadallah Azor Fakhouri Filho^I, Lorena Pinho Feijó^{II}, Kristopherson Lustosa Augusto^{III}, Maria do Patrocínio Tenório Nunes^{IV}

Universidade Federal de Uberlândia (UFU), Uberlândia (MG), Brazil

^IMD. Doctoral Student, Universidade Federal de Uberlândia (UFU), Uberlândia (MG), Brazil.

[IDorcid.org/0000-0001-8413-0619](https://orcid.org/0000-0001-8413-0619)

^{II}MD. Professional Master's Student, Centro Universitário Christus (UNICHRISTUS), Fortaleza (CE), Brazil.

[IDorcid.org/0000-0001-9638-1992](https://orcid.org/0000-0001-9638-1992)

^{III}MD, PhD. Adjunct Professor, Department of Clinical Medicine, Faculdade de Medicina da Universidade Federal do Ceará (FAMED - UFC) and Universidade de Fortaleza (UNIFOR), Fortaleza (CE), Brazil.

[IDorcid.org/0000-0001-9254-9129](https://orcid.org/0000-0001-9254-9129)

^{IV}MD, PhD. Associate Professor, Department of Internal Medicine, Faculdade de Medicina da Universidade de São Paulo (FMUSP), São Paulo (SP), Brazil.

[IDorcid.org/0000-0003-3616-515X](https://orcid.org/0000-0003-3616-515X)

KEY WORDS:

Education, medical.

Internship and residency.

Education.

ABSTRACT

BACKGROUND: There is extensive evidence, mainly from the United States and Canada, that points towards the need to train medical residents in teaching skills. Much of the “informal curriculum”, including professional values, is taught by residents when consultants are not around. Furthermore, data from the 1960s show the importance of acquiring these skills, not only for residents but also for all doctors. Teaching moments can be identified in simple daily situations, like discussing a clinical situation with patients and their families, planning patients’ care with the healthcare team or teaching peers and medical students. The aim here was to examine the significance of resident teaching courses and estimate the effectiveness of these courses and the state of the art in Brazil.

METHODS: We conducted a review of the literature, using the MEDLINE, PubMed, SciELO and LILACS databases to extract relevant articles describing residents-as-teachers (RaT) programs and the importance of teaching skills for medical residents. This review formed part of the development of a doctoral project on medical education.

RESULTS: Original articles, reviews and systematic reviews were used to produce this paper as part of a doctoral project.

CONCLUSIONS: RaT programs are important in clinical practice and as role models for junior learners. Moreover, these educational programs improve residents’ self-assessed teaching behaviors and teaching confidence. On the other hand, RaT program curricula are limited by both the number of studies and their methodologies. In Brazil, there is no such experience, according to the data gathered here, except for one master’s thesis.

INTRODUCTION

Learning medicine is not a lonely journey and is no longer a passive act. Doctors at both junior and senior levels participate in this complex process, thereby facilitating learning and consolidating and updating knowledge daily. Santos et al.¹ described facilitators as doctors who are more experienced and who thus help in the professional development of undergraduates, medical residents and their peers. These authors referred to the work of Vygotsky,² for whom “the learning process comes from outside sources and is conceived through individual interactions with the world”. Furthermore, they postulated that the role of facilitators (preceptors) is to enable some situations in which apprentices’ knowledge assimilation and production becomes transformed. The preceptor’s main role is therefore to facilitate the acquisition of theory and skills by stimulating his/her pupils to make their own discoveries.

Residency is characterized by in-service learning, i.e. training during practice within a scenario in which residents may become role models, in accordance with statements based on educational strategies for teaching and learning. Therefore, it is crucial to balance teaching, learning and healthcare assistance.

The most remarkable characteristic of medical residency is its in-service training, in which teaching is integrated with practice scenarios, so as to build a model for physicians’ ideological, ethical and professional identity. The professional competence that is expected at the end of a medical residency program needs to go beyond technical knowledge. It also encompasses skills and attitudes that show effective team capabilities, leadership, communication skills, empathy, self-control and metacognition.³ Sternszus et al.⁴ investigated the importance of resident role models in the education and career choices of medical students, in a cross-sectional survey-based study. Their study was the first to illustrate that resident role models are perceived by medical

students to be as important as role models formed by attending physicians, for their education.

Medical residents are an essential part of the workforce in most Brazilian hospitals. In teaching centers, because of the humdrum nature of daily tasks, there is an additional responsibility towards helping students and peers, to help them improve their knowledge and technical skills. Several authors have noted that the word 'doctor' comes from the Latin verb *docere*,⁵ meaning to teach, with the aim of highlighting how teaching is important for all medical professionals, including those under training.

American researchers have estimated that residents spend almost a quarter of their residency programs teaching others, even though they are undertrained for this purpose.⁶ Much of the "informal curriculum", including professional values, is taught by residents.⁷ This process focusing on peer-to-peer cooperative learning remains poorly studied.

Development of clinical acumen through good clinical teaching is a key component of medical education. Few residents will come to postgraduate training with well-developed teaching skills or a sense of their relevance to student education, or with knowledge of the principles of adult learning and its theory and practice.⁵

Pioneering studies conducted in the United States and Canada have highlighted the relevance of training medical residents to provide them with teaching skills. These studies demonstrated that 20-25% of the residents' working hours were spent on teaching activities,^{6,8,9} and that medical students attributed 30-85% of all the clinical theory that they acquired during their undergraduate programs to teaching given by their residents.^{10,11} All of this exchange of information is provided through long working hours and exhaustive periods on call. Despite all the information presented above, doctors, medical students and residents receive little or no formal instruction on how to teach.^{10,12,13}

Both the American College of Graduate Medical Education¹⁴ and the Liaison Committee on Medical Education¹⁵ have recommended that formal instruction on teaching skills should be provided for medical residents. These activities have been deemed to be so important that Louie et al.⁵ described the development of residents as teachers in terms not only of a necessary personal obligation but also of a national priority.

The aim of the present review was to assess the significance attributed to programs on teaching skills for residents that have been described in the literature. We aimed to compare these results with findings in Brazil, in order to provide up-to-date conclusions and recommendations regarding this topic.

METHODS

We conducted a review of the literature, using MEDLINE, PubMed, SciELO and LILACS databases to extract relevant articles describing residents-as-teachers (RaT) programs and the

importance of teaching skills for medical residents. This review formed part of the development of a doctoral project on medical education. We searched the literature between January 2015 and December 2017, covering the period between 1970 and 2017.

The MeSH terms used in the search were "internship and residency", "education" and "education, medical". The key words used in the search included "residents", "residents as teachers", "residents-as-teachers", "residents AND teachers", "residents AND education", "education" and "medical education". The key words used for the search in Portuguese in LILACS and SciELO included "medicos residentes", "medicos residentes AND educacao", "medicos residentes AND ensino", "educacao medica" and "ensino em medicina". The articles identified as containing information regarding teaching skills for medical residency were selected for intensive review and were analyzed by two authors (SAF and MPTN). Searches of primary publications referenced in other articles were also included. These were selected for intensive review and were analyzed by two authors (SAF and MPTN).

RESULTS

The search in LILACS and SciELO using Portuguese key words did not yield any results on the specific topic of formal residents-as-teachers programs, while the search in MEDLINE/PubMed yielded 213 articles, of which 44 were found to contain information regarding teaching skills for medical residency. The references in Portuguese that were included in this paper were extracted from the personal files of one author (MPTN).

1) Selection of relevant articles:

Table 1 describes the numbers of texts extracted from each database. The relevant publications included original articles, systematic reviews, critical reviews, randomized controlled trials, guidelines from medical education experts and material from pioneering authors in this field. **Table 2** shows the most relevant studies included in this review and summarizes their main characteristics.

2) Residents-as-teachers programs from abroad:

Specific agendas for training residents in teaching skills were first developed in the 1960s. Since then, the numbers of residents-as-teachers (RaT) programs around the world has increased and their methodologies have improved. RaT programs are nowadays part of most residency programs in the United States.

Table 1. Databases and numbers of articles extracted

Database	Number of articles
PubMed	68
MEDLINE	95
Other sources*	4

*Personal files of one author.

Table 2. Relevant studies on residents-as-teachers (RaT) programs that were included in this review

Authors	Year	Study type*	Participants (n)	Specialty	Main objectives	Conclusions
Jewett et al. ²⁰	1982	RCT	55 residents	Pediatrics	To compare residents who received clinical teaching instruction with those who did not.	The intervention group was significantly more confident as teachers and received more positive feedback on their teaching.
Sheets et al. ¹⁹	1991	RL	NA	Surgery	To describe evidence that supported the importance of training surgery residents in teaching skills.	Residency program directors and faculty members within surgery needed to acknowledge that teaching was an important component of residents' daily agenda.
Bordley et al. ²²	2000	NR	NA	Internal medicine	To describe the importance of teaching skills for residents and discuss the costs involved in these activities at different institutions.	The authors strongly supported investment in training residents as teachers.
Morrison et al. ¹³	2000	NR	NA	NA	To describe the number of RaT programs in the US and demonstrate evidence for their implementation and evaluation.	Research was needed to identify the most appropriate design for RaT programs and how they affected educational outcomes.
Furney et al. ³⁰	2001	RCT	57 second- and third-year residents	Internal medicine	To compare residents who received a one-hour intervention based on the One-Minute Preceptor, with a control group.	Intervention group residents reported statistically significant changes in all behaviors after the One-Minute Preceptor.
Morrison et al. ³⁷	2002	QS	100 medicine students, residents and faculty members	Internal medicine Pediatrics Family medicine	To describe the learning needs of residents for becoming more effective teachers, using 11 focus groups and 4 semi-structured interviews.	Residents filled important roles as practical clinical teachers and role models for junior learners.
Wamsley et al. ²⁶	2004	RL	14 articles on RaT programs	Multiple	To examine the evaluation methods for resident teaching courses and estimate the effectiveness of those teaching courses.	Resident teaching courses improved resident self-assessed teaching behaviors and teaching confidence. Further studies were needed to elucidate the best format, length, timing and content of these courses and to determine whether they influenced learner performance.
Morrison et al. ²³	2005	QS	21 third-year residents	Internal medicine Family medicine Pediatrics	To compare residents who received a 13-hour training in teaching skills in the previous year, with those who did not, through semi-structured interviews.	Intervention group residents expressed more enthusiasm for teaching, learner-centered learning and self-knowledge about teaching. Control group residents seemed easily frustrated by time constraints and often expressed cynicism and guilt toward learners.
Dewey et al. ³³	2008	SR	13 articles on RaT programs	Multiple	To identify all randomized control trials (RCTs) on residents' teaching skills programs in psychiatry.	Only one trial incorporating psychiatry residents was found to exist.
Busari et al. ²¹	2009	QS	18 residents	Pediatrics Obstetrics and gynecology	To extract recommendations from interviews, regarding how a training program for residents could be created.	Enthusiasm and enjoying teaching were good attributes of successful teachers. Reasons for poor teaching were lack of time and absence of support from attending staff.
Post et al. ²⁷	2009	SR	24 articles on RaT programs	Multiple	To provide an updated systematic review of the literature on RaT program curricula and determine the most evidence-based curriculum and evaluation strategy.	Research on RaT program curricula was limited by both the number of studies and their methodology. The results demonstrated that these curricula can significantly improve residents' teaching skills.

Continue...

Table 2. Continuation

Authors	Year	Study type*	Participants (n)	Specialty	Main objectives	Conclusions
Karani et al. ¹⁸	2014	QS	37 third-year medical students	NA	To describe what students learned from residents and teaching strategies used by excellent resident teachers.	In this study, role modeling was the most frequently classified teaching model. Residents' teaching was critically important for undergraduate students.
Owolabi et al. ³⁸	2014	QS	20 residents	Internal medicine	To evaluate the clinical teaching skills of internal medicine residents from the perspective of medical students in a tertiary-level teaching institution in Africa.	Residents' clinical teaching skills were suboptimal, particularly regarding their ability to promote understanding and retention.
Dannaway et al. ²⁸	2016	RL	12 articles on RaT programs	Multiple	To assess the current evidence regarding the efficacy of teaching skills programs for junior medical officers.	The review of the literature demonstrated many positive effects from teaching skills programs, thus supporting their use. Substantial threats of bias were present in most studies.
Ramani et al. ¹⁷	2016	FG	NA	Multiple	To guide medical educators involved in the implementation of RaT programs.	The authors highlighted the importance of congruence between formal and hidden curricula and encouraged evidence-based approaches within education.
Al Achkar et al. ¹⁶	2017	QQS	221 residency program directors	Multiple	To compare the number of RAT programs with data from a previous study in 2000 and ask for feedback about the importance of these activities.	Over 80% of the residency programs surveyed had implemented RaT programs. Program directors had realized that there was a clear need for formative training experiences for residents.
Chokshi et al. ³²	2017	IR	29 second-year residents	Pediatrics	The authors developed and evaluated an intensive one-day RaT program curriculum using a flipped classroom approach.	Residents demonstrated statistically significant improvements in performance between pre- and postworkshop evaluations through objective structural teaching evaluations and attitudinal and self-efficacy questionnaires.

*NR = narrative review; QQS = qualitative and quantitative survey; FG = framework guide; QS = qualitative study; RL = review of the literature; RCT = randomized controlled trial; SR = systematic review; IR = innovation report; NA = not applicable.

Morrison et al.¹³ reported that the prevalence of programs for developing residents' teaching skills in the United States was 55%. In 2017, the same authors published new data in which the prevalence was established as 80.54%, i.e. a 26.34% increase (95% confidence interval, CI: 20.39% to 32.29%) over the last 15 years.¹⁶ Ramani et al.¹⁷ described the potential benefits of RaT programs in the Association for Medical Education in Europe (AMEE) 2016 guide 106. Karani et al.¹⁸ reported that being a role model was the tool that residents who were teaching most frequently identified. Acquiring teaching skills was seen to involve complex conscious and unconscious activities, through observation and reflection on behaviors.

