

Avaliação da utilidade clínica de novos testes diagnósticos em tuberculose: o papel dos ensaios clínicos pragmáticos*

Evaluation of the clinical utility of new diagnostic tests for tuberculosis:
the role of pragmatic clinical trials

Gisele Huf, Afrânio Kritski

Resumo

A tuberculose é uma das principais doenças infecciosas em países em desenvolvimento, e o tempo durante o qual a cadeia de transmissibilidade é mantida tem sido apontado como um dos principais fatores de perpetuação da doença. Nesse contexto, novas ferramentas para o diagnóstico têm sido aprovadas pelas agências regulamentadoras nesses países, e sua incorporação é quase imediatamente implementada pelos programas governamentais de controle da tuberculose. As intervenções na área da saúde têm sido cada vez mais investigadas em ensaios clínicos explanatórios (que avaliam seus efeitos benéficos) e em ensaios clínicos pragmáticos (que auxiliam na tomada de decisão). Nós argumentamos que a avaliação de novas técnicas diagnósticas para a detecção da tuberculose não deveria escapar dessa mesma lógica de escrutínio.

Descritores: Tuberculose/diagnóstico; Tuberculose/prevenção e controle; Ensaio clínico controlado como tópico.

Abstract

Tuberculosis is one of the major infectious diseases in developing countries, and the length of time for which the chain of transmission is maintained has been implicated as a major factor in the perpetuation of the disease. In this context, regulatory agencies in such countries have approved new diagnostic tools, which have been almost immediately incorporated into the national tuberculosis control programs. Health interventions have been increasingly investigated in clinical trials, including explanatory trials (in order to evaluate the beneficial effects of such interventions) and pragmatic trials (in order to aid in the decision-making process). We argue that the evaluation of new diagnostic techniques for the detection of tuberculosis should not escape this same logic of evaluation.

Keywords: Tuberculosis/diagnosis; Tuberculosis/prevention & control; Controlled clinical trial as topic.

Aspectos gerais sobre a avaliação de novos testes diagnósticos para diferentes agravos à saúde

Existem várias formas e etapas na avaliação do valor potencial de um teste diagnóstico para uso clínico, mas a escolha do modelo de estudo apropriado depende da questão que se deseja responder.

A primeira questão que um novo teste suscita é saber se o resultado do teste obtido em indivíduos doentes difere do resultado

encontrado em indivíduos sadios. Para responder a essa questão, o estudo deve ser conduzido em indivíduos sabidamente com a doença e em indivíduos sadios, verificando os resultados em cada um desses grupos. Esse tipo de estudo inicia a avaliação da sensibilidade, especificidade e acurácia do novo teste diagnóstico para uma determinada enfermidade. Tal estudo não

* Trabalho realizado no Instituto Nacional de Controle de Qualidade em Saúde, Fundação Oswaldo Cruz, e no Programa Acadêmico de Tuberculose, Hospital Universitário Clementino Fraga Filho/Instituto de Doenças do Tórax, Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro (RJ) Brasil.

Endereço para correspondência: Gisele Huf. Instituto Nacional de Controle de Qualidade em Saúde, Fundação Oswaldo Cruz, Avenida Brasil, 4365 Manginhos, CEP 21040-900, Rio de Janeiro, RJ, Brasil.

Tel. 55 21 3865-5141. E-mail: gisele.huf@incqs.fiocruz.br

Apoio financeiro: Nenhum.

Recebido para publicação em 28/11/2011. Aprovado, após revisão, em 13/2/2012.

representa uma ação diagnóstica, mas é uma fase inicial, eventualmente capaz de adicionar algum conhecimento ao mecanismo de ação e ao controle da enfermidade. Nessa fase, usualmente há uma participação maior de pesquisadores das áreas básica e básica aplicada, vinculadas a laboratórios de pesquisa em universidades, institutos de pesquisa e indústrias.⁽¹⁾

A próxima pergunta sobre um teste é saber se ele é capaz de distinguir indivíduos portadores da enfermidade em estudo daqueles sadios em casos suspeitos. Nessa fase, ocorre a avaliação mais tradicional em centros de pesquisa clínica, que é baseada na comparação com um teste de referência, ou padrão ouro, para derivar medidas de acurácia diagnóstica, como sensibilidade e especificidade. Esses estudos consistem na grande maioria das avaliações de novos testes diagnósticos publicadas na literatura nos últimos anos, e essas avaliações têm sido realizadas em centros de pesquisa clínica, universidades e institutos de pesquisa com o apoio da indústria, usualmente tendo as condições estruturais necessárias para responder às demandas dos órgãos regulatórios, tais como a Agência Nacional de Vigilância Sanitária (ANVISA), no Brasil, e *Food and Drug Administration* (FDA), nos EUA.

