

Tendências em validação de métodos de ensaios qualitativos

Trends in implementing the validation of qualitative methods of analysis

RIALA6/1396

Carina de Souza GONDIM, Roberto Gonçalves JUNQUEIRA, Scheilla Vitorino Carvalho de SOUZA*

* Endereço para correspondência: Laboratório de Bromatologia/Unidade de Pesquisa Análise de Alimentos, Departamento de Alimentos - ALM, Faculdade de Farmácia - FAFAR, Universidade Federal de Minas Gerais - UFMG, Av. Antônio Carlos 6627, Campus da UFMG, Pampulha, CEP 31.270-010, Belo Horizonte, MG, Brasil. E-mail scheilla@bromatologiaufmg.com.br.
Recebido: 28.04.2011 - Aceito para publicação: 28.06.2011

RESUMO

A crescente importância das análises qualitativas e a necessidade de confiabilidade dos resultados analíticos para subsidiar as tomadas de decisões nos diversos setores da cadeia produtiva de alimentos são inquestionáveis, acarretando impactos na saúde pública, na economia e nos direitos do consumidor. Contudo, a validação desses métodos tem sido um ponto crítico na implantação de sistemas de gestão da qualidade e nos processos de acreditação de laboratórios. Apesar da existência de protocolos bem estabelecidos para validação de métodos quantitativos, ainda existe uma lacuna no desenvolvimento de abordagens para a implementação da metrologia em análises qualitativas. Neste contexto, o presente trabalho apresenta uma discussão sobre o tema validação de métodos qualitativos, abrangendo definições, delineamento experimental e análise de dados para avaliação dos parâmetros de desempenho aplicáveis como taxas de falsos resultados, taxa de confiabilidade, taxa de seletividade, taxa de sensibilidade, limite de detecção, região de perda de confiabilidade, acórdância, concordância e robustez.

Palavras-chave. validação intralaboratorial, método de ensaio qualitativo, parâmetros de desempenho.

ABSTRACT

The growing importance of qualitative analysis and the need for reliable analytical results in order to support the decision making at various sectors of food chain are unquestionable, with impacts on the economy, public health and consumer rights. However, the validation of these methods has been a critical point in the quality management systems implementation, and also in the processes for laboratory accreditation. In spite of the existence of well-established protocols for the validation of quantitative methods, there is a gap in the development of approaches for the implementation of metrology in the qualitative analysis. In this context, this paper deals with the validation of qualitative methods, including definitions, experimental design, and analysis of results for evaluating the applicable performance parameters, such as false response rates, reliability, selectivity, sensitivity, detection limit, region of unreliability, accordance, concordance and robustness.

Keywords. in-house validation, qualitative methods, performance parameters.

INTRODUÇÃO

A área de análise de alimentos é norteadada por uma crescente demanda de métodos de ensaio. Os constantes avanços tecnológicos na produção, processamento, embalagem e armazenamento de alimentos, trazem para o cotidiano dos laboratórios questões como análises de aditivos, migrantes de embalagens, contaminantes, resíduos de agrotóxico e medicamentos veterinários, adulterantes, indicadores de origem geográfica e organismos geneticamente modificados¹.

Além dos desafios decorrentes da diversidade e contínua evolução do setor de alimentos, os laboratórios especializados na análise desses produtos, necessitam demonstrar competência técnica para os ensaios que realizam², assegurando confiabilidade e comparabilidade dos resultados, os quais, na maioria das vezes, subsidiam as tomadas de decisão relativas aos aspectos econômicos, de saúde pública³ e de defesa do consumidor.

A importância das informações qualitativas vem crescendo entre laboratórios⁴ e, em diversos setores da cadeia produtiva de alimentos, métodos qualitativos têm sido frequentemente selecionados para triagens, por fornecerem resultados rápidos, objetivos, com baixo custo, simplicidade e minimização de erros, devido ao menor intervalo entre amostragem e análise^{5,6}.

No contexto de implementação de sistemas de gestão da qualidade e acreditação², a validação de métodos qualitativos tem representado um importante gargalo para o reconhecimento da competência de laboratórios de análise de alimentos^{7,8}. Apesar de haver protocolos estabelecidos sobre validação de métodos quantitativos, alguns autores apontam que pouca atenção tem sido dada para os métodos qualitativos^{6,8,9}.

Os principais documentos orientativos relacionados à validação de métodos, tanto nacionais¹⁰ quanto internacionais¹¹⁻¹³, não estão harmonizados e não detalham os principais parâmetros de desempenho e experimentos aplicáveis à validação de métodos qualitativos, o que tem gerado grande dificuldade na estruturação de procedimentos de validação adequados para esse tipo de método.

Em muitos casos, a aplicação dos princípios metroológicos tradicionais não é apropriada aos métodos qualitativos, o que denota a necessidade do desenvolvimento e da implementação de novas estratégias de garantia de confiabilidade adequadas a tais métodos^{14,15}. Embora muitos conceitos estejam

implícitos nas atividades da química analítica qualitativa desde seus primórdios, eles somente surgiram na literatura especializada recentemente, impulsionados pela organização e regulamentação metroológica internacional¹⁶.

Diante do exposto, consideramos neste trabalho os recentes desenvolvimentos conceituais e práticos relacionados à metrologia e validação de métodos qualitativos, visando fornecer uma referência no campo das ciências analíticas que possa ser utilizada para a construção do conhecimento que atenderá à atual demanda por respostas binárias confiáveis.

MÉTODOS DE ENSAIO QUANTITATIVOS E QUALITATIVOS

Os métodos de ensaio podem ser classificados em métodos quantitativos e qualitativos. Métodos quantitativos são definidos como aqueles que determinam a quantidade de determinado analito em uma amostra em valores numéricos de uma unidade de medida apropriada^{12,17}. Enquanto os métodos qualitativos são aqueles que permitem a identificação do analito com base em suas propriedades físicas, químicas ou biológicas¹². Nas análises qualitativas, a resposta binária pode vir por meio de um instrumento de medida, de *kits* de testes, que envolvem mudanças sensoriais, ou outros⁵.

Nas últimas décadas, as análises qualitativas foram consideradas em declínio no campo da ciência analítica por estarem relacionadas com identificações que utilizavam reações químicas tendo os sentidos humanos como detectores. O fato da análise instrumental também fornecer uma resposta qualitativa confiável, em adição à resposta quantitativa, significava que a análise qualitativa não estava longe do desuso. Contudo, testes qualitativos baseados em reações bioquímicas, enzimáticas e imunológicas abriram novas perspectivas para esse tipo de determinação, devido principalmente à simplicidade e confiabilidade. Nesse contexto, muitos fabricantes vêm desenvolvendo *kits* de testes altamente confiáveis para uma ampla variedade de combinações de analitos-matrizes nas mais diversas áreas de aplicação, de forma que as respostas binárias têm adquirido uma crescente importância^{14,15}.

Embora seja uma tendência, da química analítica moderna, o desenvolvimento de novas técnicas e de métodos confiáveis para identificação e quantificação de componentes em matrizes complexas relacionadas à

segurança alimentar, como técnicas acopladas, incluindo cromatografia e espectrometria de massas ou outras técnicas espectroscópicas, essas poderosas ferramentas envolvem investimentos consideráveis e requerem analistas altamente capacitados. Do ponto de vista prático, muitos usuários têm reconsiderado até que ponto é realmente necessário obter resultados quantitativos⁶. Na rotina dos laboratórios, muitas vezes torna-se inviável o processamento de milhares de amostras por métodos convencionais¹⁵, sendo adotada uma primeira etapa para determinação da presença de um ou mais analitos (métodos de vanguarda/triagens) e uma segunda etapa para estimativa do nível de concentração dos mesmos (métodos de retaguarda/confirmação e quantificação). Assim, os métodos qualitativos têm sido comumente empregados para triagens antes da confirmação e quantificação, devido ao tempo e custo analítico reduzidos desses métodos⁶.

