

O Teste de Progresso na residência médica em ginecologia e obstetrícia: a experiência nacional

The Progress Test in gynecology and obstetrics medical residency: the national experience

Gustavo Salata Romão¹  gsalata@uol.com.br

Cesar Eduardo Fernandes²  cesarefernandes_dircient@febrasgo.org.br

Aginaldo Lopes da Silva Filho³  agnaldo.ufmg@gmail.com

Marcos Felipe Silva de Sá⁴  marcosfelipe@fmrp.usp.br

RESUMO

Introdução: O Teste de Progresso (TP) é uma avaliação cognitiva abrangente e longitudinal com vantagens para o aprendiz, os programas educacionais e a sociedade. No Brasil, embora seja amplamente utilizado na graduação, existe pouca experiência do seu uso na residência médica. Este estudo visa retratar a experiência da Federação Brasileira das Associações de Ginecologia e Obstetrícia (Febrasgo) nos quatro primeiros anos de implementação do TP Individual do Residente em Ginecologia e Obstetrícia (TPI-GO).

Relato da experiência: O TPI-GO foi implementado em 2018, sendo oferecido anualmente aos residentes brasileiros. Nos dois últimos anos, o TPI-GO passou a ser oferecido no formato on-line. Para incentivar a participação, estabeleceram-se critérios que possibilitam dispensa ou bonificação na prova teórica para a obtenção do Título de Especialista em Ginecologia e Obstetrícia (Tego) para os candidatos com melhor desempenho. Os resultados são fornecidos de maneira sigilosa a cada candidato. Os resultados do desempenho dos residentes de um mesmo serviço são fornecidos ao supervisor do programa sem a identificação dos seus residentes, de modo a oferecer subsídios para autoavaliação da qualidade e indicar pontos de melhoria.

Discussão: Desde a implementação do TP, verificou-se um número crescente de candidatos inscritos com elevada taxa de aderência à prova. As medianas de desempenho do grupo de residentes evidenciaram evolução cognitiva entre os iniciantes (R1) e os concluintes (R3) dos programas.

Conclusão: A implementação do TP na residência é factível e aceitável, apresentando vantagens e benefícios. No entanto, requer preparo e envolvimento da equipe organizadora, além de apoio logístico e financeiro. O sigilo na divulgação dos resultados é recomendável e preserva o candidato. A possibilidade de bonificação e dispensa na prova teórica do Tego é um grande atrativo para a adesão e aderência dos médicos residentes ao TP. A aplicação da prova em ambiente on-line amplia o acesso, com níveis aceitáveis de segurança.

Palavras-chave: Educação Médica; Residência Médica; Avaliação Educacional; Teste de Progresso; Ginecologia e Obstetrícia.

ABSTRACT

Introduction: *The Progress Test (PT) is a comprehensive, longitudinal cognitive assessment with benefits for the learner, educational programs, and society. Although it is widely used in undergraduate courses in Brazil, there is little experience regarding its use in medical residency. This study aims to portray Febrasgo's experience in the first four years of implementing the Resident's Individual PT in ObGyn (TPI-GO).*

Experience report: *The TPI-GO was implemented in 2018, being offered annually to Brazilian residents. In the last two years, the TPI-GO has been provided online. Febrasgo established criteria that allow for exemption or bonus in the theoretical test to encourage the trainees' participation in the TPI-GO to obtain the Specialist Certification in Gynecology and Obstetrics (TEGO). Results are provided confidentially to candidates and allow self-assessment. The results of the residents' performance from the same service are provided to the program supervisor, offering subsidies for self-assessment of quality and indicating points for improvement.*

Discussion: *Since it was implemented, there has been an increasing number of candidates enrolled in the TPI-GO and a high adherence rate. The medians of performance of the two cohorts of residents studied showed cognitive evolution between beginners (R1) and graduates (R3), with differences attributed in part to the adverse effects of the Pandemic on training.*

Conclusion: *The implementation of the PT in medical residency is feasible, acceptable, and shows advantages and benefits, but it requires the commitment of the organizing team, besides logistical and financial support. The results confidentiality is recommended and preserves the candidate. The possibility of granting bonuses and exemptions in the TEGO theoretical test is an excellent attraction aiming at the adherence and adherence of residents to the PT. The TPI-GO online increases access with acceptable levels of security.*

Keywords: *Medical Education; Medical Residency; Educational Assessment; Progress Test; Obstetrics and Gynecology.*

¹Universidade de Ribeirão Preto, Ribeirão Preto, São Paulo, Brasil.

²Centro Universitário Faculdade de Medicina do ABC, Santo André, São Paulo, Brasil.

³Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brasil.

⁴Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil.

Editor: Aristides Augusto Palhares Neto.

Recebido em 07/09/22; Aceito em 04/10/22.

Avaliado pelo processo de *double blind review*.