3) Impact of residents-as-teachers programs on residents and students:

RaT programs have been particularly successful for several reasons. Medical students like to work with residents and,

additionally, appreciate their close supervision. Thus, residents are seen as a positive influence and example of professionalism.^{10,19-21} Moreover, residents spend large amounts of time together with medical students. They are close to them in practical activities and have similar ages and professional development processes. Most residents feel more satisfaction with their work while experiencing teaching duties.^{22,23} Moreover, these programs are especially attractive because they improve self-confidence in teaching.

Teaching by residents is different from and probably complementary to that of institutions' attending staff and faculty members. Residents tend to teach

1. different things (bedside skills and patient management rather than factual knowledge);
2. in a different way (as near-peer teachers); and
3. at different times (teaching while on-call).²⁴

In a qualitative analysis on a teaching initiation module that had been presented, it was concluded that it was possible to develop pedagogical skills at this stage, in a process that would be coherent with changes to health and education policies.²⁵ According to the same authors, residents needed to learn:

1. to take on leadership and provide a role model;
2. to give guidance to learners;
3. to give feedback,
4. to teach bedside skills;
5. to teach about procedures;
6. to teach about inpatients;
7. to teach about charting; and
8. to give lectures.

Most of the studies on RaT programs have shown that residents achieved significant improvement in teaching skills after some specific training. The first review of the literature on this topic was published in 2004 and analyzed 13 studies with different experimental designs.²⁶ A review conducted in 2009 assessed 24 studies on residents who were enrolled in several programs and found that in 21 of these studies there was significant enhancement of teaching performance after specific training.²⁷ A recent assessment on 39 studies, published in 2016, demonstrated that interventions based on providing teaching skills up to level 3 on Kirkpatrick's outcome scale were effective (**Table 3**).^{28,29}

The evidence supporting RaT programs appears not to correlate the time spent on training to efficiency in teaching. Even a one-hour intervention showed benefits in a study involving internal medicine residents.³⁰ A four-week elective course created for senior and family medicine residents paved the way for a paper giving advice on how to create a RaT program.³¹ In another study published in 2017, 29 participants were enrolled in a successful one-day RaT program that used a "flipped classroom" approach.³²

In another systematic review, an efficacy analysis was conducted on residents from a wide spectrum of specialties who participated in single programs.³³ On the whole, the interventions and outcomes measured were heterogeneous and the quality of the methodologies varied. The authors felt that these programs brought the opportunity to advance educational research in this field.³³

Residents who teach acquire the material that they teach more effectively than they would if they did it through self-study or through attending lectures.³⁴ Their teaching duties have been linked to greater job satisfaction.²³

In 2016, a meta-analysis revealed that most studies on residents-as-teachers programs had significant methodological flaws. Nevertheless, it was found that the main impacts of these interventions included improvement of attitudes and positive perceptions toward clinical teaching (Kirkpatrick level 2a); support for modification of knowledge or skills (Kirkpatrick level 2b); development of teaching skills (Kirkpatrick level 3) and some improvement of students' learning after the intervention (Kirkpatrick level 4b). Some studies revealed positive organizational change (Kirkpatrick level 4a).²⁸

Residents with better teaching skills seem to have greater knowledge of taught material and better clinical skills. In a retrospective study on senior residents in general surgery (covering the period from 2009 to 2013), technical ability was assessed through their performance in the Fundamentals of Laparoscopic Surgery examination. Teaching ability was assessed through evaluations among medical students on a Likert scale. There was evidence that residents who were better teachers had greater knowledge of taught material and a higher degree of laparoscopic skills.³⁵

Snell²⁴ provided information on the effects of RaT programs on residents. This author showed that those who were teaching had greater enthusiasm for teaching and greater job satisfaction.

Residents with effective teaching skills may also have a positive effect on patient care. Involvement of residents in teaching

Table 3. Kirkpatrick's²⁹ model for evaluating educational outcomes*

Kirkpatrick level	Evaluation outcome	Explanation
Level 1	Reaction	Participants' views of the learning experience and its organization, presentation, content, teaching methods and quality of instruction
Level 2A	Learning – change in attitudes	Changes in attitudes or perceptions among participant groups towards teaching and learning
Level 2B	Learning – modification of knowledge or skills	For knowledge, this relates to the acquisition of concepts, procedures and principles For skills, this relates to the acquisition of thinking and problem-solving, psychomotor and social skills
Level 3	Behavior – change in behavior	Documents the transfer of learning to the workplace or willingness of learners to apply new knowledge and skills
Level 4A	Results – change in the system or organizational practice	Refers to wider changes in the organization attributable to the educational program
Level 4B	Results – change among the participants: students and peers	Refers to improvement in medical student or peer learning or performance as a direct result of the educational intervention

*Adapted from Kirkpatrick (1994).²⁹

activities has been shown to have a positive effect on their communication skills, and good patient communication skills have been associated with better clinical outcomes.

4) Content of current RaT programs:

RaT programs have been delivered using many methods, including lectures, small-group discussion, practice with peers, videotape reflections and role-playing. The total number of hours dedicated to RaT program instruction varies widely. In the USA, RaT programs are more prevalent within pediatrics, family medicine, internal medicine, psychiatry, obstetrics/gynecology and surgery.

A thematic analysis identified five main reasons for implementing RaT programs: (1) teaching is part of residents' role; (2) learners desire formal RaT training; (3) regulatory bodies require RaT training; (4) RaT programs improve residents' education; and (5) RaT programs prepare residents for their current and future roles. There are also five reasons for not implementing RaT training: (1) lack of time and energy; (2) lack of expertise and resources; (3) newness of the program; (4) limited access to students; and (5) RaT instruction is not desired.¹⁶

5) Brazilian context:

A recent review of the literature only identified one study on RaT programs in Brazil.³⁶ The authors concluded that there was no description of formal development of teaching skills within medical residency curricula, according to the databases that were searched. Regarding the teaching-learning process (for residents in family medicine), the authors stated that the pedagogical project and teaching plan for RaT programs constituted a social complex that would need a reflective, critical and collaborative approach.

DISCUSSION

This narrative review shows that residents with better teaching skills might have greater knowledge of taught material and better clinical skills, according to some studies, which in turn enhances patient care. Students and senior doctors agree that residents have a critical role in the teaching process.³⁷ Residents themselves can also recognize their protagonist role in spreading knowledge to students, peers, patients and families.

Around the world, universities set the goals for preparing residents to teach. These objectives have been seen as an effective way of improving the entire educational experience of residents and preparing them for the future.

It is not just in developed countries that the importance of training residents in teaching skills is acknowledged. Some developing nations are also publishing their own experiences in this field. One study in Nigeria has highlighted the importance of bringing medical residents into the teaching scenario in countries with limited resources for hiring teachers.³⁸

RaT programs are delivered using many methods and a great variety of lengths of training. However, the resources available, and especially the expertise to lead the instruction, are not distributed equally among residency programs.

Data on training residents as teachers in Brazil is scarce despite the creation of new medical schools and the resulting increase in the number of students, which has brought difficult challenges to the educational system. There is an urgent need for better preparation of educators for teaching. Since 2006, Hospital das Clínicas of the Federal University of Pernambuco (Universidade Federal de Pernambuco, UFPE) has offered a 64-hour training program to medical residents within family medicine, with the aim of transforming this reality.³⁶

In the literature, it is stated that residents who teach have greater enthusiasm for teaching, with greater job satisfaction. Residents are also role models for medical students. It is possible that pedagogical development also has a positive effect on patient care, though addition of effects regarding physician communication skills, with better clinical outcomes

In an ongoing study, Brazilian residents at a public university are being trained in the One-Minute Preceptor³⁹ methodology, in comparison with controls. Our training education program combines theoretical and practical resources with a unique one-hour long dynamic. The partial results are very encouraging, with good validation scores for peer feedback based on the Stanford Faculty Development Program 26 (SFDP-26) instrument.⁴⁰ Our study also revealed that significant positive changes to the residents' teaching skills were achieved.

One innovative facet of this study was that it included scarce data on RaT programs in Brazil. We aimed to explore the reasons for implementing RaT instruction, given the fact that most Brazilian programs for residency training currently do not include RaT programs. We face the challenge of teaching pedagogical strategies for medical residents in this country. The positive effects may be of great significance for students, medical residents and patients.

The present study was susceptible to selection bias and reporting bias. We also only explored instruction of RaT program mode that had been introduced. Hence, little can be concluded regarding the effectiveness of any other form of instruction or the relevance of any other targeted skills.

CONCLUSIONS

It is necessary to add pedagogical training to the training for residents and others working in the Brazilian National Health System, regarding ethical, technical and scientific knowledge. Some successful initiatives for developing skills and attitudes within healthcare education have emerged. However, these are still insufficient.

The necessary expansion of medical residency programs and the already inflated number of medical schools in Brazil both require

qualified teachers. Implementing pedagogical training during residency training could improve clinical skills and patients' care.

An accurate diagnosis regarding the state of pedagogical attributions among medical residents in Brazil is urgently necessary. There is also a need to understand what tools are at their disposal to perform the act of teaching with quality and efficacy, cooperatively.

Considering the positive effects that have been demonstrated, it can be argued that all residency programs should require residents to undergo instruction relating to teaching skills.

Along with teacher development programs, training for medical residents as teachers has the capacity to boost the quality of healthcare education in Brazil.

REFERENCES

1. Santos EG, Ferreira RR, Mannarino VL, et al. [Assessment of preceptorship in general surgery residency in the operating room, comparison between a teaching hospital and a non-teaching hospital.] *Revista do Colégio Brasileiro de Cirurgiões*. 2012;39(6):547-52. doi: 10.1590/S0100-69912012000600017.
2. Martins JC. Vygotsky e o papel das interações sociais na sala de aula. *Reconhecer e Desvendar o Mundo*. In: Série Idéias, n. 28. São Paulo: FDE; 1997. p. 111-22.
3. Nunes MPT. Residência médica no Brasil: situação atual e perspectivas. *Cad ABEM*. 2004;1:30-2.
4. Sternszus R, Cruess S, Cruess R, Young M, Steinert Y. Residents as role models: impact on undergraduate trainees. *Acad Med*. 2012;87(9):1282-7. PMID: 22836846; doi: 10.1097/ACM.0b013e3182624c53.
5. Louie AK, Beresin EV, Coverdale J, et al. Residents as teachers. *Acad Psychiatry*. 2013;37(1):1-5. PMID: 23338863; doi: 10.1176/appi.ap.12110192.
6. Burgin S, Homayounfar G, Newman LR, Sullivan A. Instruction in teaching and teaching opportunities for residents in US dermatology programs: Results of a national survey. *J Am Acad Dermatol*. 2017;76(4):703-6. PMID: 28325391; doi: 10.1016/j.jaad.2016.08.043.
7. Stern DT. In search of the informal curriculum: when and where professional values are taught. *Acad Med*. 1998;73(10 Suppl):S28-30. PMID: 9795643.
8. Brown RS. House staff attitudes toward teaching. *J Med Educ*. 1970;45(3):156-9. PMID: 5418782.
9. Donovan A. Radiology resident teaching skills improvement: impact of a resident teacher training program. *Acad Radiol*. 2011;18(4):518-24. PMID: 21377594; doi: 10.1016/j.acra.2010.10.021.
10. Bing-You RG, Sproul MS. Medical students' perceptions of themselves and residents as teachers. *Med Teach*. 1992;14(2-3):133-8. PMID: 1406122.
11. Busari JO, Arnold A. Educating doctors in the clinical workplace: unraveling the process of teaching and learning in the medical resident as teacher. *J Postgrad Med*. 2009;55(4):278-83. PMID: 20083878; doi: 10.4103/0022-3859.58935.
12. Craig JL, Page G. Teaching in medicine: an elective course for third-year students. *Med Educ*. 1987;21(5):386-90. PMID: 3683234.
13. Morrison EH, Hafler JP. Yesterday a learner, today a teacher too: residents as teachers in 2000. *Pediatrics*. 2000;105(1 Pt 3):238-41. PMID: 10617729.
14. Education ACfGM. ACGME common program requirements. 2013. Requirement IV A. 2015;5:9. Available from: <https://medicine.umich.edu/sites/default/files/content/downloads/CPRs2013.pdf>. Accessed in 2018 (August 30).
15. Education LCoM. Functions and structure of a medical school: Standards for accreditation of medical education programs leading to the MD Degree: Liaison Committee on Medical Education; 2015. Available from: http://med.wmich.edu/sites/default/files/2015-16_Functions-and-Structure-2015-6-16.pdf. Accessed in 2018 (August 30).
16. Al Achkar M, Hanauer M, Morrison EH, Davies MK, Oh RC. Changing trends in residents-as-teachers across graduate medical education. *Adv Med Educ Pract*. 2017;8:299-306. PMID: 28496376; doi: 10.2147/AMEPS127007.
17. Ramani S, Mann K, Taylor D, Thampy H. Residents as teachers: Near peer learning in clinical work settings: AMEE Guide No. 106. *Med Teach*. 2016;38(7):642-55. PMID: 27071739; doi: 10.3109/0142159X.2016.1147540.
18. Karani R, Fromme HB, Cayea D, et al. How medical students learn from residents in the workplace: a qualitative study. *Acad Med*. 2014;89(3):490-6. PMID: 24448043; doi: 10.1097/ACM.0000000000000141.
19. Sheets KJ, Hankin FM, Schwenk TL. Preparing surgery house officers for their teaching role. *Am J Surg*. 1991;161(4):443-9. PMID: 2035763.
20. Jewett LS, Greenberg LW, Goldberg RM. Teaching residents how to teach: a one-year study. *J Med Educ*. 1982;57(5):361-6. PMID: 7069756.
21. Busari JO, Prince KJ, Scherpbier AJ, et al. How residents perceive their teaching role in the clinical setting: a qualitative study. *Med Teach*. 2002;24(1):57-61. PMID: 12098459; doi: 10.1080/00034980120103496.
22. Bordley DR, Litzelman DK. Preparing residents to become more effective teachers: a priority for internal medicine. *Am J Med*. 2000;109(8):693-6. PMID: 11099693.
23. Morrison EH, Shapiro JF, Harthill M. Resident doctors' understanding of their roles as clinical teachers. *Med Educ*. 2005;39(2):137-44. PMID: 15679680; doi: 10.1111/j.1365-2929.2004.02063.x.
24. Snell L. The Resident-as-Teacher: It's More Than Just About Student Learning. *J Grad Med Educ*. 2011;3(3):440-1. PMID: 22942984; doi: 10.4300/JGME-D-11-00148.1.
25. Morrison EH, Rucker L, Boker JR, et al. The effect of a 13-hour curriculum to improve residents' teaching skills: a randomized trial. *Ann Intern Med*. 2004;141(4):257-63. PMID: 15313741.
26. Wamsley MA, Julian KA, Wipf JE. A literature review of "resident-as-teacher" curricula: do teaching courses make a difference? *J Gen Intern Med*. 2004;19(5 Pt 2):574-81. PMID: 15109328; doi: 10.1111/j.1525-1497.2004.30116.x.