CLASSIFICAÇÃO DE MÉTODOS QUALITATIVOS

Os métodos qualitativos podem ser: identificação de analitos ou classificação de amostras. A identificação de analitos, por meio de técnicas analíticas modernas, tem sido muito utilizada. Por outro lado, as análises qualitativas de classificação têm como objetivo fornecer uma classificação de amostras rápida e confiável, em níveis de critérios previamente estabelecidos. As análises qualitativas podem ser implementadas pelo uso de ferramentas analíticas, podendo utilizar sistemas como os *kits* de testes, que produzem resultados binários sem nenhum tratamento de dados, ou sistemas baseados em técnicas analíticas, que fornecem dados brutos e requerem um tratamento posterior para sua conversão em resposta binária. O primeiro grupo é mais usado para a classificação de amostras, enquanto o segundo é mais apropriado para a identificação de analitos¹⁸.

Outras classificações são propostas para os métodos qualitativos. Com base na técnica classificam-se em químicos, físico-químicos, bioquímicos ou biológicos¹⁹, diferenciam-se de acordo com o estado físico da amostra em sólido, líquido ou gasoso²⁰ ou ainda com o sistema de detecção em sensorial ou instrumental^{16,20}. A principal característica dos métodos qualitativos baseados em detecção sensorial é que o sentido humano é utilizado para registro e interpretação da resposta. Os métodos baseados na detecção instrumental fornecem respostas instrumentais que, em muitos casos, são

medidas de absorvância embora, em princípio, qualquer instrumento possa ser utilizado. Em uma abordagem metrológica, os métodos binários instrumentais distinguem-se em instrumentais baseados em sinais univariados e instrumentais baseados na identificação de características específicas¹⁴. Os métodos instrumentais baseados em sinais univariados são basicamente métodos quantitativos utilizados como métodos qualitativos. Desse modo, o resultado final é obtido pela comparação entre as respostas de duas amostras, sendo que uma das amostras contém o analito de interesse em um nível de concentração determinado. Os métodos instrumentais baseados na identificação de características específicas são multiparamétricos, ou seja, a resposta analítica depende de vários parâmetros envolvidos na análise e a identificação dos compostos está baseada, por exemplo, no tipo ou na presença de compostos críticos da amostra.

VALIDAÇÃO DE MÉTODOS QUALITATIVOS

A Norma ISO/IEC 17025:2005 é a norma internacional que trata dos requisitos para a implementação de sistemas de gestão da qualidade em laboratórios de ensaio e calibração. A norma estabelece que métodos normalizados utilizados fora dos escopos para os quais foram concebidos, ampliados ou modificados, métodos não normalizados e métodos criados ou desenvolvidos pelos laboratórios devem ser validados. Apesar de não tratar da necessidade de validação para métodos normalizados, a referida norma define que os laboratórios devem confirmar se têm condições para operar adequadamente os métodos normalizados antes da implantação dos ensaios².

Vários são os documentos que podem ser utilizados como referência na elaboração de procedimentos de validação intralaboratorial de métodos analíticos, embora não exista uma uniformidade sobre quais características devam ser determinadas no processo de validação. O documento orientativo DOQ-CGCRE-008 do INMETRO¹⁰ é a principal referência nacional, podendo ainda ser citada a Instrução Normativa 24 de 2009, do Ministério da Agricultura, Pecuária e Abastecimento²¹, para validação de métodos no âmbito do Programa Nacional de Controle de Resíduos e Contaminantes em Produtos de Origem Animal (PNCRC). Documentos como o guia publicado pelo EURACHEM¹¹, a diretiva da União Europeia *Commission Decision* 2002/657/EC que estabelece Regulamentos sobre Desempenho de Métodos

e Interpretação de Resultados¹² e o guia harmonizado pela *Association of Official Analytical Chemists (AOAC International)*, *International Standards Organization (ISO)* e *International Union of Pure and Applied Chemistry (IUPAC)*¹³ são importantes referências internacionais que tratam da validação de métodos de ensaio.

Na diretiva *Comission Decision 2002/657/EC*¹² é feita uma distinção entre métodos de triagem e confirmação, que podem ser tanto qualitativos quanto quantitativos. Os primeiros detectam a presença de uma substância ou uma classe de substâncias no nível de interesse que têm capacidade para processar um número elevado de amostras e são utilizados para selecionar os resultados potencialmente não conformes. Os métodos confirmatórios produzem uma informação completa ou complementar que permite a identificação inequívoca da substância e, se necessário, sua quantificação no nível requerido.

Não existe um guia harmonizado para a validação de métodos qualitativos. Não obstante, há concordância de que a seletividade e o limite de detecção devam ser avaliados¹⁰⁻¹³. Estudos sobre robustez são tratados nos documentos do INMETRO¹⁰ e da União Europeia¹², embora esse parâmetro não esteja previsto no guia EURACHEM¹¹. Nesse documento são mencionados estudos das taxas de falso-positivos e falso-negativos para validação de métodos qualitativos. Os parâmetros de aplicabilidade e estabilidade são tratados como aplicáveis aos métodos qualitativos pela diretiva *Comission Decision 2002/657/EC* da União Europeia¹².

No guia harmonizado pela AOAC/ISO/IUPAC¹³ não é feita distinção entre métodos qualitativos e quantitativos ao apresentar os parâmetros de desempenho da validação. Contudo, a seletividade é tratada como um parâmetro essencialmente qualitativo e a aplicação dos parâmetros limite de detecção e robustez estão implícitos nas definições apresentadas.

Apesar de não haver documentos orientativos harmonizados para a validação de métodos qualitativos, esses métodos também precisam ser validados⁵ e, por isso, alguns documentos específicos vêm sendo publicados. Esses, no entanto, constituem propostas de validação de métodos qualitativos com aplicações específicas, que dependem das características intrínsecas do método de análise em questão, e não trazem detalhes sobre o delineamento experimental para avaliação dos parâmetros da validação.

A União Europeia possui um guia para a validação de métodos de triagem para resíduos de

drogas veterinárias baseado na diretiva *Comission Decision 2002/657/EC* e que consiste em um protocolo inicial de validação e recomendações de controle de qualidade desses métodos. Os métodos de triagem tratados pelo guia incluem os métodos qualitativos, semi-quantitativos e quantitativos. O guia sugere dois procedimentos que podem ser utilizados na validação dos métodos em questão, o primeiro baseado em uma abordagem clássica, em que os parâmetros são estudados em diferentes experimentos e, o segundo baseado em um modelo de delineamento fatorial. Em ambos os procedimentos, os parâmetros abordados são: seletividade, capacidade de detecção, limite de corte, aplicabilidade e robustez²². O *AOAC Performance Tested Program™ for test kits* foi desenvolvido para complementar o programa *Official Methods of Analysis*SM da AOAC²³ e tem como finalidade fornecer aos fabricantes de *kits* de testes uma avaliação independente da alegação de desempenho dos mesmos. Ao conceder o certificado de *Performance Tested* a um *kit* específico, a *AOAC Research Institute (AOAC-RI)* garante ao usuário do *kit* que, além de ter sido submetido a uma avaliação independente, o *kit* possui o desempenho alegado pelo fabricante dentro dos parâmetros previstos. Na avaliação desses *kits*, a seleção dos parâmetros avaliados depende do tipo de método avaliado, se qualitativo ou quantitativo. Para *kits* que envolvem determinações qualitativas, o programa prevê o estudo de limites, interferentes ou reatividade cruzada, taxas de falsos resultados e robustez²⁴. Em 2010, o CODEX ALIMENTARIUS publicou a proposta de um guia para métodos de detecção, identificação e quantificação de sequências específicas de DNA e de proteínas em alimentos obtidos por técnicas de biotecnologia, o qual trata também da validação desses métodos, incluindo a validação de métodos qualitativos de PCR (*Polymerase Chain Reaction*) e ELISA (*Enzyme-Linked Immunosorbent Assay*)²⁵. Neste documento são abordados os seguintes parâmetros: taxa de falso-positivos, taxa de falso-negativos, limite de detecção e robustez.