INTRODUÇÃO

O ensino e a avaliação do conhecimento clínico são fundamentais durante a formação médica. Entre as diferentes modalidades de avaliação cognitiva, o Teste de Progresso (TP) se destaca por ser um método abrangente, longitudinal e seriado, em que um mesmo conjunto de itens é aplicado a aprendizes que se encontram em diferentes fases ou períodos de treinamento¹. Quando comparado a outras modalidades de avaliação cognitiva, apresenta vantagens como promover a autoavaliação e avaliação formativa, permitindo reafirmar saberes prévios e identificar lacunas de aprendizagem². Por causa de sua abrangência, representa um estímulo à aprendizagem contínua e um desestímulo ao estudo de véspera de prova, conscientizando e engajando o aprendiz em seu processo metacognitivo³. Dessa forma, o TP exerce um efeito educacional favorável não somente em termos de aprendizagem funcional, como também em relação à retenção do conhecimento em longo prazo⁴.

Sob a perspectiva dos professores e preceptores, o TP fornece informações valiosas sobre o desempenho dos aprendizes e permite a identificação daqueles que necessitam de maior apoio, remediação ou progressão acelerada. Adicionalmente o TP provê maior estabilidade ao processo de avaliação cognitiva, não sendo necessário adaptá-lo às eventuais mudanças curriculares de um programa¹.

O uso do TP na residência médica tem sido descrito desde 1999, quando foi implementado na Holanda para avaliação dos médicos residentes em ginecologia e obstetrícia⁵. No Brasil, embora o TP seja amplamente utilizado na graduação em Medicina, o uso na residência tem sido bastante restrito, já que é oferecido por poucas especialidades médicas.

Em resposta aos anseios e às expectativas da sociedade por uma prática médica mais qualificada e segura, o TP apresenta maior validade preditiva da qualidade dos futuros profissionais quando comparado a outras modalidades de avaliação cognitiva pontual⁶. Por esses e outros motivos, a Federação Brasileira das Associações de Ginecologia e Obstetrícia (Febrasgo) desenvolveu o TP Individual do Residente em Ginecologia e Obstetrícia (TPI-GO) para aplicação em todo o Brasil a partir de 2018⁷. Este artigo tem por objetivo retratar a experiência dos quatro primeiros anos de implantação do TPI-GO.

RELATO DE EXPERIÊNCIA

Este é um estudo descritivo que relata a experiência da Febrasgo na implantação do TPI-GO e no acompanhamento longitudinal de duas coortes de médicos residentes em ginecologia e obstetrícia. Cada coorte representa um mesmo grupo de residentes ao longo dos três anos consecutivos de participação no TPI-GO. A primeira coorte refere-se

aos residentes que iniciaram a participação no TPI-GO em 2018 como R1, participaram da prova em 2019 como R2 e concluíram a participação em 2020 como R3. Já a segunda coorte corresponde aos residentes que iniciaram a participação no TPI-GO em 2019 como R1, participaram da prova em 2020 como R2 e concluíram a participação em 2021 como R3.

Este relato de experiência utilizou um banco de dados que não foi produzido com a finalidade de pesquisa e cujas informações são agregadas, sem possibilidade de identificação individual dos sujeitos. Sendo assim, nesse caso, este relato fica dispensado de ser registrado ou avaliado pelo sistema de Comitês de Ética em Pesquisa da Comissão Nacional de Ética em Pesquisa (CEP/Conep), conforme Resolução nº 510, de 7 de abril de 2016, do Conselho Nacional de Saúde⁸.

Implementação do TPI-GO

Em 2018, o TPI-GO foi implementado pela Comissão de Residência Médica (Coreme) da Febrasgo e disponibilizado a todos os residentes regularmente matriculados em programas nessa especialidade credenciados pelo Ministério da Educação (MEC)⁷.

A prova teórica para obtenção do Título de Especialista em Ginecologia e Obstetrícia (Tego) concedido pela Febrasgo é a base para o TPI-GO. A matriz da prova é baseada na Matriz de Competências em Ginecologia e Obstetrícia^{9,10} e assegura a distribuição balanceada entre as subáreas de atuação do ginecologista e obstetra. A prova contém 100 questões de múltipla escolha com quatro alternativas e apenas uma opção correta, sendo 50 de ginecologia e 50 de obstetrícia. As questões são elaboradas e revisadas pela Comissão Nacional do Tego (CN-Tego)¹¹.

O TPI-GO é oferecido anualmente aos residentes do primeiro (R1), segundo (R2) e terceiro (R3) anos de treinamento. A participação é voluntária, e, para estimular a adesão, foram estabelecidos critérios que possibilitam a dispensa ou bonificação na prova teórica do Tego no ano subsequente à conclusão do programa de residência médica (PRM). De acordo com esses critérios, serão dispensados da prova teórica do Tego os candidatos que participarem do TPI-GO como R1, R2 e R3, apresentarem um desempenho igual ou superior ao percentil 60 (P60) em relação aos seus pares em pelo menos duas versões da prova (incluindo necessariamente o desempenho como R3) e obtiverem uma percentagem de acertos de pelo menos 60% nas questões de ginecologia e 60% de obstetrícia na última aplicação (como R3). Serão bonificados na prova teórica do Tego os candidatos que participarem do TPI-GO como R1, R2 e R3, apresentarem um desempenho superior ao percentil 30 (P30) em relação aos seus pares em todas as versões da prova e obtiverem uma percentagem de acertos de pelo menos 50%