27. Post RE, Quattlebaum RG, Benich JJ 3rd. Residents-as-teachers curricula: a critical review. *Acad Med.* 2009;84(3):374-80. PMID: 19240450; doi: 10.1097/ACM.0b013e3181971ffe.
28. Dannaway J, Ng H, Schoo A. Literature review of teaching skills programs for junior medical officers. *Int J Med Educ.* 2016;7:25-31. PMID: 26826798; doi: 10.5116/ijme.5685.14da.
29. Kirkpatrick D. Evaluating training programs: the four levels, third edition. San Francisco: Berrett-Koehler Publishers. Available from: https://www.bkconnection.com/static/Evaluating_Training_Programs_EXCERPT.pdf. Accessed in 2018 (September 30).
30. Furney SL, Orsini AN, Orsetti KE, et al. Teaching the one-minute preceptor. A randomized controlled trial. *J Gen Intern Med.* 2001;16(9):620-4. PMID: 11556943.
31. Mann KV, Sutton E, Frank B. Twelve tips for preparing residents as teachers. *Med Teach.* 2007;29(4):301-6. PMID: 17786741; doi: 10.1080/01421590701477431.
32. Chokshi BD, Schumacher HK, Reese K, et al. A "Resident-as-Teacher" Curriculum Using a Flipped Classroom Approach: Can a Model Designed for Efficiency Also Be Effective? *Acad Med.* 2017;92(4):511-4. PMID: 28030417; doi: 10.1097/ACM.0000000000001534.
33. Dewey CM, Coverdale JH, Ismail NJ, et al. Residents-as-teachers programs in psychiatry: a systematic review. *Can J Psychiatry.* 2008;53(2):77-84. PMID: 18357925; doi: 10.1177/070674370805300202.
34. Weiss V, Needleman R. To teach is to learn twice. Resident teachers learn more. *Arch Pediatr Adolesc Med.* 1998;152(2):190-2. PMID: 9491047.
35. Falcone JL, Ferson PF, Hamad GG. S/he who can, does and teaches. S/he who cannot, doesn't. *J Surg Educ.* 2014;71(1):96-101. PMID: 24411431; doi: 10.1016/j.jsurg.2013.06.003.
36. Magalhães G. O Residente como Professor: Formação Docente no Programa de Residência em Medicina de Família e Comunidade da Universidade Federal de Pernambuco [Thesis]. São Paulo: Universidade Federal de São Paulo - Escola Paulista de Medicina; 2012. Available from: http://www2.unifesp.br/centros/cedess/mestrado/teses/tese_123_o_residente_como_professor_formacao_docente_no_programa_%20gustavo_magalhaes.pdf. Accessed in 2018 (May 29).
37. Morrison EH, Hollingshead J, Hubbell FA, et al. Reach out and teach someone: generalist residents' needs for teaching skills development. *Fam Med.* 2002;34(6):445-50. PMID: 12164622.
38. Owolabi MO, Afolabi AO, Omigbodun AO. Performance of residents serving as clinical teachers: a student-based assessment. *J Grad Med Educ.* 2014;6(1):123-6. PMID: 24701322; doi: 10.4300/JGME-D-13-00130.1.
39. Neher JO, Gordon KC, Meyer B, et al. A five-step "microskills" model of clinical teaching. *J Am Board Fam Pract.* 1992;5(4):419-24. PMID: 1496899.
40. Skeff KM, Stratos GA, Berman J, et al. Improving clinical teaching. Evaluation of a national dissemination program. *Arch Intern Med.* 1992;152(6):1156-61. PMID: 1599342.

This paper formed part of the development of a doctoral project on medical education, at the Federal University of Uberlândia (Universidade Federal de Uberlândia, UFU), Uberlândia (MG), Brazil

Sources of funding: The authors declare that they did not have any financial support

Conflict of interest: The authors declare that they did not have any conflicts of interest

Date of first submission: March 30, 2018

Last received: July 19, 2018

Accepted: August 6, 2018

Address for correspondence:

Saadallah Azor Fakhouri Filho
Av. Pará, 1.720 — Bloco 2H
Departamento de Clínica Médica, Universidade Federal de Uberlândia (UFU)
Campus Umuarama
Uberlândia (MG) — Brasil
CEP 38405-320
Tel. (+55 34) 3225-8621
E-mail: safakhouri@hotmail.com

Objective Structured Teaching Examination (OSTE): An Underused Tool Developed to Assess Clinical Teaching Skills.

A Literature Narrative Review

Key words: Education, Medical. Educational Measurement. Professional Competence

Abstract:

Background: There are plenty of options for evaluating medical students and medical residents' clinical skills. Objective structured clinical evaluations (OSCEs) emerged as a powerful and reliable tool for assessing multiple cognition domains of clinical expertise. In the same way as OSCEs emerged to assess clinical skills, the objective structured teaching evaluations (OSTEs) came to light as a promising and unbiased intervention to evaluate the act of clinical teaching. **Design and Setting:** Narrative review developed at Universidade Federal de Uberlandia, Brazil. **Methods:** We performed a literature search about OSTEs using MEDLINE, via PubMed, and LILACS, via Biblioteca Virtual em Saude, databases. SciELO library was also searched for Brazilian papers. Systematic reviews, reviews and randomized controlled trials assessing specifically the performance of OSTEs in the development of faculty and medical residents were then selected. **Results:** Our search retrieved 178 papers, from which 40 were considered eligible for intensive review. Most of the selected studies reported positive effects from OSTE activities. However, there is little quantitative data to support its impact on teaching skills improvement. **Conclusions:** Considering that OSCEs became a widespread tool to assess medical students and residents' clinical skills, it is high time to incorporate OSTEs in the way we evaluate teaching skills in Brazil. There is encouraging data from abroad to support the implementation of this assessment tool in our country. The net benefit would possibly involve medical students, residents and faculty in bringing awareness about the importance of excelling in teaching skills.

Introduction:

Teaching is a complex activity that needs multiple methods for proper evaluation. One can directly examine teaching performances or highlight students results. All methods carry their strong and weak points and any isolated assessment will provide an incomplete image of a faculty's teaching skill.

Goe et.al¹ describe 5 points to measure teaching efficacy: (1) helping with learning, (2) contributing to the enhancement of academic, social and attitudinal performance, (3) use of multiple resources to engage students in learning opportunities, (4) promoting civic values and respect to diversity, (5) collaborating with the institution, peers and families to help students succeeding, particularly those with more chances to fail.

The most used method to assess the act of teaching is through direct observation of the activity, which can be done by directors, other faculty members (supervisors), external evaluators (locally or with the aid of video recorded performances). These evaluations are, in general, more time and resource consuming when compared to estimates centered in tests based on teaching efficacy, such as national and international exams on student's performances. Standardized tests provide, in a more economical and efficient way, reliable measures of the content acquired by students.²

Evaluations are routine events in the lives of university students worldwide. Their methods take different shapes, one of them being those performed in simulated situations, developed specifically to assess clinical skills. Since the famous neurologist Howard Barrows³ used trained actors to evaluate his residents' examination skills in a simulated ambient in the 1960's, this assessment method spread fast in the academic community.

Nowadays, simulated events are used not only in medical institutions. A wide diversity of careers, especially those linked to health, embraced this resource as a means of evaluation.

The authors of this narrative review are working in the development of residents as teachers, using the OSTE as a means of evaluation. Considering the need for refinement in faculty development and the importance of medical residents' teaching skills awareness, we judged of utmost interest to introduce, report and disseminate the concept of clinical teaching skill's assessment. In this paper we describe the main evidences that give support for implementing the OSTE method as an objective, summative and formative tool.

The main objectives of this narrative review were to introduce the concept of evaluating the act of teaching, through systematic observation of a simulated activity and summarize the evidences that give support to the dissemination of the method "Objective Structured Teaching Examination/Evaluation/Exercise/Encounter- OSTE"

Methods:

We conducted a literature review using MEDLINE, via PubMed, and LILACS, via Biblioteca Virtual emSaude, databases to extract relevant articles which describe simulated activities with the specific aim of evaluating teaching skills. We have also searched SciELO library for Brazilian papers. Key words used in the literature search included "Objective Structured Teaching Examination", "Objective Structured Teaching Evaluation", "Objective Structured Teaching Encounter", "Objective Structured Teaching Exam", "Observed Structured Teaching Examination" and teaching AND evaluation. MeSH terms used in the search were educational measurement/methods, professional competence, teaching/methods and teaching /standards.

Relevant publications included original articles, systematic reviews, critical reviews, randomized controlled trials, guidelines from medical education experts and material from pioneer authors in this field. Attention was given to papers focusing on pre and post-OSTE intervention designs, validity and reliability assessments, experimental work and publications describing the process of creating OSTE scenarios.

Results:

One hundred and seventy-eight articles were identified and 40 contained information regarding the specific topic of evaluating clinical teaching skills through objective and structured simulated activities. Searches of primary publications referenced in other articles were also included. These were selected for intensive review and were analyzed by two authors, SAF and MPTN. The search in LILACS and SciELO using Portuguese key words yielded no results on the specific topic of interest. **Table 1** describes the number of texts extracted from each database.

The majority of studies included in this review reported positive responses from participants in the activities involving OSTEs, faculty and residents alike. There was a trend towards lack of robust information about quantitative data regarding the same activity. Overall, the recommendations are in favor of the dissemination of OSTEs as a means of assessing clinical teaching skills. **Table 2** describes the most relevant studies included in this review and summarizes their main objectives and results.

Discussion:

1- Defining the Instrument:

In medical education we routinely perform simulated assessments of clinical skills, known as “Objective Structured Clinical Evaluations” (OSCEs). At our institutions, OSCEs are used with summative, formative and selective purposes, such as in medical residents’ entry contests. In the same way that we can put clinical skills under scrutiny through simulated experiences, some authors hypothesized that teaching skills could also be assessed using the same principles.⁴⁻⁷

Clinical training is vital for producing medical staff capable of delivering high quality assistance. Successful clinical teachers actively engage their students in patient care, providing constructive counseling and feedback.

Objective Structured Teaching Examinations/Encounters/Evaluations/Exercises (OSTES) came to light as a promising and unbiased intervention to assess the act of teaching.

What are the reasons commonly described to use the OSTE as a means of assessing teaching skills of teachers/facilitators? This method promotes rapid and rigorous assessment, enables practice of specific skills in a realistic atmosphere with immediate feedback and potential benefits for raters.⁸

Studies using the OSTE were started during the 1990's. **Table 3** details how simulated activities designed to assess teaching skills evolved through the years.^{4-7,9-15} This table was modified from material produced in a workshop performed in 2006 by Wamsely et.al,⁸ from the University of California, which specifically described the OSTE as a novel tool designed to enhance teaching skills and techniques.

OSTE's consist of simulated scenarios of teaching environments which use standardized students with the aim of providing objective and immediate feedback to a teacher or medical resident participating in the activity.¹⁶

Most of the literature describe OSTEs as 10 to 15 minutes encounters where standardized students interact with a teacher in video recorded sessions discussed in large debriefing groups.¹⁷

It is an opportunity for the teacher to show how he/she teaches and interact with standardized students in a controlled and thus protected scenario. There is the possibility to enact daily situations and repeat them as necessary.

During an OSTE activity participants go through sequential stations, each consisting of a different clinical teaching situation. Standardized students are used, and a simulated patient can be included depending on the purpose of the specific station.

Both verbal and non-verbal communication skills are evaluated. Feedback is given about different aspects of the teacher's performance. Feedback can also be given from the standardized student. Necessary materials to perform the activity include descriptions of the situations in every detail, good training and precise orientations to the standardized student and instructors and a reliable and validated instrument to rate the OSTE. Standardization must occur because real students or trained actors follow a script with pre-defined objectives.