PARÂMETROS DE DESEMPENHO

Na Tabela 1 estão apresentados os parâmetros de desempenho estudados na validação de métodos qualitativos e os respectivos parâmetros correspondentes na validação de métodos quantitativos.

Tabela 1. Parâmetros de desempenho estudados na validação de métodos qualitativos e respectivos parâmetros correspondentes na validação quantitativa

Método Qualitativo	Método Quantitativo
Taxas de falso-positivos, de falso-negativos, de confiabilidade, rastreabilidade metrológica	Exatidão
Taxa de sensibilidade	Sensibilidade
Taxa de seletividade/ Confirmação de identidade	Seletividade
Região de perda de confiabilidade	Incerteza
Limite de detecção/ Limite de corte	Limite de detecção/Limite de quantificação/Limite de decisão/ Capacidade de detecção
Robustez	Robustez
Acordância e Concordância	Precisão

Fonte: Adaptado de Trullols et al.⁶; Ellison e Fearn⁸ e Ríos et al.¹⁴

Os parâmetros de desempenho dos métodos podem ser classificados em: principais, básicos e de produtividade¹⁷. Os principais (exatidão e representatividade) constituem os atributos do resultado analítico, relacionados às duas propriedades metrológicas clássicas – rastreabilidade e incerteza. Parâmetros analíticos básicos (precisão, robustez, sensibilidade e seletividade) são os atributos relacionados ao processo de medição analítico e que servem de suporte para as propriedades principais. Por fim, os de produtividade correspondem aos atributos do processo analítico que possuem relevância prática, como velocidade, custo/efetividade e fatores pessoais.

Na definição das propriedades analíticas dos métodos qualitativos, dois parâmetros são de suma importância: a taxa de falso-positivos e a taxa de falso-negativos. A taxa de falso-positivos é a probabilidade de obter um resultado positivo, uma vez que o analito não está presente na amostra (ou do método avaliado responder positivamente a um método de referência). A taxa de falso-negativos é a probabilidade de obtenção de um resultado negativo quando a substância está presente na amostra (ou do método avaliado responder negativamente a um método de referência)^{6,8,26}.

A confiabilidade de um método qualitativo é definida como a proporção de resultados corretos (positivos ou negativos) de uma bateria de testes independentes. Trata-se de um parâmetro positivo, que

abarca conotações negativas dos erros nas medições qualitativas. A confiabilidade é uma propriedade principal dos testes qualitativos, podendo ser relacionada à exatidão dos métodos quantitativos^{14,16}.

As taxas de sensibilidade e seletividade são parâmetros de desempenho de métodos qualitativos, intimamente relacionados com as taxas de falsos resultados. A sensibilidade, também chamada poder do teste, consiste na habilidade do método em detectar amostras verdadeiramente positivas como positivas, enquanto a taxa de sensibilidade correspondente à probabilidade de um método classificar como positiva uma amostra sabidamente positiva^{6,8,17}. No contexto de testes de hipóteses estatísticos, poder do teste é também definido como $1-\beta$, em que β representa a probabilidade de se obter resultados falso-negativos⁵.

De maneira análoga, a seletividade corresponde à habilidade em detectar amostras verdadeiramente negativas como negativas, sendo a taxa de seletividade a probabilidade de classificar como negativa uma amostra sabidamente negativa^{6,8,17}. Em testes de hipóteses estatísticos (1-seletividade) é relacionada ao nível de significância α , o qual representa a probabilidade de se obter resultados falso-positivos⁵.

Quando amostras que podem ou não conter determinado analito são analisadas, as taxas de falsos resultados, de sensibilidade e de seletividade são parâmetros que caracterizam a eficiência do método de ensaio em questão. Para amostras que contêm o analito, outros parâmetros podem ser estimados como o valor preditivo positivo e o valor preditivo negativo. No entanto, como em ensaios com amostras reais, é impossível distinguir as respostas falsas das verdadeiras, ou seja, os resultados verdadeiramente positivos dos falso-positivos e os verdadeiramente negativos dos falso-negativos, os valores preditivos podem ser estimados somente se a prevalência de amostras contendo o analito for conhecida²⁷.

Nas análises qualitativas, a incerteza não pode ser expressa do mesmo modo que nas análises quantitativas, ou seja, como um parâmetro que caracteriza a dispersão dos resultados. No caso dos métodos qualitativos, a incerteza tem uma natureza probabilística, podendo ser definida como a probabilidade de se tomar uma decisão errada. Os métodos com respostas binárias não possuem um número associado ao resultado e, assim, a incerteza não é expressa como um valor numérico, mas como uma região onde há probabilidade de ocorrência de erros. Nos

métodos qualitativos, o conceito clássico de incerteza pode ser substituído pelo conceito de região de perda de confiabilidade, termo preferencialmente utilizado. Esta região é definida pelos limites superiores e inferiores de concentração, entre os quais o método produz resultados falsos^{5,6,14,28}.

O limite de detecção, conforme definido para análises quantitativas, pode ser calculado quando a resposta é um valor numérico, sendo consideradas duas probabilidades de erro α e β ²⁹, onde α é a probabilidade de se cometer um erro do tipo I, de rejeitar H_0 (hipótese de nulidade) quando H_0 é verdadeira, que mede a taxa dos falso-positivos e β é a probabilidade de se cometer um erro do tipo II, de não rejeitar H_0 quando H_0 é falsa, que mede a taxa de falso-negativos. Quando a resposta é binária, entretanto, o desvio padrão de amostras brancas não pode ser calculado e as probabilidades de erro não podem ser consideradas ao mesmo tempo. Portanto, o limite de detecção pode ser definido como a menor concentração de um analito que pode ser detectada, confiavelmente como positiva, em uma dada matriz. Isto implica em considerar somente o erro tipo II, ou taxa de resultados falso-negativos, usualmente de 5 %. Esta definição é apresentada no contexto de verificação de um nível máximo de concentração permitido, mas se for extrapolado a determinação de uma concentração mínima permitida, a probabilidade de erro tipo I, ou taxa de falso-positivos, deve ser considerada também a 5 %. Como as duas probabilidades de erro não podem ser consideradas simultaneamente, no primeiro caso o limite de detecção coincidirá com o limite superior da região de perda de confiabilidade, na qual a taxa de sensibilidade é de 95 % e, no segundo caso, ele coincidirá com o menor limite da região de perda de confiabilidade⁶.

Além do limite de detecção do método, que é uma característica intrínseca do método de ensaio, o limite imposto pela legislação ou pelo cliente e o limite de corte do laboratório para métodos que possuem respostas instrumentais (*cut-off*) estão entre as referências quantitativas envolvidas nos resultados dos métodos qualitativos¹⁷.

A robustez é a estimativa da capacidade de um método frente às pequenas e deliberadas mudanças nas variáveis estudadas, sendo determinada em diferentes níveis de concentração do analito¹⁸. A robustez em análise qualitativa é um parâmetro analítico, assim como nas análises quantitativas, cujo ultimato é definir as fragilidades experimentais de um método pela indicação

das variáveis críticas na garantia da confiabilidade das respostas. A robustez depende da concentração dos analitos, sendo que, na faixa de perda de confiabilidade, haverá maior influência de fatores experimentais^{4,17}.

A rastreabilidade metrológica, do ponto de vista prático, assume que existe uma cadeia ininterrupta de calibrações de um sistema de medição ou comparações. Entre outras possibilidades, a rastreabilidade metrológica pode ser obtida comparando os resultados obtidos pelo método em validação com aqueles obtidos por um método de referência ou pelo uso de um material de referência certificado (MRC)^{6,14,30}. É importante ressaltar que os padrões de referência são a chave para a garantia da rastreabilidade metrológica das informações e resultados fornecidos pelos métodos qualitativos, já que as características analíticas desses padrões são utilizadas como base de comparação para compostos desconhecidos da amostra¹⁴.

