

Exames de licenciamento em Medicina – uma boa ideia para a formação médica no Brasil?

Medical licensing examinations – are they a good idea for medical education in Brazil? (abstract: p. 15)

Exámenes de licencia en Medicina – ¿Una buena idea para la formación médica en Brasil? (resumen: p. 16)

Rafael Barbosa da Silva Bica^(a)

<rafael_bica@yahoo.com.br> 

George Edward Machado Kornis^(b)

<gkornis@gmail.com> 

^(a) Pós-graduando do Programa de Pós-Graduação do Departamento de Política, Planejamento e Administração em Saúde (Doutorado), Instituto de Medicina Social (IMS), Universidade do Estado do Rio de Janeiro (UERJ). Rua São Francisco Xavier, 524, Maracanã, Bloco D, 7^º Andar. Rio de Janeiro, RJ, Brasil. 20550-900.

^(b) Departamento de Política, Planejamento e Administração em Saúde, IMS, UERJ. Rio de Janeiro, RJ, Brasil.

A sociedade brasileira vem discutindo, recentemente, a ideia de implementar um exame de licenciamento em Medicina (ELM) para médicos recém-formados no país. Para avaliar a adequação dessa proposta e sua real necessidade, analisa-se neste trabalho se os ELMs estabelecem eficazmente a aptidão dos recém-formados para exercer a Medicina e se os sistemas de avaliação vigentes no ensino médico brasileiro já o fazem. Usando o modelo baseado em competências como referencial, discute-se a experiência da comunidade internacional com os ELMs e as características das avaliações que se propõem a medir, no Brasil, a aptidão dos recém-formados. Conclui-se que, isoladamente, ambas as alternativas são ineficazes para cumprir esse objetivo. Propõe-se o desenvolvimento de um programa mais amplo de avaliação e enfatiza-se a importância da colaboração entre os atores envolvidos para aprimorar a qualidade das avaliações dos médicos recém-formados no Brasil.

Palavras-chave: Exame de licenciamento em Medicina. Anasem. Exame do Cremesp. Avaliação do estudante de Medicina.

O crescente interesse na avaliação do ensino médico

O caso brasileiro – fatores predisponentes

A graduação em Medicina no Brasil caracteriza-se pela terminalidade: o estudante recebe a licença plena para exercer Medicina ao completar o curso em uma das escolas médicas do país¹. No entanto, como demonstra Bica², desde o início da década de 1990, diversos segmentos da sociedade brasileira têm questionado esta característica e demonstrado um maior interesse pela avaliação da qualidade do ensino médico.

Um reflexo desse fenômeno é a discussão sobre a introdução de um exame de licenciamento para avaliar a aptidão dos graduandos a exercer a profissão. Assim, deixaria de ser suficiente completar a graduação. Essa discussão foi, inclusive, uma das discordâncias centrais entre as entidades que conduziram, durante os anos 1990, a Comissão Interinstitucional de Avaliação do Ensino Médico (Cinaem)³ – a iniciativa mais ampla de avaliação do ensino médico na história recente do país e que também representou, na época, uma resposta à maior demanda social por qualidade no ensino médico brasileiro.

Mudanças estruturais recentes na sociedade brasileira explicam o fenômeno. Algumas são questões particulares ao Brasil ou que ocorrem no país com especial intensidade. Outras afetam a educação médica internacionalmente.

No primeiro caso, destaca-se o aumento no número de escolas médicas. No Brasil, entre 1994 e 2015, o total passou de 83 para 275⁴. Tal expansão, ativamente estimulada pelos governos nacionais no período², motivou temores sobre perdas na qualidade do ensino por atores como conselhos profissionais⁵, associações de escolas médicas⁶ e estudantes de Medicina⁷.

Além disso, conforme Paim et al⁸ demonstram, a definição do Sistema Único de Saúde (SUS), em 1988, como ordenador da formação de profissionais para a saúde, indicou que o currículo médico deveria articular-se às demandas da saúde pública, estando, portanto, sujeito à maior regulação. Uma evidência disso foi a elaboração, pelo governo brasileiro, das Diretrizes Curriculares Nacionais (DCNs) para o curso de Medicina em 2001 e 2014, estipulando regras para direcionar a formação médica para o perfil desejado de profissional.

No segundo caso, destaca-se a maior demanda global por garantias de qualidade na prestação de serviços médicos. Irvine⁹ identifica duas causas para esse fenômeno: o caráter crescentemente consumista da prestação dos serviços e a maior divulgação de erros médicos e falhas estruturais dos sistemas de saúde. A consciência do público sobre tais problemas e sobre seu direito a melhores serviços se reflete no aumento de processos judiciais envolvendo esta matéria. No Brasil, tal crescimento tem sido explosivo: entre 2010 e 2014, houve um aumento de 140% dos processos que chegaram em caráter recursivo ao Superior Tribunal de Justiça¹⁰.

Essa tendência reflete-se na ubiquidade do discurso de “segurança do paciente” e de “confiança do público” em propostas de sistemas internacionais de certificação de qualificações médicas¹¹. Analogamente, pode-se afirmar que a interferência do Estado brasileiro sobre o desenho curricular do ensino médico – concretizada pelas DCNs – representa também uma resposta a essas demandas sociais.

As respostas dos atores envolvidos: exame do Cremesp, Anasem e a proliferação de sistemas de avaliação

Desde a formação da Cinaem, a discussão sobre a avaliação dos estudantes de Medicina e a utilidade dos ELMs têm sido protagonizadas pela Associação Brasileira de Educação Médica (Abem), pelo Conselho Federal de Medicina (CFM), pelo Conselho Regional de Medicina do Estado de São Paulo (Cremesp) e por representantes do Governo Federal, todos assumindo, em diversas instâncias, posicionamentos alinhados a seus pontos de vista e interesses sobre a questão².

