

# OS DIFERENTES OLHARES NA AVALIAÇÃO DE ALUNOS EM ESTÁGIO CLÍNICO SUPERVISIONADO

ROSÂNGELA CURVO LEITE DOMINGUES<sup>1\*</sup>, ELIANA AMARAL<sup>2</sup>, ANGÉLICA MARIA BICUDO ZEFERINO<sup>2</sup>

Trabalho realizado na Universidade Estadual de Campinas - Faculdade de Ciências Médicas

## RESUMO

**OBJETIVO.** Comparar notas do Conceito Global Itemizado (CGI) e da Avaliação Estruturada (AE), atribuídas por três especialidades - Ginecologia e Obstetrícia (GO), Clínica Médica (CM) e Pediatria (Ped) - a alunos em final de estágio clínico nas unidades básicas de saúde.

**MÉTODOS.** Cento e seis alunos do 4º ano médico foram avaliados por um docente de cada especialidade, utilizando CGI contendo seis domínios de habilidades técnicas (CGIt) e sete domínios humanísticos (CGIh), em escala 0-10, e pela AE resultante de observação de atendimento. Para análise estatística, utilizaram-se coeficientes de Cronbach, testes de Friedman e pareados de Wilcoxon, coeficientes de correlação de Spearman e Pearson e distâncias Euclidianas ( $p < 0,05$ ).

**RESULTADOS.** Houve alta consistência interna do CGI nas especialidades ( $>0,92$ ). As distribuições das notas concentraram-se no extremo superior da escala. As notas de Ped foram significativamente inferiores às de GO e CM (diferenças de medianas entre 0,50 e 0,67), com correlações baixas entre elas ( $-0,02 < R < 0,48$ ). Os domínios que tiveram maior impacto na confiabilidade do CGI foram: julgamento clínico (GO e Ped), solução de problemas (CM) e capacidade de auto-reflexão (Ped). As notas de GO e Ped mostraram a menor concordância; as de GIRT de Ped foram as mais discordantes das outras notas.

**CONCLUSÃO.** As especialidades têm visões distintas sobre como avaliar alunos, mesmo utilizando instrumentos semelhantes, o que pode ser reflexo de "cultura" de cada uma delas. Permanece o desafio de minimizar tais diferenças com atividades de desenvolvimento docente.

UNITERMOS: Avaliação educacional. Especialidades médicas. Estágio clínico.

## \*Correspondência:

Rua Luciano Venere  
Decourt, nº 858  
Cidade Universitária  
Campinas - SP  
CEP: 13083-740  
Telefone: (19) 9207-6404  
Fax: (19) 3289-2182  
rosangela@cpdec.com.br

## INTRODUÇÃO

Desenvolver competências técnicas e humanísticas para o exercício da profissão tem se tornado um aspecto central da educação médica, para a qual o estágio clínico supervisionado assume um papel relevante<sup>1</sup>. A necessidade de preparar profissionais críticos e reflexivos, capazes de atuar em diversos cenários da prática, em equipes multidisciplinares e aptos a responderem às demandas da sociedade pressupõe grandes transformações educacionais. Dentre elas, destaca-se a capacitação dos docentes/supervisores para garantir maior qualidade no estabelecimento de critérios de desempenho, nas observações da atuação dos alunos, na prática de *feedback* e no processo de avaliação das competências esperadas<sup>2, 3</sup>.

Segundo a Associação Européia de Educação Médica (AMEE), o termo supervisão é definido como o fornecimento de orientações e *feedback* relacionados ao desenvolvimento profissional e educacional do aluno, para ajudá-lo a prover um atendimento médico adequado<sup>4</sup>. A avaliação é um poderoso componente da supervisão eficaz, pois direciona a aprendizagem, serve como guia para os estudantes em relação às competências que devem

ser aprimoradas e auxilia os supervisores a identificarem oportunidades de melhoria da qualidade do ensino<sup>5, 6</sup>.