Dois outros parâmetros são ainda acrescentados às validações de métodos qualitativos, quando estruturadas em processos interlaboratoriais: acordância e concordância; que correspondem respectivamente, aos conceitos de precisão sob condições de repetitividade e de reprodutibilidade, aplicados a testes qualitativos. A acordância é a probabilidade de obter, para duas amostras idênticas, analisadas no mesmo laboratório e, sob condições de repetitividade, o mesmo resultado. Para o cálculo da acordância, calcula-se a acordância para cada laboratório, que é a probabilidade de duas amostras fornecerem o mesmo resultado e, em seguida a média das probabilidades de cada laboratório³¹. Concordância é a probabilidade de obtenção do mesmo resultado para duas amostras idênticas, testadas em laboratórios diferentes e, sob condições de reprodutibilidade. Esse parâmetro é estimado pelas proporções correspondentes de pares de resultados de dados concordantes^{8,31}. Da mesma forma que, em estudos de métodos quantitativos, experimentos de concordância podem ser conduzidos sob condições de precisão intermediária, representando o que normalmente acontece no laboratório sob condições de rotina, incluindo variações em lotes de reagentes, analistas e equipamentos que sejam representativos¹³.

Na validação de métodos qualitativos, a representatividade torna-se um componente essencial da qualidade das repostas qualitativas. Uma resposta 100 % confiável, mas não representativa das amostras teste, indica um resultado inadequado do ponto de vista da qualidade do método. Estratégias adequadas de amostragem e a

correta aplicação do protocolo de amostragem são as bases para essa propriedade¹⁷. A escolha das amostras deve fornecer estimativas de sensibilidade e seletividade, em uma dada faixa de concentração, o que implica em um delineamento que empregue amostras em uma faixa representativa de concentrações e, pelo menos um branco, além de alguns itens com interferentes. As replicatas também são essenciais na avaliação de métodos qualitativos, sem as quais as frequências não poderiam ser estimadas⁸.

ANÁLISE DE DADOS E AVALIAÇÃO DOS PARÂMETROS

Na Tabela 2 estão descritos os delineamentos experimentais e respectivas ferramentas de análises de dados propostos na literatura para avaliação dos parâmetros de desempenho de métodos qualitativos.

Para análise de dados e avaliação dos parâmetros, vários modelos são propostos. As tabelas de contingência e o Teorema de Bayes podem ser utilizados na determinação das taxas de falsos

resultados, de sensibilidade, de seletividade e de confiabilidade. As curvas de probabilidades ou curvas de desempenho e os testes de hipóteses estatísticos podem ser utilizados no estudo de limites e da região de perda de confiabilidade (incerteza)^{6,8}. No caso da precisão, a acordância e a concordância são avaliadas por cálculos de probabilidade e análise combinatória³¹. As taxas, limites, incerteza, acordância e concordância podem ser estudadas em um mesmo delineamento experimental que consiste em ensaios com amostras de concentrações conhecidas do analito.

Para os estudos de robustez podem ser utilizados experimentos fatoriais. Neste caso, o experimento inclui ensaios com amostras em um nível de concentração do analito para o qual o método apresenta 100 % de confiabilidade sob diferentes condições experimentais⁴.

A validação de métodos qualitativos exige maior número de medições para obtenção de valores para os parâmetros de validação com significado e níveis de confiança semelhantes aos resultados dos métodos quantitativos³². Isso porque um resultado qualitativo fornece menos informações que um

Tabela 2. Delineamentos experimentais e respectivas ferramentas para análise de dados propostos na literatura para validação de métodos qualitativos

Parâmetro	Delineamento experimental	Análise de dados	Referência
		<ul style="list-style-type: none"> - Cálculo da porcentagem de resultados positivos em cada nível. - Construção de gráfico: probabilidade de positivos $P(x)$ versus concentração (x). - Na região da curva assumida linear, extrapola-se a curva para os pontos $P(x) = 0$ e $P(x) = 100$, que são os limites da região de perda de confiabilidade. 	36
Região de perda de confiabilidade (incerteza) e taxas de falsos resultados	<ul style="list-style-type: none"> - Estabelecimento do nível de concentração de referência (C) - limite de corte ou <i>threshold</i>. - Análise das amostras, sendo 10 replicatas por nível e 5 níveis, selecionados da seguinte forma: 2 níveis acima de C, C e 2 níveis abaixo de C. 	<ul style="list-style-type: none"> - Estimativa da região de perda de confiabilidade conforme descrito por Song et al.³⁶. 	5,18
	<ul style="list-style-type: none"> - Análise das amostras, sendo 30 replicatas por nível. Um grande número de amostras deve ser analisado para assegurar inferência estatística. - Níveis selecionados devem estar levemente abaixo do limite de corte ou <i>threshold</i> selecionado. 	<ul style="list-style-type: none"> - Cálculo de porcentagem de resultados positivos em cada nível. - Construção de gráfico: $P(x)$ versus (x). - Limites da região de perda de confiabilidade determinados pelo cálculo das concentrações do analito que produzem 5% de probabilidade de se obter um resultado falso-positivo e 95% de probabilidade de se obter um resultado falso-negativo ou conforme descrito por Song et al.³⁶. - Aplicação do Teorema Bayes não é apropriada para propósitos de rotina: difícil desenvolvimento de equações e grande número de amostras. 	28

Cont. Tabela 2

Parâmetro	Delineamento experimental	Análise de dados	Referência
	<ul style="list-style-type: none"> - Análise aleatória de grande número de amostras (n>30), sendo MRC ou amostras brancas contendo o analito em diferentes níveis de concentração. - Escolha dos níveis: quanto maior a confiabilidade requerida, maior o número de níveis a serem estudados. - Estabelecimento de nível de concentração de referência (C) - limite de corte ou <i>threshold</i>. - Determinação da região de perda de confiabilidade: <ol style="list-style-type: none"> 1. Estimativa aproximada das zonas limite da região de perda de confiabilidade pela análise de 1 ou 2 replicatas por nível, em 10 níveis selecionados: 5 superiores a C e 5 inferiores a C. 2. Análise de 5 a 10 replicatas em 6 níveis, sendo 3 níveis na zona limite superior e 3 na zona limite inferior. 	<ul style="list-style-type: none"> - Determinação das taxas de falsos resultados para delimitação dos limites da região de perda de confiabilidade a um nível de probabilidade fixa. 	17
Região de perda de confiabilidade (incerteza) e taxas de falsos resultados	<ul style="list-style-type: none"> - Métodos para avaliação de falsos resultados: <ol style="list-style-type: none"> 1. Baseado na observação de probabilidades de falsos resultados em uma série de testes controlados. Experimentos devem estar concentrados em regiões onde há maior probabilidade de resultados falsos, pois as taxas de falsos resultados são difíceis de observar em número realista de experimentos a menos que a taxa seja alta (perto de 50%). 2. Baseado na previsão das características conhecidas da população, incluindo estatísticas das medições quantitativas e distribuições conhecidas de características de teste de amostra de uma população. Obs.: trata somente incerteza. 	<ul style="list-style-type: none"> - Estimativas Bayesianas podem ser calculadas pela combinação adequada das taxas de falsos resultados. 	38
	<ul style="list-style-type: none"> Obs.: trata somente taxas. 	<ul style="list-style-type: none"> - Estimativa das taxas por meio da aplicação do Teorema de Bayes, tabelas de contingência, curvas de desempenho ou testes estatísticos (nesse último, somente se a resposta for obtida por método instrumental). - Apresentação das fórmulas de cálculo das taxas obtidas a partir das tabelas de contingência. 	6
	<ul style="list-style-type: none"> Obs.: trata somente taxas. 	<ul style="list-style-type: none"> - Estimativa das taxas pela aplicação de fórmulas de cálculo obtidas a partir das tabelas de contingência. 	3
Limites	<ul style="list-style-type: none"> - Análise aleatória de amostras adicionadas do analito em uma faixa de níveis de concentração, sendo 10 replicatas por nível. 	<ul style="list-style-type: none"> - Construção de curva de resposta P(x) versus x. Limite de detecção é determinado de acordo com a concentração na qual o teste deixa de ser confiável. 	11
	<ul style="list-style-type: none"> Procedimento interlaboratorial: <ul style="list-style-type: none"> - 10 baterias. - 5 replicatas por bateria. 	<ul style="list-style-type: none"> - Aplicação de fórmulas deduzidas a partir de cálculos de probabilidade e análise combinatória. 	31
Acordância e concordância	<ul style="list-style-type: none"> Procedimento interlaboratorial para ensaios não microbiológicos: <ul style="list-style-type: none"> - 20 a 40 replicatas por bateria, incluindo replicatas, sendo 30 um número satisfatório. - 5 a 6 replicatas por material, podendo ser menor no caso de grande número de baterias; mas no mínimo 2 replicatas por teste em caso de grande quantidade de materiais. 	<ul style="list-style-type: none"> - Fórmulas propostas por Langtonn et al.³¹ 	8

