

**SCHEILLA VITORINO CARVALHO DE SOUZA**

**PROCEDIMENTO PARA VALIDAÇÃO  
INTRALABORATORIAL DE MÉTODOS DE ENSAIO:  
DELINEAMENTO E APLICABILIDADE EM ANÁLISES  
DE ALIMENTOS**

**FACULDADE DE FARMÁCIA DA UFMG  
BELO HORIZONTE, MG**

**2007**

**SCHEILLA VITORINO CARVALHO DE SOUZA**

**PROCEDIMENTO PARA VALIDAÇÃO  
INTRALABORATORIAL DE MÉTODOS DE ENSAIO:  
DELINEAMENTO E APLICABILIDADE EM ANÁLISES  
DE ALIMENTOS**

Tese apresentada ao Programa de Pós-Graduação em Ciência de Alimentos da Faculdade de Farmácia da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do grau de Doutor em Ciência de Alimentos.

Orientador: Prof. Dr. Roberto Gonçalves Junqueira

**FACULDADE DE FARMÁCIA DA UFMG  
BELO HORIZONTE, MG**

**2007**

## **FOLHA DE APROVAÇÃO**

[A ser substituída pela versão definitiva após a defesa, no momento da reprodução e encadernação. Será fornecida pela Secretaria do Programa.]

*Dedico este trabalho ao Ricardo,  
pelo amor que me incentiva sempre,  
pelos valores e sonhos compartilhados.*

*E à Maria,  
um novo sentido em minha vida.*

## **AGRADECIMENTOS INSTITUCIONAIS**

Ao Programa de Pós-Graduação em Ciência de Alimentos (PPGCA) da Faculdade de Farmácia da Universidade Federal de Minas Gerais (FAFAR/UFMG).

Às instituições parceiras *Central Science Laboratory* do *Department for Environment Food and Rural Affairs* (CSL/DEFRA, York, Reino Unido), Gerência da Rede Laboratorial do Instituto Mineiro de Agropecuária (GRL/IMA, Contagem - MG, Brasil) e Laboratório Nacional Agropecuário - Minas Gerais do Ministério da Agricultura, Pecuária e Abastecimento (LANAGRO-MG/MAPA, Pedro Leopoldo - MG, Brasil).

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

## **AGRADECIMENTOS PESSOAIS**

Ao meu orientador, Roberto Gonçalves Junqueira, admirável pelo caráter, orientação técnica e metodológica, pela amizade, parceria e constante incentivo durante todo o percurso deste trabalho, nunca faltando de sua parte apoio nos momentos difíceis.

Ao Prof. John Gilbert, co-orientador durante o período no CSL/DEFRA, pelo apoio incondicional ao trabalho e exemplo profissional.

Aos amigos e colaboradores, Richard Ginn (CSL/DEFRA), Carlos Temóteo Pinto, João Carlos Teodoro, Josefa Abucáter Lima, Juarez Fabiano Alckmin Filho e Ronaldo Linaris Sanches (LANAGRO-MG/MAPA), pelo engrandecimento deste trabalho com suas competências específicas.

Aos amigos da DVLB/IMA, Eliane Hooper Amaral, Alexandre Augusto Soares, Isis Roseclaire de Carvalho, Júnea Guzella Ramos Martinelli e Sueli Roberto Goulart, pelo estímulo e fornecimento dos dados experimentais para a publicação do procedimento de avaliação da linearidade, parte do Capítulo I desta tese.

Aos demais funcionários e amigos das instituições parceiras, em especial seus diretores John Gilbert (CSL/DEFRA), José Marques Neto (DVLB/IMA), Ricardo Aurélio Pinto Nascimento e Tomaz de Aquino Porfírio (LANAGRO-MG/MAPA), por oferecerem as realidades que empiricamente sustentaram o desenvolvimento dos argumentos científicos deste trabalho.

Aos professores do Departamento de Alimentos da FAFAR/UFMG, especialmente Maria Beatriz Abreu Glória e José Virgílio Coelho, pelos ensinamentos que sempre levarei comigo e pelo total apoio nesta que é a mais importante conquista de minha formação acadêmica. Ao professor David Lee Nelson pela correção dos textos em inglês.

Aos membros da comissão examinadora, Alice Momoyo Ata Sakuma, José Virgílio Coelho, Olguita Geralda Ferreira Rocha e Shirley de Mello Pereira Abrantes, por aceitarem participar da banca e pelas sugestões que contribuíram para a qualidade desta tese.

Aos colegas do PPGCA da FAFAR/UFMG pela amizade. Em especial aos meus sempre mestres Gecernir Cólen e Lúcia Helena Esteves dos Santos Laboissière.

Aos funcionários da biblioteca, da secretaria do Departamento de Alimentos e da secretaria do PPGCA da FAFAR/UFMG pelo incentivo e especial atenção.

Aos meus pais pelo amor irrestrito e pelos constantes ensinamentos de valorização do conhecimento. À minha irmã, sobrinhas e familiares pela força constante e pelo entendimento de várias horas ausente.

A Deus pela vida e por todas as oportunidades dela advindas.

## SUMÁRIO

LISTA DE TABELAS.....	13
LISTA DE FIGURAS.....	16
RESUMO.....	19
ABSTRACT.....	20
INTRODUÇÃO.....	21
OBJETIVOS.....	24
REVISÃO DA LITERATURA.....	25
1 VALIDAÇÃO DE MÉTODOS.....	25
1.1 DEFINIÇÕES.....	25
1.2 VALIDAÇÃO NO CONTEXTO ANALÍTICO.....	26
1.3 VALIDAÇÃO NO CONTEXTO DE SISTEMAS DE GESTÃO.....	28
1.4 QUANDO VALIDAR.....	30
1.5 SELEÇÃO DE MÉTODOS.....	32
1.6 COMO VALIDAR.....	33
1.7 PROCESSOS INTRA E INTERLABORATORIAIS.....	34
1.7.1 Estudo colaborativo.....	34
1.7.2 Ensaio de proficiência.....	36
1.7.3 Controle de qualidade interno.....	37
1.7.4 Validação intralaboratorial.....	38
1.7.5 Validação comparativa.....	39
1.8 ABRANGÊNCIA DA VALIDAÇÃO.....	40
2 PROCESSOS DE VALIDAÇÃO INTRALABORATORIAIS.....	43
2.1 PRINCÍPIOS BÁSICOS.....	44
2.1.1 Especificação e escopo.....	44
2.1.2 Representatividade.....	44
2.1.3 Verificação de premissas.....	45
2.1.4 Fontes de erros.....	45
2.2 PARÂMETROS DE DESEMPENHO.....	46
2.2.1 Termos e definições.....	46
2.2.2 Principais parâmetros em processos de validação.....	66
2.2.3 Procedimentos para avaliação.....	68
CAPÍTULO I - PROCEDIMENTO PARA VALIDAÇÃO INTRALABORATORIAL DE MÉTODOS DE ENSAIO QUANTITATIVOS E QUALITATIVOS EM ANÁLISES DE	



ALIMENTOS.....	95
RESUMO.....	95
ABSTRACT.....	96
1    INTRODUÇÃO.....	96
2    TEORIA.....	98
2.1  LINEARIDADE.....	98
2.1.1 Delineamento experimental.....	98
2.1.2 O modelo.....	100
2.1.3 Premissas.....	102
2.1.4 Coeficientes de determinação e correlação.....	108
2.2  EFEITOS DE MATRIZ.....	109
2.2.1 Delineamento experimental.....	109
2.2.2 Análise de dados.....	110
2.3  SELETIVIDADE.....	111
2.3.1 Delineamento experimental.....	111
2.3.2 Análise de dados.....	112
2.4  EXATIDÃO E PRECISÃO.....	113
2.4.1 Delineamento experimental.....	113
2.4.2 Análise de dados.....	115
2.5  LIMITES.....	119
2.5.1 Limites de detecção e de quantificação do método.....	119
2.5.2 Limite de decisão e capacidade de detecção.....	120
3    PROCEDIMENTO PROPOSTO.....	122
3.1  MÉTODOS QUANTITATIVOS.....	124
3.1.1 Linearidade.....	124
3.1.2 Efeitos de matriz.....	138
3.1.3 Seletividade, exatidão, precisão e limites.....	141
3.2  MÉTODOS QUALITATIVOS.....	161
3.2.1 Identificação do analito.....	161
3.2.2 Seletividade e limite de detecção.....	161
4    CONCLUSÕES.....	162
5    CONSIDERAÇÕES.....	162
CAPÍTULO II - VALIDAÇÃO INTRALABORATORIAL DE MÉTODO QUANTITATIVO PARA DETERMINAÇÃO DE SEMICARBAZIDA EM <i>BABY FOOD</i> POR CROMATOGRAFIA LÍQUIDA DE ALTA EFICIÊNCIA ACOPLADA A ESPECTROMETRIA DE MASSAS.....	163

	RESUMO.....	163
	ABSTRACT.....	163
1	INTRODUÇÃO.....	164
2	MATERIAL E MÉTODOS.....	167
2.1	MATERIAL.....	167
2.2	EQUIPAMENTOS.....	168
2.3	REAGENTES E PADRÕES.....	168
2.4	SOLUÇÕES.....	169
2.4.1	Solução padrão de SEM.....	169
2.4.2	Solução padrão interno de 1,2-[ <sup>15</sup> N <sub>2</sub> , <sup>13</sup> C] SEM.....	169
2.4.3	Solução de derivação: 2-NBA 10 mg/mL.....	169
2.4.4	Solução de extração: ácido clorídrico 0,2 mol/L.....	170
2.4.5	Soluções de neutralização: ortofosfato monoácido de potássio 0,2 mol/L e hidróxido de sódio 2 mol/L.....	170
2.4.6	Solução de fase móvel: formato de amônio 5 mmol/L.....	170
2.5	AMOSTRAS.....	170
2.6	MÉTODO DE ENSAIO.....	170
2.6.1	Princípio.....	171
2.6.2	Cuidados.....	171
2.6.3	Extração.....	171
2.6.4	Acondicionamento dos cartuchos de extração em fase sólida.....	171
2.6.5	Purificação.....	172
2.6.6	Separação, detecção e quantificação.....	172
2.7	VALIDAÇÃO.....	173
2.7.1	Linearidade e efeitos de matriz.....	174
2.7.2	Seletividade, exatidão, precisão e limites.....	176
3	RESULTADOS E DISCUSSÃO.....	177
4	CONCLUSÕES.....	189
5	CONSIDERAÇÕES.....	190
CAPÍTULO III - VALIDAÇÃO INTRALABORATORIAL DE MÉTODO QUANTITATIVO PARA DETERMINAÇÃO DE ARSÊNIO EM FÍGADO BOVINO POR ESPECTROFOTOMETRIA DE ABSORÇÃO ATÔMICA COM GERAÇÃO DE HIDRETOS E SISTEMA DE INJEÇÃO EM FLUXO.....		
	RESUMO.....	191
	ABSTRACT.....	191
1	INTRODUÇÃO.....	192
2	MATERIAL E MÉTODOS.....	194

2.1	MATERIAL.....	194
2.2	EQUIPAMENTOS.....	195
2.3	REAGENTES E PADRÕES.....	196
2.4	SOLUÇÕES.....	196
2.4.1	Soluções padrão de arsênio.....	196
2.4.2	Soluções de ácido clorídrico.....	196
2.4.3	Solução de ácido nítrico 50 % (v/v).....	196
2.4.4	Solução de borohidreto de sódio 0,2% (m/v).....	196
2.4.5	Solução de hidróxido de sódio 12,5 % (m/v).....	197
2.4.6	Solução de nitrato de magnésio 50 % (m/v).....	197
2.4.7	Solução pré-redutora: iodeto de potássio 5 % (m/v) e ácido ascórbico 5 % (m/v).....	197
2.5	AMOSTRAS.....	197
2.6	MÉTODO DE ENSAIO.....	197
2.6.1	Princípio.....	198
2.6.2	Cuidados.....	198
2.6.3	Digestão, calcinação e pré-redução.....	198
2.6.4	Preparo da curva de calibração.....	199
2.6.5	Detecção e quantificação.....	199
2.7	VALIDAÇÃO.....	200
2.7.1	Linearidade.....	200
2.7.2	Efeitos de matriz.....	201
2.7.3	Seletividade, exatidão, precisão e limites.....	201
3	RESULTADOS E DISCUSSÃO.....	203
4	CONCLUSÕES.....	216
5	CONSIDERAÇÕES.....	217
CAPÍTULO IV - VALIDAÇÃO INTRALABORATORIAL DE MÉTODO QUANTITATIVO PARA DETERMINAÇÃO MÚLTIPLA DE RESÍDUOS DE AVERMECTINAS EM LEITE BOVINO POR CROMATIOGRAFIA LÍQUIDA DE ALTA EFICIÊNCIA COM DETECÇÃO DE FLUORESCÊNCIA.....		
	RESUMO.....	218
	ABSTRACT.....	218
1	INTRODUÇÃO.....	219
2	MATERIAL E MÉTODOS.....	222
2.1	MATERIAL.....	222
2.2	EQUIPAMENTOS.....	223
2.3	REAGENTES E PADRÕES.....	223

2.4	SOLUÇÕES.....	224
2.4.1	Solução padrão de ABA.....	224
2.4.2	Soluções padrão de DOR.....	224
2.4.3	Soluções padrão de EPR.....	224
2.4.4	Solução padrão de IVE.....	224
2.4.5	Solução padrão de adição <i>pool</i> de ABA, DOR, EPR e IVE.....	224
2.4.6	Solução de derivação.....	225
2.4.7	Solução de extração.....	225
2.5	AMOSTRAS.....	225
2.6	MÉTODO DE ENSAIO.....	225
2.6.1	Princípio.....	225
2.6.2	Cuidados.....	225
2.6.3	Extração.....	226
2.6.4	Acondicionamento dos cartuchos de extração em fase sólida.....	226
2.6.5	Purificação.....	226
2.6.6	Preparo da curva de calibração.....	226
2.6.7	Derivação.....	226
2.6.8	Separação, detecção e quantificação.....	227
2.7	VALIDAÇÃO.....	227
2.7.1	Linearidade.....	227
2.7.2	Efeitos de matriz.....	228
2.7.3	Seletividade, exatidão, precisão e limites.....	229
3	RESULTADOS E DISCUSSÃO.....	231
4	CONCLUSÕES.....	247
5	CONSIDERAÇÕES.....	247
CAPÍTULO V - VALIDAÇÃO INTRALABORATORIAL DE MÉTODO QUALITATIVO PARA DETERMINAÇÃO DE SUBPRODUTOS DE ORIGEM ANIMAL EM RAÇÕES PARA RUMINANTES POR MICROSCOPIA.....		249
	RESUMO.....	249
	ABSTRACT.....	249
1	INTRODUÇÃO.....	250
2	MATERIAL E MÉTODOS.....	253
2.1	MATERIAL.....	253
2.2	EQUIPAMENTOS.....	253
2.3	REAGENTES E PADRÕES.....	254
2.4	SOLUÇÕES.....	255

2.4.1	Soluções de éter de petróleo:clorofórmio (10/90; 20/80; 30/70; 40/60 e 50/50, v/v).....	255
2.4.2	Soluções para provas químicas qualitativas.....	255
2.5	AMOSTRAS.....	256
2.6	MÉTODO DE ENSAIO.....	256
2.6.1	Princípio.....	257
2.6.2	Cuidados.....	257
2.6.3	Preparo das amostras.....	257
2.6.4	Separação.....	257
2.6.5	Detecção e identificação.....	258
2.7	OTIMIZAÇÃO E VALIDAÇÃO.....	258
2.7.1	Ensaio com materiais de referência.....	258
2.7.2	Ensaio com amostras brancas e adicionadas.....	259
3	RESULTADOS E DISCUSSÃO.....	259
4	CONCLUSÕES.....	268
5	CONSIDERAÇÕES.....	268
	CONCLUSÕES INTEGRADAS.....	269
	SUGESTÕES.....	271
	REFERÊNCIAS BIBLIOGRÁFICAS.....	272

## LISTA DE TABELAS

1	Panorama das principais referências relacionadas a processos de validação intra e interlaboratoriais (estudos colaborativos) e respectivos órgãos emissores.....	35
2	Variação das extensões de processos de validação em função do tipo de método (quantitativo ou qualitativo).....	42
3	Parâmetros de desempenho e respectivas definições apresentados por GREEN (1996).....	48
4	Parâmetros de desempenho e respectivas definições apresentados por EAL (1997).....	49
5	Parâmetros de desempenho e respectivas definições apresentados por NATA (1997).....	50
6	Parâmetros de desempenho e respectivas definições apresentados por AOAC (1998 b).....	51
7	Parâmetros de desempenho e respectivas definições apresentados por BRUCE, MINKKINEN & RIEKKOLA (1998).....	52
8	Parâmetros de desempenho e respectivas definições apresentados por EURACHEM (1998).....	53
9	Parâmetros de desempenho e respectivas definições apresentados por HUBER (1998).....	56
10	Parâmetros de desempenho e respectivas definições apresentados por EC (2002).....	57
11	Parâmetros de desempenho e respectivas definições apresentados por THOMPSON, ELLISON & WOOD (2002).....	58
12	Parâmetros de desempenho e respectivas definições apresentados por INMETRO (2003).....	59
13	Parâmetros de desempenho e respectivas definições apresentados por SHABIR (2003).....	60
14	Parâmetros de desempenho e respectivas definições apresentados por TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004).....	61
15	Recomendações para avaliação da linearidade na literatura.....	69
16	Recomendações para avaliação de efeitos de matriz na literatura.....	73
17	Recomendações para avaliação da seletividade na literatura.....	75
18	Recomendações para avaliação da exatidão na literatura.....	77
19	Recomendações para avaliação da precisão na literatura.....	81
20	Recomendações para avaliação do limite de detecção na literatura.....	85
21	Recomendações para avaliação do limite de quantificação na literatura.....	88
CAPÍTULO I		
I.1	Análise de variância para significância da regressão e do desvio da linearidade.....	137
I.2	Valores críticos tabelados para as estatísticas de Grubbs em um nível de significância de $\alpha = 0,05$ .....	148
I.3	Critérios de aceitabilidade para exatidão em função do nível de concentração.....	151
I.4	Análise de variância para estimativa da estatística F de Levene.....	154

I.5	Análise de variância para estimativa da precisão sob condições de repetitividade e reprodutibilidade parcial.....	155
I.6	Expressão das variâncias de repetitividade e reprodutibilidade parcial.....	155
I.7	Critérios de aceitabilidade para precisão, sob condições de reprodutibilidade parcial, em função do nível de concentração.....	157
<b>CAPÍTULO II</b>		
II.1	Condições para MRM.....	173
II.2	Volumes das soluções de adição de SEM para curvas de 2 ng/mL a 80 ng/mL em solventes e matrizes.....	174
II.3	Volumes das soluções de adição de SEM para curvas de 0,1 ng/mL a 1,0 ng/mL em solventes e matrizes.....	175
II.4	Representação da homogeneidade das variâncias dos resíduos da regressão das curvas de SEM (2 ng/mL a 80 ng/mL) em solventes e matrizes pelo teste de Levene modificado.....	182
II.5	Estatísticas da análise de variância da regressão, incluindo teste de desvio da linearidade, para curvas de SEM (2 ng/mL a 80 ng/mL) em solventes e matrizes.....	182
II.6	Estatísticas da análise de variância da regressão, incluindo teste de desvio de linearidade, para curvas de SEM (0,1 ng/mL a 1,0 ng/mL) em solventes e matrizes.....	184
II.7	Comparações entre as interseções e inclinações das curvas de matriz com as da curva de solventes para SEM (2 ng/mL a 80 ng/mL).....	185
II.8	Comparações entre as interseções e inclinações das curvas de matriz com as da curva de solventes para SEM (0,1 ng/mL a 1,0 ng/mL).....	185
II.9	Médias de recuperação aparente e desvios padrão relativos, sob condições de repetitividade, obtidos para amostras de purê de maçã, refeição de carne e pudim de arroz adicionadas de SEM em diferentes níveis de concentração.....	187
<b>CAPÍTULO III</b>		
III.1	Estatísticas da análise de variância da regressão, incluindo teste de desvio de linearidade, para as três curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes.....	210
III.2	Estatísticas da análise de variância da regressão, incluindo teste de desvio da linearidade, para as curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes e matriz.....	212
III.3	Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente de arsênio.....	215
III.4	Médias de recuperação aparente e desvios padrão relativos, sob condições de repetitividade e reprodutibilidade parcial, obtidos para amostras de fígado bovino adicionadas de arsênio em diferentes níveis de concentração.....	215
<b>CAPÍTULO IV</b>		
IV.1	Avaliação da linearidade para as três curvas de ABA (25 ng/mL a 150 ng/mL) em solvente.....	234
IV.2	Avaliação da linearidade para as três curvas de DOR (25 ng/mL a 150 ng/mL) em solvente.....	235
IV.3	Avaliação da linearidade para as três curvas de EPR (25 ng/mL a 150 ng/mL) em solvente.....	236
IV.4	Avaliação da linearidade para as três curvas de IVE (25 ng/mL a 150 ng/mL) em solvente.....	237

IV.5	Avaliação da linearidade para curvas de ABA e DOR (25 ng/mL a 150 ng/mL) em solvente e matriz obtidas nos ensaios de avaliação de efeitos de matriz.....	239
IV.6	Avaliação da linearidade para curvas de EPR e IVE (25 ng/mL a 150 ng/mL) em solvente e matriz obtidas nos ensaios de avaliação de efeitos de matriz.....	240
IV.7	Comparações entre as interseções e inclinações das curvas de matriz com as das curvas de solventes para ABA, DOR, EPR e IVE (25 ng/mL a 150 ng/mL).....	240
IV.8	Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente de ABA.....	243
IV.9	Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente de DOR.....	243
IV.10	Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente de EPR.....	244
IV.11	Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente de IVE.....	244
IV.12	Médias de recuperação aparente e desvios padrão relativos, sob condições de repetitividade e reprodutibilidade parcial, obtidos para amostras de leite bovino adicionadas de ABA e DOR em diferentes níveis de concentração.....	245
IV.13	Médias de recuperação aparente e desvios padrão relativos, sob condições de repetitividade e reprodutibilidade parcial, obtidos para amostras de leite bovino adicionadas de EPR e IVE em diferentes níveis de concentração.....	245
<b>CAPÍTULO V</b>		
V.1	Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de porcentagem de concentração.....	261
V.2	Resultados detectados e não detectados obtidos para amostras de ração para ruminantes brancas e adicionadas de subprodutos de origem animal em diferentes níveis de concentração.....	267



## LISTA DE FIGURAS

1	Relação entre as diferentes etapas do processo analítico (BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998 - adaptados).....	29
2	Diferentes níveis e ferramentas de garantia da qualidade na acreditação de sistemas de gestão de laboratórios (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004 - adaptado).....	31
3	Composição do erro de um resultado analítico, relacionada à validação intralaboratorial do método de ensaio (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004 - adaptado).....	47
4	Combinação da exatidão (erros sistemáticos) e da precisão (erros aleatórios) para indicação de acurácia (RSC, 2003 - adaptado).....	64
CAPÍTULO I		
I.1	Experimentos para validação de métodos quantitativos, incluindo formas de validação e abrangência em termos de parâmetros de desempenho avaliados.....	123
I.2	Experimentos para validação de métodos qualitativos, incluindo formas de validação e abrangência em termos de parâmetros de desempenho avaliados.....	124
I.3	Avaliação da linearidade: estimativa dos parâmetros pelo MMQO, inspeção visual dos dados e gráficos e tratamento de valores extremos.....	126
I.4	Avaliação da linearidade: testes das premissas para uso apropriado do MMQO.....	129
I.5	Decomposição da soma de quadrados total para estimativa da significância da regressão por análise de variância.....	135
I.6	Decomposição da soma de quadrados dos resíduos da regressão para estimativa da significância do desvio da linearidade por análise de variância.....	135
I.7	Decomposição auxiliar da soma de quadrados total.....	136
I.8	Avaliação dos efeitos de matriz.....	139
I.9	Delineamento experimental para avaliação da precisão sob condições de repetitividade.....	143
I.10	Delineamento experimental para avaliação da precisão sob condições de repetitividade e reprodutibilidade parcial.....	144
I.11	Avaliação da seletividade, exatidão, precisão e limites: estimativa da concentração de cada amostra, cálculo da recuperação aparente, inspeção visual dos dados e tratamento de valores extremos.....	146
I.12	Avaliação da seletividade: análise das recuperações aparentes, registros de detecção e quantificação e comparação com critérios de aceitabilidade.....	150
I.13	Avaliação da exatidão: estimativa das médias de recuperação aparente e comparação com critérios de aceitabilidade.....	151
I.14	Avaliação da precisão: testes das premissas para emprego da análise de variância, estimativa dos desvios padrão de repetitividade e reprodutibilidade parcial e comparação com critérios de aceitabilidade.....	153
I.15	Avaliação do limite de quantificação: estimativa das médias de recuperação aparente, desvio padrão de repetitividade e reprodutibilidade parcial e comparação com critérios de aceitabilidade.....	158
I.16	Avaliação do limite de detecção: estimativa das médias dos sinais obtidos	

	para o analito nas amostras adicionadas, nas amostra brancas ou nos brancos da curva e comparação com critérios de aceitabilidade.....	159
<b>CAPÍTULO II</b>		
II.1	Estruturas químicas da nitrofurazona, azodicarbonamida e SEM.....	165
II.2	Gráficos exploratórios dos resíduos da regressão das curvas de SEM (2 ng/mL a 80 ng/mL) em solventes e matrizes, com indicação dos respectivos valores extremos diagnosticados pelo teste de resíduos padronizados Jackknife.....	178
II.3	Gráficos de probabilidade normal das curvas de SEM (2 ng/mL a 80 ng/mL) em solventes e matrizes.....	180
II.4	Gráficos de Durbin-Watson das curvas de SEM (2 ng/mL a 80 ng/mL) em solventes e matrizes.....	181
II.5	Curvas de calibração do analito nos solventes e nas matrizes obtidas para SEM nos ensaios de avaliação de efeitos de matriz, na faixa de 2 ng/mL a 80 ng/mL, com respectivas equações e coeficientes de determinação.....	183
II.6	Estrutura química do derivado 2-NBA de SEM e cromatogramas típicos de MRM obtidos por CLAE-EM-EM.....	186
II.7	Distribuição dos resultados individuais de recuperação aparente obtidos para amostras adicionadas entre 0,10 µg/kg e 0,50 µg/kg, incluindo os níveis correspondentes aos limites de detecção (0,10 µg/kg) e quantificação (0,25 µg/kg).....	188
<b>CAPÍTULO III</b>		
III.1	Gráficos exploratórios dos resíduos da regressão das três curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes, com indicação dos respectivos valores extremos diagnosticados pelo teste de resíduos padronizados Jackknife.....	204
III.2	Gráficos de probabilidade normal das três curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes.....	206
III.3	Representação da homogeneidade das variâncias das três curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes pelo teste de Levene modificado.....	207
III.4	Gráficos de Durbin-Watson das três curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes.....	208
III.5	Curvas de calibração do arsênio nos solventes obtidas nos ensaios de avaliação de linearidade, na faixa de 0,5 ng/mL a 8 ng/mL, com respectivas equações e coeficientes de determinação.....	209
III.6	Curvas de calibração do analito nos solventes e na matriz obtidas para arsênio nos ensaios de avaliação de efeitos de matriz, na faixa de 0,5 ng/mL a 8 ng/mL, com respectivas equações e coeficientes de determinação.....	211
III.7	Registros gráficos típicos de leituras de absorvância.....	214
III.8	Distribuição dos resultados individuais de recuperação aparente obtidos para amostras adicionadas entre 15 µg/kg e 120 µg/kg, incluindo os níveis correspondentes aos limites de detecção (15 µg/kg) e quantificação (30 µg/kg).....	216
<b>CAPÍTULO IV</b>		
IV.1	Estrutura química das avermectinas.....	220
IV.2	Delineamento da distribuição de amostras de leite bovino nos ensaios de avaliação da exatidão e precisão do método para determinação de	

	avermectinas em leite.....	229
IV.3	Gráficos exploratórios dos resíduos da regressão das três curvas de ABA e DOR (25 ng/mL a 150 ng/mL) em solvente, com indicação dos respectivos valores extremos diagnosticados pelo teste de resíduos padronizados Jackknife.....	232
IV.4	Gráficos exploratórios dos resíduos da regressão das três curvas de EPR e IVE (25 ng/mL a 150 ng/mL) em solvente, com indicação dos respectivos valores extremos diagnosticados pelo teste de resíduos padronizados Jackknife.....	233
IV.5	Curvas de calibração dos analitos no solvente e na matriz obtidas para ABA, DOR, EPR e IVE nos ensaios de avaliação de efeitos de matriz, na faixa de 25 ng/mL a 150 ng/mL, com respectivas equações e coeficientes de determinação.....	238
IV.6	Estrutura química do derivado fluorescente das avermectinas e cromatogramas típicos obtidos por CLAE-FL.....	242
IV.7	Distribuição dos resultados individuais de recuperação aparente obtidos para amostras adicionadas entre 5 µg/L e 20 µg/L, incluindo os níveis correspondentes aos limites de detecção (5 µg/L) e quantificação (10 µg/L).....	246
<b>CAPÍTULO V</b>		
V.1	Perfil de porcentagem de concentração dos analitos nas diferentes misturas de solventes empregadas na etapa de separação.....	260
V.2	Fragmentos de farinha de carne e ossos.....	263
V.3	Fragmentos de farinha de ossos não calcinados.....	263
V.4	Fragmentos de farinha de sangue.....	264
V.5	Fragmentos de farinha de vísceras.....	264
V.6	Fragmentos de farinha de pena.....	265
V.7	Fragmentos de farinha de peixe.....	266

## RESUMO

Um procedimento detalhado para validação intralaboratorial de métodos quantitativos e qualitativos foi desenvolvido para aplicação em análises de alimentos. O procedimento teve como base fundamentos teóricos, considerando a eficiência, praticidade e facilidade do delineamento experimental e das ferramentas de análise de dados. A forma de validação direta foi adotada com uso de padrões ou materiais de referência. Linearidade, efeitos de matriz, seletividade, exatidão, precisão sob condições de repetitividade e reprodutibilidade parcial, limite de detecção, limite de quantificação, limite de decisão e capacidade de detecção foram os parâmetros de desempenho estudados para métodos quantitativos, enquanto seletividade e limite de detecção foram avaliados para métodos qualitativos. A aplicabilidade do procedimento proposto foi demonstrada na validação de três métodos quantitativos (determinação de semicarbazida em *baby food* por cromatografia líquida de alta eficiência acoplada a espectrometria de massas, arsênio em fígado bovino por espectrofotometria de absorção atômica com geração de hidretos e avermectinas em leite bovino por cromatografia líquida de alta eficiência com detector de fluorescência) e de um método qualitativo (determinação de subprodutos de origem animal em rações para ruminantes por microscopia). Todos os métodos validados apresentaram adequação para seus respectivos propósitos de uso.

Palavras-chave: validação intralaboratorial; quimiometria; método quantitativo; método qualitativo; análises de alimentos; adequação para uso.

## ABSTRACT

A detailed procedure for single-laboratory validation of quantitative and qualitative methods was developed to be applied to food analysis. The procedure was based on theoretical background, considering the efficiency, practicability and easiness of the experimental design and the statistical tools for analysis of the data. The direct validation was adopted, using standard and reference materials. Linearity, matrix effects, selectivity, trueness, precision under repeatability and within-reproducibility conditions, limit of detection, limit of quantification, decision limit and detection capability were the parameters studied for quantitative methods, while selectivity and detection limit were evaluated for qualitative methods. The applicability of the proposed procedure was demonstrated for three quantitative methods (semicarbazide in baby food by high performance liquid chromatography tandem mass spectrometry, arsenic in bovine liver by hydride generation atomic absorption spectrometry and avermectins in bovine milk by high performance liquid chromatography with fluorescence detection) and one qualitative method (animal meals in ruminant feeds by microscopy). All the validated method showed fitness for the respective purposes.

Key words: single-laboratory validation; chemometry; quantitative method; qualitative method; food analysis, fitness for purpose.

## INTRODUÇÃO

Milhões de medidas analíticas são feitas diariamente em vários laboratórios no mundo, dando suporte a decisões nos mais diversos segmentos da sociedade (EURACHEM, 1998). Diante da crescente demanda de confiabilidade e comparabilidade destas medições, uma forte infra-estrutura metrológica está sendo implementada, com foco no reconhecimento mútuo e na eliminação de barreiras técnicas entre os países.

Neste contexto, é internacionalmente reconhecido que os laboratórios devam tomar medidas apropriadas para assegurar que são capazes de produzir resultados com a qualidade requerida, como o uso de métodos de ensaio validados, o uso de procedimentos de controle de qualidade interno, a participação em ensaios de proficiência e a acreditação (THOMPSON, ELLISON & WOOD, 2002). Vale ressaltar que os requisitos para acreditação pela norma ISO/IEC 17025 abrangem todas as medidas consideradas acima (ISO, 2005 b), as quais ainda podem ser necessárias por questões legais, regulamentares ou de certificação.

Durante muito tempo, a seleção e o desenvolvimento de métodos de ensaio foram considerados assuntos de grande importância para os laboratórios, com freqüente negligência sobre as condições de aplicação prática dos mesmos. Possivelmente, esta ênfase aconteceu devido ao fato de que a maioria das organizações, governamentais ou de padronização, desenvolvia métodos e incorporava-os às legislações ou padrões internacionais, mas não utilizava nenhum mecanismo para verificar quão bem os métodos estavam sendo aplicados. Felizmente, tal panorama tem sido suplantado pela necessidade de demonstração de sucesso na aplicação de métodos de ensaio (WOOD, 1999). Em 1992, KING considerou a existência de evidências de que cerca de 30 % de todas as medidas químicas não estavam ajustadas aos seus propósitos de uso, o que demonstrava uma necessidade urgente em confirmar a integridade de métodos de ensaio para as diversas técnicas e áreas de aplicação (BRUCE, MINKKINEN & RIEKKOLA, 1998). Esta confirmação, por exame e evidência objetiva, de que os requisitos para uso pretendido de um método são atendidos, constitui a validação (ISO, 2005 a).

O tema validação de métodos de ensaio tem recebido considerável atenção em literaturas científicas, comitês industriais e agências regulamentadoras, com grande procura por parte dos laboratórios e organismos acreditadores por protocolos para planejamento e controle destes processos. No entanto, a diversidade de áreas de aplicação, de parâmetros de desempenho a serem avaliados, bem como de terminologia e definições, têm dificultado a harmonização dos processos de validação em termos de

determinação de indicadores de qualidade dos métodos de ensaio e das respectivas modalidades de cálculos associadas a estes indicadores (ANTIGNAC et al., 2003).

Especificamente na área de análises de alimentos, a validação de métodos está estritamente relacionada com segurança alimentar e comércio internacional. Os resultados provenientes dos métodos de ensaio muitas vezes subsidiam a formulação de políticas de saúde pública para melhoramento da qualidade dos alimentos disponibilizados à população e adequação do país, do ponto de vista sanitário, às regras do comércio internacional (BRASIL, 2006).

Validações por procedimentos interlaboratoriais estão bem estabelecidas na área de análises de alimentos por meio de estudos colaborativos, como os organizados nos moldes estabelecidos pela *Association of Official Analytical Chemists (AOAC International)*, *International Standards Organization (ISO)* e *International Union of Pure and Applied Chemistry (IUPAC)* (YODEN & STEINER, 1975; ISO, 1994 a,b,c,d,e; HORWITZ, 1995; ISO, 1998; AOAC, 2002). Entretanto, a velocidade nos avanços científicos e tecnológicos, associada à crescente necessidade de garantia de qualidade, gera dificuldades para que os estudos colaborativos acompanhem as demandas, tanto por questões técnicas e organizacionais, quanto infra-estruturais e financeiras. Outra limitação relativa às validações interlaboratoriais está no fato de que estas determinam a exatidão e a precisão dos métodos, mas não consideram outros parâmetros de desempenho importantes como linearidade, faixa de trabalho, efeitos de matriz, seletividade, limite de detecção e limite de quantificação.

Frente aos fatores limitantes das validações por estudos colaborativos, procedimentos de validação intralaboratoriais (*single-laboratory* ou *in-house validation*) têm sido considerados apropriados quando se deseja avaliar os parâmetros de desempenho de um método antes de formalizar a participação em processos interlaboratoriais, os quais envolvem um alto custo. Tais procedimentos constituem também uma alternativa internacionalmente reconhecida quando não há disponibilidade de procedimentos interlaboratoriais ou quando estes procedimentos não são aplicáveis (THOMPSON, ELLISON & WOOD, 2002). É importante observar que, mesmo quando o laboratório opta por implantar em sua rotina um método que já tenha sido objeto de estudo colaborativo, permanece a necessidade de demonstrar a capacidade do laboratório em aplicar adequadamente o referido método por meio de procedimentos de validação intralaboratoriais.

Enquanto a necessidade de métodos de ensaio validados é clara, os mecanismos para desenvolvimento de uma validação intralaboratorial adequada não estão bem

definidos, nem acordados internacionalmente. Apesar dos esforços de organismos internacionais pela harmonização destes procedimentos (THOMPSON, ELLISON & WOOD, 2002), questões de interesse comum persistem: quais parâmetros de desempenho devem ser estudados, que procedimentos específicos devem ser utilizados para avaliar cada parâmetro e quais critérios de aceitabilidade devem ser definidos para cada parâmetro em particular (JENKE, 1996; HILL & REYNOLDS, 1999).

A maior parte das publicações que tratam sobre validações intralaboratoriais de métodos de ensaio limita-se às definições dos principais parâmetros de desempenho e a roteiros genéricos de procedimentos de validação (GREEN, 1996; ICH, 1996; EAL, 1997; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003), provavelmente em decorrência da ampla diversidade de técnicas e áreas de aplicação. Embora exista consenso sobre alguns tipos de parâmetros que devem ser avaliados, há grande diversidade sobre como os experimentos de validação são delineados e sobre as respectivas estatísticas aplicáveis (GREEN, 1996), sendo observadas, freqüentemente, recomendações contraditórias ou mesmo equivocadas. Na maioria das vezes temas fundamentais são omitidos, como a necessidade de verificação de premissas relacionadas aos testes estatísticos adotados.

Isto tem resultado em grande confusão e dificuldade por parte dos laboratórios interessados em validar seus métodos de ensaio, por procedimentos intralaboratoriais, que não dispõem de referências adequadas sobre como fazê-lo. Certamente, o uso indiscriminado das referências existentes também pode representar importante fonte de erro nos processos de validação e aplicação dos métodos de ensaio na área de análises de alimentos.

Assim, a conjuntura apresentada estimulou a elaboração desta tese, com enfoque no desenvolvimento de um procedimento para validações intralaboratoriais embasado em fundamentos teóricos e contendo todo o detalhamento necessário para uma avaliação segura e consistente de métodos de ensaio, o qual tivesse aplicabilidade demonstrada em diferentes escopos quantitativos e qualitativos na área de análises de alimentos.



## OBJETIVOS

O objetivo geral deste trabalho foi desenvolver um procedimento detalhado para validação intralaboratorial de métodos de ensaio quantitativos e qualitativos, embasado em fundamentos teóricos, com o detalhamento necessário para avaliações eficazes dos métodos e aplicável em análises de alimentos.

Os objetivos específicos foram:

- realizar levantamento, comparação e análise crítica dos procedimentos descritos na literatura para validação intralaboratorial de métodos;
- delinear procedimentos de validação intralaboratorial detalhados para métodos de ensaio quantitativos e qualitativos, tendo como base fundamentos teóricos relacionados aos diferentes parâmetros de desempenho, e incluindo:
  - definição da forma de avaliar o desempenho dos métodos,
  - parâmetros de desempenho a serem estudados,
  - delineamento experimental para cada parâmetro, com determinação dos níveis de concentração dos analitos a serem estudados, número de níveis e mínimo de repetições por nível,
  - testes estatísticos aplicáveis à análise de dados de cada parâmetro, com as respectivas premissas a serem verificadas e seus testes específicos;
- aplicar os procedimentos delineados na validação de diferentes métodos de ensaio para determinação de resíduos e contaminantes em alimentos:
  - método quantitativo para determinação de semicarbazida em *baby-food* (alimentos para lactentes e crianças de primeira infância) por cromatografia líquida de alta eficiência acoplada a espectrometria de massas (CLAE-EM-EM),
  - método quantitativo para determinação de arsênio em fígado bovino por espectrofotometria de absorção atômica com geração de hidretos e injeção em fluxo (HGAAS-FIAS),
  - método quantitativo para determinação de avermectinas em leite bovino por cromatografia líquida de alta eficiência com detecção por fluorescência (CLAE-FL),
  - método qualitativo para determinação de subprodutos de origem animal em rações para ruminantes por microscopia;
- disponibilizar métodos validados de análises de alimentos ainda não implementados nas redes de laboratórios oficiais do *Department for Environment Food and Rural Affairs* (DEFRA) e do Ministério da Agricultura Pecuária e Abastecimento (MAPA).

# REVISÃO DA LITERATURA

## 1 VALIDAÇÃO DE MÉTODOS

### 1.1 DEFINIÇÕES

Por não se tratar de um termo específico, diversas definições são encontradas na literatura para validação de métodos:

- é o processo que determina se um método analítico é aceitável para seu propósito pretendido (GREEN, 1996; FAJGELJ & AMBRUS, 2000);
- implica na demonstração de que o processo ou procedimento examinado realiza o que é reivindicado ou pretendido. Trata da comparação entre o comportamento observado sob condições rigorosamente estabelecidas e as expectativas de desempenho pré-determinadas, sendo que a extensão na qual a performance observada concorda com as expectativas determina se o processo é ou não válido (JENKE, 1996);
- consiste em checagens que devem ser realizadas para garantir que os parâmetros de desempenho de um método são conhecidos e demonstram que o método é cientificamente adequado nas condições em que é aplicado. Estabelece, por meio de um conjunto de estudos laboratoriais sistemáticos, que os parâmetros de desempenho do método atendem às especificações relacionadas com a intenção de uso dos resultados analíticos (NATA, 1997);
- é a busca pela faixa confiável de um método e a aplicação contínua deste conhecimento (BRUCE, MINKKINEN & RIEKKOLA, 1998);
- é o processo de determinação de um requisito analítico e de confirmação de que o método sob consideração possui capacidade de desempenho consistente com os requisitos de aplicação. Constitui o processo de determinação dos parâmetros de desempenho e das limitações de um método, com identificação dos fatores que podem mudar estes parâmetros e em qual extensão (EURACHEM, 1998);
- é o processo que confirma que um procedimento analítico empregado para um teste específico é adequado para seu propósito de uso (HUBER, 1998);
- em um conceito amplo, envolve todas as ações feitas para estabelecer e manter a qualidade dos resultados analíticos, incluindo a avaliação inicial dos parâmetros de desempenho, diversos tipos de ensaios interlaboratoriais e controle de qualidade interno (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999);

- é o processo de determinação da conveniência de um sistema de medição para fornecer resultados analíticos favoráveis (WOOD, 1999);
- inclui a determinação dos parâmetros de desempenho de um método, estabelecendo quais influências podem alterá-los e demonstrando que o método é adequado para uso (MAC NEIL, PATTERSON & MARTZ, 2000);
- é um componente essencial das medidas que um laboratório deve implementar para garantir que produz dados confiáveis (THOMPSON, ELLISON & WOOD, 2002);
- é o processo de verificar ou demonstrar certeza ou correção (*correctness*) (HORWITZ, 2003);
- é o processo que fornece evidências documentadas de que um método é capaz de medir aquilo que pretende medir (KING, 2003);
- é a investigação do alcance do propósito analítico do método, o qual é obtido por resultados analíticos com um nível de incerteza aceitável (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004);
- é a comprovação, pelo fornecimento de evidência objetiva, de que os requisitos para uma aplicação ou uso específico pretendido foram atendidos (EAL, 1997; EC, 2002; INMETRO, 2003; ISO, 2005 a).

## **1.2 VALIDAÇÃO NO CONTEXTO ANALÍTICO**

Segundo BRUCE, MINKKINEN & RIEKKOLA (1998), um processo analítico pode ser dividido nas etapas de desenvolvimento, otimização, validação, aplicação e revalidação do método. Estes autores discutem que, na prática, a passagem de uma etapa para outra acontece suavemente, embora seja aconselhável que o método a ser validado esteja claramente descrito antes de iniciar o processo de validação. FAJGELJ & AMBRUS (2000) relatam que a validação de métodos acontece após a conclusão da etapa de desenvolvimento, assumindo que fatores como calibração, adequação do sistema de medição e estabilidade do analito tenham sido satisfatoriamente estabelecidos. Tais autores argumentam ainda que a validação deva preceder a aplicação prática dos métodos, sendo importante uma subsequente implementação de controles de qualidade internos para a continuidade do processo.

Outros pesquisadores, contudo, consideram a separação das diferentes fases do processo analítico como algo difícil de ser obtido em condições reais. INMAN et al. (1987) tratam a validação como parte integrante do desenvolvimento, indicando que o desenvolvimento do método somente poderia ser considerado completo após avaliação e

demonstração de performance analítica aceitável. Para GREEN (1996), a validação de um método também não pode ser separada do seu desenvolvimento, uma vez que o responsável pelo desenvolvimento não irá saber se as condições do método são aceitáveis até que os estudos de validação sejam realizados. Este autor, assim como HUBER (1998), caracteriza o desenvolvimento e validação de métodos como processos iterativos. O guia publicado pela EURACHEM (1998) considera a validação de métodos intimamente vinculada ao desenvolvimento de forma que, muito freqüentemente, não é possível determinar o ponto no qual o desenvolvimento de um método termina para dar início à validação. Isto porque muitos dos parâmetros de desempenho associados à validação do método são normalmente avaliados como parte do desenvolvimento. Conforme a abordagem feita por JENKE (1996), a relação entre a validação e outras fases analíticas freqüentemente não é clara. O conceito hierárquico de utilização de métodos, no qual as etapas de desenvolvimento, otimização, validação, aplicação e revalidação são distintas, seqüenciais e independentes, representa uma situação indesejável e potencialmente ineficiente. A validação de um método sem nenhuma atividade de verificação prévia pode gerar um ciclo vicioso no qual métodos promissores são submetidos a protocolos rigorosos de validação para falharem em um ou mais parâmetros, requerendo desenvolvimento e otimização adicionais, além da revalidação. Diante disto, um processo de validação dividido em dois estágios é proposto por este autor, sendo uma validação “rudimentar” ou “pré-validação” conduzida durante o desenvolvimento do método e outra validação “formal” após o desenvolvimento ter sido completado. TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004) consideram como pré-validação a validação conduzida em condições intralaboratoriais, e validação aquela feita por meio de ensaios interlaboratoriais ou estudos colaborativos.

O desenvolvimento de métodos pode ocorrer de inúmeras formas, num extremo envolvendo adaptações a um método existente por meio de mudanças mínimas, em outro extremo partindo do delineando experimental para criação de um novo método. Nesta última forma, o processo de desenvolvimento exige muito trabalho e uma dúvida inicial acerca do sucesso do método desenvolvido (EURACHEM, 1998).

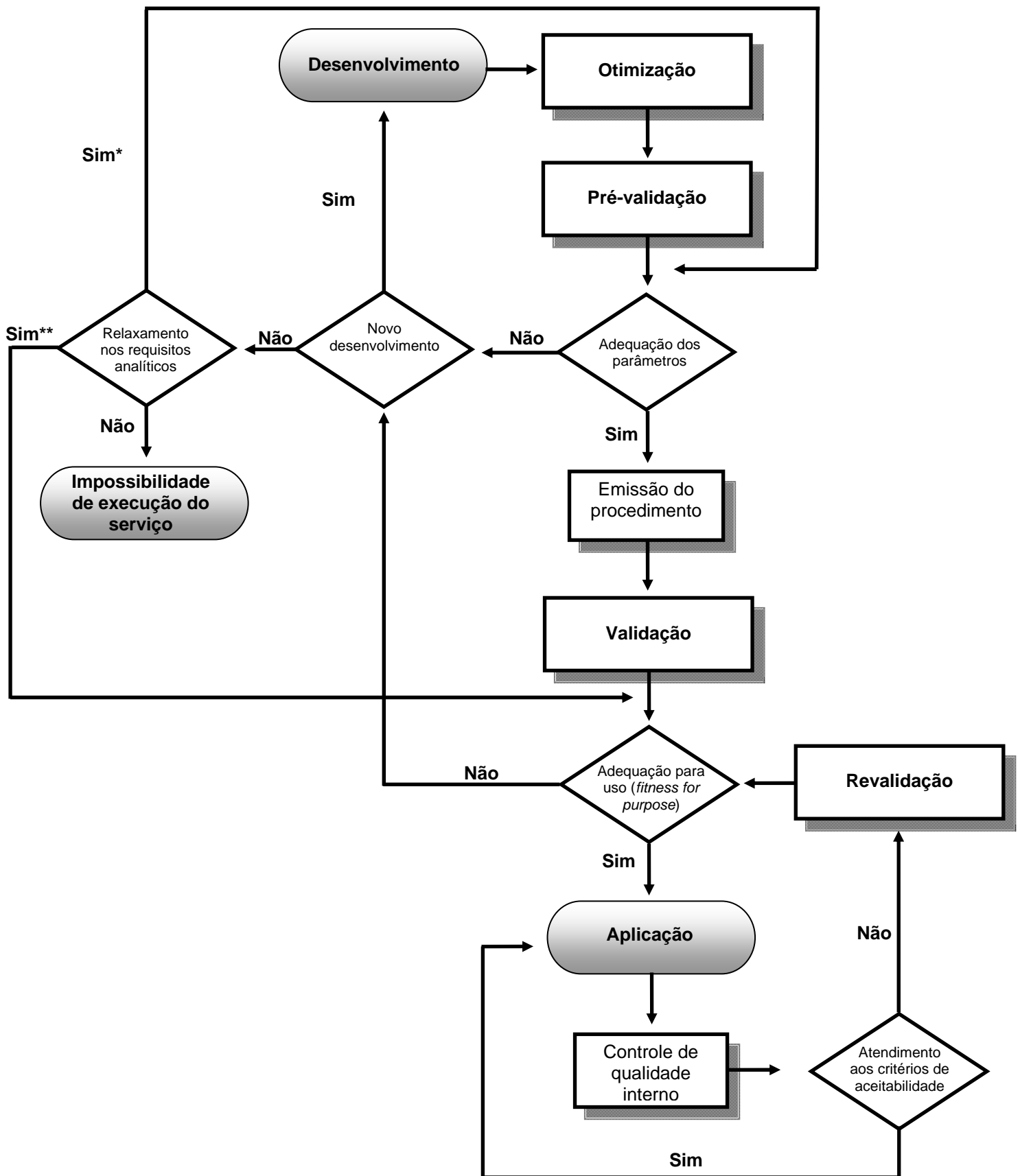
A revalidação pode ser necessária se algum aspecto operacional do método for modificado durante a rotina de sua aplicação ou se houver indicação de que o método estabelecido não é mais adequado (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). HILL & REYNOLDS (1999) sugerem que a revalidação seja restrita aos casos em que houver alterações ou extensões significativas ao método, ficando sob a responsabilidade dos analistas e do laboratório a classificação das mudanças como

significativas ou não. De acordo com HUBER (1998), as faixas de operação do método são determinadas em experimentos de robustez durante a etapa de desenvolvimento. Assim, as revalidações se fazem necessárias somente quando um método for alterado em características fora das faixas de operação previamente estabelecidas. Este autor também discute a necessidade de revalidação do método em função de resultados de controles de qualidade internos fora dos critérios de aceitabilidade especificados pelo laboratório.

A Figura 1 representa as etapas de desenvolvimento, otimização, “pré-validação”, validação, aplicação e revalidação de métodos de forma iterativa, ilustrando a relação entre estas diferentes etapas na condução do processo analítico.

### **1.3 VALIDAÇÃO NO CONTEXTO DE SISTEMAS DE GESTÃO**

A norma ISO/IEC 17025 (ISO, 2005 b) representa a síntese da experiência acumulada em todo o mundo em avaliações e na demonstração da competência de laboratórios. Esta norma contém requisitos de direção e técnicos que devem ser atendidos pelos laboratórios para implementação e manutenção de um sistema de gestão laboratorial adequado. Como requisitos de direção são apresentados: organização; sistema de gestão; controle de documentos; análise de pedidos, propostas e contratos; subcontratação; aquisições; atendimento ao cliente; reclamações; controle de trabalhos não-conformes; melhoria; ações corretivas; ações preventivas; controle de registros; auditorias internas; e análise crítica pela direção. Os requisitos técnicos abordados são: pessoal; acomodações e condições ambientais; métodos e validação de métodos; equipamentos; rastreabilidade da medição; amostragem; manuseio de itens (amostras); garantia da qualidade de resultados; e apresentação dos resultados. Todos os requisitos exigidos pela norma são tidos como essenciais e estão integrados no âmbito de sistemas de gestão laboratoriais e de processos de acreditação. Desta forma, embora os objetivos específicos de ferramentas como validação de métodos, controle de qualidade interno e participação em programas de comparação interlaboratorial ou ensaios de proficiência sejam diferentes, estas ferramentas possuem características complementares no processo de garantia da confiabilidade e comparabilidade dos resultados.



**Figura 1** - Relação entre as diferentes etapas do processo analítico (BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998 - adaptados).

\* Possibilidade de relaxamento dos requisitos analíticos quando os resultados da etapa de pré-validação indicam que os parâmetros de desempenho não estão adequados.

\*\* Possibilidade de relaxamento dos requisitos analíticos quando os resultados da etapa de validação indicam que não há adequação para uso do método.

Segundo a EAL (1997), os procedimentos para validação de métodos devem ser conduzidos de forma que sua descrição seja parte de um sistema e manual de gestão. MAC NEIL, PATTERSON & MARTZ (2000) discutem que o método validado deve estar incluído no escopo da acreditação do laboratório para que seja considerado apropriado. No procedimento harmonizado pela AOAC *International*, ISO e IUPAC para validação intralaboratorial de métodos, THOMPSON, ELLISON & WOOD (2002) relacionam a validação de métodos como requisito para acreditação pela ISO/IEC 17025. A União Européia, por meio da diretiva que estabelece normas sobre desempenho de métodos e interpretação de resultados (EC, 2002), requer que os laboratórios de seus países membros sejam acreditados em conformidade com a norma ISO/IEC 17025. O documento publicado pelo INMETRO (2003) sobre validação de métodos destaca a necessidade de que os laboratórios estejam cientes dos requisitos da norma ISO/IEC 17025 para seleção, desenvolvimento, utilização e validação de métodos. Os guias para validação de métodos da EAL (1997) e da EURACHEM (1998) também ratificam o entendimento de que a validação é parte integrante de sistemas de gestão laboratoriais quando afirmam que o uso de equipamentos dentro de especificações, funcionando corretamente e adequadamente calibrados (fatores técnicos e instrumentais), assim como a condução dos estudos por pessoal competente (fatores humanos) e em condições ambientais adequadas (fatores ambientais), devem estar implícitos nos procedimentos de validação de métodos.

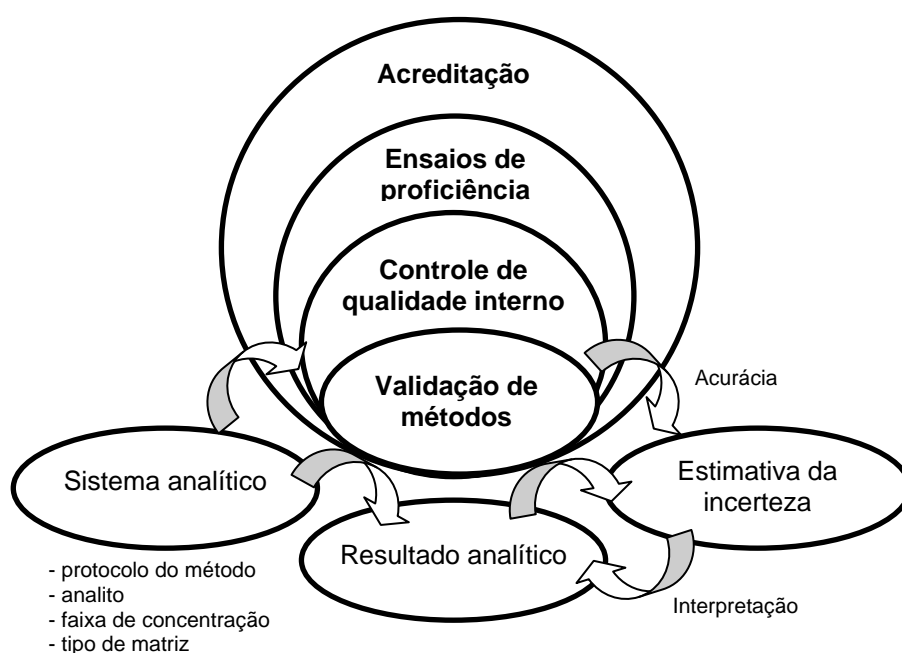
A Figura 2 ilustra a relação entre validação de métodos e outras ferramentas de garantia da qualidade na acreditação de sistemas de gestão de laboratórios.

Apesar de não cobertas pela norma ISO/IEC 17025 (ISO, 2005 b), questões como segurança e descarte de material também devem estar estruturadas nos sistemas de gestão dos laboratórios para que haja uma adequada condução das atividades analíticas, bem como dos procedimentos de validação de métodos.

#### **1.4 QUANDO VALIDAR**

A ISO/IEC 17025 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 a necessidade de validação para métodos normalizados, esta norma define que os laboratórios devem confirmar que têm condições de operar adequadamente os métodos normalizados antes da implantação dos ensaios (ISO,

2005 b). Considera-se métodos normalizados aqueles desenvolvidos por um organismo de normalização ou outras organizações, cujos métodos sejam aceitos pelo setor técnico em questão. Métodos não normalizados são aqueles desenvolvidos pelo próprio laboratório ou outras partes, ou adaptados a partir de métodos normalizados e validados (ILAC, 2002). Neste contexto, assume-se que os métodos designados como normalizados tenham sido submetidos a validações prévias por estudos colaborativos (MAC NEIL, PATTERSON & MARTZ, 2000). Frequentemente, isto procede, visto que a escolha de um método como oficial ou padrão requer que dados de precisão sejam incluídos (WOOD, 1999) e baseados em resultados obtidos de estudos colaborativos (HUND, MASSART & SMEYERS-VERBEKE, 2000).



**Figura 2** - Diferentes níveis e ferramentas de garantia da qualidade na acreditação de sistemas de gestão de laboratórios (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004 - adaptado).

WOOD (1999) discute a necessidade de validação quando os laboratórios desenvolvem seus próprios métodos ou quando fazem modificações significativas em métodos normalizados. Este autor recomenda que os laboratórios devam verificar se são capazes de alcançar os parâmetros de desempenho dados no método sempre que o método escolhido for normalizado. Se o método de referência não for validado por estudos colaborativos, uma referência do desvio padrão de reprodutibilidade pode ser estimada em função da concentração do analito pela equação de HORWITZ (1982). Quando o laboratório obtiver variâncias muito inferiores às relatadas para método de



referência, importantes fontes de incerteza podem ter sido ignoradas (THOMPSON, ELLISON & WOOD, 2002).

O INMETRO (2003) aborda a necessidade de processos de validação para métodos normalizados modificados ou métodos totalmente novos desenvolvidos pelo laboratório, além da necessidade de demonstração de que o laboratório tem condições de operar de maneira adequada os métodos normalizados (emitidos por organismos de normalização, organizações reconhecidas na sua área de atuação, publicados em livros ou periódicos de grande credibilidade na comunidade científica) antes de sua implantação.

Outras referências relacionadas à validação de métodos, por sua vez, sugerem processos de validação mesmo para métodos normalizados. Segundo a NATA (1997), se um método normalizado é utilizado, ou se um método normalizado é modificado para atender requisitos especiais, ou um método completamente novo é desenvolvido, o laboratório deve estabelecer se os parâmetros de desempenho do método atendem aos requisitos das operações analíticas pretendidas.

O guia para validação de métodos da EURACHEM (1998) também indica a necessidade de avaliação dos parâmetros de desempenho para métodos normalizados. Este documento complementa ainda que a validação deva ser realizada no desenvolvimento de novos métodos; para métodos normalizados revisados; quando o controle de qualidade interno indicar alterações de um método com o tempo; para métodos normalizados utilizados em diferentes laboratórios, por diferentes analistas ou equipamentos; e para demonstrar equivalência entre dois métodos.

THOMPSON, ELLISON & WOOD (2002) discorrem sobre a necessidade de validação de métodos normalizados submetidos a estudos colaborativos; além dos métodos normalizados submetidos a estudos colaborativos e aplicados a uma nova matriz; métodos publicados em literatura científica, com ou sem descrição de parâmetros de desempenho; métodos empíricos; análises *ad hoc*; e nos casos de alterações de analistas ou de equipamentos.

## **1.5 SELEÇÃO DE MÉTODOS**

Segundo a ISO/IEC 17025 (ISO, 2005 b), o laboratório deve selecionar e utilizar métodos de ensaio, incluindo os métodos para amostragem, que atendam às necessidades do cliente e que sejam apropriados para os ensaios que realiza. Muitas vezes isto somente é possível porque o laboratório auxilia o cliente na seleção dos

métodos, uma vez que estes freqüentemente não possuem competência analítica específica. Um pré-requisito para uma seleção de sucesso é que o propósito do ensaio e outras necessidades específicas estejam bem documentados em uma ordem ou requisição de serviço ou em um projeto de estudo (WOOD, 1999).

Quando o cliente não especificar o método a ser utilizado, o laboratório deve selecionar métodos apropriados que tenham sido publicados em normas internacionais, regionais ou nacionais, por organizações técnicas respeitáveis, em textos ou periódicos científicos relevantes ou especificados pelo fabricante do equipamento utilizado na análise, assegurando a utilização das últimas edições válidas. Também podem ser empregados métodos desenvolvidos ou adaptados pelo laboratório, se forem apropriados para o uso e se estiverem validados (ISO, 2005 b). Sempre que possível e prático, é recomendado que os laboratórios dêem preferência aos métodos cujos parâmetros de desempenho tenham sido avaliados por estudos colaborativos, de acordo com protocolos internacionais (THOMPSON, ELLISON & WOOD, 2002; ISO, 2005 b).

Cabe considerar que laboratórios que trabalham em controles oficiais nem sempre têm liberdade para selecionar os métodos, os quais podem estar estabelecidos em legislações específicas (WOOD, 1999).

## **1.6 COMO VALIDAR**

As formas de validação ou de verificação de desempenho de métodos de ensaio citadas pela ISO/IEC 17025 (ISO, 2005 b) incluem: calibração com uso de padrões ou materiais de referência; comparações com outros métodos; comparações interlaboratoriais; avaliação sistemática dos fatores que influenciam o resultado; e avaliação da incerteza dos resultados.

Segundo sugestões de MAC NEIL, PATTERSON & MARTZ (2000), a validação de métodos pode envolver: o uso de método normalizado, previamente validado por meio de estudos colaborativos; a comparação da performance do método em validação com a obtida por um método de referência reconhecido (oficial ou padrão internacional) e adequadamente documentado; o uso de materiais de referência; e a documentação da performance do método por meio de participações em ensaios de proficiência.

## 1.7 PROCESSOS INTRA E INTERLABORATORIAIS

As diferentes formas de validação recomendadas, diretas ou comparativas, podem ser estruturadas em processos intra e interlaboratoriais. Um panorama dos principais guias, protocolos e normas publicados para determinação de processos de validação intra e interlaboratoriais (estudos colaborativos) está apresentado na Tabela 1.

Apesar das vantagens e desvantagens apresentadas pelos diferentes processos, eles são complementares e não excluem uns aos outros. Se por um lado um estudo colaborativo (processo interlaboratorial) não pode ser conduzido sem que o método tenha sido previamente avaliado em ensaios de validação intralaboratoriais; por outro lado, métodos com parâmetros de desempenho adequados, avaliados em processos intralaboratoriais, não podem ser considerados como oficiais ou padrões sem estudos colaborativos complementares (WOOD, 1999; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004).

### 1.7.1 Estudo colaborativo

Estudos colaborativos, ou estudos interlaboratoriais de performance de métodos como designado pela IUPAC (HORMITZ, 1995), envolvem a análise de uma mesma amostra (ou amostras idênticas), pelo mesmo método de ensaio, em diferentes laboratórios, para determinar os parâmetros de desempenho de um método. Estes estudos cobrem medidas em condições de repetitividade e reprodutibilidade (EC, 2002).

Muitos segmentos, incluindo a área de análises de alimentos, consideram a validação de métodos por procedimentos interlaboratoriais, organizada na forma de estudos colaborativos, conforme preconizado pela ISO 5725 (ISO, 1994 a,b,c,d,e; ISO, 1998) ou pelo protocolo harmonizado pela AOAC *International*, ISO e IUPAC (YOUNG & STEINER, 1975; HORWITZ, 1995; AOAC, 2002), como um “padrão ouro de validação” (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999) ou como uma “validação completa” (WOOD, 1999).

Entretanto, muitas vezes estes estudos podem não ser praticáveis ou mesmo necessários (THOMPSON, ELLISON & WOOD, 2002), além de nem sempre serem suficientes para demonstrar adequação para uso (*fitness for purpose*) de um método (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999). Quando métodos de ensaio são utilizados por um grande número de laboratórios, para determinação de analitos estáveis, em matrizes bem definidas, a adoção deste tipo de validação é bastante encorajada (THOMPSON, ELLISON & WOOD, 2002). Na área de análises de alimentos, mesmo em

**Tabela 1** - Panorama das principais referências relacionadas a processos de validação intra e interlaboratoriais (estudos colaborativos) e respectivos órgãos emissores

Assunto	Órgão emissor	Título	Referência
Validação intralaboratorial	<i>United States Food and Drug Administration (FDA), United States Pharmacopeia (USP), International Conference on Harmonization (ICH)</i>	<i>Validation of analytical procedures: methodology</i>	ICH (1996)
	<i>European Cooperation for Accreditation of Laboratories (EAL)</i>	<i>EAL-P11. Validation of test methods: general principles and concepts</i>	EAL (1997)
	<i>National Association of Testing Authorities - Australia (NATA)</i>	<i>Technical note 17. Format and content of test methods and procedures for validation and verification of chemical test methods</i>	NATA (1997)
	<i>Association of Official Analytical Chemists (AOAC International)</i>	<i>AOAC Peer-verified methods program. Manual on policies and procedures</i>	AOAC (1998 b)
	<i>EURACHEM Working Group. A Focus for Analytical Chemistry in Europe</i>	<i>The fitness for purpose of analytical methods, a laboratory guide to method validation and related topics</i>	EURACHEM (1998)
	<i>Food and Agricultural Organization (FAO), World Health Organisation (WHO), CODEX ALIMENTARIUS Committee on Methods of Analysis and Sampling (CODEX/CCMAS)</i>	<i>CX/MAS 02/10. Single-laboratory validation. Consideration of harmonized IUPAC guidelines for the in-house validation of methods of analysis</i>	CODEX ALIMENTARIUS (2002)
	<i>The Commission of the European Communities (EC)</i>	<i>Commission decision 2002/657/EC</i>	EC (2002)
	<i>Association of Official Analytical Chemists (AOAC International), International Standards Organization (ISO) e International Union of Pure and Applied Chemistry (IUPAC)</i>	<i>Harmonized guidelines for single-laboratory validation of methods of analysis</i>	THOMPSON, ELLISON & WOOD (2002)
Instituto Nacional de Metrologia, Normalização e Qualidade Industrial - Brasil (INMETRO)	DOQ-CGCRE-008. Orientações sobre validação de métodos de ensaios químicos	INMETRO (2003)	
Estudo colaborativo	<i>Association of Official Analytical Chemists (AOAC International)</i>	<i>Statistical manual of the AOAC. Statistical techniques for collaborative tests</i>	YOU DEN & STEINER (1975)
	<i>International Standards Organization (ISO)</i>	<i>ISO 5725. Accuracy (trueness and precision) of measurement methods and results</i>	ISO (1994 a,b,c,d,e; 1998)
	<i>Association of Official Analytical Chemists (AOAC International), International Standards Organization (ISO) e International Union of Pure and Applied Chemistry (IUPAC)</i>	<i>Protocol for the design, conduct and interpretation of method-performance studies</i>	HORWITZ (1995)
	<i>Association of Official Analytical Chemists (AOAC International)</i>	<i>AOAC Official methods of analysis. Appendix D: guideline for collaborative study procedures to validate characteristics of a method of analysis</i>	AOAC (1997)

situações nas quais a validação de métodos por estudos colaborativos pode ser praticável, raramente é possível cobrir todas as combinações de analitos, concentrações e matrizes nas quais os métodos serão aplicados (HILL & REYNOLDS, 1999).

Os estudos colaborativos fornecem os procedimentos básicos para estimativa de exatidão e precisão de um método, esta última nas formas extremas de repetitividade e reprodutibilidade, podendo também estabelecer formas intermediárias de precisão dependendo das circunstâncias nas quais os procedimentos são aplicados. Segundo WOOD (1999), o objetivo destes estudos é avaliar e quantificar a precisão de um método e não sua exatidão. De qualquer forma, estes procedimentos interlaboratoriais não consideram parâmetros de desempenho como linearidade, seletividade, limites de detecção e quantificação (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999).

Outra limitação para a organização de estudos colaborativos está na recomendação de que os laboratórios participantes devem ser escolhidos aleatoriamente do universo de laboratórios que utilizam o método em validação. Mesmo desconsiderando a impossibilidade prática de seleção aleatória de laboratórios, o padrão estabelecido pela ISO 5725 (ISO, 1994 a,b,c,d,e; ISO, 1998) muitas vezes não é aplicado ou é aplicado incorretamente. Frequentemente, a premissa de que os laboratórios participantes devem trabalhar com a mesma acurácia também não é sustentada, devido à grande dificuldade em organizar um experimento interlaboratorial no qual todos os participantes operem o mesmo método de ensaio de acordo com as mesmas prescrições (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999). Vale destacar ainda os pontos críticos no desenvolvimento deste tipo de estudo em função da necessidade de que os materiais distribuídos sejam idênticos e estáveis, o que muitas vezes exige onerosos testes de homogeneidade e estabilidade, além de regulamentações para embarque de mercadorias e restrições às importações (MAC NEIL, PATTERSON & MARTZ, 2000).

### **1.7.2 Ensaio de proficiência**

Um modelo alternativo adotado para validar métodos por procedimentos interlaboratoriais consiste no uso de resultados de participações em ensaios de proficiência ao longo do tempo. Os ensaios de proficiência são considerados ferramentas de controle de qualidade externo (EURACHEM, 1998). Nestes ensaios, cujos procedimentos também são harmonizados entre *AOAC International*, ISO e IUPAC (THOMPSON, ELLISON & WOOD, 2006; ISO, 1997 a), os laboratórios participantes

analisam o material distribuído pelo organizador, utilizando o procedimento analítico adotado em suas rotinas de trabalho.

O propósito dos ensaios de proficiência é, primeiramente, a avaliação e melhoria contínua do desempenho do laboratório e de analistas, e não a validação de métodos. Porém, não existe aqui o problema observado para estudos colaborativos nos quais diferentes laboratórios adotam diferentes procedimentos analíticos (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999).

Restrições ao uso deste modelo encontram-se no fato de que os resultados de desempenho somente podem ser obtidos após participações extensivas em diversas rodadas de ensaios de proficiência, além dos elevados custos destas participações. Algumas dificuldades encontradas para os estudos colaborativos também são observadas neste caso, como a necessidade de garantia da homogeneidade, estabilidade e distribuição adequada dos materiais; e a limitação à avaliação dos parâmetros exatidão e precisão.

### **1.7.3 Controle de qualidade interno**

Com o objetivo de minimizar os custos da validação, muitos laboratórios utilizam resultados obtidos de controles de qualidade internos e compilados na forma de cartas de controle como substitutos ou complementares ao processo de validação de seus métodos (HILL & REYNOLDS, 1999).

Todavia, os verdadeiros objetivos do controle de qualidade interno são: i) monitorar a estabilidade dos parâmetros de desempenho de um método previamente validado durante seu uso em atividades de rotina; ii) manter um controle estatístico do processo de medição (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999); e iii) indicar aceitação ou rejeição de um grupo de resultados, com necessidade de repetição dos ensaios (THOMPSON & WOOD, 1995). Como o processo de validação, na prática, não consegue incluir a totalidade de variáveis analíticas encontradas durante o uso do método, ferramentas de controle de qualidade interno são requeridas para fornecerem evidências do desempenho do método durante sua aplicação (HILL & REYNOLDS, 1999). Uma distinção entre a validação de métodos e o controle de qualidade interno é apresentada por THOMPSON, ELLISON & WOOD (2002): enquanto a primeira é conduzida uma única vez ou em intervalos infreqüentes durante o tempo de uso de um método, indicando que tipo de performance pode ser esperada; o segundo é realizado com freqüência estabelecida para demonstrar o desempenho do método validado ao longo do tempo.

Processos de validação baseados em dados acumulados oriundos de cartas de controle de qualidade interno, assim como de dados provenientes de participações em ensaios de proficiência, somente permitem que o método seja considerado validado após um intervalo de tempo considerável (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999), além de serem restritos a certos parâmetros de desempenho, analitos, faixas de concentração e tipos de matrizes.

#### **1.7.4 Validação intralaboratorial**

Validações intralaboratoriais correspondem a estudos analíticos que envolvem um único laboratório, utilizando um mesmo método, para analisar a mesma ou diferentes amostras, sob diferentes condições, em um intervalo de tempo justificado (EC, 2002). Estes estudos permitem a avaliação do desempenho de métodos com um número considerável de experimentos, utilizando diferentes combinações de analitos, concentrações e matrizes, em um curto intervalo de tempo, facilmente adaptáveis à diferentes situações. As validações intralaboratoriais atendem um importante nicho de processos de validação de métodos devido à sua flexibilidade e habilidade em fornecer resultados rápidos para novos métodos (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999).

Procedimentos de validação intralaboratoriais têm sido internacionalmente aceitos, uma vez que diferentes organismos oficiais têm proposto protocolos para este tipo de validação, por exemplo NATA (1997); AOAC (1998 b); EURACHEM (1998); EC (2002); AOAC *International*, ISO e IUPAC (THOMPSON, ELLISON & WOOD, 2002); e INMETRO (2003).

Mesmo a AOAC *International*, forte partidária de que procedimentos interlaboratoriais são os meios preferidos para validação de métodos (EURACHEM, 1998), introduziu o *Peer Verified Method Program* para validação de métodos com somente mais um ou dois outros laboratórios, o qual é dividido em três estágios: i) estabelecimento dos parâmetros de desempenho por procedimentos intralaboratoriais; ii) demonstração de performance aceitável em um segundo ou terceiro laboratório; e iii) demonstração de performance satisfatória em estudo colaborativo interlaboratorial. De acordo com este programa, o sucesso no segundo estágio constitui requisito para considerar o método como AOAC *Peer verified* (AOAC, 1998 b).

O procedimento harmonizado pela AOAC *International*, ISO e IUPAC para validação intralaboratorial de métodos (THOMPSON, ELLISON & WOOD, 2002) considera este tipo de validação apropriado quando se deseja avaliar o desempenho de

um método antes de formalizar a participação em ensaios interlaboratoriais, quando não existirem ensaios interlaboratoriais disponíveis, ou nas situações em que estes não forem aplicáveis. Tal procedimento é recomendado como referência nas validações intralaboratoriais para propósitos do *CODEX ALIMENTARIUS* (CODEX ALIMENTARIUS, 2002).

Também no Brasil, iniciativas que avalizam os processos de validação intralaboratoriais vêm sendo tomadas, incluindo a tradução de documentos internacionais (ANVISA, 2007) e a criação de comissões de estudo como o CTLE 05 que gerou o documento DOQ-CGCRE-008 (INMETRO, 2003).

Processos de validação intralaboratorial permitem avaliar não somente a exatidão e a precisão de um método, mas também outros parâmetros de desempenho (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999) como linearidade, faixa de trabalho, sensibilidade, efeitos de matriz, seletividade, limites de detecção e de quantificação. Por outro lado, trabalhar isoladamente reduz, de forma inevitável, a quantidade de dados obtidos, restringindo informações como, por exemplo, àquelas relativas à comparabilidade interlaboratorial. Apesar da impossibilidade de estudo do parâmetro reprodutibilidade nas validações intralaboratoriais, o que exigiria a execução de experimentos em diferentes laboratórios, esta informação nem sempre é necessária (EURACHEM, 1998). Muitas vezes, estudos de reprodutibilidade parcial são suficientes (VAN DER VOET, VAN RHIJN & VAN DE WIEL, 1999). Também é possível ter uma estimativa da comparabilidade de resultados de medição de qualquer método com outros métodos pelo ensaio de materiais de referência certificados ou pela comparação do método em estudo com outro para o qual a validação já tenha sido feita (EURACHEM, 1998).

### **1.7.5 Validação comparativa**

As formas de validação de métodos até agora descritas podem ser classificadas como validações diretas. No entanto, também é possível validar um método de forma comparativa (EAL, 1997). Processos de validação de métodos por comparação com métodos normalizados podem ser estruturados tanto em estudos intra quanto interlaboratoriais, sendo válidas as vantagens e desvantagens apresentadas anteriormente. Cumpre considerar que na maioria das vezes em que um laboratório opta por implantar um método não normalizado, isto acontece devido às limitações para execução de métodos de referência (reconhecidos como oficial ou padrão) em sua estrutura analítica. Ou ainda, quando o método normalizado recomenda a utilização de



processos obsoletos e o laboratório dispõe de equipamentos mais modernos. De qualquer forma, se não existe a possibilidade de emprego do método normalizado, a validação por comparação também não pode ser realizada.

## **1.8 ABRANGÊNCIA DA VALIDAÇÃO**

A abrangência de processos de validação por estudos colaborativos está muito bem estabelecida no protocolo harmonizado pela AOAC *International*, ISO e IUPAC (HORWITZ, 1995), com definição do número mínimo de laboratórios participantes, materiais e replicatas, assim como os parâmetros de desempenho avaliados e as estatísticas aplicáveis na análise de dados. Em contraponto, a maior parte dos documentos que tratam sobre procedimentos para validação intralaboratorial de métodos discute os fatores que afetam a abrangência, mas não estabelece a extensão mínima necessária para este processo.

A abrangência de um procedimento de validação de métodos deve ser suficiente para atender às necessidades de uma determinada aplicação (ISO, 2005 b). O nível de detalhamento de um processo de validação é determinado por circunstâncias como necessidades, possibilidades, riscos, custos e prazos (EAL, 1997). Como a caracterização do desempenho de um método por meio de procedimentos de validação é um processo caro, inevitavelmente, fica limitado às considerações de custos e de tempo. Balancear as restrições ao processo de validação com a necessidade de um método validado é bastante difícil e, em algumas circunstâncias, torna-se mais apropriado subcontratar o serviço de outro laboratório (EURACHEM, 1998).

A validação depende do propósito de uso do método, da técnica escolhida e do procedimento em questão (BRUCE, MINKKINEN & RIEKKOLA, 1998).

Métodos são utilizados para diferentes propósitos como pesquisa, desenvolvimento, produção e processos de controle de qualidade (BRUCE, MINKKINEN & RIEKKOLA, 1998). Geralmente, processos de validação mais aprofundados e que garantam um alto grau de confiabilidade dos resultados são exigidos quando saúde, segurança e grandes valores econômicos estão envolvidos (EAL, 1997), o que corresponde à realidade da maior parte dos ensaios na área de análises de alimentos.

Cada técnica utilizada tem suas características próprias e deficiências, as quais também devem ser consideradas (BRUCE, MINKKINEN & RIEKKOLA, 1998). A extensão de estudos de validação depende, em parte, da experiência acumulada na técnica

analítica usada. Verificações mais estritas podem ser necessárias para técnicas pouco conhecidas ou menos estabelecidas (THOMPSON, ELLISON & WOOD, 2002).

Os requisitos de profundidade nos processos de validação variam ainda com as condições do método de referência, sua maturidade e prevalência de uso (EAL, 1997). Normalmente, processos mais extensos são recomendados para validações de métodos não normalizados que para métodos normalizados, sendo sugeridas verificações mais simplificadas para métodos normalizados previamente validados, validações restritas a alguns parâmetros para métodos normalizados alterados e validações aprofundadas para novos métodos (THOMPSON, ELLISON & WOOD, 2002; KING, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Infelizmente, a maioria dos métodos publicados na literatura científica fornece informações insuficientes sobre os parâmetros de desempenho dos métodos, o que dificulta a obtenção de credibilidade sobre os mesmos (WOOD, 1999) e torna necessária a realização de validações complementares mais detalhadas pelos laboratórios que pretendem aplicá-los.

A abrangência do processo da validação também é função do tipo de método quantitativo ou qualitativo. São considerados métodos quantitativos aqueles que determinam a quantidade de um determinado analito expressa na forma de um valor numérico com as unidades apropriadas; e métodos qualitativos aqueles que permitem identificar um analito com base nas suas propriedades físicas, químicas ou biológicas (EC, 2002). Métodos quantitativos são definidos também como aqueles capazes de produzir resultados quantitativos com exatidão e precisão aceitáveis; e métodos qualitativos como aqueles capazes de detectar a presença do analito, mas com informações quantitativas limitadas ou sem informações quantitativas (HILL & REYNOLDS, 1999). Para métodos quantitativos são sugeridas avaliações de parâmetros de desempenho como linearidade, faixa, sensibilidade, efeitos de matriz, seletividade, exatidão, precisão e limites. Enquanto para métodos qualitativos o conjunto de parâmetros avaliados é restrito, por exemplo, à seletividade e limite de detecção (ICH, 1996; JENKE, 1996; AOAC, 1998 b; EURACHEM, 1998; HILL & REYNOLDS, 1999; EC, 2002; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Diferentes extensões recomendadas para processos de validação de métodos quantitativos e qualitativos encontram-se ilustradas na Tabela 2. A AOAC (1998 b) apresenta alguns parâmetros adicionais para validação de métodos qualitativos como reação cruzada e incidência de resultados falsos positivos e falsos negativos. EURACHEM (1998) e HILL & REYNOLDS (1999) também consideram a relação entre resultados verdadeiros e falsos positivos (ou

**Tabela 2** - Variação das extensões de processos de validação em função do tipo de método (quantitativo ou qualitativo)

Parâmetro de desempenho	ICH (1996) JENKE (1996)		AOAC (1998 b)		EURACHEM (1998)		HILL & REYNOLDS (1999)		EC (2002)		TAVERNIERS et al. (2004)	
	Quantitativo	Qualitativo	Quantitativo	Qualitativo	Quantitativo	Qualitativo	Quantitativo	Qualitativo	Quantitativo	Qualitativo	Quantitativo	Qualitativo
Linearidade*	+	-	+	-	+	-	+	-	+	-	+	-
Faixa	+	-	NA	NA	+	-	+	-	NA	NA	+	-
Sensibilidade	NA	NA	+	-	+	-	NA	NA	NA	NA	+	-
Confirmação da identidade	NA	NA	NA	NA	+	+	NA	NA	NA	NA	NA	NA
Reação cruzada	NA	NA	-	+	NA	NA	NA	NA	NA	NA	NA	NA
Efeitos de matriz	NA	NA	NA	NA	NA	NA	+	-	NA	NA	NA	NA
Seletividade/Especificidade**	+	+	+	+	+	+	+	+	+	+	+	+
Exatidão	+	-	+	-	+	-	+	-	+	-	+	-
Precisão	+	-	+	-	+	-	+	-	+	-	+	-
Recuperação	NA	NA	+	+	+	-	NA	NA	+	-	+	-
Incidência de resultados falsos positivos e negativos	NA	NA	-	+	-	+	-	+	NA	NA	NA	NA
Limite de detecção	+	+	+	-	+	+	NA	NA	NA	NA	-	+
Limite de decisão	NA	NA	NA	NA	NA	NA	NA	NA	+	+	NA	NA
Limite de determinação/ Limite de decisão	NA	NA	+	-	NA	NA	NA	NA	NA	NA	NA	NA
Limite de quantificação	+	-	+	-	+	-	NA	NA	NA	NA	+	-
Capacidade de detecção	NA	NA	NA	NA	NA	NA	NA	NA	+	+	NA	NA
Menor nível de calibração	NA	NA	NA	NA	NA	NA	+	-	NA	NA	NA	NA
Aplicabilidade	NA	NA	NA	NA	NA	NA	NA	NA	+	+	NA	NA
Estabilidade	NA	NA	NA	NA	NA	NA	+	+	+	+	NA	NA
Robustez	+	+	NA	NA	+	+	NA	NA	+	+	+	+
Estimativa da incerteza	NA	NA	NA	NA	+	-	NA	NA	NA	NA	NA	NA

\* Parâmetro denominado linearidade/curva de calibração por AOAC (1998), calibração por HILL & REYNOLDS (1999) e curva de calibração por EC (2002).

\*\*Parâmetro denominado seletividade/especificidade por EURACHEM (1998), EC (2002) e TAVERNIERS et al. (2004), e especificidade nas demais referências.

NA = parâmetro não aplicável, + = avaliação recomendada, - = avaliação não recomendada.

falsos negativos) como uma forma de expressar a precisão para métodos qualitativos. EURACHEM (1998) aborda ainda o estudo da confirmação da identidade do analito junto à avaliação do parâmetro seletividade.

Assim, se por um lado a extensão de um processo de validação e escolha dos parâmetros de desempenho a serem avaliados depende das condições do método (normalizado ou não normalizado) e do tipo de método de ensaio (quantitativo ou qualitativo); por outro lado, o plano de validação é determinado pelos requisitos analíticos definidos com base nas necessidades dos clientes ou em determinações regulamentares (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Como o cliente ou a autoridade regulamentadora podem definir diferentes requisitos de tempo, custos e nível de qualidade em função do propósito de aplicação, dentro de um único laboratório podem existir diferentes níveis de qualidade de métodos associados a diferentes níveis de validação (BRUCE, MINKKINEN & RIEKKOLA, 1998).

## **2 PROCESSOS DE VALIDAÇÃO INTRALABORATORIAIS**

INMAN et al. (1987) consideram que o processo de validação pode ser dividido em quatro fases: i) identificação dos parâmetros de desempenho apropriados; ii) delineamento dos experimentos para avaliação de cada parâmetro; iii) determinação dos critérios de aceitabilidade para a performance de cada parâmetro; e iv) tradução dos resultados no procedimento do método.

O processo de validação recomendado pela NATA (1997) é conduzido em três estágios: i) determinação da exatidão e precisão do método com um único analista; ii) análises de amostras desconhecidas preparadas independentemente; e iii) determinação da reprodutibilidade e da robustez do método.

BRUCE, MINKKINEN & RIEKKOLA (1998) tratam a validação como um processo contínuo. Estes autores utilizam a expressão “validação original” para a validação inicial do método e decompõem o processo em quatro fases de igual importância: i) planejamento dos testes; ii) realização dos testes; iii) análise dos resultados (estatística e química); e iv) documentação. Estes autores incluem ainda uma quinta etapa de aplicação de informações obtidas durante a validação, como refinamento do método escrito, medidas de revalidação e definição de ferramentas de controle de qualidade.

CHIAP et al. (1999), por sua vez, dividem o processo de validação em duas fases: i) pré-validação; e ii) validação propriamente dita. Na primeira fase são avaliados linearidade, faixa de trabalho, limite de detecção e limite de quantificação; enquanto na

segunda fase são estabelecidos seletividade, exatidão e precisão. SHABIR (2003) também considera as fases de pré-validação e validação de método, entretanto, sua abordagem sobre estas fases difere das apresentadas por CHIAP et al. (1999). SHABIR (2003) inclui na pré-validação a avaliação de requisitos de qualificação, calibração e manutenção dos equipamentos; de estabilidade de soluções analíticas; e de otimização e verificação da adequação dos sistemas de medição. Na etapa de validação este autor considera a avaliação dos parâmetros de desempenho.

FEINBERG & RAGUÈNÈS (1999) separam o processo de validação em quatro delineamentos experimentais distintos: i) verificação da linearidade da curva de calibração e computação dos limites de detecção e quantificação; ii) controle da especificidade e dos efeitos de matriz; iii) estimativa da precisão e exatidão; e iv) cálculo da reprodutibilidade parcial, este último quando necessário.

## **2.1 PRINCÍPIOS BÁSICOS**

### **2.1.1 Especificação e escopo**

A validação intralaboratorial é única para cada método e deve ser consistente com os propósitos de uso (INMAN et al., 1987). Validações se aplicam a protocolos definidos, para determinação de analitos, em determinadas faixas de concentração e tipos de matrizes. Considerando que “sistemas analíticos” compreendem o protocolo do método de ensaio a ser validado, os analitos e suas faixas de concentração e os tipos de matrizes especificados, com rigor, pode-se dizer que os “sistemas analíticos” são validados e não os “métodos de ensaio” (THOMPSON, ELLISON & WOOD, 2002). Dai a grande importância de que, antes de qualquer processo de validação ser iniciado, o escopo seja definido, incluindo ambos sistema analítico (protocolo do método, analito, faixa de concentração e tipo de matriz) e requisitos analíticos (considerados para avaliação da adequação para uso) (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004).

### **2.1.2 Representatividade**

É essencial que os estudos de validação intralaboratorial sejam representativos, fornecendo um levantamento realístico do número e faixa de efeitos operacionais, além dos analitos, faixas de concentração e matrizes possíveis de ocorrerem nas condições de uso do método (THOMPSON, ELLISON & WOOD, 2002). Os experimentos envolvendo replicatas devem ser delineados, sempre que possível, considerando-se a importância de que as replicatas sejam verdadeiras e não somente replicatas de leituras, pois o objetivo

dos procedimentos de validação é conhecer a variabilidade real e não a variabilidade mínima dos resultados (INMAN, 1987; EURACHEM, 1998). Para métodos que envolvem multi-analitos ou multi-matrizes, bem representados na área de análises de alimentos pelos métodos de ensaios de resíduos de agrotóxicos, a validação é impraticável para todas as combinações de analitos, concentrações e matrizes que podem ser encontradas. Desta forma, recomenda-se que sejam validados analitos e matrizes representativos, para que o tempo e os custos da validação sejam viáveis antes da implantação do método nas atividades de rotina do laboratório. Subseqüentemente, os parâmetros de desempenho podem ser avaliados em todas as condições possíveis para definir a performance do método. Quando houver indicação de performance inadequada, o método deve ser modificado e revalidado (FAJGELJ & AMBRUS, 2000).

### **2.1.3 Verificação de premissas**

Além do fornecimento dos parâmetros de desempenho, que indicam a adequação para uso e têm dominado a realidade prática de processos de validação intralaboratorial, estes estudos também devem incluir uma verificação das premissas nas quais os métodos de ensaio são baseados. Como os estudos de validação são baseados em hipóteses estatísticas, uma verificação básica das premissas relacionadas aos testes estatísticos é fundamental para garantia de que os princípios destes testes não sejam afetados e para que os resultados obtidos sejam sustentados (THOMPSON, ELLISON & WOOD, 2002). Assim, previamente à realização de qualquer inferência, é necessário examinar as premissas para determinar se os dados são apropriados à aplicação dos testes (SNEDECOR & COCHRAN, 1989).

Há uma importante implicação prática neste princípio aparentemente abstruso. É mais fácil verificar aspectos globais de uma premissa do que provar que uma suposição particular está correta. Portanto, quando houver grande prática e uso próspero de uma técnica analítica em particular, para uma gama bem estabelecida de analitos, concentrações e matrizes, as verificações da validação têm caráter preventivo. Reciprocamente, quando a experiência for pequena, os processos de validação precisam prover fortes evidências de que as premissas feitas são apropriadas nos casos particulares sob estudo (THOMPSON, ELLISON & WOOD, 2002).

### **2.1.4 Fontes de erros**

Os erros das medidas analíticas provêm de diferentes fontes e de diferentes níveis de organização do processo analítico, devendo ser considerados no contexto de

validação de métodos e, posteriormente, para estimativa da incerteza da medição e verificação da adequação para uso.