Para que a avaliação do aluno em estágio clínico tenha um forte impacto educacional, é importante que haja entendimento dos objetivos a serem atingidos, consenso sobre conteúdo a ser avaliado, conhecimento do nível de competência dos alunos e *feedback* construtivo e eficaz, para que o estudante possa refletir sobre suas ações e traçar um plano de desenvolvimento profissional contínuo<sup>4</sup>.

No entanto, a definição de quais domínios da competência os alunos devem atingir nas várias etapas de sua formação pode não ser clara ou ainda não ser igualmente compreendida por supervisores de diferentes estágios clínicos. Essas diferenças tendem a se acentuar quando se contrasta a supervisão oferecida por especialistas, que podem ter distintas percepções de seu papel docente e da importância relativa de sua especialidade na formação do médico<sup>4,7,8</sup>.

O objetivo do presente estudo foi comparar as avaliações de docentes/supervisores de diferentes especialidades, realizadas com alunos em estágio clínico, durante um módulo de ensino integrado.

1. Doutora em Tocoginecologia – Área de ciências Biomédicas. Universidade Estadual de Campinas – UNICAMP, Campinas, SP

2. Professor associado da Universidade Estadual de Campinas. UNICAMP, Campinas, SP

## MÉTODOS

Trata-se de um estudo transversal por análise secundária de dados gerados a partir do processo avaliativo de alunos iniciando prática clínica em atenção primária à saúde. No módulo denominado *Atenção Integral à Saúde*, os alunos do 4º ano devem adquirir experiência clínica fora do hospital universitário, sob supervisão de docentes e tutores da própria unidade de saúde. A experiência leva ao desenvolvimento de habilidades procedimentais e raciocínio clínico, além de contribuir para criação de vínculo com o paciente e com outros profissionais da equipe de saúde.

Os estudantes são divididos em seis grupos, alocados em Unidades Básicas do Sistema Único de Saúde (SUS), onde realizam atendimentos de Ginecologia e Obstetrícia (GO), Clínica Médica (CM) e Pediatria (Ped). O estágio acontece durante nove meses de um ano letivo, três tardes por semana (12 horas), em semanas alternadas. A cada dia da semana os alunos atendem pacientes de uma das especialidades clínicas. Realizam anamnese e exame físico, indicam condutas diagnósticas e terapêuticas para as situações prevalentes, registram o atendimento, discutem o caso clínico com docentes, tutores e membros da equipe de saúde e participam de reuniões de projeto terapêutico. Os procedimentos do estágio são os mesmos nas diferentes UBS. Além das atividades práticas, há seis horas semanais de seminários sobre os temas clínicos compatíveis com as experiências esperadas neste nível de atenção à saúde, integrando conceitos de saúde de adulto, criança e saúde da mulher.

Este estudo refere-se às avaliações realizadas ao final do segundo semestre de 2005, segundo ano de oferecimento desse módulo de ensino no novo currículo de Medicina. Cada estudante foi avaliado por três docentes/supervisores, de forma independente, um de cada especialidade, por meio do Conceito Global Itemizado (CGI) e da Avaliação Estruturada (AE). O CGI resumiu a impressão que o avaliador formou do aluno ao longo do semestre em relação a 13 domínios da competência, sendo seis referentes à competência técnica (denominado a partir de agora como CGI<sub>t</sub>) - qualidade da história, exame clínico, conhecimento médico, julgamento clínico, solução de problemas, hábitos de trabalho - e sete à competência humanística (denominado a partir de agora como CGI<sub>h</sub>) - comunicação e relacionamento com pacientes e familiares, respeito, capacidade de autorreflexão, percepção do contexto, interação com colegas, interação com docentes, interação com demais profissionais. As notas atribuídas a cada domínio variaram de 0 a 10, com descritores ancorando o menor e o maior valor.

A Avaliação Estruturada (AE) teve como base a observação docente de três consultas completas, uma em cada das especialidades (GO, Ped, CM). O desempenho de cada aluno foi registrado em uma lista de verificação detalhada, que continha questões relacionadas a aspectos técnicos e humanísticos da competência: abordagem inicial da consulta, anamnese, exame físico, preenchimento dos formulários e orientações ao paciente, em número e conteúdo semelhantes, mas adaptados às especificidades de cada especialidade.