Cont. Tabela 2

Parâmetro	Delineamento experimental	Análise de dados	Referência
	<ul style="list-style-type: none"> - Identificação e seleção das principais variáveis que afetam o desempenho do método (considerar informações obtidas previamente na otimização e validação do método). - Seleção dos números de níveis a serem estudados para as variáveis, 2 em geral. - Seleção de delineamento experimental. No caso de método com pequeno número de variáveis e níveis envolvidos, pode-se estabelecer um experimento fatorial completo. - Estabelecimento de variação no valor da resposta que significa alteração no desempenho do método. - Condições de estudo de caso: <ol style="list-style-type: none"> 1. Selecionadas 3 variáveis e 2 níveis, totalizando 8 diferentes experimentos. 2. Experimentos realizados em 8 dias, sendo cada experimento realizado 5 vezes, por um período de quatro dias, totalizando 20 repetições por experimento. 3. Somente alterações de 5% nas taxas de confiabilidade, ou de falsos resultados, foram consideradas. 	<ul style="list-style-type: none"> - Apresentação de fórmulas para cálculo dos efeitos dos fatores e realização de teste t para conclusão da significância do efeito do fator sobre os resultados do método. 	6
Robustez	<ul style="list-style-type: none"> - Identificação de variáveis que têm efeito significativo sobre o desempenho do método. - Análise com materiais de referência, MRC ou amostras de composição conhecida. - Monitoramento da exatidão e precisão do método frente às alterações das variáveis. - Análise de cada <i>set</i> experimental uma vez. 	<ul style="list-style-type: none"> - Determinação do efeito de cada alteração na condição de realização do teste na média. - Classificação das variáveis de acordo com o efeito sobre o desempenho do método. 	11
	<ul style="list-style-type: none"> - Identificação de fatores que podem influenciar os resultados por estudos prévios. - Fatores alterados em ordem de grandeza coerente com os desvios geralmente encontrados entre laboratórios. - Realização de teste utilizando a abordagem de Youden (planejamento fatorial fracionário). Outros métodos podem ser utilizados, no entanto, na abordagem de Youden o tempo e esforço exigidos são mínimos. - Sempre que um fator que influencia significativamente os resultados for encontrado, novas análises devem ser realizadas para estudo dos limites de aceitabilidade deste fator. - Fatores que influenciam significativamente os resultados devem ser claramente identificados no protocolo do método. 	<ul style="list-style-type: none"> - Na abordagem de Youden, as interações entre os diferentes fatores não podem ser detectadas. 	12

resultado quantitativo, independentemente do método estatístico empregado^{28,32}. Para os métodos quantitativos, a independência dos resultados qualitativos deve ser assegurada pela análise aleatória das amostras^{8,17}.

As amostras selecionadas para realização dos estudos de validação podem ser MRC, amostras brancas e adicionadas de padrão do analito de interesse^{14,17}.

Taxas de falsos resultados, de sensibilidade, de seletividade e de confiabilidade

Tabelas de Contingência

As tabelas de contingência tratam análises qualitativas como a classificação de um problema. Havendo dois tipos de resultados possíveis no teste

qualitativo, como positivo (maior ou igual a um valor especificado) ou negativo (menor que o valor especificado), e duas possíveis situações reais, uma tabela de contingência 2 x 2 pode ser construída, como apresentado na Tabela 3. Embora as amostras sejam consideradas em dois grupos, são necessários estudos com amostras em vários níveis de concentração⁵. Estas tabelas podem ser utilizadas na determinação das taxas de falsos resultados, taxa de seletividade, taxa de sensibilidade e taxa de confiabilidade para cada nível de concentração estudado⁶.

Tabela 3. Exemplo de Tabela de Contingência 2 x 2

Resultado do teste	Analito Presente	Analito Ausente	Total
Positivo	TP	FP	TP + FP
Negativo	FN	TN	FN + TN
Total	FN + TP	FP + TN	N

FP = quantidade de resultados falso-positivos; FN = quantidade de resultados falso-negativos; TP = quantidade de resultados positivos corretos; TN = quantidade de resultados negativos corretos; N = TP+FP+FN+TN

A partir das fórmulas de cada uma das taxas, os valores de cada uma delas podem ser calculados com os dados da tabela de contingência. A taxa de falso-positivos (TFP) é definida como a razão entre o número de resultados falso-positivos (FP) e a soma desse com o número de resultados negativos corretos ou número total de negativos conhecidos (TN), multiplicada por 100 (Equação 1)^{6,8,26}:

$$TFP = \frac{FP}{FP + TN} \times 100 \quad (\text{Eq. 1})$$

De maneira análoga, a taxa de falso-negativos (TFN) é a razão entre o número de resultados falso-negativos (FN) e a soma desse com o número de resultados positivos corretos ou número total de positivos conhecidos (TP), multiplicada por 100 (Equação 2)^{6,8,26}:

$$TFN = \frac{FN}{FN + TP} \times 100 \quad (\text{Eq. 2})$$

A taxa de sensibilidade (TSB) é definida como a razão entre TP e a soma desse com FN, multiplicada por 100 (Equação 3), enquanto a taxa de seletividade (TST)

pode ser definida como a razão entre TN e a soma desse com FP, multiplicada por 100 (Equação 4)^{6,8,17}:

$$TSB = \frac{TP}{TP + FN} \times 100 \quad (\text{Eq. 3})$$

$$TST = \frac{TN}{TN + FP} \times 100 \quad (\text{Eq. 4})$$

A taxa de confiabilidade (TCF) é definida como a diferença entre o total de resultados (100 %) e a soma de TFP e TFN, como apresentado na Equação 5:

$$TCF = 100 - TFP - TFN \quad (\text{Eq. 5})$$

As tabelas de contingência são frequentemente usadas na avaliação da confiabilidade dos testes de triagem e na comparação entre diferentes ensaios, que se baseiam na sensibilidade e na seletividade de diferentes testes de triagem, quando aplicados às mesmas amostras. Nesse caso, deve ser observado que diferenças na sensibilidade e na seletividade dos testes podem ser resultado do acaso, já que situações reais e resultados estão sujeitos à erros aleatórios, especialmente quando o número de amostras é pequeno. Dependendo do objetivo da análise, é preferível que o método seja mais sensível que seletivo, ou vice-versa. Em geral, um aumento na sensibilidade acarreta um decréscimo na seletividade. Desse modo, quanto maior o número de amostras, mais informação a tabela de contingência irá apresentar e maior será a confiabilidade desses parâmetros⁵.

Uma das vantagens mais importantes das tabelas de contingência é que essas podem ser facilmente aplicáveis aos vários tipos de bioensaios existentes, como imunológicos, microbiológicos, clínicos e bioquímicos. Embora, pela avaliação das tabelas de contingência seja possível obter uma visão geral sobre o desempenho do método, tais tabelas, por si só, não fornecem informações individuais, como a probabilidade de erro para cada amostra⁶. Outra desvantagem é que o desempenho da tabela de contingência depende muito do tamanho da amostra e do delineamento experimental utilizado para construí-la⁵.