Essa mobilização produz também ações concretas. Nos últimos anos, diversos sistemas de avaliação de estudantes foram propostos à sociedade brasileira pelos atores acima mencionados, que buscam também se articular a outros grupos (entidades médicas; associações de professores e escolas de Medicina; estudantes; empregadores em saúde; e representantes do poder público) para conferir maior peso político às suas propostas.

As duas iniciativas que, com objetivos distintos, propuseram a avaliação global das competências dos estudantes de Medicina foram o exame do Cremesp, implementado em 2005, e a Avaliação Nacional Seriada dos Estudantes de Medicina (Anasem), proposta pelo governo Dilma Rousseff em 2013 e implementada em 2016. Ambas declararam, em seus documentos oficiais^{5,12}, o objetivo de avaliar todas as competências que os estudantes de Medicina devem adquirir durante a graduação. Com isso, intensificaram o debate sobre a pertinência de se introduzir um ELM para os graduandos.

Entre outros modelos recentemente propostos, destacam-se os testes de progresso: avaliações seriadas dos conhecimentos adquiridos ao longo do curso para orientar o aprendizado dos graduandos¹³. Sistemas de avaliação das escolas médicas também têm sido sugeridos como a melhor forma de garantir a formação de médicos de qualidade¹⁴.

A discussão internacional sobre ELMs: fatores conjunturais

O debate sobre ELMs não é exclusividade brasileiro e sua introdução tem sido discutida em diversos países¹⁵. Assim como no Brasil, uma possível causa é o aumento global no número de escolas médicas, fenômeno que gera, também na comunidade acadêmica internacional, a percepção de que a qualidade do ensino oferecido pelas novas escolas é insuficiente¹⁵.

Outro fator importante é a maior mobilidade internacional da força de trabalho médica, tornando inevitável que ela seja composta, em boa parte, por profissionais cuja formação se tem pouca gerência ou conhecimento¹⁵. No Brasil, esse fenômeno motiva a discussão sobre a regulamentação do Revalida¹⁶ – o ELM para egressos de escolas estrangeiras que desejam exercer a Medicina no país.

Alguns estudos recentes^{17,18} demonstram, além disso, que mesmo em países onde o ensino médico é considerado relativamente uniforme e de alta qualidade, como o Reino Unido, há significativas diferenças no desempenho dos egressos de diferentes escolas em testes padronizados e no rigor das avaliações de cada escola.

Tais fatores reforçam a percepção de que a qualidade da formação dos médicos é variável e incerta, aumentando as demandas por qualidade na prestação da assistência⁹.

Assim, apesar das particularidades do contexto brasileiro, a discussão sobre ELMs é um fenômeno internacional e tem diversos determinantes comuns. Logo, as experiências e repercussões globais a respeito do tema precisam ser consideradas quando se pensa em soluções.

Objetivo e metodologia

Considerando a dimensão do debate sobre ELMs no Brasil e a pertinência do tema, o objetivo deste trabalho é discutir a seguinte questão: a introdução de um ELM é recomendável para garantir que os recém-formados em escolas médicas brasileiras estejam capacitados a exercer a Medicina de forma segura e competente no Brasil?

A discussão se baseará em duas análises: um estudo sobre o que já se conhece a respeito da eficácia dos ELMs em países onde eles já foram introduzidos e, em seguida, se já existem, no Brasil, mecanismos de avaliação dos recém-formados em Medicina que determinem, satisfatoriamente, sua aptidão ao exercício da profissão.

Para a primeira análise, são revistos alguns conceitos essenciais em avaliação educacional. Com base em textos de referência nessa área do conhecimento e adotando-se o modelo de educação por competências como paradigma, descrevem-se as características que definem a qualidade de uma avaliação e, a partir daí, os métodos mais recomendados para aferir as competências de um estudante de Medicina. Considerando-se os formatos adotados pelos ELMs contemporâneos, discute-se sua eficácia teórica e potenciais limitações.

A eficácia prática dos ELMs é discutida a partir do estudo das experiências internacionais com esse tipo de avaliação. Com base nas referências obtidas pela recente revisão sistemática de Archer et al¹⁹ sobre o tema, analisaram-se os trabalhos que investigaram correlações entre medidas de desempenho profissional de médicos e os seus resultados em ELMs previamente realizados.

Para a segunda análise, aplicam-se os mesmos conceitos sobre avaliação educacional ao contexto do ensino médico brasileiro. Analisam-se as estruturas das duas avaliações que se propõem a medir todas as competências esperadas do recém-formado no Brasil: o exame do Cremesp e a Anasem. Para isso, utilizam-se documentos institucionais e publicações que regulamentam seus formatos e descrevem seu processo de elaboração. A partir daí, discute-se se seus respectivos desenhos permitem que elas, de fato, cumpram suas propostas de avaliação global e atinjam os objetivos, bastante distintos, que motivaram suas implementações.

Com base nessas análises, e considerando-se os fatores que vêm fortalecendo a demanda por garantias de qualidade no ensino médico brasileiro, pondera-se se existe, no Brasil, a necessidade de implementar sistemas de avaliação global dos recém-formados em Medicina e se, nesse contexto, a introdução dos ELMs seria recomendável. Analisa-se também o possível mérito de alternativas de avaliação que têm sido propostas por diversos atores. Finalmente, propõe-se um conjunto de medidas que, na visão dos autores, parece ser adequado para satisfazer a essa demanda. Espera-se, desse modo, contribuir para que o debate sobre avaliação do ensino médico brasileiro seja melhor fundamentado e, assim, permita aos interessados encontrar as melhores soluções para uma questão tão relevante.