2- Instrument Applications:

Teaching evaluations in OSTE scenarios are used with distinct purposes in American academic institutions. Some authors utilized the OSTE to evaluate teaching skills of medical preceptors, faculty from multiple schools and residents.^{7,10,11,18,19} Mc Sparron et.al²⁰ developed an OSTE with the specific aim of assessing faculty skills in orienting a procedure. The task consisted in guiding a standardized student required to place a central line in a model developed for this purpose. The activity was named "prOSTE" by the authors who reported better performances from standardized students when instructors offered positive feedback and suggestions for improvement and when the procedure was explained step by step. Another use for this method was to observe medical teachers or residents' teaching skills and try to enhance them through effective feedback. Cerrone et.al²¹ described an activity aimed at using the OSTE as a means of evaluating and promoting chief residents as emotionally intelligent leaders. The authors cite emotional intelligence as a core competence of leadership required from these residents. OSTE scenarios for this activity addressed professionalism, confidentiality, team work and

escalation issues (conflict resolution with higher hierarchy team members). The chief-residents were divided into pairs and participated on simulations of authentic incidents which were not identified. They conclude affirming that OSTEs combined with debriefing sessions post-activities can enable a powerful platform where those being evaluated (teacher, residents) would demonstrate competencies in areas frequently ignored, omitted or even considered “hidden”. Finally, the strategy was also used to evaluate interventions designed to enhance teaching skills of faculty and medical residents from different institutions.^{14,22,23} Steinert et.al²⁴ in a review of 303 publications on faculty development programs, stress specifically the importance of using OSTEs to measure the scores pre and post-interventions.

3- Using the Instrument Beyond Medicine:

OSTES are nowadays not restricted to medicine or to preceptors and residents working in large university hospitals. Mc Andrews et.al²⁵ developed an activity based on the OSTE to measure the response of faculty from the dentistry department, located at New York University, to a development program in teaching skills. Sixteen faculty members participated in a study measuring scores pre and post-intervention. Three stations were created to evaluate 15 domains of teaching. There was statistically significant improvement in all tested domains’ scores after exposure to an intervention based on faculty development. Mc Cutcheon et.al²⁶ describe a study aimed at training pharmacy preceptors in interprofessional educational activities. One OSTE station was created in such a way that two standardized students, one from nursery and one from pharmacy, interacted with a patient diagnosed with asthma. Data analysis showed statistically significant improvement in all 15 self-confidence items expressed by the participants. The authors highlight some of these items: guiding students from different specialties, facilitating a simulated activity, leading a debriefing session and discussing the core competencies in interprofessional education.

4- Rating OSTE’s Participants:

The OSTE is a very versatile modality of evaluation. It can be used with educational and assessment purposes. Currently, the OSTE is the gold standard instrument to assess medical residents’ teaching skills.²⁵ It also qualifies itself as the most accurate measure to evaluate faculty development programs.²⁴ To provide unbiased, valid and reliable scores, activities such as the OSTE, require appropriate instruments of measurement. It’s imperative that these instruments possess high measurement properties, thus, contributing with reliable and relevant feedback with respect to qualities and deficiencies demonstrated by the individuals being evaluated.

Review studies identified more than 32 different instruments developed to score simulated activities in clinical teaching. These are essential part in the process of continuous qualification in medical education and development of clinical teaching competencies.²⁷

Two scoring instruments are validated and widely used by the academic community: (1) "Stanford Faculty Development Program-SFDP26",²⁸ (2) "System for Evaluation of Teaching Qualities-SETQ".²⁹ The SFDP26 was originally developed in the United States and evaluates the following teaching skills: (1) establishing the learning environment, (2) session control, (3) communicating learning goals, (4) facilitating retention and understanding, (5) evaluation of previous knowledge, (6) promoting self-directed learning, (7) giving feedback.

5- Instrument Advantages:

Irby et.al³⁰ describe the main reasons for using the OSTE:(1) rapid and rigorous evaluation of clinical teaching skills, (2) provides time for practice of specific teaching skills,(3) creates a practical and realistic atmosphere, (4) provides immediate feedback, (5) potential benefit to standardized students who acquire additional knowledge while they rehearse for the stations and while giving feedback to the person under evaluation.

According to Wamsley et.al,³¹ teaching skill's quality is obtained, in most scenarios, through students' personal opinions. This type of evaluation is susceptible to the teacher's charisma and to her/his communication skills. In this circumstance, the picture of a good teacher can be contaminated by personal interactions in daily activities, being in a classroom or in hospital wards; this interaction bringing a positive or negative bias. The OSTE has the advantage of attenuating that potential contamination by specifically assessing the teacher's professional performance. This can be achieved using standardized situations, instruments, students and even patients.

Another attractive feature the OSTE brings is the opportunity to deliver an immediate debriefing session post-activity. The debriefing process is considered by many researchers as the most important component of any learning experience based on simulation. It's a direct function of the learner's capacity to reflect on his/her performance in the activity. Decker et al.³²affirm that the learning process is dependent on the interaction between experience and reflection. On the reflection process, the same authors ponder that it is a conscious consideration of meanings and implications that arise from an action. All these processes would lead to assimilation of knowledge, skills and attitudes that did not exist before the intervention. The authors conclude citing the necessary conditions for an effective debriefing session: (1) facilitated by a competent individual,(2) conducted in an environment that favors knowledge acquisition and that guarantee confidentiality, trust, open communication, self-analysis and reflection,(3) facilitated by someone who directly observed the

simulation experience,(4) based on a work plan specifically structured for that task,(5) congruent with participants' objectives and previously established results.

Considering the present review,there is little or no knowledge about how to apply the OSTE method in Brazil, despite all the positive remarks regarding the activity. The authors of this review managed to successfully develop OSTE scenarios in a doctoral project still in progress. Four scenarios were created simulating the discussion of a case in emergency, ambulatory and surgical ward settings and teaching a simple procedure. We translated into Portuguese and then validated the SFDP26 instrument using a specialists' panel, trained arts students as standardized medicine pupils and finally received help from voluntary medical residents who performed in the proposed simulations (data not yet published).

6- Instrument Limitations:

According to the reviewed literature, the main limitations for developing the OSTE are: scarce financial and human resources to implement the activities and lack of time attributed to tasks deemed "less important". Training standardized students or hiring professional actors for the activities can be a limiting factor in institutions where simulations are not part of the local culture. There are also no data related to the financial impact of this kind of activity nor of its cost effectiveness.

Systematic review on the use and efficacy of the OSTE, published in 2011, evidenced little quantitative data to support its impact on teaching skills improvement.³³Besides that, most participants reported beneficial effects.The same authors cite also the lack of uniformity in the scenarios' scripts. Reliability of the rating instruments was moderate to high in most studies and rating scores appear not to be related to rater background, faculty or student. Only 4 studies evaluated specifically station validity, thus, there is little evidence concerned to OSTE validity to this date.^{11,13,18,34}

7- Brazilian Studies:

Our search using Portuguese key words in LILACS database and SciELO yielded no results. To our knowledge there is no experience using the OSTE method in Brazil. This narrative review is part of doctoral project, still in progress, aimed to assess residents' clinical teaching competencies before and after an intervention based on the One Minute Preceptor technique.³⁵The OSTE method was presented, based on our previously cited experience, in the Brazilian Congress of Medical Education 2017, as a workshop.

Conclusions:

Teaching is a complex activity that requires multiple methods for complete assessment. In the present review we elected an observational and standardized method- OSTE.

Adequate assessments of teachers' performances give them the opportunity to reflect on their teaching skills and consequently evolving during the process of knowledge acquisition. Teachers desire and need feedback, not only to their teaching moments, but also to enhance educational results. Evaluation systems directed to teachers are frequently proposed to offer feedback and guide the evolution of professional practice.³⁶

Assessing faculty, preceptors and medical residents is a subject of major controversy not only in Brazil. We believe that most teachers with well consolidated careers would feel uneasy to have their skills put under scrutiny of peers or students. This could be a reason to explain why the word evaluation, contained in the acronym OSTE, is frequently changed by encounter, exercise or situation.

In the same way that we brought to our reality the OSCE as a tool to assess medical students and residents, it is also possible to develop the OSTE dynamic in our practice. Some useful publications deal specifically with the process of creating an OSTE station and are excellent starting points. Boillat et al.¹⁶ describe 12 tips to construct an activity based on simulation to assess clinical teaching skills (**Table 4**). There is potential benefit from the OSTE in all educational levels since we are moving past the evaluation of reactions, the lower level in Kirkpatrick scale,³⁷ and reaching the understanding of attitudes and behavior change.

References:

1. Goe L, Croft A. Methods of Evaluating Teacher Effectiveness. Research-to-Practice Brief. National Comprehensive Center for Teacher Quality. 2009.
2. Stecher B, Garet M, Hamilton L, et al. Improving Teaching Effectiveness: Implementation: The Intensive Partnerships for Effective Teaching Through 2013–2014 Santa Monica, CA: RAND Corporation; 2016 [Available from: https://www.rand.org/content/dam/rand/pubs/research_reports/RR1200/RR1295/RAND_RR1295.pdf.
3. Barrows HS, Abrahamson S. The programmed Patient: A Technique for Appraising Student Performance in Clinical Neurology. J Med Educ. 1964;39:802-5. PMID:14180699
4. Simpson DE, Lawrence SL, Krogull SR. Using standardized ambulatory teaching situations for faculty development. Teach Learn Med: An International Journal. 1992;4(1):58-61. doi:10.1080/10401339209539534
5. Orlander JD, Bor DH, Strunin L. A structured clinical feedback exercise as a learning-to-teach practicum for medical residents. Acad Med. 1994;69(1):18-20. PMID:8285990
6. Lesky LG, Wilkerson L. Using "standardized students" to teach a learner-centered approach to ambulatory precepting. Acad Med. 1994;69(12):955-7. PMID:7999182. doi:10.1097/00001888-199412000-00004
7. Prislun MD, Fitzpatrick C, Giglio M, Lie D, Radecki S. Initial experience with a multi-station objective structured teaching skills evaluation. Acad Med. 1998;73(10):1116-8. PMID:9795632. doi:10.1097/00001888-199810000-00024
8. Wamsley M, Julian K, Morrison E, Zabar S. Workshop F06: Introduction to the Objective Structured Teaching Evaluation (OSTE) – A Novel Tool for Evaluating Teaching Skills. 2006.

9. EK K. Multiple station exams and teaching exercises (MSEs) for teaching about teaching. In: Scherpbier AJJA vdVC, Rethans JJ, van der Steeg AFW, editor. *Advances in Medical Education*. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1997. p. 226-7.
10. Dunnington GL, DaRosa D. A prospective randomized trial of a residents-as-teachers training program. *Acad Med*. 1998;73(6):696-700. PMID:9653409. doi:10.1097/00001888-199806000-00017
11. Schol S. A Multiple-station Test of the Teaching Skills of General Practice Preceptors in Flanders, Belgium. *Acad Med*. 2001;76(2):176-80. PMID:11158840. doi:10.1097/00001888-200102000-00018
12. Gelula MH, Yudkowsky R. Microteaching and standardized students support faculty development for clinical teaching. *Acad Med*. 2002;77(9):941. PMID:12228115. doi:10.1097/00001888-200209000-00048
13. Morrison EH, Boker JR, Hollingshead J, et al. Reliability and validity of an objective structured teaching examination for generalist resident teachers. *Acad Med*. 2002;77(10):S29-S32. PMID:12377697
14. Stone S, Mazor K, Devaney-O'Neil S, et al. Development and implementation of an objective structured teaching exercise (OSTE) to evaluate improvement in feedback skills following a faculty development workshop. *Teach Learn Med*. 2003;15(1):7-13. PMID:12632702. doi:10.1207/S15328015TLM1501_03
15. Morrison EH, Rucker L, Boker JR, et al. A pilot randomized, controlled trial of a longitudinal residents-as-teachers curriculum. *Acad Med*. 2003;78(7):722-9. PMID:12857695
16. Boillat M, Bethune C, Ohle E, Razack S, Steinert Y. Twelve tips for using the objective structured teaching exercise for faculty development. *Med Teach*. 2012;34(4):269-73. PMID:22455695. doi:10.3109/0142159X.2011.599891
17. Sturpe DA, Schaivone KA. A primer for objective structured teaching exercises. *Am J Pharm Educ*. 2014;78(5):104. PMID:24954244. doi:10.5688/ajpe785104
18. Zabar S, Hanley K, Stevens DL, et al. Measuring the competence of residents as teachers. *J Gen Intern Med*. 2004;19(5p2):530-3. PMID:15109318. doi:10.11/j.1525-1497.200430219.x
19. Gaba ND, Blatt B, Macri CJ, Greenberg L. Improving teaching skills in obstetrics and gynecology residents: evaluation of a residents-as-teachers program. *Am J Obstet Gynecol*. 2007;196(1):87.e1-7. PMID:17240248. doi:10.1016/j.ajog.2006.09.037
20. McSparron JI, Ricotta DN, Moskowitz A, et al. The ProSTE: identifying key components of effective procedural teaching. *Ann Am Thorac Soc*. 2015;12(2):230-4. PMID:25564926. doi:10.1513/annalsATS.201406-237BC
21. Cerrone SA, Adelman P, Akbar S, Yacht AC, Fornari A. Using Objective Structured Teaching Encounters (OSTEs) to prepare chief residents to be emotionally intelligent leaders. *Med Educ Online*. 2017;22(1):1320186. PMID:28460596. doi:10.1080/10872981.2017.1320186
22. Morrison EH, Rucker L, Boker JR, et al. The effect of a 13-hour curriculum to improve residents' teaching skills: a randomized trial. *Ann Intern Med*. 2004;141(4):257-63. PMID:15313741.
23. Zackoff M, Jerardi K, Unaka N, Sucharew H, Klein M. An observed structured teaching evaluation demonstrates the impact of a resident-as-teacher curriculum on teaching competency. *Hosp Pediatr*. 2015;5(6):342-7. PMID:26034166. doi:10.1542/hpeds.2014-0134
24. Steinert Y, Mann K, Centeno A, et al. A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education: BEME Guide No. 8. *Med Teach*. 2006;28(6):497-526. PMID:17074699. doi:10.1080/01421590600902976
25. McAndrew M, Eidtson WH, Pierre GC, Gillespie CC. Creating an objective structured teaching examination to evaluate a dental faculty development program. *J Dent Educ*. 2012;76(4):461-71. PMID:22473558