O erro de uma medição corresponde ao resultado da medição menos o valor verdadeiro do mensurando. O erro total pode ser decomposto em erro aleatório e erro sistemático. Erro aleatório é o resultado de uma medição menos a média resultante de um infinito número de medições do mesmo mensurando, efetuadas sob condições de repetitividade. Enquanto o erro sistemático é a média resultante de um infinito número de medições do mesmo mensurando, efetuadas sob condições de repetitividade, menos o valor verdadeiro do mensurando (ISO, 1993 a,b).

Existem diversas formas possíveis de partição dos erros, podendo ser citados: i) o modelo estabelecido pela ISO 5725 (ISO, 1994 a,b,c,d,e; ISO, 1998) que combina efeitos do laboratório e efeitos do ensaio (*run*); e ii) o modelo de estimativa da incerteza determinado pelo ISO GUM (ISO, 1995) e EURACHEM/CITAC (2000) que é bastante apropriado para estabelecer e medir, separadamente, os efeitos das diversas fontes de incerteza que influenciam o resultado.

Apesar das diferentes fontes de erros não serem necessariamente independentes, uma forma útil de representar estas fontes inclui: erros aleatórios, tendências do ensaio, tendências do laboratório, tendências do método e efeitos de variação de matriz (THOMPSON, ELLISON & WOOD, 2002; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004), conforme indicado na Figura 3.

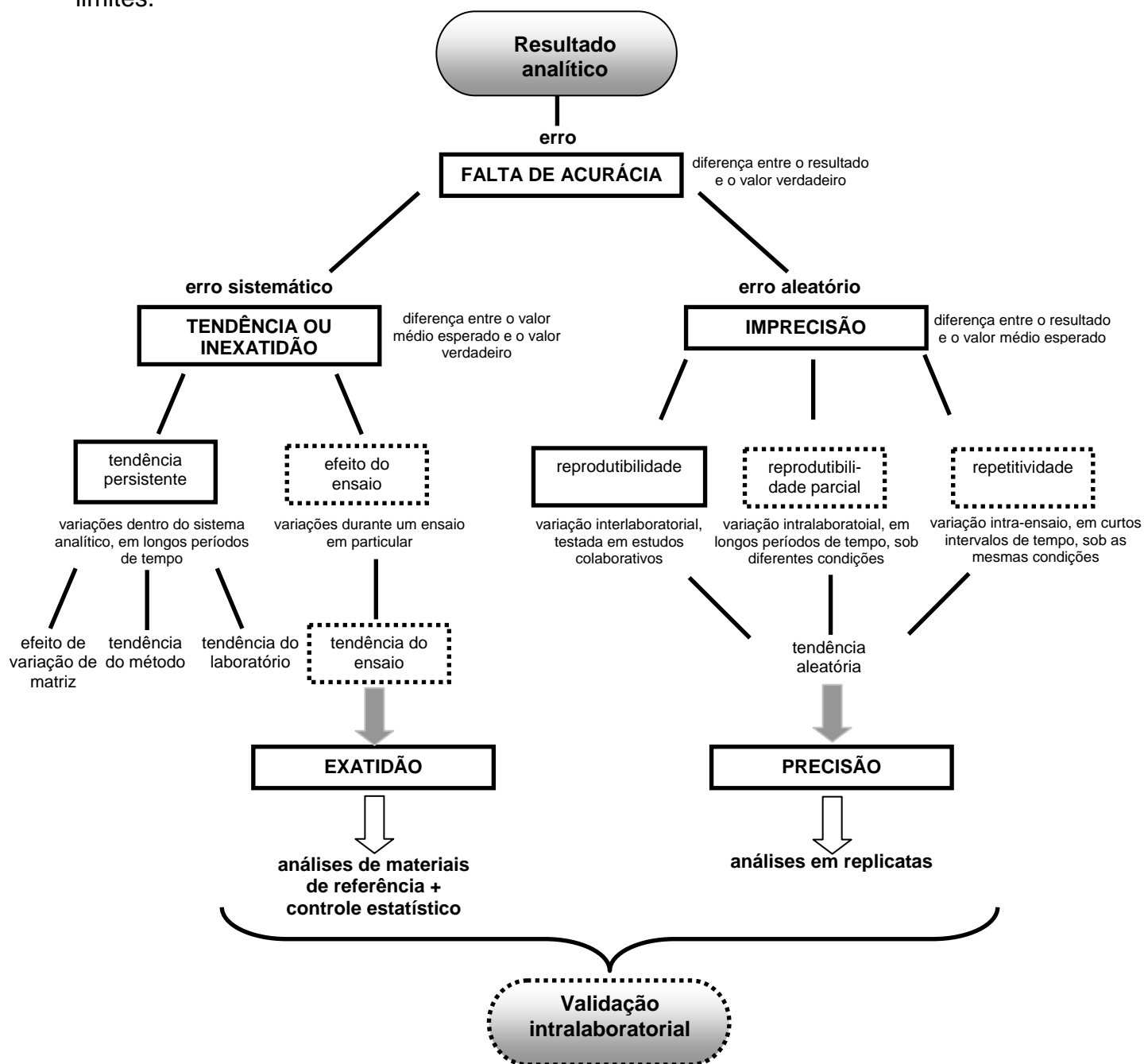
## **2.2 PARÂMETROS DE DESEMPENHO**

Parâmetro de desempenho é definido pela EC (2002) como uma qualidade funcional que pode ser atribuída a um método de ensaio.

### **2.2.1 Termos e definições**

Significativos progressos são reconhecidos no sentido de construir um esquema conceitual único e consistente para os termos relacionados à validação. Alguns termos e definições foram esclarecidos e refinados, enquanto outros, vagos ou ambíguos, estão sendo eliminados ou redefinidos (ISO, 1993 a; RSC, 2003). No entanto, divergências ainda são observadas na literatura com relação aos termos e definições dos parâmetros de desempenho avaliados em processos de validação, como pode ser observado nas Tabelas 3 a 14.

Considerando que o refinamento no uso de termos e suas respectivas definições não é somente um formalismo, e que o mau uso das palavras pode levar a sérias confusões (RSC, 2003), alguns comentários foram incluídos nos rodapés das Tabelas 3 a 14, visando harmonização da terminologia, especialmente relacionada à faixa, seletividade, exatidão, recuperação aparente, condições de avaliação da precisão e limites.



**Figura 3** - Composição do erro de um resultado analítico, relacionada à validação intralaboratorial do método de ensaio (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004 - adaptado).

Pontilhado indica os componentes dos parâmetros de desempenho exatidão e precisão que podem ser avaliados em processos de validação intralaboratorial.



**Tabela 3** - Parâmetros de desempenho e respectivas definições apresentados por GREEN (1996)

Parâmetro de desempenho	Definição
Linearidade	- Verificação de que as soluções de amostras estão em uma faixa de concentração na qual a resposta do analito é linearmente proporcional à concentração.
Faixa	- Intervalo de concentração ao longo do qual linearidade, exatidão ( <i>accuracy</i> )** e precisão aceitáveis são obtidas.
Especificidade*	- Habilidade do método em medir acuradamente a resposta do analito na presença de todos os compostos potenciais da amostra.
Exatidão ( <i>accuracy</i> )**	- Proximidade do valor medido ao valor verdadeiro para a amostra.
Precisão	- Quantidade de dispersão nos resultados obtidos de múltiplas análises de uma amostra homogênea.
Repetitividade/Precisão intra-ensaio	- Precisão obtida por análises repetidas, em um laboratório e no mesmo dia, de alíquotas de uma amostra homogênea, as quais tenham sido preparadas independentemente de acordo com o procedimento do método.
Precisão intermediária	- Precisão obtida quando o ensaio é realizado por múltiplos analistas, utilizando múltiplos equipamentos, em múltiplos dias, em um laboratório. Historicamente chamada de <i>ruggedness</i> ***.
Reprodutibilidade	- Precisão determinada pelo ensaio de amostras homogêneas em múltiplos laboratórios, freqüentemente como parte de estudos interlaboratoriais.
Limite de detecção	- Menor concentração de analito que produz resposta detectável acima do nível do ruído do sistema, tipicamente, três vezes o nível do ruído.
Limite de quantificação	- Menor nível do analito que pode ser medido exata ( <i>accurately</i> )** e precisamente.
Robustez ( <i>robustness</i> )	- Habilidade de um método não ser afetado por pequenas mudanças em parâmetros como porcentagem de conteúdo orgânico ou pH da fase móvel, concentração de tampões, temperatura e volume de injeção.
Estabilidade	NA

\* Uso do termo especificidade, em contradição à recomendação pelo emprego do termo seletividade (VESSMAN et al., 2001).

\*\* Uso do termo *accuracy* e de sua definição para tratar exatidão, em contradição à recomendação pelo emprego do termo *trueness* (ISO, 1993 b).

\*\*\* Relação do termo robustez (*ruggedness*) à precisão intermediária.

NA = definição não apresentada.

**Tabela 4 - Parâmetros de desempenho e respectivas definições apresentados por EAL (1997)**

<b>Parâmetro de desempenho</b>	<b>Definição</b>
Repetitividade (de resultados de medição)	- Proximidade de concordância entre resultados de medidas sucessivas de um mesmo mensurando, conduzidas sob as mesmas condições de medição. Condições de repetitividade incluem mesmo procedimento de medição, mesmo observador, mesmo instrumento de medição utilizado sob as mesmas condições, mesmo local e repetições em curtos intervalos de tempo. Pode ser expressa quantitativamente em termos de características de dispersão dos resultados.
Reprodutibilidade (de resultados de medição)	- Proximidade de concordância entre resultados de medidas de um mesmo mensurando, conduzidas sob condições modificadas de medição. As mudanças nas condições podem incluir princípio da medição, método de medição, observador, instrumento de medição, padrão de referência, local, condições de uso e tempo. Pode ser quantitativamente expressa em termos de características de dispersão dos resultados.
Incerteza (da medição)	- Parâmetro associado ao resultado de uma medição que caracteriza a dispersão dos valores que poderiam ser razoavelmente atribuídos ao mensurando. O parâmetro pode ser, por exemplo, desvio padrão (ou um múltiplo deste) ou intervalo de confiança.

**Tabela 5** - Parâmetros de desempenho e respectivas definições apresentados por NATA (1997)

Parâmetro de desempenho	Definição
Linearidade	- Habilidade de um método analítico produzir resultados de ensaios que sejam proporcionais à concentração do analito nas amostras, em uma dada faixa de concentração, tanto diretamente quanto por meio de uma transformação matemática bem definida.
Faixa (de aplicabilidade)	- Intervalo entre o maior e o menor nível de concentração de um analito no qual tenha sido demonstrada capacidade de determinação com precisão, exatidão ( <i>accuracy</i> )** e linearidade requeridas, sob as condições de ensaio estabelecidas.
Sensibilidade	- Inclinação da função de calibração.
Seletividade/Especificidade*/ Interferências/Efeitos de matriz	- Capacidade de determinar acurada e especificamente o analito de interesse na presença de outros componentes da matriz, sob as condições de ensaio estabelecidas.
Exatidão ( <i>accuracy</i> )**	- Grau de afastamento entre uma medida e o valor aceito.
Precisão	- Variabilidade que ocorre quando repetidas medições são feitas.
Recuperação	NA
Reprodutibilidade ( <i>ruggedness</i> )***	NA
Limite de detecção	- Menor concentração do analito em uma amostra que pode ser detectada, mas não necessariamente quantificada, sob as condições de ensaio estabelecidas.
Limite de quantificação/ Limite de reportagem	- Menor concentração do analito que pode ser determinada com exatidão ( <i>accuracy</i> )** e precisão (repetitividade) aceitáveis, sob as condições de ensaio estabelecidas.
Robustez ( <i>robustness</i> )	NA
Incerteza da medição	- Parte da expressão do resultado corrigido de uma medição, a qual define a faixa de valores dentro da qual o valor verdadeiro ou, se apropriado, o valor verdadeiro aceito, é esperado ser encontrado.

\* Uso intercambiável dos termos especificidade e seletividade, em contradição à recomendação pelo emprego do termo seletividade (VESSMAN et al., 2001).

\*\* Uso do termo *accuracy* e de sua definição para tratar exatidão, em contradição à recomendação pelo emprego do termo *trueness* (ISO, 1993 b).

\*\*\* Relação do termo robustez (*ruggedness*) à reprodutibilidade.

NA = definição não apresentada.

**Tabela 6** - Parâmetros de desempenho e respectivas definições apresentados por AOAC (1998 b)

Parâmetro de desempenho	Definição
Linearidade/Curva de calibração	- Linearidade: habilidade de um método obter resultados de ensaios proporcionais à concentração. - Curva de calibração: representação gráfica do sinal medido em função da quantidade de analito.
Sensibilidade	NA
Especificidade*	- Habilidade de um método medir somente o que é pretendido que seja medido.
Exatidão ( <i>accuracy</i> )**	- Proximidade do valor determinado em relação ao valor verdadeiro. A diferença entre o valor médio e o valor verdadeiro, ou seja, erros sistemáticos ou tendências, expressos como porcentagem do valor verdadeiro.
Recuperação	- A fração do analito adicionada a uma amostra antes dos ensaios, a qual é medida (recuperada) pelo método. Indicativo de exatidão ( <i>trueness</i> ) para métodos quantitativos.
Precisão	- Grau de concordância de medições sob condições específicas.
Repetitividade	- Componente da reprodutibilidade que pode ser determinado intralaboratorialmente.
Reprodutibilidade	- Idealmente, precisão avaliada pela performance entre laboratórios.
Limite de detecção	- Menor quantidade que pode ser medida com certeza estatística razoável.
Limite de determinação/ Limite de decisão***	- Menor quantidade de analito, se realmente presente, que será detectada e que pode ser identificada.
Limite de quantificação	- Igual ou maior ao menor ponto de concentração da curva de calibração. A menor quantidade de analito na amostra que pode ser quantitativamente determinada com precisão e exatidão ( <i>accuracy</i> )** apropriadas, para o analito e matriz considerados.
Reação cruzada	NA
Incidência de falsos positivos e falsos negativos	- Forma de expressar a precisão para métodos qualitativos.

\* Uso do termo especificidade e sua definição, em contradição à recomendação pelo emprego do termo seletividade (VESSMAN et al., 2001).

\*\* Uso do termo *accuracy* para tratar exatidão, em contradição à recomendação pelo emprego do termo *trueness* (ISO, 1993 b).

\*\*\* Os termos limite de determinação e limite de decisão aparecem freqüentemente em outras referências como sinônimos de limite de quantificação e de limite de detecção, respectivamente, e não como um limite intermediário.

NA = definição não apresentada.

**Tabela 7** - Parâmetros de desempenho e respectivas definições apresentados por BRUCE, MINKKINEN & RIEKKOLA (1998)

Parâmetro de desempenho	Definição
Linearidade	- Conhecimento de como a resposta medida depende da concentração do analito.
Faixa	- Faixa de concentração na qual a sensibilidade do método pode ser considerada uma constante.
Sensibilidade	- Parâmetro que descreve o quanto a resposta muda com mudanças na concentração do analito.
Seletividade/Especificidade	- Parâmetros relacionados ao evento da detecção. Método que produz resposta para um único analito pode ser chamado de específico. Método que produz respostas para várias espécies químicas, mas é capaz de distinguir a resposta do analito das outras respostas pode ser chamado de seletivo.
Exatidão ( <i>accuracy</i> )*	- Descreve quão próximo o valor medido está do valor verdadeiro.
Precisão	NA
Repetitividade/Precisão intra-ensaio	- Estimativa de precisão em uma situação em que a mesma amostra é analisada com o mesmo procedimento, pelo mesmo equipamento, em um curto intervalo de tempo e pelo mesmo analista.
Precisão de cotidiano ( <i>day-to-day</i> ) /Precisão inter-ensaios	- Estimativa da precisão em uma situação em que a mesma amostra é analisada pelo mesmo laboratório, mas por diferentes analistas, possivelmente por diferentes equipamentos e em tempos diferentes.
Reprodutibilidade	- Estimativa da precisão de um método numa situação em que a mesma amostra é analisada pelo mesmo procedimento, por equipamentos diferentes, em tempos diferentes, por diferentes analistas e em diferentes laboratórios.
Limite de detecção/ Limite de decisão/ Limite de discriminação	- Limite de detecção: menor concentração do analito em uma amostra que pode ser distinguida do ruído do instrumento usado para medição. Frequentemente uma razão sinal/ruído de 2 ou 3 é requerida para medidas em cromatografia e espectrofotometria. - Limite de decisão: limite de detecção expresso em unidades de resultado analítico. - Limite de discriminação: fornece a menor quantidade do analito adicionada a uma amostra que pode ser distinguida com elevada probabilidade estabelecida. No nível zero de adição de analito, o limite de discriminação é o limite de detecção.
Limite de quantificação/ Limite de determinação	- Nível mínimo de analito que pode ser medido com exatidão ( <i>accuracy</i> )* e precisão aceitáveis.
Robustez ( <i>robustness</i> ou <i>ruggedness</i> )	- Descreve a estabilidade do método. Parâmetro que estuda a praticabilidade** e transferibilidade** de um método.

\* Uso do termo *accuracy* e de sua definição para tratar exatidão, em contradição à recomendação pelo emprego do termo *trueness* (ISO, 1993 b).

\*\* Trata praticabilidade e transferibilidade associadas ao estudo da robustez.

NA = definição não apresentada.

**Tabela 8 - Parâmetros de desempenho e respectivas definições apresentados por EURACHEM (1998)**

<b>Parâmetro de desempenho</b>	<b>Definição</b>
Faixa de trabalho/Faixa linear/Linearidade	<ul style="list-style-type: none"> <li>- Faixa de trabalho: faixa de concentração do analito, ou valores de propriedades, ao longo da qual o método pode ser aplicado.</li> <li>- Faixa linear: por inferência é a faixa de concentrações do analito ao longo da qual o método fornece resultados de ensaio proporcionais à concentração do analito (AOAC - PVMC).</li> <li>- Linearidade: habilidade de um método obter resultados de ensaios proporcionais à concentração do analito (AOAC - PVMC).</li> </ul>
Confirmação da identidade	<ul style="list-style-type: none"> <li>- Determinação de que o sinal produzido no estágio de medição, ou outra propriedade medida, o qual tem sido atribuído ao analito, é somente devido ao analito e não a presença de algo química ou fisicamente similar, ou que apareça como uma coincidência.</li> </ul>
Sensibilidade	<ul style="list-style-type: none"> <li>- Gradiente da curva resposta, ou seja, a alteração na resposta do instrumento que corresponde a uma mudança na concentração do analito.</li> <li>- Mudança na resposta de um instrumento de medição dividida pela correspondente alteração no estímulo (VIM 1984 e IUPAC Orange Book).</li> </ul>
Seletividade/Especificidade*	<ul style="list-style-type: none"> <li>- Medidas que estabelecem a confiabilidade das medições na presença de interferentes. Especificidade é geralmente considerada como 100 % de seletividade, mas concordância sobre isto não é universal.</li> <li>- Seletividade quantitativa: termo utilizado junto de outro substantivo (como constante, coeficiente, índice, fator, número) para a caracterização quantitativa de interferências (IUPAC Compendium of Chemical Technology, 1985).</li> <li>- Seletividade qualitativa: a extensão na qual outras substâncias interferem com a determinação de uma substância de acordo com um dado procedimento (IUPAC Compendium of Chemical Technology, 1985).</li> <li>- Seletividade/Especificidade*: habilidade de um método em determinar acurada e especificamente o analito de interesse na presença de outros componentes na matriz, sob condições de ensaio estabelecidas (NATA Tech Note#13).</li> <li>- Especificidade**: habilidade de mensurar de forma inequívoca o analito na presença de compostos os quais pode-se esperar que estejam presentes. Tipicamente estes podem incluir impurezas, produtos de degradação, matriz, etc. (ICH Q2A, CPMP/ICH/381/95).</li> <li>- Especificidade: habilidade de um método medir somente aquilo que é pretendido medir (AOAC - PVMC).</li> </ul>
Exatidão ( <i>trueness</i> )	<ul style="list-style-type: none"> <li>- Comparação entre a média de resultados do método com valores conhecidos. Duas técnicas básicas: verificação contra valores de referência de materiais caracterizados ou de outro método caracterizado.</li> <li>- Proximidade de concordância entre o valor médio obtido de um grande conjunto de resultados de ensaios e um valor de referência aceito. Medida da exatidão (<i>trueness</i>) é geralmente expressa em termos de tendências (<i>bias</i>). A referência à exatidão (<i>trueness</i>) como acurácia (<i>accuracy</i>) não é geralmente recomendada (ISO 3534-1).</li> </ul>
Precisão	<ul style="list-style-type: none"> <li>- Proximidade de concordância entre resultados de ensaios independentes obtidos sob condições estipuladas. Depende somente da distribuição de erros aleatórios e não está relacionada ao valor verdadeiro ou valor especificado. Usualmente expressa em termos de imprecisão e computada como desvio padrão dos resultados de ensaios. Medidas quantitativas da precisão dependem crucialmente das condições estipuladas, sendo repetitividade e reprodutibilidade formas particulares de condições extremas (ISO 3534-1).</li> </ul>
Repetitividade	<ul style="list-style-type: none"> <li>- Menor precisão esperada, fornece uma idéia da variabilidade esperada quando um método é aplicado por um único analista, em um equipamento, ao longo de um curto intervalo de tempo. Por exemplo, o tipo de variabilidade esperada entre resultados de uma amostra analisada em duplicata.</li> <li>- Condições nas quais resultados de ensaios independentes são obtidos com o mesmo método, em amostras idênticas, no mesmo laboratório, pelo mesmo analista, utilizando o mesmo equipamento e em curtos intervalos de tempo (ISO 3534-1).</li> <li>- Proximidade de concordância entre resultados de medições sucessivas de um mesmo mensurando conduzidas nas mesmas condições de medição (IUPAC Orange Book).</li> </ul>
Precisão intermediária/Reprodutibilidade intralaboratorial	<ul style="list-style-type: none"> <li>- Medidas entre diferentes analistas, em intervalos de tempo extensos, em um único laboratório, mas sob condições bem estabelecidas.</li> <li>- Expressa variações intralaboratoriais: diferentes dias, diferentes analistas, diferentes equipamentos, etc. (ICH Q2A, CPMP/ICH/381/95).</li> </ul>

(continua)

Parâmetro de desempenho	Definição
Reprodutibilidade/ Reprodutibilidade interlaboratorial	<ul style="list-style-type: none"> <li>- Maior medida de precisão normalmente encontrada, embora exclua formalmente variações relativas a tempo.</li> <li>- Condições nas quais resultados de ensaios são obtidos com o mesmo método, em amostras idênticas, em diferentes laboratórios, por diferentes analistas, utilizando diferentes equipamentos. A validade da determinação da reprodutibilidade requer que as condições modificadas sejam especificadas. Pode ser expressa quantitativamente em termos de dispersão dos resultados (ISO 3534-1).</li> </ul>
Acurácia ( <i>accuracy</i> )	<ul style="list-style-type: none"> <li>- Proximidade do resultado ao valor verdadeiro. Geralmente envolve efeitos sistemáticos e aleatórios, sendo normalmente estudada como dois componentes: exatidão (<i>trueness</i>) e precisão.</li> <li>- Proximidade de concordância entre o resultado de um ensaio e o valor de referência aceito. Quando aplicado a um conjunto de resultados de ensaios, envolve a combinação de componentes aleatórios e sistemáticos (ISO 3534-1).</li> <li>- Grandeza que refere a diferença entre a média de um conjunto de resultados ou de um resultado individual e o valor o qual é aceito como verdadeiro ou correto para a grandeza medida (IUPAC Compendium of Chemical Technology, 1985).</li> </ul>
Recuperação	<ul style="list-style-type: none"> <li>- Forma de determinar a eficiência do processo de extração do analito da porção de amostra.</li> <li>- A fração de analito adicionada a amostra (fortificada ou adicionada) antes da análise, sendo que a porcentagem de recuperação é calculada como segue:  <math display="block">\% R = [(CF - CU) / CA] \times 100</math> , sendo <i>CF</i> a concentração do analito medida na amostra fortificada, <i>CU</i> a concentração do analito medida na amostra não fortificada, e <i>CA</i> a concentração do analito adicionada na amostra fortificada (AOAC - PVMC).</li> </ul>
Limite de detecção	<ul style="list-style-type: none"> <li>- Para métodos qualitativos: o limiar de concentração no qual o ensaio se torna pouco confiável</li> <li>- Para métodos quantitativos: menor concentração do analito ou valor da propriedade que pode ser confiavelmente detectada pelo método. Para propósito de validação é normalmente suficiente prover uma indicação do nível no qual a detecção se torna problemática. Para este propósito <math>\text{branco} + 3s</math> é suficiente.</li> <li>- Concentração <math>C_L</math>, ou a quantidade <math>q_L</math>, derivada da menor medição <math>x_L</math>, que pode ser detectada com certeza razoável para um dado procedimento analítico. O valor <math>x_L</math> é dado pela equação <math>x_L = x_{bi} + ks_{bi}</math>, sendo <math>x_{bi}</math> a média das medidas do branco, <math>s_{bi}</math> o desvio padrão das medidas do branco, e <math>k</math> um fator numérico escolhido de acordo com o intervalo de confiança desejado (IUPAC Compendium of Chemical Technology, 1985).</li> <li>- Concentração ou quantidade de analito no material analisado a qual irá prover com probabilidade <math>(1 - \beta)</math>, a conclusão de que a concentração de analito no material analisado é maior que o branco de matriz (ISO 11843-1).</li> <li>- Menor quantidade de analito, se realmente presente, que será detectada e que pode ser identificada (AOAC - PVMC).</li> <li>- Menor concentração do analito em uma amostra que pode ser detectado, mas não necessariamente quantificado, sob condições de ensaio estabelecidas (NATA Tech Note#13).</li> </ul>
Limite de quantificação	<ul style="list-style-type: none"> <li>- Menor concentração do analito que pode ser determinada com aceitáveis níveis de exatidão (<i>trueness</i>) e precisão (repetitividade). Também definido por várias convenções como sendo a concentração de analito correspondente ao valor da amostra branca mais cinco, seis ou dez desvios padrão da média do branco. Algumas vezes conhecido como limite de determinação.</li> <li>- Parâmetro que marca a habilidade de um processo de medição quantificar adequadamente um analito. Expresso em termos de sinal ou valor de analito (verdadeiro) que irá produzir estimativas tendo um desvio padrão relativo especificado, comumente de 10 %. Então <math>L_Q = k_Q \times \sigma_Q</math>, sendo <math>L_Q</math> o limite de quantificação, <math>\sigma_Q</math> o desvio padrão no ponto, e <math>k_Q</math> um múltiplo recíproco do desvio padrão relativo selecionado, dez para a IUPAC (IUPAC Orange Book).</li> <li>- O conteúdo igual ou maior que o menor ponto de concentração da curva de calibração (AOAC - PVMC).</li> <li>- Menor concentração de um analito que pode ser determinada com exatidão (<i>accuracy</i>) e precisão (repetitividade) aceitáveis, sob condições de ensaio estabelecidas (NATA Tech Note#13).</li> </ul>

(continua)

Parâmetro de desempenho	Definição
Robustez ( <i>ruggedness</i> ou <i>robustness</i> )	<p>- Corresponde a identificação dos estágios do método, os quais se não conduzidos com cuidado suficiente, irão causar severos efeitos na performance do método e podem até mesmo prejudicar seu bom funcionamento. Envolve provocar variações deliberadas ao método e investigar os efeitos subseqüentes na performance.</p> <p>- <i>Robustness</i>: medida da capacidade do método não ser afetado por pequenas, mas deliberadas, variações em parâmetros, fornecendo uma indicação de sua confiabilidade durante o uso normal (ICH Q2A, CPCP/ICH/381/95).</p> <p>- <i>Ruggedness</i>: estudo intralaboratorial do comportamento de um processo analítico quando pequenas mudanças em condições ambientais e/ou operacionais são feitas, semelhantes aquelas prováveis de aparecerem em diferentes ambientes de ensaios. Fornece informações sobre os efeitos de pequenas mudanças de maneira rápida e sistemática (AOAC - PVMC).</p>
Estimativa da incerteza	<p>- Parâmetro único (usualmente desvio padrão ou intervalo de confiança) que expressa a faixa de valores possíveis com base no resultado da medição.</p> <p>- Parâmetro associado ao resultado de uma medição que caracteriza a dispersão dos valores que podem razoavelmente ser atribuídos ao mensurando (VIM 1993).</p>

\* Uso intercambiável dos termos especificidade e seletividade, em contradição à recomendação pelo emprego do termo seletividade (VESSMAN et al., 2001).

\*\* Uso do termo especificidade, em contradição à recomendação pelo emprego do termo seletividade (VESSMAN et al., 2001).



**Tabela 9** - Parâmetros de desempenho e respectivas definições apresentados por HUBER (1998)

<b>Parâmetro de desempenho</b>	<b>Definição</b>
Linearidade/Curva de calibração	- Habilidade em fornecer resultados que sejam diretamente, ou por meio de transformações matemáticas bem definidas, proporcionais à concentração dos analitos nas amostras em uma dada faixa.
Faixa	- Intervalo entre o maior e menor nível (incluindo estes níveis) no qual tenha sido demonstrada linearidade, exatidão ( <i>accuracy</i> )* e precisão, utilizando o método como escrito.
Seletividade/Especificidade	- Especificidade: um método que produz resposta para um único analito somente. - Seletividade: um método o qual fornece respostas para um número de entidades químicas que podem ou não ser distinguidas umas das outras.
Exatidão ( <i>accuracy</i> )*	- Extensão na qual os resultados gerados por um método e o valor verdadeiro concordam.
Precisão	- Extensão na qual resultados de ensaios individuais de injeções múltiplas de uma série de padrões concordam.
Repetitividade	- Precisão obtida quando a análise é conduzida em um laboratório, por um analista, utilizando um equipamento e em um intervalo de tempo relativamente curto.
Precisão intermediária	- Precisão obtida comparando-se os resultados de um método, em um único laboratório, ao longo de várias semanas.
Reprodutibilidade	- Precisão obtida entre laboratórios, determinada pela análise de alíquotas de lotes homogêneos, em diferentes laboratórios, com diferentes analistas, utilizando condições operacionais e ambientais que podem diferir, mas que estejam dentro dos parâmetros especificados pelo método.
Limite de detecção	- Ponto no qual um valor medido é maior que a incerteza associada a ele. Menor concentração de analito em uma amostra que pode ser detectada mas não necessariamente quantificada. Em cromatografia, a quantidade injetada que resulta em um pico com altura pelo menos duas ou três vezes maior que o ruído da linha de base.
Limite de quantificação	- Menor quantidade injetada que fornece medições precisas. Em cromatografia, requer picos com alturas dez a vinte vezes maiores que o ruído da linha de base.
Robustez ( <i>robustness</i> )	- Exame dos efeitos de parâmetros operacionais sobre os resultados das análises.

\* Uso do termo *accuracy* para tratar exatidão, em contradição à recomendação pelo emprego do termo *trueness* (ISO, 1993 b).

**Tabela 10** - Parâmetros de desempenho e respectivas definições apresentados por EC (2002)

Parâmetro de desempenho	Definição
Curva de calibração	NA
Seletividade/Especificidade*	- Habilidade de um método distinguir entre o analito sob medição e outras substâncias. Este parâmetro é predominantemente dependente da técnica, mas pode variar de acordo com a classe do composto ou matriz.
Exatidão ( <i>trueness</i> )	- Proximidade de concordância entre o valor médio obtido de uma série de resultados de ensaios e um valor de referência aceitável. É usualmente expressa como tendência ou erro sistemático ( <i>bias</i> ).
Precisão	- Proximidade de concordância entre resultados de ensaios independentes obtidos sob condições estipuladas (predeterminadas). A medida da precisão é expressa em termos de imprecisão e computada como desvio padrão do resultado do ensaio. Menor precisão é determinada por um elevado desvio padrão.
Repetitividade	- Precisão obtida sob condições em que resultados independentes são obtidos com o mesmo método, para amostras idênticas, no mesmo laboratório, pelo mesmo analista, utilizando o mesmo equipamento.
Reprodutibilidade intralaboratorial/ Reprodutibilidade parcial Reprodutibilidade	- Precisão obtida no mesmo laboratório sob condições estipuladas (predeterminadas), relativas, por exemplo, ao método, amostras, analistas, ambiente, ao longo de intervalos de tempo justificadamente longos. - Precisão em que resultados independentes são obtidos com o mesmo método, para amostras idênticas, em diferentes laboratórios, com diferentes analistas, utilizando diferentes equipamentos.
Acurácia ( <i>accuracy</i> )	- Proximidade de concordância entre o resultado de um ensaio e um valor de referência aceitável. Determinada pela avaliação da exatidão ( <i>trueness</i> ) e precisão.
Recuperação	- Porcentagem da concentração verdadeira de uma substância recuperada durante o procedimento analítico. É determinada durante a validação quando não há material de referência certificado disponível.
Limite de decisão	- Limite no qual e acima do qual se pode concluir que uma amostra é não-conforme com uma probabilidade de erro alfa. Erro alfa é a probabilidade da amostra ensaiada ser conforme apesar de ter obtido um resultado não-conforme (falsa decisão não-conforme).
Capacidade de detecção	- Menor quantidade de uma substância que pode ser detectada, identificada e/ou quantificada numa amostra com uma probabilidade de erro beta. No caso de substâncias para as quais não haja limite permitido estabelecido, a capacidade de detecção é a concentração mais baixa na qual o método é capaz de detectar amostras verdadeiramente contaminadas com uma probabilidade de $(1 - \beta)$ . No caso de substâncias com um limite permitido estabelecido, a capacidade de detecção é a concentração na qual o método é capaz de detectar a concentração do limite permitido com uma probabilidade de $(1 - \beta)$ . Erro beta é a probabilidade da amostra ensaiada ser não-conforme, apesar de ter obtido um resultado conforme (falsa decisão conforme).
Robustez ( <i>ruggedness</i> )	- Susceptibilidade de um método de ensaio às alterações das condições experimentais, as quais podem ser expressas como tipos de amostras, analitos, condições de armazenamento, condições ambientais ou de preparo de amostra em que o método pode ser aplicado tal como apresentado ou com pequenas alterações específicas. Para todas as condições experimentais que possam, na prática, estar sujeitas a flutuações (por exemplo: estabilidade de reagentes, composição da amostra, pH, temperatura) qualquer variação que puder alterar o resultado deve ser indicada.
Estabilidade	NA
Aplicabilidade**	NA

\* Uso intercambiável dos termos especificidade e seletividade, em contradição à recomendação pelo emprego do termo seletividade (VESSMAN et al., 2001).

\*\* Trata aplicabilidade associada ao estudo da robustez.

NA = definição não apresentada.

**Tabela 11** - Parâmetros de desempenho e respectivas definições apresentados por THOMPSON, ELLISON & WOOD (2002)

Parâmetro de desempenho	Definição
Calibração/Linearidade/ Teste de efeitos de matriz	NA
Faixa	- Intervalo de concentração do analito no qual o método pode ser considerado como validado.
Sensibilidade	- Gradiente da função de calibração.
Seletividade	- Grau no qual um método pode quantificar o analito acuradamente na presença de interferentes.
Exatidão ( <i>trueness</i> )	- Proximidade de concordância entre resultados de ensaios com o valor de referência aceito para a propriedade sendo medida. Estabelecida quantitativamente em termos de tendência ( <i>bias</i> ), com menor tendência indicando maior exatidão.
Precisão	- Proximidade de concordância entre resultados de ensaios independentes obtidos sob condições estipuladas. Usualmente especificado em termos de desvio padrão ou desvio padrão relativo.
Repetitividade/Precisão intra-ensaio	NA
Precisão inter-ensaios	NA
Reprodutibilidade	NA
Limite de detecção	- Menor quantidade ou concentração de analito em uma amostra que pode ser confiavelmente distinguida de zero.
Limite de determinação/Limite de quantificação	- Concentração abaixo da qual o método analítico não pode operar com precisão aceitável. Algumas vezes esta precisão é arbitrariamente definida como 10 % do desvio padrão relativo, outras vezes o limite também é arbitrariamente tido como um múltiplo fixo (tipicamente 2) do limite de detecção.
Robustez ( <i>ruggedness</i> )	- Resistência a mudar os resultados produzidos por um método analítico quando pequenos desvios são feitos nas condições experimentais descritas no procedimento.
Estimativa de incerteza	NA
Varição de matriz	NA
Aplicabilidade	NA
Adequação para uso	- Extensão na qual a performance do método se equipara aos critérios acordados entre o analista e o usuário final dos dados.

NA = definição não apresentada.

**Tabela 12 - Parâmetros de desempenho e respectivas definições apresentados por INMETRO (2003)**

Parâmetro de desempenho	Definição
Linearidade	- Habilidade de um método analítico em produzir resultados que sejam diretamente proporcionais à concentração do analito em amostras, em uma dada faixa de concentração.
Faixa de trabalho/Faixa linear de trabalho	- Faixa de trabalho é a faixa de concentrações do analito ou valores da propriedade no qual o método pode ser aplicado. Faixa linear de trabalho é o intervalo entre os níveis inferior e superior de concentração do analito no qual foi demonstrado ser possível a determinação com precisão, exatidão e linearidade exigidas, sob as condições especificadas para o ensaio.
Sensibilidade	- Parâmetro que demonstra a variação da resposta em função da concentração do analito. Pode ser expresso pela inclinação da curva de regressão linear de calibração e é determinado simultaneamente aos testes de linearidade. Depende da natureza do analito e da técnica de detecção utilizada.
Seletividade/Especificidade	- Parâmetros relacionados ao evento da detecção. Um método específico produz resposta para apenas um analito. Um método seletivo produz respostas para vários analitos, mas pode distinguir a resposta de um analito da de outros.
Exatidão* e tendência ( <i>bias</i> )	- Concordância entre o resultado de um ensaio e o valor de referência aceito como convencionalmente verdadeiro. Quando aplicada a uma série de resultados de ensaio, implica numa combinação de componentes de erros aleatórios e sistemáticos (tendência).
Precisão	- Dispersão de resultados entre ensaios independentes, repetidos de uma mesma amostra, amostras semelhantes ou padrões, em condições definidas. É normalmente determinada para circunstâncias específicas de medição e as duas formas mais comuns de expressá-la são por meio da repetitividade e a reprodutibilidade, sendo usualmente expressa pelo desvio padrão.
Repetitividade	- Grau de concordância entre os resultados de medições sucessivas de um mesmo mensurando, efetuadas sob as mesmas condições de medição: mesmo procedimento de medição, mesmo observador, mesmo instrumento usado sob as mesmas condições, mesmo local e repetições em curto espaço de tempo. Expressa quantitativamente em termos de característica de dispersão dos resultados.
Precisão intermediária	- Precisão sob condições de reprodutibilidade em que os resultados dos ensaios são obtidos com o mesmo método, variando-se laboratórios**, operador ou equipamentos. - Precisão avaliada sobre a mesma amostra, amostras idênticas ou padrões, utilizando o mesmo método, no mesmo laboratório ou em laboratórios diferentes**, mas definindo exatamente quais condições a variar (uma ou mais), tais como: diferentes analistas, diferentes equipamentos e diferentes tempos.
Reprodutibilidade	- Grau de concordância entre os resultados das medições de um mesmo mensurando, efetuadas sob condições variadas de medição.
Recuperação	NA
Limite de detecção	- Limite de detecção do equipamento é a concentração do analito que produz um sinal de três a cinco vezes a razão ruído/sinal*** do equipamento. Limite de detecção do método é a concentração mínima de uma substância medida e declarada com 95 % ou 99 % de confiança de que a concentração do analito é maior que zero. Este limite é determinado através da análise completa de uma dada matriz contendo o analito.
Limite de quantificação/Limite de determinação	- Menor concentração do analito que pode ser determinada com um nível aceitável de precisão e veracidade ( <i>trueness</i> ). Pode ser considerado como sendo a concentração do analito correspondente ao valor da média do branco mais cinco, seis ou dez desvios padrão. Corresponde ao padrão de calibração de menor concentração (excluindo o branco).
Robustez	- Mede a sensibilidade que o método apresenta face a pequenas variações.
Incerteza da medição	NA

\* Uso da definição do termo acurácia (*accuracy*) para exatidão, em contradição à recomendação pelo emprego do termo *trueness* (ISO, 1993 b).

\*\* A precisão intermediária pode ser avaliada como parte de estudos interlaboratoriais, entretanto, este parâmetro está associado à avaliação da precisão inter-ensaios em condições intralaboratoriais.

\*\*\* A relação correta seria sinal/ruído e não ruído/sinal.

NA = definição não apresentada.

**Tabela 13** - Parâmetros de desempenho e respectivas definições apresentados por SHABIR (2003)

Parâmetro de desempenho	Definição
Linearidade	- Habilidade de um método analítico em produzir resultados que sejam diretamente proporcionais à concentração do analito em amostras, em uma dada faixa.
Faixa	- Intervalo entre os maiores e menores níveis de analito (inclusive) que tenham sido demonstrados com linearidade, exatidão ( <i>accuracy</i> )** e precisão, utilizando o método como escrito.
Especificidade*	- Habilidade de um método em medir acuradamente a resposta do analito na presença de todos os componentes potenciais da amostra.
Exatidão ( <i>accuracy</i> )**	- Proximidade dos resultados de ensaios obtidos pelo método analítico em relação ao valor verdadeiro.
Precisão	- Medida do grau de repetitividade*** de um método analítico sob condições normais de operação e normalmente expressa como a porcentagem de desvio padrão relativo para um número estatisticamente significativo de amostras.
Repetitividade/Precisão intra-ensaio	- Resultados de um método operado em um curto intervalo de tempo, sob as mesmas condições.
Precisão intermediária	- Resultados de variações intralaboratoriais devidas a eventos aleatórios como diferentes dias, analistas, equipamentos, etc.
Reprodutibilidade	- Determinação de amostras homogêneas em múltiplos laboratórios, freqüentemente parte de estudos interlaboratoriais.
Limite de detecção	- Menor concentração de um analito em uma amostra que pode ser detectada e não quantificada. Expressa como uma relação sinal/ruído de usualmente 3:1.
Limite de quantificação	- Menor concentração de um analito em uma amostra que pode ser determinada com exatidão ( <i>accuracy</i> )** e precisão aceitáveis, sob condições operacionais determinadas no método. Expressa como uma relação sinal/ruído de 10:1.
Robustez ( <i>robustness</i> )	- Habilidade de um método não ser afetado por pequenas e deliberadas variações em seus parâmetros.

\* Uso do termo especificidade, em contradição à recomendação pelo emprego do termo seletividade (VESSMAN et al., 2001).

\*\* Uso do termo *accuracy* para tratar exatidão, em contradição à recomendação pelo emprego do termo *trueness* (ISO, 1993 b).

\*\*\* Relação do termo precisão somente à condição de repetitividade, quando este parâmetro também pode ser estudado sob outras condições experimentais, como precisão intermediária ou reprodutibilidade.

**Tabela 14** - Parâmetros de desempenho e respectivas definições apresentados por TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004)

Parâmetro de desempenho	Definição
Linearidade	- Habilidade do método em obter resultados de ensaios proporcionais à concentração do analito (em uma dada faixa).
Faixa de trabalho/Faixa linear/ Limites de linearidade	- Faixa de concentrações/quantidades de analito ao longo da qual o método fornece resultados de ensaios proporcionais a concentração do analito, ou na qual um modelo linear pode ser aplicado com um nível de confiança conhecido.
Sensibilidade	- Alteração na resposta de um instrumento de medição dividida pela mudança correspondente no estímulo.
Seletividade/Especificidade	- Habilidade do método em determinar acurada e especificamente o analito de interesse na presença de outros componentes na matriz (que sejam esperados de estarem presentes na matriz) sob condições estipuladas de ensaios (especificidade = 100 % de seletividade).
Exatidão ( <i>trueness</i> )	- Proximidade de concordância entre a expectativa do resultado de um ensaio (valor médio esperado) e um valor de referência aceito (valor verdadeiro).
Precisão	- Proximidade de concordância entre resultados de ensaios independentes obtidos sob condições estipuladas.
Repetitividade/Precisão intra-ensaio	- Precisão sob condições em que resultados de ensaio independentes são obtidos com o mesmo método, em amostras idênticas, no mesmo laboratório, pelo mesmo analista, utilizando o mesmo equipamento e dentro de curtos intervalos de tempo.
Precisão intermediária/Precisão inter-ensaios	- Precisão sob condições em que resultados de ensaio independentes são obtidos com o mesmo método, em amostras idênticas, no mesmo laboratório, mas por diferentes analistas, utilizando diferentes equipamentos, ao longo de um período de tempo ampliado.
Reprodutibilidade/Precisão interlaboratorial	- Precisão sob condições em que resultados de ensaio independentes são obtidos com o mesmo método, em amostras idênticas, em diferentes laboratórios, com diferentes analistas, utilizando diferentes equipamentos.
Acurácia ( <i>accuracy</i> )	- Proximidade de concordância entre um resultado de ensaio e o valor de referência aceito.
Recuperação	- Fração do analito adicionada à amostra (fortificada ou adicionada) antes da análise, a qual é medida pelo método.
Limite de detecção	- Para métodos qualitativos: o limiar de concentração no qual o ensaio se torna pouco confiável. - Para métodos quantitativos: menor concentração ou quantidade de analito que pode ser confiavelmente distinguida e zero, ou que pode ser detectada/medida com certeza estatística razoável.
Limite de quantificação	- Menor concentração ou quantidade de analito que pode ser determinada quantitativamente com níveis aceitáveis de exatidão ( <i>trueness</i> ) e precisão (repetitividade).
Robustez	- <i>Ruggedness</i> : perfil (testado intralaboratorialmente) de um processo analítico quando pequenas mudanças são feitas em condições ambientais e/ou operacionais. - <i>Robustness</i> : medida da capacidade de um processo analítico não ser afetado por pequenas, mas deliberadas, variações nos parâmetros de performance de um método, a qual provê uma indicação da confiabilidade do método em uso normal.

Segundo EURACHEM (1998), INMETRO (2003) e TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004), a faixa de trabalho é geralmente limitada no extremo inferior pelos valores dos limites de detecção ou de quantificação, enquanto no limite superior as restrições são impostas por efeitos dependentes do sistema de detecção do equipamento. Dentro da faixa de trabalho deve existir uma faixa de resposta linear. THOMPSON, ELLISON & WOOD (2002) também relatam que a faixa de trabalho (ou de validação) não é necessariamente idêntica à faixa linear (ou de calibração) mas, ao contrário dos outros autores, destacam que dentro da faixa linear existe uma faixa de trabalho. Enquanto a calibração pode cobrir um intervalo mais amplo de concentração, a validação pode cobrir uma faixa mais restrita. Na prática, a maioria dos métodos é validada somente em um ou dois níveis de concentração.

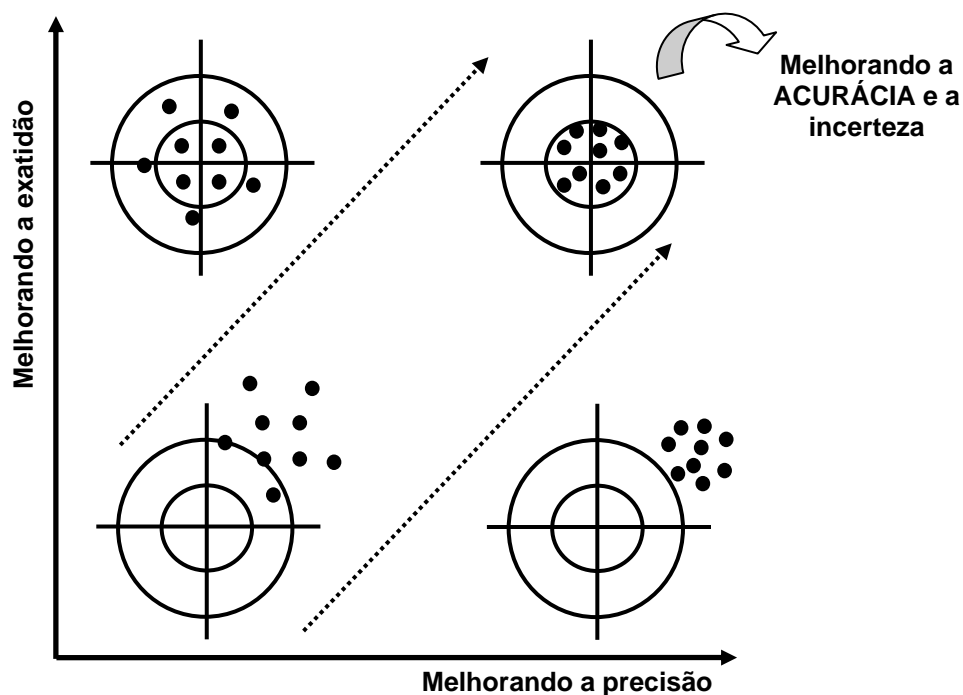
A capacidade de discriminar o analito de outros compostos interferentes tem sido, por muitos anos, expressa como seletividade de um método ou de um sistema de medição. Entretanto, o mesmo significado é comumente atribuído ao termo especificidade (INMAN et al., 1987; GREEN 1996; ICH, 1996; JENKE, 1996; AOAC, 1998 b; HILL & REYNOLDS, 1999; SHABIR, 2003). Algumas referências utilizam os dois termos, seletividade e especificidade, de forma intercambiável (NATA, 1997; EURACHEM, 1998; EC, 2002). Outras consideram ambos os termos, mas apresentam definições distintas para cada um deles (BRUCE, MINKKINEN & RIEKKOLA, 1998; HUBER, 1998; INMETRO, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Desde 1983, a IUPAC tenta esclarecer esta sobreposição com a afirmativa de que “especificidade é o ultimato da seletividade” (DEN BOEF & HULANICKI, 1983). Em 2001, foi apresentada, também pela IUPAC, uma discussão para diferenciação no uso destes termos, seguida pela recomendação de promoção do emprego do termo seletividade em química analítica para expressar a extensão na qual um método, em particular, pode ser utilizado para determinar analitos na presença de outros componentes de comportamento semelhante. O uso do termo especificidade neste contexto foi desencorajando (VESSMAN et al., 2001). É importante mencionar que em análises biológicas e microbiológicas os termos seletividade e especificidade são empregados para tratar porcentagens de resultados falsos positivos e falsos negativos, de forma diferente e conflitante com a adotada em química analítica (EURACHEM, 1998; ISO, 2003).

O termo acurácia (*accuracy*) tem sido relacionado a tantos significados que algumas organizações preferiram substituí-lo por termos com interpretações específicas como exatidão (*trueness*), tendência ou tendenciosidade (*bias*) e recuperação (*recovery*) (HORWITZ, 2003). Embora a ISO 3534-1 (ISO, 1993 b) recomende que o termo acurácia

(*accuracy*) não seja utilizado para tratar exatidão (*trueness* ou *exactness*) (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004), este tratamento é freqüentemente encontrado na literatura (GREEN, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; HUBER, 1998; INMETRO, 2003; SHABIR, 2003). A acurácia, proximidade de concordância entre um resultado de ensaio e o valor de referência aceito (ISO, 1993 a,b), está relacionada a uma combinação de erros sistemáticos e aleatórios, e não somente a erros sistemáticos. Se por um lado, a exatidão trata da concordância entre o valor médio obtido de uma grande série de resultados de ensaios e um valor de referência aceito, sendo expressa em termos de tendência (*bias*) ou erro sistemático total. E, por outro lado, a precisão faz referência à proximidade de concordância entre resultados de ensaios independentes obtidos sob condições estipuladas, dependendo somente da distribuição aleatória dos erros, sem possuir relação com o valor verdadeiro ou especificado (ISO, 1993 b; RSC, 2003). Pode-se concluir que a acurácia está relacionada a uma combinação dos parâmetros exatidão e precisão, e não somente à exatidão (EURACHEM, 1998; RSC, 2003), conforme esquematizado na Figura 4. Outro problema relacionado à terminologia atribuída a estes parâmetros encontra-se nas diferentes traduções feitas do inglês para a língua portuguesa. Por exemplo, alguns autores traduzem *accuracy* como exatidão (EC, 2002 versão em português; INMETRO, 2003) e *trueness* como veracidade (EC, 2002 versão em português). Os termos adotados neste trabalho foram exatidão (*trueness*), tendência (*bias*) e acurácia (*accuracy*).

A IUPAC sugere distinção entre os termos “recuperação” e “recuperação aparente” para evitar a confusão causada pelo uso do termo recuperação para cobrir duas situações distintas. Sendo: i) o rendimento de uma etapa de pré concentração ou extração de um processo analítico, no qual o termo “recuperação” é recomendado; e ii) a quantidade (valor observado/valor de referência) obtida por um procedimento analítico que envolve curva de calibração, em que “recuperação aparente” é recomendada (BURNS, DANZER & TOWNSHEND, 2002). Apesar da existência desta recomendação da IUPAC, na literatura consultada não foi observado o emprego do termo “recuperação aparente”, o qual foi utilizado nesta tese.





**Figura 4** - Combinação da exatidão (erros sistemáticos) e da precisão (erros aleatórios) para indicação de acurácia (RSC, 2003 - adaptado).

Diferentes termos também são apresentados na literatura para expressar as diversas condições de estimativa da precisão. Condições de repetitividade (GREEN, 1996; EAL, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004) são também denominadas precisão intra-ensaio (GREEN, 1996; BRUCE, MINKKINEN & RIEKKOLA, 1998; THOMPSON, ELLISON & WOOD, 2002; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Condições de precisão intermediária (GREEN, 1996; EURACHEM, 1998; HUBER, 1998; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004) são tratadas como precisão de cotidiano (*day-to-day*) (BRUCE, MINKKINEN & RIEKKOLA, 1998), precisão inter-ensaios (BRUCE, MINKKINEN & RIEKKOLA, 1998; THOMPSON, ELLISON & WOOD, 2002; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004), reprodutibilidade interna (FEINBERG & RAGUÈNÈS, 1999), reprodutibilidade parcial ou reprodutibilidade intralaboratorial (EURACHEM, 1998; EC, 2002). Enquanto condições de reprodutibilidade (GREEN, 1996; EAL, 1997; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS,

DE LOOSE & VAN BOCKSTAELE, 2004) aparecem ainda como reprodutibilidade interlaboratorial (EURACHEM, 1998) e precisão interlaboratorial (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Os termos repetitividade, reprodutibilidade parcial e reprodutibilidade foram os empregados neste trabalho.

A AOAC (1998 b) apresenta os termos limite de determinação e limite de decisão como sinônimos e correspondentes a uma concentração intermediária aos limites de detecção e quantificação. Todavia, outras referências empregam o termo limite de decisão como sinônimo de limite de detecção (BRUCE, MINKKINEN & RIEKKOLA, 1998) e o termo limite de determinação como sinônimo de limite de quantificação (BRUCE, MINKKINEN & RIEKKOLA, 1998; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003). A EC (2002) adota os parâmetros limite de decisão ( $CC\alpha$ ) e capacidade de detecção ( $CC\beta$ ) para ensaios em setores específicos da área de alimentos (contaminantes e resíduos orgânicos em animais vivos e produtos de origem animal), no âmbito da União Européia. Para analitos proibidos, sem limites máximos estabelecidos, estes parâmetros são comparáveis aos limites de detecção e quantificação, respectivamente, uma vez que suas concentrações correspondem a medidas acima do sinal obtido para amostras brancas (matrizes isentas do analito ou para as quais o analito não tenha sido detectado). No entanto, para substâncias com limites máximos estabelecidos, estes parâmetros não podem ser relacionados aos limites de detecção e quantificação, pois são expressos em relação aos limites. Estes parâmetros, também recomendados pela IUPAC (CURRIE, 1995) e pela ISO 11843 (ISO, 1997 b; ISO, 2000), foram baseados nos estudos preliminares de CURRIE (1968) e HUBAUX & VOS (1970). Contudo, IUPAC (CURRIE, 1995) e ISO 11843 (ISO, 1997 b; ISO, 2000) empregam símbolos diferentes ( $x_c$  para limite de decisão e  $x_d$  para capacidade de detecção) e não distinguem procedimentos para analitos permitidos e proibidos. Apesar destes documentos serem bastante completos e complexos, constituem referências de difícil compreensão pelos analistas (ANTIGNAC et al., 2003; PECORELLI et al., 2004), o que tem estimulado a criação de fóruns de harmonização (CURRIE, 1999) e de discussão sobre estes parâmetros de desempenho (DESIMONI, 2004). Nesta tese, limite de detecção, limite de quantificação, limite de decisão ( $CC\alpha$ ) e capacidade de detecção ( $CC\beta$ ) foram os termos e símbolos adotados.

Em função do quadro ainda heterogêneo sobre os termos e definições dos parâmetros de desempenho, grandes dificuldades podem ser observadas tanto na apresentação quanto na interpretação de resultados de processos de validação. Enquanto

o processo de padronização avança, análise crítica é extremamente válida na busca por clareza, objetividade e entendimento do assunto abordado.

### **2.2.2 Principais parâmetros em processos de validação**

O processo de validação demanda uma decisão sobre quais parâmetros de desempenho precisam ser avaliados (EURACHEM, 1998). JENKE (1996) avaliou a frequência dos parâmetros de desempenho relacionados a processos de validação em publicações de métodos das áreas governamental, industrial e acadêmica. Os parâmetros mais frequentes foram exatidão e precisão (repetitividade e reprodutibilidade), seguidos de seletividade, linearidade, limite de quantificação, limite de detecção, robustez e sensibilidade. Este autor argumentou que a exclusão dos parâmetros mais frequentemente aplicados nos protocolos de validação, como exatidão e precisão, implicaria na necessidade de justificativas pertinentes, enquanto a exclusão de parâmetros menos frequentes, como sensibilidade, seria menos controversa. Entretanto, é importante argumentar que a seleção de parâmetros de desempenho para um determinado processo de validação também está fortemente relacionada com a forma de validação adotada (por exemplo, a avaliação da sensibilidade pode ser essencial em protocolos de validação por comparação de métodos).

As Tabelas 3 a 14 mostram a variabilidade dos parâmetros de desempenho recomendados na literatura para processos de validação intralaboratoriais.

De maneira geral, os principais parâmetros estudados são linearidade, faixa de trabalho, sensibilidade, seletividade, exatidão, precisão, limite de detecção e limite de quantificação (GREEN, 1996; ICH, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; KING, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Muitos documentos não consideram a sensibilidade como parâmetro fundamental da validação (GREEN, 1996; EAL, 1997; HUBER, 1998; EC, 2002; SHABIR, 2003), visto que se trata de um parâmetro arbitrário que depende de ajustes no instrumento de medição (THOMPSON, ELLISON & WOOD, 2002). Os limites também são indicados somente em ensaios de traços, não sendo necessários em ensaios de analitos que correspondem ao principal componente da matriz (INMAN et al., 1987; KING, 2003).

Alguns autores incluem outros parâmetros de desempenho como parte de processos de validação como adequação do sistema analítico (ICH, 1996); estabilidade de soluções (INMAN et al., 1987; GREEN, 1996; HILL & REYNOLDS, 1999; EC, 2002);

robustez (GREEN, 1996; JENKE, 1996; NATA, 1997; BRUCE, MINKKINEN & RIEKKOLA, 1998; EC, 2002; INMETRO, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004); recuperação aparente (INMAN et al., 1987; NATA, 1997; AOAC, 1998 b; EURACHEM, 1998; EC, 2002; INMETRO, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004); amostragem e preparo das amostras (HORWITZ, 2003; KING, 2003); estimativa da incerteza (EAL, 1997; NATA, 1997; EURACHEM, 1998; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; KING, 2003); praticabilidade e transferibilidade entre laboratórios (BRUCE, MINKKINEN & RIEKKOLA, 1998); aplicabilidade (EC, 2002; THOMPSON, ELLISON & WOOD, 2002; HORWITZ, 2003); variação de matriz e adequação para uso (THOMPSON, ELLISON & WOOD, 2002). No entanto, não existe consenso sobre a inclusão destes parâmetros na validação de métodos.

SHABIR (2003) trata os estudos de estabilidade de soluções, qualificação de equipamentos e adequação das condições do sistema analítico, em uma etapa anterior à validação, a qual este autor denomina pré-validação.

O parâmetro robustez é freqüentemente considerado fora do processo de validação, como parte de estudos de desenvolvimento do método (ICH, 1996; EURACHEM, 1998; HUBER, 1998; SHABIR, 2003). Assim, o estudo da robustez em processos de validação somente é considerado necessário quando os limites para os parâmetros experimentais e seus respectivos desvios permitidos não forem previamente estabelecidos (THOMPSON, ELLISON & WOOD, 2002). De qualquer forma, a robustez é importante para fornecer informações sobre o efeito de parâmetros experimentais significativos para a estimativa da incerteza (EURACHEM/CITAC, 2000).

Muitas vezes, a recuperação aparente também não é abordada como parâmetro de desempenho em processos de validação, mas sim como uma ferramenta para avaliação dos parâmetros exatidão e precisão (HUBER, 1998; HILL & REYNOLDS, 1999; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003).

HILL & REYNOLDS (1999) debatem que processos como preparo de amostras (diferentemente do processamento de amostras) podem ser impossíveis de serem validados.

A ISO/IEC 17025 (ISO, 2005 b) analisa a necessidade da estimativa da incerteza, no contexto de métodos de ensaio, mas não como um parâmetro de desempenho em processos de validação. Esta norma considera que uma estimativa razoável da incerteza deve estar baseada no conhecimento do desempenho do método e no escopo da medição, devendo fazer uso, por exemplo, de experiência e dados de validação

anteriores. Segundo TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004) a estimativa da incerteza é um indicador chave tanto da adequação para uso de um método quanto da confiabilidade dos resultados analíticos obtidos em um laboratório. Este parâmetro cobre todas as fontes de erros do processo analítico, além daquelas obtidas nos processos de validação de métodos. Na prática, embora os dados provenientes de validações intralaboratoriais de métodos e de estudos colaborativos constituam a base da estimativa da incerteza, estes dados representam somente uma parte desta, sugerindo que a estimativa da incerteza seja considerada como mais do que simplesmente um parâmetro de desempenho de validação de método.

Aplicabilidade e adequação para uso são tratadas, na maioria das referências, como processos complementares à validação e não como parâmetros de desempenho avaliados durante este processo. Os próprios autores que citam a aplicabilidade como parâmetro de desempenho em validações de métodos se contradizem quando discorrem sobre o tema. “Após a validação, a documentação deve fornecer além da especificação dos parâmetros de desempenho, informações sobre a identidade do analito, faixa de concentração, tipos de matrizes, o protocolo do método (descrevendo equipamentos, reagentes, o procedimento e variações permissíveis), o propósito de uso e requisitos de incerteza críticos” (THOMPSON, ELLISON & WOOD, 2002; HORWITZ, 2003). Já a adequação para uso é a propriedade dos dados produzidos por um processo de medição habilitarem seus usuários a tomarem decisões tecnicamente corretas para o propósito estabelecido (THOMPSON & FEARN, 1996), o que envolve resultados da validação, da estimativa da incerteza, entre outros requisitos. No passado, a validação de métodos tendia a se concentrar somente no processo de avaliação dos parâmetros de desempenho (EURACHEM, 1998). Atualmente, sabe-se que este processo culmina na avaliação da adequação dos parâmetros estudados para o propósito de uso do método.

### **2.2.3 Procedimentos para avaliação**

As Tabelas 15 a 21 contêm alguns procedimentos recomendados na literatura para avaliação da linearidade, efeitos de matriz, seletividade, exatidão, precisão, limite de detecção e limite de quantificação. Usualmente, os procedimentos não são harmonizados em termos de delineamentos experimentais e ferramentas estatísticas para análises de dados, além de serem descritos de forma superficial, com formas de cálculos esporadicamente detalhadas. Na maioria das vezes, os princípios básicos dos processos de validação, como representatividade e verificação de premissas, não são mencionados.

**Tabela 15 - Recomendações para avaliação da linearidade na literatura**

Referência	Delineamento experimental	Análise de dados
GREEN (1996)	<ul style="list-style-type: none"> <li>- Preparar soluções padrão em cinco níveis de concentração, de 50 % a 150 % a concentração alvo do analito.</li> <li>- Analisar os padrões, no mínimo, três vezes.</li> </ul>	<ul style="list-style-type: none"> <li>- Examinar o coeficiente de correlação e o intercepto da curva de regressão linear da resposta <i>versus</i> concentração. Um coeficiente de correlação maior que 0,999 é geralmente considerado como evidência de ajuste aceitável dos dados. O intercepto deve ser menor que um pequeno percentual (2 %) da resposta obtida para o analito no nível alvo.</li> <li>Observação: esta referência destaca a importância do exame visual do gráfico da resposta <i>versus</i> a concentração.</li> </ul>
ICH (1996)	<ul style="list-style-type: none"> <li>- Mínimo de cinco níveis de concentração na faixa de trabalho.</li> <li>Observação: esta referência especifica faixas mínimas de 80 % a 120 % ou 70 % a 130 % a concentração teste dependendo da aplicação.</li> </ul>	<ul style="list-style-type: none"> <li>- Fazer a inspeção visual do gráfico do sinal em função da concentração do analito. Se houver relação linear, avaliar os resultados por métodos estatísticos apropriados como por exemplo o cálculo da curva de regressão pelo método dos mínimos quadrados. Em alguns casos os dados precisam ser submetidos à transformação matemática antes da análise de regressão.</li> <li>- Estimar o coeficiente de correlação, inclinação e interseção da curva e soma de quadrados dos resíduos da regressão.</li> <li>- Analisar os desvios dos resultados obtidos da linha de regressão (resíduos).</li> </ul>
ISO (1996)	<ul style="list-style-type: none"> <li>- Analisar materiais de referência em pelo menos três níveis, pelo menos duplicata (sendo recomendado quanto mais replicatas for possível na prática).</li> <li>- A faixa coberta pelos materiais de referência deve incluir a faixa de valores encontrados em condições de operação normais do sistema de medição.</li> <li>- A composição dos materiais de referência deve ser a mais próxima possível à composição do material a ser medido.</li> <li>- Os valores dos materiais de referência devem ser distribuídos aproximadamente eqüidistantes sobre a faixa.</li> <li>- O número de replicatas deve ser o mesmo para todos os níveis.</li> <li>- O tempo e condições nas quais as replicadas são tomadas deve cobrir uma ampla faixa para garantir que todas as condições operacionais sejam representadas.</li> <li>Observação: esta referência considera premissas i) não há erro nos valores aceitos para os materiais de referência (esta premissa não é verificada mas é considerada válida se as incertezas dos materiais de referência forem pequenas quando comparadas a magnitude dos erros nos valores medidos); ii) a equação de calibração é assumida como sendo linear; iii) as medidas repetidas do material de referência são assumidas com independentes e iv) normalmente distribuídas (estas premissas não são verificadas); v) o desvio padrão dos resíduos da regressão é assumido como constante ou proporcional ao valor aceito do material de referência.</li> </ul>	<ul style="list-style-type: none"> <li>- Plotar os dados para verificar o controle do sistema de medição durante o experimento de calibração, a premissa de linearidade e a variabilidade das medidas em função dos valores aceitos dos materiais de referência.</li> <li>- Estimar da curva de calibração linear sob a premissa de que o desvio padrão dos resíduos da regressão é constante.</li> <li>- Plotar a equação de calibração e os resíduos da regressão.</li> <li>- Avaliar o gráfico de resíduos que é um forte indicador de desvios das premissas de linearidade e desvio padrão constante. Se houver indicativo de que o desvio padrão dos resíduos da regressão não é constante, estimar a curva de calibração linear sob a premissa de que o desvio padrão dos resíduos da regressão é proporcional.</li> <li>- Avaliar o desvio da linearidade. Se a variabilidade devida ao desvio da linearidade for significativa em relação à variabilidade devida às medidas em replicatas, investigar os procedimentos adotados durante o experimento de calibração e reavaliar a premissa de linearidade. Se o desvio de linearidade persistir, adotar um modelo alternativo.</li> <li>Observação: esta referência detalha as equações do modelo e para estimativa da inclinação, interseção, resíduos da regressão, soma de quadrados dos resíduos e análise de variância para avaliação do desvio da linearidade.</li> </ul>

(continua)

Referência	Delineamento experimental	Análise de dados
NATA (1997)	<ul style="list-style-type: none"> <li>- Analisar amostras adicionadas e padrões. Pelo menos sete determinações em cada uma das cinco concentrações na faixa de trabalho.</li> </ul>	<ul style="list-style-type: none"> <li>- Plotar o gráfico dos resultados dos ensaios em função da concentração do analito. A representação visual do gráfico freqüentemente revela anomalias escondidas pelos números.</li> <li>- Calcular a curva de regressão usando o método dos mínimos quadrados para os resultados dos testes contra as concentrações do analito.</li> <li>- A inclinação da curva de regressão e sua variância fornecem uma medida matemática da linearidade e o intercepto é uma medida da potencial tendência do método.</li> </ul>
AOAC (1998 b)	<ul style="list-style-type: none"> <li>- Curva de calibração: preparar a curva com um mínimo de cinco concentrações de soluções padrão (sem o zero).</li> <li>- Linearidade: utilizar padrões externos do analito em quatro concentrações nominais 0x, ½x, 1x e 2x, em três dias separados (onde x é o nível normal, usualmente encontrado ou o nível alvo de ação).</li> </ul>	<ul style="list-style-type: none"> <li>- Testar estatisticamente e expressar a curva.</li> </ul>
BRUCE, MINKKINEM & RIEKKOLA (1998)	<ul style="list-style-type: none"> <li>- Preparar soluções do analito com ou sem matriz.</li> <li>- Diluir estas soluções em pelo menos cinco diferentes níveis de concentração, cobrindo toda a faixa de trabalho.</li> <li>- Medir as respostas destas soluções de analito.</li> </ul>	<ul style="list-style-type: none"> <li>- Ajustar os resultados por regressão linear não ponderada, sem forçar a curva passar pela origem.</li> <li>- Coeficiente de correlação é freqüentemente usado como uma medida de linearidade da curva de calibração (adequado quando maior ou igual a 0,995 para amostras biológicas ou 0,997 para amostras farmacêuticas por cromatografia líquida de alta eficiência).</li> <li>- O método pode ser considerado não tendencioso se a curva passar pela a origem, dentro de um intervalo de confiança de 95 %.</li> <li>- Um método simples para julgar a adequação da curva de calibração é a inspeção visual dos resíduos da regressão. Observação: esta referência cita que estatísticas mais sofisticadas são apresentadas para avaliar linearidade.</li> </ul>
DANZER & KURRIE (1998)	<p>Observação: esta referência explica o modelo e fundamentos da regressão linear simples, além das premissas i) média dos resíduos da regressão é zero; ii) resíduos da regressão não são correlacionados; iii) resíduos da regressão devem ser homoscedásticos (quando não forem aplicar regressão ponderada); iv) erro na variável x é considerado negligível (se o erro em x for considerável, utilizar ajuste por regressão ortogonal); v) resíduos da regressão seguem a distribuição normal (se não houver distribuição normal utilizar calibração robusta).</p>	<ul style="list-style-type: none"> <li>- Estimar inclinação, interseção, resíduos da regressão e respectivos desvios padrão.</li> <li>- Realizar inspeção visual do gráfico de resíduos da regressão para obter informações preliminares sobre as características dos resíduos.</li> <li>- Avaliar a linearidade pelo teste de F. <i>A priori</i>, sem levar em consideração nenhum modelo não linear, comparar o desvio padrão dos resíduos da regressão com o desvio padrão dos valores de y em relação as suas médias. <i>A posteriori</i>, considerando um dado modelo não linear, comparar o desvio padrão dos resíduos da regressão do modelo linear com o do modelo não linear.</li> <li>- Avaliar homoscedasticidade pelo teste de F máximo (em casos com o mesmo número de replicatas nos pontos de calibração) ou pelo teste de Bartlett (para casos mais obscuros).</li> <li>- Avaliar os parâmetros da calibração por comparação com os valores teoricamente esperados pelo teste de t.</li> </ul> <p>Observação: esta referência detalha as equações (do modelo, para estimativa da inclinação, interseção, resíduos da regressão, desvios e testes, nas condições de regressão simples, ponderada e ortogonal) e desaconselha o uso do coeficiente de correlação em calibração.</p>

(continua)

Referência	Delineamento experimental	Análise de dados
EURACHEM (1998)	<ul style="list-style-type: none"> <li>- Preparar duas curvas.</li> <li>- Curva 1: branco mais materiais de referência ou amostras brancas adicionadas do analito em várias concentrações, sendo necessárias, pelo menos, seis concentrações diferentes mais o branco.</li> <li>- Curva 2: materiais de referência ou amostras brancas adicionadas do analito em, pelo menos, seis concentrações diferentes, na faixa linear.</li> </ul> <p>Observação: as diferentes concentrações devem ser preparadas independentemente e não de alíquotas de uma mesma solução mãe.</p>	<p>Curva 1:</p> <ul style="list-style-type: none"> <li>- Plotar as respostas medidas (eixo <i>y</i>) contra as concentrações do mensurando (eixo <i>x</i>).</li> <li>- Examinar visualmente para identificar a faixa linear aproximada e os limites superior e inferior da faixa de trabalho.</li> </ul> <p>Curva 2:</p> <ul style="list-style-type: none"> <li>- Plotar as respostas medidas (eixo <i>y</i>) contra as concentrações do mensurando (eixo <i>x</i>).</li> <li>- Examinar visualmente a existência de <i>outliers</i> (valores extremos) que possam não ser refletidos na regressão. Não é seguro remover <i>outliers</i> sem primeiro verificar por meio de determinações nas proximidades das concentrações.</li> <li>- Calcular coeficiente de regressão linear.</li> <li>- Calcular e plotar o gráfico dos valores dos resíduos da regressão. Distribuição aleatória em torno da linha reta confirma linearidade. Tendências sistemáticas indicam não linearidade.</li> <li>- Se a variância das replicatas for proporcional à concentração, usar regressão ponderada preferivelmente que regressão não ponderada.</li> <li>- Em certas circunstâncias pode ser melhor tentar ajustar os dados a uma curva não linear. Funções maiores que quadráticas não são recomendadas.</li> </ul>
HUBER (1998)	<ul style="list-style-type: none"> <li>- Realizar séries de três a seis injeções de cinco ou mais padrões de concentrações de 80 % a 120 % a faixa de concentração esperada.</li> </ul>	<ul style="list-style-type: none"> <li>- Determinar a equação de regressão linear. A interseção calculada pela equação de regressão não deve ser significativamente diferente de zero. Se um intercepto diferente de zero for obtido, deverá ser demonstrado que não há efeito na exatidão do método.</li> <li>- Frequentemente linearidade é avaliada graficamente em adição ou alternativamente a uma avaliação matemática. Fazer a avaliação por inspeção visual do gráfico do sinal altura ou área do pico em função da concentração do analito. Como desvios de linearidade são difíceis de serem detectados, dois gráficos adicionais podem ser usados: i) gráfico dos resíduos da regressão <i>versus</i> a concentração ou logaritmo da concentração. Os resíduos da regressão devem estar igualmente distribuídos entre valores positivos e negativos na faixa linear; ii) resposta relativa (sinal dividido pela respectiva concentração) no eixo <i>y</i> e a concentração correspondente no eixo <i>x</i> em uma escala logarítmica. A linha obtida deve ser horizontal na faixa linear.</li> </ul>
EC (2002)	<ul style="list-style-type: none"> <li>- Pelo menos cinco níveis, incluindo o zero, para construção das curvas de calibração.</li> </ul>	<ul style="list-style-type: none"> <li>- Descrever a faixa de trabalho da curva.</li> <li>- Descrever a fórmula matemática da curva e o grau de ajuste dos dados à curva.</li> <li>- Descrever os intervalos de aceitabilidade dos parâmetros da curva.</li> </ul>
THOMPSON, ELLISON & WOOD (2002)	<ul style="list-style-type: none"> <li>- Seis ou mais padrões de calibração.</li> <li>- Padrões de calibração devem ser igualmente espaçados ao longo da faixa de concentração de interesse.</li> <li>- A faixa deve cobrir de 0 % a 150 % ou de 50 % a 150 % a concentração esperada de ser encontrada, dependendo de qual destas for mais apropriada.</li> <li>- Padrões de calibração devem ser ensaiados, pelo menos em duplicata, preferencialmente triplicata ou mais, em ordem aleatória.</li> </ul> <p>Observações: o delineamento é importante. As replicatas têm que ser independentes ou fornecer uma estimativa do erro puro.</p>	<ul style="list-style-type: none"> <li>- Não utilizar coeficiente de correlação como indicador de linearidade.</li> <li>- Fazer ajuste exploratório por regressão linear simples.</li> <li>- Examinar o perfil do gráfico dos resíduos da regressão (perfis óbvios). Perfil curvado indica desvio da linearidade.</li> <li>- Avaliar não-linearidade por teste de significância comparando a variância do desvio da linearidade com aquela devida ao erro puro.</li> <li>- Heteroscedasticidade é bastante comum em calibração analítica. Um perfil sugerindo heteroscedasticidade significa que os dados da calibração são melhores tratados por regressão ponderada.</li> <li>- Se desvio da linearidade não for significativo, verificar se o intercepto é estatisticamente diferente de zero.</li> </ul>

(continua)



Referência	Delineamento experimental	Análise de dados
INMETRO (2003)	<ul style="list-style-type: none"> <li>- Curva 1: pelo menos sete replicatas de branco adicionado de concentrações variadas do analito ou amostras brancas adicionadas de concentrações variadas do analito.</li> <li>- Curva 2: pelo menos sete replicatas de materiais de referência (diferentes concentrações) na faixa linear ou amostras brancas adicionadas de concentrações variadas do analito na faixa linear.</li> </ul> <p>Observações: preparar diferentes concentrações de forma independente e não alíquotas da mesma solução mãe.</p>	<ul style="list-style-type: none"> <li>- Curva 1</li> <li>- Definir, por observação visual, a faixa linear aproximada e limites inferior e superior da faixa de trabalho.</li> <li>- Colocar no eixo x as concentrações do analito e no eixo y as respostas das medições.</li> <li>- Curva 2</li> <li>- Colocar no eixo x as concentrações do analito e no eixo y as respostas das medições.</li> <li>- Verificar visualmente a existência de <i>outliers</i> (valores extremos) que possam interferir na regressão (antes de remover os <i>outliers</i> fazer determinações nas proximidades das concentrações).</li> <li>- Calcular coeficiente de regressão linear. Valor de coeficiente de correlação linear maior que 0,90 é usualmente requerido.</li> <li>- Calcular e fazer o gráfico dos valores dos resíduos da regressão. Distribuição aleatória em torno da linha reta confirma linearidade. Tendências sistemáticas indicam não linearidade.</li> <li>- Como desvios de linearidade são muitas vezes difíceis de serem detectados visualmente, teste de t dos resíduos de <math>x^*</math> são utilizados para verificar sua adequação.</li> </ul> <p>Observação: procedimento baseado em EURACHEM (1998).</p>
SHABIR (2003)	<ul style="list-style-type: none"> <li>- Pelo menos cinco níveis de concentração, em uma faixa especificada.</li> <li>- Para ensaios a faixa é de 80 % a 120 % a concentração alvo. Para testes de impurezas a faixa é do limite de reportagem de cada impureza até 120 % da especificação. Na literatura é freqüentemente visto de 25 % a 200 % a contração nominal do analito.</li> </ul>	<p>Observação: procedimento baseado em GREEN (1996), exceto pela abordagem de que a qualidade do ajuste dos dados à curva de regressão pode ser avaliada por um procedimento baseado na soma de quadrados dos resíduos da regressão.</p>
TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004)	<ul style="list-style-type: none"> <li>- Faixa linear: analisar branco e, no mínimo, seis níveis de concentração independentes.</li> <li>- Linearidade: escolher materiais de referência certificados ou amostras adicionadas em, no mínimo, seis níveis de concentração dentro da faixa linear estimada.</li> </ul>	<ul style="list-style-type: none"> <li>- Faixa linear: ajustar a curva de calibração e estimar a faixa linear visualmente.</li> <li>- Linearidade: calcular os resíduos dos valores de y da curva de calibração. Plotar os resíduos da regressão em função da concentração. Se houver distribuição aleatória indica linearidade. Se houver tendências sistemáticas indica não linearidade. Calcular os sinais relativos (sinal/concentração). Plotar os sinais relativos em função da concentração. Se a linha for horizontal indica linearidade.</li> </ul> <p>Observação: esta referência discute que linearidade não pode ser expressa e deve ser demonstrada.</p>

\* O correto seria resíduos de y e não resíduos de x, uma vez o erro em x é considerado negligível.