nas questões de ginecologia e de obstetrícia na última aplicação (como R3). Nesta última condição, os candidatos terão um bônus de 10% nas questões de ginecologia e de obstetrícia na prova teórica do Tego no ano subsequente¹¹. A adoção dos percentis 30 e 60 (P30 e P60) como pontos de corte na prova teórica do Tego foi estabelecida pela equipe que implementou a avaliação por meio do método de Angoff modificado¹². Essa escolha se baseou no pressuposto de que pontos de corte referenciados por norma são mais recomendados do que valores fixos (*cut-off*) por conta da variação nos níveis de dificuldade entre as diferentes versões do TP^{11,13,14}.

A aplicação do TPI-GO ocorre simultaneamente à prova teórica do Tego, sob a coordenação da Coreme-Febrasgo. Inicialmente, o TPI-GO foi oferecido presencialmente em 11 cidades: Porto Alegre (RS), Curitiba (PR), São Paulo (SP), Ribeirão Preto (SP), Rio de Janeiro (RJ), Belo Horizonte (MG), Salvador (BA), Recife (PE), Fortaleza (CE), Brasília (DF) e Belém (PA)¹¹.

A partir de 2020, passou a ser oferecido na modalidade *on-line*, o que permitiu ao candidato realizar a prova em seu domicílio utilizando computador próprio. Para garantir a segurança, os candidatos instalam um aplicativo que bloqueia o acesso ao navegador de internet e permite a monitorização contínua e automática, utilizando a câmera e o microfone do próprio computador. Além disso, os candidatos são monitorados remotamente por fiscais, sendo possível detectar sons, movimentos corporais e desvios do olhar, bem como a presença de outras pessoas ou artefatos suspeitos no ambiente. Quando há infrações ou tentativas de fraude, os candidatos são advertidos, e, em casos de persistência, a prova é interrompida¹⁵.

A divulgação dos resultados do desempenho individual é feita somente ao próprio candidato e de maneira sigilosa. São fornecidas informações sobre o seu desempenho geral e nas questões de ginecologia e de obstetrícia, e o desempenho progressivo em relação aos anos anteriores. Também são disponibilizados gráficos comparativos em relação ao desempenho geral de seus pares (R1, R2 ou R3) em nível nacional, utilizando como parâmetros a mediana, o percentil 30 (P30) e o percentil 60 (P60). Para os R3, o laudo inclui informações sobre a dispensa ou bonificação na prova teórica do Tego¹¹. Para o cálculo desses valores e a identificação dos candidatos que preenchem os critérios para bonificação ou dispensa na prova teórica do Tego no ano subsequente, utilizou-se o pacote de análise estatística do Excel com seus diversos recursos

disponíveis. Esse mesmo *software* foi utilizado na elaboração dos gráficos do desempenho dos residentes em relação ao P30, à mediana e ao P60 do desempenho geral em nível nacional. A análise estatística das diferenças de desempenho entre R1 e R2 e entre R2 e R3 em cada coorte foi realizada por meio do teste de Mann-Whitney considerando a não homogeneidade da distribuição de notas dos residentes. Em todas as análises, consideraram-se significativas as diferenças para $p < 0,05$. A consistência interna das provas de ginecologia e de obstetrícia passou a ser analisada a partir de 2020 por meio do coeficiente alfa de Cronbach, considerando-se adequados os valores de alfa acima de 0,64¹⁶. Para as análises foi utilizado o *software* IBM SPSS Statistics – V22.

O consolidado do desempenho dos residentes de um mesmo serviço é enviado ao supervisor do PRM. Com base nesses resultados, é possível avaliar a evolução do grupo ao longo do PRM e compará-lo com as curvas de desempenho nacional em cada categoria (R1, R2 e R3). Para preservar o sigilo, somente poderão receber esses resultados os supervisores de PRM em que pelo menos três residentes de cada categoria se aplicaram ao TPI-GO¹¹. Os residentes não são identificados no relatório.

Ao final de cada aplicação do TPI-GO, realiza-se a análise psicométrica dos itens quanto ao grau de dificuldade e poder discriminativo, fornecendo bases para os trabalhos da comissão de revisores no ano subsequente¹⁷.

Perfil dos residentes participantes

Desde 2018, quando o TPI-GO foi implementado, tem-se observado aumento do número de participantes (Tabela 1), com predomínio de R1 em todos os anos. Quanto ao perfil, a idade dos candidatos oscila entre 24 e 52 anos, com média de 28,3 e mediana de 28,2 anos, e predomínio de mulheres (87%). Em relação ao estado civil, 84% dos candidatos são solteiros, 15% são casados e 0,8% é divorciado. Quanto à localização dos PRM de origem dos candidatos, 117 (52%) estão no Sudeste, 42 (19%) no Sul, 32 (14%) no Nordeste, 20 (9%) no Centro-oeste e 12 (5%) no Norte.