26. McCutcheon LR, Whitcomb K, Cox CD, et al. Interprofessional objective structured teaching exercise (iOSTE) to train preceptors. *Curr Pharm Teach Learn*. 2017;9(4):605-15. PMID:29233433.doi:10.1016/j.cptl.2017.03.023
27. Fluit CR, Bolhuis S, Grol R, Laan R, Wensing M. Assessing the quality of clinical teachers. *J Gen InternMed*. 2010;25(12):1337-45. PMID:20703952.doi:10.1007/s11606-010-1458-y
28. Williams BC, Litzelman DK, Babbott SF, Lubitz RM, Hofer TP. Validation of a global measure of faculty's clinical teaching performance. *Acad Med*. 2002;77(2):177-80. PMID:11841985
29. Lombarts M, Arah O, Busch O, Heineman M. Using the SETQ system to evaluate and improve teaching qualities of clinical teachers. *Ned Tijdschr Geneesk*. 2010;154:A1222-A. PMID:20170574
30. Irby DM. Teaching and learning in ambulatory care settings: a thematic review of the literature. *Acad Med*. 1995;70(10):898-931. PMID:7575922
31. Wamsley MA, Julian KA, Vener MH, Morrison EH. Using an objective structured teaching evaluation for faculty development. *Med Educ*. 2005;39(11):1160-1. PMID:16262834.doi:10.1111/j.1365-2929.2005.02287.x
32. Decker S, Fey M, Sideras S, et al. Standards of best practice: Simulation standard VI: The debriefing process. *Clinical Simulation in Nursing*. 2013;9(6):S26-S9. doi:10.1016/j.ecns.2013.04.008
33. Trowbridge RL, Snyderman LK, Skolfield J, Hafler J, Bing-You RG. A systematic review of the use and effectiveness of the Objective Structured Teaching Encounter. *Med Teach*. 2011;33(11):893-903. PMID:22022900.doi:10.3109/0142159X.2011.577463
34. Quirk M, Mazor K, Haley H-L, et al. Reliability and Validity of Checklists and Global Ratings by Standardized Students, Trained Raters, and Faculty Raters in an Objective Structured Teaching Environment. *Teach Learn Med*. 2005;17(3):202-8. PMID:16042515.doi:10.1207/s15328015tlm1703_2
35. Neher JO, Gordon KC, Meyer B, Stevens N. A five-step "microskills" model of clinical teaching. *J Am Board Fam Pract*. 1992;5(4):419-24. PMID:1496899
36. Tucker PD, Stronge JH. Linking teacher evaluation and student learning: ASCD; 2005.
37. Kirkpatrick D. Evaluation, Training Programs: the four levels, San Francis co: Berrett-Koehler Publishers. learning transfer. *International Journal of Training and Development*. 1994;6(1):36-48.
https://www.bkconnection.com/static/Evaluating_Training_Programs_EXCERPT.pdf

Table 1: Databases and number of articles extracted

Database	Number of articles
MEDLINE	178
LILACS	0

Table 2: Relevant Studies on the OSTE* Contained in this Review

Authors	Year	Study Type	Participants (n)	Specialty	Objectives	Conclusions
Orlander et al.⁵	1994	Qualitative/experimental	23 residents 47 interns	Internal Medicine	Obtain residents' and interns' opinions after participating in a Clinical Feedback Exercise	Residents reported better attitudes over clinical teaching skills and contact with interns
Lesky et al.⁶	1994	Qualitative/Experimental	Ambulatorial Preceptors and standardized students (sample not informed)	Not informed	Develop an ambulatorial preceptorship program and help preceptors to adopt a student-centered teaching approach	Preceptors who participated in the activity reported better comprehension of their clinical teaching competencies
Schol et al.¹¹	2001	Validity and Reliability	35 residents acting as preceptors	General Practice	Evaluate the Validity and Reliability of an evaluation with 7 OSTE scenarios	This type of assessment provides important informations about the levels of competency in clinical teaching
Morrison et. al¹³	2002	Validity and Reliability	23 second-year residents	Internal Medicine, Pediatrics, Family Medicine	Evaluate the validity and reliability of an 8-station OSTE developed specifically for residents involved in teaching activities	Cronbach's alpha >0,9 for all 8 stations; 92% of the residents reported that the stations reflected with realism important clinical teaching skills
Stone et al.¹⁴	2003	Pre/ post-intervention design	40 teachers	Family Medicine Pediatrics Internal Medicine	Develop and implement an OSTE to assess a faculty development module	OSTEs can be more sensitive to detect skills more readily applicable to daily practice
Sturpe et al.¹⁷	2014	Descriptive	NA **	NA	Describe how to develop a scenario to	When used together with other teaching

					evaluate clinical skills	strategies, the OSTE can enhance faculty's teaching strategies
Cerrone etal.²¹	2017	Intervention Non-randomized	80 chief-residents	15 specialties	Develop a program capable of enhancing chief-residents' emotional intelligence	There was statistically significant improvement in OSTE's scores after intervention
Zackoff etal.²³	2015	Prospective with intervention group and blinded raters	49 residents	Pediatrics	Create an OSTE capable of detecting changes in residents' teaching skills after participation in a teaching skills' development program	The OSTE showed efficacy in detecting changes in clinical teaching skills
Steinert etal.²⁴	2006	Systematic Review	NA	Multiple	Describe the evidences that give support to programs aimed at developing teaching skills	Participants generally reported positive changes after specific interventions, but methodological limitations remain and should be addressed
Mc Andrew etal.²⁵	2011	Descriptive	12 teachers	Dentistry	Describe the creation of an OSTE to assess a faculty development program in dentistry	The use of OSTE's not only enhances the levels faculty development programs, it also enriches efforts to foster teaching in dentistry
Mc	2017	Pilot,	23	Pharmacy	Evaluate	Pharmacy

Cutcheon et.al.²⁶		qualitative, Pre/post-OSTE	preceptors		interprofessional teaching skills	preceptors demonstrated better teaching skills and self-confidence after OSTE participation
Wamsley et.al.³¹	2005	Qualitative/ Experimental	16 teachers	Internalmedicine, Pediatrics, Family Medicine	Develop teaching skills in faculty from different specialties	Teachers reported they would change their teaching styles after OSTE participating; Standardized students reported more interest in teaching after the activity
Trowbridge et.al.³³	2011	Systematic Review	NA	Multiple	Describe the use of OSTE and the evidences that give support to its use	The OSTE is a promising innovation with potential application in the assessment and promotion of clinical teaching; new research is necessary to define the appropriate stations' design

*OSTE: Objective Structured Teaching Examination

**NA: not applicable

Table 3*: OSTE Evolution in the last decades**

Year	Author	Name attributed to activity	Target audience
1992	Simpson e cols. ⁴	SATS: Standardized Ambulatory Teaching Situations	Faculty
1994	Orlander e cols. ⁵	CFE: Clinical Feedback Exercise	Residents
1994	Lesky e cols. ⁶	OSTE: Objective Structured Teaching Exercise	Faculty
1997	Kachur e cols. ⁹	Multiple Station Exam/ Teaching Exercise	Faculty
1998	Dunnington e cols. ¹⁰	OSTE: Objective Structured Teaching Evaluation	Residents
1998	Prislin e cols. ⁷	OSTE: Objective Structured Teaching Evaluation	Faculty
2001	Schol e cols. ¹¹	MSTAT: Multiple-Station Teaching Assessment Test	Faculty
2002	Gelula e cols. ¹²	Teaching Encounters	Faculty
2002	Morrison e cols. ¹³	OSTE: Objective Structured Teaching Examination	Residents
2003	Stone e cols. ¹⁴	OSTE: Objective Structured Teaching Exercise	Faculty
2003	Morrison e cols. ¹⁵	OSTE: Objective Structured Teaching Examination	Residents

*Modified from: Timeline developed by Elizabeth Kachur, PhD⁸

**OSTE: Objective Structured Teaching Examination

Table4*: Tips for producing anOSTEscenario**

1	Establish specific goalsfor the activity in advance
2	Determine context and target audience for the activity
3	Identify specific teaching skills to be evaluated
4	Set the scenario: use real cases based on situations that occurred in the place where the activity is being performed
5	Develop the assessment tool: checklists, self-evaluations and those to be performed by standardized students or peers
6	Choose the standardized student: medicine or other pupils, trained actors, faculty, institution staff
7	Train the standardized student according to the scenario and teaching skill to be assessed
8	Test the scenario before the main event to correct potential errors and maximize case fidelity
9	Protect the teacher being evaluated: follow strictly the principles of effective debriefing and feedback to reduce stress from the activity
10	Incorporate this simulation method to faculty development programs in the specific context of your institution
11	Emphasize the positive aspects of the method to enhance adherence from faculty or residents
12	Evaluate the activity after simulations' conclusion: use pre and post-activity's interviews, written evaluations or qualitative comments

*Modified from Boillat et. al¹⁶

**OSTE: Objective Structured Teaching Examination

UNIVERSIDADE FEDERAL DE
UBERLÂNDIA/MG

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Instrução Formal de Médicos Residentes em Técnicas de Ensino

Pesquisador: Saadallah Azor Fakhouri Filho

Área Temática:

Versão: 3

CAAE: 42685915.6.0000.5152

Instituição Proponente: HOSPITAL DE CLINICAS DA UNIVERSIDADE FEDERAL DE UBERLANDIA

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 1.372.171

Apresentação do Projeto:

Conforme apresenta o protocolo:

" Durante toda a historia da medicina seus praticantes de alguma forma ou de outra sempre estiveram envolvidos em atividades de ensino. Houve contribuições importantes de gregos, romanos, árabes e a partir do Renascimento a Europa passou a exercer a hegemonia relacionada ao conhecimento médico. O modelo de ensino Europeu foi a base para a pratica da medicina ate o inicio do século XX. Este modelo foi inclusive aquele utilizado nos primórdios das atividades de ensino nas faculdades de medicina brasileiras.

Metodologia Proposta: Os residentes da amostra serão submetidos a uma avaliação previa de seus conhecimentos em técnicas de ensino. A avaliação será feita através do instrumento denominado Organized and Structured Teaching Examination (OSTE). Desta forma toda a amostra será pontuada quanto ao desempenho em diferentes estações onde serão simuladas atividades de ensino. Após a avaliação inicial a amostra será dividida de forma randomizada em dois grupos. Um dos grupos será submetido a intervenção que consiste em apresentar a técnica de ensino chamada One Minute Preceptor. O grupo controle será submetido a uma intervenção que não mostra relação alguma com teorias relacionadas ao ensino. Desta forma os residentes não terão conhecimento a respeito da alocação inicial quanto aos grupos controle e intervenção. Após a conclusão das

Endereço: Av. João Naves de Ávila 2121- Bloco "1A", sala 224 - Campus Sta. Mônica

Bairro: Santa Mônica

CEP: 38.408-144

UF: MG

Município: UBERLÂNDIA

Telefone: (34)3239-4131

Fax: (34)3239-4335

E-mail: cep@propp.ufu.br

Continuação do Parecer: 1.372.171

atividades de capacitação em técnicas de ensino os residentes serão novamente submetidos a avaliação com o método OSTE e os resultados serão comparados com a pontuação inicial obtida pelos dois grupos. A segunda avaliação será efetuada 2 semanas após a intervenção inicial. Após intervalo de 6 meses o grupo controle será submetido a intervenção através do One Minute Preceptor e o grupo restante será introduzido a intervenção não relacionada ao tema em estudo. Novamente serão submetidos a avaliação com o OSTE após as intervenções com intervalo de tempo de 2 semanas. Com relação ao tamanho da amostra no Brasil consideramos que a ideia geral da pesquisa, não é partir do universo em estudo, mas do tamanho da amostra a ser analisada: Vamos considerar um $\alpha = 5\%$, $\beta = 20\%$ e poder de estudo de $1 - \beta = 0,8$. Segundo os estudiosos do assunto, como trabalharemos com grupo controle pareado com cruzamento e medidas repetidas (3 medidas do mesmo sujeito) a amostra por grupo pode variar de 23 a 51 (ou seja de 46 a 102 indivíduos no total).

Critério de Inclusão: Médicos residentes matriculados nos primeiros e segundos anos dos Programa de Residência em Clínicas Médica, Dermatologia e Moléstias Infecciosas da Faculdade de Medicina, no Hospital de Clínicas de Uberlândia da Universidade Federal de Uberlândia.

Critério de Exclusão: Serão excluídos os residentes que na época das avaliações e intervenções não puderem comparecer para as atividades."

Objetivo da Pesquisa:

Segundo o projeto:

O objetivo primário desta proposta é o de oferecer instrução formal em técnicas de ensino para médicos residentes com finalidade de formar médicos preceptores e desenvolver as técnicas de comunicação dos mesmos. Médicos mais capazes de se comunicar serão teoricamente mais eficazes na tarefa de transmitir conhecimento para seus pares ou para seus pacientes. Em última análise o que se espera é uma melhor qualidade nos cuidados em saúde e nas relações interpessoais.

O Objetivo secundário seria a inclusão das técnicas de ensino e dos métodos de avaliação nos futuros currículos dos programas de residência médica sabendo do efeito positivo que desempenham sobre o treinamento em serviço dos residentes.

Avaliação dos Riscos e Benefícios:

Segundo os pesquisadores:

Riscos: Os riscos associados ao presente estudo consistem na possível identificação dos sujeitos da pesquisa e na identificação dos resultados de desempenho obtidos por cada participante nas

Endereço: Av. João Naves de Ávila 2121- Bloco "1A", sala 224 - Campus Sta. Mônica
Bairro: Santa Mônica **CEP:** 38.408-144
UF: MG **Município:** UBERLÂNDIA
Telefone: (34)3239-4131 **Fax:** (34)3239-4335 **E-mail:** cep@propp.ufu.br

Continuação do Parecer: 1.372.171

diferentes avaliações. No sentido de minimizar os riscos, na identificação dos sujeitos, as avaliações serão realizadas por pessoas que não estarão envolvidas no processo inicial de orientação. Para os riscos na identificação dos resultados de desempenho, os mesmos, após a aplicação, serão arquivados em banco de dados com acesso exclusivo aos executores do presente estudo.