Avaliação de estudantes de Medicina: aspectos técnicos

Considerações iniciais

Por definição, o objetivo do ELM é identificar os candidatos aptos a exercer a Medicina. Para tanto, é necessário usar métodos de avaliação que meçam, com precisão aceitável, as competências necessárias para uma prática segura e qualificada.

Os ELMs são uma modalidade de avaliação somativa: realizada para fins de certificação do indivíduo ou para definir sua aptidão a progredir para a próxima etapa de sua trajetória acadêmica ou profissional²⁰. Tais avaliações são descritas como de “alta aposta” pelas consequências que implicam para o futuro do indivíduo.

As avaliações somativas não buscam interferir no aprendizado do indivíduo – quando o fazem, é um efeito não intencional. Por outro lado, as avaliações formativas têm justamente esse objetivo: medir a proficiência do indivíduo em determinadas competências para definir objetivos de aprendizado e permitir mudanças de rumo em um programa educacional. Logo, avaliações formativas e somativas não são antagônicas, mas complementares²⁰.

Dessa forma, a escolha dos métodos de avaliação exige clareza de objetivos. As opções são abundantes. Nas últimas décadas, os métodos se tornaram mais diversificados e precisos²¹. O conhecimento de suas propriedades psicométricas permitiu identificar quais devem ser usados em cada cenário. A confiabilidade indica se o resultado do teste permite inferências sobre a proficiência do estudante no domínio que está sendo avaliado²¹. Assim, são essenciais uma boa amostragem do conteúdo em questão e uma correção bem padronizada. A validade, por sua vez, mede a correspondência entre os conteúdos testados e aqueles que se deseja testar²¹. Descreve, também, o poder do teste de identificar os variados níveis de proficiência dos candidatos. Já o impacto educacional relaciona as consequências que o teste produz sobre o processo de aprendizado²⁰. A aceitabilidade indica se diferentes atores (estudantes, professores e gestores) concordam com o formato e o conteúdo da avaliação. Finalmente, a exequibilidade considera os recursos (humanos e materiais) necessários para implementar um teste²¹.

Shumway e Harden²¹ argumentam que a aplicação desses conceitos possibilita criar avaliações altamente precisas e condizentes com os recursos disponíveis em cada instituição. Permite analisar também se determinado modelo de avaliação (como os ELMs) cumpre os objetivos aos quais se propõe.

O modelo baseado em competências: o que é avaliado e como é avaliado? Implicações teóricas para os ELMs]

Para Epstein e Hundert²², a aptidão geral de um estudante de Medicina é entendida por educadores e conselhos profissionais como o desempenho satisfatório em competências bem delimitadas que, em conjunto, reúnem as dimensões de uma prática segura e qualificada. Tais competências incluem conhecimento teórico, habilidades psicomotoras, comunicação interpessoal, ética profissional e raciocínio clínico aplicado a situações da prática médica.

Os autores argumentam que, embora as definições de competências não sejam uniformes, essa definição de aptidão é praticamente universal no ensino médico e se associa, segundo Norcini et al.²³, à crescente demanda global pela adoção de metas

educacionais, cujo cumprimento vem prevalecendo sobre a avaliação dos processos educacionais.

Assim, a implementação do modelo baseado em competências implica a realização de medidas por métodos de avaliação²³. Uma abordagem largamente usada para facilitar tais medidas foi proposta por Miller²⁴, que agrupou competências semelhantes em domínios mais amplos. A figura 1 esquematiza a ideia.



Figura 1. Pirâmide de Miller

Fonte: Adaptado de Boursicot et al.²³.

O “saber” e o “saber como” são domínios de competências cognitivas²³: o conhecimento teórico do profissional. Já os domínios seguintes reúnem as competências relacionadas à aplicação prática dos conhecimentos adquiridos. “Mostrar como” indica o que o profissional é capaz de fazer – suas habilidades clínicas²⁴. Isso inclui competências psicomotoras (manobras do exame físico, procedimentos médicos, etc.) e comportamentais (comunicar-se com pacientes e colegas, cumprir procedimentos de segurança do paciente, etc.). Já o “fazer” descreve a *performance* real do médico: como ele realmente desempenha suas funções em meio às pressões de um ambiente dinâmico, com suas incertezas e subjetividades; seu comportamento ético diante de situações emocionalmente complexas e flexibilidade diante de demandas diversas.

Conhecimentos e habilidades podem ser avaliados por medidas pontuais (como as empregadas em um ELM) de modo válido e confiável²⁵. Os domínios cognitivos são melhor avaliados por testes escritos, enquanto as habilidades clínicas são

idealmente avaliadas por métodos de simulação, como o Objective Structured Clinical Examination (OSCE).

No entanto, considerando os métodos de avaliação atualmente disponíveis, o domínio da *performance* não pode ser avaliado com confiabilidade e validade aceitáveis por medidas pontuais²⁵. São necessárias observações repetidas do profissional em seu cenário de prática ao longo de uma “janela estendida de tempo” e, de preferência, por múltiplos avaliadores. Essa é a premissa de métodos como a observação a 360º, o conceito global itemizado e o Mini-CEX.

Deve-se ressaltar, ainda, que as competências são específicas: a proficiência em uma não prevê o desempenho nas demais²⁵. Assim, somente uma combinação de métodos avalia a aptidão geral de um estudante, visto que nenhum método isolado é eficaz para os três domínios.