Teorema de Bayes

O teorema formulado por Bayes, no século XVIII, vem sendo muito utilizado nas últimas décadas

por sua versatilidade e aplicação em diversos campos da ciência, para expressar e atualizar probabilidades quando estruturas dicotômicas³³ estão envolvidas, incluindo as análises qualitativas^{5,33}.

O teorema de Bayes é representado pela Equação 6:

$$P(A|T) = \frac{P(T|A) \cdot P(A)}{P(T|A) \cdot P(A) + P(T|\bar{A}) \cdot P(\bar{A})} = \frac{P(T|A) \cdot P(A)}{P(T)} \quad (\text{Eq. 6})$$

Sendo $P(A|T)$ a probabilidade de A dado T , $P(A)$ a probabilidade a priori de A , $P(T|A)$ a probabilidade de T dado A , $P(T)$ a probabilidade de T . O travessão acima de A denota o evento complementar.

No caso das análises qualitativas, A denota a presença do analito na amostra e \bar{A} corresponde à ausência do analito na amostra, enquanto T representa o resultado do teste. $P(A)$ é a probabilidade de o analito estar presente na amostra antes da realização do teste. $P(A|T)$ é a probabilidade de o analito estar presente na amostra levando em conta o resultado do teste, enquanto $P(T|A)$ é a probabilidade de ocorrer o resultado estando o analito presente na amostra, o que indica a TSB ³⁴.

Sendo T o resultado do teste e \bar{T} o evento complementar, temos que $P(T|A) = 1 - P(\bar{T}|A)$, em que $P(\bar{T}|A)$ é a probabilidade de se obter um resultado falso-negativo, ou seja, a TFN . Sendo assim, $P(T|\bar{A})$ é a probabilidade de se obter um resultado positivo na ausência de A , correspondendo à probabilidade de um resultado falso-positivo, ou seja, à TFP ³⁴.

Desse modo, o teorema de Bayes pode ser utilizado na determinação das taxas. A taxa de falso-positivos é estimada pela proporção de resultados incorretos relatados para amostras conhecidas que não contém a substância pesquisada (Equação 7). Da mesma forma, a taxa de falso-negativos é estimada pela proporção de resultados incorretos para as amostras que contém a substância pesquisada (Equação 8)^{6,8,26}:

$$P(T|\bar{A}) = \frac{\text{Resultados falso-positivos}}{\text{Total de resultados negativos}} \times 100 = \frac{FP}{FP+TN} = TFR \quad (\text{Eq. 7})$$

$$P(\bar{T}|A) = \frac{\text{Resultados falso-negativos}}{\text{Total de resultados positivos}} \times 100 = \frac{FN}{FN+TP} = TFN \quad (\text{Eq. 8})$$

A TSB é estimada a partir da proporção de resultados corretos relatados para amostras conhecidas que contém o analito (Equação 9) e a TST é estimada a

partir da proporção de resultados corretos para os itens de ensaio que não contém o analito (Equação 10)^{6,8,17}:

$$P(T|A) = \frac{\text{Resultados positivos verdadeiros}}{\text{Total de resultados positivos}} \times 100 = \frac{TP}{TP+FN} = TSB \quad (\text{Eq. 9})$$

$$P(\bar{T}|\bar{A}) = \frac{\text{Resultados negativos verdadeiros}}{\text{Total de resultados negativos}} \times 100 = \frac{TN}{TN+FP} = TSB \quad (\text{Eq. 10})$$

A relação entre as probabilidades condicionais de sensibilidade e seletividade fornece a razão de verossimilhança (*likelihood ratio* - LR) (Equação 11), que é uma medida do poder do teste³⁵, ou seja, é uma medida da alteração das probabilidades após obtenção do resultado do teste³⁴:

$$LR = \frac{TSB}{1 - TST} = \frac{1 - TFN}{TFP} = \frac{P(T|A)}{P(T|\bar{A})} \quad (\text{Eq. 11})$$

A TCF , é estimada pela proporção de resultados verdadeiros no total de ensaios realizados (Equação 12):

$$P(A|T) = \frac{\text{Total resultados verdadeiros}}{\text{Total de resultados}} \times 100 = \frac{TP+TN}{TP+TN+FP+FN} \times 100 =$$

$$P(A|T) = 100 - TFP - TFN = TCF \quad (\text{Eq. 12})$$

O Teorema de Bayes permite o cálculo de probabilidades condicionais referentes a apenas uma amostra, incorpora as informações de realizações anteriores e obtém como resultado, para a próxima realização, a probabilidade a posteriori a nova probabilidade a priori. Assim, a probabilidade da ocorrência de um resultado errôneo é dada individualmente^{5,6,34}. No entanto, para que o procedimento forneça boas estimativas, o número de testes de ensaio deve ser alto. Do ponto de vista analítico, as desvantagens do método são a complexidade da nomenclatura e a dificuldade de quantificação dos diferentes valores das probabilidades⁵.

Limites e região de perda de confiabilidade

Curvas de desempenho

As curvas de desempenho são obtidas pela construção de um gráfico de probabilidade em função da concentração. Em uma situação ideal, em que x representa a concentração verdadeira, $P(x)$ a taxa de resultados

positivos e $N(x)$ a taxa de resultados negativos, para uma determinada concentração de referência C ($x = C$), os possíveis resultados obtidos em um método binário são: não (resultados negativos), quando $x < C$, com $P(x) = 0\%$ e $N(x) = 100\%$; e sim (resultados positivos), quando $x > C$, com $P(x) = 100\%$ e $N(x) = 0\%$ ^{28,36}.

No entanto, em situações reais, concentrações próximas a C ($C_0 < C < C_1$), $P(x)$ e $N(x)$ apresentam valores entre 0% e 100%. A faixa de concentração que inclui esses valores intermediários é justamente a região de perda de confiabilidade. Assim, nesta região, são obtidas as taxas de falso-positivos ($C_0 < x < C$) e as taxas de falso-negativos ($C < x < C_1$). Para amostras contendo $x < C_0$ e $x > C_1$, são obtidos resultados negativos e positivos, respectivamente, com 100% de confiabilidade²⁸.

Na Figura 1, estão ilustradas as curvas de desempenho para a situação ideal e para uma situação real genérica, em que $\alpha = \beta = 0,05$ ^{14,28,36}. Na curva de desempenho real, o limite de detecção pode ser estabelecido em função de α ou de β no intervalo da região de perda de confiabilidade que, por sua vez, é determinada pela região de maior inclinação da curva, ou seja, pelo intervalo entre os níveis nos quais a probabilidade de resultados positivos deixa de ser de 0% e 100%^{5,6,11,28}. Assim, os limites C_0 e C_1 , que definem a região de perda de confiabilidade, podem ser determinados pelo cálculo das concentrações do analito que produzem 5% de probabilidade de obtenção de falso-positivos e 95% de resultados positivos corretos (o que corresponde a 5% de probabilidade de obtenção de falso-negativos).

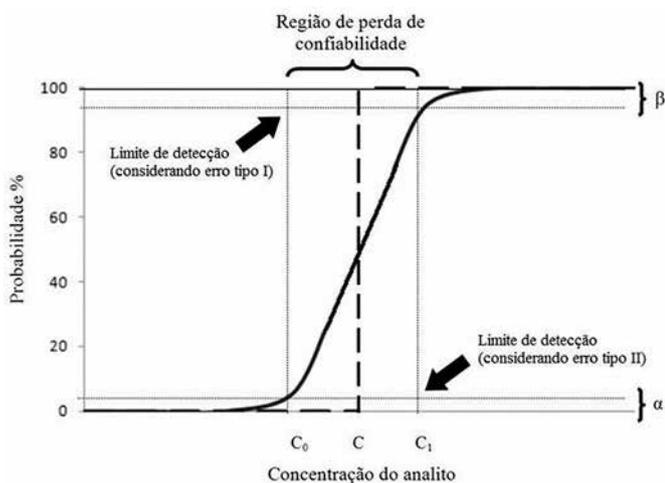


Figura 1. Curva de desempenho para determinação da região de perda de confiabilidade e limites relacionados

A principal desvantagem das curvas de desempenho é a necessidade de realização de análises em replicatas, acarretando consequências negativas, como o aumento do tempo e o custo do processo⁶. Contudo, Ellison e Fearn⁸ observaram que o estudo de um grande número de níveis com poucas repetições por nível, em vez de poucos níveis com muitas repetições, fornece estimativas de probabilidades relativamente confiáveis, o que pode ser considerada uma vantagem sob o ponto de vista prático.