Para melhor diferenciar a percepção dos docentes sobre os alunos, na análise dos dados as duas subescalas do CGI, CGI<sub>t</sub> e CGI<sub>h</sub> foram consideradas como estratégias de avaliação distintas.

Dessa forma, compararam-se as notas provenientes de CGI (média ponderada de CGI<sub>t</sub> e CGI<sub>h</sub>), CGI<sub>t</sub> (média dos domínios 1 a 6), CGI<sub>h</sub> (média dos domínios 7 a 13) e AE (média das notas de todas as questões do formulário de desempenho), atribuídas por docentes de cada uma das três especialidades clínicas aos 106 alunos do 4º ano de Medicina. A análise estatística descritiva incluiu medianas, notas mínimas e máximas e distâncias interquartis. Para investigar as diferenças nas quatro distribuições foram aplicados os testes de Friedman e as comparações múltiplas com testes de Wilcoxon (W) para dados emparelhados, com correção de Bonferroni. As correlações entre as notas das três especialidades foram estudadas por meio de coeficientes de correlação de Spearman. O coeficiente alfa de Cronbach foi calculado para investigar a confiabilidade dos questionários de CGI em cada uma das três especialidades. O coeficiente de correlação de Pearson foi usado para a análise de confiabilidade entre os 13 domínios do CGI para cada especialidade. A concordância relativa entre as notas foi calculada recorrendo a distâncias Euclidianas. O nível de significância escolhido foi de 5%. O estudo foi aprovado pelo Comitê de Ética em Pesquisa, sob o número 581/2005.

## RESULTADOS

Os resultados apresentam as notas dos 103 alunos que tiveram todas as suas avaliações completas. Participaram desse estudo 19 docentes, cada um avaliando um grupo distinto de alunos. A confiabilidade do questionário foi elevada em todas as especialidades (alfa de Cronbach de 0,933 para GO, 0,939 para CM e 0,916 para Ped).

A Tabela 1 mostra que as distribuições das notas das quatro estratégias de avaliação estão concentradas no extremo direito da escala, com diferenças significativas entre elas, nas três especialidades.

As correlações entre as notas das três especialidades foram moderadas para CGI<sub>t</sub>, mas muito baixas ou mesmo negativas para AE, CGI<sub>h</sub> e CGI (Tabela 2).

Houve diferenças significativas nas distribuições das notas em seis dos 13 domínios do CGI: um técnico (*solução de problemas*) e cinco humanísticos (capacidade de autorreflexão, percepção do contexto, interação com colegas e interação com demais profissionais), após correção de Bonferroni. Essas diferenças estiveram presentes em todos os seis domínios, entre CM e Ped, em cinco deles entre GO e Ped (exceto Interação com colegas), mas em nenhum desses domínios entre GO e CM.

O impacto dos diferentes domínios na confiabilidade do CGI para cada especialidade apresentou diferenças relevantes. A Tabela 3 mostra que respeito (GO), hábitos de trabalho (CM) e interação com colegas (Ped) foram os menos correlacionados com a soma dos outros domínios. Exame físico (GO) e hábitos de trabalho (CM e Ped) foram os menos explicados pelos outros domínios. Julgamento clínico (GO e Ped), solução de problemas (CM) e capacidade de autorreflexão (Ped) produziram a maior queda no alfa, diminuindo a confiabilidade da escala, se eliminados.

A análise da concordância relativa entre as 12 notas CGI, CGI<sub>t</sub>, CGI<sub>h</sub> e AE, (em cada uma das especialidades) mostrou que as notas mais concordantes foram encontradas dentro das especialidades e não entre elas. As notas de Clínica Médica

**Tabela 1 - Estatística descritiva das estratégias de avaliação de estágio clínico para Ginecologia-Obstetrícia (GO), Pediatria (Ped) e Clínica Médica (CM) e comparação das distribuições das notas entre especialidades\*\* (N= 103)**