Os impactos educacionais da avaliação também devem ser considerados. Embora avaliações somativas como os ELMs não visem ao processo educacional, suas influências sobre o aprendizado são amplamente descritas^{26,27}. Sua importância produz estresse e leva à adoção de estratégias de estudo específicas e, muitas vezes, cansativas e de curto prazo²⁶. Isso se reflete em questões como a escolha dos temas a estudar²⁷ e a valorização da preparação teórica em relação a atividades de cunho prático. Embora a inevitabilidade desse impacto seja reconhecida, é essencial que os avaliadores procurem, por meio da elaboração cuidadosa dos testes, minimizar efeitos deletérios sobre o aprendizado^{26,27}.

As implicações das considerações acima são claras. A impossibilidade de avaliar o domínio da *performance* em medidas pontuais como os ELMs reflete-se nos seus formatos atuais, todos baseados apenas em métodos de avaliação de conhecimento e habilidades¹⁹. Se o ELM contar apenas com questões escritas, sua validade será ainda mais comprometida, por excluir todas as competências práticas da avaliação.

Além disso, o risco de impacto educacional negativo é elevado: são avaliações que envolvem um currículo extenso e fortes implicações para o futuro profissional. Embora seja possível minimizá-lo por meio de questões que privilegiem a reflexão e o raciocínio clínico, essa é uma das suas limitações mais importantes.

Exames de licenciamento para graduandos em Medicina: onde estamos?

Em que pesem as limitações teóricas, poder-se-ia argumentar que evidências de efetividade prática justificariam a implementação dos ELMs. Nessa linha de raciocínio, alguns autores procuraram associações entre as notas dos candidatos e medidas posteriores de desempenho profissional.

Tamblyn et al²⁸ descreveram a associação positiva entre as notas obtidas no ELM em Québec, no Canadá, e parâmetros como a adequação da prescrição médica ao diagnóstico do paciente. Já Wenghofer et al²⁹ descreveram associação semelhante entre a avaliação de profissionais por seus pares e seus resultados no ELM aplicado no restante do país. Norcini et al³⁰, por sua vez, descreveram uma relação entre os resultados dos profissionais no componente escrito do ELM norte-americano e indicadores posteriores de morbimortalidade cardiovascular de pacientes sob sua responsabilidade.

No entanto, esses trabalhos têm limitações metodológicas que comprometem o alcance de seus achados. Em primeiro lugar, escolheram parâmetros empíricos de qualidade: nada indica que representem a aptidão global dos profissionais. Em segundo lugar, excluem de suas amostras os médicos reprovados nos exames em questão – exatamente aqueles que os ELMs se propõem a identificar. Finalmente, limitam-se a indicar uma associação entre notas e desempenhos, sem mostrar se, independentemente dessa relação, o desempenho está adequado – o que constitui precisamente seu objetivo.

Na Europa, Archer et al¹⁹ reuniram informações sobre as estruturas e os resultados dos ELMs em diversos países. Considerando a paucidade de dados encontrados sobre os processos de construção e validação dos ELMs, os autores não puderam avaliar sua robustez psicométrica. Além disso, a escassez de trabalhos que investigam associações entre os resultados dos ELMs e medidas de desempenho profissional (como os analisados no início desta seção) levaram os autores a concluir que não se pode afirmar que os ELMs sejam benéficos ou prejudiciais para a qualidade da assistência médica de um país, não podendo, portanto, ser recomendados com base em evidências.

Assim, a discussão internacional sobre os méritos dos ELMs segue aberta. Defensores e opositores usam argumentos contundentes, mas que ainda não foram adequadamente testados. Com isso, mantém-se o caráter especulativo do debate.

Anasem e Exame do Cremesp: o que se pode afirmar sobre sua eficácia?

Em 2005, o Cremesp introduziu uma avaliação anual⁵ dos formandos de escolas médicas paulistas: o exame do Cremesp. A prova, elaborada pela Fundação Carlos Chagas e composta por questões de múltipla escolha, é opcional e seus resultados são periodicamente divulgados pelo Cremesp, que neles se baseia para fazer inferências³¹ sobre a aptidão dos graduandos e recomendar a obrigatoriedade do ELM no Brasil. Além disso, o Cremesp busca acordos com empregadores do setor saúde, outros conselhos regionais e coordenadores de programas de residência médica para expandir a aplicabilidade do exame³².

Por constituir a primeira proposta de avaliar a aptidão dos recém-formados por meio de uma prova, a iniciativa motivou discussões e posicionamentos variados pelos diversos atores envolvidos com educação médica no Brasil². Uma possível consequência dessa maior mobilização sobre o tema foi a criação da Anasem. Em 2013, por ocasião do lançamento do programa Mais Médicos, o Governo Federal estipulou³³ que os estudantes de Medicina fariam, a partir de 2016, uma avaliação seriada de suas competências em diferentes momentos da graduação. Foi a primeira iniciativa de avaliação específica do ensino médico pelo Estado brasileiro.

O governo definiu a Anasem como um exame sem cunho somativo, cujos resultados individuais seriam divulgados apenas aos estudantes para auxiliá-los no progresso de seu aprendizado¹². A Anasem foi definida como uma prova inteiramente escrita³⁴, composta por questões de múltipla escolha e discursivas, para avaliar os “conhecimentos, habilidades e atitudes” previstos nas DCNs de 2014.

Aplicando aos desenhos de prova da Anasem e do Exame do Cremesp os conceitos mais atuais sobre avaliação e os dados disponíveis sobre a efetividade dos ELMs, o que

se pode afirmar sobre o seu potencial de, efetivamente, cumprirem as tarefas a que se propõem?