Benefícios: O projeto visa melhoria nas capacidades de comunicação e aplicação de técnicas de ensino, e na formação de futuros médicos preceptores que atuaram no ensino de outros residentes.

Comentários e Considerações sobre a Pesquisa:

Não há.

Considerações sobre os Termos de apresentação obrigatória:

Termos apresentados.

Recomendações:

Não há.

Conclusões ou Pendências e Lista de Inadequações:

As pendências apontadas no parecer consubstanciado número 1.257.422, de 04 de Outubro de 2015, foram atendidas.

De acordo com as atribuições definidas na Resolução CNS 466/12, o CEP manifesta-se pela aprovação do protocolo de pesquisa proposto.

O protocolo não apresenta problemas de ética nas condutas de pesquisa com seres humanos, nos limites da redação e da metodologia apresentadas.

Considerações Finais a critério do CEP:

Data para entrega de Relatório Final ao CEP/UFU: Junho de 2016.

OBS.: O CEP/UFU LEMBRA QUE QUALQUER MUDANÇA NO PROTOCOLO DEVE SER INFORMADA IMEDIATAMENTE AO CEP PARA FINS DE ANÁLISE E APROVAÇÃO DA MESMA.

O CEP/UFU lembra que:

a- segundo a Resolução 466/12, o pesquisador deverá arquivar por 5 anos o relatório da pesquisa e os Termos de Consentimento Livre e Esclarecido, assinados pelo sujeito de pesquisa.

Endereço: Av. João Naves de Ávila 2121- Bloco "1A", sala 224 - Campus Sta. Mônica

Bairro: Santa Mônica

CEP: 38.408-144

UF: MG

Município: UBERLÂNDIA

Telefone: (34)3239-4131

Fax: (34)3239-4335

E-mail: cep@propp.ufu.br

Continuação do Parecer: 1.372.171

b- poderá, por escolha aleatória, visitar o pesquisador para conferência do relatório e documentação pertinente ao projeto.

c- a aprovação do protocolo de pesquisa pelo CEP/UFU dá-se em decorrência do atendimento a Resolução CNS 466/12, não implicando na qualidade científica do mesmo.

Orientações ao pesquisador :

- O sujeito da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado (Res. CNS 466/12) e deve receber uma via original do Termo de Consentimento Livre e Esclarecido, na íntegra, por ele assinado.
- O pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado e descontinuar o estudo somente após análise das razões da descontinuidade pelo CEP que o aprovou (Res. CNS 466/12), aguardando seu parecer, exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade de regime oferecido a um dos grupos da pesquisa que requeiram ação imediata.
- O CEP deve ser informado de todos os efeitos adversos ou fatos relevantes que alterem o curso normal do estudo (Res. CNS 466/12). É papel de o pesquisador assegurar medidas imediatas adequadas frente a evento adverso grave ocorrido (mesmo que tenha sido em outro centro) e enviar notificação ao CEP e à Agência Nacional de Vigilância Sanitária – ANVISA – junto com seu posicionamento.
- Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas. Em caso de projetos do Grupo I ou II apresentados anteriormente à ANVISA, o pesquisador ou patrocinador deve enviá-las também à mesma, junto com o parecer aprobatório do CEP, para serem juntadas ao protocolo inicial (Res.251/97, item III.2.e).

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_472225.pdf	13/10/2015 10:21:12		Aceito
Outros	RespostaParecerConsuistnr1257422.doc	13/10/2015 10:20:41	Saadallah Azor Fakhouri Filho	Aceito
Projeto Detalhado / Brochura Investigador	ProjetoDetalhadoCorrig.doc	13/10/2015 10:19:01	Saadallah Azor Fakhouri Filho	Aceito

Endereço: Av. João Naves de Ávila 2121- Bloco "1A", sala 224 - Campus Sta. Mônica

Bairro: Santa Mônica

CEP: 38.408-144

UF: MG

Município: UBERLÂNDIA

Telefone: (34)3239-4131

Fax: (34)3239-4335

E-mail: cep@propp.ufu.br

Continuação do Parecer: 1.372.171

Outros	Resposta às pendências.doc	17/08/2015 09:11:56		Aceito
Outros	Instrumento de Coleta de Dados - Validado Port.docx	27/07/2015 17:51:46		Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE nova CEP.doc	27/07/2015 17:50:39		Aceito
Declaração de Pesquisadores	Autorização da Instituição.jpg	10/03/2015 12:03:47		Aceito
Declaração de Pesquisadores	Termo Compr Equip Exec.jpg	25/02/2015 10:31:42		Aceito
Declaração de Pesquisadores	Link dos Currículo Lattes.doc	25/02/2015 10:31:15		Aceito
Declaração de Pesquisadores	Instrumento Coleta de Dados - 02.JPG	25/02/2015 10:31:04		Aceito
Declaração de Pesquisadores	Instrumento Coleta de Dados - 01.JPG	25/02/2015 10:30:53		Aceito
Outros	Declaração Instituição.jpg	25/02/2015 10:30:36		Aceito
Folha de Rosto	Folha de Rosto.JPG	25/02/2015 10:28:20		Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

UBERLÂNDIA, 16 de Dezembro de 2015

Assinado por:

**Sandra Terezinha de Farias Furtado
(Coordenador)**

Endereço: Av. João Naves de Ávila 2121- Bloco "1A", sala 224 - Campus Sta. Mônica

Bairro: Santa Mônica

CEP: 38.408-144

UF: MG

Município: UBERLÂNDIA

Telefone: (34)3239-4131

Fax: (34)3239-4335

E-mail: cep@propp.ufu.br

USP - HOSPITAL DAS
CLÍNICAS DA FACULDADE DE
MEDICINA DA UNIVERSIDADE

PARECER CONSUBSTANCIADO DO CEP

Elaborado pela Instituição Coparticipante

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Instrução Formal de Médicos Residentes em Técnicas de Ensino

Pesquisador: Saadallah Azor Fakhouri Filho

Área Temática:

Versão: 1

CAAE: 42685915.6.3001.0068

Instituição Proponente: HOSPITAL DE CLINICAS DA UNIVERSIDADE FEDERAL DE UBERLANDIA

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 1.894.307

Apresentação do Projeto:

Este é um projeto universitário multicêntrico de pesquisa voltado para médicos residentes

Objetivo da Pesquisa:

Visa utilizar um instrumento padronizado, validado para a língua portuguesa, para avaliar a formação e capacitação dos residentes, de modo a aprimorar a metodologia e acumular informações úteis para um preparo de qualidade internacional

Avaliação dos Riscos e Benefícios:

Não haverá riscos, e as intervenções se limitarão à avaliação da qualidade e desempenho dos médicos residentes

Comentários e Considerações sobre a Pesquisa:

O projeto é atual, de qualidade, e se encontra corretamente elaborado

Considerações sobre os Termos de apresentação obrigatória:

O TCLE é adequado

Recomendações:

Não há

Conclusões ou Pendências e Lista de Inadequações:

Não há pendências.

Endereço: Rua Ovídio Pires de Campos, 225 5º andar

Bairro: Cerqueira Cesar

CEP: 05.403-010

UF: SP

Município: SAO PAULO

Telefone: (11)2661-7585

Fax: (11)2661-7585

E-mail: cappesq.adm@hc.fm.usp.br

Continuação do Parecer: 1.894.307

Considerações Finais a critério do CEP:

Em conformidade com a Resolução CNS nº 466/12 – cabe ao pesquisador: a) desenvolver o projeto conforme delineado; b) elaborar e apresentar relatórios parciais e final; c) apresentar dados solicitados pelo CEP, a qualquer momento; d) manter em arquivo sob sua guarda, por 5 anos da pesquisa, contendo fichas individuais e todos os demais documentos recomendados pelo CEP; e) encaminhar os resultados para publicação, com os devidos créditos aos pesquisadores associados e ao pessoal técnico participante do projeto; f) justificar perante ao CEP interrupção do projeto ou a não publicação dos resultados.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_839434 E1.pdf	20/12/2016 15:47:40		Aceito
Declaração de Pesquisadores	doc2.pdf	20/12/2016 15:46:06	Saadallah Azor Fakhouri Filho	Aceito
Declaração de Pesquisadores	doc1.pdf	20/12/2016 15:45:47	Saadallah Azor Fakhouri Filho	Aceito
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_839434 E1.pdf	08/12/2016 16:48:52		Aceito
Outros	RespostaParecerConsubistnr1257422.doc	13/10/2015 10:20:41	Saadallah Azor Fakhouri Filho	Aceito
Projeto Detalhado / Brochura Investigador	ProjetoDetalhadoCorrig.doc	13/10/2015 10:19:01	Saadallah Azor Fakhouri Filho	Aceito
Outros	Resposta às pendências.doc	17/08/2015 09:11:56		Aceito
Outros	Instrumento de Coleta de Dados - Validado Port.docx	27/07/2015 17:51:46		Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE nova CEP.doc	27/07/2015 17:50:39		Aceito
Declaração de Pesquisadores	Autorização da Instituição.jpg	10/03/2015 12:03:47		Aceito
Declaração de Pesquisadores	Termo Compr Equip Exec.jpg	25/02/2015 10:31:42		Aceito
Declaração de Pesquisadores	Link dos Currículo Lattes.doc	25/02/2015 10:31:15		Aceito
Declaração de Pesquisadores	Instrumento Coleta de Dados - 02.JPG	25/02/2015 10:31:04		Aceito
Declaração de	Instrumento Coleta de Dados - 01.	25/02/2015		Aceito

Endereço: Rua Ovídio Pires de Campos, 225 5º andar

Bairro: Cerqueira Cesar

CEP: 05.403-010

UF: SP

Município: SAO PAULO

Telefone: (11)2661-7585

Fax: (11)2661-7585

E-mail: cappesq.adm@hc.fm.usp.br

USP - HOSPITAL DAS
CLÍNICAS DA FACULDADE DE
MEDICINA DA UNIVERSIDADE

Continuação do Parecer: 1.894.307

Pesquisadores	JPG	10:30:53		Aceito
Outros	Declaração Instituição.jpg	25/02/2015 10:30:36		Aceito
Folha de Rosto	Folha de Rosto.JPG	25/02/2015 10:28:20		Aceito

Situação do Parecer:

Aprovado

Necessita apreciação da CONEP:

Não

SAO PAULO, 19 de Janeiro de 2017

Assinado por:
Joel Faintuch
(Coordenador)

Endereço: Rua Ovídio Pires de Campos, 225 5º andar

Bairro: Cerqueira Cesar

CEP: 05.403-010

UF: SP

Município: SAO PAULO

Telefone: (11)2661-7585

Fax: (11)2661-7585

E-mail: cappesq.adm@hc.fm.usp.br

Termo de Consentimento Livre e Esclarecido

O senhor/ senhora está sendo convidado como voluntário (a) a participar do projeto de pesquisa: Instrução de médicos residentes em técnicas de ensino.

O motivo que nos leva a estudar o problema se fundamenta na escassez de material e recursos alocados à capacitação de médicos residentes brasileiros em técnicas de ensino. A dinâmica do trabalho diário do médico residente se fundamenta no ensino em diversos momentos. Neste contexto gostaríamos de avaliar o conhecimento prévio dos residentes e capacitar os mesmos para que possam desempenhar suas funções de ensino de forma mais coesa.

Os procedimentos de coleta de dados se darão da seguinte forma: inicialmente todos os residentes serão submetidos a uma avaliação do conhecimento prévio que cada um possui no que se concerne às técnicas de ensino. A seguir serão divididos em dois grupos que passarão por dois tipos diferentes de intervenção. Após esta primeira etapa serão novamente avaliados por meio de instrumento específico semelhante ao aplicado no início da avaliação. Após 6 meses os dois grupos serão submetidos às mesmas intervenções, porém de forma cruzada, ou seja, cada grupo será submetido à avaliação que o grupo oposto recebeu na primeira etapa. Por fim, serão todos submetidos novamente a avaliação por meio de instrumento específico.

Durante as avaliações não acreditamos que haverá qualquer tipo de desconforto ou possibilidade de constrangimento dos residentes em estudo. A necessidade de gravação em vídeo das avaliações poderá ser requisitada conforme necessidades específicas do projeto e as imagens não serão exibidas a pessoas que não estejam envolvidas no processo de pontuação dos desempenhos observados.

Os benefícios previstos a todos os participantes deste projeto serão avaliados com base no aprendizado adquirido nas intervenções propostas. Como todos receberão as intervenções nas diferentes etapas não haverá prejuízo a nenhum dos participantes.

Os participantes serão esclarecidos sobre a pesquisa em qualquer aspecto que desejarem. São todos livres para se recusarem a participar, retirar seu consentimento ou interromper a participação a qualquer momento. A participação é voluntária e a recusa em participar não irá acarretar qualquer penalidade ou perda de benefícios.

Os pesquisadores irão tratar a identidade dos participantes com padrões profissionais de sigilo. Os resultados das avaliações poderão ser enviados aos residentes após a conclusão da coleta de dados e permanecerão confidenciais. O nome dos residentes e o material que indique sua participação não será liberado sem permissão. Não haverá qualquer tipo de identificação nominal em publicações que possam resultar deste estudo. Uma cópia deste consentimento será arquivada no departamento de Clínica Médica da Universidade Federal de Uberlândia e outra será fornecida a cada participante.

Declaro que concordo em participar desse estudo. Recebi uma cópia deste termo de consentimento livre e esclarecido e me foi dada a oportunidade de ler e esclarecer as minhas dúvidas.