Analisamos os candidatos em duas coortes com medidas repetidas. Na primeira, foram incluídos os candidatos que realizaram as três provas no período de 2018 a 2020, e, na segunda, os candidatos que o fizeram no período de 2019 a 2021. A aderência à prova foi avaliada por meio da proporção de candidatos que se aplicaram ao TPI-GO como R3 em relação ao número de candidatos que se aplicaram à prova como R1 no início de cada coorte (Tabela 2).

Tabela 1. Médicos residentes inscritos no TPI-GO por ano

Ano	Total de inscritos	R1	R2	R3
2018	1.202	497	360	345
2019	1.314	568	457	289
2020	1.422	628	480	314
2021	1.495	634	501	360

Fonte: Elaborada pelos autores.

Tabela 2. Aderência dos candidatos ao TPI-GO

Coorte	Total de R1 em 2018	Total de R3 em 2020	Aderência
De 2018 a 2020	497	314	63%
De 2019 a 2021	568	360	63%

Fonte: Elaborada pelos autores.

Desempenho dos residentes participantes

A evolução dos candidatos no TPI-GO reflete o ganho de conhecimento adquirido ao longo do treinamento e pode ser avaliada a partir do desempenho geral dos candidatos em nível nacional, em cada coorte (gráficos 1 e 2). Na coorte 1, houve aumento estatisticamente significativo do desempenho entre os R1 e os R2 ($p < 0,001$), sem aumento estatisticamente significativo do desempenho entre os R2 e os R3 ($p = 0,584$). Na coorte 2, houve redução estatisticamente significativa do desempenho entre os R1 e os R2 ($p < 0,001$), com aumento estatisticamente significativo do desempenho entre os R2 e os R3 ($p < 0,001$).

A consistência interna das provas, medida pelo coeficiente alfa de Cronbach para as provas de ginecologia e obstetrícia, foi 0,640 e 0,560 em 2020, e 0,647 e 0,709 em 2021, respectivamente, o que se traduz em valores adequados na prova de ginecologia em 2020 e 2021, e na prova de obstetrícia em 2021, segundo os critérios aqui utilizados.

DISCUSSÃO

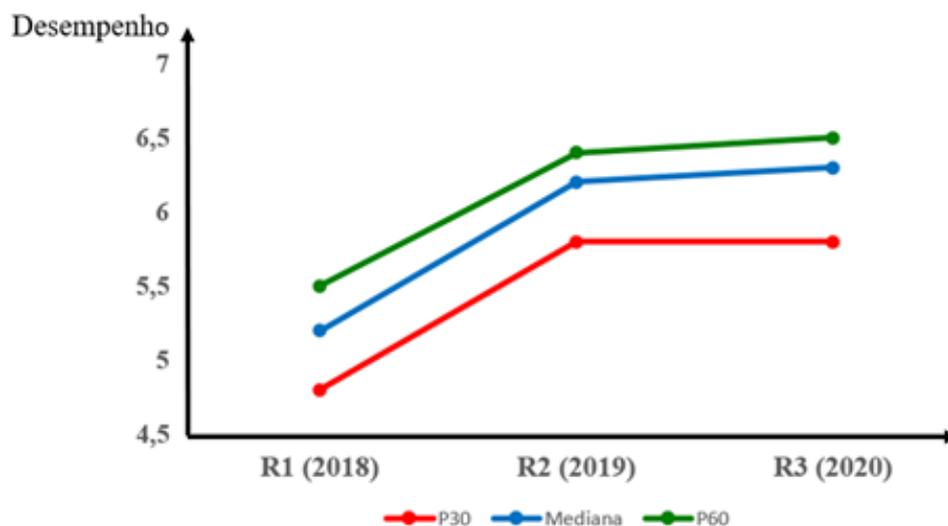
O TPI-GO é uma iniciativa da Febrasgo ante o conhecimento das vantagens e dos benefícios do teste, e diante da constatação da inexistência de uma avaliação cognitiva confiável, abrangente e unificada que pudesse ser oferecida nacionalmente a todos os médicos residentes em ginecologia e obstetrícia. O desenvolvimento e a implementação do TPI-GO envolveram um meticuloso trabalho da Coreme-Febrasgo e da CN-Tego, com o apoio da diretoria científica e da presidência da Febrasgo em duas gestões consecutivas^{7,11}.

Desde a implementação em 2018, o TPI-GO tem apresentado um número crescente e expressivo de candidatos inscritos, atingindo um total 634 candidatos iniciantes