A variável a se considerar inicialmente é a validade. A elaboração da avaliação deve acolher contribuições de atores como estudantes de Medicina, pacientes, docentes, médicos e empregadores³⁵. Não se trata apenas de promover a aceitabilidade da prova, mas também de aumentar sua validade de conteúdo com as percepções únicas de cada ator sobre o que é relevante para o exercício da profissão. Uma avaliação que prescindia dessas contribuições já se inicia em desvantagem.

Isso, infelizmente, caracteriza os dois exames em questão. O processo de construção do exame do Cremesp é mencionado vagamente em seus documentos oficiais^{5,31}, sem maiores explicações sobre como o conteúdo é elaborado ou menções à participação de outros atores. A Anasem, que originalmente contaria com diversos atores em sua Comissão Gestora, foi posteriormente colocada sob responsabilidade exclusiva do Instituto Nacional de Educação e Pesquisa (Inep), vinculado ao Ministério de Educação (MEC). As alusões ao aproveitamento da experiência de certos atores com outros tipos de avaliação são vagas, sem especificar como suas contribuições seriam contempladas. Assim, são desenhos de prova basicamente unilaterais.

A validade de conteúdo da Anasem e do exame do Cremesp tem outra limitação. Ambos os exames utilizam apenas questões escritas – avaliando adequadamente apenas o conhecimento teórico do candidato. As competências relacionadas às habilidades clínicas e à *performance* são ignoradas. Tal incoerência torna-se mais evidente diante do amplo consenso internacional existente a esse respeito. Assim, é desprovida de sentido a definição da Anasem¹² como avaliação de “conhecimentos, habilidades e atitudes” pelo Inep. Pela mesma razão, são incorretas as inferências do Cremesp sobre a aptidão geral dos estudantes a partir dos resultados de seu exame.

Para se aproximarem desses objetivos, as duas avaliações deveriam, no mínimo, incluir cenários de simulação (como OSCEs) para avaliar as habilidades clínicas: isso as nivelaria aos modelos contemporâneos de ELMs em alcance e limitações. É claro que esse desenho seria caro e trabalhoso, exigindo insumos e recursos humanos abundantes, o que comprometeria a exequibilidade das avaliações. No entanto, considerando a relevância dos objetivos colocados, esse maior investimento é essencial: não se pode avaliar amplamente a aptidão dos milhares de estudantes que se graduam anualmente em Medicina com recursos de qualidade técnica inferior. Esse é mais um argumento para a colaboração entre diversos atores, cujos recursos humanos e materiais poderiam se complementar.

Considerações sobre o possível impacto educacional dessas avaliações também são necessárias. No caso da Anasem, na qual esse impacto é intencional, seria essencial especificar como será usado o *feedback* dado aos estudantes para aprimorar sua formação. Será desenvolvido um plano de ação em conjunto com a escola médica? Ficar por conta do próprio aluno? Como se medirá, posteriormente, a melhora desejada no aprendizado? Nada disso é especificado nos documentos que formalizam a prova, sugerindo que a questão não foi propriamente considerada. A existência de um plano de ação que trace objetivos educacionais a partir do resultado de uma avaliação é considerada essencial³⁶ para que esta seja de fato formativa, e não apenas um exercício caro e improdutivo no qual o ato de avaliar é mais relevante do que os seus frutos.

Por outro lado, o Exame do Cremesp é uma avaliação de objetivos claramente somativos. Nas limitadas análises do Cremesp sobre o seu processo de elaboração, os riscos de impactos educacionais negativos, nos moldes descritos por Cilliers et al²⁷, não são problematizados. Da mesma forma, nada se diz sobre como o desenho das questões busca minimizá-los.

Diante dessas considerações, pode-se perceber que a Anasem e o exame do Cremesp são avaliações de desenho simplista e limitado alcance. Têm baixa validade, ignorando domínios inteiros de competências; não têm representatividade, excluindo diversos atores relevantes em sua elaboração; e não demonstram preocupação concreta com seus potenciais impactos educacionais.

Discussão – Existe lugar para ELMs no ensino médico brasileiro?

A indefinição sobre a eficácia dos ELMs torna a discussão internacional, por vezes, mais baseada em opiniões subjetivas do que em evidências¹⁹. Essa tendência é marcante na discussão brasileira. Opiniões assertivas sobre a utilidade dos ELMs e argumentos passionais têm prevalecido sobre análises pragmáticas sobre o que já se conhece e sobre as dúvidas que persistem².

O resultado é um empobrecimento técnico da discussão. Poucas são as referências à experiência internacional com ELMs e ao conhecimento psicométrico já estabelecido. É frequente aparecerem, nos posicionamentos dos diversos atores, confusões sobre os objetivos e usos dos diferentes testes, o que sugere desatualização no tema e, certamente, reflete-se em desenhos inadequados de prova, como discutiu-se acima a respeito da Anasem e do exame do Cremesp.

Outra característica notável no cenário brasileiro é a busca pela afirmação de determinados pontos de vista sobre avaliação do ensino médico em detrimento dos que lhe antagonizam. É o que Labra³⁷ descreve como “jogo de soma zero”: a busca de poder político e prestígio institucional às custas da exclusão dos opositores das arenas de tomada de decisão.

Isso se reflete nas características dos formatos de avaliação que vêm sendo propostos no Brasil². Modelos redundantes, mutuamente antagônicos e virtualmente nunca resultantes da cooperação ampla entre atores têm sido a norma – como pequenos feudos que procuram se expandir em detrimento dos demais. Não tem sido observado um esforço de colaboração para construir um modelo de avaliação minimamente aceitável para todos – o que exigiria concessões das partes interessadas.