**Tabela 16 - Recomendações para avaliação de efeitos de matriz na literatura**

Referência	Delineamento experimental	Análise de dados
NATA (1997)	<p>- Analisar amostras adicionadas, padrões e materiais de referência certificados. Pelo menos sete determinações de cada uma das três concentrações em cada tipo de matriz.</p> <p>- Analisar amostras nas quais concentrações conhecidas dos interferentes suspeitos tenham sido adicionadas. Estas amostras devem conter o analito na faixa de concentração esperada de ser encontrada na prática. Observação: esta referência trata seletividade e efeito de matriz no mesmo experimento.</p>	<p>-</p>
BRUCE, MINKKINEM & RIEKKOLA (1998)	<p>- Vários procedimentos e estatísticas correspondentes podem ser utilizados para estudar seletividade (ou testes para efeito de matriz) dependendo da disponibilidade de analito, de matriz sem analito e de amostras de referência adequadas nas concentrações de interesse.</p> <p>- Quando houver disponibilidade de matriz sem analito ou um conjunto adequado de amostras de referência, preparar dois grupos de amostras para ensaio, um com matriz e outro sem matriz, ambos com concentrações idênticas de analito. O número de amostras paralelas em cada nível de concentração deve ser grande o suficiente (<math>n \geq 6</math>) para facilitar a significância da comparação.</p> <p>- Quando não houver matriz sem analito, preparar dois grupos com adição de analito nas mesmas concentrações de adição sendo um com matriz e outro sem matriz. Observação: esta referência trata seletividade e efeito de matriz no mesmo experimento.</p>	<p>- Usar o teste de F para comparar as variâncias dos dois grupos (<math>s_1^2</math> e <math>s_2^2</math>), sendo <math>F = s_1^2 / s_2^2</math> ou <math>F = s_2^2 / s_1^2</math>, com a maior variância no denominador. Se F calculado não for significativo, a matriz não tem efeito sobre a precisão do método, neste nível de concentração. As variâncias dos dois grupos podem ser combinadas e as médias comparadas pelo teste de <math>t = \frac{ \bar{x}_1 - \bar{x}_2 }{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}</math>, onde <math>s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)}</math>, <math>\bar{x}_1</math> e <math>\bar{x}_2</math> são as médias das respostas dos analitos para amostras com e sem matriz no mesmo nível de concentração, <math>n_1</math> e <math>n_2</math> o número de amostras paralelas nos dois grupos. Se F calculado for significativo, as variâncias não podem ser consideradas iguais e o teste de t para variâncias diferentes deve ser adotado. Quando a faixa de concentração for relativamente estreita, ou se o erro analítico devido a concentração puder ser assumido como negligível, utilizar teste de t pareado <math>t = \frac{\bar{x}_d \sqrt{n}}{s_d}</math>, onde <math>\bar{x}_d</math> é a média das diferenças entre os pares de respostas, <math>s_d</math> o desvio padrão das diferenças, e <math>n</math> o número de pares. Se o método for utilizado em uma faixa ampla de concentração, o erro (sistemático e aleatório) pode ser dependente da concentração do analito. Neste caso, as diferenças entre os pares de amostras não seguem a distribuição normal e o teste de t não é aplicável, sendo utilizada a diferença relativa para análise dos dados.</p> <p>- Plotar os resultados dos grupos como função da concentração do analito adicionada no mesmo gráfico. Comparar as inclinações das curvas. Se as inclinações das duas curvas de regressão forem as mesmas, o único efeito de matriz presente é da interferência natural causada pelo nível básico do analito.</p>
EURACHEM (1998)	<p>- Método de adição.</p>	<p>- Interferentes podem afetar a inclinação da curva de calibração diferentemente do que seria com o analito de interesse. Este efeito tem o potencial de indicar a possível presença de interferentes desconhecidos, mas não é útil se a curva de recuperação for inerentemente não linear.</p>

(continua)

Referência	Delineamento experimental	Análise de dados
THOMPSON, ELLISON & WOOD (2002)	<ul style="list-style-type: none"> <li>- Aplicar o método de adição a uma solução teste derivada de uma amostra típica (matriz).</li> <li>- Fazer o teste de forma que forneça as mesmas diluições finais dos procedimentos normais e que as soluções adicionadas com matriz cubram a mesma faixa definida para as soluções usuais de calibração.</li> </ul>	<ul style="list-style-type: none"> <li>- Se a calibração for linear, comparar as inclinações da equação de calibração usual e do gráfico de adição de matriz. Se não houver diferença significativa, significa que não há efeito de matriz.</li> <li>- Se a calibração não for linear, métodos mais complexos são necessários para testes de significância, mas uma comparação visual com concentrações iguais usualmente basta.</li> </ul>
INMETRO (2003)	<p>Observações: esta referência trata seletividade e efeito de matriz no mesmo experimento. Procedimento baseado em BRUCE, MINKKINEM &amp; RIEKKOLA (1998), exceto pelo número de amostras paralelas em cada nível de <math>n \geq 7</math>.</p>	<p>Observação: procedimento baseado em BRUCE, MINKKINEM &amp; RIEKKOLA (1998), exceto pela informação da equação do teste de t para variâncias diferentes</p> $t = \frac{ \bar{x}_1 - \bar{x}_2 }{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ <p>e por não tratar os procedimentos utilizados para faixas amplas de concentração.</p>

**Tabela 17 - Recomendações para avaliação da seletividade na literatura**

Referência	Delineamento experimental	Análise de dados
GREEN (1996)	<ul style="list-style-type: none"> <li>- Analisar amostras contendo o analito e todos os componentes potenciais da amostra (placebo da formulação, intermediários de síntese, excipientes, produtos de degradação, impurezas de processo, etc.) e uma solução contendo somente o analito. Outros componentes potenciais são gerados expondo o analito a condições de estresse suficientes para degradá-lo a 80 % - 90 % de pureza.</li> </ul>	<ul style="list-style-type: none"> <li>- Comparar a resposta do analito em amostras contendo o analito e todos os componentes potenciais da amostra com a resposta de uma solução contendo somente o analito.</li> <li>- Avaliar o pico do analito em termos de pureza e resolução do pico eluído mais próximo. No caso de resolução aceitável para o analito e potenciais componentes da amostra, considerar os parâmetros cromatográficos ajustados como coluna, fase móvel, fluxo e modo de detecção.</li> </ul>
ICH (1996)	<ul style="list-style-type: none"> <li>- Identificação: analisar amostras contendo o analito e amostras que não contenham o analito. Testes de identificação podem ser aplicados também a materiais estruturalmente similares ou intimamente relacionados ao analito.</li> <li>- Ensaios e testes de impurezas: testar impurezas adicionando analito e outros compostos em níveis apropriados. Quando as impurezas não forem disponíveis, analisar amostras estocadas sob condições de estresse relevantes (luz, calor, umidade, hidrólise ácido-básica e oxidação) pelo método em validação e por um segundo método bem caracterizado.</li> </ul>	<ul style="list-style-type: none"> <li>- Identificação: confirmar a discriminação por resultados positivos para as amostras contendo o analito e resultados negativos para as amostras que não contenham o analito. Confirmar que respostas positivas não sejam obtidas para materiais estruturalmente similares ou intimamente relacionados ao analito.</li> <li>- Ensaios e testes de impurezas: demonstrar a separação das impurezas e de outros componentes da matriz. No caso de amostras analisadas por outro método, comparar os resultados obtidos pelos 2 métodos e o perfil das impurezas. Em procedimentos cromatográficos, utilizar cromatogramas representativos. Para separações críticas, avaliar a resolução de dois componentes que eluem próximos uns dos outros.</li> </ul>
AOAC (1998 b)	<ul style="list-style-type: none"> <li>- Analisar brancos de solventes e de campo para garantir que não há presença de interferentes.</li> <li>- Testar os resultados para o analito de interesse em diferentes condições experimentais (duas técnicas de detecção diferentes, dois princípios analíticos diferentes).</li> </ul>	<ul style="list-style-type: none"> <li>- O método deve ser capaz de distinguir o analito de materiais interferentes conhecidos.</li> <li>- O comportamento do analito durante as análises deve ser indistinguível dos padrões correspondentes em matrizes apropriadas.</li> </ul>
BRUCE, MINKKINEM & RIEKKOLA (1998)	<ul style="list-style-type: none"> <li>- Realizar pesquisa cuidadosa do conhecimento disponível na área de aplicação para encontrar os componentes que precisam ser testados.</li> <li>- Analisar analito, matriz com e sem analito, matérias-primas do processo, impurezas dos materiais iniciais ou do processo, subprodutos e produtos de degradação ou metabólitos e reagentes em branco.</li> <li>- Expor os componentes e matriz a condições extremas (calor, ácido, álcali, oxidação, radiação UV/visível, luz fluorescente) para determinar possíveis produtos de degradação.</li> <li>- Para cromatografia líquida e espectrofotometria UV/Visível, comparar os resultados variando condições de medição e análise para verificar que nenhum outro componente conhecido ou desconhecido esteja sendo determinado junto com o analito. Usar técnicas adicionais como cromatografia em camada delgada e espectrometria de massas após separação e coleta do analito.</li> <li>- Para técnicas cromatográficas, além das comparações visuais dos cromatogramas, calcular diferentes parâmetros como resolução, retenção relativa (fator de separação), fator de capacidade (fator de retenção), fator de simetria e número de pratos teóricos.</li> </ul>	-

(continua)

Referência	Delineamento experimental	Análise de dados
EURACHEM (1998)	<ul style="list-style-type: none"> <li>- Analisar amostras e materiais de referência pelo método em validação e outros métodos independentes.</li> <li>- Analisar amostras contendo vários interferentes suspeitos na presença do analito de interesse.</li> </ul>	<ul style="list-style-type: none"> <li>- Usar o resultado de técnicas confirmatórias para determinar a habilidade do método em confirmar a identidade do analito e medir o analito separadamente dos outros interferentes. Decidir que evidências são razoavelmente requeridas para suficiente confiabilidade.</li> <li>- Examinar o efeito dos interferentes pelo aumento ou inibição da detecção ou quantificação do mensurando. Se a detecção ou quantificação for inibida pelos interferentes será necessário novo desenvolvimento de método.</li> </ul>
HUBER (1998)	<ul style="list-style-type: none"> <li>- Em cromatografia, escolher colunas ideais e ajustar as condições cromatográficas, como composição de fase-móvel, temperatura da coluna, comprimento de onda do detector. Utilizar espectros durante a eluição de um cromatograma.</li> </ul>	-
EC (2002)	<ul style="list-style-type: none"> <li>- Selecionar substâncias potencialmente interferentes e amostras brancas relevantes para detectar a presença de possíveis interferentes e estimar o efeito das interferências.</li> <li>- Selecionar compostos quimicamente relacionados (metabólitos, derivados, etc.) e outras substâncias e que possam ser encontradas com o composto de interesse que possam estar presentes na amostra.</li> <li>- Analisar um número adequado de amostras brancas representativas (<math>n \geq 20</math>) e verificar a presença de interferências (sinais, picos, traços iônicos) na região de interesse onde o analito alvo é esperado eluir.</li> <li>- Analisar amostras adicionadas em uma concentração relevante com substâncias que interferem na identificação ou quantificação do analito.</li> </ul>	<ul style="list-style-type: none"> <li>- Avaliar se a presença de interferentes conduz a uma falsa identificação, se a identificação do analito alvo é impedida pela presença de um ou mais interferentes e se a quantificação é influenciada.</li> </ul>
THOMPSON, ELLISON & WOOD (2002)	<ul style="list-style-type: none"> <li>- Executar o procedimento com uma matriz branca e a mesma amostra branca adicionada com potencial interferente em uma concentração apropriada.</li> </ul>	<ul style="list-style-type: none"> <li>- Seletividade é essencialmente uma determinação qualitativa baseada em testes de interferências.</li> <li>- Uma medida quantitativa é o índice de seletividade, que é a razão entre a sensibilidade do método (inclinação da curva de calibração) e a inclinação da resposta produzida independentemente por um potencial interferente. Se não houver amostra branca disponível e um material típico for utilizado em substituição, a estimativa da inclinação da resposta pode ser feita de um experimento simples somente sob a premissa de que efeitos de matriz mútuos são ausentes.</li> </ul>
INMETRO (2003)	Observação: procedimento baseado em BRUCE, MINKKINEM & RIEKKOLA (1998).	-
SHABIR (2003)	Observação: procedimento baseado em GREEN (1996).	Observação: procedimento baseado em GREEN (1996).
TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004)	<ul style="list-style-type: none"> <li>- Adicionar possíveis interferentes em diferentes amostras. Utilizar um segundo método se houver interferentes disponíveis.</li> </ul>	<ul style="list-style-type: none"> <li>- Calcular a porcentagem de recuperação.</li> <li>Observação: esta referência discute que seletividade não pode ser expressa e deve ser demonstrada, dependendo do tipo e propósito do método.</li> </ul>

**Tabela 18 - Recomendações para avaliação da exatidão na literatura**

Referência	Delineamento experimental	Análise de dados
GREEN (1996)	<ul style="list-style-type: none"> <li>- Analisar uma amostra de concentração conhecida.</li> <li>- Analisar a amostra pelo novo método e por um método existente conhecido como exato.</li> <li>- Analisar amostras brancas adicionadas de analito, em três níveis, em triplicata, na faixa de 50 % a 150 % a concentração alvo do analito. Se não for possível preparar amostras brancas, usar o método de adição.</li> </ul>	<ul style="list-style-type: none"> <li>- Comparar o valor medido com o valor verdadeiro.</li> <li>- Comparar os resultados dos ensaios com o novo método e com os obtidos para o método existente.</li> <li>- Calcular a porcentagem de recuperação para amostras adicionadas, sendo um critério que a recuperação média seja de 100 % + 2 % em cada nível de concentração na faixa de 80 % a 120 % a concentração alvo.</li> </ul>
ICH (1996)	<ul style="list-style-type: none"> <li>- Vários métodos para determinar exatidão: aplicação de um procedimento analítico a um analito de pureza conhecida ou com a substância a ser analisada adicionada em quantidades conhecidas, comparação dos resultados obtidos pelo método em validação com os obtidos por um segundo método bem caracterizado.</li> <li>- Realizar um mínimo de nove determinações de um mínimo de três níveis de concentração, cobrindo a faixa especificada (por exemplo: três concentrações e três determinações de cada concentração).</li> </ul>	<ul style="list-style-type: none"> <li>- Reportar a exatidão como a porcentagem de recuperação da quantidade adicionada ou como a diferença entre a média e o valor verdadeiro aceito com os respectivos intervalos de confiança.</li> </ul>
NATA (1997)	<ul style="list-style-type: none"> <li>- Analisar amostras adicionadas. Pelo menos sete determinações de cada uma das três concentrações em cada tipo de matriz. Os níveis de concentração podem ser, por exemplo, próximos ao limite de detecção, próximos da concentração máxima permitida e próximos a metade da faixa de uso do método.</li> <li>- Se houver material de referência certificado disponível, realizar pelo menos sete análises replicatas de cada material.</li> <li>- Analisar padrão ou material de referência para o qual a concentração do analito seja conhecida com elevada exatidão e precisão.</li> </ul>	<ul style="list-style-type: none"> <li>- Estimar a recuperação. Uma limitação deste procedimento é que o analito adicionado não se encontra na mesma forma do analito naturalmente presente na amostra, podendo fornecer recuperações otimistas.</li> <li>- Estatísticas para avaliação de tendências do método incluem o teste de médias (teste de t), comparação de dois métodos por ajuste de mínimos quadrados, comparação da precisão de diferentes métodos (teste de F) e teste de qui-quadrado para comparar resultados experimentais com resultados esperados ou teóricos.</li> <li>- A diferença entre o valor conhecido como verdadeiro e a média de determinações replicatas com o método é devida a tendência do método (exatidão) e erros aleatórios (precisão).</li> </ul>
AOAC (1998 b)	<ul style="list-style-type: none"> <li>- Adota uma das técnicas a seguir, nesta ordem de preferência: uso de material de referência certificado, uso de método de referência com incerteza conhecida, uso de recuperação de amostras adicionadas. Observações: para amostras adicionadas, as concentrações devem cobrir a faixa de interesse, incluindo a concentração alvo, o limite inferior e superior. A quantidade adicionada deve ser uma fração substancial, maior que a quantidade presente na amostra não fortificada.</li> </ul>	<ul style="list-style-type: none"> <li>- Calcular a porcentagem de recuperação</li> </ul> $R\% = \frac{CF - CU}{CA} \times 100$ <p>onde <i>CF</i> é a quantidade medida do analito na amostra fortificada, <i>CU</i> a quantidade de analito medida na amostra não fortificada, e <i>CA</i> a quantidade de analito adicionada na amostra fortificada.</p>

(continua)

Referência	Delineamento experimental	Análise de dados
BRUCE, MINKKINEM & RIEKKOLA (1998)	<ul style="list-style-type: none"> <li>- Comparar resultados do método com resultados obtidos para a mesma amostra com um método de referência, pela análise de material de referência certificado, materiais de referência padrão, ou testes de recuperação. Selecionar um grupo de amostra que representa a provável faixa de concentração do analito. Analisar pelo método em questão e por um outro método independente.</li> <li>- Preparar amostras de ensaio em níveis relevantes de concentração do analito e processar as amostras pelo método.</li> </ul>	<ul style="list-style-type: none"> <li>- Comparar os resultados obtidos para os dois métodos por regressão linear*.</li> <li>- Calcular a porcentagem de recuperação das amostras preparadas <math>R\% = \frac{C_m - C_{nf}}{C_f} \times 100</math>, onde <math>C_m</math> é a quantidade medida do analito na amostra fortificada, <math>C_{nf}</math> a quantidade de analito medida na amostra não fortificada, e <math>C_f</math> a quantidade de analito adicionada a amostra fortificada. A porcentagem de recuperação aceitável depende da concentração e de requisitos do nível de qualidade requerido pelos clientes. Em análises de alimentos FDA recomenda faixas de 80 % a 110 % para concentrações de analito acima de 0,1 mg/kg e de 60 % a 100 % para concentrações de analito abaixo de 0,1 mg/kg, enquanto a AOAC <i>International</i> indica 70 % a 120 % para analitos em níveis de partes por bilhão.</li> </ul>
EURACHEM (1998)	<ul style="list-style-type: none"> <li>- Analisar branco de reagente e material de referência pelo método em validação, dez vezes. Sujeito a incerteza de o branco ser um branco verdadeiro e à caracterização do material de referência.</li> <li>- Analisar branco de reagente e material de referência/teste pelo método em validação e por um método independente (de preferência primário), dez vezes.</li> <li>- Analisar amostras brancas ou não fortificadas e amostras fortificadas com o analito de interesse em uma faixa de concentração, seis vezes.</li> </ul>	<ul style="list-style-type: none"> <li>- Subtrair o valor médio do branco do valor médio do material de referência. Comparar o valor medido com o valor verdadeiro aceito para o material de referência. Neste caso a tendência do método é avaliada.</li> <li>- Subtrair o valor médio do branco do valor médio do material de referência/teste. Comparar os resultados do método com medidas similares feitas usando o método independente/primário. Neste caso a tendência do método relativa ao método independente/primário é avaliada.</li> <li>- Determinar a recuperação do analito nas várias concentrações <math>R\% = \frac{C_1 - C_2}{C_3} \times 100</math>, onde <math>C_1</math> é a concentração determinada na amostra fortificada, <math>C_2</math> a concentração determinada na amostra não fortificada, e <math>C_3</math> a concentração da fortificação.</li> </ul> <p>Observação: recuperações de amostra fortificadas são usualmente melhores que as obtidas nas amostras reais nas quais o analito está mais fortemente ligado.</p>
HUBER (1998)	<ul style="list-style-type: none"> <li>- Utilizar método de referência, assumindo que a incerteza do método de referência seja conhecida.</li> <li>- Analisar amostras de concentrações conhecidas (material de referência certificado).</li> <li>- Analisar amostras brancas adicionadas, se material de referência certificado não for disponível. Os níveis de concentração devem cobrir a faixa de interesse e incluir uma concentração próxima ao limite de quantificação.</li> </ul>	<ul style="list-style-type: none"> <li>- Comparar os resultados do método com resultados de um método de referência estabelecido.</li> <li>- Comparar o valor medido com o valor verdadeiro fornecido com o material.</li> <li>- Determinar a recuperação das amostras adicionadas comparando a resposta do extrato com a resposta do analito dissolvido em solvente puro.</li> </ul> <p>Observação: esta referência apresenta uma tabela da AOAC - PVMC (1993) com uma estimativa de dados de recuperação em função da concentração do analito.</p>

(continua)

Referência	Delineamento experimental	Análise de dados
EC (2002)	<p>- Exatidão somente pode ser estabelecida por meio de material de referência certificado. O procedimento é descrito em detalhes na ISO 5725-4. Um exemplo dado: analisar seis replicatas de material de referência certificado de acordo com as instruções do método e determinar a concentração do analito presente em cada replicata.</p> <p>- Quando nenhum material de referência certificado estiver disponível, avaliar a recuperação com amostras fortificadas. Analisar dezoito alíquotas de amostras brancas em três grupos de seis alíquotas com 1; 1,5 e 2 vezes o limite mínimo de desempenho para analitos sem limites máximos estabelecidos ou 0,5; 1 e 1,5 vezes o limite máximo permitido. Calcular a concentração presente em cada amostra.</p>	<p>- Calcular média, desvio padrão e coeficiente de variação para a concentração. Calcular a exatidão dividindo a concentração média detectada pelo valor do certificado e multiplicando por 100, para expressar o resultado em porcentagem.</p> <p>- Calcular a porcentagem de recuperação para cada amostra como o valor medido multiplicado por 100 e dividido pelo nível de fortificação. Quando não se dispõe de amostras brancas deve-se considerar a diferença entre o valor obtido para a amostra fortificada e o valor obtido para a amostra sem fortificação. Calcular a média de porcentagem de recuperação e coeficiente de variação para os seis resultados de cada nível.</p> <p>Observações: esta referência apresenta uma tabela com valores aceitáveis para porcentagem de recuperação média em função da concentração do analito.</p>
THOMPSON, ELLISON & WOOD (2002)	<p>- Os valores de referência para experimentos de exatidão podem ser obtidos com materiais de referência certificados, materiais de referência, métodos de referência ou uso de adição/recuperação.</p> <p>- Material de referência certificado: utilizado para estabelecer todos os aspectos de tendência (do método, do laboratório e do ensaio), simultaneamente, assumindo que não haja efeitos de variações de matriz. É importante garantir que os valores de incerteza certificados sejam suficientemente pequenos para permitir a detecção de uma tendência de magnitude importante. Onde for apropriado e necessário, um número adequado de materiais de referência certificados, considerando matrizes e concentrações de analitos, deve ser examinado.</p> <p>- Material de referência: utilizado quando material de referência certificado não for disponível ou como adicional aos estudos de materiais de referência certificados. Materiais de referência incluem materiais caracterizados por um produtor de material de referência, mas cujos valores não são acompanhados da incerteza ou são qualificados de outra maneira, materiais caracterizados por um fabricante de material, materiais caracterizados no laboratório para uso como material de referência ou distribuídos em ensaios de proficiência. Apesar da rastreabilidade destes materiais ser questionável, é bem melhor usá-los do que não realizar estudos de tendência.</p> <p>- Método de referência: opção útil para verificação de tendências de métodos alternativos ou modificações de um método padrão já validado e em uso no laboratório. Analisar um número de amostra típico, preferencialmente cobrindo uma faixa de concentração bastante uniforme, pelos dois métodos.</p> <p>- Adição/recuperação: utilizada para avaliar tendências na ausência de materiais de referência ou para dar suporte a estudos com materiais de referência. Analisar uma amostra típica pelo método em validação em seu estado original e após adição de quantidades conhecidas do analito a uma porção teste. Calcular a concentração presente em cada amostra.</p>	<p>- Material de referência certificado: experimentos de exatidão típicos geram uma resposta média para um material de referência. Na interpretação dos resultados, a incerteza associada ao valor certificado deve ser levada em conta com a incerteza originada da variação estatística no laboratório. Onde o valor de incerteza certificado for pequeno, o teste t de Student é normalmente conduzido, utilizando o termo de precisão apropriado. Quando um número de materiais for examinado e as incertezas dos valores certificados forem menores que as dos resultados analíticos, é razoavelmente seguro usar regressão linear simples para avaliar os resultados. Neste caso a tendência pode ser expressa em função da concentração, aparecendo como um intercepto diferente de zero.</p> <p>- Material de referência: utilizado da mesma forma que os materiais de referência certificados, embora sem incerteza declarada qualquer teste de significância confia completamente na precisão dos resultados observados.</p> <p>- Método de referência: comparar os resultados ao longo da faixa estudada por teste estatístico apropriado (por exemplo: teste de t pareado, com verificação das premissas de homogeneidade de variâncias e normalidade). Verificar qualquer tendência entre os métodos.</p> <p>- Adição/recuperação: a diferença entre os resultados da amostra no estado original e adicionada do analito corresponde a uma proporção de massa adicionada. Estudos de recuperação somente permitem avaliar a tendência devida a efeitos do analito adicionado. Os mesmos efeitos não são necessariamente aplicáveis na mesma extensão à amostras naturais. Enquanto boas recuperações não são garantias de exatidão, pobres recuperações são certamente indicativas de falta de exatidão.</p>

(continua)



Referência	Delineamento experimental	Análise de dados
INMETRO (2003)	<ul style="list-style-type: none"> <li>- Material de referência certificado: realizar uma série de análises em replicata.</li> <li>- Comparações interlaboratoriais: ensaios de proficiência e processos colaborativos.</li> <li>- Comparação de métodos: comparar os resultados obtidos utilizando um método interno com os resultados conseguidos através de um método de referência. Realizar análises em replicata, utilizando os dois métodos, em separado, sobre as mesmas amostras, numa faixa restrita de concentrações ou em toda faixa de concentração em que se pretende validar o método.</li> <li>- Amostras adicionadas: preparar amostras adicionadas de quantidades conhecidas do analito, em pelo menos três níveis de concentração.</li> </ul>	<p>- Material de referência certificado: quando o valor obtido não estiver dentro do intervalo de incerteza indicado para o valor certificado, o laboratório deve procurar as causas desse desvio e tentar eliminá-las ou aceitá-las, dependendo do rigor definido para os resultados. Calcular a média e desvio padrão dos valores obtidos pelo laboratório e comparar com os valores certificados do material de referência, usando: erro relativo</p> $ER = \frac{(x_{lab} - x_v)}{x_v} \times 100$ ; teste de hipótese (como teste de F de Snedecor para comparação das variâncias ou teste t de Student para comparação das médias); índice z-escore $z = \frac{(x_{lab} - x_v)}{s}$ ; e erro normalizado $En = \frac{(x_{lab} - x_v)}{\sqrt{U_{lab}^2 + U_{ref}^2}}$ ; <p>onde <math>x_{lab}</math> é o valor obtido pelo laboratório, <math>x_v</math> o valor aceito como verdadeiro, <math>s</math> a unidade de desvio ou incerteza do material de referência certificado, <math>U_{lab}^2</math> a incerteza do resultado, e <math>U_{ref}^2</math> a incerteza associada ao valor verdadeiro. Para índice z-escore adotar os critérios de aceitabilidade do ISO Guia 43. Considerar valores de <math>z</math>, em módulo, menores ou iguais a 2 como satisfatórios, entre 2 e 3 como questionáveis, e maiores que 3 como insatisfatórios. Para erro normalizado considerar adequados valores, em módulo, menores ou iguais a 1.</p> <ul style="list-style-type: none"> <li>- Comparação de métodos: usar técnicas para comparar os resultados como testes de hipóteses, teste de regressão linear* entre os dois métodos e projeto de experimentos.</li> </ul> <p>Observações: procedimentos para cálculo das porcentagens de recuperação baseados em EURACHEM (1998).</p>
SHABIR (2003)	<p>Observação: procedimento baseado em GREEN (1996).</p> <ul style="list-style-type: none"> <li>- Analisar material de referência certificado. Quando não houver material de referência certificado adicionar uma amostra branca e uma matriz típica com quantidades conhecidas do analito.</li> <li>- Mínimo de três níveis de concentrações na faixa do método de ensaio (no limite de detecção, no meio e alto).</li> <li>- Mínimo de três repetições por concentração.</li> </ul>	<p>Observação: procedimento baseado em GREEN (1996).</p> <ul style="list-style-type: none"> <li>- Para material de referência certificado: calcular a tendência (erro ou diferença entre o valor medido e o valor verdadeiro) ou o z-escore</li> </ul> $z = \frac{X_{encontrado} - X_{certificado}}{\sqrt{\frac{SD_{encontrado}}{n_{encontrado}} + \left(\frac{SD_{certificado}}{n_{certificado}}\right)^2}}$ ou $z = \frac{X_{encontrado} - X_{certificado}}{\sqrt{\frac{SD_{encontrado}}{n_{encontrado}} + \left(\frac{CI}{2}\right)^2}}$ , onde $SD$ é o desvio padrão, $n$ o número de observações, e $CI$ o intervalo de confiança, sendo requeridos valores de z-escore menores ou iguais ao módulo de 2. <ul style="list-style-type: none"> <li>- Para amostras adicionadas calcular a porcentagem de recuperação como <math>R\% = 100 \times [spike_{matriz} / spike_{branca}]^{**}</math>, sendo requeridos valores de porcentagens de recuperação dependo do nível de concentração do analito.</li> </ul>
TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004)		

\* O correto seria testes de correlação, método dos mínimos quadrados ortogonais ou método de máxima verossimilhança, uma vez que na comparação entre dois métodos há erros nas duas variáveis.

\*\* A fórmula apresentada para a porcentagem de recuperação não considera a razão entre o valor medido para a amostra adicionada, subtraído do valor medido para a amostra branca, e o valor teórico adicionado.

**Tabela 19 - Recomendações para avaliação da precisão na literatura**

Referência	Delineamento experimental	Análise de dados
GREEN (1996)	<ul style="list-style-type: none"> <li>- Repetitividade: fazer um mínimo de dez injeções de uma solução (precisão do equipamento ou repetitividade de injeção) e analisar, em um laboratório e no mesmo dia, alíquotas de uma amostra homogênea, preparadas independentemente, de acordo com o método (repetitividade ou precisão intra-ensaio).</li> <li>- Reprodutibilidade: testar amostras homogêneas em múltiplos laboratórios como parte de estudos colaborativos.</li> </ul>	<ul style="list-style-type: none"> <li>- Calcular o desvio padrão relativo. A precisão do equipamento deve ser maior ou igual a 1 % e a precisão intra-ensaio maior ou igual a 2 %. Para um método de impurezas a precisão do equipamento deve ser menor ou igual a 5 % e a precisão intra-ensaio menor ou igual a 10 %.</li> <li>- A avaliação dos resultados de reprodutibilidade freqüentemente está focada na medida da tendência dos resultados que nas diferenças de precisão somente. Equivalência estatística é freqüentemente usada como uma medida de aceitabilidade para resultados interlaboratoriais. Como alternativa, um procedimento mais prático é o uso de equivalência analítica, na qual uma faixa de resultados aceitáveis é escolhida antes do estudo e usada para julgar os resultados obtidos de diferentes laboratórios. Os resultados obtidos nos laboratórios devem ser estatisticamente equivalentes ou a média dos resultados deve estar dentro de 2 % o valor obtido no laboratório de testes primários. Para método de impurezas os resultados obtidos nos laboratórios devem ser estatisticamente equivalentes ou a média dos resultados deve estar dentro de 10 % o valor obtido no laboratório de testes primários</li> </ul>
ICH (1996)	<ul style="list-style-type: none"> <li>- Mínimo de nove determinações cobrindo a faixa especificada pelo procedimento (por exemplo: três níveis e três repetições de cada nível) ou por um mínimo de seis determinações a 100 % da concentração alvo.</li> <li>- Condições de repetitividade, reprodutibilidade parcial e reprodutibilidade.</li> </ul>	<ul style="list-style-type: none"> <li>- Reportar desvio padrão, desvio padrão relativo (coeficiente de variação) e intervalo de confiança para cada tipo de precisão estudada.</li> </ul>
NATA (1997)	<ul style="list-style-type: none"> <li>- Repetitividade: analisar replicatas de análises de amostras, sendo pelo menos sete determinações de cada tipo de matriz.</li> <li>- Reprodutibilidade: analisar replicatas de amostras por diferentes operadores, em diferentes laboratórios, usando diferentes equipamentos (usualmente por meio de estudos interlaboratoriais).</li> <li>- Analisar padrão ou material de referência para o qual a concentração do analito seja conhecida com elevada exatidão e precisão.</li> </ul>	<ul style="list-style-type: none"> <li>- A diferença entre o valor conhecido como verdadeiro e a média das determinações replicatas é devida a tendência do método (exatidão) e erros aleatórios (precisão).</li> </ul>

(continua)

Referência	Delineamento experimental	Análise de dados
BRUCE, MINKKINEM & RIEKKOLA (1998)	<ul style="list-style-type: none"> <li>- Realizar determinações duplicatas de amostras de rotina escolhidas aleatoriamente.</li> <li>- Quando possível, amostrar de três a dez subamostras representativas de um mesmo lote e analisá-las de forma similar, no mesmo momento, com o mesmo equipamento e pelo mesmo analista.</li> <li>- Repetitividade: analisar a mesma amostra, pelo mesmo procedimento, pelo mesmo equipamento, em curto intervalo de tempo e pelo mesmo analista.</li> <li>- Reprodutibilidade parcial: analisar a mesma amostra, pelo mesmo laboratório, mas por diferentes analistas, possivelmente por diferentes equipamentos e em tempos diferentes.</li> <li>- Reprodutibilidade: analisar a mesma amostra, com o mesmo procedimento, por diferentes equipamentos, em tempos diferentes, por diferentes analistas e em diferentes laboratórios.</li> </ul>	<ul style="list-style-type: none"> <li>- Estabelecer o desvio padrão relativo como <math>s_r = \sqrt{\frac{\sum_i d_{ri}^2}{2n}}</math>, onde <math>d_{ri} = (d_i/\bar{x}_i) \times 100\%</math> são as diferenças relativas das medições em duplicatas, e <math>n</math> o número de duplicatas.</li> <li>- A precisão neste caso é dada como desvio padrão <math>s</math>, desvio padrão relativo <math>s_r</math>, ou intervalo de confiança da média <math>CI</math>. Calcular o intervalo de confiança pela equação <math>CI = \frac{t \times s}{\sqrt{n}}</math>, onde <math>n</math> é o número de resultados paralelos, e <math>t</math> um coeficiente tomado da distribuição <math>t</math> de Student com um nível de confiança de 95%.</li> <li>- Estimar a precisão nas condições de repetitividade, reprodutibilidade parcial e reprodutibilidade por análise de variância, em um dado nível de confiança, geralmente 95%, de acordo com o padrão internacional ISO 5725. Calcular o desvio padrão de repetitividade como <math>\sqrt{s_r^2} = \sqrt{\frac{\sum_{i=1}^p s_i^2}{p}}</math> e a repetitividade <math>r = 2,83\sqrt{s_r^2}</math>.</li> <li>- Calcular o desvio padrão de reprodutibilidade parcial como <math>\sqrt{s_r^2 + s_L^2} = \sqrt{\left[ \frac{p \sum_{i=1}^p X_i^2 - \left( \sum_{i=1}^p X_i \right)^2}{p(p-1)} - \frac{s_r^2}{n} \right] + s_r^2}</math> e a reprodutibilidade parcial <math>R = 2,83\sqrt{s_r^2 + s_L^2}</math>, onde <math>p</math> é o número de dias ou ensaios, <math>s_i</math> o desvio padrão dos resultados de um dia, <math>X_i</math> a média dos resultados de um dia, e <math>n</math> o número de medições paralelas.</li> <li>- Valores obtidos pela equação de Horwitz podem ser utilizados como referência para os desvios padrão relativos estimados.</li> </ul>
	<ul style="list-style-type: none"> <li>- Analisar padrões, materiais de referência ou amostras brancas fortificadas em várias concentrações, ao longo da faixa de trabalho.</li> <li>- Repetitividade: mesmo analista, equipamento, laboratório, em curto intervalo de tempo, dez repetições independentes.</li> <li>- Reprodutibilidade parcial: diferentes analistas, equipamentos, mesmo laboratório, em intervalo de tempo estendido, dez repetições independentes.</li> <li>- Reprodutibilidade: diferentes analistas, equipamentos, laboratórios, em intervalo de tempo estendido, dez repetições independentes. Requer estudo colaborativo.</li> </ul>	<ul style="list-style-type: none"> <li>- Para métodos qualitativos: expressar a precisão como porcentagem de resultados falsos positivos e porcentagem de resultados falsos negativos.</li> <li>- Para métodos quantitativos: determinar o desvio padrão para cada concentração, em cada condição.</li> </ul>

(continua)

Referência	Delineamento experimental	Análise de dados
HUBER (1998)	<p>- Repetitividade: analisar em um laboratório, por um mesmo operador, usando o mesmo equipamento e em curto intervalo de tempo, pelo menos, cinco ou seis determinações, de três diferentes matrizes, em dois ou três níveis de concentração diferentes.</p> <p>- Reprodutibilidade: analisar alíquotas de lotes homogêneos, em diferentes laboratórios, com diferentes analistas, usando diferentes equipamentos e condições operacionais, que podem diferir mas ainda estar dentro dos parâmetros especificados do método (testes interlaboratoriais).</p> <p>Observação: esta referência cita variações típicas que afetam a reprodutibilidade de um método.</p>	<p>- Calcular o desvio padrão relativo.</p> <p>Observações: esta referência cita que os critérios de aceitabilidade dependem muito do tipo de análise e apresenta uma tabela da AOAC - PVMC (1993) com uma estimativa de dados de desvio padrão relativo (intra ou inter-dias) em função da concentração do analito.</p>
EC (2002)	<p>- Repetitividade: preparar amostras fortificadas, dezoito alíquotas de amostras brancas em três grupos de seis alíquotas com 1; 1,5 e 2 vezes o limite mínimo de desempenho para analitos sem limites máximos estabelecidos ou 0,5; 1 e 1,5 vezes o limite máximo permitido. Analisar as amostras. Calcular a concentração detectada em cada amostra. Repetir pelo menos mais duas vezes.</p> <p>- Reprodutibilidade parcial: analisar amostras fortificadas, dezoito alíquotas de amostras brancas em três grupos de seis alíquotas com 1; 1,5 e 2 vezes o limite mínimo de desempenho para analitos sem limites máximos estabelecidos ou 0,5; 1 e 1,5 vezes o limite máximo permitido. Analisar as amostras. Calcular a concentração detectada em cada amostra. Repetir pelo menos mais duas vezes com operadores diferentes ou condições ambientais diferentes, por exemplo.</p> <p>- Reprodutibilidade: o laboratório deve participar de estudos colaborativos de acordo com a norma ISO 5725-2.</p>	<p>- Repetitividade: calcular a concentração média, desvio padrão e coeficiente de variação para as amostras fortificadas. Calcular a média de concentrações global e coeficiente de variação para amostras fortificadas.</p> <p>- Reprodutibilidade intralaboratorial: calcular a concentração média, desvio padrão e coeficiente de variação para as amostras fortificadas.</p> <p>Observações: esta referência apresenta uma tabela com valores aceitáveis de desvio padrão sob condições de reprodutibilidade em função da concentração do analito.</p>
THOMPSON, ELLISON & WOOD (2002)	<p>- Precisão sob condições de repetitividade ou intra-ensaio e entre ensaios. As variações nas condições entre ensaios devem representar o que normalmente acontece no laboratório sob o uso de rotina do método. Por exemplo: variações em lotes de reagentes, analistas e equipamentos devem ser representativas.</p> <p>- Usualmente estas duas fontes de erro são obtidas em resultados analíticos individuais, os quais têm uma precisão combinada dita total. Estimativas destas duas precisões podem ser obtidas de forma simplificada por análise de amostras selecionadas em duplicata em um número sucessivo de ensaios. Cada duplicata deve ser uma execução independente do procedimento aplicada as diferentes porções de ensaio separadas.</p>	<p>- Calcular a variância dos componentes separados pela aplicação de análise de variância fator único. Teste de F é apropriado para dados com distribuição normal dos resíduos.</p>

(continua)

Referência	Delineamento experimental	Análise de dados
INMETRO (2003)	<p>- Repetitividade: pode ser determinada por meio de análise de padrões, material de referência ou adição de branco em várias concentrações na faixa de trabalho. Analisar sete ou mais repetições.</p> <p>- Reprodutibilidade parcial: efetuar <math>n</math> medições em replicata, ou em ensaio único, sobre a amostra, nas condições pré-definidas, pois existem vários métodos de estudar este tipo de dispersão. Quando aplicável, repetir este procedimento sobre outras amostras, abrangendo outros níveis de concentração.</p> <p>- Reprodutibilidade: não é um componente de validação de método executado por um único laboratório.</p>	<p>- Repetitividade: calcular o desvio padrão para cada concentração. Calcular o limite de repetitividade, para um nível de confiança de 95 %, como 2,8 vezes o desvio padrão de repetitividade. Avaliar se há diferença significativa entre duplicatas de uma amostra.</p> <p>- Precisão intermediária: calcular como</p> $S_i = \sqrt{\frac{1}{t(n-1)} \sum_{j=1}^t \sum_{k=1}^n (y_{jk} - \bar{y}_j)^2}$ <p>onde <math>t</math> é o total de amostras ensaiadas, <math>n</math> é o total de ensaios efetuados por amostra, <math>j</math> o número da amostra <math>j = 1, t</math>, <math>k</math> o número do ensaio da amostra <math>j</math> <math>k = 1, n</math>, <math>y_{jk}</math> o valor do resultado <math>k</math> para a amostra <math>j</math>, e <math>\bar{y}_j</math> a média dos resultados da amostra <math>j</math>. Observação: esta referência cita a ISO 5725-3 e a ISO 5725-6.</p> <p>- Reprodutibilidade: calcular o desvio padrão para cada nível, separadamente, após eliminação dos <i>outliers</i> (valores extremos), conforme ISO 5725-2 e ASTM E178. A partir do desvio padrão obtido sob condições de reprodutibilidade calcular o limite de reprodutibilidade, para um nível de confiança de 95 %, como 2,8 vezes o desvio padrão de reprodutibilidade. Avaliar se há diferença significativa entre duplicatas de amostras.</p>
SHABIR (2003)	<p>Observação: procedimento baseado em GREEN (1996) e ICH (1996).</p> <p>- Analisar amostras em condições de repetitividade, reprodutibilidade parcial e reprodutibilidade.</p> <p>- Mínimo, de três níveis de concentrações na faixa do método de ensaio (no limite de detecção, no meio e alto).</p> <p>- Mínimo de três repetições por concentração.</p>	<p>Observação: procedimento baseado em GREEN (1996).</p> <p>- Repetitividade: calcular desvio padrão <math>SD_r</math>, desvio padrão relativo <math>RSD_r = 100 \times SD_r / \bar{X}</math>, limite de repetitividade <math>r = 2,83SD_r</math>, e intervalo de confiança <math>CI_r = \bar{X} \pm C</math>, onde <math>\bar{X}</math> é a média dos resultados, e</p> $C = \frac{s \times t}{\sqrt{n}}$ <p>com <math>t</math> o valor t de Student.</p> <p>- Reprodutibilidade parcial: calcular desvio padrão <math>SD_{Rp}</math>, desvio padrão relativo <math>RSD_{Rp} = 100 \times SD_{Rp} / \bar{X}</math>, limite de reprodutibilidade parcial <math>R_p = 2,83SD_{Rp}</math>, e intervalo de confiança <math>CI_{Rp} = \bar{X} \pm C</math>, onde <math>\bar{X}</math> é a média dos resultados, e <math>C = \frac{s \times t}{\sqrt{n}}</math> com <math>t</math> o valor t de Student. Considerar aceitáveis desvios padrão relativos de 0,5 a 0,6 vezes os valores teóricos determinados pela equação de Horwitz.</p> <p>- Reprodutibilidade: calcular desvio padrão <math>SD_R</math>, desvio padrão relativo <math>RSD_R = 100 \times SD_R / \bar{X}</math>, limite de reprodutibilidade <math>R = 2,83SD_R</math>, e intervalo de confiança <math>CI_R = \bar{X} \pm C</math>, onde <math>\bar{X}</math> é a média dos resultados, e <math>C = \frac{s \times t}{\sqrt{n}}</math> com <math>t</math> o valor t de Student. Considerar aceitáveis desvios padrão relativos de 0,5 a 2 vezes os valores teóricos determinados pela equação de Horwitz.</p>
TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004)		

**Tabela 20** - Recomendações para avaliação do limite de detecção na literatura

Referência	Delineamento experimental	Análise de dados
GREEN (1996)	- Determinar somente para ensaios de impurezas. Observar picos cromatográficos próximos do limite de detecção.	- Relação sinal/ruído de 3.
ICH (1996)	- Avaliação visual: analisar amostras com concentrações conhecidas do analito. Aplicável em métodos instrumentais e não instrumentais. - Relação sinal/ruído: analisar um número apropriado de amostras com concentrações baixas e conhecidas do analito e amostras brancas. Aplicável em procedimentos analíticos que apresentam ruído de linha de base. - Desvio padrão e inclinação: preparar uma curva de calibração do analito na faixa do limite de detecção e um número apropriado de amostras brancas.	- Avaliação visual: estabelecer o nível mínimo no qual o analito pode ser confiavelmente detectado. Apresentar cromatogramas relevantes. - Relação sinal/ruído: comparar os sinais medidos para amostras com concentrações baixas e conhecidas do analito com os sinais obtidos para amostras brancas e estabelecer a concentração mínima na qual o analito pode ser confiavelmente detectado. Uma relação sinal/ruído entre 3 ou 2:1 é geralmente considerada aceitável para estimativa do limite de detecção. - Desvio padrão e inclinação: calcular o limite de detecção $DL = \frac{3,3\sigma}{S}$ , onde $\sigma$ é o desvio padrão da resposta, e $S$ é a inclinação da curva de calibração. Calcular $\sigma$ como o desvio padrão das respostas obtidas para as amostras brancas ou como o desvio padrão dos resíduos da regressão.
NATA (1997)	- Analisar amostras brancas e adicionadas em níveis baixos. Pelo menos sete determinações em cada tipo de matriz.	- Plotar o parâmetro detectado contra a concentração de analito adicionada. - Determinar a imprecisão (desvio padrão relativo ou coeficiente de variação) para cada concentração a partir do limite de detecção até uma faixa de concentração ser estabelecida na qual a imprecisão permaneça aceitável. Por este método o menor nível no qual o analito pode ser confiavelmente detectado é estabelecido. - Uma relação sinal/ruído de 3:1 para métodos instrumentais é geralmente aceita. - Determinar o desvio padrão do <i>background</i> do ruído obtido na análise de amostras brancas e multiplicar por um fator 3.
AOAC (1998 b)	- Realizar leituras de amostras brancas ( $\geq 20$ ).	- Calcular o valor médio das leituras de brancos da matriz mais três desvios padrão da média. - Expressar em concentração do analito. Observação: esta referência cita um limite intermediário chamado limite de determinação ou de decisão correspondente a média do conteúdo medido para amostras brancas ( $\geq 20$ ) mais seis vezes o desvio padrão da média.
BRUCE, MINKKINEM & RIEKKOLA (1998)	- Analisar uma amostra branca e soluções de analito com pequenas quantidades de analito.	- Calcular a concentração do limite de detecção como $C_L = \frac{k\sigma_B}{X_A} \times C_0$ , onde $X_A$ é o sinal obtido para o analito, $k$ uma constante 2 ou 3, $\sigma_B$ o desvio padrão do <i>background</i> , e $C_0$ a concentração do analito para o sinal $X_A$ . - Frequentemente uma relação sinal/ruído de 2 ou 3 é requerida para medidas cromatográficas ou espectrofotométricas.

(continua)

Referência	Delineamento experimental	Análise de dados
EURACHEM (1998)	<p>Para métodos qualitativos:</p> <ul style="list-style-type: none"> <li>- Analisar amostras brancas adicionadas do analito em uma faixa de concentração.</li> <li>- Medir dez replicatas independentes em cada nível de concentração. As medidas devem ser aleatórias.</li> </ul> <p>Para métodos quantitativos:</p> <ul style="list-style-type: none"> <li>- Analisar dez amostras brancas independentes. Este procedimento somente é útil onde a amostra branca fornecer desvio padrão diferente de zero. Obter uma amostra branca verdadeira pode ser difícil.</li> <li>- Analisar dez amostras brancas independentes, fortificadas na menor concentração aceitável (concentração mais baixa para a qual um grau aceitável de incerteza pode ser alcançado).</li> </ul>	<p>Para métodos qualitativos:</p> <ul style="list-style-type: none"> <li>- Construir uma curva de porcentagem de resultados positivos (ou negativos) <i>versus</i> a concentração.</li> <li>- Determinar, por inspeção, o limiar de concentração no qual o método se torna pouco confiável.</li> </ul> <p>Para métodos quantitativos:</p> <ul style="list-style-type: none"> <li>- Calcular o desvio padrão <math>s</math> dos valores das amostras brancas. Expressar o limite de detecção como a concentração do analito correspondente ao valor médio do branco <math>+3s</math>.</li> <li>- Calcular o desvio padrão <math>s</math> dos valores das amostras fortificadas. Expressar o limite de detecção como a concentração do analito correspondente a <math>0 + 3s</math> ou o valor correspondente a amostra branca <math>+4,65s</math> (derivado do teste de hipóteses).</li> </ul>
HUBER (1998)	-	<ul style="list-style-type: none"> <li>- Em cromatografia, definir a quantidade de analito injetada que resulta em um pico com altura de, pelo menos, duas a três vezes maior que a linha de base do ruído.</li> </ul>
EC (2002)	<p>Observação: esta referência não descreve limite de detecção e sim limite de decisão.</p> <p>Para analitos proibidos:</p> <ul style="list-style-type: none"> <li>- Preparar uma curva de calibração de acordo com a ISO 11843. Usar amostras brancas fortificadas no nível mínimo de performance requerido e acima deste nível, em espaços eqüidistantes. Analisar as amostras. Este procedimento somente é aplicável a ensaios quantitativos.</li> <li>- Analisar pelo menos vinte amostras brancas, por matriz. Este procedimento é aplicável a ensaios quantitativos e qualitativos.</li> </ul> <p>Para analitos com limites permitidos estabelecidos:</p> <ul style="list-style-type: none"> <li>- Preparar uma curva de calibração de acordo com a ISO 11843. Usar amostras brancas fortificadas ao redor do limite permitido, em espaços eqüidistantes. Analisar as amostras.</li> <li>- Analisar pelo menos vinte amostras fortificadas com o analito no limite permitido, por matriz.</li> </ul>	<p>Para analitos proibidos:</p> <ul style="list-style-type: none"> <li>- Plotar o sinal contra a concentração adicionada. Calcular o limite de decisão como a concentração correspondente ao intercepto mais 2,33 vezes o desvio padrão de reprodutibilidade parcial do intercepto (<math>\alpha = 1\%</math>).</li> <li>- Calcular a razão sinal/ruído no intervalo de tempo na qual o analito é esperado. O limite de decisão pode ser considerado como três vezes a razão sinal/ruído.</li> </ul> <p>Para analitos com limites permitidos estabelecidos:</p> <ul style="list-style-type: none"> <li>- Plotar o sinal contra a concentração adicionada. Calcular o limite de decisão como a concentração correspondente ao limite permitido mais 1,64 vezes o desvio padrão de reprodutibilidade parcial (<math>\alpha = 5\%</math>).</li> <li>- O limite de decisão pode ser considerado como a concentração no limite permitido mais 1,64 vezes seu desvio padrão (<math>\alpha = 5\%</math>).</li> </ul>
THOMPSON, ELLISON & WOOD (2002)	<ul style="list-style-type: none"> <li>- Realizar pelo menos seis determinações independentes e completas da concentração do analito em uma amostra branca típica ou uma amostra em um nível baixo de concentração, sem restrições a resultados zero ou negativos.</li> </ul>	<ul style="list-style-type: none"> <li>- Calcular o limite de detecção aproximado como <math>3s_0</math>, onde <math>s_0</math> é a precisão estimada para as determinações.</li> </ul>
INMETRO (2003)	<ul style="list-style-type: none"> <li>- Analisar sete ou mais amostras brancas. Este limite somente é válido quando os valores dos brancos apresentarem um desvio padrão diferente de zero.</li> <li>- Analisar sete ou mais amostras brancas com adição da menor concentração aceitável do analito (concentração mais baixa para a qual um grau aceitável de incerteza pode ser alcançado).</li> </ul> <p>Observação: procedimento baseado em EURACHEM (1998), exceto pelo número de replicatas.</p>	<ul style="list-style-type: none"> <li>- Calcular a média <math>\bar{X}</math> e o desvio padrão <math>s</math> dos valores dos brancos. Calcular o limite de detecção <math>LD = \bar{X} + ts</math>, onde <math>t</math> é o valor da abscissa <math>t</math> de Student para <math>(1 - \alpha) \times 100\%</math> nível de confiança e <math>(n-1)</math> graus de liberdade.</li> <li>- Calcular o desvio padrão (<math>s</math>) dos valores dos brancos com adição. Calcular o limite de detecção <math>LD = 0 + ts</math>, onde <math>t</math> é o valor da abscissa <math>t</math> de Student para <math>(1 - \alpha) \times 100\%</math> nível de confiança e <math>(n - 1)</math> graus de liberdade.</li> </ul> <p>Observação: procedimento baseado em EURACHEM (1998).</p>

(continua)

Referência	Delineamento experimental	Análise de dados
SHABIR (2003)	<ul style="list-style-type: none"> <li>- Técnica instrumental.</li> <li>- Técnica visual e não instrumental: como cromatografia em camada delgada.</li> <li>- Técnica baseada no desvio padrão da resposta <math>SD</math> e na inclinação da curva de calibração <math>S</math>.</li> </ul>	<ul style="list-style-type: none"> <li>- Técnica instrumental: concentração com relação sinal/ruído especificada, geralmente 3:1.</li> <li>- Técnica visual e não instrumental: -</li> <li>- Técnica baseada no desvio padrão da resposta <math>SD</math> e na inclinação da curva de calibração <math>S</math>: o limite de detecção é <math>LOD = 3,3(SD / S)</math>. <math>SD</math> pode ser obtido de amostras brancas, dos resíduos da regressão e dos interceptos da curva de regressão.</li> </ul>
TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004)	<ul style="list-style-type: none"> <li>- Medir um mínimo de dez amostras brancas independentes ou dez amostras brancas adicionadas na menor concentração aceitável.</li> <li>- Escolher um número de níveis de concentrações baixos próximos ao limite de detecção esperado.</li> </ul>	<ul style="list-style-type: none"> <li>- Calcular a resposta medida <math>x_L = x_{bi} + 3s_{bi}</math> para ensaios com amostras brancas ou <math>x_L = 3s_x</math> para amostras adicionadas, onde <math>x_{bi}</math> é a média das medidas do branco, <math>s_{bi}</math> o desvio padrão das medidas do branco, e <math>s_x</math> o desvio padrão das medidas das amostras adicionadas na menor concentração aceitável.</li> <li>- Ajustar a curva de calibração e a inclinação <math>A</math>. Calcular o limite de detecção <math>LOD = x_L / A = [x_{bi} + 3s_{bi}]</math>.</li> <li>Observação: o limite de detecção é expresso como a concentração ou quantidade de analito que é derivada do sinal medido.</li> </ul>



**Tabela 21 - Recomendações para avaliação do limite de quantificação na literatura**

Referência	Delineamento experimental	Análise de dados
GREEN (1996)	<ul style="list-style-type: none"> <li>- Reduzir a concentração do analito até atingir um nível onde a precisão do método for inaceitável. Este limite é requerido somente para ensaios de impurezas.</li> </ul>	<ul style="list-style-type: none"> <li>- Desvio padrão relativo menor ou igual a 20 %. Se não for determinado experimentalmente, o limite de quantificação é freqüentemente calculado como a concentração do analito que fornece uma relação sinal/ruído de 10.</li> </ul>
ICH (1996)	<ul style="list-style-type: none"> <li>- Avaliação visual: analisar amostras com concentrações conhecidas do analito. Aplicável em métodos instrumentais e não instrumentais.</li> <li>- Relação sinal/ruído: analisar um número apropriado de amostras com concentrações baixas e conhecidas do analito e amostras brancas. Aplicável em procedimentos analíticos que apresentam ruído de linha de base.</li> <li>- Desvio padrão e inclinação: preparar uma curva de calibração do analito na faixa do limite de quantificação e um número apropriado de amostras brancas.</li> </ul>	<ul style="list-style-type: none"> <li>- Avaliação visual: estabelecer o nível mínimo no qual o analito pode ser quantificado com precisão e exatidão aceitável. Apresentar cromatogramas relevantes.</li> <li>- Relação sinal/ruído: comparar os sinais medidos para amostras com concentrações baixas e conhecidas do analito com os sinais obtidos para amostras brancas e estabelecer a concentração mínima na qual o analito pode ser confiavelmente quantificado. Uma relação sinal/ruído de 10:1 é geralmente considerada aceitável para estimativa do limite de quantificação.</li> <li>- Desvio padrão e inclinação: calcular o limite de detecção</li> </ul> $QL = \frac{10\sigma}{S}$ <p>onde <math>\sigma</math> é o desvio padrão da resposta, e <math>S</math> a inclinação da curva de calibração. Calcular <math>\sigma</math> como o desvio padrão das respostas obtidas para as amostras brancas ou como o desvio padrão dos resíduos da regressão.</p>
NATA (1997)	<ul style="list-style-type: none"> <li>- Analisar amostras brancas e adicionadas em níveis baixos. Pelo menos sete determinações em cada tipo de matriz.</li> </ul>	<ul style="list-style-type: none"> <li>- Calcular o limite de reportagem como sendo dez vezes o limite de detecção.</li> </ul>
AOAC (1998 b)	<ul style="list-style-type: none"> <li>- Realizar leituras de amostras brancas.</li> </ul>	<ul style="list-style-type: none"> <li>- Calcular o valor médio da leitura do branco da matriz mais dez desvios padrão da média.</li> <li>- Expressar em concentração do analito.</li> </ul>
BRUCE, MINKKINEM & RIEKKOLA (1998)	<ul style="list-style-type: none"> <li>- Analisar amostras brancas.</li> <li>- Analisar amostras adicionadas no limite de detecção ou acima deste.</li> </ul>	<ul style="list-style-type: none"> <li>- Selecionar um nível de concentração que produza um sinal dez vezes o desvio padrão relativo do ruído ou do ruído das amostras brancas.</li> <li>- Mediar a precisão no limite de detecção ou acima deste. Selecionar a concentração com desvio padrão relativo abaixo de 10 % a 20 %. O desvio padrão relativo das amostras neste nível deve ser próximo ao estabelecido pela equação de Horwitz.</li> </ul>
EURACHEM (1998)	<ul style="list-style-type: none"> <li>- Analisar dez amostras brancas independentes. Obter uma amostra branca verdadeira pode ser difícil.</li> <li>- Fortificar alíquotas de amostras brancas em vários níveis de concentração próximos ao limite de detecção. Medir dez replicatas independentes em cada nível de concentração.</li> </ul>	<ul style="list-style-type: none"> <li>- Calcular o desvio padrão <math>s</math> dos valores das amostras brancas. Expressar o limite de quantificação como a concentração do analito correspondente ao valor médio do branco <math>+5s</math>, <math>+6s</math> ou <math>+10s</math>.</li> <li>- Calcular o desvio padrão <math>s</math> do valor do analito em cada concentração. Plotar <math>s</math> contra a concentração e determinar o valor do limite de quantificação por inspeção. Expressar o limite de quantificação como a menor concentração do analito a qual pode ser determinada com um nível de incerteza aceitável. Normalmente o limite de quantificação faz parte do estudo de determinação da faixa de trabalho e não deve ser determinado por extrapolação abaixo da menor concentração de branco fortificado.</li> </ul>
HUBER (1998)	<ul style="list-style-type: none"> <li>- Definir a precisão esperada para o limite de quantificação.</li> <li>- Preparar amostras adicionadas diluídas em quantidades decrescentes de analito.</li> <li>- Injetar seis vezes.</li> </ul>	<ul style="list-style-type: none"> <li>- Calcular o desvio padrão relativo para cada concentração.</li> <li>- Plotar o gráfico da precisão <i>versus</i> a concentração.</li> <li>- Definir como limite de quantificação a concentração correspondente a precisão esperada.</li> </ul>

(continua)

Referência	Delineamento experimental	Análise de dados
EC (2002)	<p>Observação: esta referência não descreve limite de quantificação e sim capacidade de detecção.</p> <p>Para analitos proibidos:</p> <ul style="list-style-type: none"> <li>- Preparar uma curva de calibração de acordo com a ISO 11843. Usar amostras brancas fortificadas no nível mínimo de performance requerido e abaixo deste nível, em espaços eqüidistantes. Analisar as amostras. Este procedimento somente é aplicável a ensaios quantitativos.</li> <li>- Analisar pelo menos vinte amostras brancas fortificadas com o analito no limite de decisão, por matriz. Este procedimento é aplicável a ensaios quantitativos e qualitativos.</li> </ul> <p>Para analitos com limites permitidos estabelecidos:</p> <ul style="list-style-type: none"> <li>- Preparar uma curva de calibração de acordo com a ISO 11843. Usar amostras brancas fortificadas ao redor do limite permitido, em espaços eqüidistantes. Analisar as amostras.</li> <li>- Analisar pelo menos vinte amostras fortificadas com o analito no limite de decisão, por matriz.</li> </ul>	<p>Para analitos proibidos:</p> <ul style="list-style-type: none"> <li>- Plotar o sinal contra a concentração adicionada. Calcular a capacidade de detecção como a concentração correspondente ao limite de decisão mais 1,64 vezes o desvio padrão de reprodutibilidade parcial do conteúdo médio medido no limite de decisão (<math>\beta = 5\%</math>).</li> <li>- A capacidade de detecção pode ser considerada como o valor do limite de decisão mais 1,64 vezes seu desvio padrão (<math>\beta = 5\%</math>).</li> </ul> <p>Para analitos com limites permitidos estabelecidos:</p> <ul style="list-style-type: none"> <li>- Plotar o sinal contra a concentração adicionada. Calcular o limite de decisão como a concentração correspondente ao limite de decisão mais 1,64 vezes o desvio padrão de reprodutibilidade parcial (<math>\beta = 5\%</math>).</li> <li>- A capacidade de detecção pode ser considerada como o valor do limite de decisão mais 1,64 vezes seu desvio padrão (<math>\beta = 5\%</math>).</li> </ul>
INMETRO (2003)	<ul style="list-style-type: none"> <li>- Analisar sete ou mais amostras brancas. Obter uma amostra branca verdadeira pode ser difícil.</li> <li>- Analisar sete ou mais amostras brancas com adição de concentrações variadas do analito, próximas ao limite de detecção.</li> </ul> <p>Observação: procedimento baseado em EURACHEM (1998), exceto pelo número de replicatas.</p>	<ul style="list-style-type: none"> <li>- Calcular a média <math>\bar{X}</math> e o desvio padrão <math>s</math> dos valores dos brancos. Calcular o limite de quantificação <math>LQ = \bar{X} + 5s</math>, <math>LQ = \bar{X} + 6s</math> ou <math>LQ = \bar{X} + 10s</math>.</li> <li>- Calcular o desvio padrão <math>s</math> do valor do analito para cada concentração. Fazer o gráfico <math>s</math> versus concentração e atribuir um valor para o <math>LQ</math> por inspeção. Expressar o limite de quantificação como a concentração mais baixa do analito que pode ser determinada com um nível aceitável de incerteza.</li> </ul> <p>Observação: procedimento baseado em EURACHEM (1998).</p>
SHABIR (2003)	<ul style="list-style-type: none"> <li>- Técnica instrumental.</li> <li>- Técnica visual e não instrumental: como cromatografia em camada delgada.</li> <li>- Técnica baseada no desvio padrão da resposta <math>SD</math> e na inclinação da curva de calibração <math>S</math>.</li> </ul>	<ul style="list-style-type: none"> <li>- Técnica instrumental: concentração com relação sinal/ruído especificada, geralmente 10:1.</li> <li>- Técnica visual e não instrumental: -.</li> <li>- Técnica baseada no desvio padrão da resposta <math>SD</math> e na inclinação da curva de calibração <math>S</math>: o limite de detecção é <math>LOQ = 10(SD / S)</math>. <math>SD</math> pode ser obtido de amostras brancas, dos resíduos da regressão e dos interceptos da curva de regressão.</li> </ul>
TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004)	<ul style="list-style-type: none"> <li>- Medir um mínimo de dez amostras brancas independentes.</li> <li>- Escolher um número de níveis de concentrações baixas próximas ao limite de detecção calculado.</li> </ul>	<ul style="list-style-type: none"> <li>- Calcular a resposta medida <math>x_L = x_{bi} + 10s_{bi}</math> para ensaios com amostras brancas, onde <math>x_{bi}</math> é a média das medidas do branco, e <math>s_{bi}</math> o desvio padrão das medidas do branco.</li> <li>- Ajustar a curva de calibração e a inclinação <math>A</math>. Calcular o limite de detecção <math>LOD = x_L / A = [x_{bi} + 10s_{bi}]</math>.</li> </ul> <p>Observação: o limite de quantificação é expresso como a concentração ou quantidade de analito que é derivada do sinal medido.</p>

## Linearidade e faixa

Calibração é um procedimento que determina a diferença sistemática existente entre um sistema de medição e um sistema de referência representado por um material de referência e seus valores aceitos (ISO, 1996). Considerando que a maioria dos métodos de ensaio em análises de alimentos utiliza relações lineares, a determinação da equação de calibração é extremamente importante tanto em ensaios de rotina quanto em procedimentos de validação de métodos (RSC, 2005 a). Mais ainda, uma adequada avaliação da linearidade é fundamental para a qualidade dos demais parâmetros de desempenho estabelecidos para métodos quantitativos (como sensibilidade, efeitos de matriz, seletividade, exatidão, precisão e limites). Entretanto, exceto quando ocorrem grandes erros no preparo das curvas, os erros devidos à calibração correspondem aos menores componentes na compilação da incerteza total (THOMPSON, ELLISON & WOOD, 2002).

Os experimentos descritos para avaliação da linearidade freqüentemente envolvem preparo de curvas com ou sem matriz, em sua maioria com cinco a seis níveis de concentração, incluindo ou não o ponto zero, e com um mínimo de duas a sete replicatas por nível. O método dos mínimos quadrados ordinários (MMQO) é consenso como ferramenta estatística para avaliação da linearidade. A maior parte das referências recomenda a estimativa dos parâmetros e dos resíduos (erros) da regressão, além da inspeção visual do gráfico  $x$ - $y$  e do gráfico dos resíduos da regressão (GREEN, 1996; ICH, 1996; ISO, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; DANZER & KURRIE, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Contudo, os procedimentos para análise de dados propostos na literatura são muitas vezes controversos. Enquanto alguns autores recomendam a avaliação do coeficiente de correlação como indicador de linearidade (GREEN, 1996; ICH, 1996; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; INMETRO, 2003), outros enfatizam que esta estatística não deve ser utilizada neste contexto (DANZER & KURRIE, 1998; THOMPSON, ELLISON & WOOD, 2002; HORWITZ, 2003). Determinadas referências destacam ainda as premissas relativas ao MMQO, bem como testes estatísticos aplicáveis a algumas delas (ISO, 1996; DANZER & KURRIE, 1998; EURACHEM, 1998; THOMPSON, ELLISON & WOOD, 2002).

As faixas de trabalho e linear podem ser diferentes para diferentes matrizes, de acordo com o efeito de interferentes provenientes de cada matriz. A avaliação destas faixas é essencial para o entendimento da aplicabilidade de um método de ensaio

(INMETRO, 2003) e também útil para o planejamento do grau de calibração que será requerido na aplicação do método em atividades de rotina (EURACHEM, 1998; INMETRO, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). A determinação da faixa de trabalho é normalmente derivada dos estudos de linearidade e depende da intenção de aplicação do procedimento (ICH, 1996).

### **Efeitos de matriz e seletividade**

Uma amostra consiste do analito a ser medido, da matriz e de outros componentes que podem ter algum efeito na medição, mas que não se quer detectar ou quantificar (BRUCE, MINKKINEN & RIEKKOLA, 1998; INMETRO, 2003). A seletividade é um parâmetro de desempenho relacionado ao evento de detecção do analito na presença destes componentes. Este parâmetro possui grande importância, principalmente quando se considera a área de análises de alimentos, na qual as matrizes são misturas de constituições químicas complexas.

Dependendo de como a identidade do analito é estabelecida, interferentes podem inibir a detecção distorcendo o sinal proveniente do analito. Diferentemente, os interferentes também podem contribuir para o sinal atribuído ao analito, com efeito de aumentar a concentração do mesmo (EURACHEM, 1998). A medição pode ser alterada porque matriz, reagentes ou outros componentes afetam a sensibilidade do detector que mede o analito ou porque estes compostos afetam diretamente a resposta. Assim, efeitos de erros constantes devidos a interferentes e erros proporcionais à calibração devidos a efeitos de matriz podem estar presentes ao mesmo tempo. Uma vez detectados, estes problemas devem ser resolvidos pelo uso de curvas de matriz (ou método de adição), análises de multi-compostos, ou por alterações em etapas analíticas de preparo e tratamento de amostras, separação ou detecção (BRUCE, MINKKINEN & RIEKKOLA, 1998; INMETRO, 2003).

Apesar de extremamente relevantes, nem todos os protocolos que tratam validação intralaboratorial de métodos discorrem sobre experimentos para avaliação dos efeitos de matriz. Os estudos de efeitos de matriz são delineados com o preparo de curvas de calibração de solventes e curvas de adição de analitos em amostras brancas, sendo que o método de adição é adotado quando não é possível obter estas amostras (INMAN, 1987; NATA, 1997; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003).

Experimentos para avaliação da seletividade envolvem a distinção entre amostras brancas e adicionadas do analito, além da avaliação da capacidade de identificação do

analito de interesse na presença de interferentes, deliberadamente adicionados à amostra (INMAN, 1987; GREEN, 1996; ICH, 1996; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). No entanto, quando não há disponibilidade de interferentes, a habilidade de medir o analito é comparada com outros métodos e técnicas independentes (INMAN, 1987; AOAC, 1998 b; EURACHEM, 1998; HUBER, 1998; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004).

Alguns protocolos confundem as técnicas para avaliação de seletividade com repetitividade. Enquanto na primeira são feitas várias medições por várias técnicas, preferivelmente independentes e que adotem princípios diferentes, a segunda requer várias medições pela mesma técnica (EURACHEM, 1998).

### **Exatidão e precisão**

Exatidão e precisão são consideradas os parâmetros de validação mais importantes (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). A precisão diz respeito à distribuição dos erros aleatórios de um sistema de medição, enquanto a exatidão expressa os erros sistemáticos (ISO, 1993 b; RSC, 2003). Estes parâmetros definem a acurácia de um método e são associados à estimativa da incerteza da medição (NATA, 1997).

Para avaliação da exatidão e precisão, as referências freqüentemente sugerem experimentos com materiais de concentração conhecida do analito, em replicatas, sob condições predeterminadas (GREEN, 1996; ICH, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Os materiais utilizados são divididos em materiais certificados (materiais de referência certificados) e não certificados (materiais de referência, amostras brancas e amostras adicionadas), estes últimos também conhecidos como amostras de controle de qualidade (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004) e utilizados para estudos de recuperação aparente.

Os termos adição (*spike* ou *surrogate*) ou fortificação (*fortification*) de amostras para ensaios de recuperação aparente são utilizados de forma equivalente. Entretanto, alguns autores sugerem que a fortificação seja restrita à adição do analito de interesse enquanto a adição pode ser relacionada ao analito ou a outros compostos conhecidos como interferentes (EURACHEM, 1998; EC, 2002).

Ainda com relação a ensaios de recuperação aparente existe grande discussão na literatura sobre a correção ou não dos resultados obtidos pelo valor da recuperação aparente obtida. O NATA (1997), por exemplo, considera que seja preferível reportar os resultados não corrigidos acompanhados da recuperação aparente obtida. Alguns argumentos contra a correção dos resultados pela recuperação aparente são: i) as recuperações aparentes estimadas pelo uso de amostras adicionadas podem ser maiores que as recuperações aparentes verdadeiras obtidas com amostras nativas; ii) os fatores de recuperação aparente podem ser suspeitos pois este parâmetro varia com o tipo de matriz e concentração do analito; e iii) algumas legislações que impõem limites máximos de contaminantes são estruturadas no entendimento de que resultados não corrigidos serão usados para sanções. A favor do uso da correção é considerado que: i) tendências não corrigidas devidas a recuperações aparentes baixas significam que os resultados não serão universalmente comparáveis, transferíveis e, portanto, não apropriados a reconhecimentos mútuos; ii) correções pela recuperação aparente são matematicamente isomórficas de técnicas aceitas como o uso de padrão interno, não podendo ser suspeitas em princípio; e iii) as incertezas associadas os fatores de correção de recuperação aparente podem ser estimadas e incorporadas à incerteza combinada do resultado corrigido (RSC, 2005 b). A chave para resolução desta discussão encontra-se na recomendação publicada pela IUPAC (THOMPSON et al., 1999) que sugere que os resultados sejam corrigidos pela recuperação aparente a menos que exista alguma razão específica para não fazê-lo como, por exemplo, por exigências contratuais, regulamentares ou uso de métodos empíricos.

Outro experimento comumente sugerido para avaliação da exatidão é a comparação do método com um segundo método de referência bem estabelecido (GREEN, 1996; ICH, 1996; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004).

## **Limites**

A habilidade em detectar um composto presente na amostra, mesmo em concentrações extremamente baixas, é relacionada aos limites de detecção, quantificação, decisão e capacidade de detecção (GREEN, 1996; EC, 2002; INMETRO, 2003). Embora estes parâmetros sejam essenciais para aplicação em pesquisas, comércio internacional, saúde e segurança (CURRIE, 1995), ainda existe uma grande variedade de terminologias e fórmulas (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE,

2004). Numerosas definições, procedimentos, aplicações, interpretações e critérios de decisão têm sido propostos (GREEN, 1996; ICH, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004).

Sobre o limite de quantificação, THOMPSON, ELLISON & WOOD (2002) sugerem que, enquanto operar acima de um limite fornece alguma segurança, devemos reconhecer que há uma dicotomia artificial na escala de concentração, uma vez que medidas abaixo deste limite não são privadas de informações de conteúdo e podem ser adequadas para o propósito de uso. Estes autores não recomendam o uso deste tipo de limite, sendo preferível tentar expressar a incerteza da medição em função da concentração, comparando com os critérios acordados entre o laboratório e o cliente ou o usuário final dos resultados.

Uma observação importante é que não é verdade que a detecção ou quantificação sejam impossíveis abaixo dos limites. Acontece que, nestes níveis, a incerteza da medição pode ser maior que o próprio valor da medição (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Considerando que os resultados analíticos não são somente concentrações, mas sim estimativas de concentração propensas a erros, nas medições abaixo dos limites estas estimativas acompanhadas por suas respectivas incertezas podem levar a resultados negativos (RSC, 2001 a). Limites não são diretamente relevantes na estimativa de incerteza, porém a incerteza abaixo do limite de detecção e próxima ao limite de quantificação pode exigir cuidados especiais (EURACHEM/CITAC, 2000).

TAVERNIERS, DE LOOSE & VAN BOCKSTAELE (2004) advertem também que os parâmetros limite de detecção e sensibilidade não devem ser confundidos.

# **CAPÍTULO I - PROCEDIMENTO PARA VALIDAÇÃO INTRALABORATORIAL DE MÉTODOS DE ENSAIO QUANTITATIVOS E QUALITATIVOS EM ANÁLISES DE ALIMENTOS**

## **RESUMO**

Um procedimento detalhado para validação intralaboratorial de métodos em análises de alimentos foi proposto. Para métodos quantitativos, o delineamento experimental incluiu curvas do analito em solventes e em matriz, amostras brancas e adicionadas. Na determinação da linearidade e efeitos de matriz, inspeção visual dos dados e de gráficos e tratamento de valores extremos (teste de Jacknife) foram sugeridos. Os dados foram analisados pelo método dos mínimos quadrados ordinários, com verificação das premissas de normalidade (teste de Ryan-Joiner), homoscedasticidade (teste de Brown-Forsythe), independência (teste de Durbin-Watson) dos resíduos e ajuste ao modelo (análise de variância). As inclinações e interseções das curvas do analito em solventes e em matriz foram comparadas (teste de t), com verificação prévia da homogeneidade das variâncias (teste de F). Após inspeção das recuperações aparentes e tratamento de valores extremos (teste de Grubbs), as mesmas foram comparadas com valores críticos para verificação da falta de exatidão. Desvios padrão de repetitividade e reprodutibilidade parcial foram estimados (análise de variância), por nível de concentração, após confirmação da normalidade (teste de Ryan-Joiner) e homoscedasticidade (teste de Brown-Forsythe) dos resíduos. Limites foram estabelecidos usando os resultados de recuperação aparente e ruídos obtidos para amostras brancas. Para métodos qualitativos, ensaios com padrões, amostras brancas e adicionadas foram recomendados. Seletividade foi definida pela ausência de resultados negativos para amostras adicionadas acima do limite de detecção do método e ausência de resultados positivos para amostras brancas. Limite de detecção foi estabelecido a partir do ponto em que o método apresentou resultados não detectados para amostras adicionadas. Os delineamentos experimentais e a seleção das ferramentas para análises de dados foram baseados em fundamentos teóricos, assim como aspectos de eficiência, praticidade e facilidade. Os princípios de representatividade e verificação de premissas foram os alicerces deste processo.

Palavras-chave: validação intralaboratorial; método quantitativo; método qualitativo; análises de alimentos, quimiometria.



## ABSTRACT

A PROCEDURE FOR SINGLE-LABORATORY VALIDATION OF QUANTITATIVE AND QUALITATIVE METHODS IN FOOD ANALYSIS. A detailed procedure for single-laboratory validation in food analysis was proposed. For quantitative methods, the experimental design included solvent and matrix matched curves, blank samples and spiked samples. For linearity and matrix effects assessment, visual inspection of the data and plots and outliers treatment (Jackknife test) were suggested. Data were fitted by ordinary least squares method, with the check of the assumptions of normality (Ryan-Joiner test), homoscedasticity (Brown-Forsythe test), independence (Durbin-Watson test) of residuals and adjustment to the model (analysis of variance). The slopes and intercepts of the solvent and matrix matched curves were compared (t test), with a previous verification of the homogeneity of the variances (F test). After visual inspection of the recoveries and outliers treatment (Grubbs test), they were compared with critical values to verify lack of trueness. Standard deviation of repeatability and within-reproducibility were estimated (analysis of variance), for each concentration level, after the confirmation of residuals normality (Ryan-Joiner test) and homoscedasticity (Brown-Forsythe test). Limits were assessed using the recovery results and the noise obtained for blank samples. For qualitative methods, assays with reference standards and materials, blank samples and spiked samples were recommended. Selectivity was defined by the absence of negative results for spiked samples in levels above the detection limit and positive results for blank samples. Limit of detection was assessed. Limit of detection was established based on the point in which the method showed non detected results for spiked samples. The experimental design and the choice of statistical tools for analysis of the data were based on theoretical background, efficiency, practicability and easiness, considering representativeness and testing of assumptions.

Key words: single-laboratory validation; quantitative method; qualitative method; food analysis, chemometry.

## 1 INTRODUÇÃO

A área de análises de alimentos é norteadada por uma crescente demanda de métodos de ensaio. Os constantes avanços tecnológicos na produção, processamento,

embalagem e armazenamento de alimentos, advindos principalmente das mudanças de comportamento e hábitos alimentares da sociedade contemporânea, trazem ao cotidiano das análises de alimentos questões como análises de aditivos, migrantes de embalagens, contaminantes, resíduos de pesticidas e drogas veterinárias, adulterantes, indicadores de origem geográfica e organismos geneticamente modificados.

Além dos desafios advindos da diversidade e contínua evolução deste setor, os laboratórios de análises de alimentos necessitam demonstrar competência técnica para garantia da qualidade dos ensaios que realizam (ISO, 2005 b). Também é necessário que os laboratórios assegurem a confiabilidade e comparabilidade dos resultados que emitem, os quais, na maioria das vezes, subsidiam tomadas de decisões relativas a aspectos econômicos, de saúde pública (BRASIL, 2006) e defesa do consumidor.

No contexto de implementação de sistemas de gestão da qualidade e acreditação (ISO, 2005 b), a validação de métodos tem representado um importante gargalo para os laboratórios de análises de alimentos e, conseqüentemente, um entrave para a credibilidade nos resultados emitidos (KING, 2003). Se por um lado tais laboratórios sofrem imposições devidas à necessidade de comprovação da confiabilidade dos resultados, por outro, existem limitações em função da diversidade e complexidade de analitos e matrizes, assim como da variedade de faixas de concentração.

Enquanto a necessidade de métodos de ensaio validados é clara, os mecanismos para o desenvolvimento de uma validação intralaboratorial adequada não estão bem definidos, nem acordados internacionalmente. Apesar dos esforços de organismos internacionais pela harmonização destes procedimentos (THOMPSON, ELLISON & WOOD, 2002), questões de interesse comum incluem: quais parâmetros de desempenho devem ser estudados, que procedimentos específicos devem ser utilizados para avaliar cada parâmetro e quais critérios de aceitabilidade devem ser definidos para cada parâmetro em particular (JENKE, 1996; HILL & REYNOLDS, 1999 TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004).

A maior parte das publicações que tratam sobre validações intralaboratoriais de métodos de ensaio limita-se às definições dos principais parâmetros de desempenho e a roteiros genéricos de procedimentos de validação (GREEN, 1996; ICH, 1996; EAL, 1997; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003). Embora exista consenso sobre alguns tipos de parâmetros que devem ser avaliados, o delineamento dos experimentos de validação, bem como as estatísticas aplicadas, são diversificados (GREEN, 1996), sendo observadas, freqüentemente,

recomendações contraditórias ou mesmo equivocadas. Na maioria das vezes, temas fundamentais são omitidos, como a representatividade do delineamento experimental e a necessidade de verificação de premissas relacionadas aos testes estatísticos adotados (THOMPSON, ELLISON & WOOD, 2002).

Desta forma, o desenvolvimento de um procedimento para validação intralaboratorial de métodos aplicável a análises de alimentos consiste em uma resposta à urgente necessidade de laboratórios e analistas frente a atual realidade de busca pela confiabilidade dos resultados no âmbito de acreditação de laboratórios.

Um procedimento detalhado para validação intralaboratorial de métodos quantitativos e qualitativos em análises de alimentos foi proposto neste trabalho, tendo como base fundamentos teóricos relacionados aos delineamentos experimentais e aos testes estatísticos para análise de dados e determinação dos diferentes parâmetros de desempenho, além de eficiência, praticidade e facilidade.

## **2 TEORIA**

### **2.1 LINEARIDADE**

#### **2.1.1 Delineamento experimental**

Experimentos de calibração requerem o estabelecimento de uma faixa preliminar de trabalho, que depende dos objetivos práticos da calibração. Sempre que possível, a concentração mais freqüentemente esperada para as amostras deve estar no centro desta faixa de interesse (ISO, 1996). Os valores medidos próximos ao limite inferior da faixa de trabalho devem ser distinguidos dos brancos (INMETRO, 2003).

DRAPER & SMITH (1998) discutem o delineamento experimental para análise de regressão, em função dos graus de liberdade dos resíduos (erros) da regressão a serem alocados entre as fontes de variação erro puro e desvio da linearidade na análise de variância. Estes autores avaliam diferentes configurações, em número de níveis de concentração e número de replicatas por nível, para um mesmo número total de pontos. O delineamento com três níveis de concentração, sendo dois extremos e um central, com maior número de replicatas nos níveis extremos, foi demonstrado como o melhor para avaliar desvio da linearidade. Contudo, a maior parte das referências relacionadas à validação intralaboratorial de métodos sugere um mínimo de cinco (GREEN, 1996; ICH, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; HUBER, 1998; EC, 2002; SHABIR, 2003) ou seis (EURACHEM, 1998; THOMPSON, ELLISON &

WOOD, 2002; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004) níveis de concentração. O número mínimo de replicatas predominantemente citado varia de dois a três (GREEN, 1996; ISO, 1996; AOAC, 1998 b; HUBER, 1998; THOMPSON, ELLISON & WOOD, 2002). No que diz respeito à independência das medições, cada nível de concentração deve corresponder a uma solução padrão preparada de forma independente (EURACHEM, 1998; INMETRO, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004), sendo desaconselhado o preparo de soluções por meio de diluições sucessivas (EURACHEM, 1998; INMETRO, 2003; FEINBERG & RAGUÈNÈS, 1999).

ISO (1996) e THOMPSON, ELLISON & WOOD (2002) recomendam que os níveis de concentração devam ser equidistantes ao longo da faixa estudada. O equilíbrio da distribuição dos níveis ao longo da faixa é importante, pois pode influenciar na identificação de pontos de alavanca (*leverages*) e, conseqüentemente, de valores extremos (*outliers*) ou (BURKE, 2001a).

As referências que citam o preparo do ponto zero ou branco da curva nos procedimentos de avaliação da linearidade (AOAC, 1998 b; EURACHEM, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004) não esclarecem sobre as condições de emprego deste ponto. Embora o nível zero seja necessário para ajustar o equipamento, a inclusão deste ponto na análise de regressão somente é justificada quando é definitivamente conhecido que o modelo linear é aplicável. É notório que pontos próximos aos extremos da faixa de trabalho são sujeitos a maiores oscilações do que os pontos centrais (MILLER & MILLER, 1993; MONTGOMERY & RUNGER, 1994; DRAPER & SMITH, 1998) e que desvio da linearidade é comum no extremo inferior de uma curva de calibração (RSC, 2001). Assim, o uso de um ponto zero pouco confiável na análise de regressão pode afetar a estimativa dos parâmetros, distorcer a distribuição dos resíduos da regressão e aumentar a variância dos mesmos.

As replicatas de cada nível de calibração fornecem informações sobre a variabilidade inerente às respostas de medição ou erro puro (DRAPER & SMITH, 1998). As replicatas devem ser independentes ou repetições genuínas (ISO, 1996; THOMPSON, ELLISON & WOOD, 2002), e não somente repetições de leitura. Quando as replicatas não são verdadeiras, a variância dos resíduos da regressão tende a ser subestimada e o teste de desvio da linearidade tende a detectar, erradamente, desvios de linearidade inexistentes (DRAPER & SMITH, 1998).

Também é expressivo que as medidas sejam feitas em ordem aleatória para evitar problemas de confundir desvio da linearidade com efeitos temporais como flutuações (RSC, 2005 a; THOMPSON, ELLISON & WOOD, 2002) e também para garantir a independência dos resultados.

### 2.1.2 O modelo

O modelo e a equação de regressão linear simples encontram-se descritos nas Equações I.1 e I.2, respectivamente, descritas a seguir:

$$Y_i = \alpha + \beta X_i + \varepsilon_i, \quad \text{Equação I.1}$$

sendo

$Y_i$  = variável dependente ou resposta,

$\alpha$  = interseção,

$\beta$  = inclinação,

$X_i$  = variável independente ou preditor,

$\varepsilon_i$  = resíduo da regressão.

$$\hat{y}_i = a + bx_i, \quad \text{Equação I.2}$$

sendo

$\hat{y}_i$  = variável dependente estimada pela equação de regressão,

$a$  = estimativa da interseção  $\alpha$ ,

$b$  = estimativa da inclinação  $\beta$ ,

$x_i$  = concentração conhecida do analito.

Para um dado valor de  $X_i$ , a observação correspondente  $Y_i$  consiste no valor  $\alpha + \beta X_i$  somado a quantidade  $\varepsilon_i$ , o incremento pelo qual uma observação individual de  $Y_i$  se afasta da linha de regressão. Nestas circunstâncias, a variável  $Y_i$  é sujeita a erro, enquanto a magnitude do erro na variável  $X_i$  é considerada desprezível (DRAPER & SMITH, 1998). Este modelo é freqüentemente aplicado e sustentado em curvas de calibração, embora com os adventos tecnológicos e de automação de alguns instrumentos esta condição possa ser questionável (MILLER & MILLER, 1993). Quando a incerteza do valor de referência de  $X_i$  é considerável, o método dos mínimos quadrados ordinários (MMQO) não deve ser utilizado. Vários procedimentos têm sido detalhados para ajuste dos dados ao modelo linear com erro nas duas variáveis, como o método dos

mínimos quadrados ortogonais (DANZER & KURRIE, 1998) e o método de máxima verossimilhança (RIPLEY & THOMPSON, 1987; RIU & RIUS, 1995).

A prática de constranger a curva de calibração, forçando-a a passar pela origem, é um assunto controverso. STRONG (1979) argumenta que a representação gráfica teórica de algumas medidas analíticas, especialmente das que seguem a Lei de Lambert-Beer, se ajustam a uma reta que passa pela origem. Considerando que as correções do branco da curva tenham sido cuidadosamente feitas, este autor propõe o uso de um modelo cuja equação tenha um único parâmetro  $Y_i = \beta X_i + \varepsilon_i$ , ao invés do modelo convencional de equação com dois parâmetros  $Y_i = \alpha + \beta X_i + \varepsilon_i$ . LEARY & MESSICK (1985) também discorrem sobre a constrição ser apropriada em casos nos quais a realidade física da situação de medição dita que a curva deve passar através de pontos teoricamente fixados. CAULCUTT & BODDY (1983) consideram a possibilidade de constrição, justificada por leis físicas relacionadas à absorvância e concentração, mas ponderam que no caso de haver tendências significativas estas justificativas são anuladas. MEIER & ZUND (1993) expressam a Lei de Lambert-Beer em correspondência à Equação 1.2 como:

$$A = A_{\text{branco}} + \text{comprimento}_{\text{do}} \text{percurso} \times \text{absortividade} \times \text{concentracao} ,$$

sendo  $y_i$  a absorvância  $A$ ,  $x_i$  a concentração,  $a$  a absorvância do branco da curva  $A_{\text{branco}}$  e  $b$  o produto entre o comprimento do percurso e a absortividade. Estes autores discutem que, apesar da constrição ser justificada sob um ponto de vista puramente científico, há conseqüências práticas como o aumento da variância dos resíduos da regressão. ELLERTON (1980) também debate sobre a escolha do modelo verdadeiro e discorda com o uso da constrição tomando como base o aumento de exatidão e precisão. Se um modelo de parâmetro único é estimado a partir de um conjunto de dados, quando o modelo verdadeiro seria de dois parâmetros, o erro padrão das estimativas é influenciado pela tendência resultante e, então, não há uma medida verdadeira da precisão. SCHWARTZ (1986) sugere que a constrição não seja feita nos casos em que a regressão é utilizada para prover uma curva de calibração da qual as concentrações de amostras desconhecidas seriam estimadas. Isto porque a constrição teria um efeito marcante sobre a estimativa da incerteza do ensaio. SNEDECOR & COCHRAN (1989) sustentam a necessidade de testar a hipótese nula de que a curva é uma linha reta que passa pela origem. O primeiro passo seria ajustar os dados ao modelo de dois parâmetros. A condição necessária para que a linha passe pela origem é que o intercepto não seja significativamente diferente de zero. Estes autores consideram que adotar um modelo de dois parâmetros seria uma aproximação melhor do que uma extrapolação pouco

confiável. DRAPER & SMITH (1998) manifestam a opinião de que a omissão de um parâmetro de um modelo é uma premissa muito forte, usualmente injustificada. Constranger uma curva de calibração forçando-a a passar pela origem, com base na correção do branco da curva, não é conduta confiável, uma vez que os valores medidos para o branco são estimados e não exatamente conhecidos (WENTWORTH, 1965).

### 2.1.3 Premissas

O método utilizado para análise de dados é o MMQO. O ajuste de uma equação de calibração pelo MMQO assume várias premissas relativas aos resíduos da regressão e ao modelo: i) os resíduos são variáveis aleatórias com média zero  $E(\varepsilon_i) = 0$  e variância  $V(\varepsilon_i) = \sigma^2$  constante e desconhecida; ii) os resíduos são variáveis normalmente distribuídas  $\varepsilon_i \sim N(0, \sigma^2)$ ; iii) os resíduos são homoscedásticos, com distribuição constante ao longo dos valores de  $X_i$ ; iv) o resíduo de uma observação  $\varepsilon_i$  não é correlacionado com o resíduo em outra observação  $\varepsilon_j$ , ou seja,  $cov(\varepsilon_i, \varepsilon_j) = 0$  sendo  $i \neq j$ . Os resíduos não são apenas não correlacionados, mas independentes; e v) a relação entre  $X_i$  e  $Y_i$  é linear (DRAPER & SMITH, 1998).

Algumas possíveis violações a estas premissas encontram-se na possibilidade de haver valores extremos no conjunto de dados, dos resíduos da regressão não seguirem a distribuição normal, não apresentarem variância constante (heteroscedasticidade), não serem independentes (autocorrelação) e a função não seguir o modelo linear (SNEDECOR & COCHRAN, 1989).

### Inspeção visual de gráficos

A análise dos gráficos  $x$ - $y$  (GREEN, 1996; ICH, 1996; ISO, 1996; NATA, 1997; EURACHEM, 1998; HUBER, 1998; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003) e dos gráficos dos resíduos da regressão (ISO, 1996; BRUCE, MINKKINEN & RIEKKOLA, 1998; DANZER & KURRIE, 1998; EURACHEM, 1998; HUBER, 1998; CHIAP et al., 1999; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004) é constantemente citada nas referências sobre validação intralaboratorial de métodos. Embora a inspeção visual seja propensa a erros de interpretação, devidos principalmente às características subjetivas desta análise, a análise dos gráficos  $x$ - $y$  e dos gráficos dos resíduos da regressão é fundamental como ferramenta complementar aos testes estatísticos formais (BURKE, 2001 a; WORKMAN & MARK, 2004; MARK & WORKMAN, 2005 b).

O exemplo clássico de ANSCOMBE (1973) revela a importância da inspeção visual do gráfico  $x$ - $y$  para a interpretação dos resultados da análise de regressão. Este autor apresenta quatro grupos fictícios de dados, com onze observações cada, e características obviamente diferentes. Para todos os grupos são obtidos os mesmos valores numéricos de parâmetros como inclinação, interseção, coeficiente de determinação e soma de quadrados dos resíduos da regressão. No entanto, apesar da semelhança numérica dos parâmetros da regressão, a observação da representação gráfica de cada grupo de dados revela perfis bastante diferentes, como um caso com presença de valores extremos e outro com desvio ao modelo linear.

MEIER & ZÜND (1993) afirmam que, se somente poucos pontos de calibração são disponíveis, o gráfico dos resíduos da regressão é útil para revelar possíveis tendências. Pelo formato deste gráfico é possível avaliar o ajuste dos dados ao modelo, inferir se os resíduos da regressão têm uma variância aproximadamente constante, ou se existem valores extremos. Este gráfico é uma ferramenta de diagnóstico valiosa que fornece informações sobre a estrutura dos dados e, conseqüentemente, sobre fatores de ponderação apropriados.

### **Valores extremos e pontos de alavanca**

Apesar das excelentes características do MMQO, este método tem o inconveniente de ser muito sensível à presença de pontos de influência como valores extremos ou pontos de alavanca. Valores extremos são definidos como observações de uma amostra tão separadas das demais que sugerem que possam pertencer a uma população diferente ou serem resultado de erros na medição. Alguns autores diferenciam a terminologia identificando valores extremos detectados entre níveis de confiança de 95 % e 99 % como *stragglers* e valores extremos diagnosticados com níveis de confiança maiores que 99 % como *outliers* (BURKE, 2001 b). Neste trabalho, não se fez distinção dos termos em função do nível de confiança. Os pontos de alavanca são aqueles pontos distantes dos demais pontos ao longo do eixo  $x$ , que afetam diretamente a estimativa dos parâmetros da equação da curva de calibração, agindo como uma alavanca (ARMITAGE & BERRY, 1994; ISO, 1994 b; BURKE, 2001 a).

Enquanto a inspeção visual dos dados (SNEDECOR & COCHRAN, 1989; MILLER & MILLER, 1993), dos gráficos  $x$ - $y$  (EURACHEM, 1998; INMETRO, 2003) e dos gráficos dos resíduos da regressão (MEIER & ZÜND, 1993; MILLER & MILLER, 1993) podem ser bastante informativos para identificação de valores extremos e pontos de alavanca, estatísticas mais formais de diagnóstico como as baseadas no cálculo do resíduo



padronizado, resíduo “studentizado”, distância de Cook e resíduo padronizado Jackknife (ou resíduo externamente “studentizado”) também são úteis (BELSLEY, KUH & WELSCH, 1980; ARMITAGE & BERRY, 1994; BARRET & LEWIS, 1994; MELOUN & MILITKÝ, 2001). Considerando que cada ferramenta de diagnóstico detecta um fenômeno específico nos dados, a seleção da estatística ideal deve ser cuidadosa. O cálculo do resíduo padronizado, resíduo “studentizado” e resíduo padronizado Jackknife, por exemplo, é baseado na análise dos resíduos da regressão, enquanto a distância de Cook é baseada na medida da influência exercida sobre os parâmetros estimados (MELOUN & MILITKÝ, 2001). Se um ponto é um valor extremo, ele terá um resíduo maior, o que irá inflar a estimativa da variância dos resíduos da regressão e, conseqüentemente, subestimar as estimativas dos resíduos padronizados e “studentizado”. Paradoxalmente, este efeito dificulta a detecção de um valor extremo (BELSLEY, KUH & WELSCH, 1980), prejudicando a eficiência destes diagnósticos. O cálculo do resíduo padronizado Jackknife emprega uma estimativa da variância dos resíduos da regressão independente do ponto sob suspeita. Estes resíduos são facilmente computados para todas as observações sem necessidade de ajustar várias regressões separadamente, para cada ponto excluído (BELSLEY, KUH & WELSCH, 1980; WEISBERG, 1985). A distância de Cook (COOK, 1977), recomendada por MULHOLLAND & HIBBERT (1997) para detectar valores extremos e comumente disponível em pacotes estatísticos, mede a influência de um ponto específico na estimativa dos parâmetros do modelo, combinando o resíduo da regressão e pontos de alavanca em uma única medida. Se um ponto não afeta significativamente a estimativa dos parâmetros do modelo, o valor de Cook não será significativo. Entretanto, o ponto pode influenciar fortemente a variância dos resíduos da regressão (MELOUN & MILITKÝ, 2001), o que torna o emprego da distância de Cook, em detrimento dos outros métodos, importante somente quando não há controle sobre pontos de alavanca, ou seja, sobre a distribuição dos pontos ao longo do eixo x no delineamento experimental.

### **Normalidade**

A validade de inferências produzidas por estatísticas paramétricas em amostras finitas, considerando que o tamanho das amostras seja controlado, depende crucialmente de premissas relacionadas ao tipo de distribuição. Conseqüentemente, para validade dos testes, é fundamental avaliar se as distribuições hipotetizadas são compatíveis com os dados. Não normalidade, por exemplo, é impeditiva para inferências baseadas nos testes de t e F. Testes de normalidade são prevalentes porque a premissa de normalidade é

feita com bastante frequência em análises estatísticas (DUFOUR et al., 1998). Esta premissa é citada pela ISO (1996) e por DANZER & KURRIE (1998) em procedimentos para avaliação da linearidade, mas tais autores não discutem sobre formas ou testes para avaliá-la. As revisões elaboradas por MARDIA (1980), D'AGOSTINO & STEPHENS (1986) e BARINGHAUS, DANSCHKE & HENZE (1989) reportam numerosos exemplos de testes de normalidade, aproximadamente 40, incluindo aqueles baseados em procedimentos gráficos e estatísticas formais. Os testes gráficos incluem histogramas e gráfico de normalidade, enquanto os estatísticos abrangem testes de hipótese como Kolmogorov-Smirnov, Anderson-Darling, Shapiro-Wilk, Jarque-Bera e D'Agostino. FILLIBEN (1975) propõe um teste de normalidade baseado na correlação linear entre as estatísticas observadas ordenadas e os valores esperados de uma distribuição normal, publicando uma tabela de valores críticos. RYAN & JOINER (1976) apresentam um teste bastante simples, fácil de calcular, visto que não são necessárias tabelas especiais para sua computação, e que combina gráficos de probabilidade normal com um teste estatístico formal de correlação. Este teste é numericamente similar ao teste de Filliben, cuja potência para amostras pequenas ( $n = 20$ ) foi demonstrada (FILLIBEN, 1975), além de ser equivalente ao potente teste de Shapiro-Wilk (RYAN & JOINER, 1976).

Quando uma variável não segue a distribuição normal, é possível criar uma variável transformada e testá-la para normalidade. BOX & COX (1964) sugerem um grupo de transformações que inclui raiz quadrada, logaritmo e inverso. DANZER & KURRIE (1998) indicam o emprego de métodos não paramétricos para calibração nos casos em que não houver distribuição normal.

### **Homoscedasticidade**

Regressão linear pelo MMQO assume que cada nível da faixa tem uma variação absoluta constante (homoscedasticidade). Todavia, muitos métodos de ensaio produzem dados que são heteroscedásticos, com valor relativo dos erros constante (BOX, 1953). Além da possibilidade de verificação desta premissa pela inspeção visual dos gráficos (ISO, 1996; MULHOLLAND & HIBBERT, 1997; EURACHEM, 1998; THOMPSON, ELLISON & WOOD, 2002), há vários testes formais possíveis para avaliação da homogeneidade das variâncias. O teste de Cochran, recomendado para tratamento de valores extremos em estudos colaborativos (HORWITZ, 1995; AOAC, 1997), também é reportado para verificação da homogeneidade das variâncias dos resíduos da regressão com o objetivo de selecionar a curva de calibração mais apropriada (CHIAP et al., 1999). Este teste assume que o número de replicatas de cada nível é igual ou similar e que há

um número suficiente de replicatas para obtenção de uma estimativa razoável da variância de cada nível (BURKE, 2001 b). Testes de F máximo (Hartley) e de Bartlett também são indicados para avaliação da homoscedasticidade dos resíduos (DANZER & CURRIE, 1998), embora sejam muito sensíveis à premissa de distribuição normal (SNEDECOR & COCHRAN, 1989). O teste de F máximo requer igual número de replicatas entre os níveis e não considera todos os níveis (DANZER & CURRIE, 1998), somente os de maior e menor variância. O teste de Bartlett permite que o número de replicatas seja desigual (DANZER & CURRIE, 1998). O teste estatístico para amostras de mesmo tamanho, proposto por LEVENE (1960), foi subsequente generalizado para amostras de tamanhos desiguais (DRAPER & HUNTER, 1969). Neste teste, o resultado estatístico é obtido por análise de variância inteiramente casualizada em que cada observação é substituída por seu desvio absoluto da média do grupo. Entretanto, para amostras muito pequenas, a validade do teste fica prejudicada, devido à alta correlação entre os desvios de um mesmo grupo. BROWN & FORSYTHE (1974) apresentam uma fórmula alternativa para o teste de Levene, robusta em condições de não normalidade, que utiliza a mediana em lugar da média e o módulo dos desvios para minimizar possíveis problemas causados pela correlação entre os desvios de um mesmo grupo.

Se houver heteroscedasticidade, é possível criar uma variável transformada e testá-la novamente para homogeneidade das variâncias. O método dos mínimos quadrados ponderados (MMQP), um tipo particular de método dos mínimos quadrados generalizados (MMQG), é comumente recomendado para dados heteroscedásticos (JOHNSTON, 1984).