Tomar qualquer decisão na área de saúde, inclusive acerca do método diagnóstico a ser utilizado, significa pesar as evidências científicas disponíveis sobre os riscos e os benefícios de estratégias alternativas, mas depende também da confiança que essas evidências são capazes de inspirar. Esse reconhecimento gerou o surgimento de uma série de sistemas formais para categorizar a qualidade das evidências (por exemplo, de muito alta a muito baixa), entre os quais se destaca o sistema denominado *Grading of Recommendations Assessment, Development and Evaluation* (GRADE), adotado pela Organização Mundial da Saúde (OMS). O sistema GRADE foi proposto inicialmente no Canadá com o intuito de avaliar novas recomendações para o controle da tuberculose e passou a ser mais utilizado nos países desenvolvidos na última década.⁽²⁻⁴⁾

O sistema GRADE é composto por duas vertentes principais. A primeira considera o grau de confiabilidade das evidências em função dessas serem decorrentes de modelos de estudos mais capazes de evitar erros sistemáticos ou vieses. Essa hierarquia considera que os ensaios clínicos e as revisões sistemáticas de ensaios clínicos

produzem evidências mais confiáveis, seguidos por estudos observacionais (tipo coorte ou caso-controle) e, finalmente, por estudos descritivos, analíticos ou não, sem um grupo de comparação (estudos de casos) e por opiniões de especialistas. A segunda vertente é ditada pela robustez da recomendação, implicando que estudos com alta evidência científica não necessariamente engendram recomendações robustas ou fortes, e vice-versa. Pode ser que, embora uma evidência possa estar amparada por estudos randomizados de alta qualidade, um efeito adverso perturbador possa tornar uma opção de tratamento pouco conveniente para alguns; por outro lado, um efeito adverso comprovado por estudos observacionais pode inibir uma determinada prescrição quando há opções igualmente eficazes e mais seguras.

Embora as recomendações quanto à utilização de testes diagnósticos compartilhem uma mesma lógica fundamental com as recomendações para intervenções terapêuticas, esses testes apresentam características e desafios peculiares, os quais são postos em destaque quando da aplicação do sistema GRADE na área de testes e de estratégias diagnósticas. Ao avaliarmos os resultados de novos testes diagnósticos, observamos que uma baixa acurácia limita muito o valor clínico de um teste; no entanto, uma sensibilidade e/ou especificidade elevadas não garantem, por si só, uma melhora em desfechos considerados importantes para médicos e pacientes. Na prática, os clínicos desejam saber o quanto um teste é capaz de afetar o julgamento clínico ao ser utilizado na tomada de decisão sobre a conduta com o paciente. O que se deseja saber é se os indivíduos submetidos ao teste saem-se melhor, em termos dos desfechos subsequentes às intervenções diagnósticas e terapêuticas que o novo teste suscita, do que os indivíduos que não são avaliados pelo teste. Muito raramente esse benefício é absolutamente evidente na literatura; geralmente, como no caso de testes para a detecção precoce de uma doença assintomática, essa avaliação só pode ser feita de forma acurada pelo acompanhamento de indivíduos que foram randomizados para se submeter ao teste de interesse ou a outro (ou nenhum) teste.

A utilização de ensaios clínicos para a avaliação de procedimentos diagnósticos ainda está em seus primórdios. Na maior parte das vezes, não se sabe realmente qual estratégia diagnóstica é a melhor, pois apresentam vantagens e desvantagens;

entretanto, é necessário decidir se a nova estratégia diagnóstica deveria realmente substituir a atual — os estudos deveriam avaliar os efeitos dessas estratégias na tomada de decisão do médico, assim como nos desfechos importantes para o paciente e nos custos para a sociedade.

Uma das dificuldades que se apresenta é o fato de que os testes diagnósticos não afetam diretamente os desfechos importantes para os pacientes, mas sim as decisões subsequentes. A relação entre um teste diagnóstico e os desfechos para o paciente é predominantemente indireta. O efeito primário desejado sobre a saúde do paciente é raramente produzido pelo teste diagnóstico em si: o verdadeiro interesse é a via pela qual a informação obtida por meio do teste diagnóstico é capaz de melhorar a saúde, ou seja, o efeito desejado é obtido pelo papel mediador do teste diagnóstico nas decisões sobre outras intervenções. Mais comumente, a informação do teste diagnóstico é utilizada para amparar as decisões de iniciar, modificar ou interromper o tratamento. Por essa razão, algumas vezes, a comparação entre estratégias diagnósticas significa, de fato, uma comparação entre estratégias que incluem tanto o método diagnóstico quanto o tratamento que ele engendra. Nesse caso, não apenas o valor do teste diagnóstico está sendo avaliado, mas também o benefício da detecção precoce e do tratamento adequado da doença. A vantagem dos estudos randomizados consiste, portanto, na redução do viés devido a fatores externos na comparação entre os grupos. Entretanto, um aspecto também fundamental reside na capacidade dos resultados serem generalizáveis: quão diferente é nosso paciente daquele incluído nos estudos? Podemos aplicar nele os resultados obtidos naqueles estudos?