Para analisar a eventual utilidade dos ELMs ao ensino médico brasileiro, deve-se aplicar pragmaticamente ao contexto nacional os conhecimentos e experiências internacionais acumulados a seu respeito. Para começar, fica claro que, com os métodos de avaliação disponíveis atualmente, os ELMs não avaliam eficazmente a aptidão de um recém-formado para exercer a Medicina. A sua apresentação como panaceia para problemas relacionados à má prática médica não se justifica: mesmo os modelos mais sofisticados de prova não avaliam a *performance* no ambiente de trabalho – uma dimensão essencial da aptidão.

Por outro lado, essa limitação não justifica manter as avaliações individuais das escolas médicas como garantia de qualidade no desempenho profissional. Já existem evidências claras de que o rigor dos seus sistemas de avaliação é muito variável¹⁷

e que os desempenhos dos seus egressos em avaliações externas padronizadas são significativamente variáveis¹⁸. Essas análises, tomadas em conjunto, não permitem inferências sobre a qualidade das avaliações de escolas médicas individuais. No entanto, mostram que as qualificações dos recém-formados são diferentes mesmo em países com número pequeno e estável de escolas médicas: médicos com diversos níveis de aptidão recebem o mesmo diploma, com iguais atribuições e responsabilidades. Em um país como o Brasil, no qual a variabilidade na qualidade do ensino é muito maior, essa distorção é certamente mais grave e não pode ser ignorada.

Assim, é necessário analisar modelos alternativos que se aproximem mais de atender à demanda da sociedade brasileira por qualidade nos serviços médicos que lhe são prestados. Considerando essas dificuldades, alguns autores têm proposto combinações de métodos^{24,38}, sugerindo que avaliações somativas como os ELMs poderiam ser filtros para selecionar os candidatos com os conhecimentos e habilidades adequados – competências que elas, comprovadamente, avaliam bem. A *performance* seria, em seguida, avaliada por diversos observadores em um maior intervalo de tempo.

Essa avaliação de *performance* usa métodos formativos e, em muitos países, realiza-se após o término da graduação, durante programas de *trainee* ou de residência médica³⁹. Um exemplo é o Foundation Programme, do Reino Unido: um período de trabalho supervisionado durante o qual as competências dos recém-formados são desenvolvidas e avaliadas. No fim do primeiro ano, um desempenho favorável permite o licenciamento pleno para o exercício profissional.

Evidentemente, programas desse tipo seriam de difícil implementação no Brasil. Seriam necessárias mudanças na regulamentação do licenciamento da profissão, de modo a tornar compulsório um período de prática supervisionada prévia ao seu exercício pleno. Os programas de supervisão teriam que ser estruturados para evitar que os recém-formados se tornassem, na prática, “mão de obra barata” para as unidades de saúde, sendo que muitas das quais convivem com déficits crônicos de profissionais. Seria necessário capacitar, para a avaliação, todo o conjunto de supervisores clínicos – em um país onde pouco se pratica a avaliação sistematizada e baseada em critérios objetivos. Soma-se a tudo isso, finalmente, a dificuldade de mobilizar os recursos humanos e materiais necessários à elaboração e implementação anual de um ELM para a avaliação de conhecimentos e habilidades.

No entanto, não há soluções simples para uma questão dessa complexidade. Diante do crescimento recente das escolas médicas brasileiras, o ensino de qualidade irregular e insuficiente já é realidade, bem como os seus efeitos sobre a assistência à população. Embora faça sentido criticar esse crescimento e defender maiores restrições à abertura de novas escolas, isso não reverterá as distorções já instaladas.

Uma alternativa comumente proposta por atores como a Abem e o CFM seria avaliar as escolas médicas. Tal prática é correta e desejável, especialmente se seus resultados induzirem mudanças nas práticas de ensino e gerarem consequências, como o descredenciamento, às escolas com padrões insuficientes de qualidade. No entanto, avaliações de escolas e de estudantes são práticas complementares, com estruturas e objetivos distintos. Não existem, na experiência internacional, evidências objetivas de que a avaliação das escolas garanta a todos os formandos as qualificações necessárias a uma prática de qualidade⁴⁰.

Conclusões

Garantir às sociedades nacionais uma prática médica segura e qualificada é uma demanda que, segundo Van der Vleuten⁴¹, não pode mais ser ignorada. O autor argumenta que ela se torna ainda mais dramática em países onde uma rápida expansão das escolas médicas se associa à percepção negativa sobre a qualidade do ensino oferecido. Esse parece ser, precisamente, o caso do Brasil⁵⁻⁷. Nesse contexto, fortalece-se a argumentação favorável a garantias de qualidade por meio de avaliações externas às escolas⁴².

A complexidade dessa tarefa também deve ser apreciada. Mesmo em países onde o ensino médico é percebido como uniforme e de alta qualidade, não há consenso sobre os métodos de avaliação mais adequados¹⁹. A necessidade de se combinar métodos cuidadosamente desenhados, todavia, já não está mais em questão³⁵. Não é de surpreender, portanto, que o uso isolado de avaliações pontuais ou a manutenção do *status* de terminalidade dos cursos de Medicina, alternativas mais comumente defendidas pelos atores brasileiros, sejam medidas inadequadas para este propósito.