Este procedimento é bastante usado quando não há muito conhecimento sobre a precisão do método de triagem utilizado, por exemplo, no caso de um método analítico recém-implantado em um laboratório⁵.

Testes de Hipóteses

Quando o objetivo da análise está na determinação da presença de determinado analito na amostra, duas hipóteses são geradas: H_0 (hipótese de nulidade) - a amostra contém o analito; e H_1 (hipótese alternativa) - a amostra não contém o analito.

Como resultado, pode-se obter dois tipos de erros diferentes: erroneamente detectar a presença do analito (admitir um resultado falso-positivo - erro tipo I) ou sua ausência (admitir um resultado falso-negativo - erro tipo II). Em geral, a preferência de uma escolha ou outra é uma questão de convenção²⁷.

Os testes de hipóteses comparam a resposta da amostra com a de uma referência pré-definida, que contém o analito em um nível de concentração específico⁶. Alguns estudos têm utilizado testes de hipóteses para avaliar a incerteza dos métodos qualitativos⁵ e, embora a aplicação destes testes não resulte em novas técnicas instrumentais de análise química qualitativa, esses tornam possível a melhoria do desempenho das atuais técnicas de estimativa de incerteza, refletindo o grau de confiabilidade do procedimento escolhido.

As abordagens utilizadas para exprimir a incerteza são possíveis pela aplicação de analogia de testes estatísticos de hipóteses com a aplicação das fórmulas de Bayes³⁷. Desse modo, as principais vantagens dos testes de hipóteses decorrem da utilização da probabilidade conhecida de erro tipo I e tipo II. Esse método torna mais fácil a avaliação da incerteza quando se utiliza métodos qualitativos que fornecem uma resposta instrumental. A rastreabilidade também pode ser verificada e o limite de detecção calculado. No entanto, se o *kit* de teste não fornecer uma resposta

instrumental ou se a resposta for baseada em uma observação visual, que não pode ser quantificada, o teste de hipótese não poderá ser aplicado⁶. Quando se utilizam intervalos estatísticos, os experimentos constituem uma série de análises de amostras independentes, contendo o analito em um limite específico, sob condições de precisão intermediária, aumentando as fontes de incerteza⁵.

Robustez

O tipo de resposta binária do método qualitativo define o estudo da robustez, que por sua vez, envolve o desenvolvimento de um delineamento experimental para avaliação de fatores ou variáveis previamente escolhidos que acarretam alterações na confiabilidade e nas taxas de falsos resultados. Esses fatores são as condições experimentais intrínsecas do método⁴. Desse modo, a robustez deve ser determinada em diferentes níveis de concentração do analito para cada um dos analitos considerados¹⁸.

Assim, em termos gerais, o estudo da robustez inclui os seguintes pontos: seleção das variáveis ou fatores; seleção do número de níveis para cada variável (em geral 2 níveis) e do delineamento experimental ótimo; estabelecimento do valor de reposta que provoca alterações nos fatores e nos experimentos e do cálculo dos efeitos de cada fator e entre os fatores; e elaboração das conclusões⁴.

Uma boa opção de delineamento experimental para o estudo da robustez, quando não se tem muitas variáveis e níveis envolvidos, é um experimento fatorial completo⁴. O delineamento experimental inclui a realização de testes em nível para o qual se obtém 100% de confiabilidade, ou seja, 0% de taxa de falso-positivos e 0% de taxa de falso-negativos, com variação nos níveis dos fatores estudados. O efeito de cada fator (*DF*) pode ser calculado pela Equação 13:

$$DF = \frac{\sum TCF(+)-\sum TCF(-)}{n_j} \quad (\text{Eq. 13})$$

Sendo *TCF(+)* a confiabilidade observada no maior nível do fator, *TCF(-)* a confiabilidade observada no menor nível do fator, *n* o número de experimentos e *n_j* o número de experimentos no *j*-ésimo nível do fator estudado⁴.

Após o cálculo do efeito dos fatores estudados, o teste *t* pode ser realizado para conclusão sobre a

significância da influência dos fatores sobre as respostas (Equação 14):

$$t = \frac{|DF| \times \sqrt{n_i}}{s \times \sqrt{2}} \quad (\text{Eq. 14})$$

Sendo *s* o desvio padrão encontrado em condições de precisão intermediária e *DF* o efeito de cada fator, definido entre 0 e 1. O valor calculado deve ser comparado ao valor tabelado de distribuição *t*, considerando o nível de significância e o número de graus de liberdade associados à estimação de *s*⁴.

Acordância e concordância

Como definido, a acordância é o parâmetro de repetitividade aplicado à análise qualitativa. A acordância é a probabilidade (em porcentagem) de dois materiais de teste idênticos, analisados pelo mesmo laboratório, em condições de repetitividade, fornecerem o mesmo resultado, ou seja, ambos fornecerem dois resultados positivos ou dois negativos³¹.

Para o cálculo da acordância, calcula-se a probabilidade de cada laboratório ou de cada bateria analítica (em estudos interlaboratoriais sob condições de precisão intermediária) fornecer o mesmo resultado e, em seguida, a média das probabilidades de todos os laboratórios ou baterias analíticas. A soma da quantidade de combinações de pares obtida pelo número de resultados positivos com a quantidade de combinações de pares obtida pelo número de resultados negativos dividida pela quantidade de combinações de pares, que podem ser formados pelo total de análises realizadas, fornece a probabilidade de cada laboratório ou de cada bateria analítica, ou seja, a acordância. De uma maneira mais simples, a acordância pode ser calculada pela Equação 15.

$$ACO = \frac{\{k(k-1) + (n-k)(n-k-1)\}}{n(n-1)} \quad (\text{Eq. 15})$$

Sendo *n* o total de resultados obtidos pelo laboratório ou bateria analítica e *k* o número de resultados positivos obtidos pelo laboratório ou bateria analítica³¹.

A concordância, o parâmetro qualitativo equivalente ao parâmetro quantitativo de reprodutibilidade, é a probabilidade de dois materiais de teste idênticos provenientes de diferentes laboratórios ou de diferentes baterias analíticas (em

estudos intralaboratoriais sob condições de precisão intermediária) fornecerem o mesmo resultado, ou seja, que ambos sejam positivos ou negativos. A concordância pode ser calculada de modo intuitivo enumerando-se as associações de resultados concordantes positivos e negativos de cada laboratório ou bateria analítica com os demais resultados do conjunto. A concordância será então o resultado da divisão entre a soma de associações de resultados concordantes positivos e negativos de todos os laboratórios ou baterias e a soma das associações de resultados dos laboratórios ou baterias. De modo mais simples, a concordância pode ser calculada pela Equação 16:

$$CON = \frac{2[k(k-nb)] + nb(nb-1) - ACO[nb(n-1)]}{n^2b(b-1)} \quad (\text{Eq. 16})$$

Sendo b o número de laboratórios ou baterias analíticas e ACO a concordância³¹.