O que pode ser revelado, depois de mais de duas décadas de utilização de revisões sistemáticas e de meta-análises na área da saúde, é que, na maioria dos ensaios clínicos, os desfechos avaliados não respondiam a questões chave para a tomada de decisão sobre a incorporação ou não da tecnologia ao sistema de saúde, pelo fato de os estudos terem sido realizados em populações específicas e pouco representativas da população geral. Além disso, tais estudos, em geral, não incluíam avaliações de custo-efetividade.⁽⁵⁾ Ao final da década de 90, a distinção dos ensaios clínicos entre explanatórios e pragmáticos começou a ganhar evidência.⁽⁶⁾ Ensaios clínicos explanatórios,

realizados em centros de pesquisa clínica, procuram responder questões sobre a eficácia, ou seja, “se” e “como” uma intervenção funciona; ensaios clínicos pragmáticos são realizados para amparar uma tomada de decisão na área de saúde, e, para isso, são realizados em condições muito próximas às da rotina dos serviços e em pacientes muito semelhantes aos que irão necessitar do tratamento no futuro. A realização de ensaios pragmáticos não tem recebido a atenção da academia, das agências governamentais de fomento e, menos ainda, do parque industrial, pois seu interesse é que o produto seja registrado no órgão regulatório e comercializado. Em uma recente revisão sistemática que incluiu 168.000 ensaios clínicos randomizados realizados no período entre 1976 e 2002, observou-se que apenas 95 (0,05%) deles preenchiam os critérios de classificação para ensaios clínicos pragmáticos. Os autores enfatizaram a necessidade urgente de se priorizar a realização de ensaios pragmáticos que possam responder também questões relativas à aplicabilidade da nova tecnologia no sistema de saúde, e não apenas a questões de eficácia utilizadas nos ensaios clínicos explanatórios, cuja finalidade maior seria a obtenção do registro do produto nos órgãos regulatórios.⁽⁷⁾

A situação não é diferente na área de testes diagnósticos. Novos testes diagnósticos, avaliados com o financiamento da indústria em centros de pesquisas clínicas por meio de testes de acurácia, têm sido suficientes para a aprovação de seu registro para a comercialização nos órgãos regulatórios (FDA, nos EUA; *European Agency for the Evaluation of Medicinal Products*, na Europa; e ANVISA, no Brasil). Tais testes são implementados no sistema privado tão logo estejam disponíveis no mercado, com base em experiências com um número limitado de casos, isto é, com uma expectativa subjetiva de sua utilidade. Como resultado do lobby das indústrias e da sociedade civil (induzida pelo marketing) no sistema econômico vigente, a lógica tem sido a de que devem ser oferecidas ao indivíduo que procura o sistema de saúde todas as inovações tecnológicas produzidas com alguma evidência científica, mesmo que não haja nenhuma avaliação sistematizada de seu impacto no sistema de saúde.

Avaliação e incorporação de novas tecnologias diagnósticas em tuberculose no Brasil e no mundo

A tuberculose é uma das principais doenças infecciosas nos países em desenvolvimento e é uma importante causa de morbidade e mortalidade. A transmissão da tuberculose tem sido apontada como um dos principais fatores de perpetuação da doença no planeta, estando associada a certos fatores, como a desigualdade social, o advento da AIDS e o envelhecimento da população.⁽⁸⁾

A OMS estima que, globalmente, houve um número incidente de 9,4 milhões de casos em 2009, o que significa um aumento quando se considera a taxa de incidência de 8,3 milhões em 2000 e de 6,6 milhões de casos em 1990.⁽⁸⁾ Ainda que o número total de casos incidentes de tuberculose tenha aumentado em termos absolutos como resultado do crescimento populacional, o número de casos per capita está em decréscimo. No entanto, a taxa de declínio ainda é considerada lenta, na razão de menos de 1% de declínio ao ano.⁽⁸⁾ Em 2009, havia cerca de 440.000 novos casos de tuberculose multirresistente (TBMR) e, entre eles, 150.000 evoluíram para o óbito. Além disso, houve 1.400.000 mortes entre pacientes HIV negativos e 380.000 mortes adicionais entre aqueles HIV positivos, resultando em 4.700 casos de óbito por dia, aproximadamente.

Um dos fatores primordiais para o surgimento de novos casos é o tempo durante o qual o paciente portador de tuberculose pulmonar permanece sem diagnóstico e tratamento, mantendo a cadeia de transmissibilidade.⁽⁹⁾ O rendimento diagnóstico da baciloscopia (BAAR) é baixo (40-60%), principalmente em pacientes com baixa carga bacilar nas amostras respiratórias, como aqueles infectados por HIV ou portadores de outras doenças imunossupressoras (30%).⁽¹⁰⁾ Por trinta anos, a pesquisa deixou de ser priorizada pela OMS, mas, em 2006, por meio do *Global Plan to Stop TB* para o controle da tuberculose, a OMS passou a recomendar novamente a realização de pesquisas para tuberculose nas áreas de desenvolvimento e de avaliação de novos métodos diagnósticos, medicamentos, vacinas e estratégias de gestão.⁽¹¹⁾ Nesse plano, o aumento do diagnóstico dos casos de tuberculose paucibacilar, de coinfeção tuberculose/HIV e de TBMR passou, portanto, a ser essencial para o sucesso do programa, pois o diagnóstico rápido da tuberculose permite o

início do tratamento medicamentoso e diminui o tempo de transmissão da doença, com consequente diminuição do número de pessoas infectadas pelo indivíduo doente.⁽¹²⁾