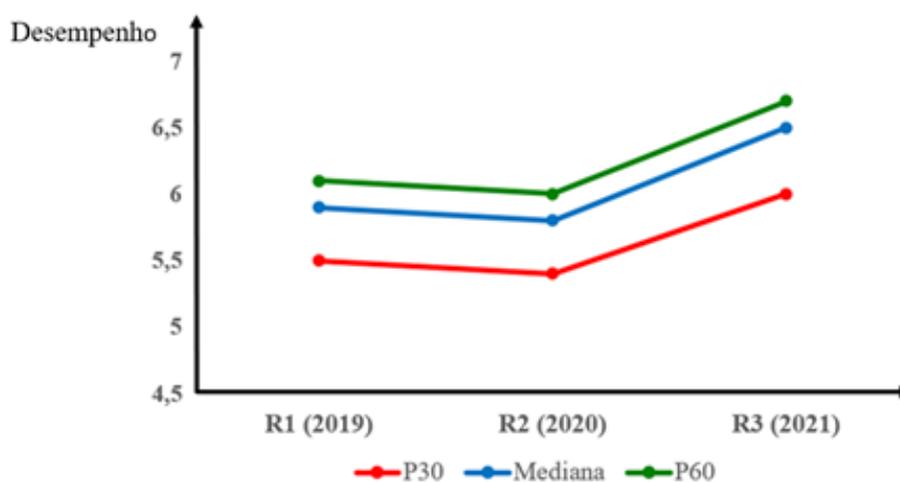
(R1) inscritos em 2021, o que corresponde a 48% do total de R1matriculados em PRM credenciados pelo MEC nessa especialidade ($n = 1.325$)¹⁸. Também são notórias as taxas de aderência dos candidatos nos três anos consecutivos de aplicação da prova, atingindo 63% nas duas coortes avaliadas. Diversas são as razões para o êxito desse modelo de avaliação. A começar pelo sigilo na divulgação dos resultados ao candidato, de modo a evitar a sua exposição perante os pares, preceptores e supervisores, e prevenir atitudes discriminativas que levem ao constrangimento, ao estresse ou à perda da autoconfiança. A avaliação formativa orientada para o aprendizado, o sigilo e o *feedback* construtivo oferecem melhores resultados quando comparados às avaliações excessivamente competitivas que expõem os aprendizes à desqualificação ou depreciação¹⁹. Diversos estudos demonstraram um ganho consistente no aprendizado e na retenção de conteúdo em longo prazo a partir da avaliação formativa e do *feedback* pós-teste²⁰⁻²².

Outro fator determinante para a adesão e aderência dos candidatos ao TPI-GO é a possibilidade de dispensa ou bonificação na prova teórica do Tego. Ser dispensado ou bonificado não significa apenas uma maior facilidade na obtenção do título, mas também o reconhecimento do seu desenvolvimento na especialidade, corroborado pelo desempenho satisfatório em uma avaliação confiável.

Também deve ser considerada a maior facilidade de acesso ao TPI-GO pela aplicação em ambiente virtual a partir de 2020. A modalidade *on-line* permite ao candidato realizar o TPI-GO utilizando seu próprio computador, com um sistema de monitorização que garante níveis aceitáveis de segurança. Dessa forma, prescinde-se da necessidade de deslocamento até os centros aplicadores, o que representa um grande desafio para os serviços que dependem da força

Gráfico 1. Evolução do desempenho dos residentes no TPI-GO na coorte de 2018 a 2020

Fonte: Elaborado pelos autores.

Gráfico 2. Evolução do desempenho dos residentes no TPI-GO na coorte de 2019 a 2021

Fonte: Elaborado pelos autores.

de trabalho dos residentes, além de reduzir os gastos com passagens e hospedagem.

No tocante ao perfil dos candidatos, verifica-se um predomínio das mulheres. Esse achado é corroborado por outros levantamentos que demonstram a inserção cada vez maior das mulheres nos PRM brasileiros¹⁸. Em relação à procedência, mais da metade é oriunda de PRM da Região Sudeste, que concentra 54,5% dos residentes em ginecologia e obstetria¹⁸.

O desempenho geral no TPI-GO é de grande valia no acompanhamento do ganho cognitivo pelos residentes ao longo do treinamento. O conhecimento é um componente essencial das competências do médico e tem impacto direto sobre o seu desempenho profissional²³. Logo, a verificação e a mensuração do conhecimento adquirido durante o PRM se

tornam essenciais. A magnitude de aquisição do conhecimento pode ser estimada pelo “tamanho do efeito”, cujo valor reflete a força da relação entre uma intervenção pedagógica e o aprendizado²⁴. Uma importante finalidade do TP é avaliar o ganho cumulativo de conhecimento e o efeito do treinamento sobre o aprendizado dos residentes²⁴.

A análise da primeira coorte de candidatos que iniciaram a aplicação no TPI-GO em 2018 como R1, participaram da prova em 2019 como R2 e completaram a participação em 2020 como R3 (coorte 2018-2020) mostra um aumento progressivo na mediana do desempenho. Entretanto, o ganho foi expressivamente maior do R1 para o R2, em que se verificou um aumento estatisticamente significativo. Já para a diferença de desempenho entre os R2 e os R3, não foi verificada significância estatística. Diversas hipóteses

poderiam justificar a estagnação do ganho cognitivo verificada entre os R2 e os R3 na primeira coorte.