A criação de sistemas eficazes de avaliação exige amplos recursos e a cooperação dos atores interessados³⁵. No caso brasileiro, porém, a verdadeira proliferação de novas propostas de avaliação, com metodologias e objetivos cada vez mais discrepantes e antagônicos entre si, demonstra que esses atores não estão dialogando efetivamente e seguem considerando apenas seus próprios pontos de vista. Assim, não se identificam em nossa história recente movimentos de concessões efetivas, de parte a parte, que permitam a elaboração de modelos mais qualificados e representativos de avaliação².

Neste texto, sugeriu-se uma combinação de abordagens de avaliação que acomoda a complexidade exigida pela tarefa e se baseia em conceitos sólidos sobre avaliação. A expectativa dos autores, no entanto, transcende esta proposta: o que se deseja, na verdade, é que os representantes de escolas médicas, estudantes, profissão médica, empregadores em saúde e sociedade civil possam construir, em conjunto, sistemas robustos de avaliação em que suas contribuições individuais, que são únicas e essenciais, permitam que o ensino médico brasileiro se aproxime o máximo possível de fornecer os profissionais qualificados de que necessita a assistência médica no país.

Contribuições dos autores

Todos os autores participaram ativamente de todas as etapas de elaboração do manuscrito.

Agradecimentos

O primeiro autor recebeu, durante um período de 24 meses, uma bolsa da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) para auxílio à condução de seu programa de doutorado. Não houve, para esta pesquisa, quaisquer outras fontes de financiamento.

Direitos autorais

Este artigo está licenciado sob a Licença Internacional Creative Commons 4.0, tipo BY (https://creativecommons.org/licenses/by/4.0/deed.pt_BR).



Referências

1. Brasil. Ministério da Educação e Cultura. Diretrizes curriculares nacionais do curso de graduação em Medicina. Diário Oficial da União. 6 Jun 2014.
2. Bica R. Exames de licenciamento em Medicina: uma análise da discussão brasileira e internacional [tese]. Rio de Janeiro: Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro; 2018.
3. Pierucci L. Mecanismos de avaliação do ensino superior: avaliação e disputa pela formação médica legítima (1990-2005). *Rev Espac Acad.* 2010; 9(107):72-9.
4. Scheffer M. Demografia Médica no Brasil 2015. São Paulo: Departamento de Medicina Preventiva, FMUSP, CREMESP, CFM; 2015.
5. Conselho Regional de Medicina do Estado de São Paulo. Exame do CREMESP: uma contribuição para a avaliação do ensino médico. São Paulo: CREMESP; 2009.
6. Edler F, Fonseca M. Proliferação das escolas médicas: aumento do número de médicos formados. *Cad ABEM.* 2006; 2:26-7.
7. Direção Executiva Nacional dos Estudantes de Medicina. Exame da ordem – Opinião da DENEM. USP Proj Tutores [Internet]. 2005 [citado 9 Jan 2016]. Disponível em: <http://www2.fm.usp.br/tutores/bom/bompt19.php>
8. Paim J, Travassos C, Almeida C, Bahia L, Macinko J. The Brazilian health system: history, advances and challenges. *Lancet.* 2011; 377(9779):1778-97.
9. Irvine DA. A short story of the GMC. *Med Educ.* 2006; 40(3):202-11.
10. Cambricoli F. Em 4 anos, o número de processos por erro médico cresce 140% no STJ. O Estado de São Paulo, Caderno Saúde. 22 Mar 2015.
11. Sheldon H, Swain D, Harriss L. The patient voice in revalidation: a discourse analysis. Oxford: Picker Institute Europe; 2011.
12. Brasil. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Documento Básico – Avaliação Nacional Seriada dos Estudantes de Medicina. Brasília: INEP/MEC; 2016.
13. Bicudo A, Maffei C, Abbade J, Hafner M, Yazakisun S. Avaliação do estudante: contribuição do teste de progresso. In: Lampert JB, Zeferino AMB organizadores. 10 anos das diretrizes curriculares nacionais do curso de graduação em Medicina. Rio de Janeiro: ABEM; 2014. p.15-22.
14. Tokashiki C, Sitnik D, Santos I, Vanzolini M, Martins M, Macerex N, et al. Proposta de instrumento e processo de avaliação para escolas médicas do Brasil – Relatório III. São Paulo: Projeto Avaliação das Escolas Médicas Brasileiras; 2014.
15. Swanson D, Roberts T. Trends in national licensing examinations in medicine. *Med Educ.* 2016; 50(1):101-14.