### **Independência**

Uma correlação seriada dos resíduos da regressão é chamada autocorrelação. O MMQO assume que os resíduos são independentes. A autocorrelação dos resíduos afeta a variância das estimativas dos parâmetros obtidos pelo método mínimos quadrados e pode subestimar a estimativa da variância da regressão e os intervalos de confiança. Em testes de hipóteses, pode levar a inferências equivocadas, indicando falsa significância dos regressores (DRAPER & SMITH, 1998). Esta premissa é citada pela ISO (1996) e por DANZER & KURRIE (1998) em procedimentos para avaliação da linearidade, mas tais autores não discutem sobre formas ou testes para avaliá-la. O teste de Durbin-Watson estuda a estrutura ou aleatoriedade dos resíduos da regressão, sendo freqüentemente aplicado para determinar autocorrelação de primeira ordem (DURBIN & WATSON, 1951; HOFFMANN & VIEIRA, 1987; DRAPER & SMITH, 1998). Esta estatística também

apresenta potencial como teste para caracterização de especificações equivocadas de modelo, sendo importante para encontrar o número ótimo de termos latentes de um modelo (RUTLEDGE & BARROS, 2002). Trata-se de um teste muito sensível para verificar não linearidade, mas não ideal para ser aplicado com este propósito, visto a sua baixa potência para detectar esta condição (HOFFMANN & VIEIRA, 1987; MARK, 2003; MARK & WORKMAN, 2005 a).

No caso de dados autocorrelacionados, MMQO não é apropriado, sendo recomendadas transformações dos dados (LITTLE & HILLS, 1975) ou uso do MMQG (JOHNSTON, 1984).

### **Ajuste ao modelo linear**

Para algumas técnicas analíticas, o modelo linear pode não ser aplicado e modelos não lineares ou polinomiais são melhor adaptados. Alguns autores citam a análise visual de gráficos como uma forma de identificar desvios de linearidade (MULHOLLAND & HIBBERT, 1997; EURACHEM, 1998; RSC, 2005 a; INMETRO, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Contudo, testes de hipóteses formais também são descritos. DANZER & KURRIE (1998) e CODEX ALIMENTARIUS (2005 a) sugerem a avaliação da linearidade por testes de F, envolvendo a comparação entre a variância dos resíduos da regressão do modelo linear com a variância dos resíduos da regressão obtidos para um modelo não linear. MARK (2003) examina a linearidade pelo teste de quantos termos devem ser incluídos em uma função ajustada, baseado no princípio de que qualquer função pode ser aproximada por um modelo polinomial. Outros autores indicam uma avaliação do desvio da linearidade por análise de variância (ISO, 1996; CHIAP et al., 1999; FEINBERG & RAGUÈNÈS, 1999; RSC, 2005 a; THOMPSON, ELLISON & WOOD, 2002). Neste caso, estritamente falando, não é a linearidade ou o ajuste ao modelo linear que é avaliado, mas sim, é demonstrado que o desvio da linearidade é tão pequeno para ser medido que pode ser considerado como não significativo sob o ponto de vista estatístico. Esta avaliação pode ser obtida por meio de medições em replicatas de cada nível da calibração, o que fornece informações sobre a variabilidade inerente das medições ou erro puro (RSC, 2005 a; THOMPSON, ELLISON & WOOD, 2002). O desvio da linearidade consiste em dividir a variação total da resposta instrumental em três fontes: i) devida ao modelo da regressão; ii) devida ao desvio da linearidade; e iii) devida ao erro puro. Em seguida, a variância devida ao desvio da linearidade é comparada com a relativa ao erro puro (DRAPER & SMITH, 1998). As características e limitações deste teste devem ser compreendidas para interpretação de

seus resultados. Se uma curva de calibração tiver curvatura significativa, a hipótese nula de linearidade será rejeitada, sendo que muito cuidado deve ser tomado para encontrar o modelo mais apropriado (MONTGOMERY & RUNGER, 1994). Uma alternativa óbvia seria o ajuste a um modelo polinomial, mas a questão sobre o quão complexo este modelo precisa ser é difícil e fundamental. Por outro lado, se a hipótese nula não é rejeitada, não significa que o modelo linear seja correto, mas somente que dados obtidos não contradizem este modelo (DRAPER & SMITH, 1998) ou que os dados são insuficientes para detectar inadequações ao modelo (MEYER & ZUND, 1993). Somado a isto, em curvas de calibração existem causas de desvio de linearidade outras que não a não linearidade, sugerindo que o teste de desvio da linearidade seja sempre considerado em conjunto a inspeções visuais de gráficos (THOMPSON, ELLISON & WOOD, 2002) e testes formais das premissas dos resíduos da regressão.

#### **2.1.4 Coeficientes de determinação e correlação**

Apesar de ainda serem amplamente utilizados e até mesmo recomendados como indicadores de linearidade (GREEN, 1996; ICH, 1996; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; INMETRO, 2003), o coeficiente de determinação  $R^2$  e o coeficiente de correlação  $r$  não são adequados para este fim e não devem ser utilizados neste contexto (DANZER & KURRIE, 1998; FEINBERG & RAGUÈNÈS, 1999; BURKE, 2001 a; RSC, 2005 a; THOMPSON, ELLISON & WOOD, 2002; HORWITZ, 2003; MARK & WORKMAN, 2005 b). Estas estatísticas não medem a adequação de um modelo pois podem ser infladas pela adição de termos ao modelo (MONTGOMERY & RUNGER, 1994; MULHOLLAND & HIBBERT, 1997). Tais coeficientes indicam o grau de ajuste dos dados à curva, independentemente do modelo. Se os pontos de uma calibração são bem ajustados à curva, o valor experimental de  $r$  será próximo de um. Contudo, valores de  $r$  próximos da unidade não pertencem necessariamente a uma relação linear, podendo resultar de pontos bem ajustados a um modelo não linear (RSC, 2005 a). Também foi demonstrado que valores elevados de  $r$  podem coexistir com erros relativos excepcionalmente grandes no extremo inferior da curva de calibração (MULHOLLAND & HIBBERT, 1997).

Além de indicarem, de forma equivocada, o uso do  $r$  para avaliação da linearidade, GREEN (1996), BRUCE, MINKKINEN & RIEKKOLA (1998) e INMETRO (2003) estabelecem valores aceitáveis de 0,999; 0,995 a 0,997; e 0,90, respectivamente, para esta estatística. Entretanto, é importante observar que a significância do  $r$  é dependente do número de observações da regressão (BURKE, 2001 a).

## **2.2 EFEITOS DE MATRIZ**

Efeitos de matriz não significativos simplificam enormemente o processo de calibração, visto que os padrões podem ser preparados como soluções simples do analito em solventes (THOMPSON, ELLISON & WOOD, 2002). Todavia, utilizar soluções padrão simples para preparo da curva de calibração, sem a realização prévia dos testes de efeitos de matriz, significa assumir que estes efeitos não existem. Em diversas áreas, esta premissa é freqüentemente inválida (MILLER & MILLER, 1993), especialmente em análises de alimentos.

### **2.2.1 Delineamento experimental**

O delineamento experimental para testes de efeitos de matriz consiste do preparo de curvas de calibração simples (curvas de analito em solventes) e curvas de materiais de referência ou de adição de analito em amostras brancas ou não (curvas de analito em matriz) (INMAN, 1987; NATA, 1997; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003), com concentrações idênticas (BRUCE, MINKKINEN & RIEKKOLA, 1998; INMETRO, 2003) ou abrangendo a mesma faixa de concentração (THOMPSON, ELLISON & WOOD, 2002). Considera-se amostras brancas aquelas de matrizes isentas do analito ou para as quais o analito não tenha sido detectado. O uso de amostras típicas de cada matriz é fundamental para eliminar a possibilidade de efeitos de matriz que diferem em magnitude de uma amostra para a outra (MILLER & MILLER, 1993). As curvas de analito em matriz são preparadas por adição do analito a soluções teste, obtidas de amostras ensaiadas nas mesmas condições dos procedimentos normais (THOMPSON, ELLISON & WOOD, 2002). Isto para que as quantidades e características dos interferentes presentes no extrato final de uma amostra sejam reproduzidas nas curvas. Cumpre observar que não somente os interferentes provenientes da matriz são avaliados nestes experimentos, mas também todos aqueles provenientes de materiais, reagentes e soluções utilizados no preparo das amostras.

Alguns aspectos importantes do delineamento experimental para avaliação de efeitos de matriz não são tratados na literatura sobre validação intralaboratorial. Primeiro, a necessidade de que o preparo e o ensaio das curvas de solventes e de matriz sejam feitos no mesmo dia ou em dias diferentes, sob condições emparelhadas. Apesar desta consideração gerar uma restrição no número de níveis de concentração ou replicatas estudados, em função de limitações operacionais do laboratório ou da técnica, ela

sustenta a validade dos experimentos de efeitos de matriz em situações nas quais características críticas das curvas de calibração, como estabilidade e variabilidade em diferentes dias, não são conhecidas (HORWITZ, 2003). O segundo aspecto é relativo ao fato de que, em algumas análises de alimentos, não é possível fazer a adição do analito na solução teste obtida no final do processo de ensaio da amostra, por exemplo, em virtude da necessidade de reações prévias de derivação e da dificuldade de obtenção de produto da reação. Nestes casos, a amostra é adicionada do analito em etapas iniciais ou intermediárias do ensaio, assumindo-se que os efeitos sobre as respostas obtidas devidos a perdas do analito durante as etapas do processo analítico não são importantes.

Segundo INMAN (1987), estudos de efeitos de matriz se espelham nos experimentos de avaliação da linearidade. Assim, as considerações sobre o delineamento experimental para comprovação da linearidade também são válidas para o preparo das curvas de calibração destinadas às verificações dos efeitos de matriz.

### **2.2.2 Análise de dados**

No caso de disponibilidade de um conjunto de amostras de referência ou de amostras brancas verdadeiras, o teste de t é recomendado para comparação entre as médias obtidas nos ensaios com e sem matriz, para cada nível de concentração estudado (BRUCE, MINKKINEN & RIEKKOLA, 1998; INMETRO, 2003). Este teste assume que os dados seguem a distribuição normal (ARMITAGE & BERRY, 1994), sendo necessária uma verificação da premissa de normalidade para que qualquer inferência seja feita. O teste de t para comparação de médias de populações com variâncias desconhecidas pode ser calculado utilizando as variâncias amostrais de cada grupo ou por meio de uma variância combinada dos dois grupos. Neste último caso, há necessidade de verificação da premissa de homogeneidade das variâncias dos resultados dos grupos comparados (SNEDECOR & COCHRAN, 1989). O teste de F é indicado por BRUCE, MINKKINEN & RIEKKOLA (1998) para esta avaliação da homoscedasticidade. Quando a faixa de concentração estudada é pequena, permitindo que o erro analítico devido à concentração seja considerado desprezível, o teste de t emparelhado é sugerido. Este teste tem a vantagem de envolver todos os níveis de concentração estudados (BRUCE, MINKKINEN & RIEKKOLA, 1998), embora também exija a verificação das premissas de normalidade e homogeneidade das variâncias (THOMPSON, ELLISON & WOOD, 2002).

No entanto, quando não houver material de referência, situação mais próxima à realidade dos laboratórios de análises de alimentos, as inclinações das curvas com e sem matriz são comparadas (BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998;

THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003). Esta análise estatística baseia-se no fato de que os componentes da matriz podem afetar a estimativa dos parâmetros da equação da curva de calibração (BRUCE, MINKKINEN & RIEKKOLA, 1998). O teste preconizado para comparação das inclinações e interseções das curvas é o teste de t (ARMITAGE & BERRY, 1994). Aqui também existe a vantagem de que todos os níveis estudados são avaliados conjuntamente. No entanto, além da premissa de normalidade, a premissa de ajuste ao modelo linear também deve ser considerada. Isto porque a comparação das inclinações e interseções por este teste somente pode ser aplicada no caso de modelos lineares (EURACHEM, 1998; THOMPSON, ELLISON & WOOD, 2002). Quando modelos não lineares são diagnosticados, métodos bastante complexos são necessários para testes de significância ou, então, a avaliação fica limitada a uma comparação visual das curvas (THOMPSON, ELLISON & WOOD, 2002). Uma vez comprovadas as premissas de normalidade e ajuste ao modelo, o teste de t pode ser aplicado nas formas com ou sem variâncias combinadas, dependendo da homogeneidade das variâncias. A rigor, o teste de F máximo de HARTLEY (1950) foi elaborado para comparar a maior e a menor variâncias de um grupo de variâncias, obtidas com mesmo número de replicatas. O teste bilateral de F, proposto por SNEDECOR & COCHRAN (1989) possui a vantagem de permitir a avaliação da homogeneidade das variâncias dos resíduos da regressão das duas curvas, mesmo com diferentes números de observações, o que pode ser uma realidade em função de problemas experimentais ou do tratamento de valores extremos.

## **2.3 SELETIVIDADE**

### **2.3.1 Delineamento experimental**

Apesar dos estudos de efeitos de matriz garantirem uma avaliação parcial da seletividade do método, em relação aos interferentes presentes na amostra e no processo analítico, outras interferências também são prováveis. Entender os diferentes mecanismos que causam interferências pode ajudar na estruturação dos experimentos complementares para avaliação da seletividade e também na identificação de soluções para os problemas encontrados (BRUCE, MINKKINEN & RIEKKOLA, 1998; INMETRO, 2003). Estudos da seletividade são usualmente limitados pelo conhecimento sobre o analito e potenciais interferentes, como produtos de degradação, cujos mecanismos de formação são muitas vezes desconhecidos (INMAN, 1987; ICH, 1996). Há casos em que interferências químicas podem ser identificadas para um método em particular, mas as



chances de serem encontradas na prática podem ser improváveis. Também é muito difícil determinar que não há interferentes, uma vez que sempre existe a possibilidade de encontrar um interferente não reconhecido previamente. Assim, os analistas têm que equilibrar, com base em tempo, custos e benefícios, a busca por interferentes no processo de validação (EURACHEM, 1998).

Experimentos para avaliação da seletividade descritos na literatura sobre validação intralaboratorial de métodos envolvem ensaios com padrões ou materiais de referência, amostras brancas, amostras contendo o analito, além da avaliação da capacidade de identificação do analito de interesse na presença de interferentes (INMAN, 1987; GREEN, 1996; ICH, 1996; AOAC, 1998; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Quando não há disponibilidade de interferentes, alguns autores sugerem a avaliação da habilidade de medição do analito por diferentes métodos, técnicas ou por meio de variações nas condições instrumentais (INMAN, 1987; AOAC, 1998 b; EURACHEM, 1998; HUBER, 1998; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004).

### **2.3.2 Análise de dados**

Testes qualitativos sobrealvam a análise de dados obtidos em experimentos para avaliação da seletividade, podendo ser citados: i) perfil indistinguível da resposta do analito na amostra em relação à resposta obtida para o analito em soluções padrão (GREEN, 1996; AOAC, 1998); ii) obtenção de resultados positivos para amostras com analito e de resultados não detectados para amostras brancas (ICH, 1996); iii) distinção entre analitos e interferentes (ICH, 1996; AOAC, 1998); iv) verificação de efeitos de inibição ou aumento da detecção ou quantificação do analito na presença dos interferentes (EURACHEM, 1998; EC, 2002).

Medidas quantitativas podem ser obtidas em cromatografia por meio de estimativas da resolução e outros parâmetros cromatográficos (GREEN, 1996; ICH, 1996). THOMPSON, ELLISON & WOOD (2002) indicam o cálculo do índice de seletividade que corresponde à razão entre a sensibilidade do método (inclinação da curva de calibração) e a inclinação da resposta produzida por um potencial interferente. Contudo, estes autores não reportam critérios para análise e aceitabilidade destes resultados, sugerindo que sejam medidas úteis em processos comparativos de validação.

## 2.4 EXATIDÃO E PRECISÃO

### 2.4.1 Delineamento experimental

Experimentos para avaliação de exatidão envolvem ensaios com materiais de referência certificados, materiais de referência, amostras brancas e amostras adicionadas, enquanto na verificação da precisão são necessários ensaios replicados de amostras idênticas ou homogêneas, sob condições de repetitividade e reprodutibilidade parcial, preestabelecidas (GREEN, 1996; ICH, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004).

Materiais de referência certificados são relacionados a todos os aspectos da tendência (do método, do laboratório e do ensaio), por apresentarem incerteza estabelecida e rastreabilidade a padrões internacionais comprovada (THOMPSON, ELLISON & WOOD, 2002). Estes materiais são considerados ferramentas úteis para alcançar a rastreabilidade em medidas analíticas, para calibrar equipamentos e métodos, para monitorar a performance do laboratório, validar métodos e permitir a comparação entre métodos (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Entretanto, os materiais de referência certificados são caros e podem não ser representativos de amostras típicas (EURACHEM, 1998), não garantindo a exatidão do método. A condição ideal seria um material de referência certificado da matriz de interesse (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). No caso da precisão, o uso de materiais de referência certificados, os quais são freqüentemente bem mais homogêneos que amostras verdadeiras, pode subestimar a variação dos resultados (THOMPSON, ELLISON & WOOD, 2002).

Na área de alimentos, materiais de referência certificados são escassos e dispendiosos, principalmente devido às dificuldades no preparo, homogeneização e garantia de estabilidade durante armazenamento e transporte. Quando não há material de referência certificado disponível, uma amostra branca da matriz de interesse pode ser adicionada de quantidades conhecidas de analito na forma de material puro (ou de pureza conhecida) e estável ou de outro material bem caracterizado (THOMPSON & WOOD, 1995), para realização de ensaios de recuperação aparente (GREEN, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Outra opção é o método

de adição, empregado nos casos em que não é possível obter amostras brancas para os analitos e matrizes em questão (GREEN 1996; SHABIR, 2003). Estes experimentos permitem o estudo de diferentes tipos de analitos, concentrações e matrizes (TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Embora amostras adicionadas não tenham incertezas associadas, estudos de recuperação aparente com tais materiais permitem avaliar tendências devidas a efeitos relacionados ao analito adicionado (THOMPSON, ELLISON & WOOD, 2002). Na maioria das vezes, amostras adicionadas não representam fielmente a forma na qual o analito se encontra na matriz em amostras naturais, fornecendo uma impressão pouco realista de eficiência dos processos de extração (EURACHEM, 1998; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Em decorrência disto, recuperações aparentes satisfatórias em ensaios com amostras adicionadas não são garantia de exatidão, mas recuperações aparentes insatisfatórias são indicativas de falta de exatidão (THOMPSON, ELLISON & WOOD, 2002).

Ensaio das mesmas amostras pelo método em validação e por um outro método de referência também podem ser empregados para avaliação da exatidão (GREEN, 1996; ICH, 1996; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Todavia, estes experimentos exigem que o laboratório tenha condições de operar adequadamente os dois métodos.

Exatidão e precisão são geralmente dependentes da concentração do analito, devendo ser determinadas para um diferente número de concentrações (EURACHEM, 1998; INMETRO, 2003; HORWITZ, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Na literatura sobre validações intralaboratoriais de métodos, um mínimo de três níveis de concentração, com três a dez replicatas por nível, é freqüentemente recomendado (GREEN, 1996; ICH, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Nos experimentos para avaliação da exatidão e precisão, as variações entre ensaios devem representar o que normalmente acontece no laboratório em condições de rotina. Os materiais usados devem ser típicos em termos de tipos de matriz, devendo representar as amostras encontradas na aplicação do método. Variações de lotes e marcas de reagentes, analistas e equipamentos também devem ser representativas no delineamento dos parâmetros repetitividade e reprodutibilidade parcial (THOMPSON, ELLISON & WOOD, 2002).

## 2.4.2 Análise de dados

### Valores extremos

Valores extremos podem ser obtidos ocasionalmente nos experimentos para avaliação da exatidão e precisão (acurácia). Tais resultados devem ser rejeitados com o objetivo de garantir que estimativas das médias e desvios padrão não sejam contaminadas por equívocos ou erros grosseiros (KELLY, 1990; ARMITAGE & BERRY, 1994). Contudo, a rejeição de um valor extremo detectado é uma questão bastante complexa. Quando uma investigação das possíveis causas indicar motivos para a ocorrência do valor extremo, a eliminação do valor fica justificada. Mas, se nenhuma causa for encontrada, é difícil determinar critérios para rejeição ou não do dado. Alguns testes estatísticos fornecem a probabilidade de resultados discrepantes ocorrerem, embora não existam testes que possam provar a presença de um valor extremo. Assim, a decisão final sobre rejeição ou não de um valor extremo identificado fica ancorada na experiência de quem faz a análise dos dados (KELLY, 1990).

Muito tem sido publicado sobre procedimentos de detecção de valores extremos. Segundo SNEDECOR & COCHRAN (1989), as principais formas de proteção contra a ocorrência destes valores são o cuidado durante a condução dos ensaios, medições e registros e a inspeção visual dos dados antes da análise. ARMITAGE & BERRY (1994) também citam a inspeção dos dados para checagens lógicas de resultados que possam ser tanto impossíveis quanto implausíveis em virtude das características da variável sob medição.

A maioria dos testes de valores extremos busca por uma medida da distância relativa de um ponto suspeito e o valor médio, como os testes de Dixon e Grubbs. A maioria dos testes busca por um único valor, entretanto, algumas vezes é possível que vários valores extremos estejam presentes no mesmo conjunto de dados (BURKE, 2001 b). Nestas circunstâncias, dois valores extremos podem mascarar um ao outro de forma que nenhum deles seja identificado por um teste de detecção de um único *valor* (KELLY, 1990). Isto pode ser contornado de duas formas: i) pela aplicação iterativa dos testes de valores extremos; ou ii) pelo uso de testes que buscam pares de valores (BURKE, 2001 b), ou seja, testes separados para pares de valores extremos vicinais (dois no extremo inferior ou dois no extremo superior) ou para um par de valores extremos polares (um *valor* no extremo inferior e outro no extremo superior) (KELLY, 1990). Outra limitação do diagnóstico único é que a cada remoção de valor extremo, seguida por novo tratamento, há perda de um grau de liberdade e, conseqüentemente, perda na confiabilidade do teste.

O teste de Dixon pesquisa a existência de um único valor extremo (BURKE, 2001 b). A grande vantagem do teste de Dixon, baseado na razão entre faixas, é a facilidade para cálculo. No entanto, neste teste as regras mudam arbitrariamente conforme aumenta o número de observações, além do fato de que a tabela original serve para testes unilaterais, embora exista a possibilidade de existência de valores nos extremos inferior e superior. Diante disto, o tratamento de valores extremos utilizando testes combinados (KELLY, 1990), como as estatísticas de Grubbs para um valor extremo ( $G_1$ ), dois valores extremos vicinais ( $G_2$ ) e dois valores extremos polares ( $G_3$ ) (BURKE, 2001 b), é bastante razoável em substituição ao teste de Dixon. A forma de cálculo das estatísticas de Grubbs varia na literatura, assim como o nível de proteção adotado. HORWITZ (1995) e AOAC (1997) sugerem um nível de significância de 0,025 bilateral, a ISO 5725 (ISO, 1994 a,b,c,d,e; ISO, 1998) adota 0,05 bilateral, enquanto BARRET & LEWIS (1994) e BURKE (2001 b) indicam 0,10 bilateral (ou 0,05 unilateral). O grupo de autores que defende o uso do teste bilateral argumenta que testes unilaterais deveriam ser restritos para a detecção de valores extremos nos casos em que somente aqueles dados situados em uma direção especificada são de interesse. Contudo, os valores extremos são definidos com relação a um conjunto de dados que desperta o interesse na pesquisa destes dispersos. Portanto, se na formulação do teste é declarado que há um valor extremo superior, como preconizado no teste de Grubbs, parece mais aceitável usar o critério unilateral (BARRET & LEWIS, 1994).

Valores extremos também podem ser identificados com relação à influência que exercem sobre a dispersão dos resultados de um grupo de dados quando comparado com outros grupos. O teste de Cochran é conduzido para comparar a variância suspeita com a soma das variâncias de todos os outros grupos, assumindo que o número de replicatas de cada grupo seja igual ou similar e que haja um número suficiente de replicatas para obtenção de uma estimativa razoável da variância de cada grupo (BURKE, 2001 b). BURKE (2001 b) não recomenda que este teste seja aplicado iterativamente, alegando que pode levar a remoção de uma grande porcentagem de dados. Porém, HORWITZ (1995) e AOAC (1997) sugerem o emprego iterativo deste teste e do teste de Grubbs determinando um limite máximo de exclusão de 22,2 % dos dados. Como regra prática, se mais que 20 % dos dados originais for identificado como valor extremo, é importante questionar sobre a premissa de distribuição dos dados ou sobre a qualidade dos dados obtidos (KELLY, 1990; BURKE, 2001 b).

Alguns autores sugerem o emprego de estatísticas robustas (ou não paramétricas) como uma alternativa conveniente para análise de dados quando há suspeita de que

existem valores extremos (MILLER & MILLER, 1993; ARMITAGE & BERRY, 1994; BURKE, 2001 b; RSC, 2001 b). Alguns exemplos são o gráfico de Box-Whiskers (TUKEY, 1977), o desvio absoluto da mediana (MAD) (BURKE, 2001 b) ou a média ajustada (*trimmed*) (ARMITAGE & BERRY, 1994). As estatísticas robustas apresentam a vantagem de não serem influenciadas por desvios da distribuição normal, o que acontece com as estatísticas paramétricas e, portanto, são muito úteis para aumentar a confiabilidade dos testes de valores extremos quando a distribuição dos dados é desconhecida (BRUKE, 2001 b). No entanto, estatísticas robustas assumem que os dados seguem uma distribuição aproximadamente normal (unimodal e simétrica) contaminada por valores extremos, podendo gerar falsos resultados se aplicados a resultados notadamente assimétricos ou multimodais ou quando uma grande proporção dos dados apresentar valor idêntico (RSC, 2001 b).

### **Exatidão**

Diversos testes estatísticos são recomendados para avaliação da exatidão por comparação dos resultados obtidos pelo método em validação com os valores verdadeiros atribuídos a materiais de referência certificados. Alguns exemplos são o teste de t (NATA, 1996; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003), erro relativo (INMETRO, 2003), índice z-escore (INMETRO, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004) e erro normalizado (INMETRO, 2003) para comparação das médias, e o teste de F para comparação das variâncias (NATA, 1996; INMETRO, 2003). A literatura supracitada não trata as premissas assumidas para o teste de F, z-escore e erro normalizado (normalidade) nem para o teste de t (normalidade e homoscedasticidade).

Quando os experimentos são estruturados para comparação dos resultados obtidos por dois métodos ou técnicas independentes, o MMQO (NATA, 1996; BRUCE, MINKKINEN & RIEKKOLA, 1998; THOMPSON, ELLISON & WOOD, 2002) ou teste de t emparelhado, com comprovação das premissas de normalidade e homoscedasticidade (THOMPSON, ELLISON & WOOD, 2002), são sugeridos. As premissas relativas ao MMQO não são citadas nestes casos.

Nos ensaios com amostras adicionadas, a estatística universalmente utilizada é a recuperação aparente (GREEN, 1996; ICH, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004), a qual é posteriormente comparada com

valores críticos estabelecidos como requisitos analíticos do método, que variam em função da concentração do analito na amostra. GREEN (1996) reporta valores aceitáveis de médias de recuperação aparente de  $100\% \pm 2\%$ , em experimentos na área farmacêutica. BRUCE, MINKKINEM & RIEKKOLA (1998) referenciam os critérios estabelecidos pelo *United States Food and Drug Administration* (FDA) de  $80\%$  a  $110\%$  e de  $60\%$  a  $100\%$  para concentrações maiores que  $0,1\text{ mg/kg}$  e menores ou iguais a  $0,1\text{ mg/kg}$ , respectivamente. Estes autores também citam o critério estabelecido pela *AOAC International* de  $70\%$  a  $120\%$  para concentrações abaixo de partes por milhão. HORWITZ (1982) recomenda recuperação aparente para análises de traços, sendo valores mínimos aceitáveis de  $60\%$  e  $80\%$  para concentrações de  $\mu\text{g/kg}$  e  $\text{mg/kg}$ , respectivamente. A EC (2002) determina faixas de médias de recuperação aparente de  $50\%$  a  $120\%$ ,  $70\%$  a  $110\%$  e de  $80\%$  a  $110\%$ , respectivamente, para níveis  $\leq 1\ \mu\text{g/kg}$ ,  $> 1\ \mu\text{g/kg}$  a  $10\ \mu\text{g/kg}$  e  $\geq 10\ \mu\text{g/kg}$ . Existem, no entanto, regulamentações específicas de valores críticos de recuperação, como a estabelecida para métodos de ensaios de micotoxinas em alimentos pela EC (2006).

## Precisão

As estatísticas para análise de dados com objetivo de avaliação da precisão, sob condições de repetitividade e reprodutibilidade, estão bem estabelecidas pela ISO 5725 (ISO, 1994 a,b,c,d,e; ISO, 1998). Esta norma recomenda a análise de variância para este propósito, sendo citada em alguns documentos que tratam validação intralaboratorial de métodos (BRUCE, MINKKINEN & RIEKKOLA, 1998; KUTTATHARMMAKUL, MASSART & SMEYERS-VERBEKE, 1999; MAROTO et al., 1999; MAROTO et al., 2001; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO, 2003). A análise de variância assume várias premissas como: i) os resíduos são variáveis aleatórias, normalmente distribuídas  $\varepsilon_i \sim N(0, \sigma^2)$ ; ii) os resíduos são homoscedásticos, com distribuição constante ao longo dos valores de  $X_i$ ; iii) o resíduo de uma observação  $\varepsilon_i$  não é correlacionado com o resíduo em outra observação  $\varepsilon_j$ ,  $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$  com  $i \neq j$ ; e iv) os efeitos dos diferentes fatores são aditivos (SNEDECOR & COCHRAN, 1989). Embora, a verificação de tais premissas não esteja tratada na literatura sobre validação intralaboratorial de métodos consultada, a teoria apresentada para verificação das premissas de normalidade e homoscedasticidade dos resíduos da regressão é aplicável de forma análoga nos estudos de precisão. Como o teste de Durbin-Watson é proposto para estudos da independência dos resíduos da regressão (DURBIN & WATSON, 1951), a garantia desta

premissa fica ancorada nos cuidados com a aleatoriedade quando do delineamento experimental. A verificação da premissa de aditividade pode ser feita pelo teste de aditividade de Tukey (SNEDECOR & COCHRAN, 1989; TUKEY, 1949).

Em experimentos delineados somente sob condições de repetitividade, a estimativa direta do desvio padrão dos resultados é considerada (GREEN, 1996; ICH, 1996; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; INMETRO, 2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004).

A função de Horwitz (HORWITZ, 1982) constitui uma generalização útil e amplamente reconhecida do desvio padrão sob condições de reprodutibilidade em análises de alimentos. Valores estimados a partir desta equação são muitas vezes utilizados como requisitos analíticos para precisão. O “trompete de Horwitz”, proposto com base em inúmeros resultados de ensaios interlaboratoriais do setor de alimentos, demonstra que conforme a concentração do analito decresce em duas ordens de grandeza, o desvio padrão de reprodutibilidade aumenta com um fator de dois (RSC, 2004). Embora a função de Horwitz funcione bem independentemente do analito, matriz e natureza do material analisado (THOMPSON & LOWTHIAN, 1995), este modelo é deficiente para níveis de concentração extremos (THOMPSON, 2000). Dois modelos complementares à função de Horwitz foram propostos por THOMPSON (2000), em virtude das diferentes faixas de concentração do analito na amostra. As pesquisas deste autor, com dados provenientes de estudos colaborativos na área de alimentos, indicam desvios padrão de reprodutibilidade sistematicamente menores que aqueles preditos pela função de Horwitz, para concentrações de analito inferiores a 100 µg/kg. Contudo, regulamentações específicas de valores críticos de desvio padrão de repetitividade e reprodutibilidade também são estabelecidas (EC, 2006).

## **2.5 LIMITES**

### **2.5.1 Limites de detecção e de quantificação do método**

Existem vários experimentos possíveis para determinar os limites, incluindo aqueles baseados na inspeção visual (para métodos instrumentais e não instrumentais), na relação sinal/ruído, na média e desvio padrão de leituras de amostras brancas, no desvio padrão, intercepto ou inclinação da curva de calibração (GREEN, 1996; ICH, 1996; NATA, 1997; AOAC, 1998 b; BRUCE, MINKKINEN & RIEKKOLA, 1998; EURACHEM, 1998; HUBER, 1998; EC, 2002; THOMPSON, ELLISON & WOOD, 2002; INMETRO,



2003; SHABIR, 2003; TAVERNIERS, DE LOOSE & VAN BOCKSTAELE, 2004). Apesar de importantes referências serem dedicadas à discussão destes parâmetros, a natureza do sinal medido, as formas de cálculos, os critérios de identificação e a significância estatística não estão sistematicamente definidos (ANTIGNAC et al., 2003).

A despeito da aparente simplicidade da idéia, a determinação dos limites é envolvida por problemas como: i) há várias possibilidades conceituais para o assunto, cada uma fornecendo uma definição diferente para limite; ii) os procedimentos dependem de uma estimativa de precisão próxima da concentração zero, mas não está claro se isto deve ser feito sob condições de repetitividade ou alguma outra condição; iii) a menos que uma excessiva quantidade de dados seja coletada, estimativas dos limites são sujeitas a uma grande variação aleatória; iv) estimativas dos limites são freqüentemente tendenciosas para o lado inferior devido a fatores operacionais; e v) inferências estatísticas relativas aos limites dependem da premissa de distribuição normal, a qual é pelo menos questionável em concentrações baixas (THOMPSON, ELLISON & WOOD, 2002). Além disto, estudos da relação sinal/ruído têm o inconveniente de serem baseados na performance do método em um único dia. Enquanto experimentos baseados em determinações de amostras brancas são dependentes da matriz (EURACHEM, 1998).

Outra consideração importante é que os limites estimados no desenvolvimento do método podem não ser idênticos em conceito ou valor numérico aos limites usados para caracterizar o método de ensaio completo. O limite de detecção instrumental ou do equipamento, como consta na literatura ou em manuais de equipamentos, é freqüentemente menor que o limite de detecção prático ou do método e inapropriado para validação de métodos (THOMPSON, ELLISON & WOOD, 2002). Em ensaios de matrizes complexas, como os alimentos, o ruído obtido das matrizes excede bastante o ruído apresentado pelo equipamento, sendo necessário estabelecer o limite com base nos procedimentos completos do método e não somente na etapa de detecção do equipamento (BRUCE, MINKKINEN & RIEKKOLA, 1998; INMETRO, 2003).

### **2.5.2 Limite de decisão e capacidade de detecção**

EC (2002) preconiza a determinação dos parâmetros de desempenho limite de decisão ( $CC\alpha$ ) e capacidade de detecção ( $CC\beta$ ) em processos de validação, os quais têm sido extensivamente estudados para aplicação em ensaios de resíduos de drogas veterinárias em alimentos de origem animal, com a intenção de estabelecer presença ou ausência de substâncias banidas ou regulamentadas (FAJGELJ & AMBRUS, 2000).

Os experimentos para determinação de  $CC\alpha$  e  $CC\beta$  envolvem ensaios com diversos níveis de concentração ou com um único nível de concentração. No primeiro caso são preparadas curvas de calibração de amostras adicionadas (ISO, 1997 b; ISO, 2000; EC, 2002). No segundo, são realizados ensaios com 20 replicatas de amostras brancas para analitos proibidos ou 20 replicatas de amostras adicionadas no limite máximo permitido para analitos permitidos para estimativa do  $CC\alpha$  e ensaios com 20 replicatas de amostras adicionadas no limite de decisão para estimativa da  $CC\beta$  (EC, 2002). Os ensaios com curvas de amostras adicionadas apresentam a vantagem de englobarem diversos níveis de concentração, além da possibilidade dos dados serem obtidos dos experimentos com amostras adicionadas para avaliação de exatidão e precisão, sem necessidade de delineamento experimental específico para estimativa de  $CC\alpha$  e  $CC\beta$ .

Há várias formas de cálculo de  $CC\alpha$  e  $CC\beta$  descritas na literatura (ISO, 1997 b; ISO, 2000; EC, 2002; ANTIGNAC et al., 2003; VAN LOCO & BEERNAERT, 2003; PECORELLI et al., 2004; VAN OVERMEIRE et al., 2004). As estimativas de  $CC\alpha$  e  $CC\beta$ , a partir de experimentos com um único nível de concentração, são obtidas pela soma da concentração correspondente ao limite de decisão mais 1,64 vezes o desvio padrão de reprodutibilidade parcial e pelo valor do limite de decisão mais 1,64 vezes seu desvio padrão, respectivamente. Já as estimativas destes parâmetros por meio de experimentos com curvas de calibração, conforme descrito pela ISO 11843 (ISO, 1997 b; ISO, 2000) e pela EC (2002), tratam as bandas em torno da regressão como se fossem retas, quando na verdade são hipérbolas. Este procedimento inclui a construção de três curvas e uso destas curvas para estimativa dos parâmetros de desempenho, sendo: i) níveis de concentração estudados *versus* sinal médio obtido em cada nível; ii) níveis de concentração estudados *versus* sinal médio obtido em cada nível mais 1,64 vezes o desvio padrão de reprodutibilidade parcial; e iii) níveis de concentração estudados *versus* sinal médio obtido em cada nível menos 1,64 vezes o desvio padrão de reprodutibilidade parcial. VAN LOCO & BEERNAERT (2003) propõem fórmulas para estimativa direta de  $CC\alpha$  e  $CC\beta$ , baseadas no trabalho original de HUBAUX & VOS (1970), que dispensam a construção de curvas das bandas em torno da regressão e não assumem que estas bandas sejam lineares.

Sobre a interpretação dos resultados de  $CC\alpha$  e  $CC\beta$  existem duas escolas. A primeira, escola baseada na relação sinal/ruído, somente considera o controle de resultados falsos positivos (erro tipo  $\alpha$ ), o que leva a uma probabilidade resultados falsos

negativos (erro tipo  $\beta$ ) de 50 %. A segunda, escola baseada em testes de hipóteses, considera valores independentes para os dois tipos de erro (CURRIE, 1999; DESIMONI, 2004). Em concordância com a segunda escola, ANTIGNAC et al. (2003) estabelecem os seguintes critérios de interpretação: i) para resultados menores que os valores estabelecidos para  $CC\alpha$ , as amostras são declaradas como conformes ou com analitos em concentrações inferiores ao limite máximo estabelecido, com um nível de confiança de  $(1-\alpha)$ ; ii) para resultados maiores ou iguais aos valores estimados para  $CC\beta$ , as amostras são declaradas como não-conformes ou com analito presente em concentrações superiores ao limite máximo, com uma probabilidade de  $(1-\beta)$ ; e iii) existe uma faixa de incerteza estatística entre os valores de  $CC\alpha$  e  $CC\beta$ . Apesar da EC (2002) levar em conta a estimativa destes parâmetros considerando ambos os tipos de erro, somente considera  $CC\alpha$  como critério de conformidade na interpretação dos resultados, sugerindo que valores menores que os estabelecidos para  $CC\alpha$  sejam declarados como conformes e maiores ou iguais a  $CC\alpha$  como não-conformes.

### 3 PROCEDIMENTO PROPOSTO

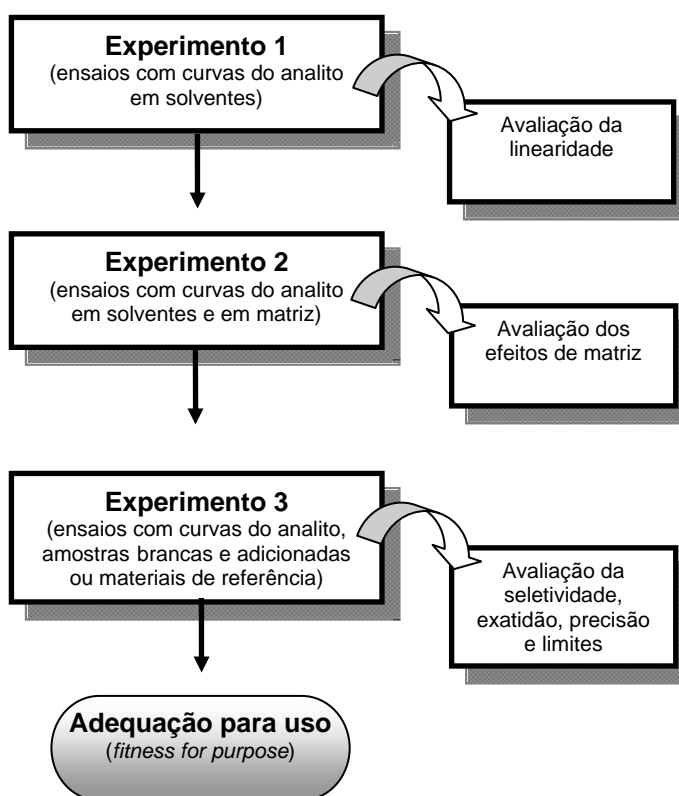
O procedimento para validação intralaboratorial de métodos quantitativos e qualitativos proposto neste trabalho, para aplicação em análises de alimentos, adotou a forma de validação direta com uso de padrões ou materiais de referência. Isto porque a validação comparativa tem a desvantagem de requerer que o laboratório tenha condições operacionais de executar um método de referência, além do método em validação.

Na validação de métodos quantitativos foram considerados parâmetros como linearidade, efeitos de matriz, seletividade, exatidão, precisão e limites. O processo foi dividido em três delineamentos experimentais distintos sendo: i) ensaios com curvas de calibração do analito em solventes para determinação da linearidade; ii) ensaios com uma curva de calibração do analito em solventes e outra do analito em matriz para verificação dos efeitos de matriz; e iii) ensaios com curvas de calibração (do analito em solventes, em matriz ou extratos sem matriz, dependendo dos resultados dos estudos de efeitos de matriz), amostras brancas e amostras adicionadas ou materiais de referência para verificação da seletividade, exatidão, precisão e limites (Figura I.1).

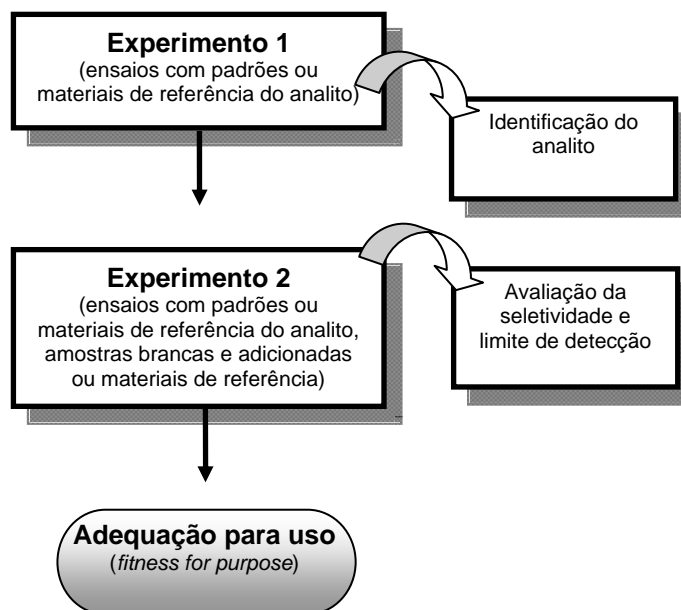
Para métodos qualitativos foram estudados identificação do analito, seletividade e limite de detecção, sendo dois delineamentos experimentais: i) ensaios com padrões ou materiais de referência para determinação das características de identificação do analito;

e ii) ensaios com padrões ou materiais de referência, amostras brancas e amostras adicionadas para verificação da seletividade e do limite de detecção (Figura I.2).

O procedimento sugerido assumiu que os laboratórios tinham sistemas de gestão da qualidade devidamente implementados, com base nos requisitos da norma ISO/IEC 17025 (ISO, 2005 b). O estabelecimento dos delineamentos experimentais e a seleção das ferramentas para análises de dados foram balizados pela fundamentação teórica apresentada no item 2, assim como por aspectos de eficiência, praticidade e facilidade. Os delineamentos experimentais sugeridos basearam-se em princípios de representatividade, enquanto as análises de dados sugeridas foram fundamentadas na verificação prévia das premissas relativas aos testes estatísticos selecionados. O nível de significância adotado nos testes de hipóteses foi  $\alpha = 0,05$ .



**Figura I.1** - Experimentos para validação de métodos quantitativos, incluindo formas de validação e abrangência em termos de parâmetros de desempenho avaliados.



**Figura I.2** - Experimentos para validação de métodos qualitativos, incluindo formas de validação e abrangência em termos de parâmetros de desempenho avaliados.

### 3.1 MÉTODOS QUANTITATIVOS

#### 3.1.1 Linearidade

##### Delineamento experimental

A primeira etapa do procedimento sugerido para avaliação da linearidade, para cada analito do escopo ou para analitos eleitos como representativos, é o delineamento experimental que se segue: i) determinação da faixa de concentração de interesse, considerando a aplicação do método e que as concentrações esperadas nas amostras devam estar próximas do centro desta faixa; ii) preparo de uma curva de calibração com soluções padrão em pelo menos seis níveis de concentração, igualmente espaçados, preparados independentemente, sendo no mínimo três replicatas independentes de cada nível. Inclusão do ponto zero (branco da curva) como uma ferramenta de controle de qualidade, avaliação do ruído e ajuste do zero instrumental; e iii) leitura das soluções em ordem aleatória, com programação do tempo e condições instrumentais entre leituras para evitar contaminação cruzada.

Nos casos em que não for possível preparar soluções em níveis igualmente espaçados, a presença de pontos de alavanca deve ser avaliada junto com a investigação de valores extremos.

Para métodos com objetivo de determinação de múltiplos analitos, as condições analíticas e instrumentais para separação, detecção e quantificação destes compostos são previamente ajustadas para que uns analitos não prejudiquem a detecção ou quantificação dos outros. Assim, as soluções padrão podem ser preparadas na forma de *pool*.

O experimento pode ser repetido em dias diferentes para observação da estabilidade das soluções e variabilidade das curvas em diferentes dias.

### **MMQO**

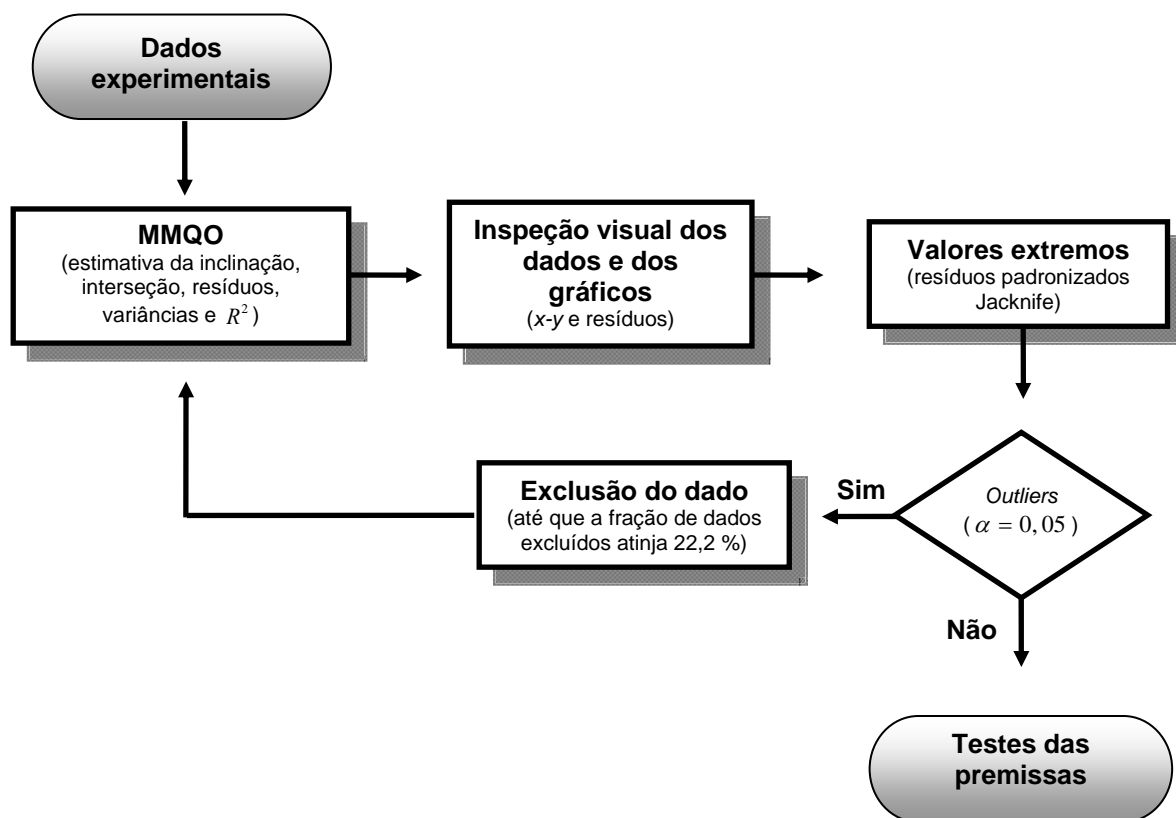
A segunda etapa da avaliação da linearidade consiste na aplicação do MMQO, incluindo estimativa dos parâmetros da regressão e tratamento de valores extremos: i) aquisição dos dados experimentais, ou seja, das respostas obtidas para as soluções padrão preparadas; ii) estimativa da inclinação, interseção, resíduos da regressão, respectivas variâncias e  $R^2$ ; iii) inspeção visual dos dados; iv) construção e inspeção visual do gráfico  $x$ - $y$  das respostas *versus* a concentração do analito; v) construção e inspeção visual do gráfico dos resíduos da regressão *versus* a concentração do analito; vi) investigação e exclusão de valores extremos pelo método dos resíduos padronizados Jackknife. Uma representação esquemática desta etapa está apresentada na Figura I.3.

Para métodos com objetivo de determinação de múltiplos analitos, os cálculos são feitos para cada analito separadamente.

#### *Estimativa dos parâmetros pelo MMQO*

Primeiramente é feita uma inspeção visual dos dados experimentais e construído um gráfico das respostas medidas *versus* as concentrações do analito para observação do comportamento dos resultados. Em seguida, os parâmetros da regressão  $\alpha$  e  $\beta$  são estimados pelo MMQO como  $a$  e  $b$ , respectivamente, correspondentes às quantidades que minimizam a soma de quadrados dos resíduos da regressão  $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ , sendo  $\hat{y}_i$  o valor estimado para  $Y_i$  para um dado valor de  $X_i$  quando  $a$  e  $b$  são determinados (Equação I.2).

As estimativas da inclinação  $b$  e da interseção  $a$  da reta ajustada estão demonstradas nas Equações I.3 e I.4, respectivamente.



**Figura I.3** - Avaliação da linearidade: estimativa dos parâmetros pelo MMQO, inspeção visual dos dados e gráficos e tratamento de valores extremos.

MMQO = método dos mínimos quadrados ordinários,  $R^2$  = coeficiente de determinação,  $y_i$  = resposta medida,  $x_i$  = concentração conhecida do analito,  $\alpha$  = significância.

$$b = \frac{S_{xy}}{S_{xx}},$$

Equação I.3

$$a = \bar{y} - b\bar{x},$$

Equação I.4

sendo

$$S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i / n,$$

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 / n,$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n},$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n},$$

$x_i$  = concentração conhecida do analito,

$y_i$  = resposta medida,

$n$  = número de pontos da curva de calibração.

Para cada valor  $x_i$ , no qual uma resposta  $y_i$  seja disponível, o resíduo  $e_i$  é dado pela Equação I.5.

$$e_i = y_i - \hat{y}_i, \quad \text{Equação I.5}$$

sendo

$\hat{y}_i$  = variável dependente estimada pela equação de regressão.

As variâncias da inclinação  $s_b^2$ , da interseção  $s_a^2$  e dos resíduos da regressão  $s_{res}^2$  são calculadas pelas Equações I.6, I.7 e I.8, respectivamente.

$$s_b^2 = \frac{s_{res}^2}{S_{xx}}, \quad \text{Equação I.6}$$

$$s_a^2 = s_{res}^2 \frac{\sum_{i=1}^n x_i^2}{nS_{xx}}, \quad \text{Equação I.7}$$

$$s_{res}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}. \quad \text{Equação I.8}$$

O coeficiente de determinação do ajuste  $R^2$  é definido pela Equação I.9.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad \text{Equação I.9}$$

Os valores estimados para a inclinação e interseção, assim como suas respectivas variâncias, são utilizados para construir a equação a ser utilizada para estimar as concentrações de amostras, com respectivas incertezas associadas. Os resíduos da regressão computados são empregados nos testes subseqüentes. Apesar das conhecidas limitações do  $R^2$ , esta estatística é avaliada para entendimento da proporção da variação total em torno da média das respostas que pode ser explicada pela regressão (DRAPER & SMITH, 1998).



### Tratamento de valores extremos

O gráfico dos resíduos da regressão *versus* os níveis de concentração é construído. Duas linhas pontilhadas horizontais correspondentes a  $\pm t_{(1-\alpha/2;n-2)} S_{res}$  são utilizadas para indicar uma faixa de variação aceitável para os resíduos da regressão representados no gráfico, sendo os casos de pontos fora destes limites percebidos como tendências. O formato da distribuição dos pontos no gráfico dos resíduos da regressão é comparado com formatos padrão para identificar valores extremos ou deficiências no modelo original (MEYER & ZUND, 1993). Formatos cônicos são indicativos de heteroscedasticidade, enquanto formas em “U” ou “U invertido” sugerem desvio de linearidade. Esta avaliação qualitativa é posteriormente confirmada por diagnósticos formais de valores extremos, homoscedasticidade e ajuste ao modelo.

Os valores extremos são avaliados pelo método dos resíduos padronizados Jacknife, cuja estatística é o resíduo padronizado Jacknife  $J_{ei}$ , calculado para cada ponto da curva de calibração, conforme descrito pela Equação I.10.

$$J_{ei} = r_i \sqrt{\frac{n-p-1}{n-p-r_i^2}}, \quad \text{Equação I.10}$$

sendo

$p$  = número de parâmetros do modelo,

$r_i = \frac{e_i}{s_{e_i}}$ , resíduo padronizado,

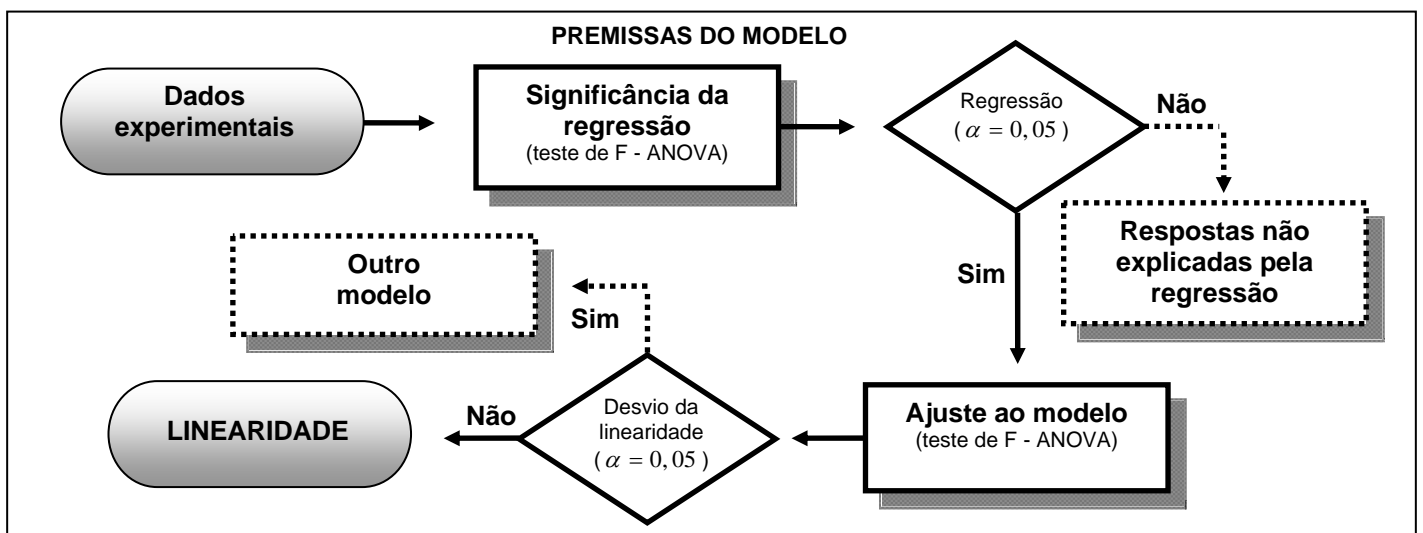
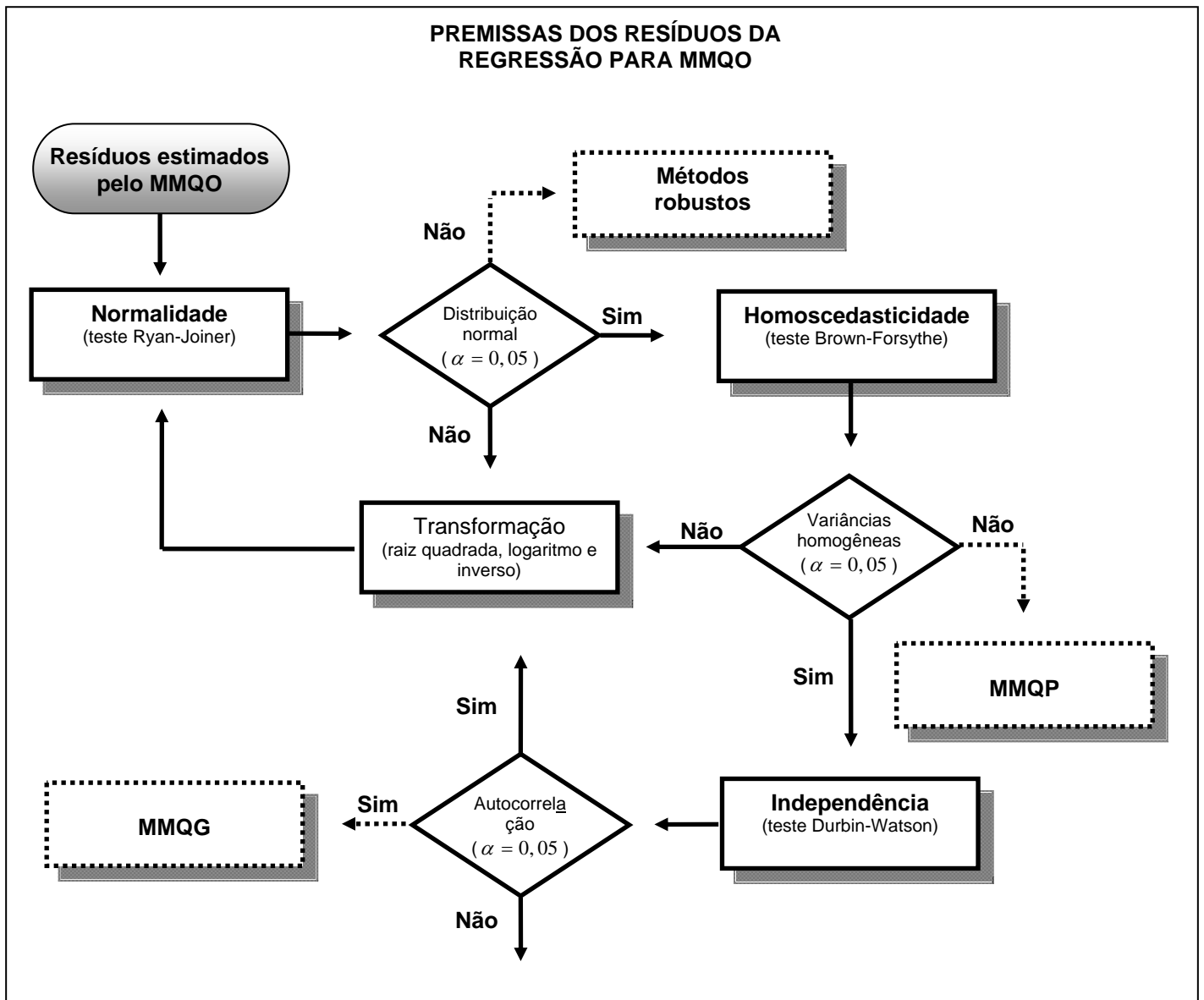
$s_{e_i} = s_{res} \sqrt{1-h_i}$ , erro padrão do resíduo,

$h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}$ , ponto de alavanca.

Os resíduos padronizados Jacknife seguem a distribuição de  $t_{(1-\alpha/2;n-p-1)}$ . Valores de  $J_{ei}$  maiores que o valor do  $t$  crítico são considerados valores extremos (BELSLEY, KUH & WELSCH, 1980) e removidos, exceto quando a porcentagem de dados tratados for superior a 22,2 % do número original de dados (HORWITZ, 1995) ou quando o ponto for a terceira e última replicata de um nível de concentração estudado. Para cada exclusão o MMQO é novamente aplicado aos dados.

### Testes de premissas

A terceira etapa do procedimento para avaliação da linearidade encontra-se ilustrada na Figura I.4. A validação do uso do MMQO é obtida por meio da verificação das



**Figura I.4 - Avaliação da linearidade: testes das premissas para uso apropriado do MMQO.**

MMQO = método dos mínimos quadrados ordinários,  $\alpha$  = significância, MMQP = método dos mínimos quadrados ponderados, MMQG = método dos mínimos quadrados generalizados, ANOVA = análise de variância.

Pontilhado indica caminhos não considerados pelo procedimento proposto.

premissas relativas aos resíduos da regressão e ajuste ao modelo linear: i) normalidade dos resíduos pelo teste de Ryan-Joiner; ii) homoscedasticidade dos resíduos pelo teste de Brown-Forsythe; iii) independência dos resíduos pelo teste de Durbin-Watson; e iv) verificação da significância da regressão e do ajuste ao modelo linear por análise de variância.

Para métodos com objetivo de determinação de múltiplos analitos, os cálculos são feitos para cada analito separadamente.

### *Teste de normalidade*

Normalidade dos resíduos da regressão é verificada pelo teste de Ryan-Joiner, que testa a hipótese nula de que os resíduos seguem a distribuição normal (os desvios da normalidade não são significativos) contra a hipótese alternativa de que os resíduos seguem outra distribuição de probabilidade (os desvios da normalidade são significativos). A estatística deste teste é o coeficiente de correlação do gráfico de probabilidade normal ou coeficiente de correlação de Ryan-Joiner  $R$ . A hipótese nula é rejeitada quando coeficientes de correlação calculados são menores que os valores críticos estabelecidos para estes coeficientes.

Primeiramente, os resíduos são ordenados. Em seguida, é construído um gráfico dos valores dos resíduos, dispostos em ordem crescente, *versus* os valores dos percentis estimados para uma distribuição normal reduzida (quantis normais). Denominando os resíduos ordenados de uma amostra de tamanho  $n$  como  $e_i$ , o gráfico de probabilidade normal, também conhecido como gráfico quantil-quantil (QQ), é construído com  $e_i$  *versus*  $q_i$ , sendo  $q_i$  correspondente aos pontos percentis de uma distribuição normal reduzida. Os quantis normais são obtidos pela Equação I.11.

$$q_i = \phi^{-1} \left\{ \frac{(i-3/8)}{(n+1/4)} \right\}, i = 1, \dots, n, \quad \text{Equação I.11}$$

sendo

$q_i$  = valor normal esperado,

$\phi^{-1}$  = inverso de uma função de distribuição normal padrão.

O coeficiente de correlação entre com  $e_i$  e  $q_i$  é computado de forma análoga ao coeficiente de correlação linear, conforme descrito na Equação I.12.

$$R_{eq} = \frac{S_{eq}}{\sqrt{S_{ee} \times S_{qq}}}, \quad \text{Equação I.12}$$

sendo

$$S_{eq} = \sum_{i=1}^n (e_i - \bar{e})(q_i - \bar{q}) = \sum_{i=1}^n e_i q_i - \sum_{i=1}^n e_i \sum_{i=1}^n q_i / n,$$

$$S_{ee} = \sum_{i=1}^n (e_i - \bar{e})^2 = \sum_{i=1}^n e_i^2 - \left( \sum_{i=1}^n e_i \right)^2 / n,$$

$$S_{qq} = \sum_{i=1}^n (q_i - \bar{q})^2 = \sum_{i=1}^n q_i^2 - \left( \sum_{i=1}^n q_i \right)^2 / n,$$

$$\bar{e} = \frac{\sum_{i=1}^n e_i}{n},$$

$$\bar{q} = \frac{\sum_{i=1}^n q_i}{n}.$$

Se os dados pertencerem a uma distribuição normal, o gráfico é obtido como uma linha reta. No entanto, se os dados forem provenientes de uma outra distribuição de probabilidade, algum grau de curvatura ou distribuição aleatórias dos dados são observados. Coeficientes de correlação abaixo de um valor crítico são traduzidos como rejeição da hipótese nula de normalidade dos resíduos. Coeficientes de correlação críticos para gráficos de probabilidade normal em diferentes níveis de significância podem ser obtidos, por interpolação polinomial, pelas Equações I.13, I.14 e I.15 (RYAN & JOINER, 1976).

$$R_{crit}(n) \approx 1.007 - \frac{0.1371}{\sqrt{n}} - \frac{0.3682}{n} + \frac{0.7780}{n^2} \text{ para } \alpha = 0,10, \quad \text{Equação I.13}$$

$$R_{crit}(n) \approx 1.0063 - \frac{0.1288}{\sqrt{n}} - \frac{0.6118}{n} + \frac{1.3505}{n^2} \text{ para } \alpha = 0,05, \quad \text{Equação I.14}$$

$$R_{crit}(n) \approx 0.9963 - \frac{0.0211}{\sqrt{n}} - \frac{1.4106}{n} + \frac{3.1791}{n^2} \text{ para } \alpha = 0,01. \quad \text{Equação I.15}$$

Quando o teste de Ryan-Joiner indicar que os dados apresentam desvio da normalidade, variáveis transformadas, como raiz quadrada, logaritmo e inverso da variável dependente, nesta ordem de prioridade, são novamente avaliadas (BOX & COX, 1964). Se o desvio de linearidade persistir após aplicação do teste de Ryan-Joiner às variáveis transformadas, métodos robustos são recomendados.

#### *Teste de homoscedasticidade*

O teste de Levene (LEVENE, 1960) modificado por BROWN & FORSYTHE (1974) é utilizado para avaliar homoscedasticidade, confrontando a hipótese nula de que as

variâncias dos resíduos da regressão não diferem entre si (há homoscedasticidade) contra a hipótese alternativa de que as variâncias dos resíduos são diferentes (há heteroscedasticidade). A estatística deste teste é o F de Levene  $F_L$ , mas no caso específico de 2 grupos, ou seja, para 1 grau de liberdade de tratamentos, em que  $t = \sqrt{F}$ , pode-se calcular a estatística t de Levene  $t_L$ . A hipótese nula é rejeitada quando a estatística  $t_L$  calculada for maior que o t crítico  $t_{(1-\alpha/2; n_1+n_2-2)}$ .

Os resíduos são divididos em dois grupos de tamanho  $n_1$  e  $n_2$ , sempre que possível equilibrando a distribuição dos pontos entre os grupos em função dos níveis de concentração  $x_i$  estudados, sem separar as replicatas de um mesmo nível em grupos diferentes. As medianas dos resíduos de cada grupo  $\tilde{e}_1$  e  $\tilde{e}_2$  são calculadas. Os valores absolutos das diferenças entre os resíduos e as medianas dos respectivos grupos  $d_{1j} = |\tilde{e}_1 - e_{1j}|$  e  $d_{2j} = |\tilde{e}_2 - e_{2j}|$  são obtidos. A média das diferenças  $\bar{d}_k$  e a soma dos quadrados dos desvios  $SQD_k$  dos valores  $d_{kj}$  de cada grupo  $k=1$  e  $k=2$  são calculados. A estatística  $t_L$  é obtida pela Equação I.16.

$$t_L = \frac{(\bar{d}_1 - \bar{d}_2)}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) s_p^2}}, \quad \text{Equação I.16}$$

sendo

$\bar{d}_k = \sum |\tilde{e}_k - e_{kj}| / n_k$ , média dos módulos das diferenças entre o j-ésimo resíduo e a mediana de cada grupo  $k=1$  e  $k=2$ ,

$n_k$  = número de observações em cada grupo,

$\tilde{e}_k$  = mediana de cada grupo,

$s_p^2 = \frac{(SQD_1 + SQD_2)}{(n_1 + n_2 - 2)}$ , variância combinada,

$SQD_k = \sum_{j=1}^{n_k} (\tilde{e}_k - e_{kj})^2$ , soma dos quadrados dos desvios entre cada j-ésimo resíduo e a mediana de seu k-ésimo grupo para cada grupo.

Se a estatística  $t_L$  calculada não exceder o valor crítico  $t_{(1-\alpha/2; n_1+n_2-2)}$ , há indicação de homoscedasticidade, não havendo razões para rejeitar a hipótese nula e acreditar que a variância dos resíduos não seja constante (LEVENE, 1960; BROWN & FORSYTHE, 1974).

Contudo, se os dados apresentarem heteroscedasticidade, o teste de Levene modificado é repetido para variáveis transformadas, como raiz quadrada, logaritmo e inverso dos resíduos da regressão (BOX & COX, 1964). Se a variância dos resíduos não for constante também nestes casos, o MMQP é indicado.

### Teste de independência

Autocorrelação dos resíduos da regressão é verificada pelo teste de Durbin-Watson, que confronta a hipótese nula de que não há autocorrelação dos resíduos (os resíduos são independentes) com a hipótese alternativa de que há autocorrelação dos resíduos (os resíduos não são independentes). Assumindo que os resíduos  $e_i$  são variáveis independentes, a autocorrelação entre os resíduos  $\rho_s = 0$ . A hipótese nula  $\rho_s = 0$  é testada contra a alternativa em que  $\rho_s = \rho$  ( $\rho \neq 0$  e  $|\rho| < 1$ ). A estatística deste teste é a estatística de Durbin-Watson  $d$ , definida pela Equação I.17.

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} . \quad \text{Equação I.17}$$

Para cada conjunto de dados, há dois limites críticos  $d_L$  (limite inferior) e  $d_U$  (limite superior). Se o valor de  $d$  estiver entre estes limites, o teste é inconclusivo. Valores de  $d < d_L$  indicam autocorrelação e rejeição da hipótese nula a um nível de significância de  $2\alpha$ , enquanto valores de  $d > d_U$  indicam independência dos resíduos e não rejeição da hipótese nula. O valor da estatística de Durbin-Watson varia de zero a quatro com média de dois. Se o valor calculado converge para dois, significa que não há autocorrelação, ou seja, os resíduos são independentes. Quando o valor se afasta de dois para zero ou quatro, a autocorrelação aumenta. Valores de Durbin-Watson iguais a zero indicam perfeita autocorrelação positiva, enquanto valores iguais a quatro indicam perfeita autocorrelação negativa. Geralmente, valores 1,5 e 2,5 podem ser utilizados como pontos de corte inferior e superior (DURBIN & WATSON, 1951). Tabelas que fornecem os limites inferior e superior de  $d$  dependentes do nível de significância  $\alpha$ , do número de pontos e do número de variáveis preditas são disponíveis. Para evitar o uso de tabelas, os valores limites de  $d$  podem ser estimados pelas Equações I.18 a 1.23 propostas a seguir, considerando diferentes níveis de significância e uma variável predita.

$$d_L \approx 1.9693 - \frac{2.8607}{\sqrt{n}} - \frac{3.4148}{n} + \frac{16.6400}{n^2} \text{ para } \alpha = 0,05 , \quad \text{Equação I.18}$$

$$d_U \approx 1.9832 - \frac{3.0547}{\sqrt{n}} + \frac{1.3862}{n} + \frac{16.3662}{n^2} \text{ para } \alpha = 0,05, \quad \text{Equação I.19}$$

$$d_L \approx 1.9845 - \frac{3.6875}{\sqrt{n}} - \frac{2.6136}{n} + \frac{20.6393}{n^2} \text{ para } \alpha = 0,025, \quad \text{Equação I.20}$$

$$d_U \approx 1.9480 - \frac{3.1647}{\sqrt{n}} - \frac{0.6472}{n} + \frac{31.5772}{n^2} \text{ para } \alpha = 0,025, \quad \text{Equação I.21}$$

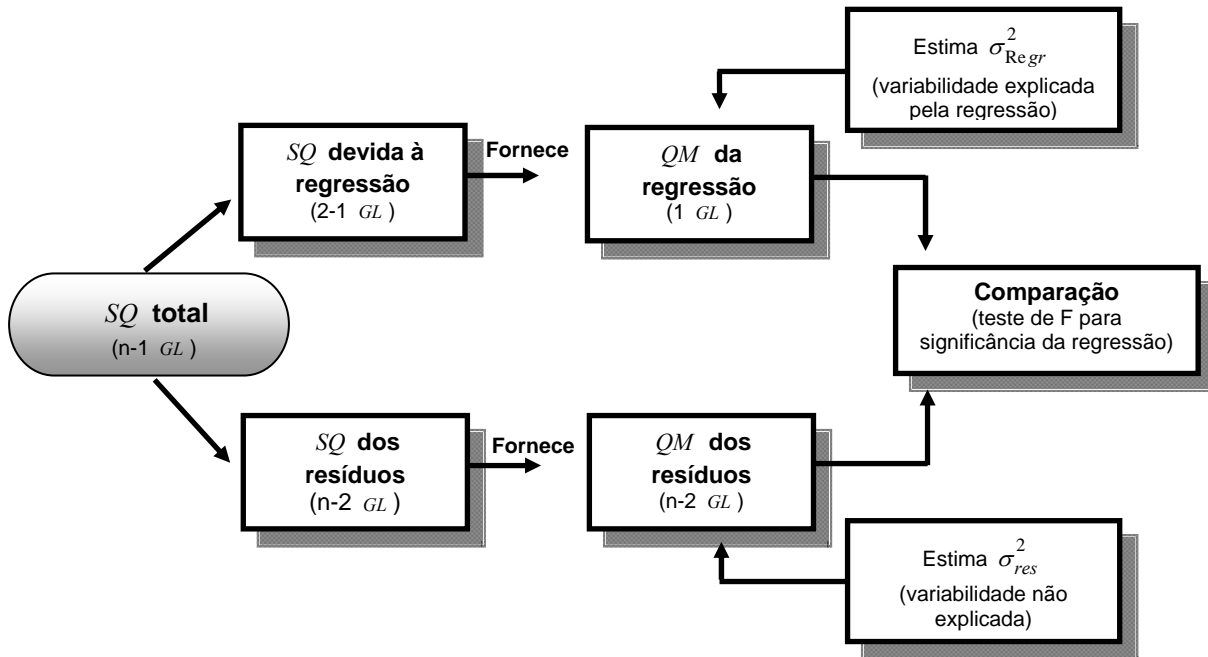
$$d_L \approx 1.9934 - \frac{4.5929}{\sqrt{n}} - \frac{1.3228}{n} + \frac{20.2288}{n^2} \text{ para } \alpha = 0,01, \quad \text{Equação I.22}$$

$$d_U \approx 1.9784 - \frac{4.2974}{\sqrt{n}} + \frac{1.0812}{n} + \frac{29.6862}{n^2} \text{ para } \alpha = 0,01. \quad \text{Equação I.23}$$

A independência também pode ser demonstrada graficamente construindo-se um gráfico dos valores de  $e_i$  versus seus valores imediatamente anteriores  $e_{i-1}$ . Um perfil de distribuição aleatória dos resíduos nos quatro quadrantes, indica independência. Concentração dos pontos no primeiro e terceiro quadrantes indica correlação positiva, enquanto predominância dos pontos no segundo e quarto quadrantes indica correlação negativa.

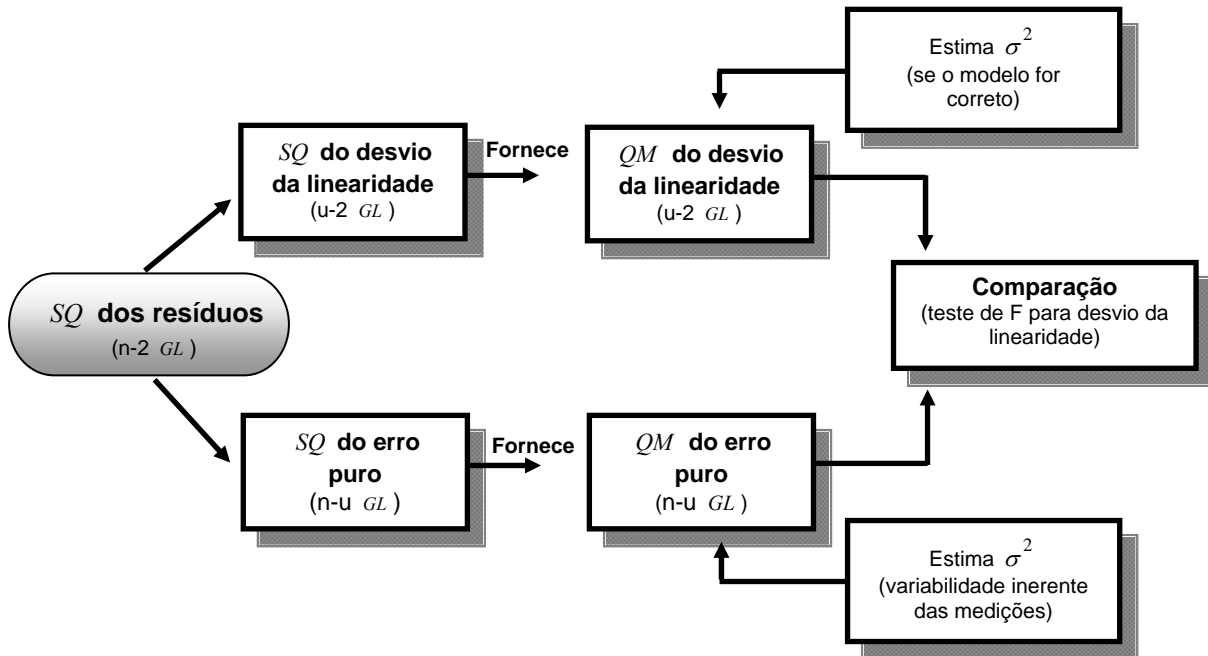
#### *Teste de significância da regressão e do desvio da linearidade*

Para realizar o teste da significância da regressão e do desvio da linearidade, a variabilidade total das respostas é decomposta em soma de quadrados dos resíduos da regressão (em torno da regressão) e soma dos quadrados devida à regressão (Figura I.5). A soma de quadrados dos resíduos da regressão é então separada em soma de quadrados do desvio da linearidade (falta de ajuste ao modelo) e soma de quadrados do erro puro (Figura I.6). Uma partição da variabilidade total das respostas, em soma de quadrados entre níveis e soma de quadrados do erro puro (dentro dos níveis), também é feita para auxiliar os cálculos (Figura I.7).



**Figura I.5** - Decomposição da soma de quadrados total para estimativa da significância da regressão por análise de variância.

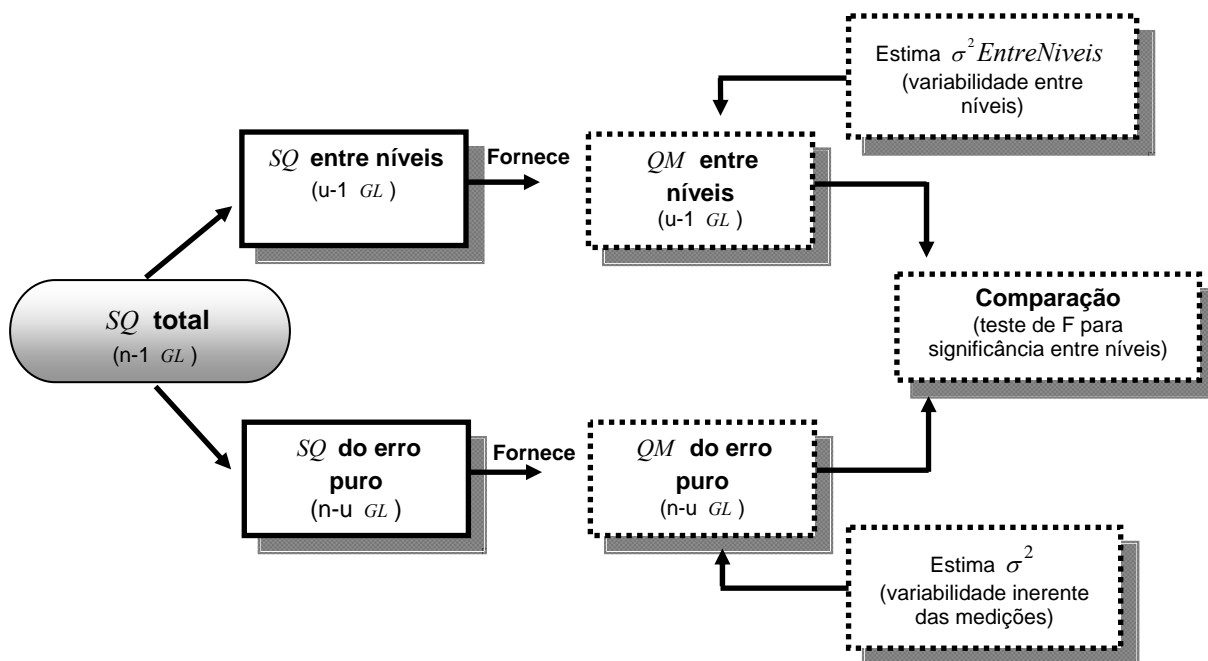
$SQ$  = soma de quadrados,  $n$  = número de pontos da curva de calibração,  $GL$  = graus de liberdade,  $QM$  = quadrado médio,  $\sigma^2$  = variância.



**Figura I.6** - Decomposição da soma de quadrados dos resíduos da regressão para estimativa da significância do desvio da linearidade por análise de variância.

$SQ$  = soma de quadrados,  $n$  = número de pontos da curva de calibração,  $GL$  = graus de liberdade,  $u$  = número de níveis de concentração,  $QM$  = quadrado médio,  $\sigma^2$  = variância.





**Figura I.7** - Decomposição auxiliar da soma de quadrados total.

$SQ$  = soma de quadrados,  $n$  = número de pontos da curva de calibração,  $GL$  = graus de liberdade,  $u$  = número de níveis de concentração,  $QM$  = quadrado médio,  $\sigma^2$  = variância.

Pontilhado indica etapas da partição que não são necessárias na avaliação da significância do desvio da linearidade.

A tabela da análise de variância para este teste é construída a partir das equações representadas na Tabela I.1. Contudo, o uso de todas as equações descritas nesta Tabela não é necessário. A soma de quadrados dos resíduos pode ser obtida pela diferença entre a soma de quadrados total e a soma de quadrados devida à regressão. A soma de quadrados do erro puro é obtida pela diferença entre a soma de quadrados total e a soma de quadrados entre níveis. Similarmente, a soma de quadrados do desvio da linearidade pode ser calculada: i) subtraindo a soma de quadrados do erro puro da soma de quadrados dos resíduos; ou ii) subtraindo a soma de quadrados entre níveis da soma de quadrados devida à regressão. A estatística deste teste é a razão entre as variâncias, que segue a distribuição de  $F$  com os graus de liberdade correspondentes.

**Tabela I.1** - Análise de variância para significância da regressão e do desvio da linearidade

<i>FV</i>	<i>GL</i>	<i>SQ</i>	<i>QM</i>	<i>F</i>
Regressão	1	$\frac{\left[ \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \right]^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$	$\frac{SQ_{Regr}}{GL_{Regr}}$	$\frac{QM_{Regr}}{S_{res}^2}$
Resíduos	$n - 2$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SQ_{res}}{GL_{res}} = s_{res}^2$	
Desvio da linearidade	$u - 2$	$\sum_{k=1}^u n_k (\hat{y}_k - \bar{y}_k)^2$	$\frac{SQ_{Desvio}}{GL_{Desvio}}$	$\frac{QM_{Desvio}}{QM_{ErroPuro}}$
Entre níveis	$u - 1$	$\sum_{k=1}^u n_k (\bar{y}_k - \bar{y})^2$		
Erro puro	$n - u$	$\sum_{k=1}^u \sum_{j=1}^{n_k} (y_{kj} - \bar{y}_k)^2$	$\frac{SQ_{ErroPuro}}{GL_{ErroPuro}}$	
Total	$n - 1$	$\sum_{i=1}^n (y_i - \bar{y})^2$	$\frac{SQ_{Total}}{GL_{Total}}$	

*FV* = fonte de variação, *GL* = graus de liberdade, *SQ* = soma de quadrados, *QM* = quadrado médio, *F* = razão entre variâncias, *n* = número de *i* pontos da curva de calibração, *u* = número de níveis de concentração, *n<sub>k</sub>* = número de pontos *j* em cada nível de concentração *k*, *y<sub>i</sub>* = resposta medida, *x<sub>i</sub>* = concentração conhecida do analito, *y<sub>i</sub>* = variável dependente estimada pela equação de regressão,  $\bar{y}$  = média das respostas medidas,  $\bar{x}$  = média das concentrações conhecidas, *y<sub>kj</sub>* = *j*-ésima resposta medida do *k*-ésimo nível de concentração,  $\bar{y}_k$  = média das respostas medidas do *k*-ésimo nível de concentração,  $\hat{y}_k$  = variável dependente estimada pela equação de regressão para o *k*-ésimo nível de concentração.

Para avaliação da significância da regressão, é testada a hipótese nula de que a regressão linear não é significativa (variação dos valores de *y<sub>i</sub>* não é explicada pela regressão) contra a hipótese alternativa de que a regressão linear é significativa (variação dos valores de *y<sub>i</sub>* é explicada pela regressão). A hipótese nula é rejeitada quando a estatística *F* estimada for maior que o valor crítico  $F_{(1-\alpha;1;n-2)}$ . Para avaliação da significância do desvio da linearidade, a hipótese nula de que o modelo linear simples é correto (não há desvio da linearidade) é confrontada com a hipótese alternativa de que o modelo linear não é correto (há desvio da linearidade). A hipótese nula é rejeitada quando a estatística *F* estimada for maior que o valor crítico  $F_{(1-\alpha;u-2;n-u)}$ .

Um desvio da linearidade significativo indica que o modelo é inadequado. Atenção deve então ser dedicada para descobrir onde e como esta inadequação ocorre. Desvio da linearidade não significativo indica que não existem razões para duvidar da adequação do modelo e que ambos, soma de quadrados do erro puro e do desvio da linearidade, podem ser utilizados como estimativas de  $\sigma^2$  (DRAPER & SMITH, 1998).

### **3.1.2 Efeitos de matriz**

#### **Delineamento experimental**

A primeira etapa do procedimento sugerido para avaliação dos efeitos de matriz, para cada combinação de analito e matriz do escopo ou para analitos e matrizes eleitos como representativos, é o delineamento experimental que se segue: i) preparo de duas curvas, sendo uma curva do analito em solventes, elaborada com soluções padrão de calibração, e outra curva do analito em matriz, construída com soluções padrão de calibração derivadas de soluções teste, obtidas de amostras ensaiadas nas mesmas condições dos procedimentos normais; ii) preparo independente de pelo menos seis níveis de concentração, igualmente espaçados, sendo no mínimo três replicatas independentes de cada nível. Inclusão do ponto zero (branco da curva) como uma ferramenta de controle de qualidade, avaliação do ruído e ajuste do zero instrumental; e iii) leitura das soluções das curvas do analito em solventes e matriz em ordem aleatória, com programação do tempo e condições instrumentais entre leituras para evitar contaminação cruzada.

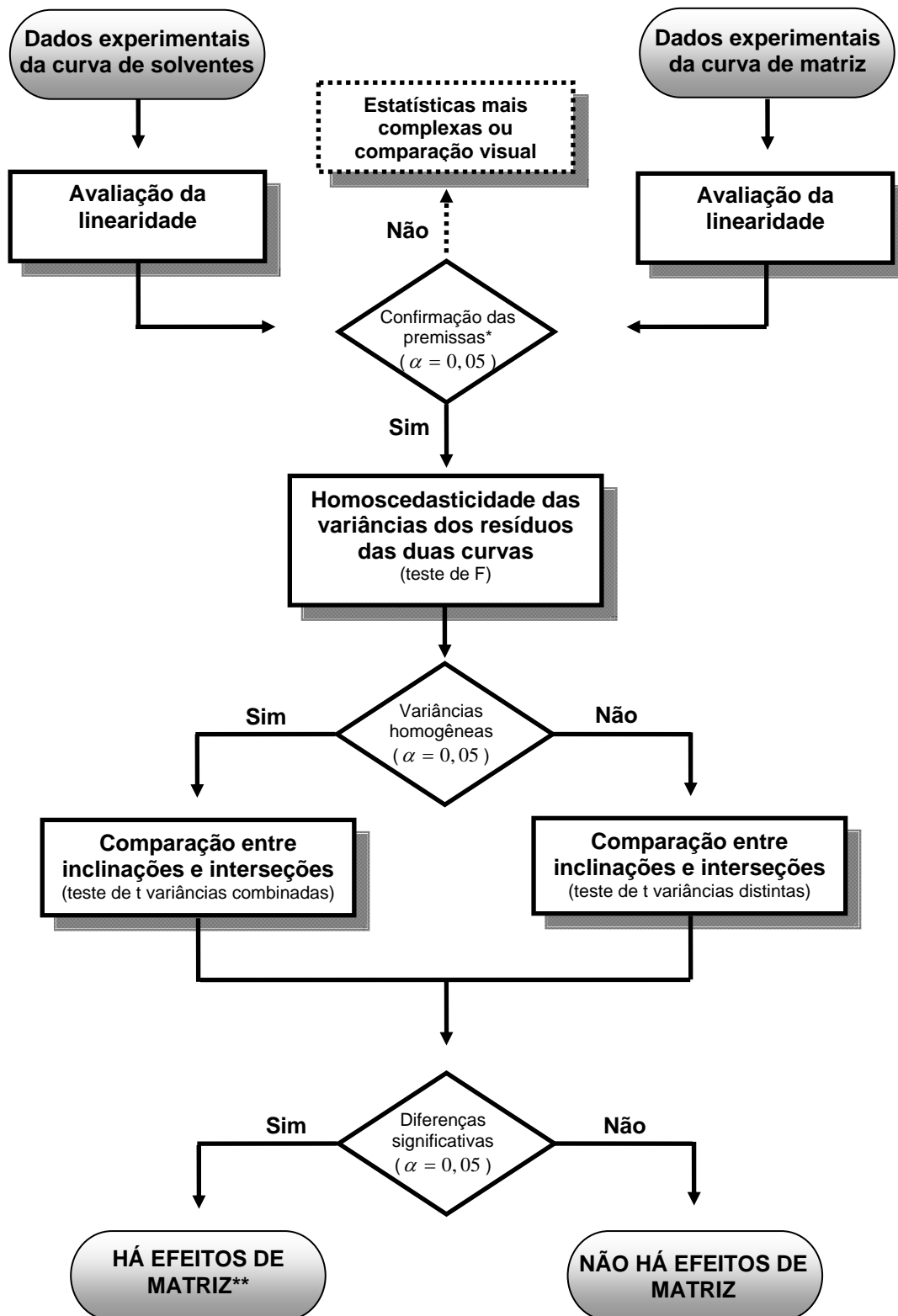
O preparo e leitura das curvas de solventes e matriz é feito no mesmo dia ou em dias diferentes, sob condições emparelhadas. O número de níveis de concentração ou o número de replicatas por nível de concentração pode ser reduzido, em casos excepcionais, para que seja possível realizar o experimento nestas condições.

Para métodos com objetivo de determinação de múltiplos analitos, as condições analíticas e instrumentais para separação, detecção e quantificação destes compostos são previamente ajustadas para que uns analitos não prejudiquem a detecção ou quantificação dos outros. Assim, as soluções padrão podem ser preparadas na forma de *pool*.

#### **Análise de dados**

A segunda etapa da avaliação dos efeitos de matriz inclui a análise de dados e os teste de premissas representados na Figura I.8.

Os dados experimentais obtidos para as curvas do analito em solventes e matriz são analisados pelo MMQO, com verificação das premissas, conforme descrito no item 3.1.1. Uma vez observados normalidade dos resíduos da regressão e ajuste ao modelo linear, as inclinações e interseções estimadas para as duas curvas são comparadas. Se estas premissas não forem confirmadas, estatísticas mais complexas ou uma restrição da análise à comparação visual das curvas são recomendadas.



**Figura I.8 - Avaliação dos efeitos de matriz.**

$\alpha$  = significância.