Primeiramente, é preciso considerar as potenciais diferenças no grau de dificuldade da prova entre as diversas versões do TPI-GO. Considerando que a Febrasgo não dispõe de um banco de itens para elaboração da prova, não é possível assumir que não existam oscilações nesse quesito. Os consórcios internacionais recomendam o armazenamento de questões pré-testadas quanto ao poder discriminativo e ao grau de dificuldade em um banco de itens. Cada nova versão da prova deve conter pelo menos 30% das questões provenientes do banco (questões-âncoras) para garantir a estabilidade da avaliação. Para reduzir a margem de acertos por memorização de questões, os itens devem permanecer no banco por um período mínimo de três anos para que possam ser reutilizadas. No entanto, a legislação brasileira exige a divulgação das questões da prova do Tego imediatamente após a sua aplicação, visando assegurar maior transparência e *feedback* aos candidatos. Essa divulgação, embora vantajosa, dificulta a reutilização de questões e limita o valor do banco de itens. Apesar dessas dificuldades, a Febrasgo está se empenhando na construção de um banco de questões validadas após sua aplicação no Tego.

Outro aspecto a ser considerado se refere à heterogeneidade das condições de aplicação do TPI-GO, que foi presencial em 2018 e 2019 e *on-line* a partir de 2020. A utilização do formato *on-line* representou uma mudança significativa no ambiente e formato da avaliação. Algumas normas de segurança foram estabelecidas, destacando-se o fato de que o candidato precisa completar uma questão para visualizar a próxima e não pode retornar a uma questão já respondida. É possível que a primeira aplicação digital tenha representado uma fonte de estresse e insegurança aos candidatos, muito embora alguns consórcios internacionais (como o consórcio alemão) tenham verificado maior preferência dos candidatos pela modalidade *on-line* do TP²⁵.

Finalmente, é preciso considerar os efeitos adversos da pandemia Covid-19 sobre o treinamento dos médicos residentes, particularmente em 2020. Durante o período mais crítico, diversas adaptações nos PRM foram necessárias, tanto para atender às exigências sanitárias quanto para reforçar as equipes no atendimento aos pacientes acometidos²⁶. Entre diversas mudanças, destacam-se o adiamento de consultas rotineiras, a suspensão de cirurgias eletivas e o deslocamento dos residentes para unidades de pronto atendimento, enfermarias e unidades de terapia intensiva²⁷. Ademais, houve desfalque nas equipes de supervisores em decorrência dos afastamentos por doença ou por enquadramento nos grupos de alto risco²⁷. As atividades teóricas foram adaptadas para o

ensino remoto emergencial, o que representou um desafio aos preceptores e supervisores²⁷. No treinamento em assistência à saúde da mulher, embora tenham sido mantidos os serviços essenciais nas maternidades e urgências, houve um déficit significativo nas outras áreas de aprendizagem²⁷. É plausível a hipótese de que todas essas adaptações durante a pandemia tenham impactado negativamente o desempenho dos residentes no TPI-GO de 2020, repercutindo negativamente na mediana de desempenho dos R3 da primeira coorte.

Na segunda coorte, em que os candidatos iniciaram a aplicação no TPI-GO em 2019 como R1, participaram da prova em 2020 como R2 e completaram a participação em 2021 como R3 (coorte 2019-2021), observa-se uma redução estatisticamente significativa do desempenho do R1 para o R2 com aumento estatisticamente do desempenho do R2 para o R3. Um dos fatores relacionados à redução da mediana dos R1 para o R2 nessa coorte foi a pandemia da Covid-19. Como já foi discutido anteriormente, essa emergência mundial de saúde pública exerceu impacto negativo sobre o treinamento dos residentes em diversas especialidades médicas, incluindo a ginecologia e obstetrícia²⁷. Talvez esse fator seja um dos responsáveis pelo desempenho abaixo das expectativas dos R2 que se aplicaram à prova em 2020, que foi sabidamente o ano mais afetado pela pandemia. Outros possíveis fatores relacionados ao baixo desempenho dos R2 no ano de 2020 foram a transição da versão presencial para a versão *on-line* da prova e as variações no grau de dificuldade da prova entre suas diferentes versões, conforme discutido anteriormente.

Em relação à consistência interna, verificaram-se níveis adequados do coeficiente alfa de Cronbach nas provas de ginecologia em 2020 e 2021 e de obstetrícia em 2021. O alfa representa uma estimativa de confiabilidade que é explicada pela variância do teste em cada ano de aplicação. Diversos fatores podem influenciar o valor de alfa, como o número de itens de teste, a interrelação entre os itens e a dimensionalidade do construto²⁸. Na literatura, encontramos diferentes relatos sobre os valores satisfatórios de alfa que variam de 0,60 a 0,95, sendo os valores ideais entre 0,70 e 0,90^{16,28,29}. No presente estudo, optamos por considerar adequados valores de alfa acima de 0,64¹⁶ e verificamos que, para a prova de obstetrícia de 2020, o valor de alfa não alcançou a faixa esperada. Esse achado pode ser atribuído a um baixo número de perguntas, à baixa interrelação entre itens ou a construtos heterogêneos^{16,28}. Resultados semelhantes foram relatados por Dijksterhuis et al.⁵ em uma amostra de residentes de ginecologia e obstetrícia que se aplicaram ao TP na Holanda de 1999 a 2008. Os valores de alfa variaram de 0,2 a 0,81 com uma média de 0,65, que foi considerada uma confiabilidade moderada nesse estudo. Segundo os autores, a explicação mais provável para a confiabilidade encontrada

seria a grande diversidade de subdomínios abordados no TP, a qual teve como consequência um número insuficiente de itens para amostrar adequadamente o domínio completo, resultando em perda de sinal mensurável⁵. No relato de experiência aqui apresentado, é preciso considerar também as dificuldades de elaboração e validação das questões de prova durante o período da pandemia, que impossibilitou a realização de reuniões presenciais dos revisores para validação das questões.