CONCLUSÃO

A garantia da qualidade em análises qualitativas deve estar baseada na avaliação dos parâmetros de desempenho relacionados às propriedades analíticas desses métodos. Contudo, nem sempre os parâmetros conhecidos das análises quantitativas podem ser aplicados às análises qualitativas. Na validação de métodos quantitativos, alguns dos parâmetros estudados são: exatidão; sensibilidade; seletividade; incerteza; limites; e precisão. No campo das análises qualitativas, esses parâmetros possuem correspondência com as taxas de falsos resultados, sensibilidade e confiabilidade, seletividade, região de perda de confiabilidade, limite de detecção e concordância e concordância, respectivamente. Tabelas de contingência e o teorema de Bayes são adotados na avaliação das taxas de falsos resultados, sensibilidade, seletividade e confiabilidade. Os limites e região de perda de confiabilidade são determinados por meio de curvas de probabilidade. Concordância e concordância são estimadas a partir de conceitos de análise combinatória. A falta de referências bem estabelecidas para a metrologia química em relação às análises qualitativas e a crescente demanda por respostas binárias confiáveis, principais motivadores desse trabalho, sinalizam para o fato de que a validação adequada desses métodos deva ser considerada um importante tópico do presente e do futuro das ciências analíticas.

AGRADECIMENTOS

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

REFERÊNCIAS

1. Souza SVC. Procedimento para validação intralaboratorial de métodos de ensaio: delineamento e aplicabilidade em análises de alimentos [tese de doutorado]. Belo Horizonte (MG): Universidade Federal de Minas Gerais; 2007.
2. Associação Brasileira de Normas Técnicas – ABNT. ISSO (International Standard Organization). ABNT NBR ISO/IEC 17025. Requisitos gerais para a competência de laboratório de ensaio e calibração. Rio de Janeiro: ABNT; 2005. 31p. (b)
3. Brasil. Ministério da Agricultura, Pecuária e Abastecimento/Secretaria de Defesa Agropecuária. Instrução normativa número 8, de 29 de abril de 2010. Diário Oficial [da] União. Brasília, DF. 03 mai 2010; Seção 1, p. 27-30.
4. Aguilera E, Lucena R, Cárdenas S, Valcárcel M, Trullols E, Ruisánchez I. Robustness in qualitative analysis: a practical approach. *Trends Anal Chem*. 2006; 26(6):621-7.
5. Pulido A, Ruisánchez I, Boqué R, Rius FX. Uncertainty of results in routine qualitative analysis. *Trends Anal Chem*. 2003; 22(10):647-54.
6. Trullols E, Ruisánchez I, Rius FX. Validation of qualitative analytical methods. *Trends Anal Chem*. 2004; 23(2):137-45.
7. King B. In-house method validation. A guide for chemical laboratories. LGC Limited; 2003.
8. Ellison SLR, Fearn T. Characterising the performance of qualitative analytical methods: Statistics and terminology. *Trends Anal Chem*. 2005; 24(6):468-76.
9. Peters FT, Drummer OH, Musshoff F. Validation of new methods. *Forensic Sci Int*. 2007; 165:216-24.
10. Instituto Nacional de Metrologia, Normalização e Qualidade Industrial - INMETRO. DOQ-CGCRE-008. Orientações sobre validação de métodos de ensaios químicos. Rio de Janeiro: INMETRO; 2010. 35p.
11. EURACHEM. The fitness for purpose of analytical methods, a laboratory guide to method validation and related topics. Teddington: LGC; 1998. 61p.
12. European Commission - EC. Commission decision 2002/657/EC of 12 August 2002. Implementing Council Directive 96/23/EC concerning performance of analytical methods and the interpretation of results. *Official Journal of the European Communities*; 2002, L221/8.
13. Thompson M, Ellison SLR, Wood R. Harmonized guidelines for single-laboratory validation of methods of analysis. *Pure Appl Chem*. 2002; 74:835-55.
14. Ríos A, Téllez H. Reliability of binary analytical responses. *Trends Anal Chem*. 2005; 24(6):509-15.

15. Valcárcel M, Cárdenas S. Modern qualitative analysis. *Trends Anal Chem*. 2005; 24(6):467.
16. Valcárcel M, Ríos A. Traceability in analytical chemistry. *Analyst*. 1995; 120:2291-7.
17. Cárdenas S, Valcárcel M. Analytical features in qualitative analysis. *Trends Anal Chem*. 2005; 24(6):477-87.
18. Ríos A, Barceló D, Buydens L, Cárdenas S, Heydorn K, Karlberg B et al. Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'. *Accred Qual Assur*. 2003; 8:68-77.
19. Unger-Heumann M. Strategy of analytical tests kits. *Fresenius' J Anal Chem*. 1996; 354(7-8):803-6.
20. Valcárcel M, Cárdenas S, Gallego M. Sample screening systems in analytical chemistry. *Trends Anal Chem*. 1999; 18(11):685-94.
21. Brasil. Ministério da Agricultura, Pecuária e Abastecimento/Secretaria de Defesa Agropecuária. Instrução normativa número 24, de 14 de julho de 2009. Diário Oficial [da] União. Brasília, DF, 22 jul 2009; Seção 1, p. 7-15.
22. European Commission – EC. Community Reference Laboratories residues (CRLs). Guidelines for the validation of screening methods for residues of veterinary medicines (initial validation and transfer). [acesso 2011 Jan 06]. Disponível em: [http://ec.europa.eu/food/food/chemicalsafety/residues/Guideline_Validation_Screening_en.pdf].
23. AOAC (AOAC Research Institute). AOAC® Performance Tested MethodsSM Program, Policies and procedures. [acesso 2011 Jan 20]. Disponível em: [www.aoc.org/testkits/Policies%20&%20Procedures.pdf].
24. Raugel P. Rapid food analysis and hygiene monitoring: kits, instruments, and systems. Berlin: Springer-Verlag; 1999.
25. CODEX ALIMENTARIUS. ALINORM 10/33/23. Joint FAO/WHO Food Standards Programme Codex Alimentarius Commission Thirty-third Session, Genebra, 2010 – Report of the thirty-fist session of the Codex Committee on methods of analysis and sampling, Budapeste, 2010.[acesso 2011 Jan 06]. Disponível em:[www.codexalimentarius.net/download/report/738/al33_23e.pdf].
26. Trullols E, Ruisánchez I, Rius FX, Huguet J. Validation of qualitative methods of analysis that use control samples. *Trends Anal Chem*. 2005; 24(6):516-24.
27. Milman BL, Konopelko LA. Uncertainty of Qualitative Chemical Analysis: General Methodology and Binary Test Methods. *J Anal Chem*. 2004; 59(12):1128-41.
28. Simonet BM, Ríos A, Valcárcel M. Unreliability of screening methods. *Anal Chim Acta*. 2004; 516:67-74.
29. Currie LA. Nomenclature in evaluation of analytical methods including detection and quantification capabilities. *Pure Appl Chem*. 1995; 67:1699-723.
30. Valcárcel M, Ríos A. Is traceability an exclusive property of analytical results? An extended approach to traceability in chemical analysis. *Fresenius' J Anal Chem*. 1997; 359: 473-5.
31. Langtonn SD, Chevennement R, Nagelkerke N, Lombard, B. Analysing collaborative trials for qualitative microbiological methods: concordance and concordance. *Int J Food Microbiol*. 2002; 75:175-81.
32. Gowik P. The validation of methods for regulatory purposes in the control of residues. *J Chromatogr A*. 2009; 1216:8051-8.
33. Armstrong N, Hibbert DB. An introduction to Bayesian methods for analyzing chemistry data. Part 1: An introduction to Bayesian theory and methods. *Chemom Intell Lab Syst*. 2009; 97:194-210.
34. Ellison SLR, Gregory S, Hardcastle WA. Quantifying uncertainty in qualitative analysis. *Analyst*. 1998; 123:1155-61.
35. Henderson AR. Assessment of clinical enzyme methodology: a probabilistic approach. *Clin Chim Acta*. 1997; 257:25-40.
36. Song R, Schlecht PC, Ashley K. Field screening test methods: performance criteria and performance characteristics. *J Hazardous Mater*. 2001; 83: 29-39.
37. Milman BL, Konopelko LA. Identification of chemical substances by testing and screening of hypotheses. *Fresenius' J Anal Chem*. 2000; 367:621-8.
38. Ellison SRL. Uncertainties in qualitative testing and analysis. *Accred Qual Assur*. 2000; 5:346-8.