Nome..... Assinatura do participante.....

Assinatura do pesquisador responsável.....

Data.....

OSTE- Estações de Avaliação

Estação 1: Ambulatório Didático

- Discussão de um caso com alunos de internato
- Alunos desinteressados e dispersos
- Poucas perguntas, apatia facilmente perceptível
- Ambulatório de geriatria - múltiplas queixas, história complexa, polifarmácia (queixa do aluno)
- Aluno levanta hipótese de Doença de Alzheimer e coloca conceitos errados (Feedback negativo)
- Avaliação da capacidade do residente de atrair a atenção do aluno, sintetizar o caso de forma mais atraente e corrigir conceitos errados de forma educada e respeitosamente
- O residente encorajou outros alunos a participarem da discussão?
- O residente estimulou a formulação de perguntas relevantes ao caso?

Estação 2- Enfermaria de Clínica Médica

- Orientação de um aluno do sexto ano na prescrição de um paciente internado por IC descompensada
- Avaliara capacidade do residente de citar princípios gerais da terapêutica com finalidade didática ('teach a general principle')
- Pergunta difícil - o residente sabe responder 'eu não sei'?
- O residente estimula o aluno a buscar leitura adicional em casa?
- Avaliara capacidade do residente em deixar claro o conteúdo que o aluno precisa dominar - objetivo do encontro

Estação 3- Unidade de Emergência

- Orientação de um aluno da graduação iniciando o curso de semiologia
- Muitas perguntas, dúvidas e grande interesse na prática clínica inicial
- Avaliara capacidade do residente de escutar o aluno
- Discussão de um caso de diarreia aguda autolimitada
- O aluno descreve corretamente os principais agentes etiológicos envolvidos na doença (Feedback positivo)
- Residente discute a terapêutica com profundidade adequada ao conhecimento do aluno?
- Aluno pergunta se pode procurar o residente em outras ocasiões para acompanhamento

Estação 4: Orientando um procedimento

- Orientação de um aluno da graduação na preparação para um procedimento
- Ator como observador, postura mais passiva
- Orientação de uma punção venosa periférica com um modelo de membro superior e materiais necessários
- Avaliação de comunicação de objetivos, prioridade para os objetivos e aceitação dos objetivos pelo aluno
- Capacidade do residente evitar digressões e retomar os objetivos conforme a necessidade

Estação 1: Ambulatório Didático

Orientações aos atores:

Esta estação visa avaliar como o residente se comporta diante de uma discussão de caso com alunos do internato. O cenário consiste de um ambulatório de geriatria onde um residente orienta dois internos após o atendimento.

O foco principal é avaliar o comportamento do residente frente a alunos desmotivados. Teremos 2 atores representando o papel dos estudantes. Um deles terá a tarefa de passar o caso e o outro acompanhará a discussão.

Inicialmente o residente será encaminhado a uma cadeira para se sentar e os alunos atores deverão permanecer em pé a menos que o residente peça que se sentem também.

O ator que vai passar o caso deve inicialmente se queixar da complexidade do caso. Sugere-se dizer que o paciente tem 91 anos e diagnóstico de Doença de Alzheimer há 8 anos e toma um “monte” de remédios. Caso o residente pergunte algo mais sobre a história o ator deve dizer que não sabe, pois, o paciente não coopera e o acompanhante não é o cuidador principal. Caso pergunte a respeito das medicações diga que não anotou e não se lembra dos nomes que são estranhos. Se o residente pedir que você volte a conversar com o paciente para melhorar a história diga que o mesmo foi ao laboratório colher um exame. Se ele(a) lhe perguntar o que você acha do caso ou se ele ficar perdido sem saber o que fazer comece a dar suas impressões sobre a Doença de Alzheimer. Coloque um conceito claramente errado. Questione o diagnóstico dizendo que a maioria dos pacientes não sobrevive mais de 3 anos após o início dos sintomas e aguarde o residente se manifestar.

O outro aluno deve se concentrar durante todo o tempo em seu celular. Se perguntado a respeito do caso deve dizer que não tem nenhuma informação adicional. Caso seja encorajado a participar da discussão guarde o celular. Se o residente pedir que você guarde o celular em qualquer momento, faça-o. Qualquer outra pergunta deve ser respondida com: “ eu não sei”. Se insistir em lhe fazer perguntas responda com: “ ele que atendeu”, apontando para o outro ator.

Duração da Sessão: 8 minutos

Competências Avaliadas:

1. Ambiente de ensino:
 - ✓ Estimulou o interesse dos alunos no tópico
 - ✓ Escutou os alunos
 - ✓ Estimulou os alunos a participarem ativamente da discussão
 - ✓ Demonstrou respeito pelos alunos
 - ✓ Evitou ridiculizar ou intimidar os alunos
2. Promoção da compreensão e retenção:
 - ✓ Respondeu às questões dos alunos claramente
3. Promoção de aprendizado auto direcionado:
 - ✓ Encorajou explicitamente estudo adicional
 - ✓ Encorajou os alunos a fazerem leitura fora da instituição
4. FEEDBACK:

- ✓ Forneceu feedback negativo (corretivo) aos alunos
- ✓ Explicou aos alunos o porquê de estarem certos/errados

Estação 2: Enfermaria de Clínica Médica

Orientações aos atores:

Esta estação visa avaliar como o residente se comporta diante de uma discussão de caso com um aluno do sexto ano em uma enfermaria de clínica médica.

O ator espera o residente com uma prescrição já pronta de um paciente internado por IC descompensada há 3 dias. O ator que está no papel de um aluno do sexto ano deve abordar o médico residente com a prescrição a ser avaliada. Sugere-se iniciar a estação com: “ Vamos ver a prescrição do paciente? ”. Após esta pergunta dê um tempo para que ele se manifeste. Tempo sugerido para que o residente avalie a prescrição e se manifeste: 2 minutos. Neste intervalo qualquer pergunta direcionada a você deve ser respondida com: “vamos continuar discutindo a prescrição...”. Após este momento você deverá fazer uma pergunta difícil: “ A amiodarona compete com a digoxina na excreção renal? Se a resposta for “ não sei” deixe de lado essa etapa. Se o residente responder “sim” ou “não” pergunte onde ele obteve essa informação: “Onde você leu isso? ”

Após esta etapa o ator deve perguntar se é isso o que ele precisa saber para a prescrição desse paciente: ‘É isso o que eu tenho que saber sobre tratamento de IC?’ Se ele responder que é somente isso finalize a tarefa. Caso ele explique com pormenores o tratamento e incentive leitura adicional aguarde até o tempo limite da sessão.

Duração da Sessão: 8 minutos

Competências Avaliadas:

1. Ambiente de ensino:
 - ✓ Estimulou o interesse dos alunos no tópico
 - ✓ Encorajou os alunos a participarem ativamente das discussões
 - ✓ Demonstrou respeito pelos alunos
 - ✓ Esteve pronto a responder “ eu não sei”
2. Promoção da compreensão e retenção:
 - ✓ Enfatizou o que queria que os alunos se lembrassem
3. Promoção de aprendizado auto direcionado:
 - ✓ Encorajou explicitamente estudo adicional
 - ✓ Encorajou os alunos a fazerem leitura fora da instituição
 - ✓ Motivou os alunos a aprenderem sozinhos
4. Comunicação dos objetivos:
 - ✓ Apresentou o nível de competência esperado
5. Feedback:
 - ✓ Ofereceu aos alunos sugestões para melhorias

Estação 3: Unidade de Emergência

Orientações aos atores:

Esta estação visa avaliar como o residente se comporta diante de um aluno durante discussão de um caso em uma unidade de emergência. O cenário consiste de da sala de reunião dos médicos preceptores e residentes.

O foco principal é avaliar a capacidade de o residente discutir de forma objetiva um caso simples de diarreia aguda e conter a empolgação de alunos no início do curso.

O ator representando um aluno do terceiro ano aborda o residente com um caso “muito legal” de diarreia aguda. O residente é informado no início da sessão que o plantão está muito cheio e com pacientes graves.

O aluno passa o caso de um jovem de 17 anos com um quadro de diarreia aguda que coincidiu com uma festa onde comeu feijoada. Outras pessoas que comeram a feijoada têm sintomas semelhantes. Se o residente perguntar por qualquer sinal de instabilidade (perfusão ruim, hipotensão, taquicardia, taquipneia ou alteração da consciência) você deve dizer que o exame físico está normal, exceto por discreta dor à palpação do abdome.

Você deve descrever sua principal hipótese diagnóstica como uma intoxicação alimentar: “É provável que o alimento estava contaminado com toxinas bacterianas que provocaram a diarreia”. Espere pelo feedback do residente após esta afirmação.

Após esta etapa tente mudar completamente de assunto comentando sobre o caso de um amigo que há 3 anos teve uma diarreia durante a Páscoa devido a excesso de ingestão de chocolates. Prolongue o assunto descrevendo características desse amigo (“ muito doido “ , “ contador de piadas”, “ a muito tempo que não o vê “ , “ está com saudades”, “ iria procura-lo no FACEBOOK”). Não pare de falar desse amigo a menos que o residente o interrompa ou que haja uma interrupção externa. Outro ator fará uma entrada com interrupção externa. Ele entrará na sala e perguntará sobre a data de uma prova. A conversa deverá se prolongar até que o residente os interrompa ou que se esgote o tempo da sessão.

Duração da Sessão: 8 minutos

Competências Avaliadas:

1. Ambiente de ensino:
 - ✓ Escutou os alunos
 - ✓ Demonstrou respeito pelos alunos
2. Promoção da compreensão e retenção:
 - ✓ Avaliou o nível de conhecimento dos alunos antes da sessão
3. Controle da sessão:
 - ✓ Fez uso eficiente do tempo para ensino
 - ✓ Atraiu atenção para o tempo
 - ✓ Evitou digressões
 - ✓ Desencorajou interrupções externas
4. Feedback:
 - ✓ Forneceu feedback positivo aos alunos
 - ✓ Explicou aos alunos o porquê de estarem corretos/incorretos
 - ✓ Ofereceu aos alunos sugestões para melhorias

Estação 4: Orientação de Procedimento

Orientações aos atores:

Esta estação visa avaliar como o residente se comporta durante a orientação de um procedimento. O cenário consiste de um laboratório de habilidades onde o residente se presta a demonstrar como é feita uma punção venosa superficial em um modelo de membro superior.

O residente entra na sala e encontra o ator sentado aguardando a orientação. Neste momento o ator aguarda que o residente se manifeste com um cumprimento ou pelo menos sua identificação. Caso isso não ocorra o ator deve iniciar a sessão chamando o residente para a tarefa.

Nesta tarefa o ator terá participação mais como um expectador, ouvindo a explicação do médico residente. Será dada maior ênfase à capacidade de os residentes exporem os objetivos do encontro, situar a relevância do tema e dar prioridade aos objetivos.

Uma forma de avaliar a capacidade de priorizar os objetivos e evitar digressões ao mesmo tempo seria perguntar sobre um tema que foge ao objetivo principal. Assim que o residente terminar sua explicação ou quando surgir alguma brecha ele poderia ser interrompido com a seguinte pergunta: “Um procedimento que me dá muito medo é a intubação, o que você me diz a respeito disso”? Deixe que ele tome o controle da sessão em direção ao tema inicial ou que fale sobre a intubação.

Duração da Sessão: 6 minutos

Competências Avaliadas:

1. Ambiente de ensino:
 - ✓ Escutou os alunos
 - ✓ Demonstrou respeito pelos alunos
2. Promoção da compreensão e retenção:
 - ✓ Avaliou o nível de conhecimento dos alunos antes da sessão
3. Controle da sessão:
 - ✓ Cobriu todos os tópicos previstos
 - ✓ Evitou digressões
4. Comunicação dos objetivos:
 - ✓ Apresentou os objetivos de forma clara e concisa
 - ✓ Apresentou a relevância dos objetivos aos alunos
 - ✓ Verificou a aceitação dos objetivos pelos alunos
 - ✓ Reestabeleceu os objetivos conforme a necessidade

Material de apoio aos atores:

Estação 1:

Descrição do caso: Paciente 91 anos, sexo feminino, com diagnóstico de Doença de Alzheimer há cerca de 8 anos com múltiplas doenças associadas e em uso de um grande número de medicações. Se o residente insistir em saber quais doenças o paciente tem diga que se lembra de terem sido citadas hipertensão, diabetes e constipação intestinal. Quanto às medicações cite insulina e aspirina. Caso pergunte a respeito das outras doenças e medicações diga que ficou perdido na história e que não conseguiu anotar ou memorizar as medicações e doenças. A dificuldade do caso se encontra na complexidade do paciente e no despreparo do aluno em coletar os dados de forma adequada e apresenta-los da melhor forma possível. Somada a essa dificuldade observamos os alunos muito desinteressados na discussão e com problemas para se concentrar. Um deles inclusive se concentra todo o tempo no celular. A partir da segunda semana de atividades vamos observar que alguns residentes irão tentar obter mais informações a respeito da paciente e observar as impressões do aluno sobre o caso. Manteremos o mesmo roteiro de respostas. O objetivo é que no final o residente dê feedback a suas impressões e corrija de forma positiva aos conceitos errados que foram colocados. Caso o residente lhe pergunte qual sua opinião a respeito do caso sugere-se que você diga: “Então... (suspiro de preguiça) é uma paciente de 91 anos com diagnóstico de Doença de Alzheimer. Muito difícil colher uma boa história e colocar anos de evolução da doença em uma consulta.”