Apesar de a detecção de BAAR no escarro possuir baixa sensibilidade, este permanece como um dos exames mais utilizados para o diagnóstico de tuberculose pulmonar sensível às drogas em países em desenvolvimento.⁽⁸⁾ Entretanto, na maioria desses países, na prática, a cultura para micobactérias, cuja sensibilidade é maior (80-85%) que a da baciloscopia no diagnóstico de tuberculose pulmonar sensível às drogas, é realizada em meio sólido Löwenstein-Jensen e está indicada apenas em casos clínicos selecionados (casos de falência ao tratamento, pacientes com baciloscopia negativa persistente ou formas extrapulmonares). O maior problema da cultura em Löwenstein-Jensen é o longo tempo de incubação (4-6 semanas), sendo que o teste de sensibilidade é realizado a partir da cultura, e não do espécime clínico, o que requer várias semanas adicionais para a obtenção dos resultados.⁽¹³⁾

Em pacientes HIV positivos e em crianças, tem sido inadequada a estratégia proposta pela OMS de priorizar a avaliação de pacientes sintomáticos respiratórios (tosse por mais de 2-3 semanas) para a busca de casos de tuberculose pulmonar. Recentemente, um grupo de autores identificou 267 casos de tuberculose em 1.748 pacientes HIV positivos com suspeita de tuberculose e observou que a tosse com duração de 2-3 semanas não auxiliava no diagnóstico. A presença de tosse e febre de qualquer duração e de sudorese por mais de 3 semanas apresentou uma sensibilidade de 93% e uma especificidade de 36%.⁽¹⁴⁾ Em um artigo de revisão, um grupo de autores propôs o desenvolvimento e a avaliação de escores clínicos e clinicorradiológicos no diagnóstico de tuberculose pulmonar em adultos e crianças em diferentes cenários epidemiológicos.⁽¹⁵⁾ Isso evidencia a urgente necessidade de avaliação de novas abordagens diagnósticas que promovam um maior impacto nas regiões mais afetadas pela coinfeção tuberculose/HIV e pela TBMR.⁽¹⁶⁾ Entretanto, na avaliação de novas tecnologias diagnósticas, é necessário considerar diferentes estratégias de detecção que incluam também a análise dos fatores associados ao acesso ao sistema de saúde, que estão vinculados ao paciente e/ou ao sistema de saúde.⁽¹⁷⁾

A OMS tem proposto testes que aumentem a sensibilidade da baciloscopia. Entretanto, na prática, até 2007, com o uso universal apenas da baciloscopia de escarro, 30-40% dos pacientes atendidos nos serviços de saúde em países em desenvolvimento eram tratados para tuberculose sem a confirmação bacteriológica. Naquele mesmo ano, no intuito de responder de modo mais efetivo à emergência da coinfeção tuberculose/HIV e da TBMR em nível global, a OMS recomendou o uso de cultura líquida para a detecção de *Mycobacterium tuberculosis* e de testes de sensibilidade aos fármacos, baseando-se em uma revisão das evidências científicas disponíveis e em uma consulta a um painel de especialistas.⁽¹⁸⁻²⁰⁾

Em 2008, a OMS recomendou o uso de testes moleculares para a triagem rápida de pacientes suspeitos de TBMR. Essa recomendação foi baseada em revisões sistemáticas, na opinião de especialistas e em resultados preliminares de efetividade obtidos em projetos demonstrativos (de fase III/IV) realizados em centros de pesquisa clínica. Tais testes deveriam ser utilizados apenas em amostras respiratórias com baciloscopia positiva ou com cultura positiva para micobactérias.^(21,22)

Em uma revisão sistemática de estudos que avaliaram novos testes de diagnóstico para tuberculose, foi demonstrada a falta de rigor metodológico na maioria dos estudos.⁽²³⁾ Os autores enfatizaram que resultados enviesados oriundos de estudos mal delineados poderiam levar à adoção de novos testes diagnósticos precocemente, propiciando pouco ou nenhum benefício.⁽²³⁾

Nos últimos anos, foram publicadas orientações na padronização de modelos de estudos na área de doenças infecciosas que, além de avaliar a acurácia do novo teste diagnóstico, deveriam também incluir avaliações de diferentes algoritmos (não apenas de testes individuais), assim como suas contribuições relativas ao sistema de cuidados de saúde, valor incremental, impacto na prática clínica sobre as escolhas de tomada de decisão na terapêutica, estudos de custo-efetividade em condições de rotina e impacto dos resultados das novas ferramentas para o paciente e para a sociedade.⁽²⁴⁻²⁶⁾

Em 2008, um grupo de autores, ao analisar dados publicados na literatura, concluiu que as novas recomendações incluídas nos manuais de normas ou nas diretrizes em tuberculose de países desenvolvidos, utilizando-se do melhor grau de

evidência científica (GRADE), eram rapidamente incorporadas nas revisões de manuais ou nas diretrizes clínicas de países em desenvolvimento com poucas mudanças, sem a sua adequação às necessidades da área de saúde de cada país⁽²⁷⁾; além disso, outro grupo de autores relatou que a maioria das diretrizes clínicas realizadas naqueles países não apresentava certificação pelo *Advancing Guideline Development, Reporting and Evaluation in Health Care* (AGREE).⁽²⁸⁾ Em 2009, por ocasião da publicação das diretrizes da Sociedade Brasileira de Pneumologia e Tisiologia, foram incluídas 24 recomendações sobre tuberculose com base no nível das evidências científicas disponíveis. Na área diagnóstica, apesar da inexistência de ensaios clínicos randomizados que confirmassem a relação custo-efetividade e o impacto clínico de novos testes, aquelas diretrizes seguiram a proposta da OMS (uso de cultura líquida em sistemas automatizados não radiométricos) e do FDA (uso de testes moleculares comercializados) em amostras respiratórias paucibacilares na investigação diagnóstica de tuberculose pulmonar.⁽²⁹⁾