Algumas medidas têm sido implementadas para melhorar a confiabilidade do TPI-GO, tais como o aumento do número total de questões de 100 para 120, a readequação da matriz e prova, e as reuniões presenciais de revisores para discutir a qualidade dos itens. Futuramente, pretende-se desenvolver um banco de itens com questões pré-testadas quanto ao índice discriminativo e ao grau de dificuldade, o que deverá contribuir ainda mais para o aprimoramento dessa avaliação.

No TPI-GO, os candidatos recebem um *feedback* individual e sigiloso que traz informações sobre o número e a proporção de acertos totais, a proporção de acertos nas questões de ginecologia e de obstetrícia, o desempenho relativo aos seus pares em nível nacional (em relação à mediana, ao P30 e ao P60) e a evolução do seu desempenho em relação aos anos anteriores. Os candidatos do último ano de aplicação (R3) recebem informações adicionais sobre a dispensa e bonificação na prova teórica do Tego do ano seguinte com base em seu desempenho, segundo os critérios previstos em edital. Essas informações constituem a base para a autoavaliação e avaliação formativa dos residentes, de modo a permitir a reafirmação de conhecimento prévio e a identificação de lacunas de aprendizagem, e potencializando o efeito educacional do TPI-GO.

CONCLUSÃO

A partir da experiência da Febrasgo com o TPI-GO, é possível concluir que a implementação do TP na residência é factível e aceitável, e apresenta uma série de vantagens e benefícios para os médicos residentes, para os preceptores e supervisores, para os programas e para a sociedade em geral. A implementação do teste requer preparo, envolvimento e comprometimento dos membros da equipe organizadora, além de apoio logístico e financeiro. O sigilo na divulgação dos resultados é recomendável porque preserva o candidato de uma exposição desnecessária e que lhe possa ser prejudicial. A possibilidade de bonificação e dispensa na prova teórica do Tego é um grande atrativo para a adesão e aderência. A aplicação da prova em ambiente remoto amplia o acesso com níveis aceitáveis de segurança.

O TP, quando utilizado na residência médica, oferece uma grande oportunidade para a avaliação formativa,

promovendo ganho consistente no aprendizado e retenção de conhecimento em longo prazo. Também permite acompanhar o ganho cognitivo dos residentes ao longo do treinamento.

AGRADECIMENTOS

Agradecemos aos membros da Coreme da Febrasgo, da Comissão do Título de Especialista, à secretaria da Febrasgo e aos médicos residentes, preceptores e supervisores de PRM que se aplicaram ao TP da Febrasgo.

CONTRIBUIÇÃO DOS AUTORES

Gustavo Salata Romão participou da concepção do projeto, do delineamento metodológico, da coleta e análise estatística dos dados, da interpretação dos resultados e da redação e aprovação da versão final do artigo. Marcos Felipe Silva de Sá, Cesar Eduardo Fernandes e Agnaldo Lopes da Silva Filho participaram da concepção do projeto, do delineamento metodológico, da interpretação dos resultados e da redação do manuscrito.

CONFLITO DE INTERESSES

Declaramos não haver conflito de interesses.

FINANCIAMENTO

Projeto financiado pela Febrasgo.

REFERÊNCIAS

1. Freeman A, Van Der Vleuten C, Nouns Z, Ricketts C. Progress testing internationally. *Med Teach*. 2010;32(6):451-5. doi: 10.3109/0142159x.2010.485231.
2. Dion V, St-Onge C, Bartman I, Touchie C, Pugh D. Written-based Progress Testing: a scoping review. *Acad Med*. 2022;97(5):747-57. doi: 10.1097/acm.0000000000004507.
3. Van Berkel HJM, Nuy HJP, Geerligs T. The influence of progress tests and block tests on study behaviour. *Instructional Science*. 1994;22(4):317-33. doi: 10.1007/BF00891784.
4. Ricketts C, Freeman AC, Coombes LR. Standard setting for progress tests: combining external and internal standards. *Med Educ*. 2009;43(6):589-93. doi: 10.1111/j.1365-2923.2009.03372.x.
5. Dijksterhuis MG, Scheele F, Schuwirth LW, Essed GG, Nijhuis JG, Braat DD. Progress testing in postgraduate medical education. *Med Teach*. 2009;31(10):e464-8. doi: 10.3109/01421590902849545.
6. Schuwirth LW, van der Vleuten CP. The use of progress testing. *Perspect Med Educ*. 2012;1(1):24-30. doi: 10.1007/s40037-012-0007-2.
7. Romão GS, Fernandes CE, de Sá MFS. Teste de Progresso Individual do residente em GO: relato da experiência do primeiro ano de implantação no Brasil. *Femina*. 2019;282-7.
8. Brasil. Conselho Nacional de Saúde. Resolução nº 510, de 7 de abril de 2016. *Diário Oficial da República Federativa do Brasil*; 2016.
9. Romão GS, dos Reis FJC, Cavalli RC, de Sá MFS. Matriz de competência em ginecologia e obstetrícia: um novo referencial para os programas de residência médica no Brasil. *Femina*. 2017;45(3):172-7.
10. Romão GS, de Sá MFS. Competency-based training and the competency framework in gynecology and obstetrics in Brazil. *Rev Bras Ginecol Obstet*. 2020;42(5):272-88. doi: 10.1055/s-0040-1708887.