Material de apoio aos atores:

Estação 2:

Nesta estação vamos discutir o caso de um paciente de 65 anos com diagnóstico de constipação intestinal crônica de etiologia chagásica que está internado na enfermaria de cardiologia com diverticulite aguda. A causa da diverticulite não é conhecida. O paciente não tem outras doenças associadas e não tomava medicações. Está internado há 2 dias com melhora progressiva. A história inicial é de 5 dias de dor abdominal contínua de moderada a forte intensidade em fossa ilíaca esquerda associada a picos febris não aferidos. O diagnóstico foi realizado com base nos sintomas e na tomografia de abdome. Essas informações são relevantes porque a partir da segunda semana de atividades alguns residentes farão menos perguntas e solicitarão a opinião do aluno a respeito do caso. A impressão que se deve passar é que as medicações estão corretamente prescritas e que a melhora dos sintomas denota boas condutas. O paciente permanece afebril nas últimas 24 horas, a dor está melhor e o paciente tem tolerado bem a dieta. Se perguntado a respeito de critérios de alta diga que vamos esperar 48 horas com o paciente afebril para dar alta com antibióticos por via oral. Manteremos as perguntas difíceis com o objetivo de obter sinceridade do residente quanto ao desconhecimento do assunto. Iremos mudar a pergunta a cada semana porque alguns residentes podem procurar as respostas em casa. Você terá a cópia de uma prescrição genérica de um paciente internado com as mesmas características para apresentar durante a sessão. O objetivo final é que o residente escute a impressão do aluno, faça comentários pertinentes ao caso e ao final dê feedback apropriado.

Material de apoio aos atores:

Estação 3:

Nesta estação vamos testar a capacidade do residente de orientar um aluno novato no ambiente de emergência ao mesmo tempo que tem que se atentar ao plantão que está muito tumultuado com pacientes graves e com maior prioridade de atendimento. Você irá abordar o residente para discussão de um caso de uma paciente do sexo feminino de 17 anos, com um quadro de diarreia aguda após ter comido uma feijoada. A diarreia é líquida com cerca de 6 evacuações ao dia e está associada a dor abdominal difusa, tipo colica, de leve intensidade. A paciente nega febre, presença de sangue ou muco nas fezes e também nega doenças associadas. Faz uso apenas de anticoncepcional. Essas informações são úteis caso o residente comece a fazer perguntas a respeito do caso. Se ele perguntar especificamente a sua impressão sobre a paciente você deve dizer que é um caso de uma paciente jovem com diarreia aguda por provável intoxicação alimentar já que há casos semelhantes em familiares que também comeram a feijoada e que a resolução deve ser espontânea com o tempo. Adicione que o tratamento deve ser baseado em hidratação e sintomáticos. A partir desse momento você deve começar a comentar sobre o caso semelhante que observou em um amigo até que seja interrompido pelo residente ou até que ocorra uma interrupção externa. Faremos a interrupção na forma de uma mensagem por whatsapp de um colega de turma que precisa saber a data de uma prova. Você então pede um tempo para responder ao colega.

O objetivo final da sessão é que o residente controle o tempo de forma que o caso seja passado de forma adequada e sem interrupções e que descreva conceitos construtivos a respeito da doença com feedback na finalização.

Material de apoio aos atores:

Cenário 4:

Nesta sessão teremos como objetivo verificar como o residente se comporta na orientação de um procedimento simples, mas que nunca foi realizado por um aluno. A sessão terá início com o ator sentado aguardando o residente se identificar e chama-lo para orientação. Espera-se que o residente solicite que o ator se manifeste a respeito do procedimento, perguntando sobre as indicações, possíveis complicações e técnica de punção. Caso o residente não pergunte nada deixe que ele discorra a respeito do procedimento. Se ele perguntar o que você sabe sobre a punção venosa superficial diga que é um procedimento simples para cateterização de veia periférica em casos de coleta de sangue ou administração de medicações. Diga que o procedimento tem poucas complicações como hematoma local e que a maioria das complicações tem abordagem expectante. Após essa abordagem tente mudar o foco da sessão mesmo que o residente esteja demonstrando o procedimento. Diga que o procedimento que mais lhe assusta é a intubação orotraqueal e que você gostaria de treinar mais antes de se formar. Tente fazer com que o residente se esqueça da punção venosa e discorra sobre a intubação. Caso ele o interrompa, escute o que tem a dizer e aguarde o final da sessão.

AVALIAÇÃO DE DESEMPENHO DE ENSINO DE MÉDICOS RESIDENTES
Stanford Faculty Development Program Questionnaire (SFDPQ)

PROJETO DE PESQUISA APROVADO PELO COMITÊ DE ÉTICA EM PESQUISA DA FACULDADE DE MEDICINA DA
 UNIVERSIDADE FEDERAL DE UBERLÂNDIA, RESPONSÁVEL Saadallah Fakhouri.

Nome do Avaliador:					
	Discordo totalmente	Discordo	Indiferente (não é possível avaliar)	Concordo (com alguma ressalva)	Concordo totalmente
1. Ambiente de ensino					
Escutou os alunos com atenção (não consultou celular, acompanhou cenas ao redor, etc)					
Estimulou (de algum modo) o interesse dos alunos no tópico					
Estimulou (de algum modo) os alunos a participarem ativamente da discussão					
Demonstrou respeito pelos alunos.					
Evitou ridiculizar ou intimidar os alunos					
2. Promoção da compreensão e retenção					
Avaliou o nível de conhecimento prévio dos alunos					
Respondeu às questões dos alunos claramente					
Respondeu de modo firme e educado às perguntas dos alunos					
3. Promoção de aprendizado auto direcionado					
Encorajou explicitamente estudo adicional (a buscarem e aprendizagem para além da ofertada na disciplina)					
Encorajou de modo educado, os alunos a fazerem leitura fora da instituição					
Encorajou de modo firme, os alunos a fazerem leitura fora da instituição					

AVALIAÇÃO DE DESEMPENHO DE ENSINO DE MÉDICOS RESIDENTES
Stanford Faculty Development Program Questionnaire (SFDPQ)

PROJETO DE PESQUISA APROVADO PELO COMITÊ DE ÉTICA EM PESQUISA DA FACULDADE DE MEDICINA DA
 UNIVERSIDADE FEDERAL DE UBERLÂNDIA, RESPONSÁVEL Saadallah Fakhouri.

4. Feedback:					
Forneceu feedback negativo (corretivo) aos alunos					
Explicou aos alunos o porquê de estarem certos/errados					
Forneceu feedback apropriado, considerando que:					
1. Planejou o que dizer					
2. Escolheu ambiente apropriado (<i>neste caso, apenas com a presença do avaliador</i>)					
3. Foi específico com as palavras (<i>utilizou expressões como na situação tal...</i>)					
4. Comentou apenas o que presenciou					
5. Comentou situações positivas					
6. Comentou situações negativas					
7. Ofereceu plano de melhoria dos aspectos negativos (<i>orientou como poderia fazer</i>)					
8. Certificou-se que o estudante entendeu					

Em uma escala de 0 (professor limitado) a 100 (professor ideal), que nota você atribui ao residente? ____

AVALIAÇÃO DE DESEMPENHO DE ENSINO DE MÉDICOS RESIDENTES
Stanford Faculty Development Program Questionnaire (SFDPPQ)

PROJETO DE PESQUISA APROVADO PELO COMITÊ DE ÉTICA EM PESQUISA DA FACULDADE DE MEDICINA DA
UNIVERSIDADE FEDERAL DE UBERLÂNDIA, RESPONSÁVEL: SAADALLAH FAKHOURI

Nome do Avaliador:					
	Discordo totalmente	Discordo	Indiferente (não é possível avaliar)	Concordo (com alguma ressalva)	Concordo totalmente
1. Ambiente de ensino					
Escutou os alunos com atenção (não consultou celular, acompanhou cenas ao redor, etc)					
Estimulou (de algum modo) o interesse dos alunos no tópico					
Estimulou (de algum modo) os alunos a participarem ativamente da discussão					
Demonstrou respeito pelos alunos					
Esteve pronto a responder “eu não sei” diante de perguntas difíceis					
2. Promoção da compreensão e retenção					
Enfatizou o que queria que os alunos se lembrassem (em discussões futuras ou em casos semelhantes)					
3. Promoção de aprendizado auto direcionado					
Encorajou explicitamente estudo adicional (a buscarem aprendizagem para além da ofertada na disciplina)					
Encorajou de modo educado, os alunos a fazerem leitura fora da instituição					
Motivou os alunos a estudarem sozinhos (buscando conhecimento ou procurando leitura adicional)					

AValiação de desempenho de ensino de médicos residentes
Stanford FacultyDevelopmentProgramQuestionnaire (SFD PQ)

PROJETO DE PESQUISA APROVADO PELO COMITÊ DE ÉTICA EM PESQUISA DA FACULDADE DE MEDICINA DA
 UNIVERSIDADE FEDERAL DE UBERLÂNDIA, RESPONSÁVEL: SAADALLAH FAKHOURI

4. Comunicação dos Objetivos:					
Apresentou o nível de competência esperado (explicitou o que o aluno deveria saber a respeito do tema em discussão)					
5. Feedback					
Ofereceu aos alunos sugestões para melhorias (estimulou procura de casos semelhantes, referências relevantes, indicou especialista para consulta, etc.)					
Fornecer feedback apropriado, considerando que:					
1. Planejou o que dizer					
2. Escolheu ambiente apropriado (<i>neste caso, apenas com a presença do avaliador</i>)					
3. Foi específico com as palavras (<i>utilizou expressões como na situação tal...</i>)					
4. Comentou apenas o que presenciou					
5. Comentou situações positivas					
6. Comentou situações negativas					
7. Ofereceu plano de melhoria dos aspectos negativos (<i>orientou como poderia fazer</i>)					
8. Certificou-se que o estudante entendeu					

Em uma escala de 0 (professor limitado) a 100 (professor ideal), que nota você atribui ao residente? ____

AValiação de desempenho de ensino de médicos residentes
Stanford Faculty Development Program Questionnaire (SFDPPQ)

PROJETO DE PESQUISA APROVADO PELO COMITÊ DE ÉTICA EM PESQUISA DA FACULDADE DE MEDICINA DA
UNIVERSIDADE FEDERAL DE UBERLÂNDIA, RESPONSÁVEL Saadallah Fakhouri.

Nome do Avaliador:					
	Discordo totalmente	Discordo	Indiferente (não é possível avaliar)	Concordo (com alguma ressalva)	Concordo totalmente
1.Ambiente de ensino					
Escutou os alunos com atenção (não consultou celular, acompanhou cenas ao redor, etc.)					
Demonstrou respeito pelos alunos.					
2. Promoção da compreensão e retenção					
Avaliou o nível de conhecimento prévio dos alunos					
3.Controle da Sessão					
Fez uso eficiente do tempo para o ensino (tentou utilizar todo o tempo que dispunha na tarefa de transmitir conhecimento)					
Atraiu atenção para o tempo (deixou claro o tempo que dispunham para exercer a atividade proposta)					
Evitou digressões (de forma respeitosa evitou que outros assuntos atrapalhassem a sessão)					
Desencorajou interrupções externas					

AValiação de desempenho de ensino de médicos residentes
Stanford Faculty Development Program Questionnaire (SFDPQ)

PROJETO DE PESQUISA APROVADO PELO COMITÊ DE ÉTICA EM PESQUISA DA FACULDADE DE MEDICINA DA
 UNIVERSIDADE FEDERAL DE UBERLÂNDIA, RESPONSÁVEL Saadallah Fakhouri.

4.Feedback:					
Fornecer feedback negativo (corretivo) aos alunos					
Explicar aos alunos o porquê de estarem certos/errados					
Oferecer ao aluno sugestões para melhorias					
Fornecer feedback apropriado, considerando que:					
1. Planejar o que dizer					
2. Escolher ambiente apropriado <i>(neste caso, apenas com a presença do avaliador)</i>					
3. Ser específico com as palavras <i>(utilizar expressões como na situação tal...)</i>					
4. Comentar apenas o que presenciou					
5. Comentar situações positivas					
6. Comentar situações negativas					
7. Oferecer plano de melhoria dos aspectos negativos <i>(orientar como poderia fazer)</i>					
8. Certificar-se que o estudante entendeu					

Em uma escala de 0 (professor limitado) a 100 (professor ideal), que nota você atribui ao residente? ____

AValiação de desempenho de ensino de médicos residentes
Stanford Faculty Development Program Questionnaire (SFD PQ)

PROJETO DE PESQUISA APROVADO PELO COMITÊ DE ÉTICA EM PESQUISA DA FACULDADE DE MEDICINA DA
 UNIVERSIDADE FEDERAL DE UBERLÂNDIA, RESPONSÁVEL Saadallah Fakhouri.

Nome do Avaliador:	Discordo totalmente	Discordo	Indiferente (não é possível avaliar)	Concordo (com alguma ressalva)	Concordo totalmente
1. Ambiente de ensino					
Escutou os alunos com atenção (não consultou celular, acompanhou cenas ao redor, etc.)					
Demonstrou respeito pelos alunos.					
Evitou ridicularizar ou intimidar os alunos					
2. Promoção da compreensão e retenção					
Avaliou o nível de conhecimento prévio dos alunos					
3. Controle da Sessão					
Cobriu todos os tópicos previstos (objetivos explicitados no início da sessão)					
Evitou digressões (de forma respeitosa evitou que outros assuntos atrapalhassem a sessão)					
4. Comunicação dos Objetivos					
Apresentou os objetivos de forma clara e concisa (antes do início da sessão apresentou os tópicos da discussão)					
Apresentou a relevância dos objetivos aos alunos (deixou claro desde o início o porquê da importância do assunto)					
Verificou a aceitação dos objetivos pelos alunos (tentou obter de alguma forma que os alunos concordassem com os objetivos propostos)					
Reestabeleceu os objetivos conforme a necessidade (retornou aos objetivos iniciais em caso de dúvida do aluno ou imprevisto no decorrer da sessão)					

Em uma escala de 0 (professor limitado) a 100 (professor ideal), que nota você atribui
ao residente? ____