Recentemente, em um inquérito realizado em 16 países com alta carga de tuberculose sobre a adoção de sete novas ferramentas para o diagnóstico de tuberculose aprovadas pela OMS desde 2007, foi confirmada a implementação de novos testes diagnósticos nas políticas de controle de tuberculose, pois aproximadamente metade dos países investigados adotou tais ferramentas.⁽³⁰⁾ Curiosamente, em nenhum dos sete países foi realizada uma avaliação do impacto da incorporação daquelas novas tecnologias, tal como proposto pela OMS e pelo *Stop TB Initiative*.⁽³¹⁾

Na avaliação de novos testes diagnósticos em tuberculose e TBMR, testes que utilizam cultura líquida – por exemplo, Bactec 960 (Bactec) – ou testes moleculares – EMTD (Gen-Probe), Amplicor (Roche), TB Test (Biometrix), MTBDRplus (Hain Life Sciences) e, mais recentemente, GeneXpert (Cepheid) – já foram recomendados pelo FDA e pelos órgãos correspondentes na União Europeia (e comercializados naqueles países). Esses testes também têm sido comercializados na área privada em países de nível intermediário de desenvolvimento econômico, como no Brasil. Apesar de não haver relatos na literatura sobre ensaios clínicos pragmáticos e sobre a relação custo-efetividade quanto ao uso desses testes na abordagem diagnóstica de tuberculose ou

de TBMR em países em desenvolvimento, o teste Xpert™ MTB/RIF foi recomendado pela OMS em dezembro de 2010.⁽³²⁾ Esse é um teste molecular totalmente automatizado, com um modelo de processamento integrado e concebido para purificar, concentrar, amplificar e identificar sequências alvo do gene *rpoB* para o diagnóstico de resistência à rifampicina. O teste fornece resultados em 120 min a partir de amostras de escarro, sem a necessidade da presença de um especialista em biologia molecular. Os resultados obtidos nos estudos de demonstração (fase III) confirmam a elevada especificidade para o diagnóstico de tuberculose e de tuberculose resistente à rifampicina. A sensibilidade do teste foi de 72% em amostras de escarro de pacientes com baciloscopia negativa, semelhante à de outros testes moleculares, como Amplicor e EMTD. Entretanto, esse teste pode ser descentralizado para as unidades de saúde de nível secundário, pois não necessita de um laboratório de biologia molecular para a sua execução.^(33,34)

Em nível global, nos últimos anos, tornou-se consenso entre os formuladores de políticas públicas que países emergentes, como o Brasil, deveriam liderar os processos na área de avaliação de tecnologia em saúde e na análise do impacto da incorporação de novas tecnologias, enfocando a habilidade da nova tecnologia em melhorar ou manter um estado de saúde. Os testes diagnósticos não deveriam escapar desse princípio.

Nos últimos anos, entre os países emergentes, apenas no Brasil foi priorizada a avaliação de tecnologia em saúde, por meio da criação de uma comissão especial para a incorporação de tecnologia do Ministério da Saúde, que estabeleceu normas para o setor.⁽³⁵⁾ Novas tecnologias poderão ser incorporadas ao sistema de saúde público ou privado somente se os estudos forem conduzidos em condições de campo, em diferentes regiões do país; tenham utilizado o desenho de estudo mais apropriado; tenham incluído uma avaliação de impacto para o sistema de saúde; e possam fornecer subsídios para amparar uma política de saúde.

A partir de 2009, foi iniciada uma parceria entre as faculdades de medicina da Universidade Federal do Rio de Janeiro, Universidade de São Paulo, Universidade Federal de Minas Gerais e da Universidade Federal do Rio Grande do Sul, em conjunto com o Instituto de Comunicação e Informação Científica e Tecnológica em Saúde

da Fundação Oswaldo Cruz e a Secretaria de Assistência a Saúde do Ministério da Saúde, denominada PROQUALIS, tendo como metas difundir informações e tecnologias para os serviços e profissionais de saúde e servir de apoio para as temáticas da qualidade do cuidado de saúde por meio de uma rede colaborativa entre as universidades públicas, universidades privadas e os Ministérios da Saúde, da Educação e da Ciência e Tecnologia.⁽³⁶⁾ O PROQUALIS também deveria subsidiar a prática clínica com diretrizes e literatura técnico-científica atual e relevante para os profissionais brasileiros. A informação deveria ser selecionada por uma rede nacional de especialistas. Para surpresa geral, à medida que as discussões e as trocas de informações foram ocorrendo, foram percebidos os seguintes fatos: os técnicos que participam das avaliações de tecnologias em saúde não interagem com os técnicos que elaboram ou modificam normas, manuais e diretrizes clínicas (exceto na Inglaterra, Espanha, Canadá, Holanda e Nova Zelândia); a maioria dos dirigentes das sociedades de classe (médicas ou não) no Brasil desconhece os critérios utilizados internacionalmente para a certificação da qualidade de diretrizes (AGREE) e tem diferentes visões acerca da elaboração e do monitoramento de mudanças em diretrizes clínicas.