11. de Sá MFS, Romão GS, Fernandes CE, da Silva Filho AL. The Individual Progress Test of Gynecology and Obstetrics Residents (TPI-GO): the Brazilian experience by Febrasgo. *Rev Bras Ginecol Obstet.* 2021;43(6):425-8. doi: 10.1055/s-0041-1731803.
12. Bandaranayake RC. Setting and maintaining standards in multiple choice examinations: AMEE Guide No. 37. *Med Teach.* 2008;30(9-10):836-45.
13. Wrigley W, van der Vleuten CP, Freeman A, Muijtjens A. A systemic framework for the progress test: strengths, constraints and issues: AMEE Guide No. 71. *Med Teach.* 2012;34(9):683-97. doi: 10.3109/0142159x.2012.704437.
14. Muijtjens AMM, Hoogenboom RJJ, Verwijnen GM, Van der Vleuten CPM. Relative or absolute standards in assessing medical knowledge using progress tests. *Adv Health Sci Educ.* 1998;3(2):81-7.
15. Romão GS, de Sá MFS, Fernandes CE, da Silva Filho AL. Comment on "progress testing anytime, anywhere – does a mobile-learning approach enhance the utility of a large-scale formative assessment tool?". *Med Teach.* 2021 Aug;43(8):976-977. doi: 10.1080/0142159x.2020.1854453. Epub 2020 Dec 2.
16. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. *Res Sci Educ.* 2018;48(6):1273-96.
17. Coombes L, Roberts M, Zahra D, Burr S. Twelve tips for assessment psychometrics. *Med Teach.* 2016;38(3):250-4. doi: 10.3109/0142159x.2015.1060306.
18. Scheffer M, Cassenote A, Guerra A, Guilloux AGA, Brandão APD, Miotto BA. *Demografia médica no Brasil 2020.* São Paulo: FMUSP, CFM; 2020.
19. Butler AC, Roediger HL. Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Mem Cognit.* 2008;36(3):604-16.
20. Kromann CB, Jensen ML, Ringsted C. The effect of testing on skills learning. *Med Educ.* 2009;43(1):21-7.
21. Roediger HL, Karpicke JD. The power of testing memory: Basic research and implications for educational practice. *Perspect Psychol Sci.* 2006;1(3):181-210.
22. Larsen DP. When I say... test-enhanced learning. *Med Educ.* 2013;47(10):961.
23. Wenghofer E, Klass D, Abrahamowicz M, Dauphinee D, Jacques A, Smees S, et al. Doctor scores on national qualifying examinations predict quality of care in future practice. *Med Educ.* 2009;43(12):1166-73. doi: 10.1111/j.1365-2923.2009.03534.x.
24. Görlich D, Friederichs H. Using longitudinal progress test data to determine the effect size of learning in undergraduate medical education – a retrospective, single-center, mixed model analysis of progress testing results. *Med Educ Online.* 2021;26(1):1972505. doi: 10.1080/10872981.2021.1972505.
25. Karay Y, Reiss B, Schaubert SK. Progress testing anytime and anywhere – does a mobile-learning approach enhance the utility of a large-scale formative assessment tool? *Med Teach.* 2020;42(10):1154-62. doi: 10.1080/0142159x.2020.1798910.
26. Romão GS, Schreiner L, Laranjeiras CLS, Bella ZIKJD, Coelho RA, Simões MDCR et al. Medical residency in gynecology and obstetrics in times of Covid-19: recommendations of the National Specialized Commission on Medical Residency of Febrasgo. *Rev Bras Ginecol Obstet.* 2020;42(7):411-4. doi: 10.1055/s-0040-1715147.
27. Brito LGO, Romão GS, Fernandes CE, Silva-Filho AL. Impact of Covid-19 on Brazilian medical residencies in obstetrics and gynecology. *Int J Gynaecol Obstet.* 2020;150(3):411-2. doi: 10.1002/ijgo.13283.
28. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ.* 2011;2:53-5. doi: 10.5116/ijme.4dfb.8dfd.
29. Souza AC, Alexandre NMC, Guirardello EB. Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. *Epidemiol Serv Saúde.* 2017;26:649-59.



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.