\* Premissas de normalidade dos resíduos da regressão e ajuste ao modelo linear são requeridas.

\*\* Para inclinações significativamente diferentes, usar curvas de adição em amostras ou em extratos sem amostras nas demais etapas da validação e nas atividades de rotina. Quando as interseções forem significativamente diferentes, estimar a concentração prévia do analito (possibilidade de não ser uma amostra branca verdadeira) ou interferente da matriz para ser descontado nos cálculos posteriores (método de adição).

Pontilhado indica caminhos não considerados pelo procedimento proposto.

Primeiramente, o teste de F (SNEDECOR & COCHRAN, 1989) é utilizado para avaliar a premissa de homogeneidade das variâncias dos resíduos da regressão, confrontando a hipótese nula de que as variâncias dos resíduos não diferem entre si (há homoscedasticidade) contra a hipótese alternativa de que as variâncias dos resíduos são diferentes (há heteroscedasticidade). A estatística do teste é a razão entre as variâncias  $F$ , sendo a maior variância no numerador e a menor no denominador (Equação I.24). A hipótese nula é rejeitada quando a estatística  $F$  calculada for maior que o valor crítico  $F_{(1-\alpha/2; n_1-2; n_2-2)}$  (SNEDECOR & COCHRAN, 1989).

$$F = \frac{s_{res1}^2}{s_{res2}^2} . \quad \text{Equação I.24}$$

Quando o teste indicar homoscedasticidade, as inclinações e interseções das curvas de solventes e de matriz são comparadas pelo teste de t com variâncias combinadas (Equações I.25 e I.26). Este teste confronta a hipótese nula de que as inclinações ou interseções não diferem entre si contra a hipótese alternativa de que as inclinações ou interseções são diferentes. A hipótese nula é rejeitada quando a estatística  $t_b$  ou  $t_a$  calculada for maior que o valor crítico  $t_{(1-\alpha/2; n_1+n_2-4)}$  (ARMITAGE & BERRY, 1994).

$$t_b = \frac{|b_1 - b_2|}{\sqrt{\frac{s_p^2}{Sxx_1} + \frac{s_p^2}{Sxx_2}}} , \quad \text{Equação I.25}$$

$$t_a = \frac{|a_1 - a_2|}{\sqrt{\frac{s_p^2 \sum x_1^2}{n_1 Sxx_1} + \frac{s_p^2 \sum x_2^2}{n_2 Sxx_2}}} , \quad \text{Equação I.26}$$

sendo

índices 1 e 2 diferenciam as curvas do analito em solventes e em matriz,

$$s_p^2 = \frac{(n_1 - 2)s_{res1}^2 + (n_2 - 2)s_{res2}^2}{n_1 + n_2 - 4} , \text{ variância combinada.}$$

Se houver heteroscedasticidade, o teste de t é empregado utilizando as variâncias amostrais de cada curva (Equações I.27 e I.28). As estatísticas do teste são a estatística t para contrastes entre inclinações  $t_b'$  e a estatística t para contrastes entre interseções  $t_a'$ .

$$t_b' = \frac{|b_1 - b_2|}{\sqrt{\frac{s_{res1}^2}{Sxx_1} + \frac{s_{res2}^2}{Sxx_2}}} , \quad \text{Equação I.27}$$

$$t_a' = \frac{|a_1 - a_2|}{\sqrt{\frac{S_{res1}^2 \sum x_1^2}{n_1 S_{xx1}} + \frac{S_{res2}^2 \sum x_2^2}{n_2 S_{xx2}}}}. \quad \text{Equação I.28}$$

A hipótese nula de que as inclinações ou interseções não diferem entre si é confrontada com a alternativa de que as inclinações ou interseções são diferentes, rejeitando-se a hipótese nula quando a estatística  $t_b'$  ou  $t_a'$  calculada for maior que o valor crítico  $t_{(1-\alpha/2; \nu)}$ , sendo  $\nu$  definido pela Equação I.29. Na comparação de médias de duas amostras independentes, as variâncias das médias têm  $n_1 - 1$  e  $n_2 - 1$  graus de liberdade (SNEDECOR & COCHRAN, 1989). Nesta comparação de inclinações e interseções de modelos lineares, os graus de liberdade são  $n_1 - 2$  e  $n_2 - 2$ , respectivamente, correspondendo ao número de pontos da calibração subtraído do número de parâmetros estimados em cada reta.

$$\nu = \frac{\left( \frac{S_{res1}^2}{n_1} + \frac{S_{res2}^2}{n_2} \right)}{\frac{(S_{res1}^2/n_1)^2}{n_1 - 2} + \frac{(S_{res2}^2/n_2)^2}{n_2 - 2}}. \quad \text{Equação I.29}$$

No caso de diferenças significativas, pesquisas complementares podem ser conduzidas comparando-se as inclinações e interseções da curva do analito em solventes com as obtidas de curvas do analito em extratos sem matriz. Isto, visando verificar se os efeitos observados são provenientes da matriz em si ou de materiais, reagentes e soluções utilizados no preparo das amostras.

Quando houver indicações de que as inclinações são significativamente diferentes, curvas de adição em amostras ou em extratos sem amostras são empregadas nas demais etapas da validação e nas atividades de rotina. Se as interseções forem significativamente diferentes, estimar a concentração prévia do analito (possibilidade de não ser uma amostra branca verdadeira) ou do interferente da matriz para uso nos cálculos posteriores (método de adição).

Para métodos com objetivo de determinação de múltiplos analitos ou matrizes, os cálculos são feitos para cada combinação analito e matriz, separadamente.

### 3.1.3 Seletividade, exatidão, precisão e limites

Os delineamentos descritos a seguir permitem avaliar os parâmetros seletividade (experimento complementar ao conduzido para efeitos de matriz), exatidão, precisão e limites. Contudo, duas formas distintas são propostas: i) delineamentos sob condições de

repetitividade; e ii) delineamentos sob condições de repetitividade e reprodutibilidade parcial. A opção pelo delineamento somente sob condições de repetitividade deve ser feita em situações específicas, por exemplo, nos casos de métodos que serão posteriormente submetidos a validações por estudos colaborativos e para os quais não será necessário estimar  $CC\alpha$  e  $CC\beta$ . Isto porque a estimativa destes parâmetros, proposta por este trabalho, considera os resultados estimados para amostras adicionadas sob condições de reprodutibilidade parcial.

## **Delineamento experimental**

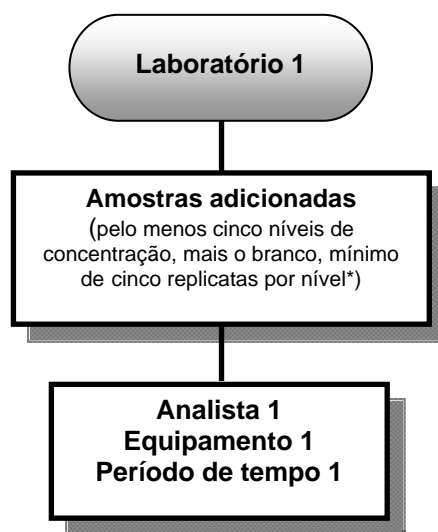
### *Condições de repetitividade*

A primeira etapa do procedimento sugerido para avaliação da seletividade, exatidão, precisão (repetitividade) e limites, para cada combinação de analito e matriz do escopo ou para analitos e matrizes eleitos como representativos, é o delineamento experimental que se segue: i) preparo de amostras brancas e adicionadas do analito ou amostras de concentração conhecida, em pelo menos cinco níveis de concentração, independentes, na faixa em que a linearidade foi demonstrada; ii) preparo de no mínimo cinco replicatas independentes de cada nível, completando um total de dez replicatas nos níveis correspondentes aos limites de detecção e quantificação; iii) análise das replicatas em ordem aleatória, sob condições de repetitividade (mesmo laboratório, mesmo analista, mesmo equipamento, mesmos reagentes e curto período de tempo), conforme indicado na Figura 1.9; e iv) leitura das amostras junto com a curva de calibração (do analito em solventes, em matriz ou extratos sem matriz, dependendo dos resultados dos estudos de efeitos de matriz), em ordem aleatória, com programação do tempo e condições instrumentais entre leituras para evitar contaminação cruzada.

Para métodos com objetivo de determinação de múltiplos analitos, as condições analíticas e instrumentais para separação, detecção e quantificação destes compostos são previamente ajustadas para que uns analitos não prejudiquem a detecção ou quantificação dos outros. Assim, as soluções padrão podem ser preparadas na forma de *pool*.

Objetivando evitar problemas de homogeneidade durante o preparo das amostras adicionadas, as amostras (brancas ou não) são previamente homogeneizadas, fracionadas em porções condizentes com as quantidades descritas pelo método de ensaio e, então, adicionadas de solução padrão do(s) analito(s).

Este procedimento se aplica às situações em que o método sob validação intralaboratorial será posteriormente submetido a estudos colaborativos, nos quais a precisão sob condições de reprodutibilidade será avaliada.



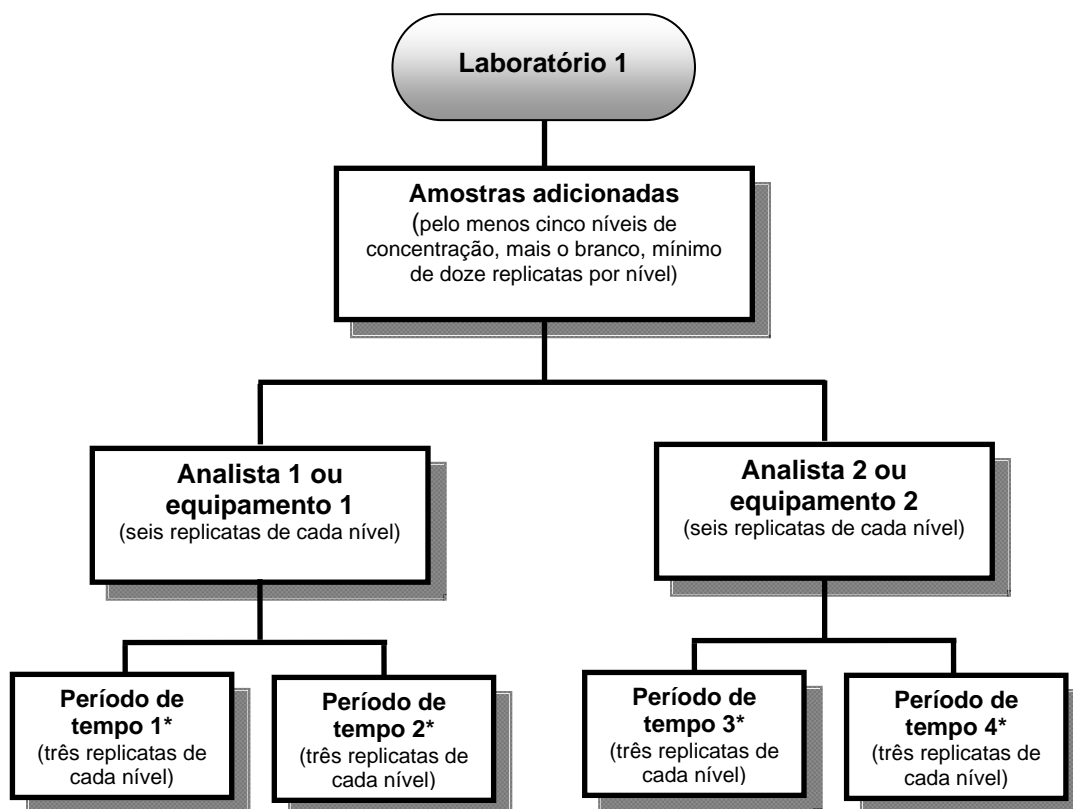
**Figura I.9** - Delineamento experimental para avaliação da precisão sob condições de repetitividade.

\* Completando dez replicatas nos níveis correspondentes aos limites de detecção e quantificação.

#### *Condições de repetitividade e reprodutibilidade parcial*

A primeira etapa do procedimento sugerido para avaliação da seletividade, exatidão, precisão (repetitividade e reprodutibilidade parcial) e limites, para cada combinação de analito e matriz do escopo ou para analitos e matrizes eleitos como representativos, é o delineamento experimental que se segue: i) preparo de amostras brancas e adicionadas do analito ou amostras de concentração conhecida, em pelo menos cinco níveis de concentração, independentes, na faixa em que a linearidade foi demonstrada; ii) preparo de no mínimo doze replicatas independentes de cada nível; iii) análise das replicatas em ordem aleatória, sob condições de repetitividade e reprodutibilidade parcial (mesmo laboratório, diferentes analistas ou diferentes equipamentos dependendo de condições infra-estruturais do laboratório, diferentes períodos de tempo). Um exemplo encontra-se ilustrado na Figura I.10; e iv) leitura das amostras junto com a curva de calibração (do analito em solventes, em matriz ou extratos sem matriz, dependendo dos resultados dos estudos de efeitos de matriz), em ordem aleatória, com programação do tempo e condições instrumentais entre leituras para evitar contaminação cruzada.





**Figura I.10** - Delineamento experimental para avaliação da precisão sob condições de repetitividade e reprodutibilidade parcial.

\* Dentro de cada ensaio, em que são combinados os componentes período de tempo, analistas e equipamentos, as amostras são analisadas sob condições de repetitividade.

Neste delineamento fatorial completamente aninhado (*fully-nested design*), assume-se que cada resultado obtido por um método em particular seja a soma dos seguintes componentes: i) a média global esperada para os resultados; ii) efeitos aleatórios causados pela mudança de analista; iii) efeitos aleatórios causados pela mudança de equipamentos; iv) efeitos aleatórios causados pelo fato das medidas serem feitas em diferentes períodos de tempo; e v) erro aleatório decorrente de medição sob condições de repetitividade. Os componentes considerados em cada ensaio (analista, equipamento e período de tempo) e erro aleatório sob condições de repetitividade são selecionados neste procedimento, uma vez que constituem as principais fontes de contribuição para a variabilidade dos resultados em um laboratório.

Para métodos com objetivo de determinação de múltiplos analitos, as condições analíticas e instrumentais para separação, detecção e quantificação destes compostos são previamente ajustadas para que uns analitos não prejudiquem a detecção ou quantificação dos outros. Assim, as soluções padrão podem ser preparadas em *pool*.

Objetivando evitar problemas de homogeneidade durante o preparo das amostras adicionadas, as amostras (brancas ou não) são previamente homogeneizadas, fracionadas em porções condizentes com as quantidades descritas pelo método de ensaio e, então, adicionadas de solução padrão do(s) analito(s).

Este procedimento se aplica às circunstâncias em que o método sob validação intralaboratorial não será posteriormente submetido a estudos colaborativos, sendo necessário conhecer a precisão do método sob condições de repetitividade e reprodutibilidade parcial.

#### *Ensaio complementares para avaliação da exatidão*

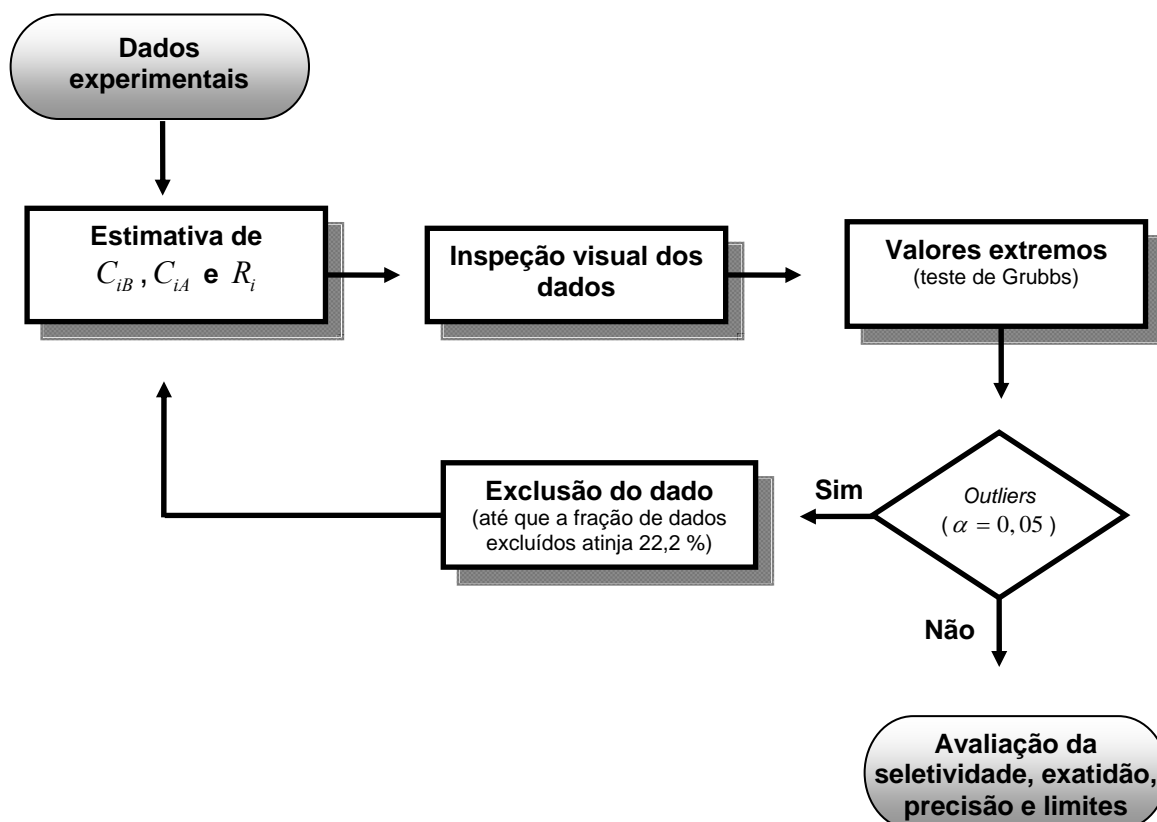
Quando houver materiais de referência certificados da matriz/analito ou de matrizes/analitos representativos do escopo de validação do método, estes materiais são analisados juntamente com as amostras brancas e adicionadas. Nos casos em que tais materiais não são disponíveis, cabe ao laboratório adotar ferramentas que permitam a avaliação da exatidão por meio de referências externas, como por exemplo: participando de ensaios de proficiência ou analisando materiais de referência certificados de matrizes similares.

#### *Ensaio complementares para avaliação da seletividade*

Quando houver conhecimento prévio sobre potenciais interferentes e disponibilidade dos mesmos, estes são adicionados às amostras junto ao analito

#### **Análise de dados**

A segunda etapa do procedimento para avaliação da seletividade, exatidão, precisão e limites inclui: i) estimativa da concentração das amostras; ii) cálculo da recuperação aparente para cada amostra; iii) inspeção visual dos dados; e iv) tratamento de valores extremos pelo teste de Grubbs, como representado na Figura I.11.



**Figura I.11** - Avaliação da seletividade, exatidão, precisão e limites: estimativa da concentração de cada amostra, cálculo da recuperação aparente, inspeção visual dos dados e tratamento de valores extremos.

$C_{iB}$  = concentração estimada para amostras brancas,  $C_{iA}$  = concentração estimada para amostras adicionadas,  $R_i$  = recuperação aparente,  $\alpha$  = significância.

As concentrações do analito nas amostras brancas  $C_{iB}$  e nas amostras adicionadas ou materiais de referência  $C_{iA}$  são estimadas por meio das curvas de calibração do analito em solventes, matriz ou extratos sem matriz (Equação I.30).

$$C_{iA} = C_{iB} = \frac{y_i - a}{b}. \quad \text{Equação I.30}$$

A recuperação aparente  $R_i$  é calculada, para cada amostra adicionada, conforme estabelecido na Equação I.31. Esta equação considera a possibilidade de não haver amostras verdadeiramente brancas na realização dos experimentos.

$$R_i = \frac{(C_{iA} - C_{iB}) \times 100}{C_{iT}}, \quad \text{Equação I.31}$$

sendo

$C_{iT}$  = concentração teórica adicionada.

Após uma análise inicial dos resultados, os valores de recuperação aparente de cada nível de concentração estudado são dispostos em ordem crescente para investigação de possíveis valores extremos pelo teste de Grubbs. As estatísticas  $G_1$  (Equações I.32 e I.33),  $G_2$  (Equação I.34) e  $G_3$  (Equações I.35 e I.36) são computadas simultaneamente para os grupos de resultados de recuperação aparente de cada nível de concentração.

$$G_{1L} = \frac{|\bar{R} - R_1|}{s}, \quad \text{Equação I.32}$$

$$G_{1U} = \frac{|\bar{R} - R_n|}{s}, \quad \text{Equação I.33}$$

$$G_2 = \frac{R_n - R_1}{s}, \quad \text{Equação I.34}$$

$$G_{3L} = 1 - \left( \frac{(n-3)s_{n-2}^2}{(n-1)s^2} \right), \quad \text{Equação I.35}$$

$$G_{3U} = 1 - \left( \frac{(n-3)s_{n-2}^2}{(n-1)s^2} \right), \quad \text{Equação I.36}$$

sendo

$$\bar{R} = \frac{\sum_{i=1}^n R_i}{n}, \text{ média de recuperação aparente, no nível de concentração estudado,}$$

$n$  = número de observações do nível de concentração estudado,

$R_1$  = menor recuperação aparente, no nível de concentração estudado,

$R_n$  = maior recuperação aparente, no nível de concentração estudado,

$s = \sqrt{\frac{\sum (R_i - \bar{R})^2}{n-1}}$ , desvio padrão das recuperações aparentes, no nível de concentração estudado,

$s_{n-2}^2 = \frac{\sum (R_i - \bar{R})^2}{n-2-1}$ , variância das recuperações aparentes, excluindo os dois valores extremos (inferiores e superiores), no nível de concentração estudado,

$s^2 = \frac{\sum (R_i - \bar{R})^2}{n-1}$ , variância das recuperações aparentes, no nível de concentração estudado.

Valores das estatísticas de Grubbs maiores que os valores críticos tabelados (Tabela I.2) são sinalizados como valores extremos. A indicação por, pelo menos, uma

das estatísticas de Grubbs é suficiente para considerar um resultado como valor extremo (GRUBBS, 1969; BARRET & LEWIS, 1994; BURKE 2001 b). Após remoção do valor extremo o teste de Grubbs é sucessivamente aplicado até que novos *valores* não sejam detectados ou até uma exclusão máxima de 22,2 % (HORWITZ, 1995) no número original de resultados.

Para métodos com objetivo de determinação de múltiplos analitos ou matrizes, os cálculos são feitos para cada combinação analito e matriz, separadamente.

**Tabela I.2** - Valores críticos tabelados para as estatísticas de Grubbs em um nível de significância de  $\alpha = 0,05$

n	$G_1$	$G_2$	$G_3$
3	1,153	2,00	-
4	1,463	2,43	0,9992
5	1,672	2,75	0,9817
6	1,822	3,01	0,9436
7	1,938	3,22	0,8980
8	2,032	3,40	0,8522
9	2,110	3,55	0,8091
10	2,176	3,68	0,7695
12	2,285	3,91	0,7004
13	2,331	4,00	0,6705
15	2,409	4,17	0,6182

$n$  = número de observações do nível de concentração estudado,  $G_1$ ,  $G_2$  e  $G_3$  = estatísticas de Grubbs.

### *Seletividade*

Neste experimento complementar ao estudo dos efeitos de matriz, a seletividade é avaliada qualitativamente por inspeção dos resultados calculados bem como dos perfis dos registros de detecção e quantificação do analito (como cromatogramas, espectros, etc.), segundo o esquema apresentado na Figura I.12. O método é considerado seletivo quando: i) o perfil de resposta obtido para o analito na amostra adicionada é indistinguível do perfil obtido para o analito em soluções da curva de calibração; ii) o analito é detectado nas amostras com analito (em níveis acima do que será considerado o limite de detecção do método), com uma relação sinal da amostra adicionada / ruído da amostra branca ou do branco da curva  $\geq 3$ ; iii) o analito não é detectado nas amostras brancas, com relação sinal da amostra branca / ruído do branco da curva  $< 3$ ; e iv) no caso de métodos destinados ao ensaio de mais de um analito, um analito não prejudica a detecção ou

quantificação do(s) outro(s). As evidências dos testes de avaliação da seletividade são mantidas na forma de registros de detecção ou quantificação representativos (como cromatogramas, espectros, etc.).

Assim, as médias dos sinais medidos para o analito nas replicatas de amostras adicionadas de cada nível de concentração  $\bar{y}_A$ , a média dos ruídos ou sinais obtidos para as amostras brancas  $\bar{y}_B$  e a média dos ruídos obtidos para os brancos da curva de calibração  $\bar{y}_{BC}$  e são estimadas (Equações I.37, I.38 e I.39).

$$\bar{y}_A = \frac{\sum_{i=1}^n y_{iA}}{n_A}, \quad \text{Equação I.37}$$

$$\bar{y}_B = \frac{\sum_{i=1}^n y_{iB}}{n_B}, \quad \text{Equação I.38}$$

$$\bar{y}_{BC} = \frac{\sum_{i=1}^n y_{iBC}}{n_{BC}}, \quad \text{Equação I.39}$$

sendo

$y_{iA}$  = sinal medido para amostra adicionada,

$n_A$  = número de amostras adicionadas analisadas,

$y_{iB}$  = ruído ou sinal medido para amostra branca,

$n_B$  = número de amostras brancas analisadas,

$y_{iBC}$  = ruído medido para branco da curva,

$n_{BC}$  = número de brancos da curva.

Os resultados são considerados detectados para amostras adicionadas quando  $\bar{y}_A \geq 3\bar{y}_B$  ou  $\bar{y}_A \geq 3\bar{y}_{BC}$  e não detectados para amostras brancas quando  $\bar{y}_B < 3\bar{y}_{BC}$ .



**Figura I.12** - Avaliação da seletividade: análise das recuperações aparentes, registros de detecção e quantificação e comparação com critérios de aceitabilidade.

$R_i$  = recuperação aparente.

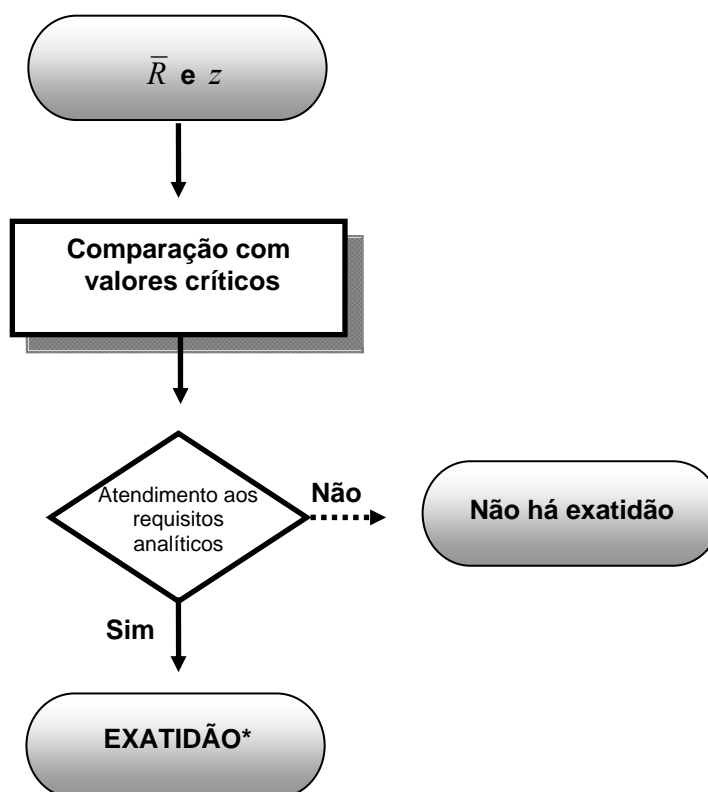
### Exatidão

Exatidão é investigada por meio de: i)  $\bar{R}$ , a média das recuperações aparentes obtidas para as replicatas de amostras adicionadas em cada nível de concentração; e ii)  $z$ , índice z-escore (Equação I.40) obtido na avaliação da concordância entre o valor obtido para a concentração do analito no material de referência  $C_{iA}$  e o valor certificado  $C_{iC}$  (Figura I.13).

$$z = \frac{(C_{iA} - C_{iC})}{s_C}, \quad \text{Equação I.40}$$

sendo

$s_C$  = incerteza do material de referência certificado.



**Figura I.13** - Avaliação da exatidão: estimativa das médias de recuperação aparente e comparação com critérios de aceitabilidade.

$\bar{R}$  = média de recuperação aparente,  $z$  = índice z-escore.

\* Quando não houver material de referência certificado disponível e a exatidão for avaliada somente com base nos resultados de recuperação, a conclusão será “não há falta de exatidão”.

Os resultados médios de recuperação aparente são confrontados com os valores discriminados na Tabela I.3 (EC, 2002) ou com valores regulamentados para casos específicos. Médias de recuperação aparente dentro das faixas estabelecidas indicam que não há falta de exatidão. Exatidão do método é demonstrada por  $z \geq |2|$ .

**Tabela I.3** - Critérios de aceitabilidade para exatidão em função do nível de concentração

Fração de massa	Faixa de $\bar{R}$ aceitável (%)
$c \leq 1,0 \times 10^{-9}$	50 a 120
$1,0 \times 10^{-9} < c < 1,0 \times 10^{-8}$	70 a 110
$c \geq 1,0 \times 10^{-8}$	80 a 110

$\bar{R}$  = média de recuperação aparente,  $c$  = concentração do analito adicionado à matriz expressa como fração de massa (por exemplo  $10^{-9}$  para 1  $\mu\text{g}/\text{kg}$  ou 1  $\mu\text{g}/\text{L}$ ).



### Precisão

O procedimento para avaliação da precisão encontra-se ilustrado na Figura I.14.

Nos experimentos conduzidos sob condições de repetitividade, a precisão é expressa em termos de desvio padrão relativo  $DPR_r$  (Equação I.41) dos resultados de recuperação aparente obtidos para as replicatas de amostras adicionadas em cada nível de concentração.

$$DPR_r = \frac{100 \times s}{\bar{R}}. \quad \text{Equação I.41}$$

Quando os experimentos são delineados sob condições de repetitividade e reprodutibilidade parcial, a precisão é estimada por análise de variância. Este teste requer verificação das premissas relativas aos resíduos das recuperações aparentes: i) normalidade dos resíduos pelo teste de Ryan-Joiner; e ii) homoscedasticidade dos resíduos pelo teste de Brown-Forsythe.

Os resíduos das recuperações aparentes, representados como  $e'_i$ , são estimados para cada nível de concentração conforme estabelecido pela Equação I.42.

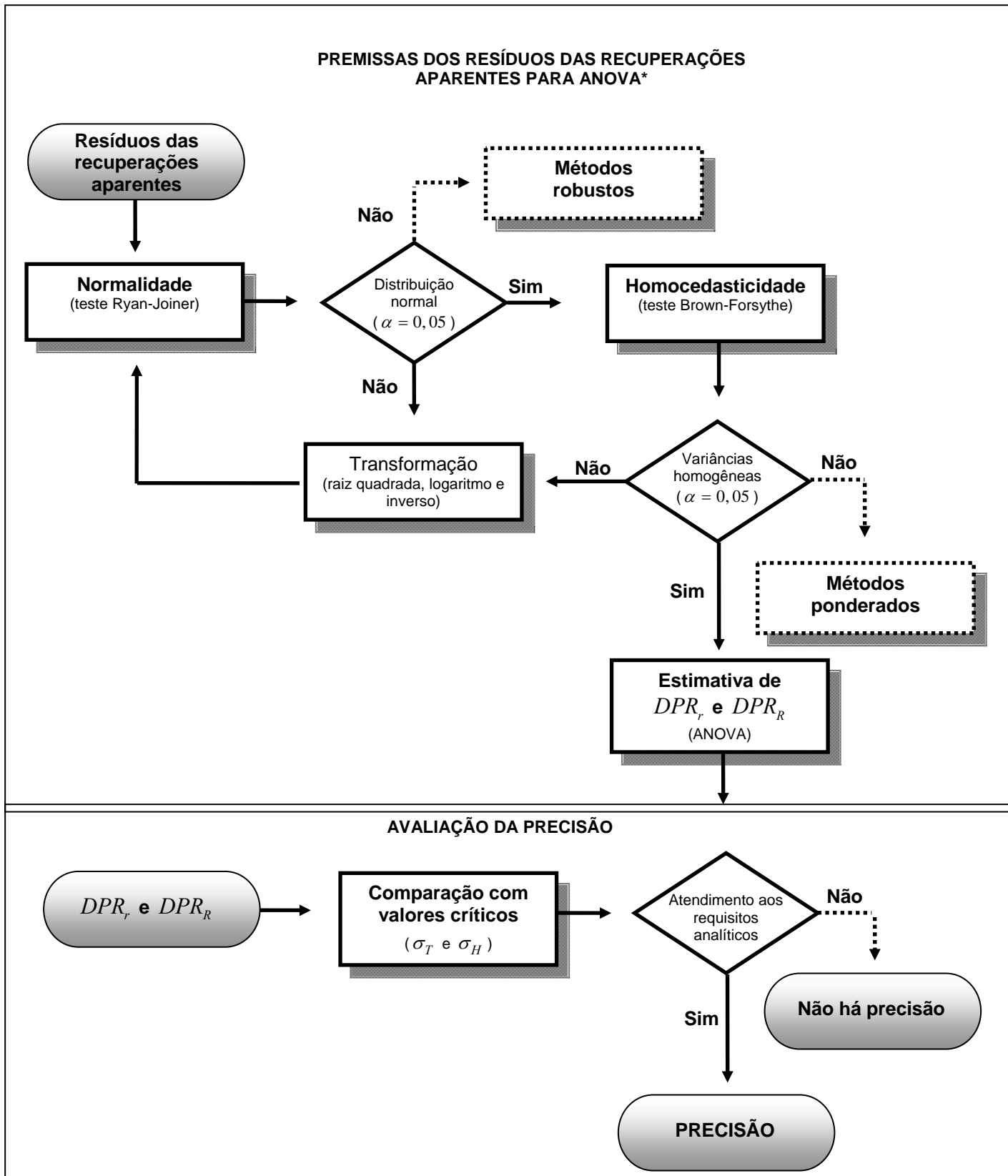
$$e'_i = R_{kj} - \bar{R}_k, \quad \text{Equação I.42}$$

sendo

$R_{kj}$  = j-ésima recuperação aparente do k-ésimo ensaio, no nível de concentração estudado,

$\bar{R}_k$  = média das recuperações aparentes de cada ensaio  $k$ , no nível de concentração estudado.

O teste de normalidade dos resíduos é conduzido exatamente como descrito em 3.1.1 (RYAN & JOINER, 1976). O mesmo não ocorre com o teste de homogeneidade das variâncias, visto que a análise de dados envolve mais de dois grupos, ou seja, tantos grupos quantos forem os ensaios  $k$  das condições de reprodutibilidade parcial estabelecidas no delineamento experimental. Neste caso, a estatística do teste é F de Levene  $F_L$  e não  $t_L$  como determinado em 3.1.1. Homoscedasticidade é verificada confrontando a hipótese nula de que as variâncias dos resíduos das recuperações aparentes não diferem entre si (há homoscedasticidade) contra a hipótese alternativa de que as variâncias dos resíduos são diferentes (há heteroscedasticidade). A hipótese nula é rejeitada quando a estatística  $F_L$  estimada for maior que o valor crítico  $F_{(1-\alpha; u-1; n-u)}$ .



**Figura I.14** - Avaliação da precisão: testes das premissas para emprego da análise de variância, estimativa dos desvios padrão de repetitividade e reprodutibilidade parcial e comparação com critérios de aceitabilidade.

$\alpha$  = significância,  $DPR_r$  = desvio padrão relativo sob condições de repetitividade,  $DPR_R$  = desvio padrão relativo sob condições de reprodutibilidade parcial, ANOVA = análise de variância,  $\sigma_T$  = desvio padrão de reprodutibilidade estimado pela equação de THOMPSON (2000),  $\sigma_H$  = desvio padrão de reprodutibilidade estimado pela equação de HORWITZ (1982).

\* Experimentos conduzidos sob condições de repetitividade e reprodutibilidade parcial. Pontilhado indica caminhos não considerados pelo procedimento proposto.

Os resíduos das recuperações aparentes são divididos em grupos correspondentes aos ensaios  $k$  para cada nível de concentração. As medianas dos resíduos de cada ensaio  $\tilde{e}'_k$  são calculadas. Os módulos das diferenças entre os resíduos e as medianas dos respectivos ensaios  $d'_{kj} = |\tilde{e}'_k - e'_{kj}|$  são obtidos. A estatística  $F_L$  é calculada por análise de variância, conforme indicado na Tabela I.4 (LEVENE, 1960; BROWN & FORSYTHE, 1974), em que a variabilidade total é dividida em soma de quadrados entre ensaios e soma de quadrados dentro dos ensaios. Contudo, não é necessário o uso de todas as equações descritas nesta Tabela. A soma de quadrados dentro dos ensaios é obtida pela diferença entre a soma de quadrados total e a soma de quadrados entre ensaios.

**Tabela I.4** - Análise de variância para estimativa da estatística F de Levene

$FV$	$GL$	$SQ$	$QM$	$F_L$
Entre ensaios	$u - 1$	$\sum_{k=1}^u n_k (\bar{d}'_k - \bar{d}')^2$	$\frac{SQ_{EntreEnsaos}}{GL_{EntreEnsaos}}$	$\frac{QM_{EntreEnsaos}}{QM_{DentroEnsaos}}$
Dentro dos ensaios	$n - u$	$\sum_{k=1}^u \sum_{j=1}^{n_k} (d'_{jk} - \bar{d}'_k)^2$	$\frac{SQ_{DentroEnsaos}}{GL_{DentroEnsaos}} = s_r^2$	
Total	$n - 1$	$\sum_{i=1}^n (d'_i - \bar{d}')^2$	$\frac{SQ_{Total}}{GL_{Total}}$	

$FV$  = fonte de variação,  $GL$  = grau de liberdade,  $SQ$  = soma de quadrados,  $QM$  = quadrado médio,  $F_L$  = estatística F de Levene,  $n$  = número de  $i$  diferenças estimadas,  $u$  = número de ensaios,  $n_k$  = número de diferenças  $j$  em cada ensaio  $k$ ,  $d'_i$  = diferença entre resíduo de recuperação aparente e mediana dos resíduos do respectivo ensaio,  $\bar{d}' =$  média das diferenças,  $d'_{jk} =$   $j$ -ésima diferença do  $k$ -ésimo ensaio,  $\bar{d}'_k =$  média das diferenças do  $k$ -ésimo ensaio.

Confirmadas as premissas, a análise de variância das recuperações aparentes obtidas para cada nível de concentração é conduzida, conforme detalhado na Tabela I.5 (ISO, 1994 a,b,c,d,e; ISO, 1998; KUTTATHARMMAKUL, MASSART & SMEYERS-VERBEKE, 1999). Esta análise fornece as estimativas da variância de repetitividade  $s_r^2$  e da variância do ensaio  $s_{Ensaio}^2$ , que inclui os componentes de variação confundidos (analista + equipamento + dia). Nem todas as equações apresentadas na Tabela I.5 são utilizadas, visto que a soma de quadrados de repetitividade é obtida pela diferença entre a soma de quadrados total e a soma de quadrados do ensaio. Equações para estimativa da variância dos componentes (analista, equipamentos e dia) individualmente não são apresentadas na referida Tabela. Isto porque, uma vez que o número de analistas e

equipamentos dentro de um laboratório é geralmente limitado, espera-se que poucos graus de liberdade estejam associados às estimativas individuais das variâncias destes componentes (KUTTATHARMMAKUL, MASSART & SMEYERS-VERBEKE, 1999).

**Tabela I.5** - Análise de variância para estimativa da precisão sob condições de repetitividade e reprodutibilidade parcial

<i>FV</i>	<i>GL</i>	<i>SQ</i>	<i>QM</i>	<i>F</i>
Ensaio	$u - 1$	$\sum_{k=1}^u n_k (\bar{R}_k - \bar{R})^2$	$\frac{SQ_{Ensaio}}{GL_{Ensaio}} = s_{Ensaio}^2$	$\frac{QM_{Ensaio}}{QM_r}$
Repetitividade	$n - u$	$\sum_{k=1}^u \sum_{j=1}^{n_k} (R_{kj} - \bar{R}_k)^2$	$\frac{SQ_r}{GL_r} = s_r^2$	
Total	$n - 1$	$\sum_{i=1}^n (R_i - \bar{R})^2$	$\frac{SQ_{Total}}{GL_{Total}}$	

*FV* = fonte de variação, *GL* = graus de liberdade, *SQ* = soma de quadrados, *QM* = quadrado médio, *F* = razão entre variâncias, *n* = número de *i* recuperações aparentes estimadas, *u* = número de ensaios, *n<sub>k</sub>* = número de recuperações aparentes *j* em cada ensaio *k*, *R<sub>i</sub>* = recuperação aparente,  $\bar{R}$  = média das recuperações aparentes, *R<sub>kj</sub>* = j-ésima recuperação aparente do k-ésimo ensaio,  $\bar{R}_k$  = média das recuperações aparentes do k-ésimo ensaio.

A partir das estimativas de  $s_r^2$  e  $s_{Ensaio}^2$  é possível calcular a variância entre ensaios  $s_{EntreEnsaio}^2$  e a variância de reprodutibilidade parcial  $s_R^2$  (Tabela I.6) (MAROTO et al, 1999; MAROTO et al, 2001).

**Tabela I.6** - Expressão das variâncias de repetitividade e reprodutibilidade parcial

Variância	Expressão
Repetitividade, $s_r^2$	$QM_r$
Entre ensaios, $s_{EntreEnsaio}^2$	$\frac{QM_{Ensaio} - QM_r}{\bar{n}}$
Reprodutibilidade parcial, $s_R^2$	$s_r^2 + s_{EntreEnsaio}^2$

$s_r^2$  = variância de repetitividade,  $s_{EntreEnsaio}^2$  = variância entre ensaios,  $s_R^2$  = variância de reprodutibilidade parcial,  $QM_r$  = quadrado médio de repetitividade,  $QM_{Ensaio}$  = quadrado médio de ensaio,  $\bar{n}$  = número médio de replicatas estudadas nos diferentes ensaios.

Conforme descrito na Tabela I.6, usa-se o número médio de replicatas estudadas nos *k* ensaios para obter  $s_{EntreEnsaio}^2$ . Quando o número de replicatas for o mesmo nos diferentes ensaios  $\bar{n} = n_k$ . No entanto, quando o número de replicatas não for o mesmo,

quer por problemas operacionais ou pela detecção de valores extremos,  $\bar{n}$  é obtido conforme estabelecido na Equação I.43.

$$\bar{n} = \frac{\sum_{k=1}^u n_k}{u} . \quad \text{Equação I.43}$$

Outra observação importante é que se a estimativa de  $s_r^2$  for maior que  $s_{Ensaio}^2$ , o valor zero é atribuído a  $s_{EntreEnsaio}^2$ , evitando a obtenção de um valor negativo para a estimativa desta variância (ISO, 1994; KUTTATHARMMAKUL, MASSART & SMEYERS-VERBEKE, 1999). Neste caso, as estimativas  $s_r^2$  e de  $s_R^2$  são consideradas equivalentes.

Vale destacar ainda que a estimativa de  $s_R^2$  inclui o cálculo de  $s_{Ensaio}^2$ , o qual é obtido pela soma dos quadrados das diferenças entre as médias de cada dia e a média global. Isto pode subestimar os efeitos de analistas e de equipamentos se estes componentes não forem alterados em cada dia. No entanto, trata-se da melhor forma de estimar a precisão intermediária em situações experimentais de laboratórios, nas quais há uma quantidade restrita de analistas e equipamentos (KUTTATHARMMAKUL, MASSART & SMEYERS-VERBEKE, 1999).

Os desvios padrão relativos sob condições de repetitividade  $DPR_r$  e reprodutibilidade parcial  $DPR_R$ , são então calculados pelas Equações I.44 e I.45, respectivamente.

$$DPR_r = \frac{\sqrt{s_r^2}}{\bar{R}} \times 100 , \quad \text{Equação I.44}$$

$$DPR_R = \frac{\sqrt{s_R^2}}{\bar{R}} \times 100 . \quad \text{Equação I.45}$$

Os resultados de  $DPR_r$  e  $DPR_R$  são comparados com os valores de referência estimados pelas funções de Horwitz ou Thompson, em função do nível de concentração (Tabela I.7) (HORWITZ, 1982; THOMPSON, 2000), ou com valores regulamentados para casos específicos. Na ausência de valores de referência específicos para precisão sob condições de reprodutibilidade parcial, esta precisão é indicada por valores de  $DPR_R$  menores ou iguais aos desvios padrão de reprodutibilidade total estimados pelas equações de HORWITZ (1982) ou THOMPSON (2000). Em condições de repetitividade, são considerados aceitáveis valores de  $DPR_r$  correspondentes a, no máximo, dois terços os valores estimados pela referidas equações (EC, 2002).

**Tabela I.7** - Critérios de aceitabilidade para precisão, sob condições de reprodutibilidade parcial, em função do nível de concentração

Fração de massa	$DPR_R$ máximo aceitável (%)
$c < 1,2 \times 10^{-7}$	$\sigma_T = 0,22c$
$1,2 \times 10^{-7} \leq c \leq 0,138$	$\sigma_H = 0,22c^{0,8495}$
$c > 0,138$	$\sigma_T = 0,01c^{0,5}$

$DPR_R$  = desvio padrão sob condições de reprodutibilidade parcial,  $c$  = concentração teórica adicionada do analito expressa como fração de massa (por exemplo  $10^{-9}$  para 1 µg/kg ou 1 µg/L),  $\sigma_T$  = desvio padrão de reprodutibilidade estimado pela equação de THOMPSON (2000),  $\sigma_H$  = desvio padrão de reprodutibilidade estimado pela equação de HORWITZ (1982).

### Limites

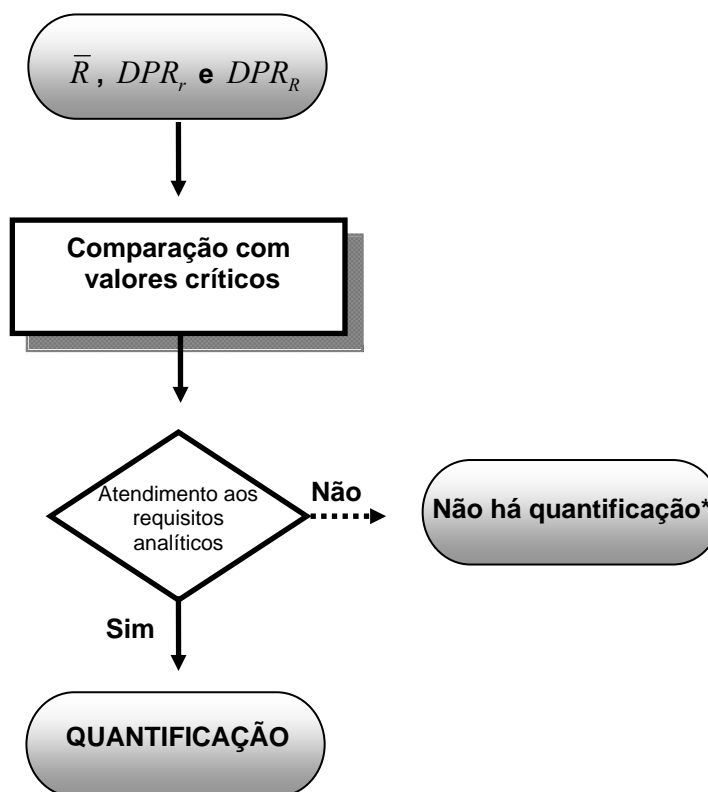
O procedimento para determinação dos limites de detecção e quantificação encontra-se ilustrado nas Figuras I.15 e I.16.

O limite de detecção do método é o menor nível de concentração em que a média dos sinais obtidos para o analito nas replicatas de amostras adicionadas for distinguida de zero com uma relação sinal da amostra adicionada / ruído da amostra branca ou do branco da curva  $\geq 3$ , mas não necessariamente quantificada.

O limite de detecção corresponde, então, ao menor nível de concentração adicionado às amostras no qual  $\bar{y}_A \geq 3\bar{y}_B$  ou  $\bar{y}_A \geq 3\bar{y}_{BC}$ .

O limite de quantificação é estabelecido como o nível de concentração estudado abaixo do qual o método não opera com exatidão e precisão aceitáveis. Este limite é determinado com base nos resultados médios de recuperação aparente e desvios padrão relativos de repetitividade e reprodutibilidade parcial, obtidos nas avaliações de exatidão e precisão, para as replicatas de amostras adicionadas nos diferentes níveis de concentração. Estes resultados são comparados com os critérios de aceitabilidade estabelecidos de forma que o limite de quantificação seja correspondente ao menor nível de concentração adicionado em que a média de recuperação aparente, desvios padrão de repetitividade e de reprodutibilidade parcial estejam dentro dos critérios.

A avaliação destes limites é aplicável para ensaios de elementos traços, resíduos, contaminantes e migrantes de embalagens.



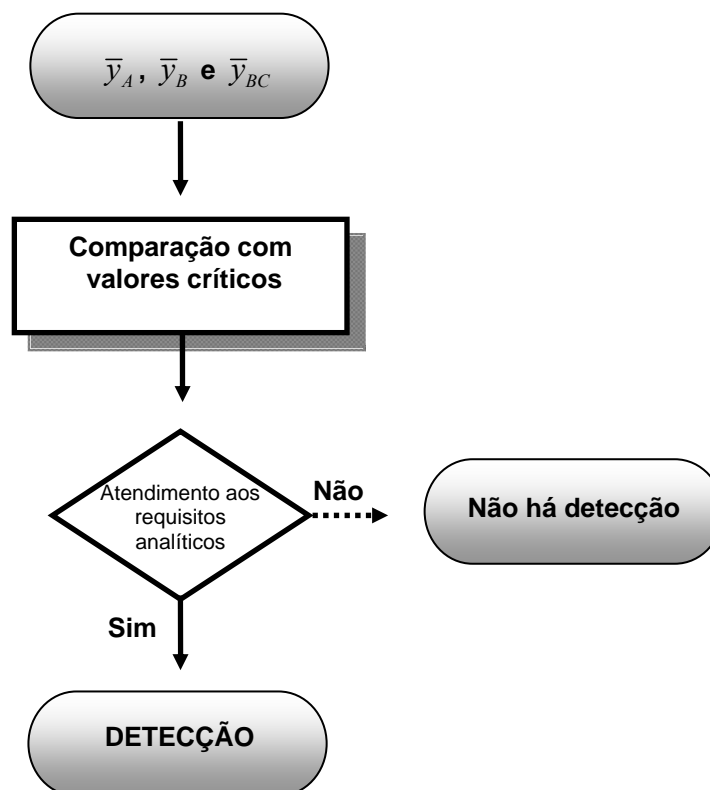
**Figura I.15** - Avaliação do limite de quantificação: estimativa das médias de recuperação aparente, desvio padrão de repetitividade e reprodutibilidade parcial e comparação com critérios de aceitabilidade.

$\bar{R}$  = média de recuperação aparente,  $DPR_r$  = desvio padrão relativo sob condições de repetitividade,  $DPR_R$  = desvio padrão relativo sob condições de reprodutibilidade parcial.

\* Avaliar a capacidade de detecção do analito no nível de concentração considerado.

$CC\alpha$  e  $CC\beta$  são estimados conforme estabelecido pela ISO 11843 (ISO, 1997 b; ISO, 2000) e VAN LOCO & BEERNAERT (2003), utilizando os dados experimentais obtidos nos ensaios de precisão sob condições de repetitividade e reprodutibilidade parcial (EC, 2002).

Os dados experimentais (respostas medidas ou concentrações estimadas para as amostras adicionadas *versus* concentrações teóricas adicionadas) são analisados pelo MMQO, excluído o nível correspondente ao limite de detecção. Como os dados utilizados são provenientes de estudos de precisão realizados em diferentes dias, recomenda-se utilizar as concentrações estimadas nos casos em que não houver comprovação da estabilidade das respostas medidas nos diferentes dias do experimento. Inclinação  $b$ , interseção  $a$  e variância dos resíduos da regressão  $s_{res}^2$  são estimadas conforme descrito em 3.1.1.  $CC\alpha$  é calculado a partir da Equação I.46.



**Figura I.16** - Avaliação do limite de detecção: estimativa das médias dos sinais obtidos para o analito nas amostras adicionadas, nas amostra brancas ou nos brancos da curva e comparação com critérios de aceitabilidade.

$\bar{y}_A$  = média dos sinais medidos para o analito nas replicatas de amostras adicionadas,  $\bar{y}_B$  = média dos ruídos ou sinais obtidos para as amostras brancas,  $\bar{y}_{BC}$  = média dos ruídos obtidos para os brancos da curva.

\* Avaliar a capacidade de detecção do analito no nível de concentração considerado.

$$CC\alpha = \frac{y_c - a}{b},$$

Equação I.46

sendo

$$y_c = a + bx_{MRL} + s_{res} t_{(\alpha; n-2)} \sqrt{1 + \frac{1}{n} + \frac{(x_{MRL} - \bar{x})^2}{\sum (x - \bar{x})^2}}, \text{ valor crítico da variável resposta,}$$

$x_{MRL}$  = concentração nominal no limite máximo de resíduo para analitos permitidos ou zero para analitos proibidos,

$t_{(\alpha; n-2)}$  = estatística t de Student com nível de significância 0,05 para analitos permitidos e 0,01 para analitos proibidos.

$CC\beta$  é estimada posteriormente por resolução da Equação 1.47 de forma iterativa.



$$y_c = a + bCC\beta + s_{res} t_{(\beta;n-2)} \sqrt{1 + \frac{1}{n} + \frac{(CC\beta - \bar{x})^2}{\sum (x - \bar{x})^2}}, \quad \text{Equação I.47}$$

sendo

$t_{(\beta;n-2)}$  = estatística t de Student com nível de significância 0,05 para analitos permitidos e proibidos.

Na rotina analítica, quando são obtidos resultados menores que os valores estabelecidos para  $CC\alpha$ , as amostras são declaradas como conformes ou com analitos em concentrações inferiores ao limite máximo estabelecido, com um nível de confiança de  $(1-\alpha)$ . No caso de resultados maiores ou iguais aos valores estimados para  $CC\beta$ , as amostras são declaradas como não-conformes ou com analito presente em concentrações superiores ao limite máximo, com uma probabilidade de  $(1-\beta)$ . Há, entretanto, uma faixa de incerteza estatística entre os valores de  $CC\alpha$  e  $CC\beta$  (ANTIGNAC et al., 2003). A determinação destes parâmetros é aplicável nos ensaios de contaminantes e resíduos regulamentados em produtos de origem animal. A 95% de confiança, para substâncias com limite máximo de resíduos definido, em  $CC\alpha$  a proporção de resultados falsos não-conformes (erro  $\alpha$ ) será de 5%, mas de falsos conformes (erro  $\beta$ ) será de 50%. Por outro lado, em  $CC\beta$ , a proporção de amostras falsamente declaradas não-conformes será muito baixa e a proporção de amostras verdadeiramente não-conformes apresentarem um resultado conforme será igual a 5%. Assim,  $CC\beta$  deverá ser empregado se o método aqui validado for utilizado como um método de triagem, concebido para selecionar, dentro de um grande número de amostras, os resultados potencialmente não-conformes e evitar a obtenção de resultados falsos conformes (EC, 2002; VAN OVERMEIRE et al., 2004). Por outro lado, se o método for utilizado com propósitos confirmatórios, nos quais a identificação inequívoca da presença de uma substância é pretendida, tanto  $CC\alpha$  como  $CC\beta$  têm aplicação (EC, 2002) e o método será capaz de identificar os analitos em 95% dos casos em  $CC\beta$  e em 50% de todos dos casos em  $CC\alpha$ .

A avaliação destes parâmetros é aplicável para ensaios de elementos traços, resíduos e contaminantes regulamentados.

## **3.2 MÉTODOS QUALITATIVOS**

### **3.2.1 Identificação do analito**

A primeira etapa do procedimento sugerido para validação de métodos qualitativos, para cada analito e matriz do escopo ou para analitos e matrizes eleitos como representativos, inclui a análise de padrões ou materiais de referência para identificação das propriedades (como características físicas, químicas, morfológicas) que podem ser atribuídas ao analito, visando confirmação de sua identidade.

### **3.2.2 Seletividade e limite de detecção**

#### **Delineamento experimental**

O delineamento experimental do procedimento proposto para avaliação da seletividade e limite de detecção de métodos qualitativos inclui: i) preparo de amostras brancas e adicionadas do analito ou materiais de referência, em pelo menos cinco níveis de concentração, independentes, na faixa de concentração de interesse, considerando a aplicação do método e que as concentrações esperadas nas amostras devam estar próximas do centro desta faixa; ii) preparo de no mínimo três replicatas independentes de cada nível, completando um total de dez replicatas no nível correspondente ao limite de detecção; e iii) análise das replicatas, cegas para o analista, em ordem aleatória, frente a padrões ou materiais de referência do analito.

#### **Seletividade**

A seletividade é avaliada qualitativamente por inspeção dos resultados obtidos bem como dos perfis dos registros de detecção ação do analito (como fotos). O método é considerado seletivo nos seguintes casos: i) o perfil de resposta obtido para o analito na amostra adicionada é indistinguível do perfil obtido para o analito em padrões ou materiais de referência; ii) o analito é detectado nas amostras adicionadas em níveis acima do que será considerado o limite de detecção do método (ausência de resultados falsos negativos); iii) o analito não é detectado nas amostras brancas (ausência de resultados falsos positivos); e iv) no caso de métodos destinados ao ensaio de mais de um analito, um analito não prejudica a detecção do(s) outro(s). As evidências dos testes de avaliação da seletividade são mantidas na forma de registros de detecção representativos.

#### **Limite de detecção**

Uma tabela contendo os números de resultados detectados e não detectados, em função do nível de concentração de analito, é elaborada. Em seguida, a tabela é

analisada para determinação do limiar de perda de confiabilidade do método, ou seja, do ponto em que o método passa a apresentar resultados não detectados para amostras adicionadas. O menor nível de concentração estudado no qual o analito foi detectado para todas as replicatas é considerado como limite de detecção do método.

## 4 CONCLUSÕES

O procedimento proposto para validação intralaboratorial de métodos de ensaio emprega a forma direta de validação, com uso de padrões e materiais de referência, cobrindo os parâmetros de desempenho necessários para avaliação da adequação para uso de métodos quantitativos e qualitativos.

A fundamentação do delineamento experimental e dos testes de hipóteses nas características de cada parâmetro de desempenho e nas particularidades das análises de alimentos proporciona eficiência ao processo de validação, bem como facilidade e praticidade de aplicação.

A estratégia de verificação de premissas básicas antes da aplicação das ferramentas estatísticas utilizadas garante confiabilidade ao procedimento de validação proposto.

O princípio da representatividade em termos de analitos, matrizes e níveis de concentração estudados contribui para a aplicabilidade do procedimento proposto.

A forma de apresentação dos testes estatísticos de modo a serem compreendidos por analistas de alimentos, permite a implementação dos mesmos em planilhas de cálculo comerciais, sem a necessidade de uso de programas estatísticos específicos.

## 5 CONSIDERAÇÕES

A referência da publicação de parte deste trabalho encontra-se descrita abaixo. Nesta referência foram apresentados dados experimentais fornecidos pela Gerência da Rede Laboratorial (GRL) do Instituto Mineiro de Agropecuária (IMA), em Contagem - Minas Gerais, Brasil.

SOUZA, S.V.C.; JUNQUEIRA, R.G. A procedure to assess linearity by ordinary least squares method. *Anal. Chim. Acta*, v. 552, p. 25-35, 2005.

# **CAPÍTULO II - VALIDAÇÃO INTRALABORATORIAL DE MÉTODO QUANTITATIVO PARA DETERMINAÇÃO DE SEMICARBAZIDA EM *BABY FOOD* POR CROMATOGRAFIA LÍQUIDA DE ALTA EFICIÊNCIA ACOPLADA A ESPECTROMETRIA DE MASSAS**

## **RESUMO**

Um procedimento detalhado para validação intralaboratorial de métodos foi desenvolvido, empregando curvas do analito em solventes e nas matrizes, amostras brancas e adicionadas de padrão. Este trabalho representa a aplicação do procedimento proposto à validação intralaboratorial de um método para ensaio de semicarbazida em *baby food* por cromatografia líquida de alta eficiência acoplada a espectrometria de massas. Linearidade foi demonstrada nas faixas de concentração de 0,1 ng/mL a 1,0 ng/mL e de 2 ng/mL a 80 ng/mL. Efeitos de matriz significativos foram observados para purê de maçã e pudim de arroz, enquanto efeitos não significativos foram determinados para *baby food* de refeição de carne, em ambas as faixas estudadas. Na faixa quantificável dos experimentos para avaliação da falta de exatidão e precisão do método, as médias de recuperação aparente variaram de 94,1 % a 107,2%; 96,6 % a 104,1 % e 87,8 % a 106,5 %, com desvios padrão relativos sob condições de repetitividade entre 0,7 % e 7,7 %; 0,2 % e 9,1 % e 0,7 % e 8,1 %, respectivamente para purê de maçã, refeição de carne e pudim de arroz. Os limites de detecção e quantificação experimentais do método foram 0,1 µg/kg e 0,25 µg/kg, respectivamente. As premissas relacionadas aos testes estatísticos adotados pelo procedimento de validação intralaboratorial delineado foram checadas e confirmadas.

Palavras-chave: validação intralaboratorial; semicarbazida; *baby food*; cromatografia líquida de alta eficiência; espectrometria de massas.

## **ABSTRACT**

SINGLE-LABORATORY VALIDATION OF A QUANTITATIVE METHOD FOR DETERMINATION OF SEMICARBAZIDE IN BABY FOOD BY HIGH PERFORMANCE LIQUID CHROMATOGRAPHY TANDEM MASS SPECTROMETRY. A detailed procedure

for single-laboratory method validation using solvent and matrix matched curves, blank and spiked samples was developed. The work reported here represents the application of this purposed approach to single-laboratory validation of a quantitative method for analysis of semicarbazide in baby food by high performance liquid chromatography tandem mass spectrometry. Linearity was demonstrated in the ranges from 0.1 ng/mL to 1.0 ng/mL and from 2 ng/mL to 80 ng/mL. Significant matrix effects were observed for apple and rice-based, while non significant effects were determined for meat-based baby food, in both ranges. In the experiments for trueness and precision assessment, in the quantification range, mean of apparent recoveries ranged from 94.1 % to 107.2%, 96.6 % to 104.1 % and 87.8 % to 106.5 %, with relative standard deviation under repeatability conditions between 0.7 % and 7.7 %, 0.2 % and 9.1 % and 0.7 % and 8.1 %, respectively for apple puree, meat based meal and rice pudding. Experimental limits of detection and quantification were 0.10 µg/kg and 0.25 µg/kg, respectively. The assumptions related to the statistics adopted by the single-laboratory validation procedure were checked and confirmed.

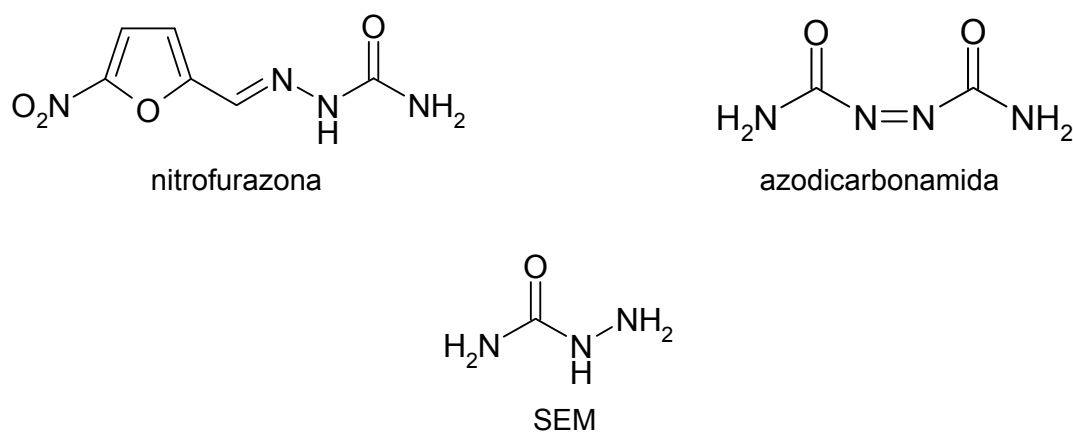
Key words: single-laboratory validation; semicarbazide; baby food; high performance liquid chromatography; mass spectrometry.

## 1 INTRODUÇÃO

Os nitrofuranos, como nitrofurazona, furazolidona, furaltadona e nitrofurantoina, constituem um grupo de antimicrobianos sintéticos, de amplo espectro, com atividade antiprotozoária, que foram amplamente utilizados em avicultura e suinocultura para controle de salmoneloses, colibaciloses e coccidioses. Estas drogas veterinárias também eram utilizadas como aditivos não nutricionais em rações para promoção de crescimento e aumento da eficiência alimentar (RUPP, MUNNS & LONG, 1993; KUMAR, TOOTHILL & HO, 1994). Devido aos potenciais efeitos tóxicos dos nitrofuranos, incluindo mutagenicidade, carcinogenicidade e possibilidade de reações de hipersensibilidade em indivíduos sensíveis, o uso destas substâncias em animais produtores de alimentos foi banido em diversos países (EC, 1995; BRASIL, 2006). Um monitoramento efetivo do uso ilegal dos nitrofuranos foi estabelecido pela pesquisa de resíduos de metabólitos dos nitrofuranos ligados às proteínas, os quais, diferentemente das drogas mães, apresentam estabilidade e podem ser detectados por períodos prolongados de tempo, mesmo após cessado o tratamento (LEITNER, ZÖLLNER & LINDNER, 2001).

A semicarbazida (SEM), um metabólito da nitrofurazona, foi utilizada como indicador do uso ilegal deste nitrofurano em animais criados para produção de alimentos (LEITNER, ZÖLLNER & LINDNER, 2001; CONNEELY et al., 2003). Entretanto, traços de SEM foram detectados em produtos de constituição puramente vegetal, demonstrado que se tratava de um indicador equivocado (EFSA, 2003 a). Pesquisas posteriores comprovaram que a SEM poderia ter sua origem associada a outras fontes, tais como contaminação ambiental e aquelas associadas ao processamento e materiais de embalagens dos alimentos (VAN RHIJN et al., 2003; PEREIRA, DONATO & DE NUCCI, 2004).

Suspeita-se que a origem de SEM nos alimentos pode estar estruturalmente relacionada a azodicarbonamida (Figura II.1), um agente químico de sopro utilizado na produção de selos plásticos para vedação de tampas de potes de vidro (VAN RHIJN et al., 2003; STADLER et al., 2004). Este aditivo é empregado para melhorar as propriedades dos selos plásticos, prevenindo vazamentos e contaminações microbiológicas dos conteúdos dos potes. Em alguns países, como os integrantes do MERCOSUL e Estados Unidos da América, a azodicarbonamida é utilizada como agente melhorador de farinhas (JOINER, VIDAL & MARKS, 1963; PEREIRA, DONATO & DE NUCCI, 2004).



**Figura II.1** - Estruturas químicas da nitrofurazona, azodicarbonamida e SEM.

Embora até o momento poucas pesquisas tenham sido desenvolvidas sobre a ocorrência de SEM em alimentos embalados em potes ou jarras de vidro, níveis de SEM variando da faixa de não detectado (menores que 1 µg/kg) a maiores que 25 µg/kg foram reportados para produtos como *baby food* (alimento para lactentes e crianças de primeira infância), geléias, conservas, pickles, maionese, mostardas, pimentas, temperos e ketchup.

Os *baby food* são os produtos reportados com concentrações mais elevadas de SEM, provavelmente devido a uma maior razão entre a área do selo plástico da tampa e a massa de alimento contida nas pequenas embalagens normalmente utilizadas para acondicionamento destes alimentos (EFSA, 2003 b,c). Atenção especial tem sido dada a estes produtos, não somente devido aos elevados níveis de SEM encontrados, mas também pela maior susceptibilidade dos consumidores aos quais os *baby food* se destinam.

A presença de SEM em alimentos tornou-se um tema de interesse público pelo fato deste composto pertencer à família das hidrazinas, conhecida por causar câncer em animais de laboratório. Existem relatos de evidências de fraca atividade carcinogênica da SEM em fêmeas de ratos, mas não em machos (TOTH & SHIMIZU, 1974; TOTH, SHIMIZU & ERICKSON, 1975). Alguns estudos sugeriram atividade genotóxica da SEM *in vitro*, com ligação a nucleosídeos de citosina de RNA, de deoxicitosina de DNA e à citosina e deoxicitosina (HAYATSU, TAKEISHI & UKITA, 1966; HAYATSU & UKITA, 1966). Entretanto, ABRAMSSON-ZETTERBERG & SVENSSON (2005) não observaram genotoxicidade da SEM *in vivo*, indicando que os efeitos carcinogênicos apresentados anteriormente para ratos não teriam relação com aberrações cromossômicas. HIRAKAWA et al. (2003) demonstraram danos ao DNA causados pela SEM na presença de  $\text{Cu}^{+2}$ . Diante das atuais limitações e contradições dos estudos até então conduzidos, não é possível concluir sobre os riscos que SEM pode causar a humanos (EFSA, 2003 c). Projetos de pesquisa vêm sendo desenvolvidos e incentivados no sentido de se avaliar a real toxicidade deste composto. Desde fevereiro de 2005, amparada pelo princípio da precaução, a União Européia suspendeu o uso de azodicarbonamida como agente de sopro na produção de selos plásticos (EC, 2004; COPPENS, DA SILVA & PETTMAN, 2006).

Considerando a alta reatividade da SEM com compostos carbonílicos e o fato de que estes grupos funcionais estão presentes nos alimentos, parte ou a totalidade da SEM nos alimentos pode estar ligada e não livre. Etapas de hidrólise ácida e derivação (HORNE et al., 1996; LEITNER, ZÖLLNER & LINDNER, 2001; CONNEELY et al., 2003) tornam-se necessárias nos métodos de ensaio de SEM em alimentos para que ambos, resíduos livres e reversivelmente ligados de SEM, sejam determinados. Também, devido a grande variedade de *baby food* disponíveis no mercado, um método de ensaio capaz de determinar SEM em uma grande variedade de matrizes é requerido.

Independentemente do método utilizado, a confiabilidade dos resultados relativos à exposição aos resíduos de SEM, incluindo dados de ocorrência, deve ser garantida por procedimentos de validação de métodos (ISO, 2005 b).

As validações de métodos têm sido reconhecidas como um dos pontos críticos nas cadeias de qualidade laboratoriais (KING, 2003). Processos de validação que fornecem parâmetros de desempenho, indicando adequação para uso de um determinado método, têm dominado as aplicações práticas. Todavia, considerando que estudos de validação são fundamentalmente baseados em testes de hipóteses estatísticas, procedimentos de validação também devem fornecer uma verificação básica de que as premissas feitas com relação aos princípios destes testes não sejam violadas (THOMPSON, ELLISON & WOOD, 2002).

Este trabalho representa a aplicação de um procedimento proposto para validação intralaboratorial de métodos baseado em um delineamento experimental simples e testes estatísticos robustos, incluindo verificações das premissas relacionadas a cada teste. O método validado foi para determinação de SEM em *baby food* por cromatografia líquida de alta eficiência acoplada a espectrometria de massas (CLAE-EM-EM).

## 2 MATERIAL E MÉTODOS

### 2.1 MATERIAL

- Abridor de latas.
- Bastão de vidro.
- Béquer de 50; 100 e 250 mL.
- Bulbo conta-gotas.
- Cartucho de extração em fase sólida de 200 mg/3 mL, de polímero de estireno divinilbenzeno SDB-L de 100 µm, 260 A (Phenomenex).
- Coluna de fase reversa para CLAE de octadecilsilano (C-18), dimensões 50 mm x 2,1 mm, tamanho de partícula 3 µm (Thermo Hypersil - Keystone).
- Espátula de aço inox.
- Filme flexível (Parafilm).
- Filtro de 0,45 µm Minisart RC 15 (Sartorius).
- Frasco âmbar de fundo cônico de 4 mL e 1,1 mL (Chromachol).
- Pérola de vidro.



- Pipeta tipo Pasteur.
- Pote plástico de 100 g com tampa.
- Recipiente de extração em fase sólida com adaptadores para os cartuchos.
- Seringa descartável de 10 mL sem agulha (Terumo).
- Tubo de centrifuga de 50 mL de polipropileno, com tampa de rosca (Falcon).

## 2.2 EQUIPAMENTOS

- Agitador tipo Vortex Mixer (Clifton Cyclone).
- Balança analítica com resolução de 0,0001 g (Sartorius).
- Balança com resolução de 0,01 g (Mettler P 1200).
- Balão volumétrico de 5; 10; 20 e 1000 mL.
- Banho maria com agitação com capacidade para  $40\text{ }^{\circ}\text{C} \pm 3\text{ }^{\circ}\text{C}$  (Clifton Digital Shaking).
- Bomba de vácuo.
- Centrífuga com capacidade de 3900 x g (Juan CR 4 22).
- Chapa aquecedora com capacidade para  $40\text{ }^{\circ}\text{C} \pm 3\text{ }^{\circ}\text{C}$  (Pierce Reacti-Therm III), acoplado a aparato para nitrogênio (Pierce Reacti-Vap III).
- Congelador e refrigerador (Beko Fridge).
- Macropipetador de volume variável de 1 mL a 10 mL (Brand).
- Micropipeta de volume variável de 100  $\mu\text{L}$  a 5000  $\mu\text{L}$  (Biohit E-pet).
- Micropipeta de volume variável de 50  $\mu\text{L}$  a 1000  $\mu\text{L}$  (Biohit E-1000).
- Misturador (Crypto Peerless).
- Proveta de 20 mL.
- Sistema cromatográfico CLAE-EM-EM constituído por um cromatógrafo líquido de alta eficiência Waters 2695 Separations Module (Milford, Massachusetts, EUA), acoplado via uma interface de ionização por eletro spray a um espectrômetro de massas Quattro Ultima Pt (Micromass, Wythenshawe, Reino Unido).

## 2.3 REAGENTES E PADRÕES

Todos dos reagentes utilizados eram de grau de pureza apropriado.

Acetato de etila (grau CLAE), metanol (grau CLAE), água (grau CLAE fluorescência) e ácido clorídrico concentrado 32 % (grau para análise) foram obtidos da Fisher Scientific (Loughborough, Leicestershire, Reino Unido). 2-nitrobenzaldeído (2-NBA) e formato de amônio (grau para análise) foram adquiridos da Sigma-Aldrich Co. Ltd.

(Gillingham, Dorset, Reino Unido). Ortofosfato monoácido de potássio anidro e hidróxido de sódio (grau para análise) foram obtidos da BDH (Poole, Dorset, Reino Unido).

O padrão de cloreto de SEM e o padrão interno de cloreto de 1,2-[<sup>15</sup>N<sub>2</sub>, <sup>13</sup>C] SEM foram fornecidos pela Sigma-Aldrich Co. Ltd. (Gillingham, Dorset, Reino Unido) e Witega (Berlin, Alemanha), respectivamente.

## **2.4 SOLUÇÕES**

### **2.4.1 Soluções padrão de SEM**

A solução estoque de SEM (1 mg/mL de SEM livre) foi preparada por dissolução de 14,9 mg ± 0,1 mg de cloreto de SEM para 10 mL de metanol. A solução intermediária (10 µg/mL) foi obtida por diluição de 100 µL da solução estoque para 10 mL de metanol. Soluções de adição de 1; 20 e 200 ng/mL foram preparadas. A solução de 200 ng/mL foi obtida por diluição de 400 µL da solução intermediária para 20 mL de metanol, enquanto a solução de 20 ng/mL foi preparada por diluição de 500 µL da solução de adição de 200 ng/mL para 5 mL com metanol e a solução de 1 ng/mL por diluição de 500 µL da solução de adição de 20 ng/mL para 10 mL com metanol. As soluções estoque, intermediária e de adição foram armazenadas sob refrigeração entre 2 °C e 8 °C e preparadas em frequências anual, mensal e semanal, respectivamente.

### **2.4.2 Soluções padrão interno de 1,2-[<sup>15</sup>N<sub>2</sub>, <sup>13</sup>C] SEM**

A solução estoque de 1,2-[<sup>15</sup>N<sub>2</sub>, <sup>13</sup>C] SEM (1 mg/mL de SEM livre) foi preparada por dissolução de 14,9 mg ± 0,1 mg de 1,2-[<sup>15</sup>N<sub>2</sub>, <sup>13</sup>C] SEM para 10 mL de metanol. A solução intermediária (10 µg/mL) foi obtida por diluição de 100 µL da solução estoque para 10 mL de metanol, enquanto a solução de adição (100 ng/mL) foi obtida por diluição de 100 µL da solução intermediária para 10 mL de metanol. As soluções estoque, intermediária e de adição foram armazenadas sob refrigeração entre 2 °C e 8 °C e preparadas em frequências anual, mensal e semanal, respectivamente.

### **2.4.3 Solução de derivação: 2-NBA 10 mg/mL**

100 mg de 2-NBA foram dissolvidos para 10 mL de metanol. Esta solução foi preparada a cada uso.

#### **2.4.4 Solução de extração: ácido clorídrico 0,2 mol/L**

Ácido clorídrico concentrado a 32 % (20 mL) foi diluído com 980 mL de água. Esta solução foi preparada em frequência semanal.

#### **2.4.5 Soluções de neutralização: ortofosfato monoácido de potássio 0,2 mol/L e hidróxido de sódio 2 mol/L**

34,88 g de ortofosfato monoácido de potássio e 80,00 g de hidróxido de sódio foram diluídos, separadamente, para 1000 mL de água. Estas soluções foram preparadas em frequência semanal.

#### **2.4.6 Solução de fase móvel: formato de amônio 5 mmol/L**

0,32 g de formato de amônio foram diluídos para 1000 mL de água. Esta solução foi preparada em frequência semanal.

### **2.5 AMOSTRAS**

Amostras brancas de purê de maçã (embalagens plásticas de 120 g), refeição de carne (embalagens de metal de 128 g) e pudim de arroz de três diferentes variedades (embalagens de metal de 425 g) foram adquiridas em supermercados na cidade de York (Yorkshire, Reino Unido). Considera-se amostras brancas aquelas matrizes isentas do analito ou para as quais o analito não tenha sido detectado. Estes três tipos de *baby food* foram escolhidos para os estudos de validação do método como representantes de grupos de produtos ácidos, ricos em lipídeos e proteínas e ricos em carboidratos, respectivamente.

As embalagens das amostras de cada tipo de *baby food* foram abertas e o conteúdo homogeneizado em misturador por 15 h. As amostras homogeneizadas foram acondicionadas em potes plásticos com tampa. Os potes foram devidamente lacrados, identificados e armazenados sob congelamento entre -18 °C e -20 °C, até o momento dos ensaios.

### **2.6 MÉTODO DE ENSAIO**

O método validado foi baseado em procedimentos previamente descritos (GIPS et al., 1993; HOOGENBOOM & POLMAN, 1993; HORNE et al., 1996; LEITNER, ZÖLLNER & LINDNER, 2001).

### 2.6.1 Princípio

O método de ensaio validado fundamentou-se na extração de SEM das amostras de *baby food* com solução de ácido clorídrico 0,2 mol/L e na sua derivação com 2-NBA. O extrato foi neutralizado, purificado por extração em fase sólida. O derivado de 2-NBA foi eluído do com acetato de etila e evaporado até *secura*. O resíduo foi redissolvido em uma mistura de água:metanol para determinação da SEM por CLAE-EM-EM.

### 2.6.2 Cuidados

Devido às características de toxicidade dos reagentes e padrões empregados neste trabalho, equipamentos de proteção individual (avental, óculos, luvas e máscaras) e equipamentos de proteção coletiva (como capelas de exaustão química) foram utilizados durante a realização dos ensaios.

### 2.6.3 Extração

As amostras de *baby food* foram descongeladas até atingirem temperatura ambiente. Alíquotas de 2,00 g  $\pm$  0,03 g foram pesadas em tubos de centrifuga, com auxílio de espátula. 10 mL  $\pm$  5 % da solução de extração, 100  $\mu$ L  $\pm$  2 % da solução de adição do padrão interno 1,2-[<sup>15</sup>N<sub>2</sub>, <sup>13</sup>C] SEM, 240  $\mu$ L  $\pm$  2 % da solução de derivação e 3 a 5 pérolas de vidro foram adicionados a cada tubo. Os tubos foram fechados, homogeneizados em agitador tipo Vortex Mixer durante 30 s e colocados em banho maria a 40 °C  $\pm$  3 °C, sob agitação, durante 15 h. Os extratos foram resfriados a temperatura ambiente. O pH dos extratos foi ajustado com as soluções de neutralização (10 mL  $\pm$  5 % da solução de ortofosfato monoácido de potássio 0,2 mol/L e 800  $\mu$ L  $\pm$  2 % da solução de hidróxido de sódio 2 mol/L). Após homogeneização em agitador tipo Vortex Mixer durante 30 s, os extratos foram centrifugados a 3900 x g durante 15 min. Brancos de solventes, sem adição de amostra, foram preparados em cada bateria analítica.

### 2.6.4 Acondicionamento dos cartuchos de extração em fase sólida

Cartuchos de extração em fase sólida adaptados às seringas descartáveis foram conectados ao recipiente de extração em fase sólida acoplado à bomba de vácuo. Os cartuchos foram lavados com 3 mL  $\pm$  5 % de acetato de etila, sob fluxo da gravidade. Esta limpeza teve caráter preventivo, visando garantir que possíveis compostos interferentes não fossem co-extraídos com o acetato de etila na etapa final de eluição. Em seguida, foram aplicados 3 mL  $\pm$  5 % de metanol e 5 mL  $\pm$  5 % de água, seqüencialmente, também

sob fluxo da gravidade. Uma pequena porção de água (aproximadamente 0,5 mL) foi mantida nos cartuchos até o momento da aplicação dos extratos das amostras.

### 2.6.5 Purificação

Os extratos de amostras e brancos de solventes foram transferidos para as seringas adaptadas aos cartuchos de extração em fase sólida, com auxílio de pipetas tipo Pasteur, deixando gotejar sob fluxo da gravidade. Os cartuchos foram lavados com 5 mL  $\pm$  5 % de água, sob fluxo da gravidade. Em seguida, os cartuchos foram secos sob vácuo por, pelo menos, 2 min. Os derivados 2-NBA de SEM foram eluídos com 3 mL  $\pm$  5 % de acetato de etila para frascos de 4 mL, utilizando vácuo suficiente para um fluxo 1 mL/min. O solvente foi evaporado até *secura* em chapa aquecedora a 40 °C  $\pm$  3 °C, sob fluxo de nitrogênio. Os extratos purificados foram retomados em 1000  $\mu$ L  $\pm$  2 % de água:metanol (60/40, v/v) e filtrados através de filtros de 0,45  $\mu$ m.

### 2.6.6 Separação, detecção e quantificação

As condições de operação do sistema CLAE-EM-EM foram programadas e controladas pelo *software* MassLynx.

As condições da CLAE foram:

- amostrador automático a 15 °C  $\pm$  5 °C;
- volume de injeção de 10  $\mu$ L;
- fase móvel composta por solução de formato de amônio 5 mmol/L e metanol;
- sistema gradiente programado para começar com 20 % de metanol e aumentar para 95 % até 7 min. Esta proporção foi mantida por 1 min e, então, retornou para a condição inicial;
- tempo total de corrida de 16 min;
- fluxo de 0,2 mL/min.

O espectrômetro de massas foi ajustado para os seguintes parâmetros:

- fonte a 120 °C;
- eletro spray capilar a 3 kV;
- temperatura de desolvatação de 450 °C;
- gás de secagem e nebulização, nitrogênio, sob fluxo de 700 L/h e 100 L/h, respectivamente;
- entrada e saída da célula de colisão a 0 eV;
- gás de colisão, argônio, a uma pressão de 10<sup>-4</sup> mbar na célula;

- espectros para SEM e 1,2-[<sup>15</sup>N<sub>2</sub>, <sup>13</sup>C] SEM obtidos na faixa de  $m/z$  50 a 300, no modo EM;
- íons selecionados por monitoramento de reações múltiplas (MRM) (Tabela II.1).

**Tabela II.1** - Condições para MRM

Componente	Íon molecular [M + H] <sup>+</sup> ( $m/z$ )	Íon produto ( $m/z$ )	Energia de colisão (eV)	Tempo de Dwell (s)
SEM	209 ± 0,5	134 ± 0,5	10	0,10
SEM	209 ± 0,5	166 ± 0,5	10	0,10
SEM	209 ± 0,5	192 ± 0,5	10	0,10
1,2-[ <sup>15</sup> N <sub>2</sub> , <sup>13</sup> C] SEM	212 ± 0,5	168 ± 0,5	10	0,10

As razões entre as respostas (áreas dos picos) obtidas para o produto  $m/z$  209 → 166 de SEM e para o produto  $m/z$  212 → 168 de 1,2-[<sup>15</sup>N<sub>2</sub>, <sup>13</sup>C] SEM foram calculadas. Curvas de calibração da razão entre as respostas *versus* a concentração de SEM foram construídas. As equações das curvas foram estimadas pelo método dos mínimos quadrados ordinários (MMQO) e utilizadas para cálculo da concentração de SEM nas amostras.

Para confirmação, as concentrações de SEM calculadas pelos fragmentos de confirmação ( $m/z$  209 → 134 e  $m/z$  209 → 192) foram divididas pela concentração obtida pelo fragmento de quantificação  $m/z$  209 → 166, sendo consideradas aceitáveis razões entre 0,80 e 1,20.

## 2.7 VALIDAÇÃO

Os parâmetros de desempenho do método foram estabelecidos por procedimentos intralaboratoriais com soluções padrão, amostras brancas e amostras a adicionadas de padrão. Linearidade, efeitos de matriz, seletividade, exatidão, precisão, limites de detecção e de quantificação experimentais foram os parâmetros estudados. A adequação do método foi avaliada com base nos resultados dos parâmetros determinados, considerando os critérios de aceitabilidade. Por questões de capacidade operacional, a validação foi realizada em duas baterias analíticas, cobrindo duas faixas de concentração. O nível de significância adotado nos testes de hipóteses foi  $\alpha = 0,05$ . Os cálculos foram feitos em planilhas comerciais, cujos resultados foram previamente validados frente aqueles obtidos em diferentes softwares de análise estatística.

### 2.7.1 Linearidade e efeitos de matriz

Linearidade e efeitos de matriz foram avaliados em ensaios com curvas de calibração preparadas com o analito nos solventes e com o analito nos extratos das matrizes (método de adição).

Seis curvas de calibração (solvente, purê de maçã, refeição de carne, pudim de arroz variedade 1, pudim de arroz variedade 2 e pudim de arroz variedade 3) foram preparadas nos níveis de concentração 2; 10; 20; 40 e 80 ng/mL (correspondentes a 1; 5; 10; 20 e 40 µg/kg de SEM em *baby food*), sendo cada nível em três replicatas independentes. Como o derivado 2-NBA de SEM não encontrava-se disponível no mercado para preparo das curvas de calibração, as mesmas foram preparadas por adição de soluções de SEM em amostras brancas ou brancos de solventes na etapa de extração, conforme descrito na Tabela II.2. Estas soluções foram analisadas em ordem aleatória, no mesmo dia.

**Tabela II.2** - Volumes das soluções de adição de SEM para curvas de 2 ng/mL a 80 ng/mL em solventes e matrizes

Concentração da solução (ng/mL)	Concentração correspondente de SEM em <i>baby food</i> (µg/kg)	Solução de adição de SEM de 20 ng/mL (µL)	Solução de adição de SEM de 200 ng/mL (µL)
0	0	0	0
2	1	100	0
10	5	0	50
20	10	0	100
40	20	0	200
80	40	0	400

Curvas de solventes, purê de maçã, refeição de carne e pudim de arroz nos níveis de 0,1; 0,2; 0,3; 0,4 e 1,0 ng/mL (correspondentes a 0,05; 0,10; 0,15; 0,20 e 0,50 µg/kg de SEM em *baby food*) foram preparadas e analisadas conforme descrito anteriormente (Tabela II.3).

Branco da curva foram preparados como ferramenta de controle de qualidade, avaliação do ruído e ajuste do zero instrumental, mas não foram incluídos nos cálculos.

**Tabela II.3** - Volumes das soluções de adição de SEM para curvas de 0,1 ng/mL a 1,0 ng/mL em solventes e matrizes

Concentração da solução (ng/mL)	Concentração correspondente de SEM em <i>baby food</i> (µg/kg)	Solução de adição de SEM de 1 ng/mL (µL)	Solução de adição de SEM de 20 ng/mL (µL)
0	0	0	0
0,1	0,05	100	0
0,2	0,10	200	0
0,3	0,15	300	0
0,4	0,20	400	0
1,0	0,50	0	50

Após ajuste exploratório dos dados de cada curva pelo MMQO, gráficos de resíduos (erros) da regressão foram examinados para investigação de perfis óbvios, sendo indicados como valores extremos (*outliers*) os pontos fora do intervalo  $\pm t_{(1-\alpha/2; n-2)} s_{res}$ , de maneira semelhante à recomendada por MEYER & ZUND (1993), sendo  $s_{res}$  o desvio padrão dos resíduos da regressão. Os valores extremos foram confirmados e tratados pelo teste dos resíduos padronizados Jackknife, o qual foi aplicado sucessivamente até que novos valores extremos não fossem detectados ou até uma exclusão máxima de 22,2 % no número original de resultados (BELSLEY, KUH & WELSCH, 1980; HORWITZ, 1995). Violações das premissas relacionadas à análise de regressão foram avaliadas: normalidade (RYAN & JOINER, 1976), homoscedasticidade (LEVENE, 1960; BROWN & FORSYTHE, 1974) e independência dos resíduos da regressão (DURBIN & WATSON, 1951). Testes de F foram conduzidos para verificar o ajuste ao modelo linear por meio da avaliação das significâncias da regressão e do desvio da linearidade (DRAPER & SMITH, 1998).

Uma vez comprovada distribuição normal dos resíduos da regressão e linearidade das curvas, a avaliação dos efeitos de matriz foi feita por comparação das inclinações e interseções obtidas para a curva do analito em solventes com as inclinações e interseções estimadas para a curva em cada matriz, nas respectivas faixas de concentração. Teste de F foi utilizado para testar a premissa de homogeneidade das variâncias dos resíduos das curvas (SNEDECOR & COCHRAN, 1989). No caso de resultado não significativo, variâncias homogêneas, as inclinações e interseções foram comparadas pelo teste de t com as variâncias combinadas (ARMITAGE & BERRY, 1994).



### 2.7.2 Seletividade, exatidão, precisão e limites

Seletividade, exatidão, precisão e limites de detecção e quantificação experimentais foram definidos em ensaios com amostras (purê de maçã, refeição de carne e pudim de arroz) brancas e adicionadas nos níveis 0,10; 0,15; 0,20; 0,25; 0,50; 1; 10 e 30  $\mu\text{g}/\text{kg}$ , em três replicatas independentes. Considerando os resultados obtidos nos estudos de efeitos de matriz, curvas de matriz foram preparadas para calcular as concentrações de SEM nas respectivas amostras adicionadas. Concentrações das amostras adicionadas de 0,10  $\mu\text{g}/\text{kg}$  a 0,50  $\mu\text{g}/\text{kg}$  foram estimadas pelas curvas das respectivas matrizes na faixa de 0,1 ng/mL a 1,0 ng/mL, enquanto as concentrações das amostras adicionadas entre 1  $\mu\text{g}/\text{kg}$  e 30  $\mu\text{g}/\text{kg}$  foram calculadas pelas curvas de matrizes na faixa de 2 ng/mL a 80 ng/mL.

Considerando que a técnica detecção e quantificação do método (CLAE-EM-EM) garante uma seletividade raramente comprometida e freqüentemente requerida em situações regulamentares e legais, a ausência de resultados falsos positivos para as amostras brancas, associada a um perfil indistinguível para SEM entre amostras adicionadas e padrões, foram considerados suficientes para determinação da seletividade.