Como qualquer mudança nas diretrizes clínicas ou nos manuais de normas constituiria uma incorporação tecnológica que poderia ser acrítica e maléfica para o indivíduo e/ou para o sistema de saúde, passou a ser prioritária uma aproximação entre pesquisadores atuantes na área de avaliação de tecnologias em saúde e aqueles profissionais que elaboram normas, manuais e diretrizes, principalmente aqueles vinculados à área de ensino das escolas biomédicas, tendo em vista modificações curriculares em médio e longo prazo.

Em 2009, pesquisadores da Rede Brasileira de Pesquisa em Tuberculose (Rede TB) exerceram atuações relevantes na elaboração da nova versão do Manual de Normas de Controle de Tuberculose do Ministério da Saúde. Foi aceita a proposta da Rede TB de se considerar qualquer modificação no manual de normas como uma incorporação tecnológica. Portanto, qualquer modificação no manual requereria uma avaliação de seu impacto no sistema nacional de saúde. Considerou-se estratégico aproximar os profissionais que desenvolvem diretrizes com aqueles que avaliam tecnologias em saúde. Em 2009-2010, foram

realizadas inúmeras atividades conjuntas entre técnicos da Secretaria de Ciência e Tecnologia e da Secretaria de Vigilância em Saúde do Ministério da Saúde, pesquisadores do Centro de Referência Professor Hélio Fraga, do Instituto de Pesquisa Clínica da Fundação Oswaldo Cruz, do Programa Acadêmico de Tuberculose da Universidade Federal do Rio de Janeiro e da Universidade Estadual do Rio de Janeiro, pesquisadores da Rede TB, assim como representantes da Fundação Melinda Gates, da *International Union Against Tuberculosis and Lung Disease* e da *United States Agency for International Development*. Essas atividades multidisciplinares propiciaram o desenvolvimento de dois projetos de envergadura nacional para avaliar a incorporação de novos testes diagnósticos para tuberculose e TBMR já comercializados no país e disponíveis no sistema privado (GeneXpert e MTBDRplus, respectivamente). Esses projetos de pesquisa operacional priorizaram a condução de ensaios clínicos randomizados pragmáticos, incluíram a avaliação de custos para o paciente e para o sistema de saúde, bem como a análise de *scale-up*. Em 2010, iniciou-se o auxílio financeiro do Ministério da Saúde e, a partir de 2011, ocorreu o repasse financeiro internacional. A plataforma acordada entre as instituições participantes tem como maior objetivo responder às demandas da Comissão de Incorporação Tecnológica do Ministério da Saúde na expectativa de identificar em quais cenários esses testes diagnósticos podem ser incorporados ao sistema público.

Em paralelo, no intuito de auxiliar a discussão sobre esse tema, a *Union International Union Against Tuberculosis and Lung Disease* e seus parceiros regionais, como a Rede TB, publicaram recentemente uma nova proposta de plataforma para avaliar o impacto de novas tecnologias diagnósticas em tuberculose, que será utilizada pela OMS a partir de 2011.^(37,38)

Por meio dessa nova plataforma de avaliação tecnológica, além de se priorizar a pesquisa operacional por meio de ensaios clínicos randomizados pragmáticos, análise de custos e *scale-up*, serão incluídos aspectos relativos à equidade e ao acesso, assim como estudos qualitativos com usuários, profissionais de saúde, gestores e representantes de indústrias locais ou internacionais que identifiquem barreiras ou facilidades para a incorporação da nova tecnologia nos diferentes sistemas de saúde. Nessa plataforma, pretende-se que a discussão

sobre a incorporação ou a disseminação de novos métodos diagnósticos em tuberculose responda pelo menos uma simples pergunta: isto será melhor para os pacientes e/ou para o sistema de saúde vigente no país? Além disso, as novas técnicas são cada vez mais comparadas em relação ao montante de recursos que utilizam. A saúde em si não é mais o único critério no qual as decisões no setor são baseadas; é necessário também levar em conta o custo social, o desprazer pessoal e o gasto de tempo do usuário e/ou do profissional de saúde.

Como mencionado anteriormente, novos instrumentos diagnósticos são muitas vezes imediatamente incorporados à rotina dos serviços tão logo sejam aprovados para a comercialização pelas agências regulamentadoras, com base em sua performance por meio da análise da sensibilidade, especificidade ou curvas ROC, em um frenesi que, se por um lado não se justifica pelo fato de que tal instrumento possa significar qualquer avanço significativo do ponto de vista dos pacientes, certamente agrada aos anseios de inovação dos pesquisadores clínicos e de sua comercialização pelos produtores ou laboratórios.