Estratégias de avaliação <sup>#</sup>	GO			CM			Ped Me-diana	Ped			F **	P***
	Me-diana	GO Inter quartis	GO Min-máx	Me-diana	CM Inter quartis	CM Min-máx		Ped Inter quartis	Ped Min-máx			
CGI	9,38	8,77-9,69	7,38-10	9,31	8,85-9,73	7,08-9,92	8,77	8,23-9,23	6,69-10	13,96	0,001	
CGIt	9,00	8,50-9,50	6,83-10	8,83	8,33-9,50	6,00-9,83	8,33	7,75-9,00	5,83-10	10,14	0,006	
CGIh	9,71	8,86-10	7,43-10	9,79	9,14-10	7,00-10	9,14	8,57-9,57	6,29-10	19,65	<0,001	
AE	9,60	9,20-10	6,70-10	9,40	9,00-9,75	7,00-10	9,00	8,25-9,50	6,00-10	36,35	<0,001	

<sup>#</sup> Conceito Global Itemizado (CGI), Conceito Global Itemizado-domínios técnicos (CGIt), Conceito Global Itemizado-domínios humanísticos (CGIh) e Avaliação Estruturada (AE)

\*\* F = valor da estatística qui-quadrado do teste de Friedman, GL=2

\*\*\* Nível de significância alcançado no teste de Friedman.

Observaram-se diferenças significativas entre as distribuições das notas de Ped e CM para CGI (W=5,16), CGIt (W=4,04), CGIh (W=5,34) e AE (W=4,42), e entre as notas de Ped e GO para CGI (W=4,08), CGIt (W=3,95), CGIh (W=3,63) e AE (W=6,24), todas com p<0,001. As notas de Ped foram significativamente inferiores às de GO e CM para todas as estratégias. As diferenças entre GO e CM foram significativas para AE (W=3,10, P=0,002), mas não o foram para CGI (W=0,09, P=0,927), CGIt (W=0,85, P=0,395) ou CGIh (W=0,77, P=0,443).

**Tabela 2 - Correlações de Spearman entre as notas dos docentes de GO, CM e Ped, para cada estratégia de avaliação (N= 103)**

Especialidade	CGI		CGIt		CGIh		AE	
	GO	CM	GO	CM	GO	CM	GO	CM
CM	0,19		0,24*		0,11		-0,03	
Ped	-0,02	0,48**	0,21*	0,48**	-0,23*	0,33**	0,19	0,20*

\* Significativa para P = 0,05 - teste bilateral de correlação populacional de Spearman

\*\* Significativa para P = 0,01 - teste bilateral de correlação populacional de Spearman

exibiram maior concordância relativa entre si (distância média = 8,05) do que as três notas de GO (distância média = 8,30) ou de Pediatria (distância média = 8,47). As menores distâncias foram observadas entre CM e Ped para CGI e CGIh e entre GO e CM para CGIt e AE, enquanto as notas de GO e Ped exibiram as maiores discordâncias.

## DISCUSSÃO

O estágio clínico representa um momento especial para a formação global do estudante, pois o ajuda a ampliar seu conhecimento teórico e prático, além de contribuir para que compreenda que deve dar atendimento a pessoas inseridas em contextos socioculturais diversos, com necessidades distintas, que esperam ser compreendidas, respeitadas e cuidadas<sup>8</sup>. Mas isso acarreta aos docentes/supervisores um desafio adicional, ao serem solicitados a avaliar os estudantes em um cenário complexo, com alto impacto educacional.

Neste estudo, observou-se que as notas de CGI, CGIt, CGIh e AE ficaram concentradas no extremo superior da escala, com medianas mais elevadas no CGIh, nas três especialidades. Esse resultado pode demonstrar que a maioria dos estudantes atingiu satisfatoriamente o nível de competência esperada no Módulo Atenção Integral à Saúde ou ainda sugerir um efeito de leniência entre os supervisores/docentes, especialmente para os domínios humanísticos. De fato, a literatura aponta para uma tendência entre os docentes de supervalorizarem as habilidades técnicas e serem mais condescendentes na avaliação das habilidades humanísticas<sup>9</sup>. Além disso, parece ser mais difícil justificar notas baixas em aspectos não cognitivos do que apontar e explicar o que precisa ser melhorado dentre os atributos técnicos<sup>10</sup>.