Valores extremos foram investigados nos resultados de recuperação aparente obtidos para cada nível de concentração pelo teste de Grubbs (GRUBBS, 1969; BARRET & LEWIS, 1994; BURKE 2001 b). As estatísticas de Grubbs para um valor extremo ( $G_1$ ), dois valores vicinais ( $G_2$ ) e dois valores polares ( $G_3$ ) foram calculadas simultaneamente. A indicação de valor extremo por pelo menos uma destas estatísticas foi suficiente para remoção do dado. Após eliminação do valor extremo, o teste de Grubbs foi aplicado sucessivamente até que novos valores não fossem detectados ou até uma exclusão máxima de 22,2 % no número original de resultados (HORWITZ, 1995).

Exatidão foi investigada por meio da média de recuperação aparente obtida para as três replicatas de amostras adicionadas em cada nível de concentração. Os critérios adotados para considerar os resultados como satisfatórios foram recuperações de 50 % a 120 %, 70 % a 110 % e 80 % a 110 % para concentrações  $\leq 1 \mu\text{g}/\text{kg}$ ,  $> 1 \mu\text{g}/\text{kg}$  a  $10 \mu\text{g}/\text{kg}$  e  $\geq 10 \mu\text{g}/\text{kg}$ , respectivamente (EC, 2002).

Precisão, sob condições de repetitividade, foi expressa em termos de desvio padrão relativo ( $DPR_r$ ) obtido para as replicatas de amostras adicionadas em cada nível de concentração. Este parâmetro foi considerado aceitável quando dentro de dois terços (EC, 2002) do desvio padrão de reprodutibilidade calculado pela equação de HORWITZ

(1982) modificada por THOMPSON (2000). Para a faixa de concentração estudada, o desvio estimado pela equação de THOMPSON foi 22 %, sendo considerados aceitáveis valores de  $DPR_r \leq 14,7 \%$ .

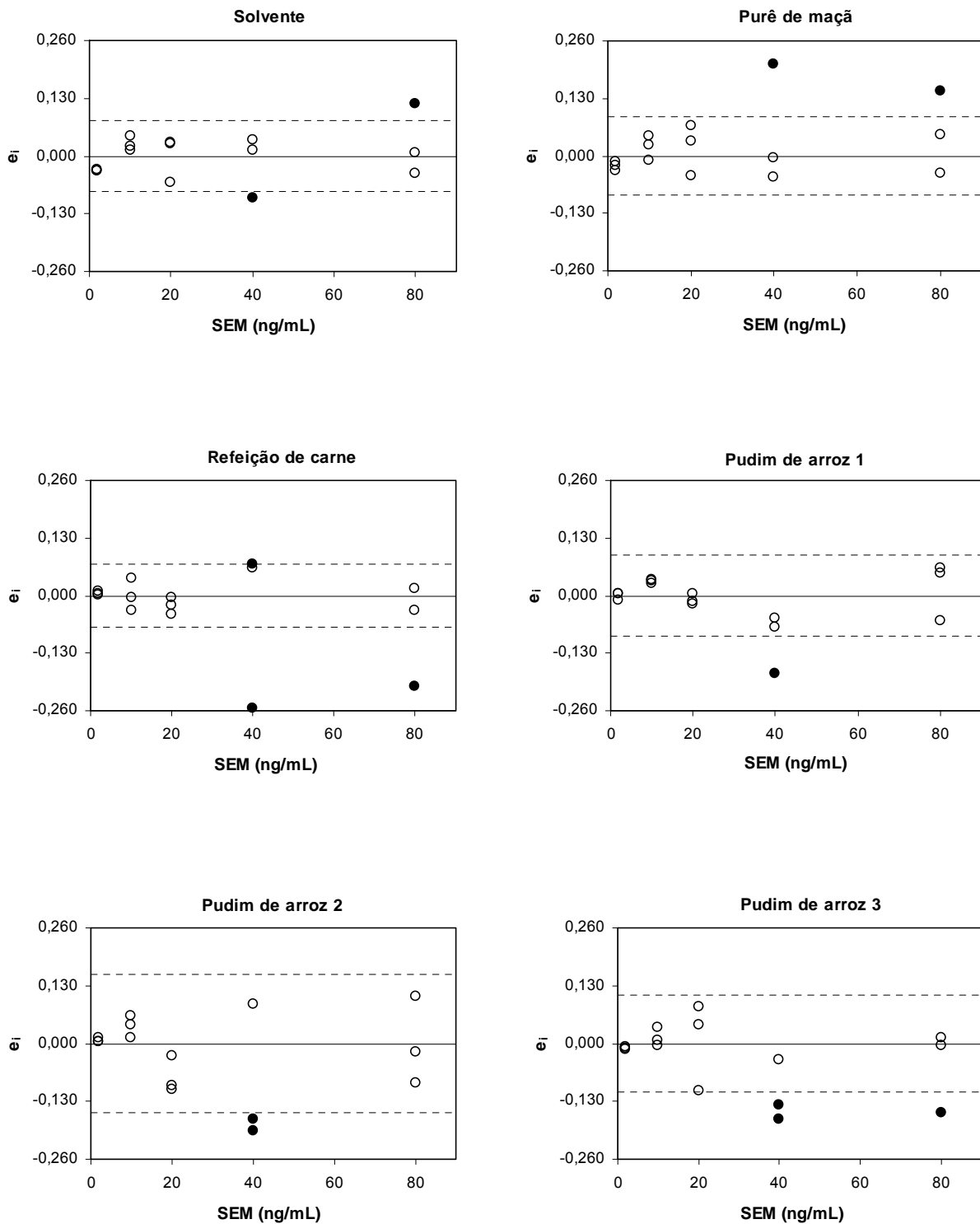
O limite de quantificação foi estabelecido como a concentração abaixo da qual o método não pôde operar com exatidão e precisão aceitáveis. O limite de detecção foi a menor concentração detectada nas amostras adicionadas, distinguida de zero (sinal da amostra adicionada / ruído da amostra branca  $\geq 3$ ), mas não necessariamente quantificada. Estes limites foram determinados com base nos resultados médios de recuperação aparente e desvio padrão relativo obtidos para as replicatas de amostras adicionadas nos diferentes níveis de concentração estudados.

### 3 RESULTADOS E DISCUSÃO

Linearidade foi confirmada para as curvas de solventes e de extratos de diferentes matrizes, nas duas faixas de concentração estudadas.

Todos os valores extremos identificados pelo critério  $\pm t_{(1-\alpha/2; n-2)} S_{res}$  foram confirmados pelo teste de resíduos padronizados Jackknife. Os gráficos dos resíduos da regressão, com indicação dos valores extremos detectados e removidos, estão apresentados na Figura II.2. Dois valores extremos foram detectados nas curvas de solventes e de purê de maçã, nos níveis 40 ng/mL e 80 ng/mL. Três valores extremos foram identificados nas curvas de refeição de carne e pudim de arroz variedade 3, sendo dois no nível 40 ng/mL e um no nível 80 ng/mL. Um valor extremo foi detectado na curva de pudim de arroz variedade 1 e dois na curva de pudim de arroz variedade 2, todos no nível 40 ng/mL. Nas curvas de refeição de carne e pudim de arroz variedade 3 um quarto valor extremo foi indicado pelo teste de resíduos padronizados Jackknife, mas o teste foi interrompido. O número máximo de valores extremos removidos foi três, correspondente ao limite de 22,2 % dos quinze dados originais. Tendências não foram observadas pela inspeção visual dos gráficos exploratórios dos resíduos da regressão, após o tratamento dos valores extremos.

A premissa de que os resíduos da regressão devem seguir a distribuição normal foi confirmada pelo teste de Ryan-Joiner. Os coeficientes de correlação obtidos foram 0,9575; 0,9667; 0,9754; 0,9855; 0,9793 e 0,9486, com valores críticos de 0,9454; 0,9454;



**Figura II.2** - Gráficos exploratórios dos resíduos da regressão das curvas de SEM (2 ng/mL a 80 ng/mL) em solventes e matrizes, com indicação dos respectivos valores extremos diagnosticados pelo teste de resíduos padronizados Jackknife.

$e_i$  = resíduo da regressão, • = valor extremo, - - - - =  $\pm t_{(1-\alpha/2; n-2)} s_{res}$ .

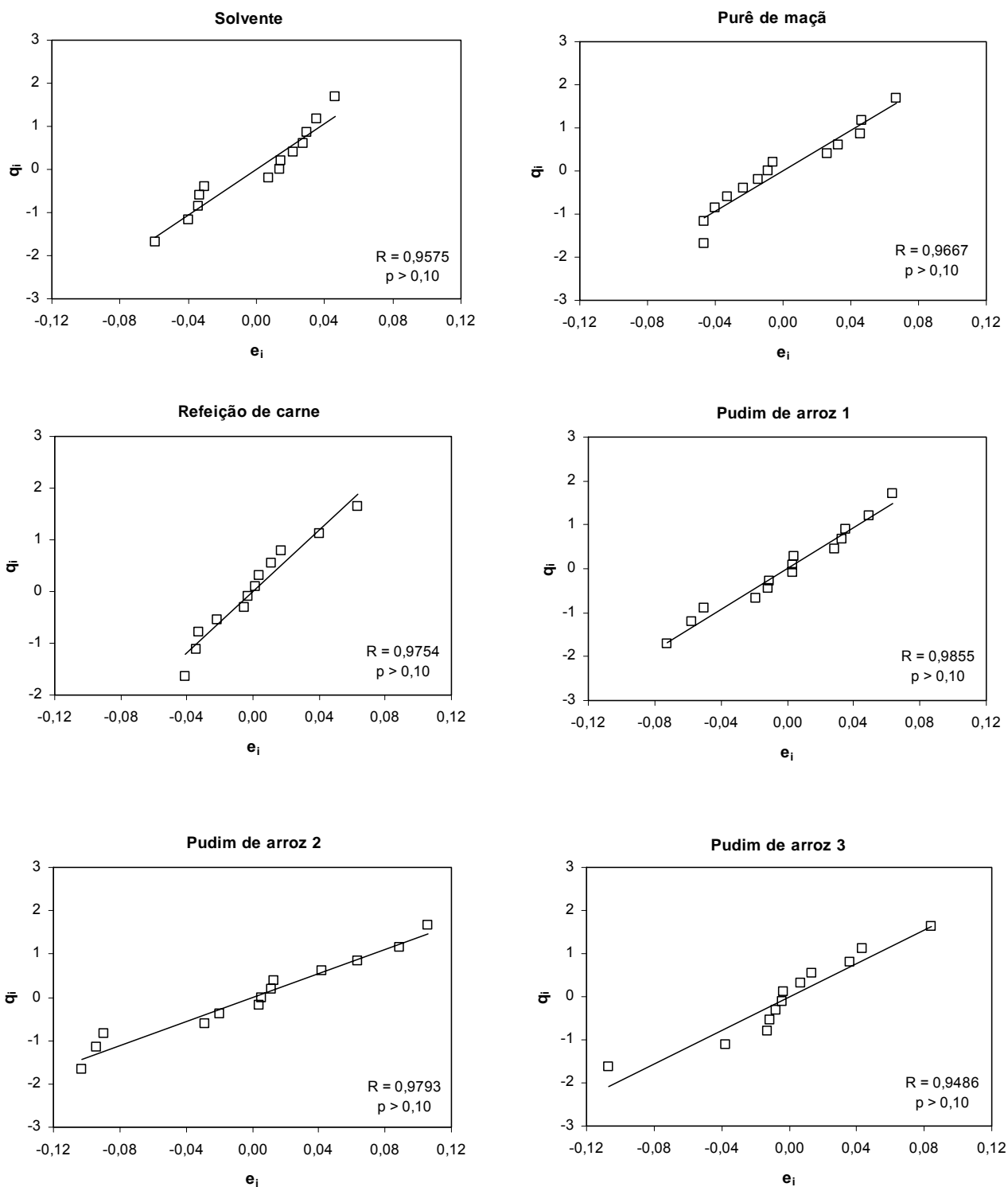
0,9422; 0,9481; 0,9454 e 0,9422 para as curvas de solventes, purê de maçã, refeição de carne, pudim de arroz variedade 1, pudim de arroz variedade 2 e pudim de arroz variedade 3, respectivamente. Estes resultados indicaram que o desvio da normalidade não foi significativo ( $p > 0,10$ ), permitindo o uso dos testes de hipóteses de t e F. Os gráficos de probabilidade normal obtidos para as curvas examinadas encontram-se ilustrados na Figura II.3.

A variabilidade dos resíduos da regressão ao longo dos níveis de concentração estudados foi constante, demonstrando homoscedasticidade. A estatística t de Levene não foi significativa ( $p > 0,05$ ) em todas as curvas examinadas (Tabela II.4). Independência dos resíduos da regressão foi evidenciada pelo teste de Durbin-Watson. O perfil homogêneo de distribuição dos pontos nos quatro quadrantes dos gráficos de Durbin-Watson sugeriu que não houve tendências positivas ou negativas (Figura II.4). As estatísticas calculadas foram 2,041 para a curva de solventes; 2,236 para purê de maçã; 1,689 para refeição de carne; 1,591 para pudim de arroz variedade 1; 2,053 para pudim de arroz variedade 2 e 2,170 para pudim de arroz variedade 3; denotando que não houve autocorrelação ( $p > 0,10$ ). A ratificação destas duas premissas (homoscedasticidade e independência) indicou adequação do uso do MMQO, sem ponderações, para estimativa dos parâmetros da regressão.

Alta significância da regressão ( $p < 0,001$ ) com desvio da linearidade não significativo ( $p > 0,05$ ) foram observados para todas as curvas (Tabela II.5), sugerindo ajuste ao modelo linear na faixa pesquisada de 2 ng/mL a 80 ng/mL. As curvas de calibração com suas respectivas equações e coeficientes de determinação estão demonstrados na Figura II.5.

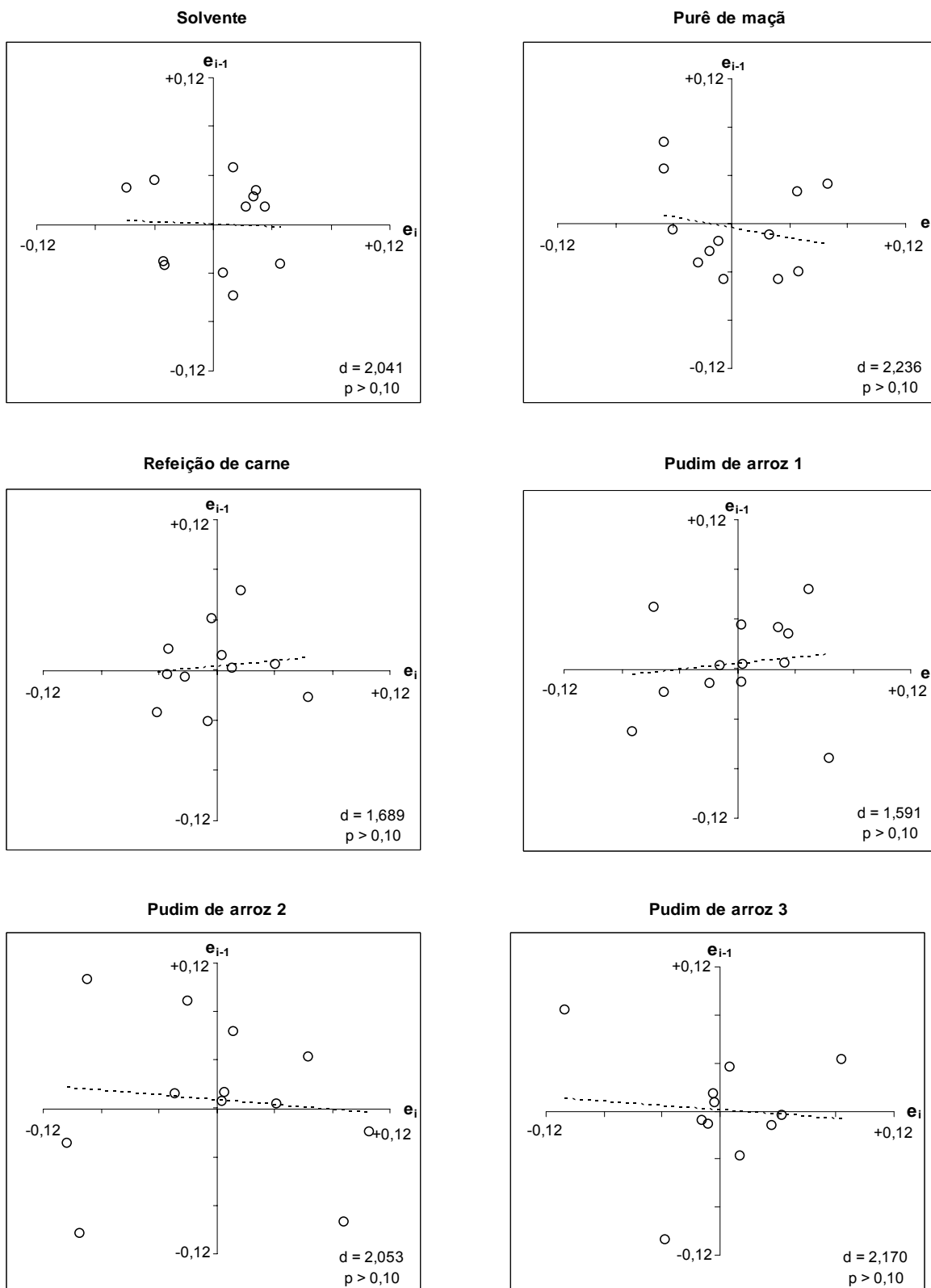
O diagnóstico de valores extremos permitiu tratar pontos que afetavam claramente o ajuste ao modelo. Heteroscedasticidade ( $p > 0,05$ ) e desvio da linearidade ( $p > 0,05$ ) seriam detectados para a curva de pudim de arroz 1 se os valores extremos fossem mantidos.

Todas as premissas relativas ao MMQO foram testadas e confirmadas para as curvas de 0,1 ng/mL a 1 ng/mL, indicando linearidade também nesta faixa. Valores extremos não foram detectados pelo teste de resíduos padronizados Jackknife para as curvas de solventes, purê de maçã e refeição de carne. Um valor extremo foi identificado na curva de pudim de arroz, no nível 0,4 ng/mL. Os coeficientes de correlação de Ryan-Joiner foram 0,9849; 0,9747; 0,9645 e 0,9763 para as curvas de solventes, purê de maçã, refeição de carne e pudim de arroz, respectivamente, demonstrando distribuição normal



**Figura II.3** - Gráficos de probabilidade normal das curvas de SEM (2 ng/mL a 80 ng/mL) em solventes e matrizes.

$e_i$  = resíduo da regressão,  $q_i$  = valor normal esperado,  $R$  = coeficiente de correlação de Ryan-Joiner,  $p$  = significância.



**Figura II.4** - Gráficos de Durbin-Watson das curvas de SEM (2 ng/mL a 80 ng/mL) em solventes e matrizes.

$e_i$  = resíduo da regressão,  $d$  = estatística de Durbin-Watson,  $p$  = significância.

**Tabela II.4** - Representação da homogeneidade das variâncias dos resíduos da regressão das curvas de SEM (2 ng/mL a 80 ng/mL) em solventes e matrizes pelo teste de Levene modificado

Estatística	Solvente	Purê de maçã	Refeição de carne	Pudim de arroz 1	Pudim de arroz 2	Pudim de arroz 3
$n$	13	13	12	14	13	12
$t_L$	0,772	-0,002	-1,120	-1,802	-1,533	0,734
$p$	0,46	1,00	0,29	0,10	0,15	0,48

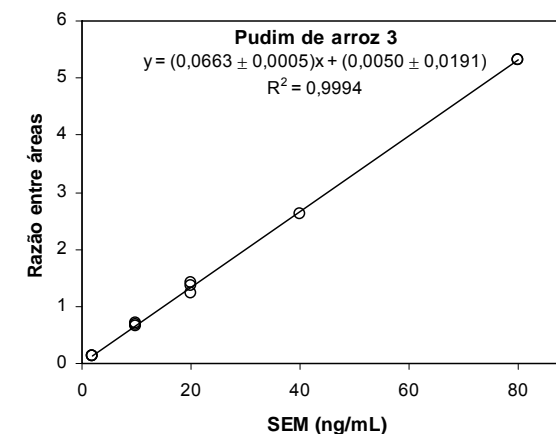
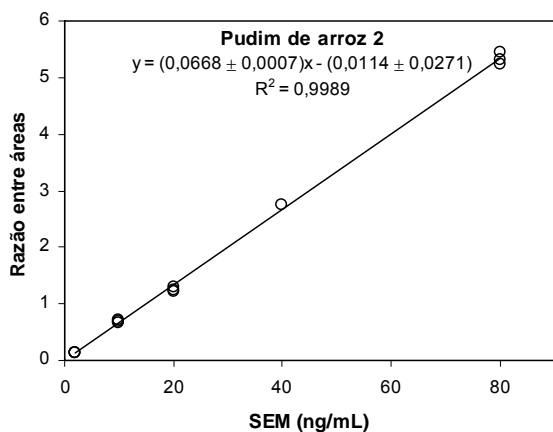
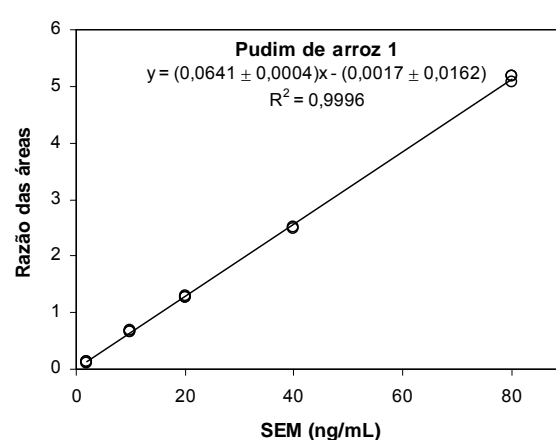
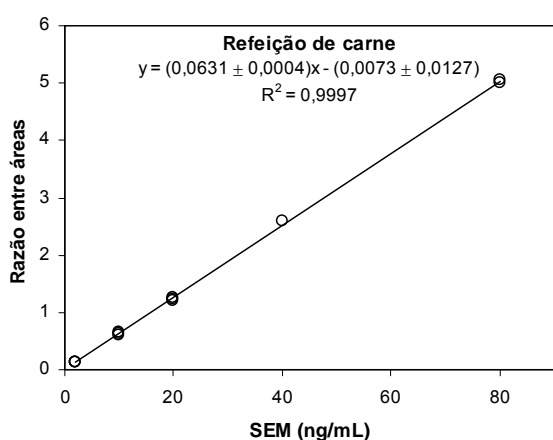
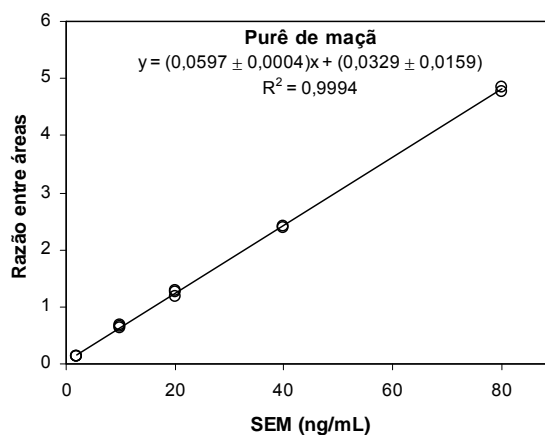
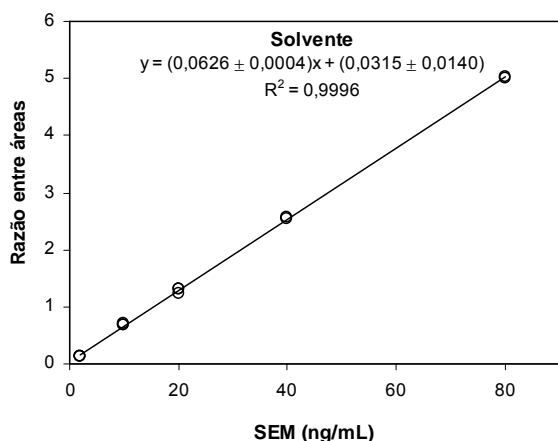
$n$  = número de observações,  $t_L$  = estatística t de Levene,  $p$  = significância.

**Tabela II.5** - Estatísticas da análise de variância da regressão, incluindo teste de desvio da linearidade, para curvas de SEM (2 ng/mL a 80 ng/mL) em solventes e matrizes

Estatística	Solvente	Purê de maçã	Refeição de carne	Pudim de arroz 1	Pudim de arroz 2	Pudim de arroz 3
<i>Regressão</i>						
$F$	$2,684 \times 10^{+4}$	$1,889 \times 10^{+4}$	$3,263 \times 10^{+4}$	$2,668 \times 10^{+4}$	$1,041 \times 10^{+4}$	$1,600 \times 10^{+4}$
$p$	$5,49 \times 10^{-20}$	$3,79 \times 10^{-19}$	$6,64 \times 10^{-19}$	$1,86 \times 10^{-21}$	$1,00 \times 10^{-17}$	$2,34 \times 10^{-17}$
<i>Desvio da linearidade</i>						
$F$	2,714	1,076	2,830	3,811	3,268	0,269
$p$	0,12	0,41	0,12	0,05	0,08	0,85

$F$  = razão entre variâncias,  $p$  = significância.

dos resíduos da regressão ( $p > 0,10$ ). As estatísticas t de Levene foram 0,161; 0,311; -1,169 e -0,998 para as curvas de solventes, purê de maçã, refeição de carne e pudim de arroz, respectivamente, confirmando homoscedasticidade ( $p > 0,05$ ). As estatísticas de Durbin-Watson indicaram independência dos resíduos da regressão para estas curvas, com valores de 2,580 ( $p > 0,10$ ) para a curva de solventes; 2,891 ( $p > 0,02$ ) para purê de maçã; 2,522 ( $p > 0,10$ ) para refeição de carne e 1,970 ( $p > 0,10$ ) para pudim de arroz. Regressões significativas ( $p < 0,001$ ) e desvios da linearidade não significativos ( $p > 0,05$ ) foram observados pelos testes de F (Tabela II.6).



**Figura II.5** - Curvas de calibração do analito nos solventes e nas matrizes obtidas para SEM nos ensaios de avaliação de efeitos de matriz, na faixa de 2 ng/mL a 80 ng/mL, com respectivas equações e coeficientes de determinação.

Razão entre áreas = resposta em (área  $m/z$  209  $\rightarrow$  166 de SEM / área  $m/z$  212  $\rightarrow$  168 de 1,2- $[^{15}\text{N}_2, ^{13}\text{C}]$  SEM),  $y$  = resposta,  $x$  = concentração de SEM em ng/mL,  $R^2$  = coeficiente de determinação.



**Tabela II.6** - Estatísticas da análise de variância da regressão, incluindo teste de desvio de linearidade, para curvas de SEM (0,1 ng/mL a 1,0 ng/mL) em solventes e matrizes

Estadística	Solvente	Purê de maçã	Refeição de carne	Pudim de arroz
<i>Regressão</i>				
<i>F</i>	4,870 x 10 <sup>+3</sup>	3,210 x 10 <sup>+1</sup>	8,140 x 10 <sup>+1</sup>	3,149 x 10 <sup>+1</sup>
<i>p</i>	2,31 x 10 <sup>-23</sup>	7,62 x 10 <sup>-4</sup>	4,20 x 10 <sup>-5</sup>	1,37 x 10 <sup>-3</sup>
<i>Desvio da linearidade</i>				
<i>F</i>	0,635	0,179	0,140	0,081
<i>p</i>	0,60	0,91	0,93	0,97

*F* = razão entre variâncias, *p* = significância.

O teste de t com variâncias combinadas foi adotado para comparar os parâmetros interseção e inclinação nos ensaios de avaliação de efeitos de matriz, uma vez que as variâncias dos resíduos de todos os pares de curvas comparadas foram consideradas homogêneas pelo teste de F ( $p > 0,05$ ). Não foram identificadas diferenças significativas entre o intercepto estimado para a curva de solventes com aqueles estimados para as curvas de matrizes ( $p > 0,05$ ), na faixa de concentração de 2 ng/mL a 80 ng/mL. Na faixa de 0,1 ng/mL a 1,0 ng/mL, diferenças significativas foram observadas entre os interceptos da curva de solventes e os das curvas de purê de maçã e pudim de arroz ( $p < 0,05$ ), embora não tenham sido detectadas diferenças para a curva de refeição de carne ( $p > 0,05$ ). Não houve significância na comparação das inclinações das curvas de refeição de carne com as inclinações das respectivas curvas de solventes ( $p > 0,05$ ). Contudo, diferenças significativas ( $p < 0,05$ ) foram obtidas quando as inclinações das curvas de purê de maçã e pudim de arroz foram comparadas com as obtidas para as curvas de solventes, em ambas as faixas estudadas (Tabelas II.7 e II.8).

Com base nestes resultados foi possível concluir que as curvas de solventes analisadas pelo procedimento descrito neste trabalho forneceram respostas que não diferiram significativamente daquelas obtidas para curvas feitas com matriz de refeição de carne, nos mesmos níveis de concentração. Por outro lado, efeitos de matriz foram evidenciados para curvas de purê de maçã e pudim de arroz, indicando a necessidade do uso de curvas de matriz para estimar concentrações de SEM em produtos com estas composições. Considerando a ampla variedade de *baby food* disponíveis no mercado, foi estipulado como necessário o uso de curvas de matriz ou a condução de testes de efeitos de matriz para avaliar cada caso especificamente.

**Tabela II.7** - Comparações entre as interseções e inclinações das curvas de matriz com as da curva de solventes para SEM (2 ng/mL a 80 ng/mL)

Estatística	Purê de maçã	Refeição de carne	Pudim de arroz 1	Pudim de arroz 2	Pudim de arroz 3
<i>Comparação da interseção com a curva de solvente</i>					
$t_a$	0,068	2,046	1,542	1,400	1,128
$p$	0,95	0,05	0,14	0,18	0,27
<i>Comparação da inclinação com a curva de solvente</i>					
$t_b$	4,951	1,110	2,870	5,414	5,889
$p$	$5,94 \times 10^{-5}$	0,28	$8,65 \times 10^{-3}$	$1,94 \times 10^{-5}$	$7,61 \times 10^{-6}$

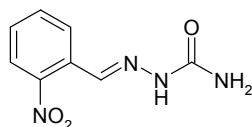
$t_a$  = estatística t para contrastes entre interseções,  $t_b$  = estatística t para contrastes entre inclinações,  $p$  = significância.

**Tabela II.8** - Comparações entre as interseções e inclinações das curvas de matriz com as da curva de solventes para SEM (0,1 ng/mL a 1,0 ng/mL)

Estatística	Purê de maçã	Refeição de carne	Pudim de arroz
<i>Comparação da interseção com a curva de solvente</i>			
$t_a$	7,697	0,026	2,815
$p$	$1,38 \times 10^{-6}$	0,98	0,01
<i>Comparação da inclinação com a curva de solvente</i>			
$t_b$	4,019	0,144	2,299
$p$	$1,12 \times 10^{-3}$	0,89	0,04

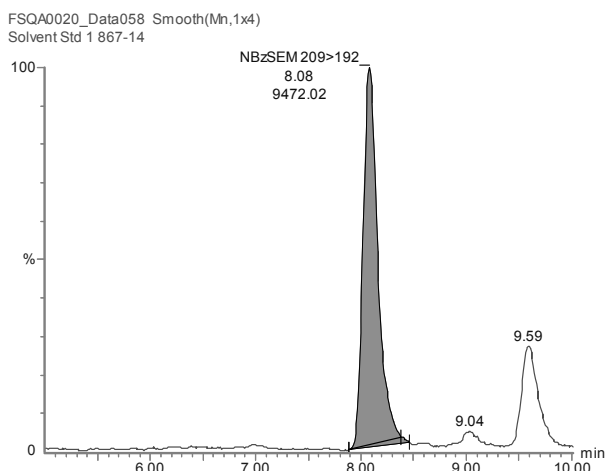
$t_a$  = estatística t para contrastes entre interseções,  $t_b$  = estatística t para contrastes entre inclinações,  $p$  = significância.

A estrutura química do derivado 2-NBA de SEM e alguns cromatogramas típicos obtidos para soluções padrão, amostras de *baby food* brancas e adicionadas estão mostrados na Figura II.6. O tempo de retenção médio para SEM foi 8,08 min  $\pm$  0,03 min. Todas as amostras brancas analisadas apresentaram resultados não detectados (sinal da amostra branca / ruído do branco da curva  $<$  3) para SEM, sem interferentes significativos no tempo de retenção deste analito. Os cromatogramas obtidos para as amostras adicionadas apresentaram-se livres de interferências, com picos de formato agudo e perfis indistinguíveis dos picos de padrões de concentrações correspondentes, indicando seletividade do método.

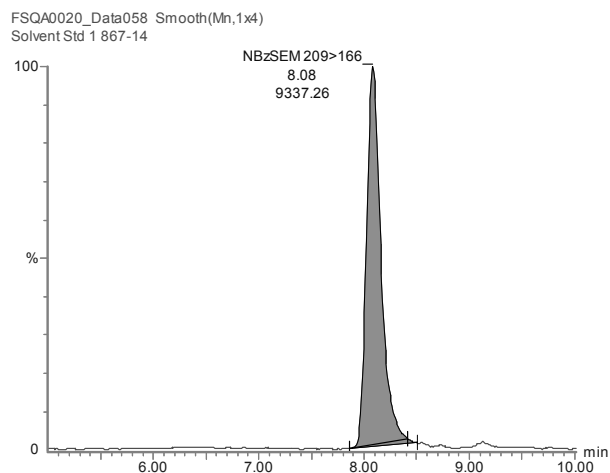


derivado 2-NBA de SEM

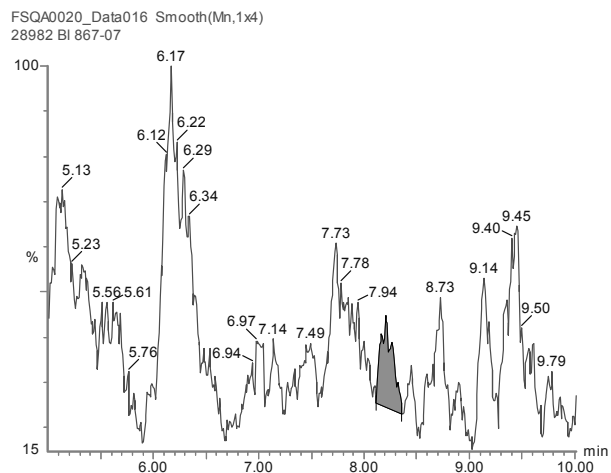
A



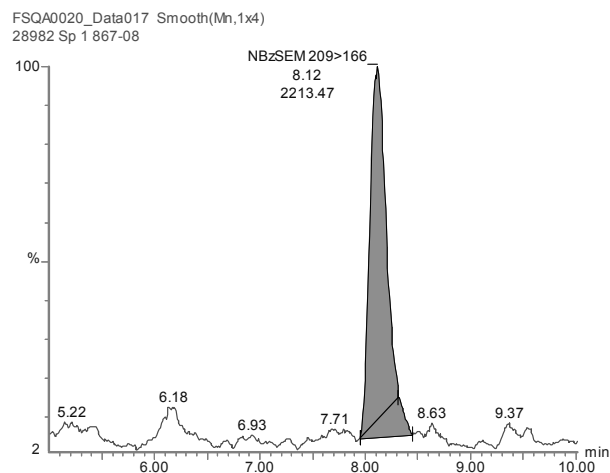
B



C



D



**Figura II.6** - Estrutura química do derivado 2-NBA de SEM e cromatogramas típicos de MRM obtidos por CLAE-EM-EM.

A =  $m/z$  209  $\rightarrow$  192 (confirmação) para solução padrão de SEM de 2 ng/mL.

B =  $m/z$  209  $\rightarrow$  166 (quantificação) para solução padrão de SEM de 2 ng/mL.

C =  $m/z$  209  $\rightarrow$  166 (quantificação) para amostra branca de refeição de carne.

D =  $m/z$  209  $\rightarrow$  166 (quantificação) para amostra de refeição de carne adicionada de 1  $\mu$ g/kg.

Valores extremos não foram detectados pelo teste de Grubbs ( $p > 0,05$ ) para os resultados de recuperação aparente em nenhum dos níveis de concentração estudados. As médias de recuperação aparente variaram de 93,8 % a 107,2 %; de 96,6 % a 104,1 % e de 87,8 % a 112,9 % para amostras de purê de maçã, refeição de carne e pudim de arroz, respectivamente. Os valores  $DPR_r$  estiveram entre 0,7 % e 24,6 % para amostras de purê de maçã; 0,2 % e 9,3 % para amostras de refeição de carne e 0,7 % e 24,7 % para amostras de pudim de arroz (Tabela II.9). Médias de recuperação aparente dentro dos critérios de aceitabilidade foram atingidas para todas as amostras e níveis de concentração avaliados. No entanto,  $DPR_r$  acima do valor aceitável de 14,7 % foram obtidos para amostras de purê de maçã adicionadas de 0,1  $\mu\text{g}/\text{kg}$  e 0,15  $\mu\text{g}/\text{kg}$  e para amostras de pudim de arroz adicionadas de 0,15  $\mu\text{g}/\text{kg}$  e 0,2  $\mu\text{g}/\text{kg}$ . Estes resultados sugeriram que não houve falta de exatidão e adequada precisão do método, sob condições de repetitividade, na faixa de concentração entre 0,25  $\mu\text{g}/\text{kg}$  e 30  $\mu\text{g}/\text{kg}$ , com médias de recuperação aparente variando de 87,8 % a 107,2 % e  $DPR_r$  entre 0,2 % e 9,1 %. Como o método em questão seria avaliado em um estudo colaborativo organizado pela *European Commission, Directorate General for Health and Consumer Protection* (DG SANCO), optou-se por não examinar a reprodutibilidade parcial neste estudo.

**Tabela II.9** - Médias de recuperação aparente e desvios padrão relativos, sob condições de repetitividade, obtidos para amostras de purê de maçã, refeição de carne e pudim de arroz adicionadas de SEM em diferentes níveis de concentração

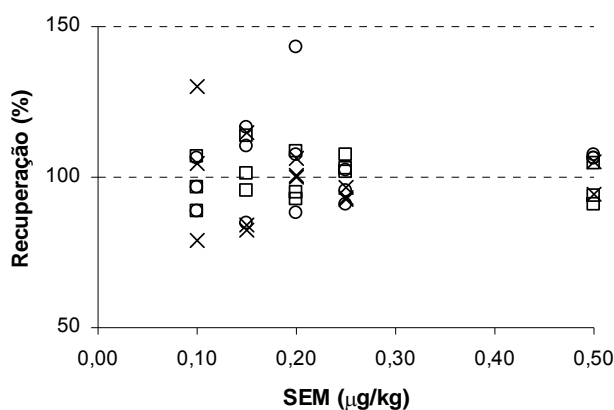
Concentração ( $\mu\text{g}/\text{kg}$ )	Purê de maçã		Refeição de carne		Pudim de arroz	
	$\bar{R}$ (%)	$DPR_r$ (%)	$\bar{R}$ (%)	$DPR_r$ (%)	$\bar{R}$ (%)	$DPR_r$ (%)
0,10	104,6	24,6	97,4	9,3	97,0	9,1
0,15	93,8	19,2	103,5	9,2	104,0	16,2
0,20	102,3	3,5	98,7	8,8	112,9	24,7
0,25	94,1	2,3	104,0	3,0	96,0	5,9
0,50	99,7	7,7	96,6	7,6	106,5	0,7
1,0	102,8	0,7	103,7	0,2	98,9	8,1
10	105,1	2,7	104,1	6,0	102,2	7,1
30	107,2	3,5	101,0	9,1	87,8	5,1

$\bar{R}$  = média de recuperação aparente,  $DPR_r$  = desvio padrão relativo de repetitividade.

Crítérios de aceitabilidade para  $\bar{R}$ : 50 % a 120 % (para amostras adicionadas  $\leq 1 \mu\text{g}/\text{kg}$ ), 70 % a 110 % (para amostras adicionadas  $> 1 \mu\text{g}/\text{kg}$  a  $10 \mu\text{g}/\text{kg}$ ), 80 % a 110 % (para amostras adicionadas  $\geq 10 \mu\text{g}/\text{kg}$ ).

Crítério de aceitabilidade para  $DPR_r$ :  $\leq 14,7$  %.

SEM foi detectada na totalidade das amostras adicionadas de 0,1  $\mu\text{g}/\text{kg}$  e este nível de concentração foi estabelecido como o limite de detecção do método, adotando o critério para detecção de sinal da amostra adicionada / ruído da amostra branca  $\geq 3$ . O limite de quantificação do método foi 0,25  $\mu\text{g}/\text{kg}$ , correspondente ao menor nível de concentração no qual os resultados de exatidão e precisão foram apropriados (média de recuperação entre 50 % e 120 % e  $DPR_r \leq 14,7$  %, respectivamente). O perfil de distribuição dos resultados individuais de recuperação aparente obtidos para amostras adicionadas entre 0,1  $\mu\text{g}/\text{kg}$  e 0,5  $\mu\text{g}/\text{kg}$  (Figura II.7) indicam a transição entre as faixas de detecção e quantificação do método.



**Figura II.7** - Distribuição dos resultados individuais de recuperação aparente obtidos para amostras adicionadas entre 0,10  $\mu\text{g}/\text{kg}$  e 0,50  $\mu\text{g}/\text{kg}$ , incluindo os níveis correspondentes aos limites de detecção (0,10  $\mu\text{g}/\text{kg}$ ) e quantificação (0,25  $\mu\text{g}/\text{kg}$ ).

x = purê de maçã, □ = refeição de carne, o = pudim de arroz.

O método publicado por HORNE et al. (1996) para determinação de metabólitos de furazolidona (AOZ) e furaltadona (AMAZ) em fígado suíno por CLAE-EM-EM apresentou linearidade na faixa de 0,01  $\mu\text{g}/\text{mL}$  a 2,0  $\mu\text{g}/\text{mL}$ , com recuperações  $\geq 70$  %,  $DPR$  intra-ensaios ( $n = 5$ ) variando de 6,2 % a 9,7 % e  $DPR$  inter-ensaios ( $n = 5$ ) entre 12,7 % e 37,0 % para amostras adicionadas de 0; 20 e 100  $\mu\text{g}/\text{kg}$ . Um limite de determinação de 10  $\mu\text{g}/\text{kg}$  foi obtido para os analitos estudados. O método descrito por LEITNER, ZÖLLNER & LINDNER (2001) para determinação de metabólitos de nitrofuranos em tecidos animais demonstrou, para o metabólito SEM, linearidade na faixa de 10  $\mu\text{g}/\text{kg}$  a 800  $\mu\text{g}/\text{kg}$ , limite de detecção de 3  $\mu\text{g}/\text{kg}$  ajustado pela relação sinal/ruído = 3 e limite de

quantificação de 10 µg/kg definido como o limite inferior da faixa linear. Ensaio de eficiência das etapas de derivação e de purificação foram realizados com amostras adicionadas em 10; 50; 100; 250 e 500 µg/kg, em duplicata, apresentando médias de recuperação aparente de 66,3 % ± 5,2 % e 91,8 % ± 6,8 %, respectivamente. Detalhes sobre delineamento experimental, bem como cálculos e critérios de aceitabilidade não foram descritos nestes trabalhos. Além disto, outros parâmetros de desempenho, como por exemplo efeitos de matriz, não foram examinados. Não foram encontrados na literatura relatos de métodos de ensaio para determinação de SEM em *baby food*.

Na validação interlaboratorial por estudo colaborativo organizado pela DG SANCO, materiais de referência de *baby food* de purê de maçã, pudim de arroz e refeição de carne, adicionados de aproximadamente 3; 10 e 30 µg/kg, respectivamente, foram enviados para vinte laboratórios de doze diferentes países europeus, sendo recebidos resultados de dezessete destes participantes. Médias de recuperação aparente entre 88,8 % e 106,1 % foram obtidas, com desvios padrão relativos sob condições de repetitividade e reprodutibilidade entre 4,2 % e 6,9 % e entre 16,6 % e 24,3 %, respectivamente (GINN et al., 2006). Os resultados do estudo colaborativo foram concordantes com os obtidos na validação intralaboratorial do método apresentados neste trabalho.

## 4 CONCLUSÕES

O procedimento utilizado foi adequado para validação intralaboratorial do método de ensaio quantitativo de SEM em *baby food* por CLAE-EM-EM. A aplicação prática do referido procedimento foi simples, tanto no que diz respeito aos experimentos laboratoriais quanto à análise dos dados, a qual pôde ser facilmente conduzida com planilhas de cálculo comerciais, sem necessidade de uso de programas estatísticos específicos.

Uma faixa linear foi estabelecida, com efeitos de matriz significativos para alguns tipos de *baby food*. Seletividade, exatidão e precisão atenderam aos requisitos analíticos estabelecidos, com limites de detecção e quantificação suficientes para determinação de SEM em *baby food*. Os parâmetros de desempenho necessários para avaliação da eficiência do método indicaram adequação para uso em atividades de controle de SEM em *baby food*.

## 5 CONSIDERAÇÕES

Os experimentos descritos neste capítulo foram realizados no *Central Science Laboratory* (CSL) do *Department for Environment Food and Rural Affairs* (DEFRA), em York, Reino Unido.

A referência da publicação deste trabalho encontra-se descrita a seguir:

SOUZA, S.V.C.; JUNQUEIRA, R.G.; GINN, R. Analysis of semicarbazide in baby food by liquid chromatography tandem mass spectrometry (LC-MS-MS) - In-house method validation. *J. Chromatogr. A*, v. 1077, p. 151-158, 2005.

# **CAPÍTULO III - VALIDAÇÃO INTRALABORATORIAL DE MÉTODO QUANTITATIVO PARA DETERMINAÇÃO DE ARSÊNIO EM FÍGADO BOVINO POR ESPECTROFOTOMETRIA DE ABSORÇÃO ATÔMICA COM GERAÇÃO DE HIDRETOS E SISTEMA DE INJEÇÃO EM FLUXO**

## **RESUMO**

Um procedimento detalhado para validação intralaboratorial de métodos foi proposto, incluindo delineamento experimental, cálculo das estatísticas e avaliação das premissas relativas a cada teste estatístico. Este procedimento foi aplicado à validação de um método para determinação de arsênio em fígado bovino por espectrofotometria de absorção atômica com geração de hidretos e sistema de injeção em fluxo. Os parâmetros de desempenho do método foram estabelecidos em ensaios com curvas do analito em solventes e em matriz, amostras brancas e amostras adicionadas. Linearidade, efeitos de matriz, seletividade, exatidão, precisão, limite de detecção e limite de quantificação foram avaliados. Linearidade foi observada entre 0,5 ng/mL e 8 ng/mL (correspondente a 15 µg/kg a 240 µg/kg de arsênio em fígado bovino), sendo que efeitos de matriz significativos não foram observados. As médias de recuperação aparente obtidas para amostras adicionadas em níveis de 30; 60; 120 e 240 µg/kg variaram de 88,9 % a 99,9 %, com desvios padrão relativos sob condições de repetitividade entre 3,7 % e 7,2 % e de reprodutibilidade parcial entre 3,7 % e 7,3 %, indicando que não houve falta de exatidão e adequada precisão do método. Os limites experimentais de detecção e quantificação do método foram de 15 µg/kg e 30 µg/kg, respectivamente.

Palavras-chave: validação intralaboratorial; arsênio; fígado bovino; espectrofotometria de absorção atômica; geração de hidretos.

## **ABSTRACT**

SINGLE-LABORATORY VALIDATION OF A QUANTITATIVE METHOD FOR ARSENIC DETERMINATION IN BOVINE LIVER BY FLOW INJECTION HYDRIDE GENERATION ATOMIC ABSORPTION SPECTROMETRY. A detailed procedure for single-laboratory method validation including experimental design, statistics calculation



and evaluation of assumptions was proposed. This procedure was applied to the validation of a method for arsenic determination in bovine liver by flow injection hydride generation atomic absorption spectrometry. The performance characteristics of the method were established by assays with solvent and matrix matched curves, sample blanks and spiked samples. Linearity, matrix effects, selectivity, trueness, precision, detection and quantification limits were evaluated. Linearity was observed between 0.5 ng/mL and 8 ng/mL (corresponding to 15 µg/kg to 240 µg/kg of arsenic in bovine liver) and no significant matrix effects were observed. Mean of apparent recoveries obtained for spiked samples at 30; 60; 120 and 240 µg/kg ranged from 88.9 % to 99.9 %, with relative standard deviation under repeatability conditions between 3.7 % and 7.2 % and under within-reproducibility conditions between 3.7 % and 7.3 %, indicating no lack of trueness and precision of the method. The experimental limits of detection and quantification were 15 µg/kg and 30 µg/kg, respectively.

Key words: single-laboratory validation; arsenic; bovine liver; atomic absorption spectrometry; hydride generation.

## 1 INTRODUÇÃO

Arsênio é um elemento tóxico e sua toxicidade depende crucialmente da espécie química na qual é encontrado, sendo a forma inorgânica mais tóxica que a orgânica. Pesquisas demonstram que o arsênio inorgânico é um indutor genético caracterizado como agente carcinogênico (FRUMKIN & THUN, 2001; HEI & FILIPIC, 2004). A forma inorgânica trivalente do arsênio reage com grupos sulfidríla de proteínas afetando a atividade de muitas enzimas que contém este grupo. Efeitos inibitórios do  $As^{+3}$  na respiração mitocondrial determinam efeitos tóxicos críticos (WEBB, 1966). A ingestão de altos níveis de arsênio está associada com injúria do fígado, neuropatias periféricas e escurecimento da pele, com o aparecimento de queratoses usualmente nas palmas das mãos e nas solas dos pés, conhecidas como “doença do pé-preto” (CHI & BLACKWELL, 1968; DOULL, KLAASSEN & AMDUR, 1980; MCSHEEHY et al., 2003). Devido a sua similaridade química com o fósforo, o arsênio participa de muitas reações celulares. Compostos organo-arsenicais específicos são formados em alguns organismos e arsênio tem sido reportado como substituto do fósforo em grupos fosfato de DNA (PATRA et al.,

2004). Há controvérsia sobre o caráter oligoelemental do arsênio com funções fisiológicas desconhecidas (FOWLER et al., 1979).

Arsênio pode contaminar animais destinados à produção de alimentos por meio de rações, água e pastagens contaminadas. A contaminação de alimentos de origem animal por arsênio é um assunto de interesse particular, devido ao amplo consumo destes produtos e sua importância nutricional (MCSHEEHY et al., 2003). Muitos países formalizaram regulamentações com respeito a limites de tolerância para arsênio em tecidos de origem animal, criando demandas de métodos de ensaio confiáveis. No Brasil, o monitoramento deste elemento em tecidos de origem animal (fígado, rim e músculo) é regulamentado pelo Plano Nacional de Controle de Resíduos (PNCRB), do Ministério da Agricultura, Pecuária e Abastecimento (MAPA), o qual estabelece 1000 µg/kg como o limite máximo permitido para este contaminante em fígado bovino (BRASIL, 2006). Apesar de ter sido descrito como uma matriz difícil (MINDAK & DOLAN, 1999), o fígado foi o tecido eleito para monitoramentos iniciais pelo PNCRB devido a metilação do arsênio neste órgão durante o processo de metabolização e detoxificação (VAHTER & MARAFANTE, 1988).

Muitos métodos têm sido descritos para determinação de arsênio em tecidos de origem animal. A determinação do conteúdo total de arsênio inorgânico é feita após a conversão de todas as espécies de arsênio na forma trivalente. O pré-redutor mais comum do  $As^{+5}$  a  $As^{+3}$  é o iodeto de potássio, o qual é eficiente em meios fortemente ácidos e utilizado em associação com ácido ascórbico para prevenção de sua oxidação a triiodeto pelo ar (HUNG, NEKRASSOVA & COMPTON, 2004). Tem-se demonstrado que a espectrofotometria de absorção atômica com geração de hidretos (HGAAS) é uma das técnicas mais eficientes e bem estabelecidas para determinação deste analito. HGAAS possui acessórios relativamente baratos, fornece baixos limites de detecção e pode ser facilmente automatizada (TSALEV, 2000). Desde a introdução desta técnica por HOLAK (1969), sistemas de injeção seqüencial e de injeção em fluxo contínuo têm sido freqüentemente utilizados. Sistemas de injeção em fluxo (FIAS) são adequados para análises *on-line*, apresentando um consumo reduzido de reagentes e amostras, além de grande freqüência de amostragem e capacidade de automação (RUZICKA & HANSEN, 1988).

A HG é essencialmente uma separação química do elemento que está sendo analisado da solução teste de matriz, a qual freqüentemente fornece interferentes significativos. Elementos, como arsênio, selênio e antimônio, rendem hidretos com características covalentes mais pronunciadas e relevante estabilidade, sendo gerados e

transportados até a célula de atomização com alta eficiência. Considerando que HG envolve reações químicas *on-line* bastante complexas, altos rendimentos de produtos de reação específicos são requeridos. HG associada à atomização em célula de quartzo, garante limites de detecção menores que aqueles obtidos por atomização em forno de grafite. Devido às condições específicas da HG para analitos individuais, a capacidade desta técnica para análises de vários elementos é limitada. (DEDINA & TSALEV, 1995; TSALEV, 2000).

As diversas etapas envolvidas no processo analítico de determinação de arsênio por HGAAS são passíveis de erros que podem distorcer o resultado do que seria o valor real. Independentemente da técnica empregada, a confiabilidade dos resultados gerados deve ser garantida por procedimentos de validação. A validação de métodos é um componente essencial das medidas que os laboratórios devem adotar e implementar visando a garantia da qualidade dos resultados emitidos (THOMPSON, ELLISON & WOOD, 2002; ISO, 2005 b).

Especificamente, a validação intralaboratorial de métodos é crítica e fundamental para a aplicação de qualquer método, para garantia da viabilidade de um método antes da formalização de um estudo colaborativo e para fornecimento de evidências de que um método de ensaio é confiável quando estudos colaborativos não são disponíveis (THOMPSON, ELLISON & WOOD, 2002). Além disto, a validação intralaboratorial de métodos cobre parâmetros de desempenho não avaliados em estudos interlaboratoriais, como linearidade, efeitos de matriz, seletividade e limites.

Este trabalho apresenta um procedimento detalhado para validação intralaboratorial de métodos baseado em um delineamento experimental simples, na aplicação de testes estatísticos consistentes e na verificação das premissas relacionadas a cada teste. O procedimento proposto foi aplicado na validação de um método clássico para determinação de arsênio em fígado bovino por HGAAS-FIAS.

## **2 MATERIAL E MÉTODOS**

### **2.1 MATERIAL**

- Bandeja de aço inox.
- Bastão de vidro.
- Béquer de 50; 100 e 500 mL.
- Cadinho de vidro borossilicato de 50 mL.

- Espátula de aço inox.
- Faca de aço inox.
- Filme flexível (Parafilm).
- Funil de vidro.
- Papel de filtro qualitativo.
- Pinça de aço inox.
- Recipiente para banho de gelo.
- Saco plástico com dimensões aproximadas de 15 cm x 30 cm e 0,10 mm de espessura.
- Tubo graduado de polipropileno de 15 mL, com tampa de rosca (Falcon).

Antes do uso, todo o material foi lavado com detergente neutro, deixado de molho em solução de ácido clorídrico 3 % por 12 h, enxaguado pelo menos dez vezes com água destilada e desionizada, e seco a temperatura ambiente.

## 2.2 EQUIPAMENTOS

- Agitador tipo Vortex Mixer (Sybron).
- Balança analítica com resolução de 0,0001 g (Mettler AK 160).
- Balança analítica com resolução de 0,01 g (Mettler).
- Balão volumétrico de 25; 100; 250; 500 e 1000 mL.
- Chapa aquecedora com capacidade para 95 °C ± 5 °C (Sybron Thermolyne).
- Congelador (Brastemp).
- Forno mufla com capacidade para 500 °C ± 50 °C (Sybron Thermolyne Furnatrol).
- Macropipetador de volume variável de 1 mL a 10 mL (Boeco).
- Máquina de moer (Arbel).
- Micropipeta automática de volume variável de 100 µL a 1000 µL (Eppendorf).
- Micropipeta automática de volume variável de 10 µL a 100 µL (Eppendorf).
- Pipeta volumétrica de 1; 2 e 5 mL.
- Proveta de 50; 100 e 500 mL.
- Seladora de embalagens plásticas (MSE-Jundiai Mecânica de Sistemas de Embalagens Ltda.).
- Sistema espectrofotométrico HGAAS-FIAS constituído por espectrofotômetro de absorção atômica Perkin Elmer Analyst 100, acoplado a um FIAS Injection Analysis System Perkin Elmer FIAS 400, amostrador automático Autosampler Perkin Elmer AS 90 e lâmpada de descarga elétrica de arsênio operada por meio de uma fonte externa Perkin Elmer EDL System 2 380mA (Norwalk, CT, EUA).

## 2.3 REAGENTES E PADRÕES

Todos dos reagentes utilizados eram de grau de pureza apropriado.

Ácido ascórbico, hidróxido de sódio e iodeto de potássio (grau para análise) foram adquiridos da Vetec (Rio de Janeiro, Brasil). Ácido nítrico 65 % (grau para análise) e borohidreto de sódio (para espectrofotometria de absorção atômica) foram obtidos da Spectrosol BDH Chemicals Ltd. (Poole, Reino Unido). Ácido clorídrico concentrado 37 % e nitrato de magnésio hexahidratado (grau para análise) foram da Biotec (Brasil) e Analar BDH Chemicals Ltd. (Poole, Reino Unido), respectivamente. Água destilada e desionizada foi preparada no laboratório.

A solução padrão estoque de arsênio de 1000  $\mu\text{g/mL}$  (para absorção atômica) foi adquirida da Riedel de Hain Fixanal (Seelze, Alemanha).

## 2.4 SOLUÇÕES

### 2.4.1 Soluções padrão de arsênio

Solução de adição de 0,5  $\mu\text{g/mL}$  foi preparada por diluição de 50  $\mu\text{L}$  da solução estoque para 100 mL com água destilada e desionizada. Esta solução foi armazenada a temperatura ambiente e preparada em frequência semanal.

### 2.4.2 Soluções de ácido clorídrico

Soluções 3 %, 10 %, 44,4 % e 50 % (v/v) foram preparadas diluindo 30 mL, 100 mL, 444 mL e 500 mL de ácido clorídrico para 1000 mL com água destilada e desionizada, respectivamente.

### 2.4.3 Solução de ácido nítrico 50 % (v/v)

500 mL de ácido nítrico foram diluídos para 500 mL com água destilada e desionizada.

### 2.4.4 Solução de borohidreto de sódio 0,2% (m/v)

0,50 g de borohidreto de sódio foram dissolvidos com água destilada e desionizada e adicionados de 1 mL de solução de hidróxido de sódio 12,5 %. A solução foi filtrada através de papel de filtro qualitativo para um balão volumétrico de 250 mL, completando o volume com água destilada e desionizada. Esta solução foi preparada a cada uso.

#### **2.4.5 Solução de hidróxido de sódio 12,5 % (m/v)**

Hidróxido de sódio (12,50 g) foi dissolvido com água destilada e desionizada, em banho de gelo, e o volume completado para 100 mL.

#### **2.4.6 Solução de nitrato de magnésio 50 % (m/v)**

250 g de nitrato de magnésio hexahidratado foram dissolvidos para 500 mL com água destilada e desionizada.

#### **2.4.7 Solução pré-redutora: iodeto de potássio 5 % (m/v) e ácido ascórbico 5 % (m/v)**

1,25 g de iodeto de potássio e 1,25 g de ácido ascórbico foram dissolvidos com água destilada e desionizada para 25 mL. Esta solução foi preparada imediatamente antes de cada uso.

### **2.5 AMOSTRAS**

Amostras brancas de fígado bovino foram adquiridas em açougues de diferentes mercados e supermercados, na cidade de Belo Horizonte (Minas Gerais, Brasil). Considera-se amostras brancas aquelas matrizes isentas do analito ou para as quais o analito não tenha sido detectado.

As amostras foram desembaladas, colocadas em uma bandeja de aço inox e limpas, com auxílio de uma faca, retirando-se películas e veias. Em seguida, as amostras foram cortadas em pedaços menores, os quais foram homogeneizados. Uma pequena porção foi transferida para a máquina de moer, processada e descartada adequadamente. O restante foi então transferido para a máquina de moer, processado e recolhido em sacos plásticos identificados. O ar contido nas embalagens plásticas foi retirado, as embalagens foram devidamente lacradas em seladora e armazenadas sob congelamento entre -14°C e -20°C, até o momento dos ensaios.

### **2.6 MÉTODO DE ENSAIO**

O método validado foi baseado em procedimentos previamente descritos (PIERCE & BROWN, 1976; USDA, 2001).

### **2.6.1 Princípio**

O método de ensaio validado fundamentou-se na destruição da matéria orgânica por via seca com o auxílio de nitrato de magnésio. As cinzas obtidas foram dissolvidas em ácido clorídrico originando uma solução que foi tratada para redução do  $\text{As}^{+5}$  a  $\text{As}^{+3}$  e, em seguida, para gerar hidretos de arsênio em um sistema FIAS para detecção e quantificação de arsênio por AAS.

### **2.6.2 Cuidados**

Equipamentos de proteção individual (avental, óculos, luvas e máscaras) e equipamentos de proteção coletiva como capelas de exaustão química foram utilizados durante a realização dos ensaios.

### **2.6.3 Digestão, calcinação e pré-redução**

As amostras de fígado bovino foram descongeladas a temperatura ambiente até atingirem consistência de corte. Alíquotas de  $5,0000 \text{ g} \pm 0,0001 \text{ g}$  foram pesadas em cadinhos de vidro borossilicato, com auxílio de espátula ou pinça de inox. 4 mL de solução de nitrato de magnésio 50 % foram adicionados a cada amostra e homogeneizados com bastão de vidro, para prevenção da volatilização do arsênio. As amostras foram secas em chapa aquecedora a  $95 \text{ }^\circ\text{C} \pm 5 \text{ }^\circ\text{C}$ , até total carbonização. As amostras carbonizadas foram transferidas para um forno mufla a aproximadamente  $80 \text{ }^\circ\text{C}$  e a temperatura do forno foi aumentada para  $500 \text{ }^\circ\text{C} \pm 50 \text{ }^\circ\text{C}$ . Após 12 h de calcinação, as cinzas das amostras foram resfriadas, cuidadosamente adicionadas de 2 mL de solução de ácido nítrico 50 % e evaporadas até secura em chapa aquecedora  $95 \text{ }^\circ\text{C} \pm 5 \text{ }^\circ\text{C}$ . Em seguida, foram retornadas ao forno mufla a aproximadamente  $80 \text{ }^\circ\text{C}$  e a temperatura do forno foi aumentada para  $500 \text{ }^\circ\text{C} \pm 50 \text{ }^\circ\text{C}$ , permanecendo nesta temperatura por 2 h. Após resfriamento, as cinzas foram adicionadas de 10 mL de solução de ácido clorídrico 44,4 % e evaporadas em chapa aquecedora a  $95 \text{ }^\circ\text{C} \pm 5 \text{ }^\circ\text{C}$ , até completa dissolução e redução do volume para aproximadamente 5 mL. As soluções dos cadinhos foram transferidas quantitativamente para tubos graduados, utilizando porções de solução de ácido clorídrico 44,4 % até um volume final de 15 mL. Uma alíquota de  $1 \text{ mL} \pm 5 \%$  foi tomada de cada extrato de amostra e transferida para outro tubo de ensaio graduado. A cada alíquota foram adicionados  $2 \text{ mL} \pm 5 \%$  de solução de ácido clorídrico 50 % e  $1 \text{ mL} \pm 5 \%$  da solução de iodeto de potássio 5 % e ácido ascórbico 5 %, homogeneizando em agitador tipo Vortex Mixer durante 30 s. Foram aguardados 45 min para permitir a completa redução do  $\text{As}^{+5}$  a  $\text{As}^{+3}$ . O volume foi completado para 10 mL com água destilada e

desionizada e homogeneizado em agitador tipo Vortex Mixer durante 30 s. Brancos de solventes, sem adição de amostra, foram preparados em cada bateria analítica.

#### 2.6.4 Preparo da curva de calibração

Alíquotas da solução padrão de adição foram pipetadas e transferidas para balões volumétricos de 25 mL. Em seguida, foram adicionados 5 mL  $\pm$  5 % da solução de ácido clorídrico 50 % e 2,5 mL  $\pm$  5 % da solução de iodeto de potássio 5 % e ácido ascórbico 5 %, homogeneizando. Após 45 min, os volumes foram completados para 25 mL com água destilada e desionizada e homogeneizados em agitador tipo Vortex Mixer durante 30 s.

#### 2.6.5 Detecção e quantificação

Tubos contendo amostras e padrões foram colocados na bandeja do FIAS. As condições ajustadas neste sistema foram:

- solução carreadora de ácido clorídrico 10 % a um fluxo de 10 mL/min, programada para entrar no sistema antes da solução redutora;
- solução redutora de boroidreto de sódio 0,2 % a 5 mL/min;
- volume de injeção da amostra de 500  $\mu$ L;
- tubos de silicone de 1,14 mm; 1,52 mm e 3,17 mm de diâmetro interno para as soluções de borohidreto de sódio 0,02 %, ácido clorídrico 10 % e injeção de amostra e descartes, respectivamente. Demais tubulações foram de politetrafluoretileno com 1 mm de diâmetro interno;
- gás de transporte argônio;
- tubo de quartzo aquecido eletricamente como atomizador.

A programação do AAS foi:

- temperatura da célula de quartzo 900 °C;
- comprimento e diâmetro interno da célula de quartzo 180 mm x 7,5 mm;
- comprimento de onda 193,7 nm;
- fenda 0,7 nm;
- sinal absorvância;
- tempo de integração 20 s;
- técnica de integração altura dos picos;
- corrente da lâmpada 380 mA;
- sem correção de *background*.



Curvas de calibração da resposta em unidades de absorvância *versus* a concentração de arsênio foram construídas. As equações das curvas foram estimadas pelo método dos mínimos quadrados ordinários (MMQO) e utilizadas para cálculo da concentração de arsênio nas amostras.

## 2.7 VALIDAÇÃO

A adequação para uso do método foi determinada com base nos resultados dos parâmetros estabelecidos (linearidade, efeitos de matriz, seletividade, exatidão, precisão, limites de detecção e de quantificação experimentais), considerando os critérios de aceitabilidade determinados. O nível de significância adotado nos testes de hipóteses foi  $\alpha = 0,05$ . Os cálculos foram feitos em planilhas comerciais, cujos resultados foram previamente validados frente aqueles obtidos em diferentes softwares de análise estatística.

### 2.7.1 Linearidade

Linearidade foi avaliada por ensaios com soluções padrão. Três curvas de calibração foram preparadas, em diferentes dias, com níveis de concentração 0,5; 1; 2; 3; 4; 5; 6; 7 e 8 ng/mL (correspondentes a 15; 30; 60; 90; 120; 150; 180; 210 e 240  $\mu\text{g}/\text{kg}$  de arsênio em fígado bovino), sendo três replicatas independentes de cada nível. Alíquotas de 25; 50; 100; 150; 200; 250; 300; 350 e 400  $\mu\text{L}$ , respectivamente, da solução padrão de adição foram utilizadas para preparo das curvas conforme descrito no método de ensaio. As soluções foram analisadas em ordem aleatória dentro de cada dia. Brancos da curva foram preparados, em triplicata, como ferramenta de controle de qualidade, avaliação do ruído e ajuste do zero instrumental, mas não foram incluídos nos cálculos.

O MMQO, que minimiza a soma dos quadrados dos resíduos (erros) da regressão, foi utilizado para calcular as estimativas dos parâmetros da regressão. Este método assume que os resíduos seguem a distribuição normal, têm variância constante ao longo de do eixo  $x$  e são independentes (DRAPER & SMITH, 1998).

Após ajuste exploratório dos dados de cada curva pelo MMQO, gráficos de resíduos da regressão foram examinados para investigação de perfis óbvios, sendo valores extremos (*outliers*) indicados por pontos fora do intervalo  $\pm t_{(1-\alpha/2;n-2)}s_{res}$ , adaptado do intervalo proposto por MEYER & ZUND (1993), sendo  $s_{res}$  o desvio padrão dos resíduos da regressão. Os valores extremos foram formalmente tratados pelo teste de

resíduos padronizados Jackknife, o qual foi aplicado sucessivamente até que novos valores não fossem detectados ou até uma exclusão máxima de 22,2 % no número original de resultados (BELSLEY, KUH & WELSCH, 1980; HORWITZ, 1995). Violações das premissas relacionadas à análise de regressão foram avaliadas: normalidade (RYAN & JOINER, 1976), homoscedasticidade (LEVENE, 1960; BROWN & FORSYTHE, 1974) e independência dos resíduos da regressão (DURBIN & WATSON, 1951). Testes de F foram conduzidos para verificar o ajuste ao modelo linear por meio da avaliação das significâncias da regressão e do desvio da linearidade (DRAPER & SMITH, 1998).

### **2.7.2 Efeitos de matriz**

Efeitos de matriz foram checados pelo método de adição. Duas curvas de calibração foram preparadas com níveis 0,5; 1; 2; 4; 6 e 8 ng/mL, sendo três replicatas independentes de cada nível, analisadas em ordem aleatória, no mesmo dia. A curva do analito nos solventes foi preparada conforme descrito nos estudos de linearidade. Para a curva do analito na matriz utilizou-se alíquotas de 2,5 mL de soluções derivadas de análises de amostras brancas de fígado bovino. A estas alíquotas foram adicionados 25; 50; 100; 200; 300 e 400 µL da solução de adição de arsênio, 5 mL da solução de ácido clorídrico 50 % e 2,5 mL da solução de iodeto de potássio 5 % e ácido ascórbico 5 %, homogeneizando. Após 45 min, os volumes foram completados para 25 mL com água destilada e desionizada e homogeneizados em agitador tipo Vortex Mixer durante 30 s. Brancos da curva, preparados em triplicata, foram utilizados como ferramenta de controle de qualidade, avaliação do ruído e ajuste do zero instrumental, mas não foram incluídos nos cálculos.

Os parâmetros da regressão foram calculados pelo MMQO. Todas as premissas avaliadas para linearidade também foram checadas para estas curvas. Uma vez comprovados a distribuição normal dos resíduos da regressão e ajuste ao modelo linear, a inclinação e a interseção obtidas para a curva de solventes foram comparadas com a inclinação e interseção estimadas para a curva de matriz. As inclinações e interseções foram comparadas pelo teste de t com as variâncias combinadas (ARMITAGE & BERRY, 1994), após verificação da premissa de homogeneidade das variâncias dos resíduos das curvas pelo teste de F (SNEDECOR & COCHRAN, 1989).

### **2.7.3 Seletividade, exatidão, precisão e limites**

Seletividade, exatidão, precisão (sob condições de repetitividade e reprodutibilidade parcial) e limites de detecção e quantificação experimentais foram

estabelecidos por ensaios com amostras brancas e adicionadas em níveis correspondentes a 15; 30; 60; 120 e 240 µg/kg, em pelo menos dez replicatas independentes de cada nível. As amostras foram divididas em três lotes, os quais foram analisados em três dias diferentes por dois analistas, sendo dia 1/analista 1, dia 2/analista 2 e dia 3/analista 1. Considerando os resultados dos testes de efeitos de matriz, curvas de solventes foram preparadas para calcular as concentrações de arsênio nas amostras adicionadas.

A ausência de resultados falsos positivos para as amostras brancas e um perfil indistinguível entre as respostas de arsênio para amostras adicionadas e padrões foram considerados suficientes para determinação da seletividade.

Valores extremos foram investigados nos resultados de recuperação aparente pelo teste de Grubbs (GRUBBS, 1969; BARRET & LEWIS, 1994; BURKE 2001 b). As estatísticas de Grubbs para um valor extremo ( $G_1$ ), dois valores vicinais ( $G_2$ ) e dois valores polares ( $G_3$ ) foram calculadas simultaneamente para os resultados de cada nível de concentração. A indicação de valor extremo por, pelo menos, uma destas estatísticas foi suficiente para eliminação do dado. Após remoção do valor extremo, o teste de Grubbs foi aplicado sucessivamente até que novos valores não fossem detectados ou até uma exclusão máxima de 22,2 % no número original de dados (HORWITZ, 1995).

Exatidão foi investigada por meio da média de recuperação aparente obtida para as replicatas de amostras adicionadas em cada nível de concentração. De acordo com os critérios estabelecidos pela EC (2002), para níveis de concentração  $\geq 10$  µg/kg, recuperações entre 80 % e 110 % foram consideradas aceitáveis.

Precisão, sob condições de repetitividade e reprodutibilidade parcial, foi estimada por análise de variância (ISO, 1994 a,b,c,d,e; ISO, 1998; KUTTATHARMMAKUL, MASSART & SMEYERS-VERBEKE, 1999) e expressa em termos de desvios padrão relativos obtidos para as replicatas de amostras adicionadas em cada nível de concentração. Um delineamento fatorial completamente aninhado (*fully-nested design*) forneceu as estimativas da variância de repetitividade ( $s_r^2$ ); da variância do ensaio ( $s_{Ensaio}^2$ ), que engloba as fontes de variação confundidas do dia e do analista (dia + analista); da variância entre ensaios [ $s_{entre\ ensaios}^2 = (s_{Ensaio}^2 - s_r^2) / n$ ], sendo  $n$  o número médio de replicatas em cada nível de concentração; e da variância de reprodutibilidade parcial ( $s_R^2 = s_r^2 + s_{entre\ ensaios}^2$ ) (MAROTO et al, 1999; MAROTO et al, 2001). Quando a estimativa da variância de repetitividade foi maior que a variância do ensaio,  $s_{entre\ ensaios}^2$  foi considerada como sendo igual a zero, evitando a obtenção de um valor negativo para a

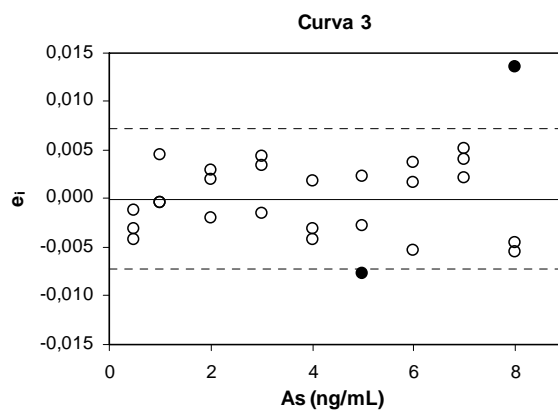
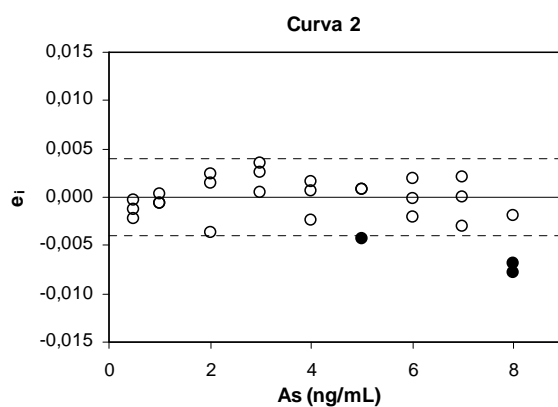
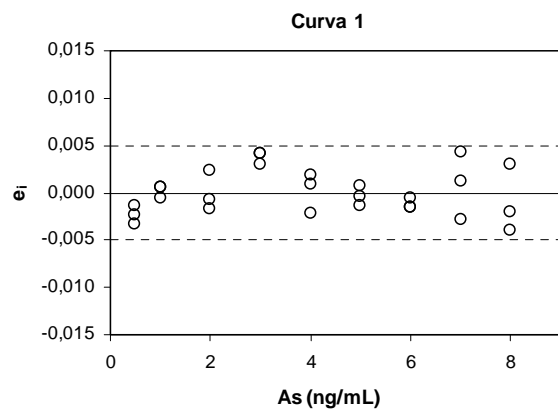
estimativa desta variância (ISO, 1994; KUTTATHARMMAKUL, MASSART & SMEYERS-VERBEKE, 1999). Neste caso, as estimativas das variâncias de repetitividade e de reprodutibilidade parcial foram consideradas equivalentes.

Para realização da análise de variância, as premissas relacionadas os teste de F foram previamente testadas: normalidade (RYAN & JOINER, 1976) e homoscedasticidade dos resíduos de recuperação aparente (LEVENE, 1960, BROWN & FORSYTHE, 1974). Desvios padrão relativos de reprodutibilidade parcial ( $DPR_r$ )  $\leq 22\%$  e desvios padrão relativos de repetitividade ( $DPR_r$ )  $\leq 14,7\%$  foram considerados aceitáveis. Estes valores corresponderam valor de precisão interlaboratorial estimado pela equação de HORWITZ (1982) modificada por THOMPSON (2000) e a dois terços (EC, 2002) do valor desta precisão, respectivamente.

O limite de quantificação foi estabelecido como a concentração abaixo da qual o método não pôde operar com exatidão e precisão aceitáveis. O limite de detecção foi a menor concentração detectada nas amostras adicionadas, mas não necessariamente quantificada, distinguida de zero (sinal da amostra adicionada / ruído da amostra branca  $\geq 3$ ). Estes limites foram estabelecidos com base nos resultados médios de recuperação aparente e desvios padrão relativos de repetitividade e reprodutibilidade parcial obtidos para as replicatas de amostras adicionadas nos diferentes níveis de concentração estudados.

### 3 RESULTADOS E DISCUSÃO

Os gráficos exploratórios dos resíduos da regressão indicaram a presença de valores extremos pelo critério do intervalo entre  $\pm t_{(1-\alpha/2; n-2)} S_{res}$ , os quais foram confirmados pelo teste formal de resíduos padronizados Jackknife. Valores extremos não foram detectados na curva 1. Três valores extremos foram observados na curva 2, sendo um no nível 5 ng/mL e dois no nível 8 ng/mL, enquanto dois valores extremos foram diagnosticados na curva 3 nos mesmos níveis. Em nenhum caso foram atingidos seis valores extremos, correspondentes ao limite máximo de pontos passíveis de serem tratados ou 22,2 % dos vinte e sete dados originais. Esta etapa permitiu a remoção de pontos que poderiam apresentar alguma influência na equação de regressão ajustada. A Figura III.1 ilustra os valores extremos tratados e demonstra que não houve tendências óbvias na distribuição dos resíduos da regressão com a eliminação dos mesmos.



**Figura III.1** - Gráficos exploratórios dos resíduos da regressão das três curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes, com indicação dos respectivos valores extremos diagnosticados pelo teste de resíduos padronizados Jackknife.

$e_i$  = resíduo da regressão, ● = valor extremo, - - - =  $\pm t_{(1-\alpha/2; n-2)} s_{res}$ .

Os gráficos de Ryan-Joiner ilustrados na Figura III.2 indicam correlação significativa entre os resíduos da regressão obtidos para cada curva e os valores normais teoricamente esperados, denotando que o desvio da distribuição normal não foi significativo ( $p > 0,10$ ). Coeficientes de correlação de 0,9824; 0,9930 e 0,9718 foram detectados para as curvas 1, 2 e 3, com valores críticos de 0,9681; 0,9651 e 0,9662, respectivamente.

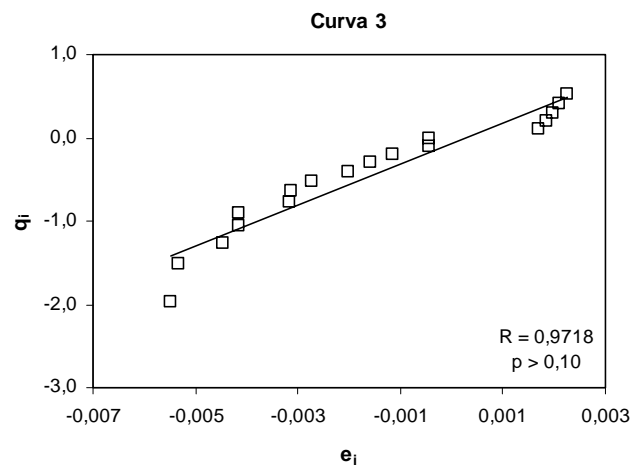
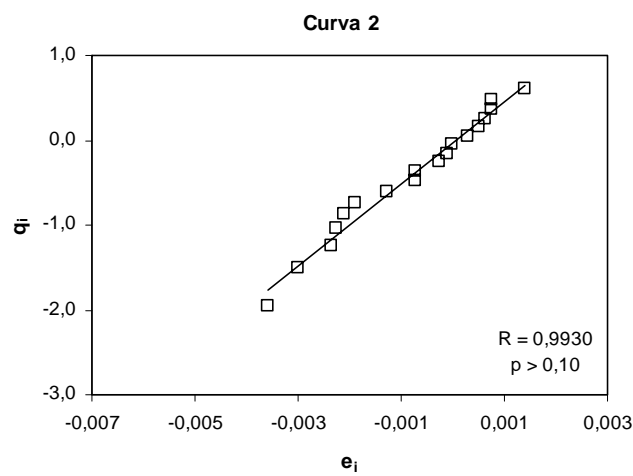
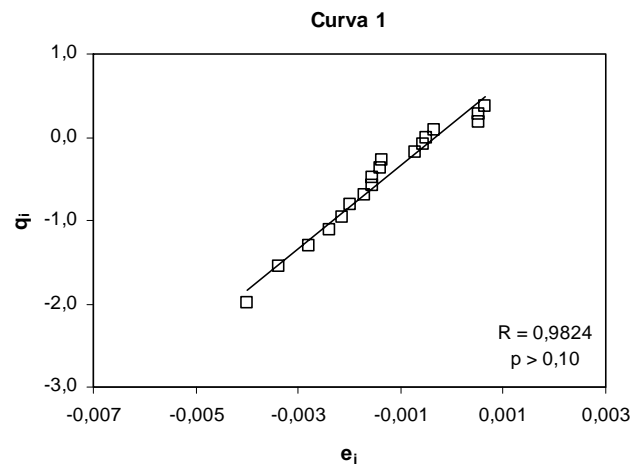
A premissa de homoscedasticidade dos resíduos da regressão foi confirmada. Estatísticas t de Levene de 0,266; 0,426 e -0,913, com significâncias de 0,79; 0,67 e 0,37, estimadas para as curvas 1, 2 e 3, respectivamente, não foram significativas como demonstrado na Figura III.3.

Houve independência dos resíduos da regressão. As estatísticas de Durbin-Watson calculadas de 1,633 para a curva 1; 1,759 para a curva 2; e 1,923 para a curva 3, indicaram que não houve autocorrelação ( $p > 0,10$ ) (Figura III.4). Quando os dados das curvas 1, 2 e 3 foram combinados, a estatística de Durbin-Watson calculada foi de 1,871 ( $p > 0,10$ ), ratificando independência dos resíduos da regressão. Segundo DRAPER & SMITH (1998), uma inferência mais segura pode ser feita com  $n \geq 50$ , devido à largura da banda dos limites críticos da estatística de Durbin-Watson.

A confirmação das três premissas indicou a possibilidade de uso de testes de hipóteses para análise dos dados obtidos e sugeriu a estimativa dos parâmetros da regressão pelo MMQO, sem necessidade de ponderações.

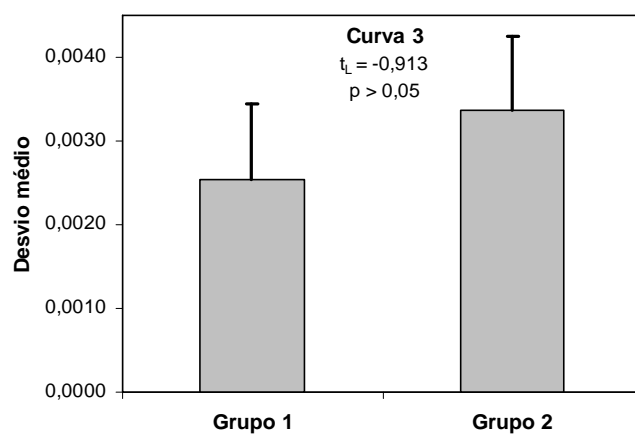
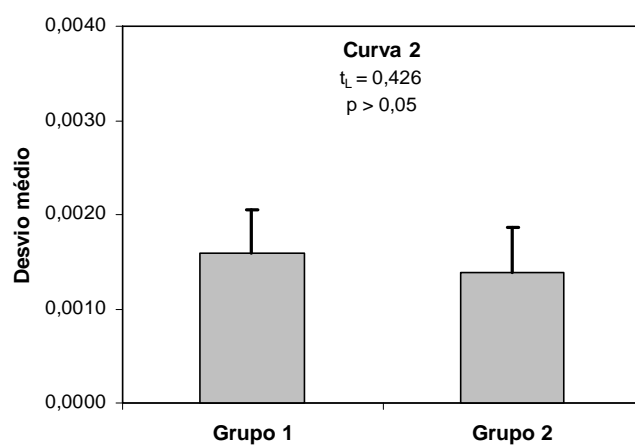
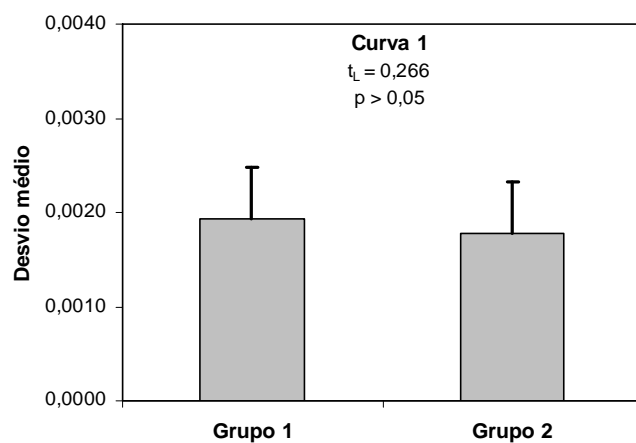
A regressão apresentou-se altamente significativa ( $p < 0,001$ ) enquanto desvios de linearidade não significativos ( $p > 0,05$ ) foram observados para as curvas 1, 2 e 3 (Tabela III.1), indicando linearidade na faixa estudada de 0,5 ng/mL a 8 ng/mL. As curvas de calibração com as respectivas equações e coeficientes de determinação estão ilustradas na Figura III.5.

A faixa linear avaliada neste trabalho foi similar a 0 µg/L a 10 µg/L ( $n = 6$ ) descrita por NÄYKKI et al. (2001) e menor que 0,25 µg/L a 2000 µg/L detectada por RUPASINGHE et al. (2004). Contudo, estes autores somente avaliaram o coeficiente de determinação ( $R^2$ ) como indicador de linearidade. Apesar da ampla utilização deste coeficiente, MULHOLLAND & HIBBERT (1997) demonstraram que erros grosseiros podem co-existir com valores aceitáveis de  $R^2$ . ANSCOMBE (1973), por meio de seu clássico exemplo, também sugeriu que valores numéricos como  $R^2$  não são apropriados para avaliação da linearidade. MINDAK & DOLAN (1999) empregaram curvas de branco



**Figura III.2** - Gráficos de probabilidade normal das três curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes.

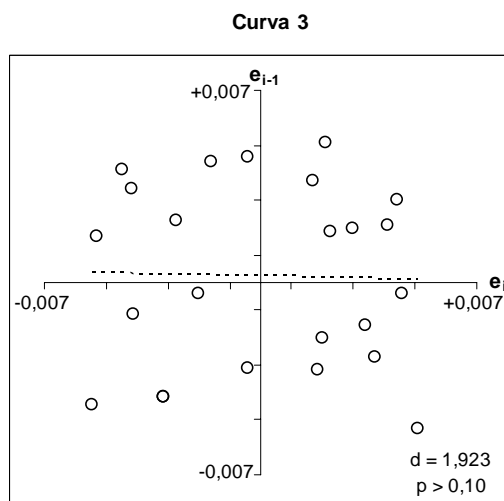
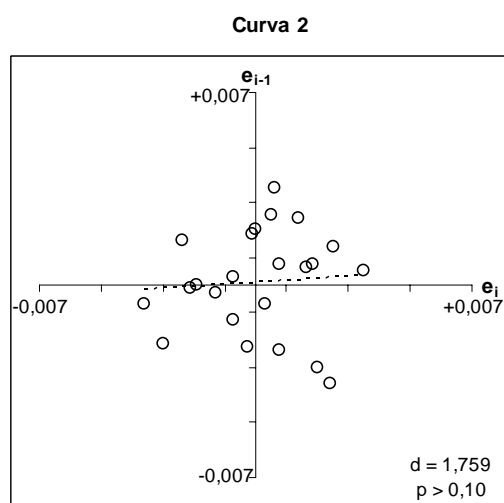
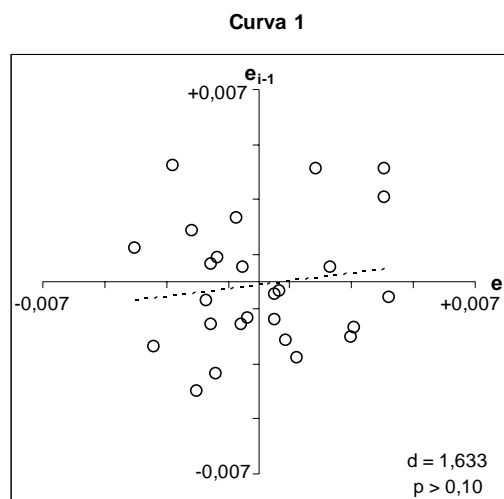
$e_i$  = resíduo da regressão,  $q_i$  = valor normal esperado,  $R$  = coeficiente de correlação de Ryan-Joiner,  $p$  = significância do desvio da distribuição normal.



**Figura III.3** - Representação da homogeneidade das variâncias das três curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes pelo teste de Levene modificado.

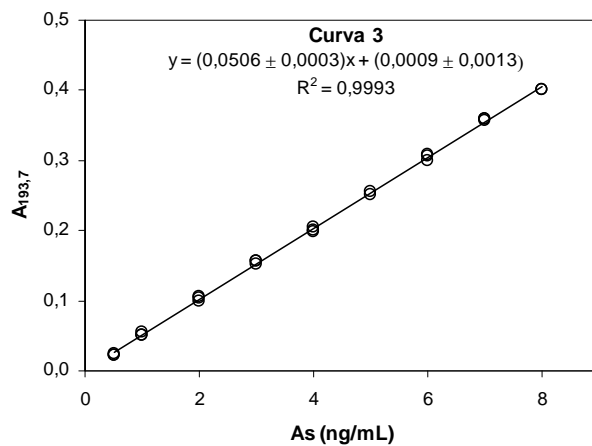
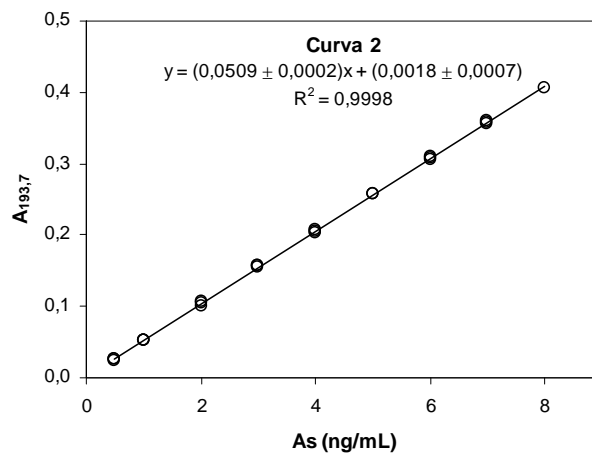
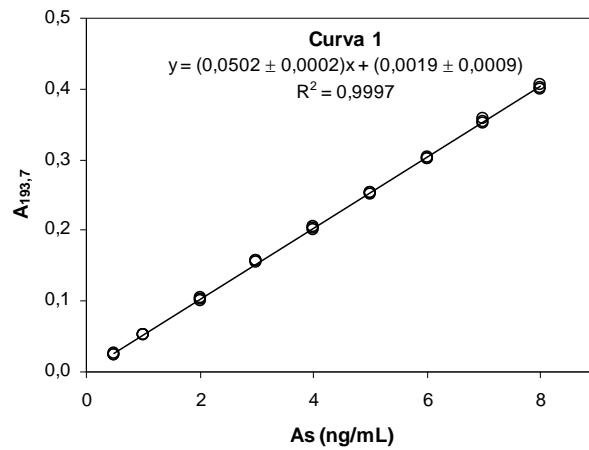
$t_L$  = estatística t de Levene,  $p$  = significância, Grupo 1 = 0,5; 1; 2; 3 e 4 ng/mL, Grupo 2 = 5; 6; 7 e 8 ng/mL.