Uma das críticas aos ensaios clínicos pragmáticos é a de que eles priorizam as questões relativas à aplicabilidade e à capacidade de generalização dos resultados em detrimento da validade interna do estudo. Entretanto, é inevitável que, ao introduzirmos a perspectiva dos pacientes nas avaliações, tornemos as medidas de desfecho do estudo mais subjetivas e, portanto, mais sujeitas a vieses. Contudo, é preciso reconhecer que independentemente do que estejamos avaliando, os resultados serão sempre coloridos pelas nossas percepções; nossas observações provavelmente falam mais sobre a forma como definimos um problema e categorizamos os possíveis desfechos do que qualquer fenômeno real que possamos observar. Essa abordagem pragmática traz como consequência o fato de que não é possível conduzir uma investigação como essa em um ambiente “puramente experimental” – se é que isso seja desejável em qualquer momento. A pesquisa básica aplicada, a pesquisa clínica, a pesquisa operacional e a prática clínica tornam-se processos entrelaçados, e devemos sempre ter como principal desfecho a saúde do paciente e a maior efetividade das ações do sistema de saúde nas áreas em que a nova tecnologia será incorporada.

Referências

- Sackett DL, Haynes RB. The architecture of diagnostic research. *BMJ*. 2002;324(7336):539-41. PMID:11872558. PMCID:1122451. <http://dx.doi.org/10.1136/bmj.324.7336.539>
- Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328(7454):1490. PMID:15205295. PMCID:428525. <http://dx.doi.org/10.1136/bmj.328.7454.1490>
- Schünemann HJ, Jaeschke R, Cook DJ, Bria WF, El-Solh AA, Ernst A, et al. An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. *Am J Respir Crit Care Med*. 2006;174(5):605-14. PMID:16931644. <http://dx.doi.org/10.1164/rccm.200602-1975T>
- The periodic health examination. Canadian Task Force on the Periodic Health Examination. *Can Med Assoc J*. 1979;121(9):1193-254. PMID:115569. PMCID:1704686.
- Laupacis A, Straus S. Systematic reviews: time to address clinical and policy relevance as well as methodological rigor. *Ann Intern Med*. 2007;147(4):273-4. PMID:17638716.
- Coutinho Eda S, Huf G, Bloch KV. Pragmatic clinical trials: an option in the construction of health-related evidence [Article in Portuguese]. *Cad Saude Publica*. 2003;19(4):1189-93. PMID:12973582.
- Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. *Trials*. 2009;10:37. PMID:19493350. PMCID:2700087. <http://dx.doi.org/10.1186/1745-6215-10-37>
- World Health Organization. Global tuberculosis control: epidemiology, strategy, financing. Geneva: World Health Organization; 2009.
- Dye C, Williams BG, Espinal MA, Raviglione MC. Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis. *Science*. 2002;295(5562):2042-6. PMID:11896268. <http://dx.doi.org/10.1126/science.1063814>
- Sester M, Giehl C, McNerney R, Kampmann B, Walzl G, Cuchi P, et al. Challenges and perspectives for improved management of HIV/ Mycobacterium tuberculosis co-infection. *Eur Respir J*. 2010;36(6):1242-7. PMID:21119204. <http://dx.doi.org/10.1183/09031936.00040910>
- World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Jun 16]. The Stop TB Strategy - Building on and enhancing DOTS to meet the TB-related Millennium Development Goals. [Adobe Acrobat document, 24p.]. Available from: http://whqlibdoc.who.int/hq/2006/WHO_HTM_STB_2006.368_eng.pdf
- Sherman LF, Fujiwara PI, Cook SV, Bazerman LB, Frieden TR. Patient and health care system delays in the diagnosis and treatment of tuberculosis. *Int J Tuberc Lung Dis*. 1999;3(12):1088-95. PMID:10599012.
- Fair E, Hopewell PC, Pai M. International Standards for Tuberculosis Care: revisiting the cornerstones of tuberculosis care and control. *Expert Rev Anti Infect Ther*. 2007;5(1):61-5. PMID:17266454. <http://dx.doi.org/10.1586/14787210.5.1.61>
- Fair E, Hopewell PC, Pai M. International Standards for Tuberculosis Care: revisiting the cornerstones of tuberculosis care and control. *Expert Rev Anti Infect Ther*. 2007;5(1):61-5. PMID:17266454. <http://dx.doi.org/10.1586/14787210.5.1.61>
- Marais BJ, Raviglione MC, Donald PR, Harries AD, Kritski AL, Graham SM, et al. Scale-up of services and research priorities for diagnosis, management, and control of tuberculosis: a call to action. *Lancet*. 2010;375(9732):2179-91. [http://dx.doi.org/10.1016/S0140-6736\(10\)60554-5](http://dx.doi.org/10.1016/S0140-6736(10)60554-5)
- Lawn SD, Ayles H, Egwaga S, Williams B, Mukadi YD, Santos Filho ED, et al. Potential utility of empirical tuberculosis treatment for HIV-infected patients with advanced immunodeficiency in high TB-HIV burden settings. *Int J Tuberc Lung Dis*. 2011;15(3):287-95. PMID:21333094.
- Sreeramareddy CT, Panduru KV, Menten J, Van den Ende J. Time delays in diagnosis of pulmonary tuberculosis: a systematic review of literature. *BMC Infect Dis*. 2009;9:91. PMID:19519917. PMCID:2702369. <http://dx.doi.org/10.1186/1471-2334-9-91>
- World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Jun 16]. TB diagnostics and laboratory strengthening - WHO policy - The use of liquid medium for culture and DST, 2007. Available from: http://www.who.int/tb/laboratory/policy_liquid_medium_for_culture_dst/en/index.html
- Dinnes J, Deeks J, Kunst H, Gibson A, Cummins E, Waugh N, et al. A systematic review of rapid diagnostic tests for the detection of tuberculosis infection. *Health Technol Assess*. 2007;11(3):1-196.
- Cruciani M, Scarparo C, Malena M, Bosco O, Serpelloni G, Mengoli C. Meta-analysis of BACTEC MGIT 960 and BACTEC 460 TB, with or without solid media, for detection of mycobacteria. *J Clin Microbiol*. 2004;42(5):2321-5. PMID:15131224. PMCID:404614. <http://dx.doi.org/10.1128/JCM.42.5.2321-2325.2004>
- World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Jun 16]. Tuberculosis (TB) - Moving research findings into new WHO policies. Available from: http://www.who.int/tb/advisory_bodies/research_to_policy/en/index.html
- World Health Organization. The Global MDR-TB & XDR-TB Response Plan 2007-2008. Geneva: World Health Organization; 2007.
- Pai M, Ramsay A, O'Brien R. Comprehensive new resource for evidence-based TB diagnosis. *Expert Rev Mol Diagn*. 2009;9(7):637-9. PMID:19817547. <http://dx.doi.org/10.1586/erm.09.48>
- Peeling RW, Smith PG, Bossuyt PM. A guide for diagnostic evaluations. *Nat Rev Microbiol*. 2006;4(12 Suppl):S2-6. <http://dx.doi.org/10.1038/nrmicro1568>
- World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Mar 20]. New Technologies for Tuberculosis Control: A framework for their adoption, introduction and implementation. [Adobe Acrobat document, 120p.]; Available from: http://whqlibdoc.who.int/publications/2007/9789241595520_eng.pdf
- World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Jun 16]. Tuberculosis (TB) - TB diagnostics and laboratory strengthening. Available from: <http://www.who.int/tb/dots/laboratory/policy/en/index4.html>
- Wei X, Walley JD, Liang X, Liu F, Zhang X, Li R. Adapting a generic tuberculosis control operational guideline and scaling it up in China: a qualitative case study. *BMC Public Health*. 2008;8:260. PMID:18662410. PMCID:2515317. <http://dx.doi.org/10.1186/1471-2458-8-260>

28. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ*. 2010;182(18):E839-42. PMID:20603348. PMCID:3001530. <http://dx.doi.org/10.1503/cmaj.090449>
29. Conde MB, Melo FA, Marques AM, Cardoso NC, Pinheiro VG, Dalcin Pde T, et al. III Brazilian Thoracic Association Guidelines on tuberculosis. *J Bras Pneumol*. 2009;35(10):1018-48. PMID:19918635.
30. van Kampen SC, Ramsay AR, Anthony RM, Klatser PR. Retooling national TB control programmes (NTPs) with new diagnostics: the NTP perspective. *PLoS One*. 2010;5(7):e11649. PMID:20661463. PMCID:2908286. <http://dx.doi.org/10.1371/journal.pone.0011649>
31. World Health Organization. New Laboratory Diagnostic Tools for Tuberculosis. Geneva: World Health Organization; 2008.
32. World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Jun 16]. Tuberculosis (TB) - Strategic and Technical Advisory Group for Tuberculosis (STAG-TB). Available from: http://www.who.int/tb/advisory_bodies/stag/en/index.html
33. Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med*. 2010;363(11):1005-15. PMID:20825313. PMCID:2947799. <http://dx.doi.org/10.1056/NEJMoa0907847>
34. Van Rie A, Page-Shipp L, Scott L, Sanne I, Stevens W. Xpert(®) MTB/RIF for point-of-care diagnosis of TB in high-HIV burden, resource-limited countries: hype or hope? *Expert Rev Mol Diagn*. 2010;10(7):937-46. PMID:20964612. <http://dx.doi.org/10.1586/erm.10.67>
35. Brasil. Portaria n. 2.587, de 30 de outubro de 2008. Dispõe sobre a Comissão de Incorporação de Tecnologias do Ministério da Saúde e vincula sua gestão à Secretaria de Ciência, Tecnologia e Insumos Estratégicos. *Diário Oficial da União, Brasília*, nº 212, p 94, 31 de out. de 2008. Seção 1.
36. Proqualis - Aprimorando as práticas de saúde [homepage on the Internet]. Rio de Janeiro: Fundação Oswaldo Cruz, Ministério da Saúde, c2009 [cited 2008 Jun 16]. Available from: www.proqualis.net
37. Mann G, Squire SB, Bissell K, Eliseev P, Du Toit E, Hesseling A, et al. Beyond accuracy: creating a comprehensive evidence base for TB diagnostic tools. *Int J Tuberc Lung Dis*. 2010;14(12):1518-24. PMID:21144235.
38. Squire SB, Ramsay AR, van den Hof S, Millington KA, Langley I, Bello G, et al. Making innovations accessible to the poor through implementation research. *Int J Tuberc Lung Dis*. 2011;15(7):862-70. PMID:21682960. <http://dx.doi.org/10.5588/ijtld.11.0161>

Sobre os autores

Gisele Huf

Pesquisadora. Instituto Nacional de Controle de Qualidade em Saúde, Fundação Oswaldo Cruz, Rio de Janeiro (RJ) Brasil.

Afrânio Kritski

Vice-Diretor. Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro (RJ) Brasil.