Os docentes das diferentes especialidades não atribuíram notas que discriminassem facilmente os alunos com

desempenho superior daqueles com desempenho inferior. O efeito teto observado pode indicar alguma falha no preenchimento dos formulários, talvez resultante da falta de capacitação dos docentes como avaliadores, ou ainda da ausência de referências consensuais sobre o que se espera do estudante nessa fase do curso médico. Embora nossos dados não nos permitam descartar nenhuma dessas alternativas para explicar as notas elevadas, tal tendência está relatada em outros estudos<sup>11,12,13</sup>.

As notas de Ped foram significativamente inferiores às de GO e CM nas quatro estratégias de avaliação, enquanto a única diferença estatisticamente significante entre GO e CM foi nas notas da AE. As notas de GO e Ped foram as mais discordantes. As notas de CM mostram maior concordância relativa com as de GO no CGIt e na AE, mas não no CGIh. Tais achados sugerem que os docentes de Ped parecem ser mais rigorosos, apresentando uma forma particular de avaliar os alunos e que há certa similaridade entre a percepção dos docentes de GO e CM sobre o desempenho esperado neste módulo de ensino clínico. As diferenças encontradas entre as especialidades na maneira de perceber e avaliar os alunos podem estar apoiadas nas características particulares de cada especialidade.

Diversos estudos sugerem que há uma relação entre atitudes e características pessoais e especialidade médica<sup>14,15,16</sup>. Um exemplo dessa relação é apresentado em um estudo chileno que toma como base os tipos psicológicos do Inventário de Indicadores de Tipos Psicológicos de Myers Briggs (MBTI)<sup>17,18</sup>. Para os autores a maior parte dos clínicos gerais apresenta características predominantemente do tipo denominado INTP (pessoas que procuram e desenvolvem uma explicação lógica para tudo; têm habilidade especial para solucionar problemas, são lógicas, céticas, analíticas e críticas, geralmente quietas, contidas, com maior interesse nas ideias do que na interação social, são flexíveis

**Tabela 3 - Confiabilidade do conceito global itemizado (CGI) e seus domínios para Ginecologia e Obstetrícia (GO), Clínica Médica (CM) e Pediatria (Ped) (N= 103)**

Domínios	r*			R <sup>2**</sup> (%)			Alfa#		
	GO	CM	Ped	GO	CM	Ped	GO	CM	Ped
1. Qualidade da história	0,747	0,790	0,717	64,1	83,3	86,9	0,926	0,932	0,906
2. Exame físico	0,635	0,720	0,776	56,6	75,1	88,2	0,929	0,934	0,904
3. Conhecimento médico	0,660	0,669	0,712	68,1	78,7	77,9	0,929	0,935	0,907
4. Julgamento clínico	0,788	0,573	0,785	71,0	72,3	80,9	0,924	0,938	0,903
5. Solução de problemas	0,673	0,849	0,642	62,4	86,8	63,6	0,928	0,929	0,910
6. Hábitos de trabalho	0,645	0,526	0,603	60,4	56,4	47,3	0,929	0,939	0,911
7. Comunicação	0,703	0,809	0,585	64,4	79,6	72,8	0,927	0,930	0,912
8. Respeito	0,633	0,668	0,549	71,6	72,5	76,9	0,930	0,935	0,913
9. Capacidade de autorreflexão	0,712	0,736	0,815	61,6	65,8	75,7	0,927	0,933	0,903
10. Percepção do contexto	0,676	0,798	0,693	75,1	75,4	72,9	0,928	0,931	0,907
11. Interação com colegas	0,737	0,683	0,449	82,6	77,4	54,3	0,926	0,935	0,916
12. Interação com docentes	0,781	0,770	0,455	79,2	75,8	73,9	0,926	0,933	0,916
13. Interação com demais profissionais	0,676	0,712	0,575	69,3	81,9	67,0	0,928	0,934	0,912