**Figura III.4** - Gráficos de Durbin-Watson das três curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes.

$e_i$  = resíduo da regressão,  $d$  = estatística de Durbin-Watson,  $p$  = significância.



**Figura III.5** - Curvas de calibração do arsênio nos solventes obtidas nos ensaios de avaliação de linearidade, na faixa de 0,5 ng/mL a 8 ng/mL, com respectivas equações e coeficientes de determinação.

$A_{193,7}$  = leitura de absorvância no comprimento de onda 193,7 nm,  $y$  = resposta,  $x$  = concentração de As em ng/mL,  $R^2$  = coeficiente de determinação.

**Tabela III.1** - Estatísticas da análise de variância da regressão, incluindo teste de desvio de linearidade, para as três curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes

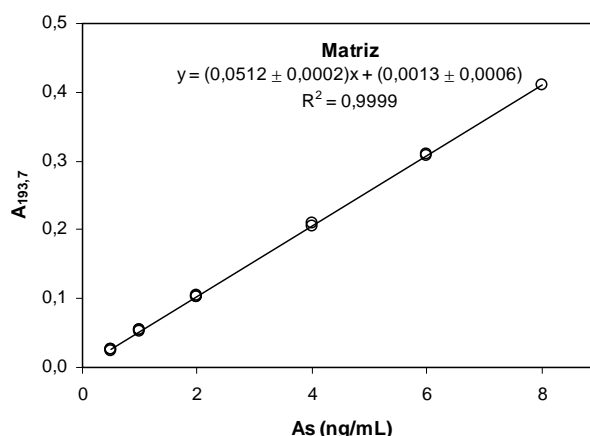
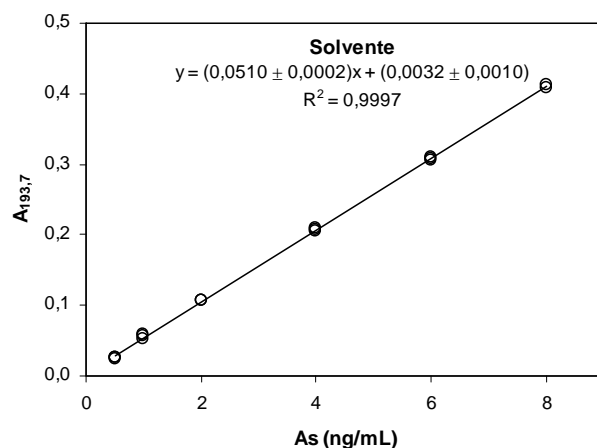
<b>Estatística</b>	<b>Curva 1</b>	<b>Curva 2</b>	<b>Curva 3</b>
<i>Regressão</i>			
<i>F</i>	7,404 x 10 <sup>+4</sup>	9,252 x 10 <sup>+4</sup>	3,117 x 10 <sup>+4</sup>
<i>p</i>	6,35 x 10 <sup>-45</sup>	2,30 x 10 <sup>-41</sup>	1,56 x 10 <sup>-37</sup>
<i>Desvio da linearidade</i>			
<i>F</i>	2,419	0,927	2,409
<i>p</i>	0,06	0,51	0,07

*F* = razão entre variâncias, *p* = significância.

de solventes e soluções padrão de 0,5 ng/mL a 10 ng/mL, sendo que as áreas dos picos de absorvância integradas eram convertidas a unidades de concentração utilizando o ajuste ao modelo não linear fornecido pelo fabricante do equipamento. No entanto, não houve registros de estudos sobre linearidade, assim como sobre ajustes a modelos não lineares neste trabalho.

Testes iniciais de linearidade foram realizados em faixas de concentração mais amplas para a validação aqui apresentada, tendo como referência os intervalos indicados no método proposto pelo USDA (2001) e os limites regulamentados pelo BRASIL (2006). No entanto, as premissas da regressão não foram atendidas nestas faixas. Mesmo no intervalo de 0,5 ng/mL a 10 ng/mL, no qual houve detecção de 2 valores extremos no nível 10 ng/mL e confirmação de homoscedasticidade ( $p > 0,05$ ), foram observados desvio da normalidade ( $p < 0,01$ ), autocorrelação ( $p < 0,01$ , pelo teste simplificado) e desvio da linearidade ( $p < 0,001$ ). Estes resultados preliminares levaram a um estreitamento da faixa de trabalho para o intervalo de 0,5 ng/mL a 8 ng/mL demonstrado. Embora esta faixa esteja abaixo da concentração correspondente ao limite estabelecido pelo BRASIL (2006) para arsênio em fígado bovino, existe a possibilidade de diluições das amostras caso concentrações superiores à faixa linear encontrada sejam detectadas. Cabe considerar que a opção por uma faixa linear inferior também se justifica visto a tendência mundial de rigidez em aspectos relativos à segurança alimentar, com freqüentes reduções dos limites máximos permitidos estabelecidos.

As curvas de solventes e de matriz obtidas nos estudos de efeitos de matriz estão apresentadas na Figura III.6. Dois valores extremos foram detectados na curva de solventes nos níveis 2 ng/mL e 8 ng/mL, enquanto quatro valores foram diagnosticados na curva de matriz nos níveis 4; 6 e 8 ng/mL, pelo teste de resíduos padronizados Jackknife.



**Figura III.6** - Curvas de calibração do analito nos solventes e na matriz obtidas para arsênio nos ensaios de avaliação de efeitos de matriz, na faixa de 0,5 ng/mL a 8 ng/mL, com respectivas equações e coeficientes de determinação.

$A_{193,7}$  = leitura de absorvância no comprimento de onda 193,7 nm,  $y$  = resposta,  $x$  = concentração de As em ng/mL,  $R^2$  = coeficiente de determinação.

Todas as premissas relativas ao MMQO foram confirmadas. Distribuição normal dos resíduos da regressão ( $p > 0,10$ ) foi observada com coeficientes de correlação de Ryan-Joiner de 0,9630 e 0,9649 para as curvas de solventes e de matriz, sendo os valores críticos 0,9529 e 0,9481, respectivamente. Homoscedasticidade foi obtida com estatísticas t de Levene de 1,442 para a curva de solventes e de 0,270 para a curva de matriz ( $p > 0,05$ ). Independência dos resíduos da regressão foi indicada pelas estatísticas de Durbin-Watson de 1,689 e 1,378 ( $p > 0,10$ ) para as curvas de solventes e matriz, respectivamente. Regressões significativas ( $p < 0,001$ ) e desvios da linearidade não significativos ( $p > 0,05$ ) foram observados pelos testes de F (Tabela III.2). Estes

resultados sinalizaram a possibilidade de comparação das inclinações e interseções pelo teste de t para avaliação dos efeitos de matriz.

**Tabela III.2** - Estatísticas da análise de variância da regressão, incluindo teste de desvio da linearidade, para as curvas de arsênio (0,5 ng/mL a 8 ng/mL) em solventes e matriz

Estatística	Solvente	Matriz
	<i>Regressão</i>	
<i>F</i>	4,687 x 10 <sup>+4</sup>	9,028 x 10 <sup>+4</sup>
<i>p</i>	4,43 x 10 <sup>-26</sup>	1,24 x 10 <sup>-24</sup>
	<i>Desvio da linearidade</i>	
<i>F</i>	2,504	3,616
<i>p</i>	0,11	0,06

*F* = razão entre variâncias, *p* = significância.

Segundo WELZ & STAUSS (1993) interferências espectrais, de memória e de matriz podem ser observadas em HGAAS. No entanto, na HG interferências espectrais são menos freqüentes, devido à volatilização do analito e sua separação da matriz. A limpeza do equipamento e o tempo estabelecido para o ciclo de análise de cada amostra garantem a eliminação dos interferentes de memória que poderiam gerar resultados falso-positivos. Por outro lado, a interferência de componentes da matriz pode inibir a formação de hidretos do analito ou alterar a cinética desta reação e induzir a formação de bolhas no separador gás-líquido gerando problemas no transporte do analito. A presença de outros elementos formadores de hidretos na matriz pode ainda suprimir ou distorcer a resposta do analito com picos múltiplos de absorção.

Efeitos de matriz não foram detectados na faixa de 0,5 ng/mL a 8 ng/mL. O teste de F indicou homogeneidade ( $p > 0,05$ ) das variâncias dos resíduos das curvas de solventes e de matriz, permitindo o uso do teste de t com variâncias combinadas para comparação dos parâmetros interseção e inclinação. Não foram observadas diferenças significativas ( $p > 0,05$ ) quando as inclinações ( $t_b = 0,826$ , com  $p = 0,42$ ) e interseções ( $t_a = 1,599$ , com  $p = 0,12$ ) estimadas para a curva de solventes e para a curva de matriz foram comparadas. Considerando estes resultados, foi possível concluir que soluções padrão de solventes preparadas como descrito neste trabalho forneceram respostas, em unidades de absorvância, que não diferiram daquelas obtidas para amostras de fígado bovino contendo concentrações correspondentes de arsênio. Conseqüentemente, curvas

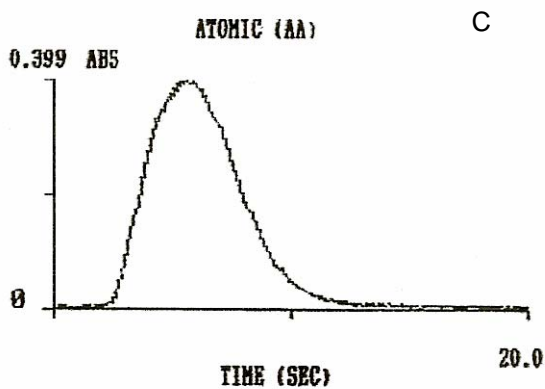
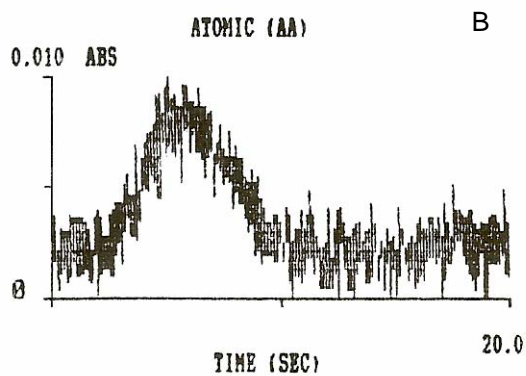
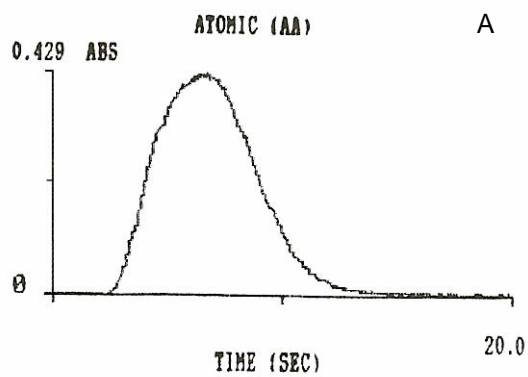
de solventes foram utilizadas nas etapas subseqüentes da validação para estimar as concentrações de arsênio em amostras de fígado bovino.

Na Figura III.7 estão ilustrados registros gráficos típicos de leituras de absorvância obtidas para soluções padrão de arsênio, amostras brancas e amostras adicionadas. Todas as amostras brancas analisadas apresentaram resultados não detectados (sinal da amostra branca / ruído do branco da curva  $< 3$ ) para arsênio. Os registros obtidos para amostras adicionadas foram indistinguíveis daqueles obtidos para as soluções padrão de arsênio correspondentes.

Não foram identificados valores extremos quando os resultados de recuperação aparente obtidos em cada nível de concentração foram analisados pelo teste de Grubbs ( $p > 0,05$ ). Normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente foram observadas pelos testes de Ryan-Joiner e Levene modificado (Tabela III.3), respectivamente, permitindo a estimativa dos desvios padrão relativos de repetitividade e reprodutibilidade parcial por análise de variância. Neste caso, os resíduos foram estimados como a diferença entre a média de recuperação aparente e os valores individuais de recuperação aparente obtidos em cada dia para cada nível de concentração.

Os valores médios de recuperação aparente,  $DPR_r$  e  $DPR_R$  calculados para cada nível de concentração encontram-se apresentados na Tabela III.4. As médias de recuperação aparente variaram entre 88,9 % e 99,9 %, com  $DPR_r$  e  $DPR_R$  entre 3,7 % e 42,9 %. Observou-se que os valores de  $DPR_r$  e  $DPR_R$  estimados para os níveis de 15; 120 e 240  $\mu\text{g}/\text{kg}$  foram os mesmos. Isto porque, a estimativa da variância de repetitividade foi maior que a do ensaio (dia+analista), sendo atribuído o valor de zero para a variação entre ensaios (ISO, 1994; ISO, 1998).

A precisão não foi homogênea ao longo da faixa de concentração na qual o método foi validado. Conforme afirmado por THOMPSON, ELLISON & WOOD (2002), muito freqüentemente a precisão varia com a concentração do analito. No nível 15  $\mu\text{g}/\text{kg}$  foram alcançados valores de  $DPR_r$  e  $DPR_R$  maiores que os valores críticos estabelecidos de 14,7 % e 22%, respectivamente. Valores médios de recuperação aparente encontraram-se dentro dos critérios de aceitabilidade em toda a faixa examinada (15  $\mu\text{g}/\text{kg}$  a 240  $\mu\text{g}/\text{kg}$ ). Com base nos critérios designados, pôde-se concluir que não houve falta de exatidão e que a precisão do método foi adequada na faixa de concentração de 30  $\mu\text{g}/\text{kg}$  a 240  $\mu\text{g}/\text{kg}$ .



**Figura III.7** - Registros gráficos típicos de leituras de absorvância.

A = solução padrão de arsênio de 8 ng/mL.

B = amostra de fígado bovino branca.

C = amostra de fígado bovino adicionada de arsênio em 240  $\mu\text{g}/\text{kg}$ .

**Tabela III.3** - Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente de arsênio

Estatística	Concentração (µg/kg)				
	15	30	60	120	240
<i>Normalidade</i>					
<i>R</i>	0,9884	0,9866	0,9616	0,9290	0,9871
<i>p</i>	p > 0,10	p > 0,10	p > 0,10	p > 0,05	p > 0,10
<i>Homoscedasticidade</i>					
<i>F<sub>L</sub></i>	0,708	0,630	0,253	0,286	0,223
<i>p</i>	0,51	0,55	0,78	0,76	0,81

*R* = coeficiente de correlação de Ryan-Joiner, *p* = significância, *F<sub>L</sub>* = estatística F de Levene.

**Tabela III.4** - Médias de recuperação aparente e desvios padrão relativos, sob condições de repetitividade e reprodutibilidade parcial, obtidos para amostras de fígado bovino adicionadas de arsênio em diferentes níveis de concentração

Concentração (µg/kg)	Fígado bovino			
	<i>n</i>	$\bar{R}$ (%)	<i>DPR<sub>r</sub></i> (%)	<i>DPR<sub>R</sub></i> (%)
15	16	89,2	42,9	42,9
30	16	99,9	7,2	7,3
60	11	99,3	6,4	6,6
120	11	88,9	4,5	4,5
240	11	91,1	3,7	3,7

*n* = número de observações,  $\bar{R}$  = média de recuperação aparente, *DPR<sub>r</sub>* = desvio padrão relativo de repetitividade, *DPR<sub>R</sub>* = desvio padrão relativo de reprodutibilidade parcial.

Critério de aceitabilidade para  $\bar{R}$ : de 80 % a 110 %.

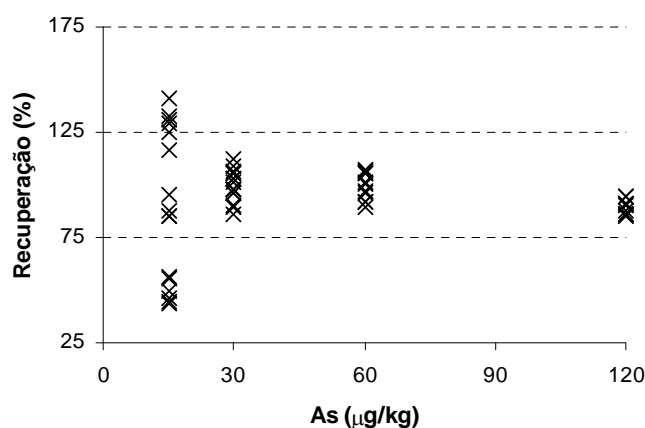
Critério de aceitabilidade para *DPR<sub>r</sub>*: ≤ 14,7 %.

Critério de aceitabilidade para *DPR<sub>R</sub>*: ≤ 22 %.

Em todas as replicatas de amostras adicionadas no nível 15 µg/kg (correspondente a 0,5 ng/mL) foi detectado arsênio com uma relação sinal da amostra adicionada / ruído da amostra branca ≥ 3. Esta concentração foi então estabelecida como o limite de detecção do método. O limite de quantificação foi o nível 30 µg/kg (correspondente a 1 ng/mL), menor concentração em que os resultados de exatidão e precisão atingiram os critérios de aceitabilidade. A Figura III.8 apresenta a distribuição dos resultados individuais de recuperação aparente obtidos para amostras adicionadas entre 15 µg/kg e 120 µg/kg, ilustrando a transição entre as faixas de detecção e quantificação do método.



Limites de detecção e quantificação de 0,037 ng/mL e 0,089 ng/mL, respectivamente, foram reportados por NÄYKKI et al. (2001). RUPASINGHE et al. (2004) obtiveram limite de detecção de 0,18 ng/mL. MINDAK & DOLAN (1999) descreveram um limite de detecção de 0,05 ng/mL e estimaram 10 µg/kg como limite de quantificação do método com base em alíquotas de 1 g de amostras. Estes limites foram menores que os encontrados neste trabalho, entretanto, os autores determinaram os limites de detecção e quantificação como  $(\bar{x}_{branco} + 3s_{branco})$  e  $(\bar{x}_{branco} + 10s_{branco})$  respectivamente, sendo  $\bar{x}_{branco}$  a média dos resultados obtidos para uma solução aquosa branca e  $s_{branco}$  o desvio padrão destes resultados. Os limites aqui apresentados foram determinados experimentalmente em ensaios de recuperação com amostras de fígado bovino adicionadas de solução padrão de arsênio, o que certamente forneceu limites mais realistas para o método em validação. Quando amostras adicionadas são utilizadas, os limites são estabelecidos envolvendo todas as etapas analíticas e suas respectivas perdas de analito e não somente a etapa de detecção e quantificação.



**Figura III.8** - Distribuição dos resultados individuais de recuperação aparente obtidos para amostras adicionadas entre 15 µg/kg e 120 µg/kg, incluindo os níveis correspondentes aos limites de detecção (15 µg/kg) e quantificação (30 µg/kg).

## 4 CONCLUSÕES

O procedimento proposto cobriu os parâmetros de desempenho necessários para avaliação da eficiência de métodos quantitativos, tendo como base estatísticas precisas,

além de ter apresentado adequada aplicação prática no processo de validação intralaboratorial do método de ensaio de arsênio em fígado bovino por HGAAS-FIAS.

Uma faixa linear foi determinada, sem efeitos de matriz significativos. Seletividade, exatidão e precisão contemplaram os critérios estabelecidos como requisitos analíticos. Os limites de detecção e quantificação estabelecidos permitiram a determinação de arsênio em fígado bovino considerando o limite regulamentado no Brasil. Estes parâmetros de desempenho do método validado indicaram sua adequação para atividades de controle de arsênio em fígado bovino, além da possibilidade de avaliações futuras da precisão sob condições de reprodutibilidade por meio de estudos colaborativos.

## **5 CONSIDERAÇÕES**

Os experimentos descritos neste capítulo foram realizados no Laboratório de Referência Agropecuária de Minas Gerais (LANAGRO-MG) do MAPA, em Pedro Leopoldo, Minas Gerais, Brasil.

A referência da publicação deste trabalho encontra-se descrita a seguir:  
SOUZA, S.V.C.; PINTO, C.T.; JUNQUEIRA, R.G. In-house method validation: application in arsenic analysis. *J. Food Compos. Anal.*, v. 20, p 241-247, 2007.

# **CAPÍTULO IV - VALIDAÇÃO INTRALABORATORIAL DE MÉTODO QUANTITATIVO PARA DETERMINAÇÃO MÚLTIPLA DE RESÍDUOS DE AVERMECTINAS EM LEITE BOVINO POR CROMATOGRAFIA LÍQUIDA DE ALTA EFICIÊNCIA COM DETECÇÃO DE FLUORESCÊNCIA**

## **RESUMO**

Um procedimento detalhado para validação intralaboratorial de métodos em análises de alimentos foi proposto. Este procedimento foi aplicado à validação de um método para ensaio de abamectina, doramectina, eprinomectina e ivermectina em leite bovino por cromatografia líquida de alta eficiência com detector de fluorescência. Linearidade e efeitos de matriz foram avaliados em ensaios com curvas dos analitos em solvente e na matriz, na faixa de 25; 50; 75; 125 e 150 ng/mL. Seletividade, exatidão, precisão, limites de detecção e quantificação, limite de decisão e capacidade de detecção foram estabelecidos em ensaios com amostras de leite bovino brancas e adicionadas a 5; 10; 15; 20 e 30 µg/L. As premissas relacionadas às estatísticas empregadas foram avaliadas e confirmadas. Linearidade foi obtida entre 25 ng/mL e 150 ng/mL. Não foram observados efeitos de matriz significativos nesta faixa. As médias de recuperação aparente variaram de 87,2 % a 101,4 %. Desvios padrão relativos sob condições de repetitividade estiveram entre 3,3 % e 11,2 %, enquanto os desvios padrão relativos de reprodutibilidade parcial foram de 7,4 % a 14,7 %. Estes resultados indicaram que não houve falta de exatidão e que a precisão do método foi adequada entre 10 µg/L e 30 µg/L, para os quatro analitos estudados. Os limites de detecção e quantificação experimentais foram 5 µg/L e 10 µg/L, respectivamente. Limites de decisão e capacidades de detecção foram estimados, assumindo um limite máximo de resíduo de 10 µg/L.

Palavras-chave: validação intralaboratorial; abamectina; doramectina; eprinomectina; ivermectina; leite bovino; cromatografia líquida de alta eficiência.

## **ABSTRACT**

SINGLE-LABORATORY VALIDATION OF A MULTI-RESIDUE METHOD FOR DETERMINATION OF AVERMECTINS RESIDUES IN BOVINE MILK BY HIGH

## PERFORMANCE LIQUID CHROMATOGRAPHY WITH FLUORESCENCE DETECTION.

A detailed procedure for single-laboratory method validation in food analysis was proposed. This procedure was applied to the validation of a method for the assay of abamectin, doramectin, eprinomectin and ivermectin in bovine milk by high performance liquid chromatography with fluorescence detection. Linearity and matrix effects were assessed by solvent and matrix-matched calibration curves at 25; 50; 75; 125 and 150 ng/mL. Selectivity, trueness, precision, experimental limits of detection and quantification, decision limit and detection capability were established by assays with milk blanks and spiked at 5; 10; 15; 20 and 30 µg/L. The assumptions related to the statistics employed were evaluated and confirmed. Linearity was obtained between 25 ng/mL and 150 ng/mL. No significant matrix effects were observed in this range. Mean of apparent recoveries varied from 87.2 % to 101.4 %. The relative standard deviations under repeatability conditions ranged from 3.3 % to 11.2 %, while the within-reproducibility relative standard deviations were from 7.4 % to 14.7 %. These results indicated no lack of trueness and precision between 10 µg/L and 30 µg/L, for the four analytes studied. The experimental limits of detection and quantification were 5 µg/L and 10 µg/L, respectively. Decision limits and detection capabilities were estimated, assuming the maximum residue level of 10 µg/L.

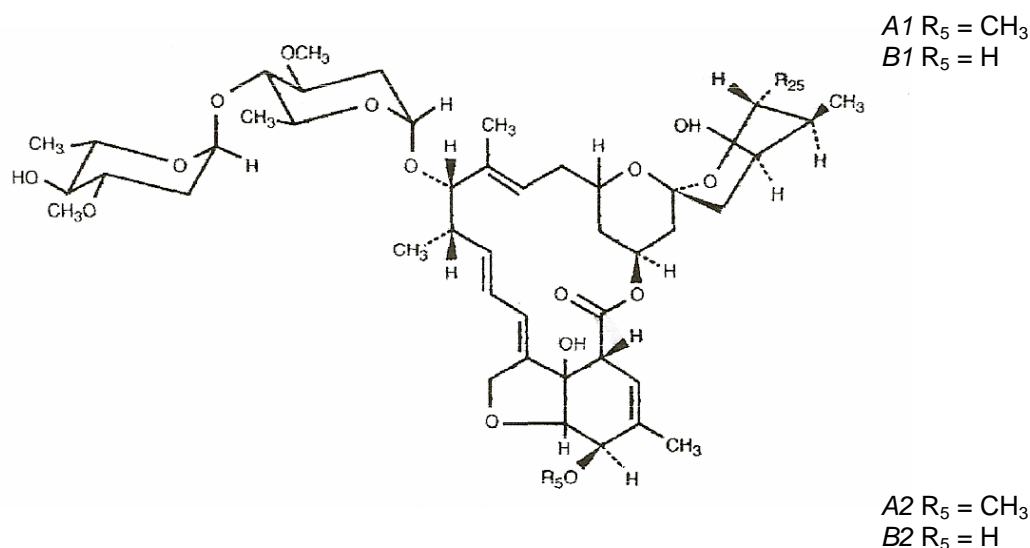
Key words: single-laboratory validation; abamectin; doramectin; eprinomectin; ivermectin; bovine milk; high performance liquid chromatography.

## 1 INTRODUÇÃO

As avermectinas constituem um importante grupo de agentes antiparasitários, amplamente utilizados em bovinos criados para produção de alimentos. Estes compostos são altamente efetivos em doses extremamente baixas contra um grande espectro de nematóides e artrópodes, muitos dos quais são significantes endoparasitas e ectoparasitas animais (CAMPBELL et al., 1983).

As avermectinas (Figura IV.1) são lactonas macrolíticas, estruturalmente relacionadas, produzidas pela fermentação de actinomicetos *Streptomyces avermitilis*. No processo de fermentação são produzidos oito tipos de avermectinas: *A1a*, *A1b*, *A2a*, *A2b*, *B1a*, *B1b*, *B2a* e *B2b*. Os compostos da série *A* são metoxilados no átomo de carbono da posição 5 enquanto os compostos da série *B* têm um grupo hidroxil nesta posição. A sub-

série 1 possui uma dupla ligação olefínica entre os carbonos 22 e 23, a qual é hidratada na sub-série 2, resultando em um grupo hidroxila na posição 23. Os homólogos *a* e *b* diferem por seus substituintes na posição do carbono 25 (CHABALA et al., 1980; SAMS, 1993). As avermectinas da série *A* são, geralmente, menos efetivas que aquelas da série *B* e, por esta razão, não foram desenvolvidas para uso comercial (CAMPBELL et al., 1983). A similar atividade biológica dos homólogos *a* e *b* não justifica os custos práticos de sua separação em processos industriais (VALENZUELA et al., 2001). Portanto, os produtos comerciais são uma mistura de não menos que 90 % do homólogo *a* e não mais que 10 % do homólogo *b* (CODEX ALIMENTARIUS, 2005 b).



**Figura IV.1** - Estrutura química das avermectinas.

Homólogo *a*  $R_{25} = \text{CH}(\text{CH}_3)\text{C}_2\text{H}_5$ .

Homólogo *b*  $R_{25} = \text{CH}(\text{CH}_3)_2$ .

Apesar do grande número de investigações, o mecanismo de ação das avermectinas ainda não está completamente elucidado. Estes compostos ativam os canais de cloreto dos neurotransmissores inibitórios ácido gama-aminobutírico e glutamato, provocando a paralisia dos parasitas (DUCE & SCOTT, 1985; TURNER & SCHAEFFER, 1989; ARENA et al., 1995). Como o mecanismo de ação não é específico, as avermectinas podem afetar organismos não alvos (MC KELLAR, 1997). As

avermectinas não apresentam atividades mutagênicas ou carcinogênicas, mas podem ser embriotóxicas em animais de laboratório (ROUDAUT, 1998). O CODEX ALIMENTARIUS (2005 b) recomenda limites máximos de resíduos de 10 µg/L e 20 µg/L para ivermectina (IVE) e eprinomectina (EPR) em leite, respectivamente, não estabelecendo valores de referência para abamectina (ABA) e doramectina (DOR) nesta matriz. No Brasil, o monitoramento de resíduos destas drogas veterinárias em leite é regulamentado pelo Plano Nacional de Controle de Resíduos (PNCRB), um programa de monitoramento, investigação e fiscalização de resíduos e contaminantes em produtos de origem animal do Ministério da Agricultura, Pecuária e Abastecimento (MAPA). Este programa estabeleceu os limites máximos sugeridos pelo CODEX ALIMENTARIUS para monitorar resíduos avermectinas em leite bovino (BRASIL, 2006).

A confiabilidade dos estudos relativos à presença de resíduos de avermectinas em leite depende crucialmente da eficiência do método de ensaio empregado. Geralmente, os métodos para determinação de avermectinas incluem extração baseada em partição líquido-líquido e purificação por extração em fase sólida. Métodos fundamentados na técnica de cromatografia líquida (CL) com detecção de ultravioleta (UV), fluorescência (FL) e espectrometria de massas (EM) têm sido reportados. Entretanto, as detecções por FL e EM oferecem maior sensibilidade e seletividade que os baseados em detecção UV. CL-FL possui um alto grau de especificidade, devido à combinação da seletividade cromatográfica e de grupos cromogênicos, além de adequada sensibilidade, devida à forte fluorescência dos derivados. CL-FL tem sido preferida também devido à vantagem de ser econômica para uso em atividades de rotina, em detrimento dos altos custos dos métodos que empregam EM (VALENZUELA et al., 2001).

A validação de métodos é o processo que fornece evidências de que um dado método de ensaio produz resultados que são adequados para o propósito de uso (ISO, 2005 b). Independentemente de quão adequado seja o desempenho de um método em um estudo de validação já estabelecido, os analistas necessitam confirmar que o método é válido quando aplicado em seus laboratórios por meio de procedimentos de validação intralaboratoriais. Considerando que processos de validação interlaboratorial principalmente avaliam exatidão e precisão sob condições de repetitividade e reprodutibilidade (HORWITZ, 1995), validações intralaboratoriais são também essenciais para determinação de parâmetros de desempenho como linearidade, efeitos de matriz, seletividade e limites.

Diversos procedimentos de validação intralaboratoriais são descritos na literatura (ICH, 1996; EAL, 1997; NATA, 1997; EURACHEM, 1998; EC, 2002; THOMPSON,

ELLISON & WOOD, 2002, INMETRO, 2003). Embora nestes documentos haja concordância sobre quais parâmetros devam ser avaliados, existe uma grande diversidade e até falta de informações sobre como estes experimentos são delineados e avaliados, em termos de estatísticas e respectivas premissas que necessitam ser verificadas.

Considerando que processos de validação incompletos ou equivocados são importantes fontes de erro em análises de alimentos, o presente trabalho apresenta um procedimento detalhado para validação intralaboratorial, aplicado a um método quantitativo para ensaio de ABA, DOR, EPR e IVE em leite por cromatografia líquida de alta eficiência com detecção de fluorescência (CLAE-FL).

## **2 MATERIAL E MÉTODOS**

### **2.1 MATERIAL**

- Bulbo conta-gotas.
- Cartucho de extração em fase sólida de 500 mg/6 mL, de octadecilsilano C-18, Accubond II ODS (Agilent Technologies).
- Coluna de fase reversa para CLAE de octadecilsilano C-18, dimensões 250 mm x 4 mm, tamanho de partícula 5 µm, Lichrospher 100 RP18 (Merck).
- Dosador tipo bico de papagaio de 5 mL e 20 mL.
- Filme flexível (Parafilm).
- Frasco de vidro com tampa de rosca de 20 mL.
- Frasco âmbar de fundo cônico de 2 mL.
- Filtro de 0,45 µm Millex (Millipore).
- Grade para tubos de centrífuga de 10 mL e 50 mL.
- Microseringa de 100 µL.
- Pesa padrão ou similar.
- Pipeta tipo Pasteur.
- Reservatório de polipropileno tipo seringa de 50 mL.
- Tubo de centrífuga de 10 mL de vidro, com fundo redondo e tampa.
- Tubo de centrífuga de 15 mL e 50 mL de polipropileno, com tampa de rosca (Falcon).

## 2.2 EQUIPAMENTOS

- Agitador tipo Vortex Mixer (Sybron Thermoline).
- Agitador mecânico horizontal (Ebberbach).
- Balança analítica com resolução de 0,0001 g (Mettler).
- Balão volumétrico de 50; 100 e 200 mL.
- Banho maria com agitação com capacidade para  $40\text{ }^{\circ}\text{C} \pm 5\text{ }^{\circ}\text{C}$ ,  $45\text{ }^{\circ}\text{C} \pm 5\text{ }^{\circ}\text{C}$  e  $60\text{ }^{\circ}\text{C} \pm 5\text{ }^{\circ}\text{C}$ , equipado com sistema de evaporação sob fluxo de nitrogênio (Yamato).
- Chapa aquecedora com capacidade para  $90\text{ }^{\circ}\text{C}$  a  $100\text{ }^{\circ}\text{C}$  (Kontes).
- Congelador (Brastemp) e refrigerador (Electrolux).
- Micropipeta automática de volume variável de 100  $\mu\text{L}$  a 1000  $\mu\text{L}$  (Eppendorf).
- Micropipeta automática de volume variável de 50  $\mu\text{L}$  a 500  $\mu\text{L}$  (Eppendorf).
- Pipeta volumétrica de 2; 5 e 10 mL.
- Proveta de 1000 mL.
- Sistema de cromatografia líquida de alta eficiência Gilson (Villiers le Bel, França), composto por bomba isocrática tipo CG-305, injetor manual Rheodyne 7125 com *loop* de 20  $\mu\text{L}$ , detector de fluorescência tipo CG-432 FL, sistema de aquisição de dados e integrador CG-300.
- Sistema de vácuo para cartuchos de extração em fase sólida tipo (Varian Vac Elut).

## 2.3 REAGENTES E PADRÕES

Todos dos reagentes utilizados eram de grau de pureza apropriado.

Acetonitrila e metanol (grau CLAE) foram obtidos da Vetec (Rio de Janeiro, Brasil). Trietilamina e hexano (grau para análise) foram adquiridos da Merck (Darmstadt, Alemanha). 1-metilimidazol (grau para análise) foi da Acros Organic (New Jersey, EUA) e anidrido acético e N,N-dimetilformamida (grau para análise) da Fisher (New Jersey, EUA). Água destilada e desionizada foi preparada no laboratório.

Os padrões de ABA, EPR e IVE foram fornecidos pela Merck (Darmstadt, Alemanha), enquanto o padrão de DOR foi da Pfizer (Groton, EUA).



## 2.4 SOLUÇÕES

### 2.4.1 Solução padrão de ABA

A solução estoque de ABA de 10 µg/mL foi preparada por dissolução de 60,6 mg ± 0,1 mg do padrão de ABA (0,819 % m/m *B1a*) para 50 mL de metanol. Esta solução foi armazenada sob congelamento entre -14 °C e -20 °C e preparada em frequência anual.

### 2.4.2 Soluções padrão de DOR

A solução estoque de DOR de 100 µg/mL foi preparada por dissolução de 10,4 mg ± 0,1 mg do padrão de DOR (95,6 %) para 100 mL de metanol. Uma solução intermediária de 10 µg/mL foi preparada por diluição de 10 mL da solução estoque para 100 mL com metanol. Estas soluções foram armazenadas sob congelamento entre -14 °C e -20 °C e preparadas em frequência anual e semestral, respectivamente.

### 2.4.3 Soluções padrão de EPR

A solução estoque de EPR de 100 µg/mL foi preparada por dissolução de 10,30 mg ± 0,1 mg do padrão de EPR (97,3 %) para 100 mL de metanol. Uma solução intermediária de 10 µg/mL foi preparada por diluição de 10 mL da solução estoque para 100 mL com metanol. Estas soluções foram armazenadas sob congelamento entre -14 °C e -20 °C e preparadas em frequência anual e semestral, respectivamente.

### 2.4.4 Solução padrão de IVE

A solução estoque de IVE de 10 µg/mL foi preparada por dissolução de 35,95 mg ± 0,1 mg do padrão de IVE (1,39 % m/m *H2B1a*) para 50 mL de metanol. Esta solução foi armazenada sob congelamento entre -14 °C e -20 °C e preparada em frequência anual.

### 2.4.5 Solução padrão de adição *pool* de ABA, DOR, EPR e IVE

A solução de adição *pool* das avermectinas de 250 ng/mL foi preparada por diluição de 5 mL da solução estoque de ABA, 5 mL da solução intermediária de DOR, 5 mL da solução intermediária de EPR e 5 mL da solução estoque de IVE para 200 mL de metanol. Esta solução foi armazenada sob congelamento entre -14 °C e -20 °C e preparada em frequência semestral.

#### **2.4.6 Solução de derivação**

Foram misturados 200 µL de 1-metilimidazol a 600 µL de anidrido acético e 900 µL de N,N-dimetilformamida e homogeneizados. Esta solução foi preparada imediatamente antes de cada uso.

#### **2.4.7 Solução de extração**

Foram misturados 700 mL de água destilada e desionizada com 300 mL de acetonitrila e 1 mL de trietilamina.

### **2.5 AMOSTRAS**

Amostras brancas de leite bovino foram obtidas de ordenhas de animais sabidamente não tratados com avermectinas na fazenda experimental do MAPA / Escola de Veterinária da UFMG, na cidade de Pedro Leopoldo (Minas Gerais, Brasil). Considera-se amostras brancas aquelas matrizes isentas do analito ou para as quais o analito não tenha sido detectado.

As amostras foram homogeneizadas, acondicionadas em frascos de vidro com tampa de rosca devidamente identificados e armazenadas sob congelamento entre -14°C e -20°C, até o momento dos ensaios.

### **2.6 MÉTODO DE ENSAIO**

O método validado foi baseado em procedimentos previamente descritos (SALISBURY, 1993; ROUDAUT, 1998; SOUZA et al., 2003).

#### **2.6.1 Princípio**

O método de ensaio validado fundamentou-se na extração de resíduos de ABA, DOR, EPR e IVE do leite com uma mistura acetonitrila:água destilada e desionizada:trietilamina, seguida de purificação do extrato e eluição das avermectinas em cartucho de C-18 com acetonitrila. Após a derivação, o extrato foi solubilizado em metanol e os resíduos foram separados dos componentes co-extraídos, detectados e quantificados por CLAE com detector FL.

#### **2.6.2 Cuidados**

Equipamentos de proteção individual (avental, óculos, luvas e máscaras) e equipamentos de proteção coletiva como capelas de exaustão química foram utilizados durante a realização dos ensaios.

### **2.6.3 Extração**

As amostras de leite foram descongeladas e homogeneizadas. Alíquotas de 5 mL  $\pm$  5 % foram pipetadas e transferidas para tubos de centrífuga de 50 mL de polipropileno, sendo adicionados 20 mL da solução de extração. Os tubos foram colocados em agitador mecânico horizontal ajustado na velocidade máxima, aproximadamente 180 oscilações/min, durante 5 min. Os extratos homogeneizados foram aquecidos em banho maria a 40 °C  $\pm$  5 °C, sob agitação, para completa dissolução da gordura. Brancos de solventes, sem adição de amostra, foram preparados em cada bateria analítica.

### **2.6.4 Acondicionamento dos cartuchos de extração em fase sólida**

Os cartuchos de extração adaptados aos reservatórios tipo seringa de 50 mL foram conectados ao sistema de vácuo. Em seguida, foram lavados com 5 mL  $\pm$  5 % de acetonitrila e com 5 mL  $\pm$  5 % da solução de extração, sob fluxo de 0,3 mL/min.

### **2.6.5 Purificação**

Os extratos de amostras e brancos de solventes foram transferidos para os reservatórios tipo seringas adaptados aos cartuchos de extração em fase sólida, com auxílio de pipetas tipo Pasteur, a um fluxo de 0,3 mL/min. Os reservatórios e cartuchos foram lavados com 5 mL  $\pm$  5 % da solução de extração, deixados secar, e novamente lavados com 2 mL  $\pm$  5 % de hexano, sob fluxo de 0,3 mL/min. Os resíduos das avermectinas foram eluídos com 5 mL  $\pm$  5 % de acetonitrila para tubos de 10 mL de vidro, sob fluxo da gravidade. Os extratos purificados foram evaporados até secura em banho maria a 60 °C  $\pm$  5 °C, sob agitação e fluxo de nitrogênio.

### **2.6.6 Preparo da curva de calibração**

Alíquotas da solução padrão de adição foram pipetadas e transferidas para tubos de 10 mL de vidro. As soluções foram evaporadas até secura em banho maria a 45 °C  $\pm$  5 °C, sob agitação e fluxo de nitrogênio.

### **2.6.7 Derivação**

Os resíduos (de amostras e de soluções padrão da curva de calibração) foram retomados com 200  $\mu$ L  $\pm$  2 % da solução de derivação. Os tubos foram fechados, homogeneizados em agitador tipo Vortex Mixer durante 30 s e colocados em chapa aquecedora entre 90 °C e 100 °C durante 1 h. Após resfriamento à temperatura ambiente,

os resíduos derivados foram retomados em  $1000 \mu\text{L} \pm 2 \%$  de metanol e filtrados através de filtros de  $0,45 \mu\text{m}$  para frascos de 2 mL.

### 2.6.8 Separação, detecção e quantificação

As condições de operação do sistema CLAE-FL foram:

- volume de injeção de  $20 \mu\text{L}$ ;
- fase móvel composta por metanol;
- sistema isocrático;
- fluxo de  $0,9 \text{ mL/min}$ ;
- comprimento de onda de excitação e emissão de  $366 \text{ nm}$  e  $426 \text{ nm}$ , respectivamente;
- tempo total de corrida de  $12 \text{ min}$ ;
- integração pelo parâmetro da altura dos picos.

Curvas de calibração da resposta (altura dos picos) *versus* a concentração de cada avermectina foram construídas. As equações das curvas foram estimadas pelo método dos mínimos quadrados ordinários (MMQO) e utilizadas para cálculo da concentração de avermectinas nas amostras.

## 2.7 VALIDAÇÃO

Os parâmetros linearidade, efeitos de matriz, seletividade, exatidão, precisão, limite de detecção, limite de quantificação, limite de decisão ( $CC\alpha$ ) e capacidade de detecção ( $CC\beta$ ) foram estabelecidos em ensaios com soluções padrão, amostras brancas e adicionadas. A adequação para uso do método foi avaliada em função dos parâmetros estudados e respectivos critérios de aceitabilidade definidos. O nível de significância adotado nos testes de hipóteses foi  $\alpha = 0,05$ . Os cálculos foram feitos em planilhas comerciais, cujos resultados foram previamente validados frente aqueles obtidos em diferentes softwares de análise estatística.

### 2.7.1 Linearidade

Três curvas de calibração foram preparadas, em diferentes dias, com níveis de concentração  $25$ ;  $50$ ;  $75$ ;  $100$ ;  $125$  e  $150 \text{ ng/mL}$  (correspondentes a  $5$ ;  $10$ ;  $15$ ;  $20$ ;  $25$  e  $30 \mu\text{g/L}$  de cada avermectina em leite bovino), sendo três replicatas independentes de cada nível. Alíquotas de  $100$ ;  $200$ ;  $300$ ;  $400$ ;  $500$  e  $600 \mu\text{L}$ , respectivamente, da solução padrão *pool* de adição foram utilizadas para preparo destas curvas conforme descrito no método de ensaio. As soluções foram analisadas em ordem aleatória dentro de cada dia.

Branco da curva foram preparados, em triplicata, como ferramenta de controle de qualidade, avaliação do ruído e ajuste do zero instrumental, mas não foram incluídos nos cálculos.

Os parâmetros da regressão foram estimados pelo MMQO. Em seguida, gráficos dos resíduos (erros) da regressão foram construídos e examinados para investigação de perfis óbvios que demonstrassem heteroscedasticidade ou desvio de linearidade, sendo indicados como valores extremos (*outliers*) os pontos fora do intervalo  $\pm t_{(1-\alpha/2;n-2)} s_{res}$ , baseado no método proposto por MEYER & ZUND (1993), sendo  $s_{res}$  o desvio padrão dos resíduos da regressão. Os valores extremos foram formalmente diagnosticados pelo teste de resíduos padronizados Jackknife (BELSLEY, KUH & WELSCH, 1980). Este teste foi aplicado sucessivamente até que novos valores extremos não fossem detectados ou até uma exclusão máxima de 22,2 % no número original de resultados (HORWITZ, 1995). As premissas relativas à análise de regressão foram verificadas: normalidade (RYAN & JOINER, 1976), homoscedasticidade (LEVENE, 1960; BROWN & FORSYTHE, 1974) e independência dos resíduos da regressão (DURBIN & WATSON, 1951). Testes de F foram adotados para verificar a adequação do ajuste ao modelo linear por meio das significâncias da regressão e do desvio da linearidade avaliado contra o erro puro (DRAPER & SMITH, 1998).

### 2.7.2 Efeitos de matriz

O método de adição foi empregado para avaliação dos efeitos de matriz. Duas curvas de calibração foram estudadas em níveis de 25; 50; 75; 100; 125 e 150 ng/mL, sendo cada nível em duas replicatas independentes, analisadas em ordem aleatória, no mesmo dia. A curva do analito no solvente (metanol) foi preparada conforme descrito nos experimentos de avaliação da linearidade. Enquanto a curva do analito na matriz utilizou extratos purificados de amostras brancas de leite bovino adicionados de 100; 200; 300; 400; 500 e 600  $\mu\text{L}$  da solução padrão *pool* de adição. Branco da curva foram preparados e utilizados como ferramenta de controle de qualidade, avaliação do ruído e ajuste do zero instrumental, mas não foram incluídos nos cálculos.

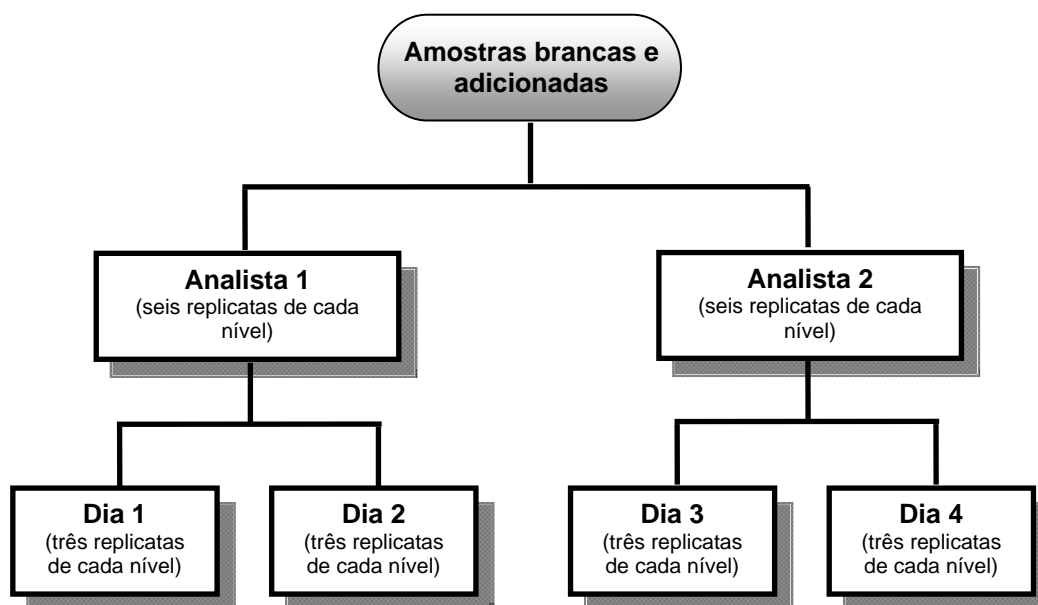
Os parâmetros da regressão foram estimados e as premissas relativas ao MMQO também foram verificadas para estas curvas. Uma vez comprovadas as premissas de normalidade a ajuste ao modelo linear, a inclinação e a interseção obtidas para a curva de solventes foram comparadas com a inclinação e interseção calculadas para a curva de matriz. A premissa de homogeneidade das variâncias dos resíduos das curvas foi

avaliada pelo teste de F (SNEDECOR & COCHRAN, 1989), para verificar se as inclinações e interseções seriam comparadas pelo teste de t com as variâncias combinadas (ARMITAGE & BERRY, 1994).

### 2.7.3 Seletividade, exatidão, precisão e limites

Seletividade, exatidão, precisão (sob condições de repetitividade e reprodutibilidade parcial), limites de detecção e quantificação,  $CC\alpha$  e  $CC\beta$  foram pesquisados em ensaios com amostras brancas e adicionadas em níveis correspondentes a 5; 10; 15; 20 e 30  $\mu\text{g}/\text{kg}$ , sendo doze replicatas independentes de cada nível. As amostras foram divididas em quatro lotes, analisados em quatro diferentes dias por dois analistas (Figura IV.2). As concentrações de avermectinas nas amostras adicionadas foram previstas pelas equações das curvas de calibração de solventes, com base nos resultados dos estudos de efeitos de matriz.

A ausência de resultados falsos positivos para as amostras brancas associada à capacidade do método separar, detectar e quantificar os quatro resíduos de avermectinas (ABA, DOR, EPR e IVE) sem interferência mútua, foram consideradas suficientes para estabelecimento da seletividade.



**Figura IV.2** - Delineamento da distribuição de amostras de leite bovino nos ensaios de avaliação da exatidão e precisão do método para determinação de avermectinas em leite.

Os resultados de recuperação aparente obtidos foram analisados quanto a presença de valores extremos pelo teste de Grubbs (GRUBBS, 1969; BARRET & LEWIS,

1994; BURKE 2001 b). As estatísticas de Grubbs para um valor extremo ( $G_1$ ), dois valores vicinais ( $G_2$ ) e dois valores polares ( $G_3$ ) foram calculadas simultaneamente para os resultados obtidos para cada nível de concentração. A indicação de valor extremo por, pelo menos, uma destas estatísticas foi suficiente para excluir o dado. Após um primeiro ciclo, o teste de Grubbs foi sucessivamente aplicado até que novos valores extremos não fossem detectados ou até uma exclusão máxima de 22,2 % no número original de resultados de recuperação aparente (HORWITZ, 1995).

Exatidão foi investigada por meio das médias de recuperação aparente obtida para as doze replicatas de amostras adicionadas em cada nível de concentração. Foram consideradas aceitáveis médias de recuperação aparente entre 70 % e 110 % para amostras adicionadas de 5 µg/L e de 80 % a 110 % para as amostras adicionadas nos níveis de 10; 15; 20 e 30 µg/L (EC, 2002).

As precisões sob condições de repetitividade e reprodutibilidade parcial foram expressas em termos de desvios padrão relativos e estimadas por análise de variância dos resultados de recuperação aparente obtidos para as doze replicatas de amostras adicionadas em cada nível de concentração (ISO, 1994 a,b,c,d,e; ISO, 1998; KUTTATHARMMAKUL, MASSART & SMEYERS-VERBEKE, 1999). Foi adotado um delineamento fatorial completamente aninhado (*fully-nested design*) que forneceu as estimativas da variância de repetitividade ( $s_r^2$ ) e da variância do ensaio ( $s_{Ensaio}^2$ ), que inclui as fontes de variação confundidas do dia e do analista (dia + analista). A variância entre ensaios [ $s_{entre\ ensaios}^2 = (s_{Ensaio}^2 - s_r^2) / n$ ] foi estimada, sendo  $n$  o número médio de replicatas em cada nível de concentração, para cálculo da variância de reprodutibilidade parcial ( $s_R^2 = s_r^2 + s_{entre\ ensaios}^2$ ) (MAROTO et al, 1999; MAROTO et al, 2001). Quando a estimativa da variância de repetitividade foi maior que a variância do ensaio,  $s_{entre\ ensaios}^2$  foi considerada como sendo igual a zero, evitando a obtenção de um valor negativo para a estimativa desta variância (ISO, 1994; KUTTATHARMMAKUL, MASSART & SMEYERS-VERBEKE, 1999). Neste caso, as estimativas das variâncias de repetitividade e de reprodutibilidade parcial foram consideradas equivalentes.

As premissas de normalidade (RYAN & JOINER, 1976) e homoscedasticidade dos resíduos dos resultados de recuperação aparente (LEVENE, 1960, BROWN & FORSYTHE, 1974), relativas à análise de variância, foram previamente testadas. Desvios padrão relativos de repetitividade ( $DPR_r$ ) foram considerados aceitáveis quando dentro de dois terços (EC, 2002) o valor de precisão interlaboratorial estimado pela equação de HORWITZ (1982) modificada por THOMPSON (2000). Já os desvios padrão relativos de

reprodutibilidade parcial ( $DPR_R$ ) foram considerados aceitáveis quando não excederam o valor estimado pelas equações de HORWITZ (1982) e THOMPSON (2000). Desta forma,  $DPR_r$  e  $DPR_R$  menores ou iguais a 14,7 % e 22 %, respectivamente, foram os critérios de aceitabilidade adotados.

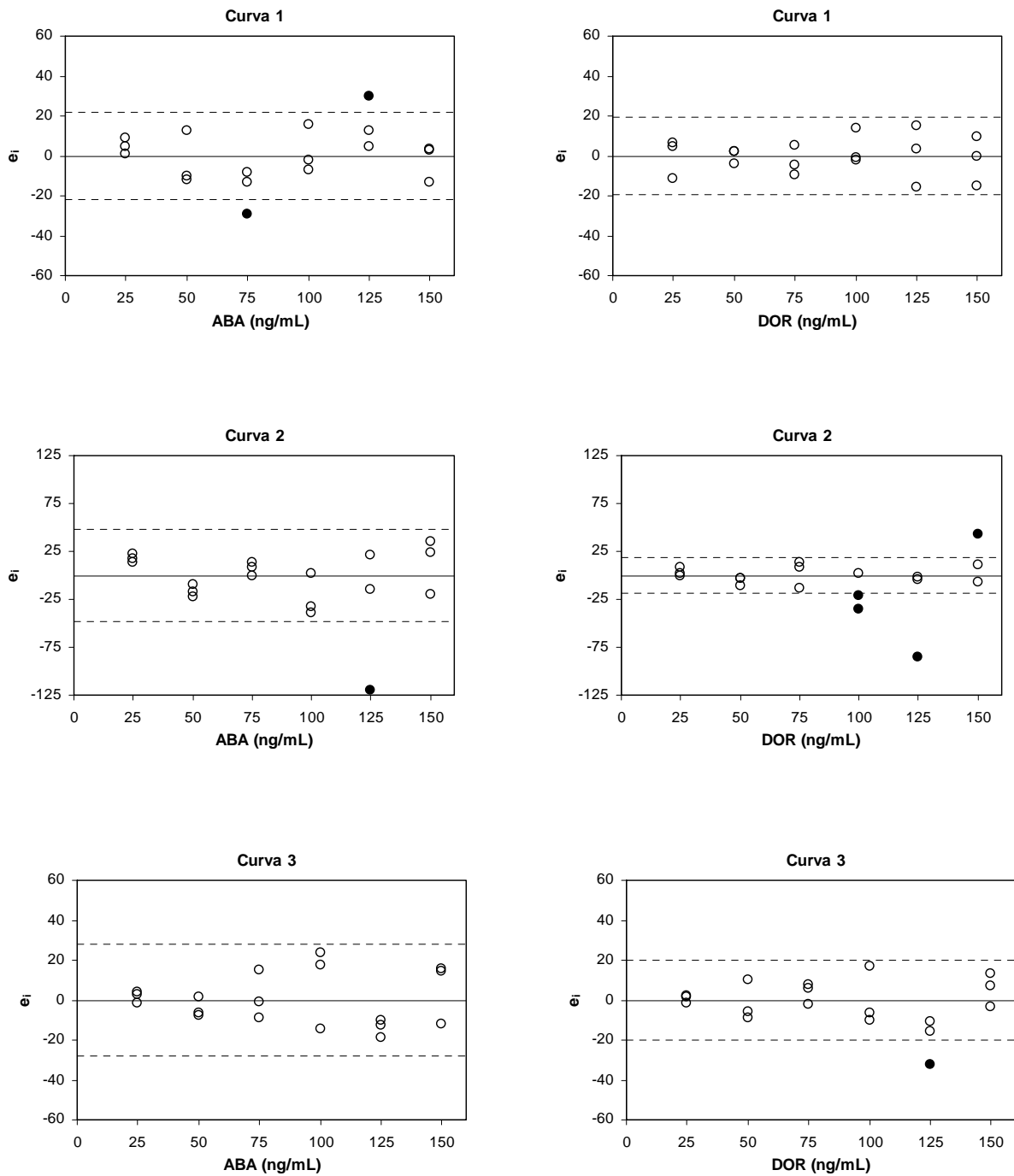
O limite de quantificação foi estabelecido como a concentração abaixo da qual o método não pôde operar com exatidão e precisão aceitáveis. O limite de detecção foi a menor concentração detectada nas amostras adicionadas mas não necessariamente quantificada, distinguida de zero (sinal da amostra adicionada / ruído da amostra branca  $\geq 3$ ). Estes limites foram estabelecidos com base nos resultados médios de recuperação aparente e desvios padrão relativos de repetitividade e reprodutibilidade parcial obtidos para as replicatas de amostras adicionadas nos diferentes níveis de concentração estudados.

$CC\alpha$  e  $CC\beta$  foram estimados (ISO, 1997 b; ISO, 2000; VAN LOCO & BEERNAERT, 2003) utilizando as concentrações estimadas e as respectivas concentrações teóricas das amostras adicionadas dos experimentos de avaliação da reprodutibilidade parcial (EC, 2002). O limite máximo de resíduo de 10  $\mu\text{g/L}$  foi adotado para o cálculo destes parâmetros para ABA, DOR, EPR e IVE.

### 3 RESULTADOS E DISCUSSÃO

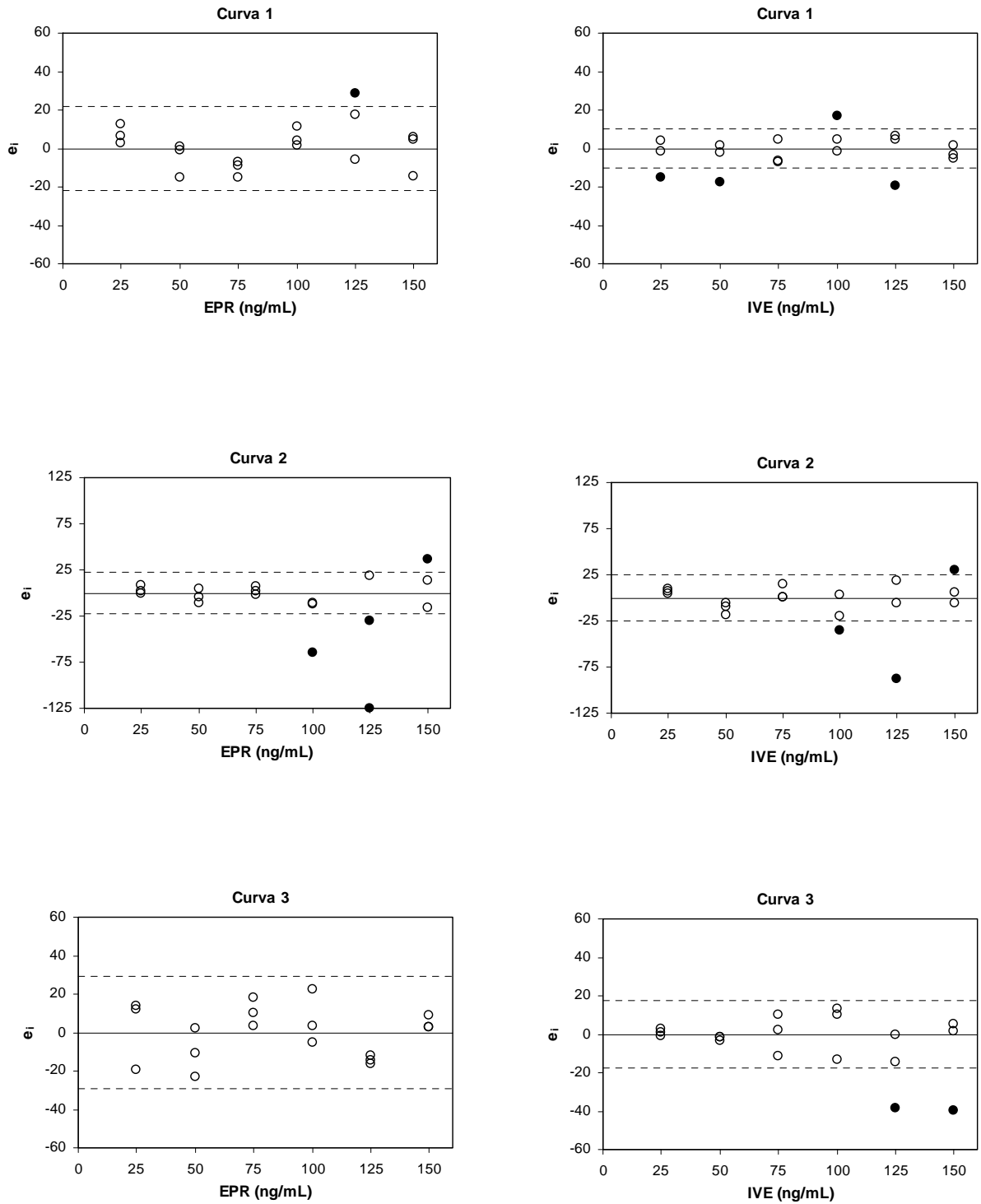
Dois valores extremos foram detectados pelo teste de resíduos padronizados Jackknife para ABA na curva 1 a 75 ng/L e 125 ng/mL e um na curva 2 a 125 ng/mL. Não foram identificados valores extremos na curva 3 para este composto. Para DOR, não foram observados valores extremos na curva 1, quatro valores foram detectados na curva 2 a 100; 125 e 150 ng/mL e um na curva 3 a 125 ng/mL. No caso da EPR, um valor extremo foi observado na curva 1 a 125 ng/mL, quatro na curva 2 a 100; 125 e 150 ng/mL, sendo que não foram identificados valores extremos na curva 3. Para IVE, quatro valores extremos foram tratados na curva 1 a 25; 50; 100 e 125 ng/mL, três na curva 2 a 100; 125 e 150 ng/mL e dois na curva 3 a 125 ng/mL e 150 ng/mL. Estes resultados confirmaram as indicações de valores extremos pelo critério  $\pm t_{(1-\alpha/2; n-2)} S_{res}$ . Em nenhum caso houve indicação de um número maior que quatro valores extremos, correspondente ao limite máximo de 22,2 % dos dezoito dados originais. Os perfis dos gráficos exploratórios apresentados nas Figuras IV.3 e IV.4 indicaram que não houve tendências óbvias na distribuição dos resíduos da regressão após tratamento dos valores extremos.





**Figura IV.3** - Gráficos exploratórios dos resíduos da regressão das três curvas de ABA e DOR (25 ng/mL a 150 ng/mL) em solvente, com indicação dos respectivos valores extremos diagnosticados pelo teste de resíduos padronizados Jackknife.

$e_i$  = resíduo da regressão, • = valor extremo, - - - =  $\pm t_{(1-\alpha/2; n-2)} s_{res}$ .



**Figura IV.4** - Gráficos exploratórios dos resíduos da regressão das três curvas de EPR e IVE (25 ng/mL a 150 ng/mL) em solvente, com indicação dos respectivos valores extremos diagnosticados pelo teste de resíduos padronizados Jackknife.

$e_i$  = resíduo da regressão, • = valor extremo, - - - =  $\pm t_{(1-\alpha/2; n-2)} s_{res}$ .

Os resultados obtidos nos experimentos para avaliação da linearidade para ABA, DOR, EPR e IVE estão demonstrados nas Tabelas IV.1, IV.2, IV.3 e IV.4, respectivamente. As premissas de que os resíduos da regressão seguem a distribuição normal, são homoscedásticos e independentes foram confirmadas, garantindo uma aplicação segura dos testes de hipóteses de t e F, além do emprego adequado da análise de variância sem ponderações para estimativa da significância da regressão e do desvio da linearidade. Os coeficientes de correlação de Ryan-Joiner indicaram que os desvios da normalidade não foram significativos ( $p > 0,10$ ). A variância dos erros ao longo dos níveis de concentração estimada pelo teste de Levene modificado também não foi significativa ( $p > 0,05$ ), sugerindo homoscedasticidade. Autocorrelação dos resíduos da regressão não foi observada ( $p > 0,10$ ), visto que as estatísticas de Durbin-Watson demonstraram independência dos mesmos.

**Tabela IV.1** - Avaliação da linearidade para as três curvas de ABA (25 ng/mL a 150 ng/mL) em solvente

Estatística	Curva 1	Curva 2	Curva 3
<i>Número de observações</i>			
$n$	16	17	18
<i>Normalidade</i>			
$R$	0,9754	0,9849	0,9740
$p$	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Homoscedasticidade</i>			
$t_L$	0,690	-1,959	-1,975
$p$	0,50	0,07	0,07
<i>Independência</i>			
$d$	2,139	1,803	1,525
$p$	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Regressão</i>			
$F$	$2,435 \times 10^{+3}$	$3,870 \times 10^{+2}$	$1,886 \times 10^{+3}$
$p$	$4,18 \times 10^{-17}$	$4,01 \times 10^{-12}$	$4,95 \times 10^{-18}$
<i>Desvio da linearidade</i>			
$F$	1,365	3,169	1,715
$p$	0,31	0,06	0,21

$n$  = número de observações,  $R$  = coeficiente de correlação de Ryan-Joiner,  $p$  = significância,  $t_L$  = estatística t de Levene,  $d$  = estatística de Durbin-Watson,  $F$  = razão entre variâncias.

**Tabela IV.2** - Avaliação da linearidade para as três curvas de DOR (25 ng/mL a 150 ng/mL) em solvente

Estatística	Curva 1	Curva 2	Curva 3
<i>Número de observações</i>			
<i>n</i>	18	14	17
<i>Normalidade</i>			
<i>R</i>	0,9898	0,9899	0,9921
<i>p</i>	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Homoscedasticidade</i>			
<i>t<sub>L</sub></i>	-1,139	0,733	-1,622
<i>p</i>	0,27	0,48	0,13
<i>Independência</i>			
<i>d</i>	2,342	2,497	1,980
<i>p</i>	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Regressão</i>			
<i>F</i>	$2,091 \times 10^{+3}$	$1,143 \times 10^{+3}$	$2,160 \times 10^{+3}$
<i>p</i>	$2,18 \times 10^{-18}$	$2,86 \times 10^{-13}$	$1,25 \times 10^{-17}$
<i>Desvio da linearidade</i>			
<i>F</i>	0,177	0,563	1,560
<i>p</i>	0,95	0,70	0,25

*n* = número de observações, *R* = coeficiente de correlação de Ryan-Joiner, *p* = significância, *t<sub>L</sub>* = estatística t de Levene, *d* = estatística de Durbin-Watson, *F* = razão entre variâncias.

Os dados obtidos foram avaliados como bem ajustados ao modelo linear. Significância da regressão ( $p < 0,001$ ) e desvios de linearidade não significativos ( $p > 0,05$ ) indicaram linearidade na faixa de 25 ng/mL a 150 ng/mL para as curvas 1, 2 e 3 das quatro avermectinas examinadas. Se os testes de avaliação das premissas fossem conduzidos sem tratamento prévio de valores extremos, alguns problemas seriam observados na aplicação do MMQO. O ajuste ao modelo foi claramente afetado pelos valores extremos detectados e removidos. Heteroscedasticidade ( $p > 0,05$ ) seria observada para a curva 2 de EPR e para a curva 3 de IVE e desvio da linearidade ( $p > 0,05$ ) seria detectado para a curva 1 de ABA se os valores fossem mantidos.

O uso de critérios equivocados como o coeficiente de correlação (*r*) (THOMPSON, ELLISON & WOOD, 2002) é predominante na avaliação da linearidade de métodos para determinação de avermectinas, assim como a omissão de estudos de efeitos de matriz. POLLMEIER et al. (2002) indicaram linearidade de um método para ensaio de EPR em

**Tabela IV.3** - Avaliação da linearidade para as três curvas de EPR (25 ng/mL a 150 ng/mL) em solvente

Estatística	Curva 1	Curva 2	Curva 3
<i>Número de observações</i>			
<i>n</i>	17	14	18
<i>Normalidade</i>			
<i>R</i>	0,9846	0,9909	0,9836
<i>p</i>	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Homoscedasticidade</i>			
<i>t<sub>L</sub></i>	0,329	-1,579	0,520
<i>p</i>	0,75	0,14	0,61
<i>Independência</i>			
<i>d</i>	1,950	2,802	1,212
<i>p</i>	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Regressão</i>			
<i>F</i>	$3,117 \times 10^{+3}$	$1,669 \times 10^{+3}$	$2,062 \times 10^{+3}$
<i>p</i>	$8,12 \times 10^{-19}$	$2,99 \times 10^{-14}$	$2,43 \times 10^{-18}$
<i>Desvio da linearidade</i>			
<i>F</i>	2,473	2,125	2,930
<i>p</i>	0,11	0,17	0,07

*n* = número de observações, *R* = coeficiente de correlação de Ryan-Joiner, *p* = significância, *t<sub>L</sub>* = estatística t de Levene, *d* = estatística de Durbin-Watson, *F* = razão entre variâncias.

leite bovino por CLAE-FL na faixa de 2,3 ng/mL a 51 ng/mL ( $r = 0,999$ ), utilizando amostras adicionadas. ROUDAUT (1998) também determinou linearidade empregando CLAE-FL para ABA, DOR, IVE e moxidectina (MOX) utilizando amostras de fígado bovino adicionadas na faixa de 7,5 µg/kg a 200 µg/kg ( $r \geq 0,993$ ). WU et al. (2001) estabeleceram linearidade para ABA ( $r = 0,9980$ ) e IVE ( $r = 0,9996$ ) por CLAE-EM na faixa de 10 ng/mL a 500 ng/mL, enquanto SALISBURY (1993) estudou IVE por CLAE-FL entre 25 ng/mL e 150 ng/mL, ambos empregando curvas de solventes.

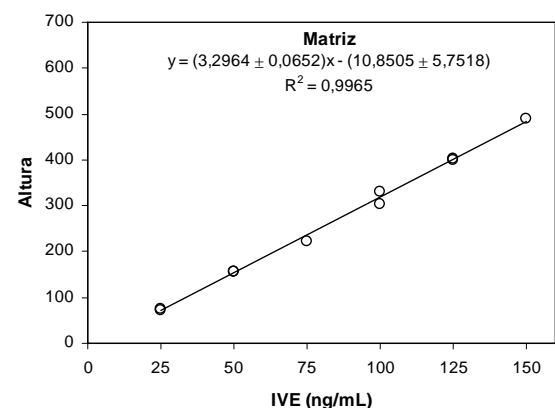
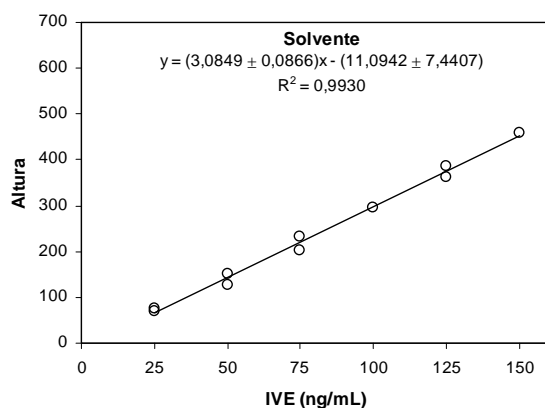
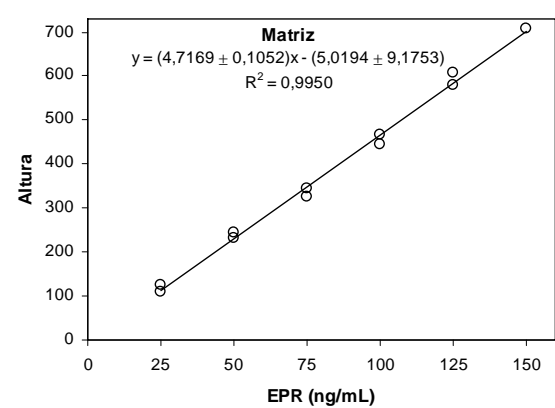
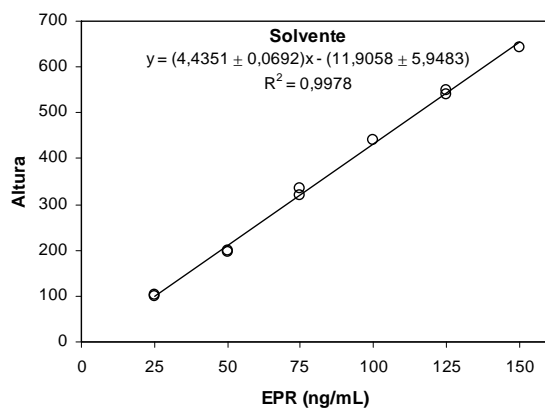
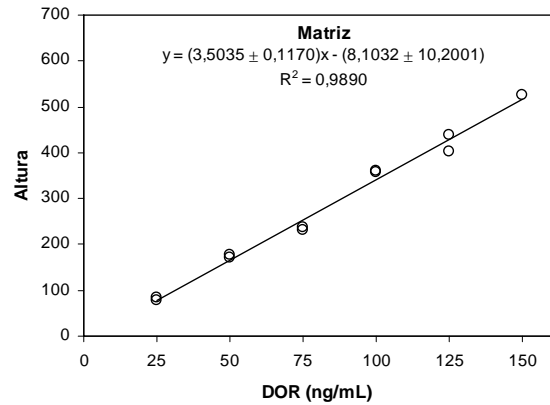
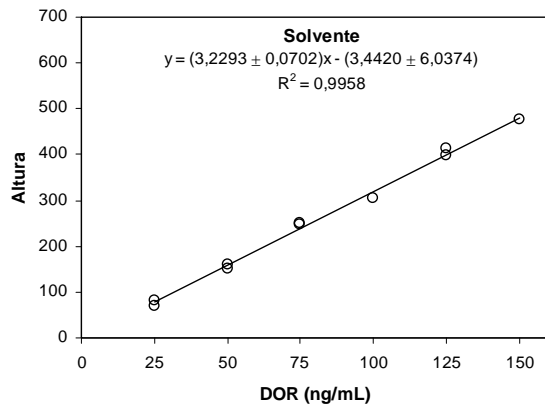
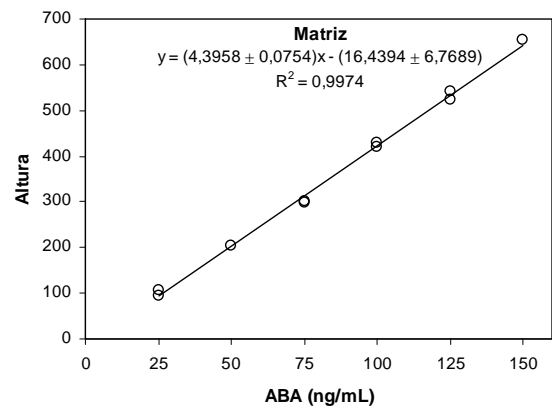
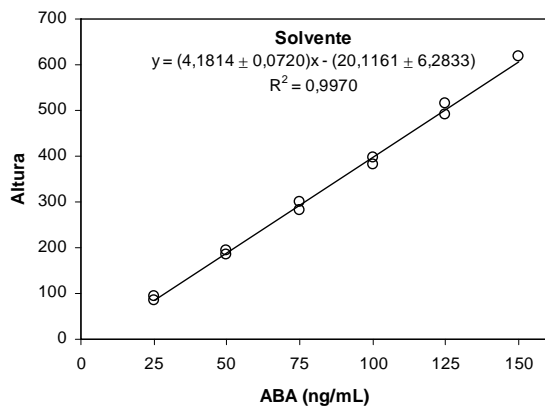
As curvas de solventes e de matriz, com respectivas equações e coeficientes de determinação, obtidas nos estudos de efeitos de matriz estão apresentadas na Figura IV.5. Pelo teste de resíduos padronizados Jackknife foram detectados um valor extremo na curva de solventes (150 ng/mL) e dois na curva de matriz (50 ng/mL e 150 ng/mL) para ABA; dois valores extremos na curva de solventes (100 ng/mL e 150 ng/mL) e um na curva de matriz (150 ng/mL) para DOR e EPR;

**Tabela IV.4** - Avaliação da linearidade para as três curvas de IVE (25 ng/mL a 150 ng/mL) em solvente

Estatística	Curva 1	Curva 2	Curva 3
<i>Número de observações</i>			
<i>n</i>	14	15	16
<i>Normalidade</i>			
<i>R</i>	0,9763	0,9889	0,9710
<i>p</i>	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Homoscedasticidade</i>			
<i>t<sub>L</sub></i>	$1,322 \times 10^{-14}$	-0,625	-1,592
<i>p</i>	1,00	0,54	0,13
<i>Independência</i>			
<i>d</i>	1,455	1,719	2,114
<i>p</i>	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Regressão</i>			
<i>F</i>	$5,322 \times 10^{+3}$	$6,445 \times 10^{+2}$	$2,237 \times 10^{+3}$
<i>p</i>	$2,93 \times 10^{-17}$	$1,84 \times 10^{-12}$	$7,55 \times 10^{-17}$
<i>Desvio da linearidade</i>			
<i>F</i>	1,496	2,256	0,580
<i>p</i>	0,29	0,14	0,68

*n* = número de observações, *R* = coeficiente de correlação de Ryan-Joiner, *p* = significância, *t<sub>L</sub>* = estatística t de Levene, *d* = estatística de Durbin-Watson, *F* = razão entre variâncias.

e dois valores extremos na curva de solventes (100 ng/mL e 150 ng/mL) e dois na curva de matriz (75 ng/mL e 150 ng/mL) para IVE. Todas as premissas relativas ao MMQO também foram ratificadas para as curvas de solventes e matriz dos quatros analitos (Tabelas IV.5 e IV.6). Distribuição normal dos resíduos da regressão foi observada pelos coeficientes de correlação de Ryan-Joiner. Homoscedasticidade foi obtida com estatísticas t de Levene não significativas ( $p > 0,05$ ). Independência dos resíduos da regressão foi indicada pelas estatísticas de Durbin-Watson ( $p > 0,10$ ). Regressões significativas ( $p < 0,001$ ) e desvios de linearidade não significativos ( $p > 0,05$ ) foram diagnosticados pelos testes de F da análise de variância. Estes resultados de confirmação do modelo linear indicaram a possibilidade de comparação das inclinações e interseções pelo teste de t para avaliação dos efeitos de matriz.



**Figura IV.5** - Curvas de calibração dos analitos no solvente e na matriz obtidas para ABA, DOR, EPR e IVE nos ensaios de avaliação de efeitos de matriz, na faixa de 25 ng/mL a 150 ng/mL, com respectivas equações e coeficientes de determinação.

$y$  = resposta em altura dos picos,  $x$  = concentração de avermectina em ng/mL,  $R^2$  = coeficiente de determinação.

**Tabela IV.5** - Avaliação da linearidade para curvas de ABA e DOR (25 ng/mL a 150 ng/mL) em solvente e matriz obtidas nos ensaios de avaliação de efeitos de matriz

Estatística	ABA		DOR	
	Solvente	Matriz	Solvente	Matriz
<i>Número de observações</i>				
<i>n</i>	11	10	10	11
<i>Normalidade</i>				
<i>R</i>	0,9807	0,9808	0,9784	0,9370
<i>p</i>	$p > 0,10$	$p > 0,10$	$p > 0,10$	$p > 0,05$
<i>Homoscedasticidade</i>				
<i>t<sub>L</sub></i>	-1,351	0,550	-0,298	-0,043
<i>p</i>	0,21	0,60	0,77	0,97
<i>Independência</i>				
<i>d</i>	1,683	1,466	2,314	2,038
<i>p</i>	$p > 0,10$	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Regressão</i>				
<i>F</i>	$3,031 \times 10^{+3}$	$3,025 \times 10^{+3}$	$1,878 \times 10^{+3}$	$8,074 \times 10^{+2}$
<i>p</i>	$1,08 \times 10^{-12}$	$1,33 \times 10^{-11}$	$8,86 \times 10^{-11}$	$4,03 \times 10^{-10}$
<i>Desvio da linearidade</i>				
<i>F</i>	0,537	2,147	2,239	2,891
<i>p</i>	0,72	0,24	0,23	0,14

*n* = número de observações, *R* = coeficiente de correlação de Ryan-Joiner, *p* = significância, *t<sub>L</sub>* = estatística t de Levene, *d* = estatística de Durbin-Watson, *F* = razão entre variâncias.

O teste de F indicou homogeneidade das variâncias dos resíduos das curvas de solventes e matrizes e a possibilidade do uso do teste de t com variâncias combinadas para avaliação dos efeitos de matriz. Assim, não houve diferenças significativas ( $p > 0,05$ ) quando as inclinações e interseções das curvas de solventes foram comparadas com aquelas estimadas para as curvas de matriz para ABA, DOR, EPR e IVE (Tabela IV.7). Com base nestes resultados, foi possível inferir que curvas de solventes de ABA, DOR, EPR e IVE forneceram as mesmas respostas que amostras de leite contendo as avermectinas estudadas nas mesmas faixas de concentração.



**Tabela IV.6** - Avaliação da linearidade para curvas de EPR e IVE (25 ng/mL a 150 ng/mL) em solvente e matriz obtidas nos ensaios de avaliação de efeitos de matriz

Estatística	EPR		IVE	
	Solvente	Matriz	Solvente	Matriz
<i>Número de observações</i>				
<i>n</i>	10	11	10	10
<i>Normalidade</i>				
<i>R</i>	0,9818	0,9739	0,9323	0,9099
<i>p</i>	$p > 0,10$	$p > 0,10$	$p > 0,05$	$p > 0,01$
<i>Homoscedasticidade</i>				
<i>t<sub>L</sub></i>	0,166	-0,234	0,279	-0,496
<i>p</i>	0,87	0,82	0,79	0,63
<i>Independência</i>				
<i>d</i>	1,549	1,725	2,034	1,819
<i>p</i>	$p > 0,10$	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Regressão</i>				
<i>F</i>	$3,650 \times 10^{+3}$	$1,809 \times 10^{+3}$	$1,129 \times 10^{+3}$	$2,273 \times 10^{+3}$
<i>p</i>	$6,26 \times 10^{-12}$	$1,10 \times 10^{-11}$	$6,73 \times 10^{-10}$	$4,15 \times 10^{-11}$
<i>Desvio da linearidade</i>				
<i>F</i>	3,202	1,357	0,147	0,808
<i>p</i>	0,14	0,37	0,95	0,58

*n* = número de observações, *R* = coeficiente de correlação de Ryan-Joiner, *p* = significância, *t<sub>L</sub>* = estatística t de Levene, *d* = estatística de Durbin-Watson, *F* = razão entre variâncias.

**Tabela IV.7** - Comparações entre as interseções e inclinações das curvas de matriz com as das curvas de solventes para ABA, DOR, EPR e IVE (25 ng/mL a 150 ng/mL)

Estatística	ABA	DOR	EPR	IVE
<i>Comparação da interseção com a curva de solvente</i>				
<i>t<sub>a</sub></i>	0,360	0,351	0,556	0,023
<i>p</i>	0,72	0,73	0,59	0,98
<i>Comparação da inclinação com a curva de solvente</i>				
<i>t<sub>b</sub></i>	1,946	1,873	2,090	1,842
<i>p</i>	0,07	0,08	0,05	0,08

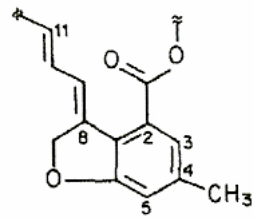
*t<sub>a</sub>* = estatística t para contrastes entre interseções, *t<sub>b</sub>* = estatística t para contrastes entre inclinações, *p* = significância.

Cromatogramas típicos obtidos para soluções padrão de avermectinas, amostras brancas e amostras adicionadas encontram-se apresentados na Figura IV.6, indicando seletividade do método. Os tempos de retenção médios observados foram de 8,37 min  $\pm$  0,37 min para ABA, 9,83 min  $\pm$  0,46 min para DOR, 6,79 min  $\pm$  0,26 min para EPR e 11,17 min  $\pm$  0,56 min para IVE. Resultados não detectados (sinal da amostra branca / ruído do branco da curva  $< 3$ ) foram obtidos para todas as amostras brancas analisadas, não havendo interferências significativas nos tempos de retenção para as avermectinas pesquisadas. Os perfis dos cromatogramas obtidos para amostras adicionadas foram indistinguíveis daqueles obtidos para as soluções padrão de avermectinas de concentrações correspondentes. A resolução dos picos indicou capacidade do método em determinar ABA, DOR, EPR e IVE sem interferências mútuas.

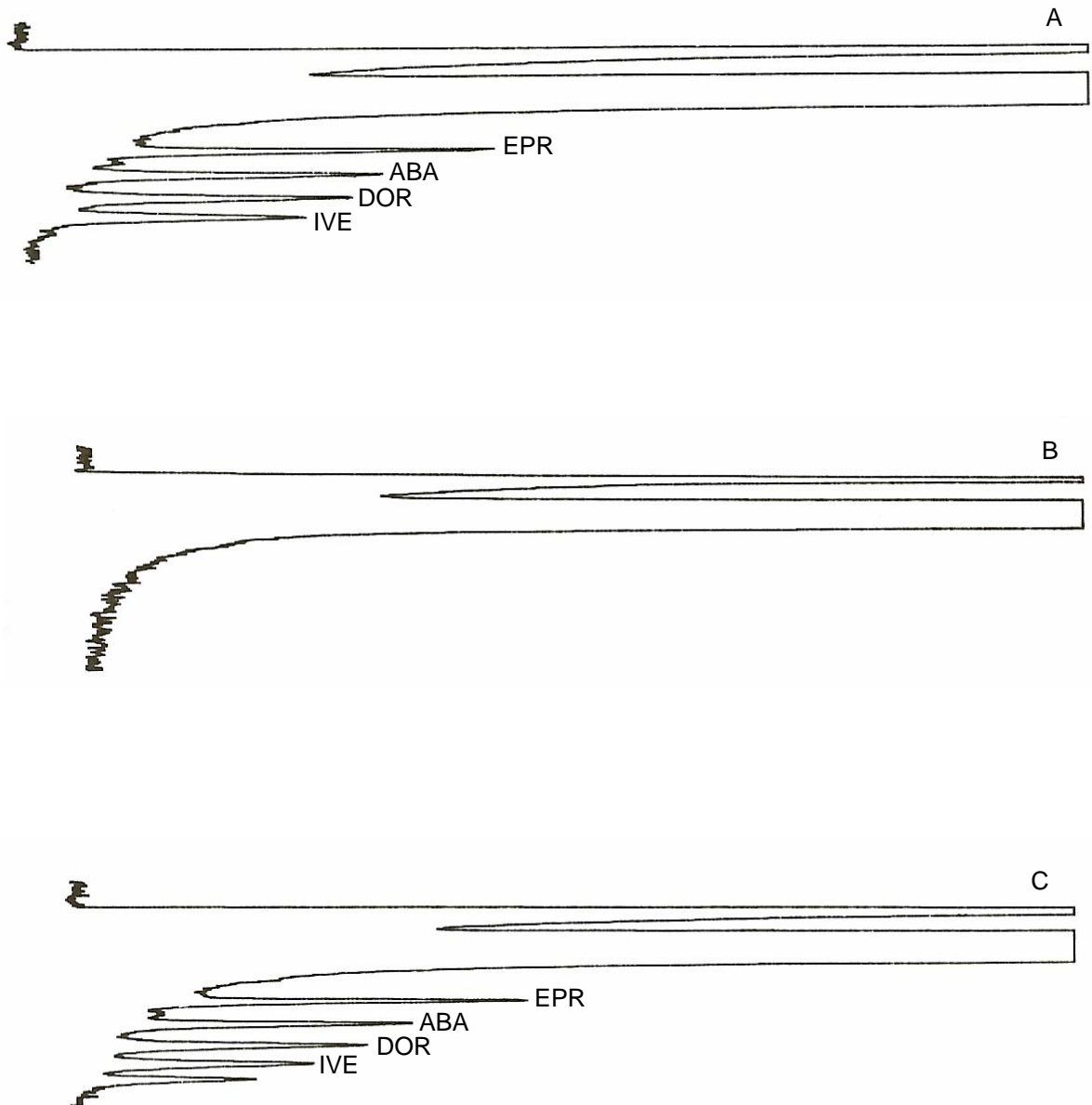
O teste de Grubbs não indicou valores extremos ( $p > 0,05$ ) entre os resultados de recuperação aparente obtidos para amostras adicionadas dos quatro analitos em todos os níveis de concentração estudados. Os resíduos, obtidos pela diferença entre a recuperação aparente média e os valores individuais de recuperação aparente obtidos em cada dia, para cada nível de concentração, apresentaram distribuição normal ( $p > 0,10$ ) e homoscedasticidade ( $p > 0,05$ ) (Tabelas IV.8, IV.9, IV.10 e IV.11), permitindo a estimativa de  $DPR_r$  e  $DPR_R$  por análise de variância.

As médias de recuperação aparente alcançadas para amostras adicionadas entre 5  $\mu\text{g/L}$  e 30  $\mu\text{g/L}$  variaram entre 98,1 % e 101,8 % para ABA; 86,2 % e 98,0 % para DOR; 81,2 % e 96,9 % para EPR; e 84,0 % e 94,8 % para IVE. Os  $DPR_r$  foram de 4,1 % a 13,9 % para ABA; de 3,6% a 21,3 % para DOR; de 5,0 % a 20,2 % para EPR; e de 4,4 % a 16,7 % para IVE. Os valores de  $DPR_R$  estiveram entre 8,0 % e 14,3 % para ABA; 8,7 % e 21,4 % para DOR; 7,4 % e 20,2 % para EPR; e 10,2 % e 19,5 % para IVE (Tabelas IV.12 e IV.13). Valores de  $DPR_r$  e  $DPR_R$  idênticos foram estimados para EPR no nível de 5  $\mu\text{g/L}$  uma vez que a estimativa da variância de repetitividade foi maior que a do ensaio (dia+analista), sendo atribuído o valor de zero para a variação entre ensaios (ISO, 1994 a,b,c,d,e; ISO, 1998).

Médias de recuperação aparente aceitáveis foram atingidas na faixa de 5  $\mu\text{g/L}$  a 30  $\mu\text{g/L}$  para ABA, DOR, EPR e IVE. Todavia,  $DPR_r$  acima de 14,7 % foram observados para DOR, EPR e IVE no nível 5  $\mu\text{g/L}$ . Estes resultados sinalizaram acurácia do método validado na faixa de 10  $\mu\text{g/L}$  a 30  $\mu\text{g/L}$ , considerando os critérios de aceitabilidade determinados.



derivado fluorescente  
das avermectinas



**Figura IV.6** - Estrutura química do derivado fluorescente das avermectinas e cromatogramas típicos obtidos por CLAE-FL.

A = solução padrão de avermectinas de 150 ng/mL.

B = amostra de leite bovino branca.

C = amostra de leite bovino adicionada de avermectinas em 30 µg/L.