16. Conceição JC, Real GCM. Revalidação de diplomas de cursos de graduação: uma análise da política em construção. *Eccos*. 2015; (38):129-44.
17. Mccrorie P, Boursicot K. Variations in medical school graduating examinations in the UK: are clinical competence standards comparable? *Med Teach*. 2009; 31(3):223-9.
18. Mcmanus I, Elder A, Champlain A, Dacre J, Mollon J, Chis L. Graduates of different UK medical schools show substantial differences in performances on MRCP (UK) Part 1, Part 2 and PACES examinations. *BMC Med*. 2008; 6:5.
19. Archer J, Lynn N, Roberts M, Coombes L, Bere S. A systematic review on the impact of licensing examinations for doctors in countries comparable to the UK. Plymouth: Plymouth University; 2015.
20. Wass V, Bowden R, Jackson N. The principles of assessment design. In: Jackson N, Jamieson A, Khan A, editors. *Assessment in medical education and training*. Oxford: Radcliffe-Oxford; 2007. p. 11-26.
21. Shumway JM, Harden RM. AMME Guide No. 25: the assessment of learning outcomes for the competent and reflective physician. *Med Teach*. 2003; 25(6):569-84.
22. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002; 287(2):226-35.
23. Norcini J, Lippner RJ, Grosso LI. Assessment in the context of licensure and certification. *Teach Learn Med*. 2013; 25 Suppl 1:62-7.
24. Boursicot K, Etheridge L, Setna Z, Sturrock A, Ker J, Smee S, et al. Performance in assessment: consensus statement and recommendations from the Ottawa conference. *Med Teach*. 2011; 33(5):370-83.
25. Van Der Vleuten C, Schuwirth L, Scheele F, Driessen E, Hodges B. The assessment of professional competence: building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol*. 2010; 24(6):703-19.
26. Al-Kadri H, Al-Moamary M, Roberts C, Van Der Vleuten C. Exploring assessment factors contributing to students' study strategies: a literature review. *Med Teach*. 2012; 34 Supl 1:42-50.
27. Cilliers F, Schuwirth L, Herman N, Adendorff H, Van Der Vleuten C. A model of the pre-assessment learning effects of summative assessments in medical education. *Adv Health Sci Educ Theory Pract*. 2012; 17(1):39-53.
28. Tamblyn R, Abrahamowicz M, Dauphinee W, Hanley J, Norcini J, Girard N, et al. Association between licensure examination scores and practice in primary care. *JAMA*. 2002; 288(23):3019-26.
29. Wenghofer E, Klass D, Abrahamowicz A, Dauphinee D, Jacques A, Smee S, et al. Doctor scores on national qualifying examinations predict quality of care in future practice. *Med Educ*. 2009; 43(12):1166-73.
30. Norcini J, Boulet J, Opalek A, Dauphinee W. The relationship between licensing examination performance and the outcomes of care by IMGs. *Acad Med*. 2014; 89(8):1157-62.
31. Conselho Regional de Medicina do Estado de São Paulo. Exame do CREMESP: agora é obrigatório. São Paulo: CREMESP; 2014.
32. Conselho Regional de Medicina do Estado de São Paulo. Exame do CREMESP: avaliação de recémgraduados passa a ter valor de mercado para hospitais e planos de saúde. CREMESP Not [Internet]. 2015 [citado 25 Set 2015]. Disponível em: <https://www.cremesp.org.br/?siteAcao=NoticiasC&id=3681>



33. Brasil. Presidência da República. Lei nº 12871/2013, de 22 de Out de 2013. Institui o Programa Mais Médicos e dá outras providências. Diário Oficial da União. 23 Out 2013.
34. Brasil. Ministério da Educação e Cultura. Portaria nº 982/2016, de 25 Ago de 2016. Institui a Avaliação Nacional Seriada dos Estudantes de Medicina. Diário Oficial da União. 26 Ago 2016.
35. Norcini J, Anderson B, Bollela V, Burch V, Costa M, Duvivier R, et al. Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 conference. *Med Teach*. 2011; 33(3):206-14.
36. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach*. 2007; 29(9):855-71.
37. Labra ME. Associativismo no setor saúde brasileiro e organizações de interesse do empresariado médico. *Physis*. 1993; 3(2):193-225.
38. Schuwirth L. The need for national licensing examinations. *Med Educ*. 2007; 41(11):1022-3.
39. Wijnen-Meijer M, Burdick W, Alofs L, Burgers C, Ten Cate O. Stages and transitions in medical education around the world: clarifying structures and terminology. *Med Teach*. 2013; 35(4):301-7.
40. Van Zanten M, Mckinley D, Montiel I, Pijano C. Medical education accreditation in Mexico and the Phillipines: impact on student outcomes. *Med Educ*. 2012; 46(6):586-92.
41. Van Der Vleuten C. National, European licensing examinations or none at all? *Med Teach*. 2009; 31(3):189-91.
42. Van Der Vleuten C. National licensing examinations and their challenges. *J Heal Spec*. 2013; 1(1):1-2.

The introduction a medical licensing examination (MLE) for the country's newly qualified doctors has been the subject of discussion recently in Brazil. To assess the suitability of and real need for this test, this article analyzes whether the MLE provides an effective measure of the aptitude of newly qualified doctors and whether the evaluation system currently in place in medical education already does this. Using a competency-based model, we discuss experiences with MLEs in other countries and the features of the examinations proposed in Brazil to measure the aptitudes of newly qualified doctors. It is concluded that, in isolation, both alternatives are inadequate. We propose the development of a broader assessment program and stress the importance of collaboration between the different actors involved in order to enhance the quality of the examinations proposed to test the aptitude of newly qualified doctors in Brazil.

Keywords: Medical licensing examination. ANASEM. The CREMESP examination. Medical student assessment.



La sociedad brasileña ha discutido recientemente la idea de implementar un examen de licencia (ELM) para médicos recién graduados en el país. Para evaluar la adecuación de esta propuesta y su real necesidad, se analiza en este trabajo si los ELMs establecen eficazmente la aptitud de los recién graduados para ejercer la Medicina y si los sistemas de evaluación vigentes en la enseñanza médica brasileña ya lo hacen. Usando el modelo basado en competencias como factor referencial, se discute la experiencia de la comunidad internacional con los ELMs y las características de las evaluaciones que se proponen medir, en Brasil, la aptitud de los recién graduados. Se concluye que, aisladamente, ambas alternativas son ineficaces para cumplir ese objetivo. Se propone el desarrollo de un programa más amplio de evaluación y se enfatiza la importancia de la colaboración entre los actores envueltos para perfeccionar la calidad de las evaluaciones de los recién graduados en Brasil.

Palabras clave: Examen de licencia en Medicina. Anasem. Examen del Cremesp. Evaluación del estudiante de Medicina.

Submetido em 24/09/18.
Aprovado em 01/05/19.