\*r= correlação de Pearson de cada domínio com a soma dos outros

\*\*R<sup>2</sup>= porcentagem da variação explicada pelas notas dos outros domínios

# alfa = alfa após a eliminação do domínio.

e facilmente adaptáveis). Os pediatras parecem ser preponderantemente ESFJ, cujas características mais marcantes são o cuidado e a atenção com as pessoas e o desejo de criar um clima harmonioso no ambiente de trabalho. São pessoas empáticas, leais, colaboradoras, organizadas, detalhistas, gostam de trabalhar em equipe, de cumprir prazos e de serem reconhecidas e valorizadas. Nesse estudo, os autores não fazem referência específica à especialidade de GO, mas à de Cirurgia, com a qual os ginecologistas poderiam ter maior aproximação. Observaram que os cirurgiões são geralmente do tipo ESTJ, pessoas práticas, decididas, realistas e diretas, voltadas para a implantação das decisões, focadas em resultados, cuidadosas com detalhes da rotina e gostam de passar rapidamente para a próxima etapa, pensando nos fatos de forma pragmática<sup>19</sup>.

Ao apresentarem práticas de supervisão realizadas em diferentes especialidades, Kilminster et al.<sup>4</sup> também sugeriram que parece haver características particulares de cada uma delas, mais uma vez sem dados específicos sobre GO. Na cirurgia foi valorizado o desempenho em procedimentos essenciais que um aluno deve realizar, enquanto em CM foi dada ênfase a diagnóstico, qualidade das anotações feitas e das prescrições. Em Ped, auto-controle emocional, relações interpessoais e habilidade de trabalhar em equipe foram mencionados. De forma geral, os profissionais que optam por especialidades com grande atuação em atenção primária (dentre elas, Ginecologia e Obstetrícia, Clínica Médica e Pediatria) valorizam a oportunidade de fornecer atendimento direto ao paciente e de se envolver com os aspectos psico-sociais do cuidado médico e apresentam atitudes mais centradas no paciente<sup>15,16,20</sup>. Mas a área de Ped parece ter uma orientação humanística mais acentuada, enquanto GO e CM parecem colocar mais foco em resultados. Essas características podem nos ajudar a entender a relativa proximidade entre as notas de GO e CM e a evidente particularidade das notas de Ped.

A elevada consistência interna do CGI mostrou que os docentes das três especialidades percebem uma homogeneidade

entre os domínios do questionário. Entretanto, ao compararmos as distribuições das 13 notas, observamos que a maior parte das diferenças entre as três especialidades esteve concentrada nos domínios humanísticos, especificamente nos domínios capacidade de autorreflexão, percepção do contexto e interação com demais profissionais. Uma possível explicação para esses resultados é que como os domínios humanísticos são mais passíveis de interpretações do que os técnicos, sua avaliação pode ser influenciada mais marcadamente pela “cultura” ou “personalidade” de cada especialidade<sup>21</sup>.

A avaliação estruturada foi o método no qual as diferenças significativas entre as especialidades se tornaram mais evidentes. Esse resultado reflete possivelmente a diferença de visão entre os docentes/supervisores sobre o que constitui a gama de conhecimentos e habilidades essenciais para os estudantes em estágio clínico, apesar dos objetivos instrucionais terem sido compartilhados na fase preparatória dessa atividade<sup>7</sup>.

Embora não seja possível ignorar as diferenças entre as especialidades, vale ressaltar necessidade de se buscar uma uniformização no processo de avaliação, particularmente num módulo integrado<sup>13</sup>. O fato de diferentes docentes avaliarem o mesmo aluno com critérios muito distintos pode acarretar consequências educacionais negativas tanto para o estudante, levando-o a ter dificuldade de elaborar um plano de melhoria e desenvolvimento, quanto para o corpo docente, prejudicando a distinção clara entre desempenho aceitável e inaceitável<sup>22</sup>.

Faz-se necessário, portanto, estabelecer critérios de referência para a avaliação do desempenho esperado dos estudantes ao longo do curso. Sem dúvida, essa tarefa se torna mais fácil quando se avaliam as competências cognitivas que as humanísticas, especialmente em um cenário em que múltiplas especialidades estão envolvidas. Nesse caso, o sistema de avaliação poderia ser composto de duas partes. A primeira englobando atributos comuns às diferentes áreas médicas, mensuráveis, relevantes para a prática clínica e uma segunda com um núcleo de conteúdos e domínios específicos para cada especialidade.

Esse conjunto permitiria a elaboração de uma matriz de desempenho do aluno que fosse mais facilmente comparável entre as disciplinas<sup>23</sup>.

Uma limitação do nosso estudo é que ele foi realizado em uma turma de alunos, avaliados por um pequeno número de supervisores/docentes. Portanto, não sabemos se esses dados podem ser generalizados para outros contextos. No entanto, são raras as publicações que comparam avaliações geradas por docentes de diferentes especialidades clínicas. Assim sendo, esse estudo representa uma contribuição inicial para procurar compreender como a “cultura” intrínseca de cada especialidade pode refletir no processo de avaliação realizado por docentes/supervisores.

## CONCLUSÃO

Há uma tendência dos docentes em atribuírem notas elevadas para os alunos, dificultando a discriminação entre o melhor e o pior desempenho. Os docentes de pediatria têm uma forma particular de avaliar os alunos, enquanto os docentes de ginecologia e obstetrícia e clínica médica os avaliam de forma semelhante. Enquanto os docentes de ginecologia e obstetrícia e clínica médica parecem valorizar mais os domínios técnicos, os de Pediatria parecem ter um olhar mais voltado para os humanísticos, o que pode estar de acordo com as “culturas” das três especialidades e com a tipologia psicológica dos profissionais que atuam em cada uma delas.

A qualificação dos docentes para a aplicação e interpretação de múltiplas estratégias de avaliação pode minimizar o impacto dos diferentes olhares na avaliação dos estudantes em um módulo clínico integrado.

**Conflito de interesse:** não há

## SUMMARY

### THE DIFFERENT PERSPECTIVES IN THE ASSESSMENT OF LEARNERS DURING SUPERVISED CLINICAL CLERKSHIP

**OBJECTIVE.** To compare medical students' global itemized ratings (GIR) and real-case structured clinical assessment (RC-SCA), generated by faculty members from three different specialties (Gynecology-O&G, Internal Medicine-IM, Pediatrics-Ped).

**METHOD.** 106 Y4 learners were assessed by one faculty member from each specialty, who filled in GIR, consisting of 6 technical domains (mean score GIRT) and 7 humanistic domains (mean score GIRh), on a 0-10 scale, and resultant RC-SCA, from direct attendance observation. Statistical analyses used Cronbach coefficient, Friedman and Wilcoxon paired tests, Pearson and Spearman correlation coefficients, Euclidean distances. Significance level=5%.

**RESULTS.** High internal consistency was observed in the three GIR (> 0.92). Ratings were negatively skewed. Ped scores were significantly lower than O&G and IM (median differences between 0.50 and 0.67), with low correlations between them (-0.02<R<0.48). The domains that had greater impact on the reliability of GIR were: clinical judgment (O&G and Ped), problem-solving (IM), and self-reflective skills (Ped). O&G and Ped scores showed the smallest agreement; GIRT Ped scores showed the greatest disagreement with all the other scores.

**CONCLUSION.** The specialties have different views on how to evaluate students' skills, inspite of using similar instruments, which can be a reflection of their “culture”. The challenge remains to minimize these differences through faculty development activities. [Rev Assoc Med Bras 2009; 55(4): 458-62]

**KEY WORDS:** Educational measurement. Specialties medical. Clinical clerkship.

## REFERÊNCIAS

1. Daelmans HE, Hoogenboom RJ, Donker AJ, Scherpbier AJ, Stehouwer CD, Van der Vleuten CP. Effectiveness of clinical rotations as a learning environment for achieving competences. *Med Teach.* 2004;6:305-12.
2. Kilminster S, Jolly B. Effective supervision in clinical practice: a literature review. *Med Educ.* 2000;34:827-9.
3. Brasil. Ministério da Educação. Conselho Nacional de Educação. Câmara de Educação Superior. Resolução CNE/CES n.4, 7 de nov 2001. - Diretrizes Nacionais para Curso de Graduação em Medicina. Brasília (DF): Legislação Oficial; 2001.
4. Kilminster S, Cottrell D, Grant J, Jolly B. AMEE Guide nº 27: Effective educational and clinical supervision. *Med Teach.* 2007;29:2-19.
5. Daelmans HE, Overmeer RM, Van der Hem-Stoktoos HH, Scherpbier AJ, Stehouwer CD, Van der Vleuten CP. In-training assessment: qualitative study of effects on supervision and feedback in undergraduate clinical rotation. *Med Educ.* 2006;40:51-8.
6. Swick S, Hall S, Beresin E. Assessing the ACGME competencies in psychiatry training programs. *Acad Psychol.* 2006;30:330-51.
7. Koens F, Rademakers JJDMJ, Cate OTJT. Validation of core medical knowledge by postgraduates and specialists. *Med Educ.* 2005;39:911-7.
8. Ribeiro MMF, Amaral, CFS. Medicina centrada no paciente e ensino médico: a importância do cuidado com a pessoa e o poder do médico. *Rev Bras Educ Med.* 2008;32:90-7.
9. Misch DA. Evaluating physicians professionalism and humanism: the case for humanism “connoisseurs”. *Acad Med.* 2002;77:489-95.
10. Pulito AR, Donnelly MB, Plymale M. Factors in faculty evaluation of medical students' performance. *Med Educ.* 2007;41:667-75.
11. Bergovec M, Kuzman T, Rojnic M, Makovic A. Is there grade inflation at medical schools? Case study of the Zagreb University School of Medicine. *Croat Med J.* 2003;44:92-7.
12. Speer AJ, Solomon DJ, Fincher RM. Grade inflation in internal medicine clerkship: results of a national survey. *Teach Learn Med.* 2000;12:112-6.
13. Nahum GG, Facs F. Evaluating medical student obstetrics and gynecology clerkship performance: which assessment tools are most reliable? *Am J Obstet Gynecol.* 2004;191:1762-71.
14. Fincher RME, Lewis LA, Jackson TW. Why students choose primary care or nonprimary care career. *Am J Med.* 1994;97:410-7.
15. Batenburg V, Smal JA, Lodder A, Melker RA. Are professional attitudes related to gender and medical specialty? *Med Educ.* 1999;33:489-93.
16. Haidet P, Dains JE, Paterniti DA, Hechet L, Chang T, Tseng E, et al. Medical student's attitudes toward the doctor-patient relationship. *Med Educ.* 2002;36:568-74.
17. Myers IB, MacCaulley M, Queen N, Hammer A. MBTI Manual: a guide in the development and use of the Myers-Briggs Type Indicator. Palo Alto: Consulting Psychologist Press Inc; 1998.
18. Calegari ML, Gemignani OH. Temperamento e carreira. São Paulo: Summus Editorial; 2006.
19. Bitran MC, Zuniga D, Lafuente MG, Viviani PG, Mena BC. Influencia da personalidad y el estilo de aprendizaje en la elección de especialidad médica. *Rev Med Chile.* 2005;133:1191-9.
20. Kassler WJ, Wartman SA, Silliman RA. Why medical students choose primary care careers. *Acad Med.* 1991;66:41-3.
21. Bland KI, Isaacs G. Contemporary trends in student selection of medical specialties. *Arch Surg.* 2002;137:259-67.
22. Bouriscot KAM, Roberts TE, Pell G. Standard setting for clinical competence at graduation from medical school: a comparison of passing scores across five medical schools. *Adv Health Sci Edu Theory Pract.* 2006;11:173-83.
23. Southgate L, Hays RB, Norcini J, Mulholland H, Ayres B, Wooliscroft J, et al. Setting performance standards for medical practice: a theoretical framework. *Med Educ.* 2001;35:474-81

Artigo recebido: 31/08/08  
Aceito para publicação: 03/02/09