**Tabela IV.8** - Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente de ABA

Estatística	Concentração (µg/kg)				
	5	10	15	20	30
<i>Normalidade</i>					
<i>R</i>	0,9741	0,9782	0,9710	0,9539	0,9862
<i>p</i>	p > 0,10	p > 0,10	p > 0,10	p > 0,05	p > 0,10
<i>Homoscedasticidade</i>					
<i>F<sub>L</sub></i>	0,874	1,299	0,337	0,155	0,707
<i>p</i>	0,49	0,34	0,80	0,92	0,57

*R* = coeficiente de correlação de Ryan-Joiner, *p* = significância, *F<sub>L</sub>* = estatística F de Levene.

**Tabela IV.9** - Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente de DOR

Estatística	Concentração (µg/kg)				
	5	10	15	20	30
<i>Normalidade</i>					
<i>R</i>	0,9631	0,9710	0,9633	0,9915	0,9842
<i>p</i>	p > 0,10	p > 0,10	p > 0,10	p > 0,05	p > 0,10
<i>Homoscedasticidade</i>					
<i>F<sub>L</sub></i>	1,179	0,473	0,048	0,278	0,643
<i>p</i>	0,38	0,71	0,98	0,84	0,61

*R* = coeficiente de correlação de Ryan-Joiner, *p* = significância, *F<sub>L</sub>* = estatística F de Levene.

Em 2002, POLLMEIER et al. indicaram exatidão e precisão de um método para determinação de EPR em leite bovino pela análise de amostras adicionadas entre 2,3 ng/mL e 51 ng/mL, sendo cinco a sete replicatas por nível, analisadas por dois laboratórios. Não ficou explicitado neste trabalho se as replicatas foram independentes ou simples repetições de leituras. A estimativa da precisão sob repetitividade e reprodutibilidade não foi feita adequadamente. Estes autores discutiram a precisão do método por meio do *DPR* calculado diretamente para os resultados obtidos por nível de concentração em cada laboratório e pelo *DPR* obtido para todos os resultados. SCHENK & LAGMAN (1999) alcançaram médias de recuperação aparente de 85,5 % a 113,8 %, com *DPR* entre 0,6 % e 23,7 %, na determinação de ABA, DOR, EPR e MOX em leite por CLEA-FL, na faixa de 1 ng/mL a 30 ng/mL, sendo quatro replicatas de cada nível. WU et

al. (2001) obtiveram médias de recuperação aparente de 74 % a 97 % para ABA e de 64 % a 87 % para IVE em experimentos empregando CL-MS com amostras de fígado suíno adicionadas entre 5 µg/kg e 100 µg/kg, sendo triplicata de cada nível. Os valores de desvios padrão reportados situaram-se entre 0,3 e 6,5 para ABA e entre 0,4 e 7,2 para IVE. ROUDAUT (1998) estimou adequadamente exatidão e precisão de um método para ensaio de ABA, DOR, IVE e MOX em fígado bovino por CLAE-FL. Na faixa de 15 µg/kg a 100 µg/kg este autor obteve médias de recuperação aparente variando de 77,5 % a 90,8 %, com  $DPR_r$  e  $DPR_R$  entre 3,80 % e 9,40 % e entre 4,05 % e 9,61 %, respectivamente. As condições experimentais assim como os critérios de aceitabilidade estipulados não foram discutidos nos trabalhos supracitados.

**Tabela IV.10** - Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente de EPR

Estatística	Concentração (µg/kg)				
	5	10	15	20	30
<i>Normalidade</i>					
$R$	0,9804	0,9514	0,9772	0,9526	0,9884
$p$	$p > 0,10$	$p > 0,10$	$p > 0,10$	$p > 0,05$	$p > 0,10$
<i>Homoscedasticidade</i>					
$F_L$	0,942	1,076	0,468	0,574	0,530
$p$	0,46	0,41	0,71	0,65	0,67

$R$  = coeficiente de correlação de Ryan-Joiner,  $p$  = significância,  $F_L$  = estatística F de Levene.

**Tabela IV.11** - Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de recuperação aparente de IVE

Estatística	Concentração (µg/kg)				
	5	10	15	20	30
<i>Normalidade</i>					
$R$	0,9259	0,9851	0,9752	0,9809	0,9486
$p$	$p > 0,01$	$p > 0,10$	$p > 0,10$	$p > 0,05$	$p > 0,10$
<i>Homoscedasticidade</i>					
$F_L$	0,028	0,757	0,413	0,391	0,758
$p$	0,99	0,55	0,75	0,76	0,55

$R$  = coeficiente de correlação de Ryan-Joiner,  $p$  = significância,  $F_L$  = estatística F de Levene.

**Tabela IV.12** - Médias de recuperação aparente e desvios padrão relativos, sob condições de repetitividade e reprodutibilidade parcial, obtidos para amostras de leite bovino adicionadas de ABA e DOR em diferentes níveis de concentração

Concentração ( $\mu\text{g/L}$ )	ABA			DOR		
	$\bar{R}$ (%)	$DPR_r$ (%)	$DPR_R$ (%)	$\bar{R}$ (%)	$DPR_r$ (%)	$DPR_R$ (%)
5	101,8	13,9	14,3	86,2	21,3	21,4
10	98,9	6,7	9,6	91,9	7,9	12,0
15	98,1	9,5	11,5	90,1	11,2	11,7
20	101,4	5,8	8,6	98,0	4,7	9,1
30	99,9	4,1	8,0	96,6	3,6	8,7

$\bar{R}$  = média de recuperação aparente,  $DPR_r$  = desvio padrão relativo de repetitividade,  $DPR_R$  = desvio padrão relativo de reprodutibilidade parcial.

Crítérios de aceitabilidade para  $\bar{R}$ : 70 % a 110 % (para amostras adicionadas  $>1 \mu\text{g/L}$  a  $10 \mu\text{g/L}$ ), 80 % a 110 % (para amostras adicionadas  $\geq 10 \mu\text{g/L}$ ).

Crítério de aceitabilidade para  $DPR_r$ :  $\leq 14,7$  %.

Crítério de aceitabilidade para  $DPR_R$ :  $\leq 22$  %.

**Tabela IV.13** - Médias de recuperação aparente e desvios padrão relativos, sob condições de repetitividade e reprodutibilidade parcial, obtidos para amostras de leite bovino adicionadas de EPR e IVE em diferentes níveis de concentração

Concentração ( $\mu\text{g/L}$ )	EPR			IVE		
	$\bar{R}$ (%)	$DPR_r$ (%)	$DPR_R$ (%)	$\bar{R}$ (%)	$DPR_r$ (%)	$DPR_R$ (%)
5	81,2	20,2	20,2	84,0	16,7	19,5
10	92,7	8,7	9,0	94,8	10,7	14,7
15	90,6	6,7	7,4	87,2	9,0	13,9
20	94,0	8,0	8,8	94,6	6,9	10,2
30	96,9	5,0	8,0	92,1	4,4	11,0

$\bar{R}$  = média de recuperação aparente,  $DPR_r$  = desvio padrão relativo de repetitividade,  $DPR_R$  = desvio padrão relativo de reprodutibilidade parcial.

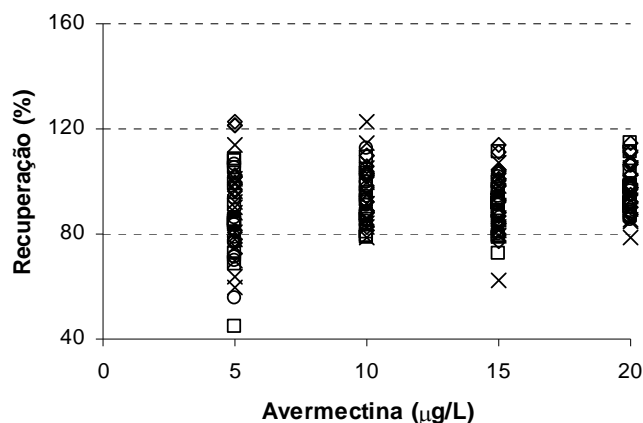
Crítérios de aceitabilidade para  $\bar{R}$ : 70 % a 110 % (para amostras adicionadas  $>1 \mu\text{g/L}$  a  $10 \mu\text{g/L}$ ), 80 % a 110 % (para amostras adicionadas  $\geq 10 \mu\text{g/L}$ ).

Crítério de aceitabilidade para  $DPR_r$ :  $\leq 14,7$  %.

Crítério de aceitabilidade para  $DPR_R$ :  $\leq 22$  %.

O limite de detecção experimental do método foi estabelecido como  $5 \mu\text{g/L}$ , visto que as quatro avermectinas foram detectadas (sinal da amostra adicionada / ruído da amostra branca  $\geq 3$ ) em todas as amostras adicionadas neste nível. O limite de quantificação para estes compostos, correspondente ao menor nível de concentração no

qual os experimentos indicaram exatidão e precisão do método, foi de 10  $\mu\text{g/L}$ . A transição entre as faixas de detecção e quantificação do método pode ser observada pela distribuição dos resultados de recuperação aparente individuais obtidos para amostras adicionadas (Figura IV.7). Limites inferiores aos alcançados neste processo de validação foram reportados por autores que adotaram como critério a relação sinal/ruído obtida para replicatas de amostras brancas. POLLMEIER et al. (2002) obtiveram limite de detecção de 0,25 ng/mL e limite de quantificação de 2 ng/mL para EPR em leite por CLAE-FL. Limite de detecção idêntico ao estabelecido por este trabalho foi publicado por WU et al. (2001) para ABA e IVE. No entanto, nada foi apresentado com relação aos limites de quantificação. ROUDAUT (1998) estabeleceu limites de detecção de 0,8; 1,5; 2,5 e 2,5  $\mu\text{g/L}$  para ABA, DOR, IVE e MOX, respectivamente, utilizando amostras brancas de fígado bovino e sistema CLAE-FL. Limite de quantificação de 7,5  $\mu\text{g/kg}$  foi apresentado por este autor, mas nenhuma descrição do experimento e critérios para determinação deste parâmetro foi feita. Embora superiores, os limites apresentados neste trabalho são certamente mais realísticos, pois foram estabelecidos sob critérios mais rigorosos em experimentos de recuperação com amostras de leite bovino adicionadas de avermectinas.



**Figura IV.7** - Distribuição dos resultados individuais de recuperação aparente obtidos para amostras adicionadas entre 5  $\mu\text{g/L}$  e 20  $\mu\text{g/L}$ , incluindo os níveis correspondentes aos limites de detecção (5  $\mu\text{g/L}$ ) e quantificação (10  $\mu\text{g/L}$ ).

$\diamond$  = ABA,  $\square$  = DOR,  $\circ$  = EPR,  $\times$  = IVE.

Considerando um limite máximo de resíduo de 10  $\mu\text{g/L}$  para as avermectinas pesquisadas,  $CC\alpha$  de 12,8; 12,9; 12,6 e 13,7  $\mu\text{g/L}$  e  $CC\beta$  de 15,6; 15,9; 15,1 e 17,4  $\mu\text{g/L}$

foram estimados para ABA, DOR, EPR e IVE, respectivamente. Assim, quando obtidos resultados menores que os valores estabelecidos para  $CC\alpha$  (concentração  $C < CC\alpha$ ), as amostras podem ser declaradas como conformes ou com analitos em concentrações inferiores ao limite máximo de resíduos estabelecido, com um nível de confiança de  $(1-\alpha)$ . No caso de resultados maiores ou iguais aos valores estimados para  $CC\beta$  (concentração  $C \geq CC\beta$ ), as amostras podem ser declaradas como não-conformes ou com analito presente em concentrações superiores ao limite máximo de resíduos, com uma probabilidade de  $(1-\beta)$ . Há, entretanto, uma faixa de incerteza estatística entre os valores de  $CC\alpha$  e  $CC\beta$ , de forma que a aplicação prática destes parâmetros depende de decisões políticas que envolvem a proteção do consumidor e do produtor de alimentos (ANTIGNAC et al., 2003).

## 4 CONCLUSÕES

O procedimento proposto foi satisfatoriamente aplicado ao processo de validação intralaboratorial do método quantitativo para ensaio de avermectinas em leite bovino por CLAE-FL. Tal procedimento permitiu a avaliação da linearidade, efeitos de matriz, seletividade, exatidão, precisão e limites, com a garantia de que as premissas relativas aos testes estatísticos aplicados não foram violadas.

Linearidade foi determinada, sem efeitos de matriz significativos, em uma faixa de concentração compatível com os níveis regulamentados para as avermectinas em leite. Seletividade, exatidão e precisão foram comprovadas. Os limites de detecção e quantificação indicaram a capacidade do método em determinar os analitos dentro dos limites regulamentados pelo Brasil e sugeridos pelo Codex Alimentarius. Os parâmetros de desempenho estudados indicaram adequação para uso do método nas atividades de monitoramento de multirresíduos de avermectinas em leite bovino, assim como a possibilidade de estudos complementares de precisão em experimentos interlaboratoriais.

## 5 CONSIDERAÇÕES

Os experimentos descritos neste capítulo foram realizados no Laboratório de Referência Agropecuária de Minas Gerais (LANAGRO-MG) do MAPA, em Pedro Leopoldo, Minas Gerais, Brasil.

A referência da publicação deste trabalho encontra-se descrita a seguir:



SOUZA, S.V.C.; LIMA, J.A.; TEODORO, J.C.; JUNQUEIRA, R.G. Validação intralaboratorial de método quantitativo para determinação múltipla de resíduos de avermectinas em leite bovino por cromatografia líquida de alta eficiência com detecção de fluorescência. *Ciênc. Technol. Aliment.* (Aceito para publicação em 19.07.2007)

# **CAPÍTULO V - VALIDAÇÃO INTRALABORATORIAL DE MÉTODO QUALITATIVO PARA DETERMINAÇÃO DE SUBPRODUTOS DE ORIGEM ANIMAL EM RAÇÕES PARA RUMINANTES POR MICROSCOPIA**

## **RESUMO**

Um procedimento para validação intralaboratorial de métodos qualitativos foi delineado e aplicado na validação de um método para detecção e identificação de subprodutos de origem animal em rações para ruminantes por microscopia. Os analitos pesquisados foram farinhas de carne e ossos, ossos não calcinados, sangue, vísceras, pena e peixe. Seletividade, limite de detecção e adequação para uso do método foram avaliados por procedimentos empregando materiais de referência dos analitos, amostras brancas e adicionadas de cada analito a 0,025 %; 0,05 %; 0,10 %; 0,20 %; 0,40 % e 0,80 %, em pelo menos três replicatas independentes. O método apresentou seletividade suficiente para distinção entre os diferentes subprodutos de origem animal proibidos, sem interferência de componentes da matriz. Os limites de detecção foram 0,05 % para farinhas de carne e ossos, ossos não calcinados, sangue, pena e peixe e 0,10 % para farinha de vísceras.

Palavras-chave: validação intralaboratorial; subprodutos de origem animal, ração, microscopia.

## **ABSTRACT**

SINGLE-LABORATORY VALIDATION OF A QUALITATIVE METHOD FOR DETECTION OF ANIMAL MEALS IN RUMINANT FEEDS BY MICROSCOPY. A procedure for single-laboratory validation of qualitative methods was proposed and applied to the validation of a method for detection and identification of animal meals in ruminant feeds by microscopy. The studied analytes were meat and bone, steamed bone, blood, poultry offal, feather and fish meals. Selectivity, detection limit and fitness for purpose were evaluated by procedures with reference materials, sample blanks and samples spiked with each analyte at 0.025 %; 0.05 %; 0.10 %; 0.20 %; 0.40 % and 0.80 %, in at least three independent replicates. The method selectivity was sufficient for distinguish the prohibited animal meals, without matrix interferences. Detection limits were 0.05 % for

meat and bone, steamed bone, blood, feather and fish meals and 0.10 % for poultry offal meal.

Key words: single-laboratory validation; animal meals; feed; microscopy.

## 1 INTRODUÇÃO

Encefalopatias espongiformes transmissíveis (TSE) são responsáveis por doenças neurodegenerativas fatais em seres humanos e animais. TSE humanas incluem a doença de Creutzfeldt-Jakob (CJD), insônia familiar fatal (FFI), a síndrome de Gerstmann-Sträussles-Scheinker (GSS) e *kuru*, uma doença transmitida por meio de rituais de canibalismo. Em animais, TSE envolvem a encefalopatia espongiforme de ovinos e caprinos ou *scrapie*, encefalopatia espongiforme bovina (BSE) ou “doença da vaca louca”, encefalopatia espongiforme felina (FSE), encefalopatia espongiforme de marta (TME) e encefalopatia espongiforme de cervídeos ou *chronic wasting disease* (CWD) (DORMONT, 2002; CONCEPCION & PADLAN, 2003). O período médio de incubação destas doenças é bastante elevado, podendo variar de quatro a seis anos para BSE em bovinos (KIMBERLIN, 1993) e correspondendo a, pelo menos, nove anos para CJD em humanos (GHANI et al., 1998), com o agravante de que não há indicações visíveis das doenças durante estes períodos.

A hipótese mais aceita sobre o agente infeccioso das TSE está relacionada aos *prions*, partículas de proteína desprovidas de material genético (DNA e RNA) e presentes em diversos tipos de células. Embora a natureza destes agentes infecciosos e seus mecanismos de ação não estejam completamente compreendidos, evidências consideráveis sugerem que as TSE sejam provocadas por transformações na conformação de *prions* normais (estrutura predominante alfa-hélice) a *prions* de conformação anormal (estrutura predominante beta-hélice). Estas estruturas de conformação anormal são resistentes à ação de proteases e à altas temperaturas, formando agregados em organelas celulares e provocando danos ao tecido nervoso (PRUSINER et al., 1981; PERETZ et al., 1997; HUR et al., 2002).

Os primeiros casos de BSE foram diagnosticados no Reino Unido em 1986. Na época, acreditava-se que a BSE era uma forma de *scrapie* adquirida por bovinos após a ingestão de subprodutos de ovinos contaminados, os quais eram adicionados às rações (KIMBERLIN, 1993). Outras hipóteses sugerem que a BSE tenha sido adquirida por ingestão de *prions* humanos contidos em subprodutos de mamíferos (ossos inteiros,

farinhas de ossos, carcaças, etc.) importados pelo Reino Unido em estado bruto (COLCHESTER & COLCHESTER, 2005). Uma nova variante de CJD (vCJD) foi descrita cerca de dez anos após a primeira identificação de BSE, também no Reino Unido, em março de 1996. Em contraste com a forma tradicional de CJD, esta nova variante afetava pacientes jovens (WILL et al., 1996 a; WILL et al., 1996 b; BROWN, 1997; FISHBEIN, 1998). Foi sugerido que a vCJD pudesse ser causada pela transmissão da BSE a humanos por meio do consumo de carne contaminada (BRUCE et al., 1997; SCOTT et al., 1999; CONCEPCION & PADLAN, 2003).

A ocorrência de BSE em bovinos e sua transmissibilidade a humanos tornaram-se uma preocupação internacional, principalmente por não existirem terapias aplicáveis para TSE em humanos. Este problema exigiu uma mobilização mundial no sentido de minimizar a exposição da população humana ao agente infeccioso da BSE por meio de práticas de controle de qualidade de alimentos e do uso seguro de produtos biológicos de origem bovina (DORMONT, 2002).

As recomendações de medidas para proteção da saúde pública acordadas entre Organização Mundial da Saúde (OMS), Organização Mundial de Saúde Animal (OIE) e Organização das Nações Unidas para a Agricultura e a Alimentação (FAO) incluem: i) nenhuma parte ou produto de animal o qual tenha apresentado sinais de TSE ou tecidos que provavelmente contenham BSE devem entrar nas cadeias alimentares (humana e animal); ii) todos os países devem estabelecer vigilância e notificação compulsória de ocorrências de BSE; iii) todos os países devem banir o uso de tecidos de ruminantes em rações para ruminantes; iv) leite e derivados são considerados seguros; v) gelatina e sebo somente são considerados seguros se processos de produção efetivos forem adotados; e vi) produtos medicinais e acessórios médicos devem ser obtidos de países sem casos esporádicos de BSE e com medidas recomendadas para minimizar os riscos (WHO, 1996).

Relevantes mudanças foram introduzidas na nutrição animal em decorrência desta crise da BSE, com o reconhecimento de que a indústria de rações e seus fornecedores de ingredientes fazem parte da cadeia alimentar e são responsáveis pela segurança destes produtos (DEN HARTOG, 2003). Neste contexto, novos ingredientes estão sendo estudados e sugeridos, além da implementação de processos de controle e garantia da qualidade dos produtos, como Boas Práticas de Fabricação (BPF) e Análise de Perigos e Pontos Críticos de Controle (APPCC) (DEN HARTOG, 2003; BRAMBILLA & DE FILLIPPIS, 2005). Sob o ponto de vista regulamentar, muitos países formalizaram a proibição do uso de proteínas derivadas de subprodutos de origem animal em rações para

alimentação ruminantes. No Brasil, o Ministério da Agricultura, Pecuária e Abastecimento (MAPA) proibiu a produção, comercialização, utilização e importação de produtos destinados à alimentação para ruminantes que contenham em sua composição proteínas e gorduras de origem animal, exceto leite e produtos lácteos, farinha de ossos calcinados, gelatina e colágeno preparados exclusivamente a partir de couros e peles (BRASIL, 2004 a,b). Tais determinações criaram a demanda por um suporte analítico muito bem estruturado, com métodos de ensaio confiáveis, que garantam um controle efetivo da segurança e rastreabilidade das cadeias alimentares relacionadas à BSE.

Muitos métodos têm sido descritos para determinação de material animal em ração baseados na detecção de proteínas como focagem isoelétrica (WINTERO, THOMSEN & DAVIES, 1990; JEMMI & SCHLOSSER, 1993), imunoensaios (ANSFIELD, 1994; MORALES et al., 1994; HSIEH, SHEU & BRIDGMAN, 1998) ou gel difusão (WAIBLINGER, WEBER & KLEINERT, 1998). Técnicas baseadas de determinação de seqüências de DNA também são apresentadas como DNA hibridização (EBBEHOJ & THOMSEN, 1991; HUNT, PARKES & LUMLEY, 1997) e reação em cadeia da polimerase (PCR), que pode ser espécie específica para bovino (MEYER et al, 1995; MATSUNAGA et al., 1999; BRODMANN & MOOR, 2003), suíno (MEYER, CANDRIAN & LÜTHY, 1994; WOLF & LÜTHY, 2001) ou ovino (CHIKUNI et al., 1994). Uma detecção espécie específica de tecidos de sistema nervoso central também foi demonstrada por meio da determinação de esfingolipídeos por cromatografia gasosa acoplada a espectrometria de massas (NIEDERER & BOLLHALDER, 2001). A microscopia de rações permite a identificação de tecidos e é uma técnica extensivamente utilizada na determinação de material de origem animal em rações para ruminantes. A detecção e identificação são baseadas na observação da conformação morfológica de fragmentos com um microscópio estereoscópico e no exame de estruturas histológicas de partículas mais finas com microscópio óptico (AOAC, 1998 a).

Independente da técnica aplicada e do tipo de método empregado, qualitativo ou quantitativo, a confiabilidade dos resultados emitidos deve ser assegurada por meio de procedimentos de validação do método (EURACHEM, 1998; THOMPSON, ELLISON & WOOD, 2002 ISO, 2005 b). Especificamente no caso dos métodos qualitativos, a avaliação de parâmetros como seletividade e limite de detecção é fundamental para o conhecimento de seu potencial aplicativo e limitações.

Os objetivos deste trabalho foram otimizar e validar, segundo procedimento proposto, um método qualitativo para detecção de subprodutos de origem animal (farinhas

carne e ossos, ossos não calcinados, sangue, vísceras, pena e peixe) em ração para ruminantes por microscopia.

## 2 MATERIAL E MÉTODOS

### 2.1 MATERIAL

- Almofariz e pistilo.
- Béquer de 100; 200 e 500 mL.
- Cápsula porcelana ou vidro borossilicato de 250 mL.
- Conjunto de peneiras manual (malhas de 0,25; 0,5; 1,0 e 2,0 mm)
- Espátula.
- Estilete.
- Filme flexível (Parafilm).
- Filtro com porosidade de 0,1 mm a 0,2 mm.
- Frasco conta-gotas.
- Frasco kitazato de 2 L.
- Funil de Büchner com diâmetro mínimo de 12 cm.
- Funil de separação de 250 mL.
- Lâminas e lamínulas.
- Pinça pontiaguda.
- Placa de Petri.
- Placa tipo *spot test*.
- Quarteador tipo Johnes.
- Recipiente para banho de gelo.
- Saco plástico.
- Suporte e garra para funil de separação.

### 2.2 EQUIPAMENTOS

- Balança analítica com resolução de 0,0001 g (Mettler).
- Balança analítica com resolução de 0,01 g (Mettler).
- Balão volumétrico de 50; 100; 200; 500; 1000 e 2000 mL.
- Bomba de vácuo.
- Chapa aquecedora (Sybron Thermolyne).

- Estufa com ventilação e capacidade para 65 °C (Fanem).
- Microscópio composto, com aumentos de 40 x a 600 x e polarização (Carl Zeiss Jena Docuval, Alemanha).
- Microscópio de imunofluorescência (Olimpus, Japão).
- Microscópio estereoscópico, com aumentos de 10 x a 100 x, luz refletida, transmitida e polarização (Carl Zeiss Jena Docuval, Alemanha).
- Proveta de 10; 50; 100; 200; 500 e 1000 mL.

## 2.3 REAGENTES E PADRÕES

Todos dos reagentes utilizados foram de grau de pureza apropriado.

Acetona, clorofórmio e éter de petróleo (grau para análise) foram obtidos da Fisher (New Jersey, EUA). Ácido clorídrico concentrado, ácido sulfúrico concentrado, acetato de chumbo e bissulfito de sódio (grau para análise) foram adquiridos da Vetec (Rio de Janeiro, Brasil). Molibdato de amônio e oxalato de amônio (grau para análise) foram da Ecibra (São Paulo, Brasil). Ácido 1,2,4-amino naftol sulfônico (grau para análise) foi da Kodak (New York, EUA), hidróxido de amônio (grau para análise) da Pro Analyti (Rio de Janeiro, Brasil), hidróxido de sódio (grau para análise) da Nuclear (São Paulo, Brasil) e sulfito de sódio (grau para análise) da Reagen (Rio de Janeiro, Brasil). Água destilada e desionizada foi preparada no laboratório.

Materiais de referência de farinhas carne e ossos, ossos não calcinados, sangue, vísceras, pena e peixe foram obtidos de indústrias sob o serviço de inspeção oficial do MAPA. Os materiais de referência de cada analito foram coletados de pelo menos três diferentes estabelecimentos cujos processos de produção atendiam aos padrões brasileiros para tamanho de partículas (ANFAR, 1998). Farinhas de ossos não calcinados, sangue, vísceras, pena e peixe foram provenientes de processos com um máximo de 5 % de retenção quando passadas através de peneiras de 2 mm, enquanto farinhas de carne e ossos foram coletadas de processos com retenção máxima de 5 % e 10 % quando passadas através de peneiras de 2 mm e 1,68 mm, respectivamente. Estes materiais foram considerados materiais de referência tomando como base a definição apresentada pelo ISO Guide 30 (ISO, 1992) de que material de referência é um material ou substância que tem um ou mais valores de propriedades que são suficientemente homogêneos e bem estabelecidos para ser usado na calibração de um aparelho, na avaliação de um método ou atribuição de valores a materiais. Neste guia são considerados os casos de alguns materiais de referência que têm propriedades as quais não podem ser

determinadas por métodos físicos e químicos exatamente definidos, sendo citado o exemplo de materiais de referência biológicos e tecnológicos.

## 2.4 SOLUÇÕES

### 2.4.1 Soluções de éter de petróleo:clorofórmio (10/90; 20/80; 30/70; 40/60 e 50/50, v/v)

As soluções de éter de petróleo:clorofórmio foram preparadas pela mistura dos respectivos solventes para volumes finais de 4 L.

### 2.4.2 Soluções para provas químicas qualitativas

As soluções para provas químicas qualitativas foram preparadas conforme descrito por AAFM (1992), EC (1998) e HUSS (1975) e acondicionadas em frascos conta-gotas devidamente identificados.

**Solução de ácido clorídrico 50 % (v/v):** 100 mL de ácido clorídrico foram diluídos para 200 mL com água destilada e desionizada.

**Solução de hidróxido de amônio 50 % (v/v):** hidróxido de amônio (100 mL) foi dissolvido com água destilada e desionizada para 200 mL.

**Solução saturada de oxalato de amônio:** 140,00 g de oxalato de amônio foram dissolvidos para 2000 mL com água destilada e desionizada e aquecidos até completa dissolução, em chapa aquecedora.

**Solução de molibdato de amônio 2,5 % (m/v):** 12,50 g de molibdato de amônio foram dissolvidos em 100 mL água destilada e desionizada e em 150 mL de solução de ácido sulfúrico 10 N. O volume foi completado para 500 mL com água destilada e desionizada. Esta solução foi armazenada em frasco âmbar, sob refrigeração.

**Solução de ácido sulfúrico 10 N:** 266,5 mL de ácido sulfúrico foram diluídos para 1000 mL com água destilada e desionizada.



**Solução de ácido 1,2,4-amino naftol sulfônico:** 250 mg de ácido 1,2,4-amino naftol sulfônico foram umedecidos com 1 ou 2 gotas de solução de bissulfito de sódio 15 %. Em seguida foram adicionados de 97 mL da solução de bissulfito de sódio 15% e de 2,5 mL de solução de sulfito de sódio 20 %. Após o período de 12 h em repouso, a solução foi filtrada e armazenada em frasco âmbar.

**Solução de bissulfito de sódio 15 % (m/v):** 15,00 g de bissulfito de sódio foram dissolvidos com água destilada e desionizada para 100 mL.

**Solução de sulfito de sódio 20 % (m/v):** 10,00 g de sulfito de sódio foram dissolvidos com água destilada e desionizada para 50 mL.

**Solução de acetato de chumbo 2 % (m/v):** 2,00 g de acetato de chumbo foram diluídos para 100 mL de solução hidróxido de sódio 10 %.

**Solução de hidróxido de sódio 10 % (m/v):** hidróxido de sódio (10,00 g) foi dissolvido com água destilada e desionizada, em banho de gelo, e o volume completado para 100 mL.

## 2.5 AMOSTRAS

Amostras brancas de rações para ruminantes foram obtidas de estabelecimentos produtores inspecionados pelo MAPA cujos processos produtivos não continham subprodutos de origem animal em suas formulações, prevenindo desta forma, os inconvenientes de possíveis contaminações cruzadas. Considera-se amostras brancas aquelas matrizes isentas do analito ou para as quais o analito não tenha sido detectado.

As amostras foram homogeneizadas e quarteadas em quarteador tipo Johnes. Em seguida foram acondicionadas em sacos plásticos devidamente identificados, os quais foram lacrados e armazenados a temperatura ambiente, protegidos de luz e umidade, até o momento dos ensaios.

## 2.6 MÉTODO DE ENSAIO

O método validado foi baseado em procedimentos previamente descritos (FERRANDO & HENRI, 1966; HUSS, 1975; AAFM, 1992; EC, 1998).

### **2.6.1 Princípio**

O método de ensaio validado fundamentou-se na separação de subprodutos de origem animal da ração por partição líquido-líquido com misturas de éter de petróleo:clorofórmio, seguida de detecção e identificação microscópica dos componentes por meio da avaliação de tamanho, forma, cor, textura, dureza, brilho e outras características de seus fragmentos, com auxílio de provas químicas e físicas qualitativas.

### **2.6.2 Cuidados**

Equipamentos de proteção individual (avental, óculos, luvas e máscaras) e equipamentos de proteção coletiva como capelas de exaustão química foram utilizados durante a realização dos ensaios.

### **2.6.3 Preparo das amostras**

Foram pesados  $15,0000 \text{ g} \pm 0,0001 \text{ g}$  de amostra peletizada, com auxílio de espátula, transferidos para uma cápsula e acrescentados de quantidade de água destilada e desionizada suficiente para umedecer a amostra. A mistura foi agitada ocasionalmente até completa desintegração dos *pellets*. A amostra foi filtrada quantitativamente através filtro (porosidade de 0,1 mm a 0,2 mm), adaptado a funil de Büchner e frasco kitazato, sob vácuo. Lavagens sucessivas foram feitas com água destilada e desionizada até o clareamento do filtrado. Em seguida, foi feita uma filtração com acetona para secagem do material retido. O material foi transferido para um almofariz, aguardando a evaporação do solvente remanescente. A amostra seca foi fragmentada com pistilo, o suficiente para desprender as partículas menores que aderem às maiores. O material foi então transferido para um funil de separação.

Amostras fareladas foram pesadas ( $15,0000 \text{ g} \pm 0,0001 \text{ g}$ ) e transferidas diretamente para um funil de separação, sem necessidade de preparo prévio.

### **2.6.4 Separação**

Foram adicionados 50 mL de clorofórmio à amostra contida no funil de separação e agitados. Após separação das fases, a fração precipitada foi drenada sobre uma cápsula e o solvente remanescente no funil de separação descartado. O excesso de solvente da cápsula foi cuidadosamente succionado e descartado. À fração remanescente no funil de separação foram adicionados 50 mL de solução de éter de petróleo:clorofórmio (10/90, v/v) e agitados, aguardando a separação das fases. A fração de maior densidade precipitada foi drenada sobre outra cápsula e o solvente remanescente no funil de

separação foi descartado. O excesso de solvente nesta cápsula foi cuidadosamente succionado e descartado. Os procedimentos de separação foram repetidos com soluções de éter de petróleo: clorofórmio (20/80; 30/70; 40/60 e 50/50, v/v), até a completa separação das fases de diferentes densidades. Após completa evaporação dos solventes restantes nas frações drenadas sobre as cápsulas, as mesmas foram secas em estufa a 65 °C. Em seguida, as frações foram peneiradas separadamente em conjunto de peneiras com malhas de 0,25; 0,5; 1,0 e 2,0 mm e as porções retidas nas peneiras foram transferidas para placas de Petri previamente identificadas.

### **2.6.5 Detecção e identificação**

Cada porção contida nas placas de Petri foi levada ao microscópio estereoscópico, iniciando o ensaio microscópico a partir da fração mais grossa. A pesquisa foi feita de uma extremidade a outra da placa de Petri para que todas as partículas fossem observadas em camada fina. Microscópio composto foi utilizado para avaliação das partículas finamente moídas. Com o auxílio de pinça, estilete, lâminas e lamínulas, cada constituinte examinado foi separado e identificado pelo tamanho, forma, cor, textura, dureza, brilho e outras características morfológicas, por meio de comparação com materiais de referência. Nos casos em que houve necessidade de provas químicas qualitativas para comprovação, a partícula foi retirada com pinça pontiaguda ou estilete e transferida para placas tipo *spot test*.

## **2.7 OTIMIZAÇÃO E VALIDAÇÃO**

Os parâmetros seletividade e limite de detecção foram estabelecidos em ensaios com materiais de referência, amostras brancas e adicionadas. A adequação para uso do método foi avaliada em função dos parâmetros estudados e respectivos critérios de aceitabilidade definidos. O nível de significância adotado nos testes de hipóteses foi  $\alpha = 0,05$ . Os cálculos foram feitos em planilhas comerciais, cujos resultados foram previamente validados frente aqueles obtidos em diferentes softwares de análise estatística.

### **2.7.1 Ensaios com materiais de referência**

Objetivando avaliar a distribuição dos analitos nas diferentes misturas de solventes empregadas na etapa de separação das frações, 15,0000 g  $\pm$  0,0001 g de pelo menos três materiais de referência de cada analito foram submetidos ao tratamento com éter de

petróleo:clorofórmio (0/100; 10/90; 20/80; 30/70; 40/60 e 50/50, v/v). Em seguida, as frações separadas em cada mistura de solventes foram submetidas ao ensaio microscópico para estabelecer as propriedades (características morfológicas) que poderiam ser atribuídas a cada analito, visando confirmação da identidade.

Os resultados de porcentagem de concentração das partículas obtidos, para cada analito e mistura de solvente, foram analisados quanto à presença de valores extremos (*outliers*) pelo teste de Grubbs (GRUBBS, 1969; BARRET & LEWIS, 1994; BURKE 2001 b). As porcentagens de concentração obtidas para as replicatas de cada analito nas diferentes misturas de solventes foram analisadas pelo teste de Tuckey (SNEDECOR & COCHRAN, 1989). As premissas de distribuição normal (RYAN & JOINER, 1976) e homoscedasticidade dos resíduos (erros) das porcentagens de concentração (LEVENE, 1960, BROWN & FORSYTHE, 1974) foram avaliadas antes da aplicação deste teste de hipóteses.

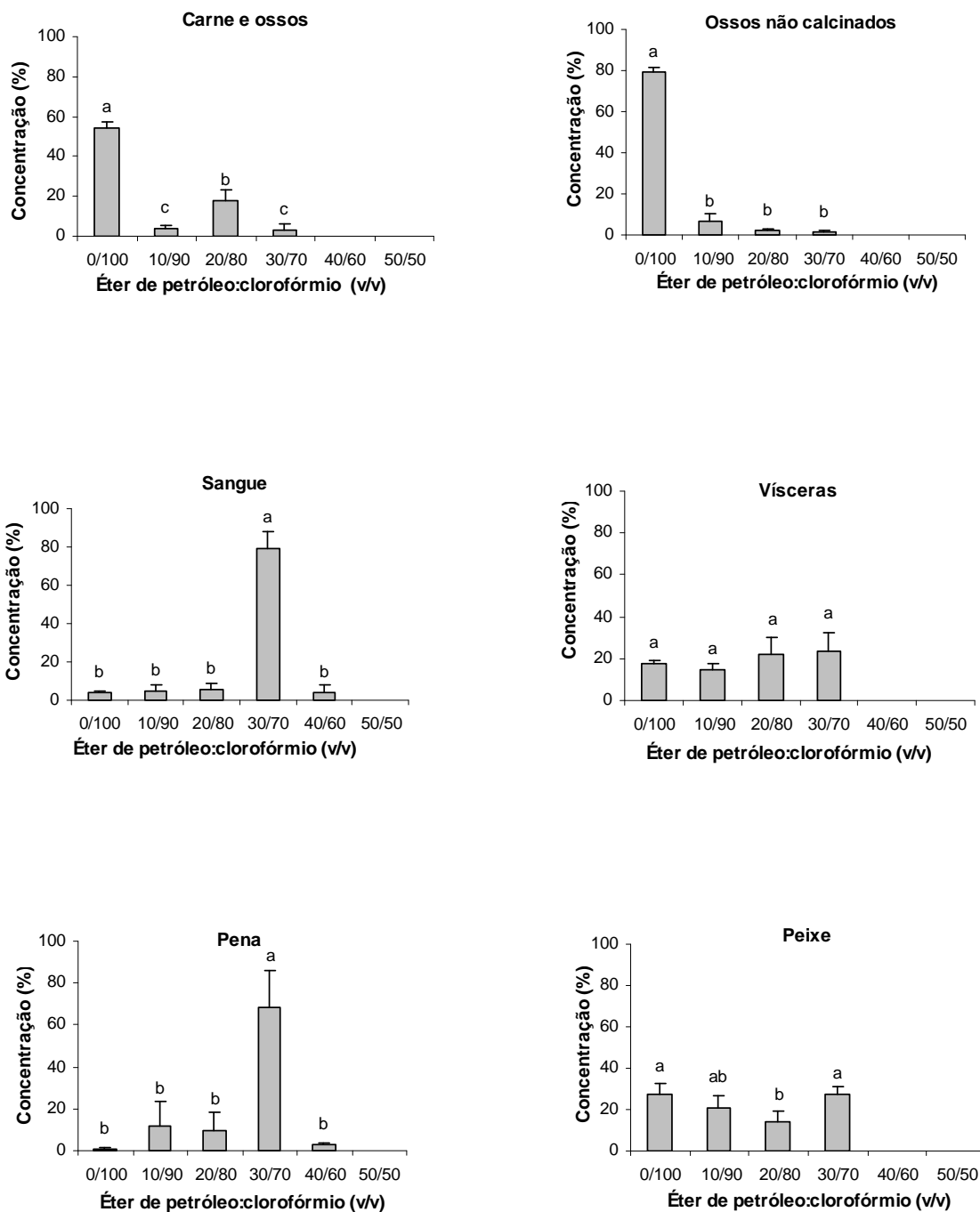
### **2.7.2 Ensaio com amostras brancas e adicionadas**

Amostras de ração para ruminantes brancas e adicionadas de cada analito em 0,025; 0,05; 0,10; 0,20; 0,40 e 0,80 %, sendo pelo menos três replicatas independentes de cada nível, foram analisadas. Objetivando avaliar a seletividade, por meio da determinação da capacidade de distinção entre os diferentes analitos de subprodutos de origem animal proibidos em ração para ruminantes, assim como pela habilidade do método em determinar cada analito sem interferência de outros componentes da matriz, amostras cegas foram distribuídas em ordem aleatória para realização dos ensaios.

Limites de detecção foram definidos para cada analito separadamente. Resultados detectados *versus* resultados não detectados foram avaliados para determinar o limiar de perda de confiabilidade do método. O menor nível de concentração estudado onde o analito foi detectado para todas as replicatas foi considerado como limite de detecção. Nestes níveis, experimentos complementares foram realizados até que fosse alcançado um mínimo de dez replicatas.

## **3 RESULTADOS E DISCUSSÃO**

Os resultados dos estudos de concentração de partículas dos analitos nas diferentes misturas de solventes encontram-se ilustrados na Figura V.1.



**Figura V.1** - Perfil de porcentagem de concentração dos analitos nas diferentes misturas de solventes empregadas na etapa de separação.

Médias indicadas pela mesma letra não diferiram significativamente pelo teste de Tuckey com um nível de significância de 0,05.

As barras de erros representam os desvios padrão obtidos para as replicatas de materiais de referência de cada analito.

Valores extremos não foram identificados entre os resultados de porcentagem de concentração pelo teste de Grubbs ( $p > 0,05$ ).

Os testes de Ryan-Joiner e de Levene modificado demonstraram que os resíduos destes resultados seguiam a distribuição normal e tinham variâncias homogêneas, permitindo o emprego do teste de Tuckey (Tabela V.1).

**Tabela V.1** - Avaliação das premissas de normalidade e homoscedasticidade dos resíduos dos resultados de porcentagem de concentração

Estatística	Analito					
	Carne e ossos	Ossos não calcinados	Sangue	Vísceras	Pena	Peixe
<i>Normalidade</i>						
$R$	0,9889	0,9792	0,9574	0,9891	0,9570	0,9780
$p$	$p > 0,10$	$p > 0,10$	$p > 0,10$	$p > 0,10$	$p > 0,10$	$p > 0,10$
<i>Homoscedasticidade</i>						
$F_L$	1,129	1,230	0,683	0,467	0,673	0,170
$p$	0,38	0,36	0,62	0,71	0,63	0,91

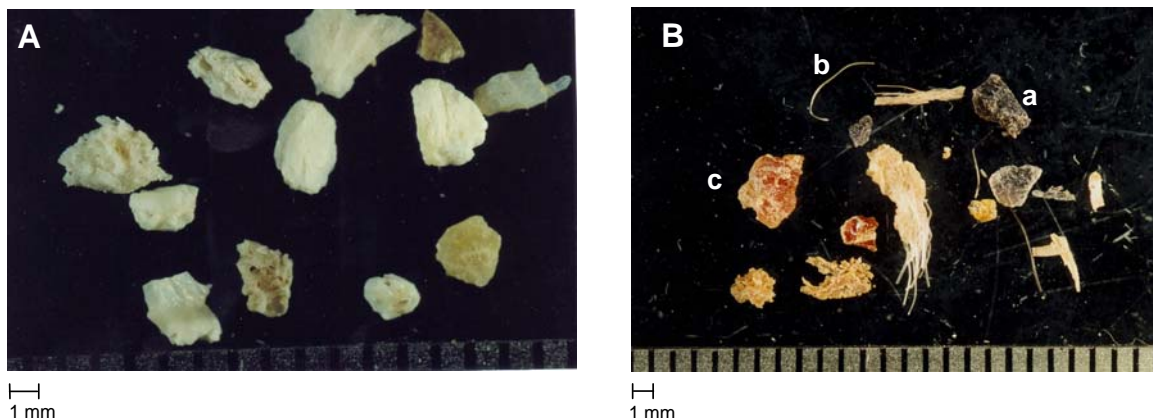
$R$  = coeficiente de correlação de Ryan-Joiner,  $p$  = significância,  $F_L$  = estatística F de Levene.

Para farinha de carne e ossos foram estudados materiais de referência de quatro estabelecimentos diferentes, sendo observadas porcentagens de concentração médias de  $(53,9 \pm 3,3)\%$ ,  $(3,5 \pm 2,2)\%$ ,  $(17,6 \pm 5,3)\%$  e  $(3,1 \pm 2,8)\%$  para as misturas de éter de petróleo:clorofórmio (0/100; 10/90; 20/80 e 30/70, v/v), respectivamente. Para os três materiais de referência de farinha de ossos não calcinados avaliados, as porcentagens de concentração médias foram de  $(79,4 \pm 2,2)\%$ ,  $(6,9 \pm 3,7)\%$ ,  $(2,4 \pm 0,7)\%$  e  $(1,2 \pm 1,1)\%$ , respectivamente. Considerando estes resultados, pôde-se observar que houve uma maior concentração dos analitos farinha de carne e ossos e farinha de ossos não calcinados quando a mistura de solventes éter de petróleo:clorofórmio (0/100, v/v) foi utilizada ( $p < 0,05$ ). No caso das farinhas de sangue e de pena, para as quais foram avaliados materiais de três diferentes procedências, foram obtidas porcentagens de concentração médias de  $(4,0 \pm 0,9)\%$ ,  $(5,2 \pm 2,8)\%$ ,  $(5,5 \pm 3,1)\%$ ,  $(78,9 \pm 9,5)\%$  e  $(4,3 \pm 3,4)\%$  e de  $(1,0 \pm 0,8)\%$ ,  $(12,1 \pm 11,4)\%$ ,  $(9,7 \pm 8,5)\%$ ,  $(68,2 \pm 17,8)\%$  e  $(3,3 \pm 0,6)\%$ , respectivamente, para misturas de éter de petróleo:clorofórmio (0/100; 10/90; 20/80; 30/70 e 40/60, v/v). Esta distribuição indicou uma maior concentração das partículas de farinha de sangue e farinha de pena na mistura éter de petróleo:clorofórmio (30/70, v/v) ( $p < 0,05$ ). No entanto, para as farinhas de vísceras e peixe não houve concentração das

partículas em nenhuma das misturas de solventes pesquisadas ( $p > 0,05$ ). As porcentagens de concentração médias variaram de  $(14,8 \pm 1,5)\%$  a  $(23,2 \pm 9,1)\%$  para os três materiais de farinha de vísceras e de  $(14,4 \pm 3,7)\%$  a  $(27,6 \pm 6,3)\%$  para os quatro materiais de farinha de peixe examinados.

As partículas de farinha de carne e ossos tiveram aparência granular. Partículas de várias cores foram encontradas. Na fração predominante proveniente da mistura de solventes éter de petróleo:clorofórmio (0/100, v/v), ocorreram fragmentos de tecido ósseo, com superfícies rugosas, opacos e resistentes à fragmentação, de tamanho variando de pequenos fragmentos a pó, e coloração branco amarelada (Figura V.2 A). Algumas provas químicas foram feitas confirmando a presença de tecido ósseo. Os fragmentos apresentaram efervescência lenta com solução de ácido clorídrico 50 % indicativa da presença de tecido ósseo. Quando foram tratados com solução de hidróxido de amônio 50 % (para neutralizar o pH do meio), seguido de solução saturada de oxalato de amônio (caracterizando a presença do elemento cálcio), houve também a formação de precipitado branco, provavelmente oxalato de cálcio. A presença de fósforo foi indicada pela adição de solução de molibdato de amônio 2,5 % e de ácido 1,2,4-amino naftol sulfônico. O fósforo teria formado ácido fosfomolibdico, o qual teria sido reduzido pelo ácido 1,2,4-amino naftol sulfônico a óxido de molibdênio, produzindo uma coloração azul. Já na fração obtida a partir da mistura de éter de petróleo:clorofórmio (20/80, v/v), partículas de carne e tendão apareceram como pedaços pequenos, duros, semi-transparentes, de cor amarela a amarela-amarronzada ou enegrecidas em virtude de aquecimento excessivo, de formatos irregulares e com superfícies rugosas e lisas. Pedaços de casco, chifre, pêlo e couro foram observados (Figura V.2 B). Contaminação por fibra vegetal também poderia ocorrer indicando a presença de conteúdo ruminal.

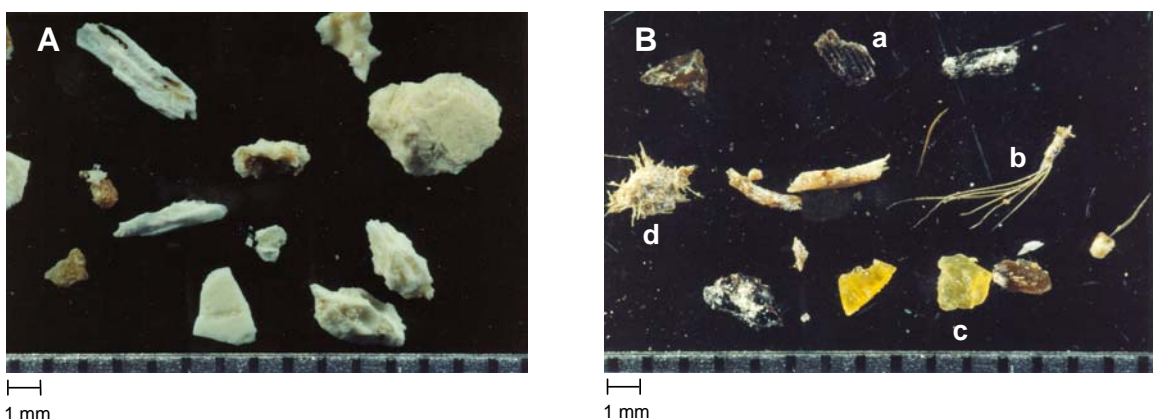
Partículas de ossos não calcinados foram observadas na fração predominante drenada da mistura de solventes éter de petróleo:clorofórmio (0/100, v/v), variando de tamanho pequeno a um pó, de coloração branco amarelada, com superfícies lisas ou ligeiramente rugosas, resistentes à fragmentação (Figure V.3 A). A presença de tecido ósseo foi sugerida por provas químicas qualitativas. Os fragmentos apresentaram efervescência lenta com solução de ácido clorídrico 50 %, houve indicação da presença de cálcio (prova com solução saturada de oxalato de amônio formando precipitado branco de oxalato de cálcio) e de fósforo (prova com solução de molibdato de amônio 2,5 % e de ácido 1,2,4-amino naftol sulfônico, formando coloração azul). Pequenas quantidades de casco, pêlo, couro e fibras vegetais semi digeridas estiveram presentes na fração proveniente da mistura éter de petróleo:clorofórmio (20/80, v/v) (Figura V.3 B).



**Figura V.2** - Fragmentos de farinha de carne e ossos.

A = fração obtida da mistura éter de petróleo:clorofórmio (0/100, v/v) - partículas de ossos.

B = fração obtida com a mistura éter de petróleo:clorofórmio (20/80, v/v) - casco / chifre (a), pêlo (b) e couro (c).



**Figura V.3** - Fragmentos de farinha de ossos não calcinados.

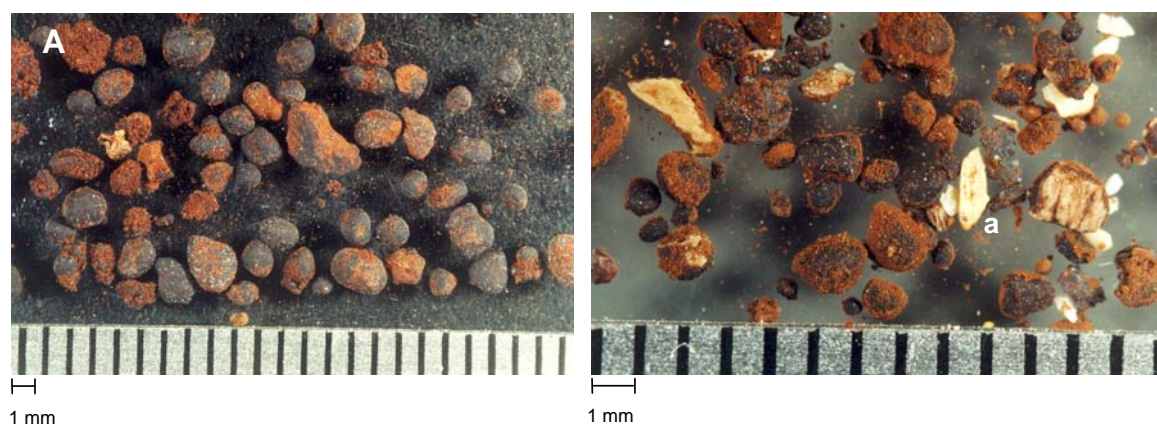
A = fração obtida da mistura éter de petróleo:clorofórmio (0/100, v/v) - partículas maiores.

B = fração obtida com a mistura éter de petróleo:clorofórmio (20/80, v/v) - casco / chifre (a), pêlo (b), couro (c) e fibra vegetal (d).

Partículas de sangue apresentaram-se principalmente na fração éter de petróleo:clorofórmio (30/70, v/v). Estes fragmentos consistiram de partículas finas e esféricas, de formato irregular, tamanho variado, coloração vermelha escura a negra púrpura, duras de quebrar, com superfícies lisas e opacas ou sem lustro (Figura V.4 A). Como de provas químicas qualitativas confirmatórias, os fragmentos foram tratados com ácido sulfúrico concentrado e analisados em microscópio de imunofluorescência. A presença de sangue foi confirmada pela transformação da hemoglobina do sangue em protoporfirina, a qual apresenta fluorescência vermelho escarlate. As partículas de farinha de sangue observadas a partir da fração mais densa de éter de petróleo:clorofórmio



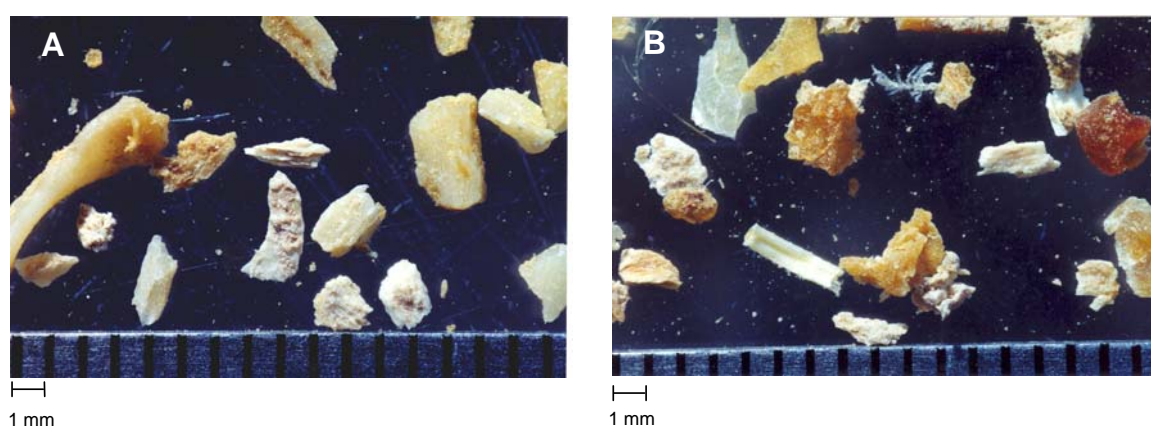
(0/100, v/v) apresentaram, algumas vezes, ossos e, ocasionalmente, pequenas quantidades de fibra vegetal e de pêlos (Figura V.4 B).



**Figura V.4** - Fragmentos de farinha de sangue.

A = fração obtida da mistura éter de petróleo:clorofórmio (30/70, v/v) - partículas esféricas.  
B = fração obtida com a mistura éter de petróleo:clorofórmio (0/100, v/v) - ossos (a).

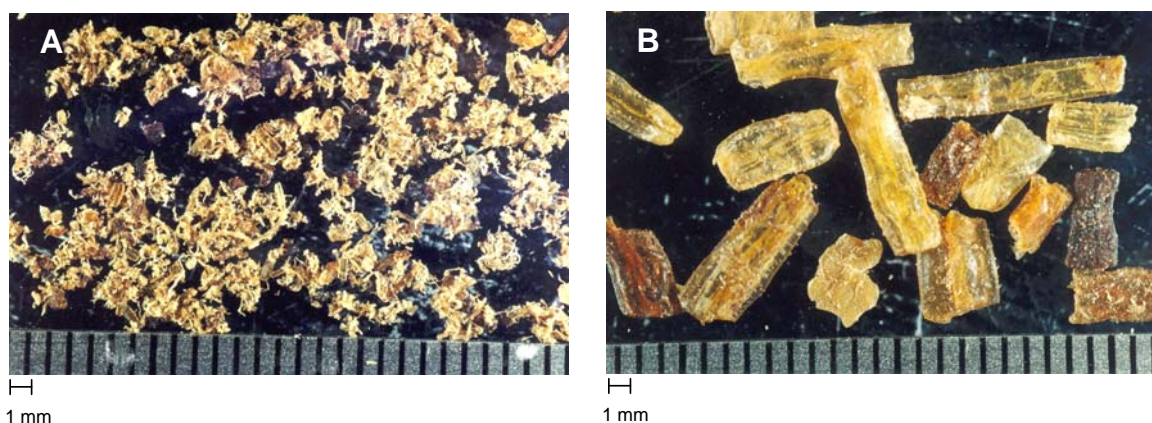
Os fragmentos de farinha de vísceras foram obtidos em diversas frações de solventes como éter de petróleo:clorofórmio (0/100; 10/90; 20/80 e 30/70, v/v). Apresentaram-se na forma de pequenas partículas, irregulares, ligeiramente gordurosas e com superfície rugosa. Os fragmentos de ossos apresentaram-se em lascas e com bordas irregulares, sendo que suas superfícies internas eram semelhantes a favos de mel e não eram tão grossas como as observadas para ossos de bovinos e suínos (Figura V.5 A e B).



**Figura V.5** - Fragmentos de farinha de vísceras.

A = fração obtida da mistura éter de petróleo:clorofórmio (0/100, v/v).  
B = fração obtida com a mistura éter de petróleo:clorofórmio (30/70, v/v).

Os fragmentos de penas foram obtidos em sua maioria da fração éter de petróleo:clorofórmio (30/70, v/v), sendo que alguns se assemelhavam a cabelos (Figura V.6 A), enquanto outros pareciam com tubos plásticos comprimidos (Figura V.6 B). As partículas tinham aparência vítrea, variando a cor de amarela ouro a marrom, com visualização de bordas serrilhadas nos canhões (hastes) retidos no processo de tamização. Uma prova confirmatória com solução de acetato de chumbo 2 % foi conduzida, indicando a presença de cistina a qual é encontrada em altos teores nas penas. Algumas partículas de bico e unhas foram detectadas. Estas tinham texturas semelhantes às dos ossos, mas com um lustro perolado similar ao dos dentes. A hidrólise completa removeu todas as evidências de uma pena normal, entretanto, raramente, ocorreu a detecção de penas não tratadas.

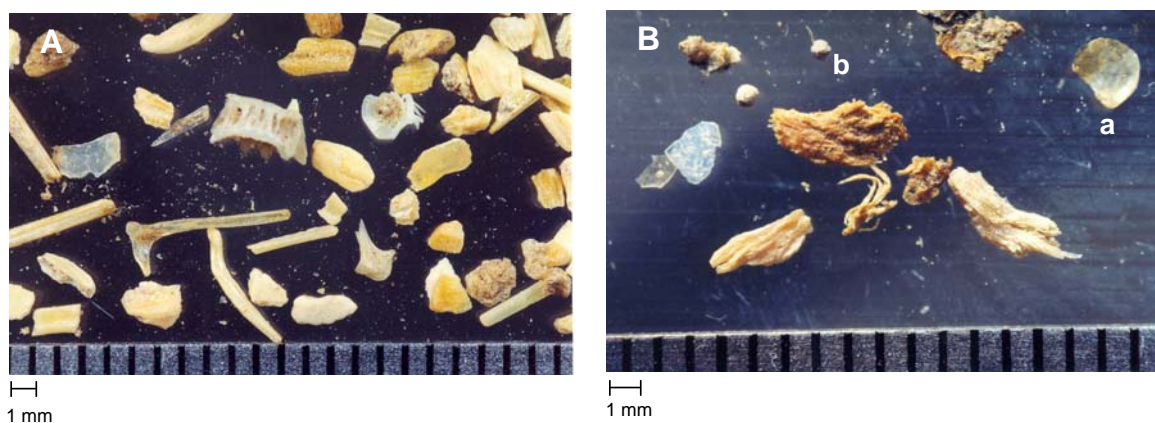


**Figura V.6** - Fragmentos de farinha de pena.

A = fração obtida da mistura éter de petróleo:clorofórmio (30/70, v/v). - aparência de cabelo.

B = fração obtida com a mistura éter de petróleo:clorofórmio (30/70, v/v) - aparência de tubos plásticos.

Fragmentos da farinha de peixe tinham aparência granular. O tecido ósseo proveniente da fração éter de petróleo:clorofórmio (0/100, v/v) apresentou-se opaco e perolado, às vezes translúcido, com algumas partículas finas e pontiagudas e outras em formato de vértebra, textura lisa em sua maioria ou rugosa, coloração branca amarelada e resistente à fragmentação (Figura V.7 A). Já na fração obtida a partir da mistura éter de petróleo:clorofórmio (30/70, v/v) foram observadas escamas translúcidas, pouco espessas, chatas ou curvas e com anéis concêntricos na superfície externa. Também foram detectados cristalinos de olhos como partículas esféricas, semi-transparentes, resistentes e, quando fragmentadas, divididas em camadas esféricas superpostas (Figura V.7 B).



**Figura V.7** - Fragmentos de farinha de peixe.

A = fração obtida da mistura éter de petróleo:clorofórmio (0/100, v/v). - partículas de ossos.  
 B = fração obtida com a mistura éter de petróleo:clorofórmio (30/70, v/v) - escama (a) e cristalino (b).

Todas as amostras brancas analisadas apresentaram resultados não detectados para os analitos pesquisados, indicando suficiente seletividade do método para distinguir os analitos dos componentes interferentes da matriz. Nos ensaios com amostras adicionadas entre 0,05 % e 0,80 % para farinhas de carne e ossos, ossos não calcinados, sangue, pena e peixe, e entre 0,10 % e 0,80 % para farinha de vísceras, todas as amostras apresentaram resultados detectados (Tabela V.2). Em nenhum caso houve identificação equivocada dos analitos. Estes resultados também confirmaram seletividade do método para distinguir os diferentes subprodutos proibidos (farinhas de carne e ossos, ossos não calcinados, sangue, vísceras, pena e peixe) em rações para ruminantes.

Quando este trabalho foi conduzido, somente estava em vigor no Brasil a portaria que proíbe o uso de proteínas de ruminantes na alimentação para ruminantes (BRASIL, 1997), de forma que os analitos farinhas de carne e ossos, ossos não calcinados e sangue eram proibidos, enquanto farinhas de vísceras, pena e peixe eram permitidos. No entanto, em 2004, houve a extensão desta proibição para fontes de proteínas e gorduras de origem animal (BRASIL, 2004 b).

Os menores níveis de concentração estudados nos quais os analitos foram detectados em todas as replicatas foram 0,05 % para farinhas de carne e ossos, ossos não calcinados, sangue, pena e peixe e 0,10 % para farinha de vísceras. Estes níveis foram estabelecidos como os limites de detecção do método para os respectivos analitos. Os limites alcançados foram inferiores ao limite de 1 % descrito por BRODMAN & MOOR (2003) para identificação de tecidos de mamíferos em ração por PCR e que o limite de 0,10 % obtido em um estudo interlaboratorial da União Européia, com vinte e oito

participantes, para detecção de farinha de ossos não calcinados em ração por microscopia (MOMCILOVIC & RASOOLY, 2000).

**Tabela V.2** - Resultados detectados e não detectados obtidos para amostras de ração para ruminantes brancas e adicionadas de subprodutos de origem animal em diferentes níveis de concentração

Concentração (%)	Analito											
	Carne e ossos		Ossos não calcinados		Sangue		Vísceras		Pena		Peixe	
	<i>n</i>	<i>D/ND</i>	<i>n</i>	<i>D/ND</i>	<i>n</i>	<i>D/ND</i>	<i>n</i>	<i>D/ND</i>	<i>n</i>	<i>D/ND</i>	<i>n</i>	<i>D/ND</i>
0,80	3	3/0	3	3/0	3	3/0	3	3/0	3	3/0	3	3/0
0,40	3	3/0	3	3/0	3	3/0	3	3/0	3	3/0	3	3/0
0,20	3	3/0	3	3/0	3	3/0	3	3/0	3	3/0	3	3/0
0,10	10	10/0	3	3/0	3	3/0	12	12/0	3	3/0	3	3/0
0,05	10	10/0	10	10/0	10	10/0	3	2/1	10	10/0	10	10/0
0,025	3	2/1	3	2/1	3	2/1	-	-	3	2/1	3	2/1
0,00	6	0/6	4	0/4	6	0/6	4	0/4	5	0/5	4	0/4

*n* = número de replicatas, *D/ND* = número de resultados detectados / número de resultados não detectados, - = não avaliado.

Os processos produtivos de rações envolvem tratamento térmico que desnatura proteínas e DNA. Esta é uma limitação significativa das técnicas baseadas na detecção de proteínas e DNA (MOMCILOVIC & RASOOLY, 2000). Técnicas imunológicas são simples, convenientes e permitem detecção de quantidades muito pequenas de proteínas. Entretanto, a maioria destas técnicas permite reações cruzadas com certos materiais como gelatina, fornecendo altos índices de resultados falsos positivos. Vários melhoramentos vêm sendo sugeridos no sentido de superar este problema, embora sem dados disponíveis sobre reprodutibilidade (ANSFIELD, 1994; HSIEH, SHEU & BRIDGMAN, 1998; MORALES et al., 1994). As técnicas baseadas em DNA apresentam limites de detecção aceitáveis e são espécie específicas, no entanto, não distinguem os diferentes tipos de tecidos, uma vez que as células de todos os tecidos de um organismo possuem DNA idêntico (MOMCILOVIC & RASOOLY, 2000). Embora a microscopia apresente dificuldades na distinção entre ossos originários de mamíferos e de aves, esta técnica é a única reconhecida pela União Européia (EC, 1998). A microscopia possui ainda inúmeras vantagens na detecção de subprodutos de origem animal em rações para ruminantes como não ser afetada pelo tratamento térmico das amostras, fornecer limites de detecção aceitáveis, baixos custos, simplicidade e rapidez, principalmente quando o

analista tem experiência na identificação de estruturas de origem animal (MOMCILOVIC & RASOOLY, 2000).

## **4 CONCLUSÕES**

O procedimento proposto foi adequado na validação intralaboratorial do método de determinação de subprodutos de origem animal em ração para ruminantes por microscopia, apresentando fácil aplicação prática e avaliação dos parâmetros de desempenho necessários para avaliação da eficiência de métodos qualitativos.

A seletividade do método atendeu aos requisitos analíticos estabelecidos, com limites de detecção suficientemente baixos. Tais parâmetros de desempenho indicaram adequação para uso do método validado em atividades de controle de qualidade de rações.

## **5 CONSIDERAÇÕES**

Os experimentos descritos neste capítulo foram realizados no Laboratório de Referência Agropecuária de Minas Gerais (LANAGRO-MG) do MAPA, em Pedro Leopoldo, Minas Gerais, Brasil.

A referência da publicação deste trabalho encontra-se descrita a seguir:  
SANCHES, R.L.; ALKMIN-FILHO, J.F.; SOUZA, S.V.C.; JUNQUEIRA, R.G. In-house validation of a method for detection of animal meals in ruminant feeds by microscopy. *Food Control*, v. 17, p. 85-72, 2006.

## CONCLUSÕES INTEGRADAS

Os protocolos de validação intralaboratorial de métodos de ensaio descritos na literatura não apresentaram harmonização em relação aos parâmetros de desempenho avaliados, terminologia, definições, delineamento experimental e análise de dados, com casos de recomendações equivocadas e contraditórias.

O procedimento proposto para validação intralaboratorial, desenvolvido empregando a forma direta de validação com uso de padrões e materiais de referência, permitiu a validação de métodos quantitativos e qualitativos em análises de alimentos.

Os parâmetros de desempenho linearidade, efeitos de matriz, seletividade, exatidão, precisão sob condições de repetitividade e reprodutibilidade parcial, limite de detecção, limite de quantificação, limite de decisão e capacidade de detecção foram apropriados para a avaliação da adequação para uso dos métodos estudados.

A fundamentação do delineamento experimental e dos testes de hipóteses nas características de cada parâmetro de desempenho e nas particularidades das análises de alimentos proporcionou eficiência ao processo de validação, bem como facilidade e praticidade de aplicação.

A estratégia de verificação de premissas básicas antes da aplicação das ferramentas estatísticas utilizadas garantiu confiabilidade ao procedimento de validação proposto.

O princípio da representatividade em termos de analitos, matrizes e níveis de concentração estudados contribuiu para a aplicabilidade do procedimento proposto.

A forma de apresentação dos testes estatísticos de modo a serem compreendidos por analistas de alimentos, permitirá a implementação dos mesmos em planilhas de cálculo comerciais, sem a necessidade de uso de programas estatísticos específicos.

Aplicabilidade do procedimento foi demonstrada para validações intralaboratoriais de métodos de análises de alimentos, em diferentes laboratórios e para diferentes escopos (analitos, níveis de concentração, matrizes e técnicas analíticas).

O tratamento de valores extremos para a remoção de dados discrepantes foi fundamental para uma aplicação correta do método dos mínimos quadrados ordinários (MMQO).

Os métodos validados pelo procedimento proposto apresentaram adequação para seus respectivos propósitos de uso em atividades de monitoramentos oficiais do *Department for Environment Food and Rural Affairs* (DEFRA) e do Ministério da Agricultura Pecuária e Abastecimento (MAPA).

## SUGESTÕES

Adoção do procedimento proposto como referência em processos de validação intralaboratorial de métodos de análises de alimentos.

Realização de estudos complementares das etapas não consideradas neste procedimento como:

- situações experimentais com erros nas duas variáveis;
- dados provenientes de distribuições diferentes da normal, heteroscedásticos, autocorrelacionados ou que seguem modelos não lineares;
- estimativa de incerteza da medição.



## REFERÊNCIAS BIBLIOGRÁFICAS

- AAFM (American Association of Feed Microscopists). *Manual of microscopic analysis of feedstuffs*. AAFM, 1992. p. 113-145.
- ABRAMSSON-ZETTERBERG, L.; SVENSSON, K. Semicarbazide is not genotoxic in the flow cytometry-based micronucleus assay in vivo. *Toxicol. Lett.*, v. 155, p. 211-217, 2005.
- ANFAR (Associação Nacional dos Fabricantes de Rações). *Compêndio Brasileiro de Alimentação Animal*. Matérias Primas. São Paulo: ANFAR, 1998. p. 26-27.
- ANSCOMBE, F.J. Graphs in statistical analysis. *Am. Stat.*, v. 27, p. 17-21, 1973.
- ANSFIELD, M. Production of sensitive immunoassay for detection of ruminant proteins in rendered animal material heated to > 130°C. *Food Agric. Immunol.*, v. 6, p. 419-433, 1994.
- ANTIGNAC, J.P.; LE BIZEC, B.; MONTEAU, F.; ANDRE, F. Validation of analytical methods based on mass spectrometric detection according to the “2002/657/EC” european decision: guideline and application. *Anal. Chim. Acta*, v. 483, p. 325-334, 2003.
- ANVISA (Agência Nacional de Vigilância Sanitária). REBLAS (Rede Brasileira de Laboratórios Analíticos em Saúde). Tradução dos Guias da EURACHEM. Métodos adequados ao uso - Guia de laboratório para validação de métodos e tópicos relacionados Disponível em: <<http://www.anvisa.gov.br/reblas/eurachem/index.htm>>. Acesso em: abril de 2007.
- AOAC (Association of Official Analytical Chemists). *AOAC Official methods of analysis*. Official method 964.07 - Microscopy of animal feed. Washington: AOAC, 1998. (a)
- AOAC (Association of Official Analytical Chemists). *AOAC Peer-verified methods program*. *Manual on polices and procedures*. Gaithersburg, Maryland: AOAC, 1998. 35 p. (b)

AOAC (Association of Official Analytical Chemists). *AOAC Official methods of analysis*. Appendix D: guidelines for collaborative study procedures to validate characteristics of a method of analysis. Washington: AOAC, 2002.

ARENA, J.P.; LIU, K.K.; PARESS, P.S.; FRAZIER, E.G.; CULLY, D.F.; MROZIK, H.; SCHAEFFER, J.M. The mechanism of action of avermectins in *Caenorhabditis elegans*: correlation between activation of glutamate-sensitive chloride current, membrane binding, and biological activity. *J. Parasitol.*, v. 81, p. 286-294, 1995.

ARMITAGE, P.; BERRY, G. *Statistical methods in medical research*. Oxford: Blackwell Science Ltd., 1994. 620 p.

BARINGHAUS, L.; DANSCHKE, R.; HENZE, N. Recent and classical tests for normality - a comparative study. *Commun. Stat. Simul. Comput.*, v. 18, p. 363-379, 1989.

BARRET, V.; LEWIS, T. *Outliers in statistical data*. 3 ed. New York: John Wiley, 1994. 604 p.

BELSLEY, D.A.; KUH, E.; WELSCH, R.E. *Regression diagnostics: identifying influential data and sources of collinearity*. New York: Wiley, 1980. 292 p.

BOX, G.E.P. Non-normality and tests on variances. *Biometrika*, v. 40, p. 318-335, 1953.

BOX, G.E.P.; COX, D.R. An analysis of transformations. *J. Roy. Statist. Soc.*, v. 26, p. 211-243, 1964.

BRAMBILLA, G.; DE FILLIPPIS, S. Trends in animal feed composition and the possible consequences on residue tests. *Anal. Chim. Acta*, v. 529, p. 7-13, 2005.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Portaria número 290 de 16 de julho de 1997. Proíbe o uso de qualquer fonte de proteína de ruminantes na alimentação de ruminantes. *Diário Oficial da União*. Brasília, 17 julho 1997, Seção 1, p. 151.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Instrução normativa número 7 de 17 de março de 2004. Proíbe a importação de ruminantes, seus produtos e subprodutos destinados a quaisquer fins, e de produtos para uso veterinário que contenham em sua composição insumos oriundos de ruminantes, quando originários ou procedentes de países que registraram casos autóctones de EEV, e de outros países considerados de risco pela SDA. *Diário Oficial da União*. Brasília, 18 março 2004, Seção 1, p. 3. (a)

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Instrução normativa número 8 de 25 de março de 2004. Proíbe em todo território nacional a produção, comercialização e utilização de produtos destinados à alimentação de ruminantes que contenham em sua composição proteínas e gorduras de origem animal. *Diário Oficial da União*. Brasília, 26 março 2004, Seção 1, p. 5. (b)

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Portaria número 50 de 20 de fevereiro de 2006. Aprova os Programas de Controle de Resíduos em Carne, Leite, Mel, Ovos e Pescado do exercício de 2006. *Diário Oficial da União*. Brasília, 03 março 2006, Seção 1, p. 15.

BRODMANN, P.D.; MOOR, D. Sensitive and semi-quantitative TaqMan™ real-time polymerase chain reaction systems for the detection of beef (*Bos taurus*) and the detection of the family *Mammalia* in food and feed. *Meat Sci.*, v. 65, p. 599-607, 2003.

BROWN, M.B.; FORSYTHE, A.B. Robust tests for the equality of variances. *J. Am. Stat. Assoc.*, v. 69, p. 364-367, 1974.

BROWN, P. The risk of bovine spongiform encephalopathy (“mad cow disease”) to human health. *J. Am. Med. Assoc.*, v. 278, p. 1008-1011, 1997.

BRUCE, B.; MINKKINEN, P.; RIEKKOLA, M.L. Practical method validation: validation sufficient for an analysis method. *Mikrochim. Acta*, v. 128, p. 93-106, 1998.

BRUCE, M.E.; WILL, R.G.; IRONSIDE, J.W.; MCCONNELL, I.; DRUMMOND, D.; SUTTIE, A.; MCCARDLE, L.; CHREE, A.; HOPE, J.; BIRKETT, C.; COUSENS, S.; FRASER, H.;

- BOSTOCK, C.J. Transmissions to mice indicate that “new variant” CJD is caused by the BSE agent. *Nature*, v. 389, p. 498-501, 1997.
- BURKE, S. Regression and calibration. *LC GC*, p. 13-18, 2001. (a)
- BURKE, S. Missing values, outliers, robust statistics & non-parametric methods. *LC GC*, p. 19-24, 2001. (b)
- BURNS, D.T.; DANZER, K.; TOWNSHEND, A. Use of the terms “recovery” and “apparent recovery” in analytical procedures. *Pure Appl. Chem.*, v. 74, p. 2201-2205, 2002.
- CAMPBELL, W.C.; FISHER, M.H.; STAPLEY, E.O.; ALBERS-SCHOENBERG, G.; JACOB, T.A. Ivermectin: a potent new antiparasitic agent. *Science*, v. 221, p. 823-828, 1983.
- CAULCUTT, R.; BODDY, R. *Statistics for analytical chemists*. New York: Chapman and Hall, 1983. 253 p.
- CHABALA, J.C.; MROZIK, H.; TOLMAN, R.L.; ESKOLA, P.; LUSI, A.; PETERSON, L.H.; WOODS, M.F.; FISHER, M.H. Ivermectin, a new broad-spectrum antiparasitic agent. *J. Med. Chem.*, v. 23, p. 1134-1136, 1980.
- CHI, I.C.; BLACKWELL, R.Q. A controlled retrospective study of blackfoot disease, an endemic peripheral gangrene disease in Taiwan. *Am. J. Epidemiol.*, v. 88, p. 7-24, 1968.
- CHIAP, P.; HUBERT, Ph.; BOULANGER, B.; CROMMEN, J. Validation of an automated method for the liquid chromatographic determination of atenolol in plasma: application of a new validation protocol. *Anal. Chim. Acta*, v. 391, p. 227-238, 1999.
- CHIKUNI, K.; TABATA, T.; KOSUGIYAMA, M.; MONMA, M.; SAITO, M. Polymerase chain reaction assay for detection of sheep and goat meats. *Meat Sci.*, v. 37, p. 337-345, 1994.

CODEX ALIMENTARIUS. Codex Alimentarius Commission. Codex Committee on Methods of Analysis and Sampling. (FAO/WHO). CX/MAS 02/10. *Single-laboratory validation. Consideration of harmonized IUPC guidelines for the in-house validation of methods of analysis.* Agenda item 8a) on the 24th session. Budapest, Hungary, 2002.

CODEX ALIMENTARIUS. Codex Alimentarius Commission. Codex Committee on Methods of Analysis and Sampling. (FAO/WHO). CL 2005/44 MAS. *Draft Guidelines for evaluating acceptable methods of analysis.* Rome, Italy, 2005. (a)

CODEX ALIMENTARIUS. Codex Alimentarius Commission. Toxicological evaluation of certain veterinary drug residues in food. Disponível em: <<http://www.codexalimentarius.net/web/jecfa.jsp>>. Acesso em: 28 de abril de 2005. (b)

COLCHESTER, A.C.F.; COLCHESTER, N.T.H. The origin of bovine spongiform encephalopathy: the human prion disease hypothesis. *Lancet*, v. 366, p. 856-861, 2005.

CONCEPCION, G.P.; PADLAN, E.A. Are humans getting “mad-cow disease” from eating beef, or something else? *Med. Hypotheses*, v. 60, p. 699-701, 2003.

CONNELLY, A.; NUGENT, A.; O'KEEFE, M.; MULDER, P.P.J.; VAN RHIJN, J.A.; KOVACSICS, L.; FODOR, A.; MC CRACKEN, R.J.; KENNEDY, D.G. Isolation of bound residues of nitrofurans from tissue by solid-phase extraction with determination by liquid chromatography with UV and tandem mass spectrometric detection. *Anal. Chim. Acta*, v. 483, p. 91-98, 2003.

COOK, R.D. Detection of influential observation in linear regression. *Technometrics*, v. 19, p. 15-18, 1977.

COPPENS, P.; DA SILVA, M.F.; PETTMAN, S. European regulations on nutraceuticals, dietary supplements and functional foods: a framework based on safety. *Toxicology*, v. 221, p. 59-74, 2006.

CURRIE, L.A. Limits for qualitative detection and quantitative determination. *Anal. Chem.*, v. 40, p. 586-593, 1968.

- CURRIE, L.A. Nomenclature in evaluation of analytical methods including detection and quantification capabilities. *Pure Appl. Chem.*, v. 67, p. 1699-1723, 1995.
- CURRIE, L.A. Detection and quantification limits: origins and historical overview. *Anal. Chim. Acta*, v. 391, p. 127-134, 1999.
- D'AGOSTINO, R.B.; STEPHENS, M.I.A. *Goodness-of-fit techniques*. New York: Marcel Dekker, 1986. p. 367-419.
- DANZER, K.; CURRIE, L.A. Guidelines for calibration in analytical chemistry. Part 1. Fundamentals and single component calibration. *Pure Appl. Chem.*, v. 70, p. 993-1014, 1998.
- DEDINA, J.; TSALEV, D.L. *Hydride generation atomic absorption spectrometry*. Chichester: Wiley, 1995. 526 p.
- DEN BOEF, G.; HULANICKI, A. Recommendations for the usage of selective, selectivity and related terms in analytical chemistry. *Pure Appl. Chem.*, v. 55, p. 553-556, 1983.
- DEN HARTOG, J. Feed for food: HACCP in the animal feed industry. *Food Control*, v. 14, p. 95-99, 2003.
- DESIMONI, E. About  $CC\alpha$  and  $CC\beta$  as introduced by the Commission Decision of 12 august 2002 implementing council directive 96/23/EC. *Accred. Qual. Assur.*, v. 9, p. 724-725, 2004.
- DORMONT, D. Prions, BSE and food. *Int. J. Food Microbiol.*, v. 78, p. 181-189, 2002.
- DOULL, J.; KLAASSEN, C.D.; AMDUR, M.O. *Toxicology: the basic science of poisons*. New York: Macmillan, 1980. p. 435-438.
- DRAPER, N.R.; HUNTER, W.G. Transformations: some examples revisited. *Technometrics*, v. 11, p. 23-40, 1969.

- DRAPER, N.R.; SMITH, H. *Applied regression analysis*. New York: Wiley. 1998. 706 p.
- DUCE, I.R.; SCOTT, R.H. Actions of dihydroavermectin B1a on insect muscle. *Br. J. Pharmac.*, v. 85, p. 395-401, 1985.
- DUFOUR, J.M.; FARHAT, A.; GARDIOL, L.; KHALAF, L. Simulation-based finite sample normality tests in linear regressions. *Econom. J.*, v. 1, p. 154-173, 1998.
- DURBIN, J.; WATSON, G.S. Testing for serial correlation in least squares regression ii. *Biometrika*, v. 38, p. 159-178, 1951.
- EAL (European Cooperation for Accreditation of Laboratories). *EAL-P11. Validation of test methods: general principles and concepts*. Paris: EAL, 1997. 13 p.
- EBBEHOJ, K.F.; THOMSEN, P.D. Differentiation of closely related species by DNA hybridization. *Meat Sci.*, v. 30, p. 359-366, 1991.
- EC (European Commission). Commission directive number 1442/95. *Official Journal of the European Communities*, 1995, L 143, p. 26-30.
- EC (Commission of the European Communities). Commission Directive 98/99/EC of 13 November 1998 establishing guidelines for the microscopic identification and estimation of constituents of animal origin for the official control of feedingstuffs. *Official Journal*, 1998, L 318, p. 0045-0050.
- EC (European Commission). 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, L 221/8.
- EC (European Commission). Commission regulation number 2004/1/EC of 6 January 2004. Amending Directive 2002/72/EC as regards the suspension of the use of azodicarbonamide as blowing agent. *Official Journal of the European Communities*, 2004, L 007, p. 45-46.

EC (European Commission). Commission regulation number of 23 February 2006. Laying down the methods of sampling and analysis for the official control of the levels of mycotoxins in foodstuffs. *Official Journal of the European Communities*, 2006, L 70, p. 12-34.

EFSA (European Food Safety Authority). Panel on food additives, flavourings, processing aids and materials in contact with food. Preliminary advice on the possible occurrence of semicarbazide in packaged foods. Press release 28 July 2003. Disponível em: <<http://www.efsa.eu.int/AFC/adhoc SEM1.html>>. Acesso em: dezembro de 2003. (a)

EFSA (European Food Safety Authority). Panel on food additives, flavourings, processing aids and materials in contact with food. Additional advice on semicarbazide, in particular related to baby food. Press release 9 October 2003. Disponível em: <[http://www.efsa.eu.int/p\\_foodadd\\_en.html](http://www.efsa.eu.int/p_foodadd_en.html)>. Acesso em: dezembro de 2003. (b)

EFSA (European Food Safety Authority). Background information on semicarbazide found in foods packaged in glass jars and bottles. Press release 15 October 2003. Disponível em: <[http://www.efsa.eu.int/pdf/pressreal20031015\\_en.pdf](http://www.efsa.eu.int/pdf/pressreal20031015_en.pdf)>. Acesso em: dezembro de 2003. (c)

ELLERTON, R.R.W. Comments on regression through the origin. *Anal. Chem.*, v. 52, p. 1151-1152, 1980.

EURACHEM. *The fitness for purpose of analytical methods, a laboratory guide to method validation and related topics*. Teddington: LGC, 1998. 61 p.

EURACHEM/CITAC. *Quantifying uncertainty in analytical measurements*. Teddington: LGC, 2000. 120 p.

FAJGELJ, A.; AMBRUS, A. Guidelines for single-laboratory validation of analytical methods for trace-level concentrations of organic chemicals. In: FAJGELJ, A.; AMBRUS, A. (Ed.) *Principles and practices of method validation*. Cambridge: Royal Society of Chemistry, 2000. p. 179-252.



- FEINBERG, M.; RAGUÈNÈS, N. Development and application of a standardized validation procedure for food chemistry laboratories. *Anal. Chim. Acta*, v. 391, p. 239-252, 1999.
- FERRANDO, R.; HENRI, N. *Determinacion microscópica de los componentes de los piensos*. Zaragoza: Editorial Acribia, 1966. 14 p.
- FILLIBEN, J.J. The probability plot correlation coefficient test for normality. *Technometrics*, v. 17, p. 111-117, 1975.
- FISHBEIN, L. Transmissible spongiform encephalopathies, hypotheses and food safety: an overview. *Sci. Total Environ.*, v. 217, p. 71-82, 1998.
- FOWLER, B.A.; ISHINISHI, N.; TSUCHIYA, T.; VAHTER, M. In: FRIBERG, L.; NORDBERD, G.F.; VOAK, V.B. (Ed.) *Handbook of the toxicology of metals*. Amsterdam: Elsevier, 1979. 293 p.
- FRUMKIN, H.; THUN, M.J. Arsenic. *Cancer J. Clin.*, v. 51, p. 254-262, 2001.
- GHANI, A.C.; FERGUSON, N.M.; DONNELLY, C.A.; HAGENAARS, T.J.; ANDERSON, R.M. Epidemiological determinants of the pattern and magnitude of the vCJD epidemic in Great Britain. *Proc. R. Soc. Lond. B*, v. 265, p. 2443-2452, 1998.
- GINN, R.; WILSON, L.; SOUZA, S.V.C.; DE LA CALLE, M.B. Determination of semicarbazide in baby food by liquid chromatography tandem mass spectrometry: Interlaboratory validation study. *J. AOAC Int.*, v. 89, p. 728-734, 2006.
- GIPS, M.; BRIDZY, M.; BAREL, S.; SOBACK, S. A high performance liquid chromatography method for detection of chloramphenicol, nitrofurazone and furazolidone residues in muscle, liver and kidney. In: EURO RESIDUE, 2, 1993, Veldhoven, The Netherlands. *Anais...* Veldhoven: Euro Residue, 1993. p. 313-317.
- GREEN, J.M. A practical guide to analytical method validation. *Anal. Chem.*, v. 68, p. 305A-309A, 1996.

- GRUBBS, F. Procedures for detecting outlying observations in samples. *Technometrics*, v. 11, p. 1-21, 1969.
- HARTLEY, H.O. The maximum F-ratio as a shortcut test for heterogeneity of variance. *Biometrika*, v. 37, p. 308-312, 1950.
- HAYATSU, H.; TAKEISHI, K.; UKITA, T. The modification of nucleosides and nucleotides. III. A selective modification of cytidine with semicarbazide. *Biochim. Biophys. Acta*, v. 123, p. 445-457, 1966.
- HAYATSU, H.; UKITA, T. The modification of nucleosides and nucleotides. IV. The reactions of semicarbazide with nucleic acids. *Biochim. Biophys. Acta*, v. 123, p. 458-470, 1966.
- HEI, T.K.; FILIPIC, M. Role of oxidative damage in the genotoxicity of arsenic. *Free Radic. Biol. Med.*, v. 37, p. 574-581, 2004.
- HILL, A.R.C.; REYNOLDS, S.L. Guidelines for in-house validation of analytical methods for pesticide residues in food and animal feeds. *Analyst*, v. 124, p. 953-958, 1999.
- HIRAKAWA, K.; MIDORIKAWA, K.; OIKAWA, S.; KAWANISHI, S. Carcinogenic semicarbazide induces sequence-specific DNA damage through the generation of reactive oxygen species and the derived organic radicals. *Mutat. Res.*, v. 536, p. 91-101, 2003.
- HOFFMANN, R.; VIEIRA, S. *Análise de regressão. Uma introdução à econometria*. São Paulo: Editora Hucitec, 1987. 379 p.
- HOLAK, W. Gas-sampling technique for arsenic determination by atomic absorption spectrophotometry. *Anal. Chem.*, v. 41, p. 1712-1713, 1969.
- HOOGENBOOM, L.A.P.; POLMAN, TH.H.G. Simultaneous detection of protein bound residues of the nitrofurans furazolidone, furaltadone, nitrofurazone and nitrofurantoin. In: EURO RESIDUE, 2, 1993, Veldhoven, The Netherlands. *Anais...* Veldhoven: Euro Residue, 1993. p. 376-381.

- HORNE, E.; CADOGAN, A.; O'KEEFFE, M.; HOOGENBOOM, L.A.P. Analysis of protein-bound metabolites of furazolidone and furaltadone in pig liver by high-performance liquid chromatography and liquid chromatography-mass spectrometry. *Analyst*, v. 121, p. 1463-1468, 1996.
- HORWITZ, W. Evaluation of analytical methods used for regulation of food and drugs. *Anal. Chem.*, v. 54, p. 67A-76A, 1982.
- HORWITZ, W. Protocol for the design, conduct and interpretation of method-performance studies. *Pure Appl. Chem.*, v. 67, p. 331-343, 1995.
- HORWITZ, W. *Validation: an invisible component of measurement*. 2003, 10 p. Disponível em: <<http://www.aoac.org/dietsupp6/Dietary-Supplement-web-site/HorwitzValid.pdf>>. Acesso em: dezembro de 2003.
- HSIEH, Y.H.; SHEU, S.C.; BRIDGMAN, R.C. Development of a monoclonal antibody specific to cooked mammalian meats. *J. Food Prot.*, v. 61, p. 476-481, 1998.
- HUBAUX, A.; VOS, G. Decision and detection limits for linear calibration curves. *Anal. Chem.*, v. 42, p. 849-855, 1970.
- HUBER, L. Validation of analytical methods: review and strategy. *LC/GC Int.*, Feb., p. 96-105, 1998.
- HUND, E.; MASSART, D.L.; SMEYERS-VERBEKE, J. Inter-laboratory studies in analytical chemistry. *Anal. Chim. Acta*, v. 423, p. 145-165, 2000.
- HUNG, D.Q.; NEKRASSOVA, O.; COMPTON, R.G. Analytical methods for inorganic arsenic in water: a review. *Talanta*, v. 64, p. 269-277, 2004.
- HUNT, D.J.; PARKES, H.C.; LUMLEY, I.D. Identification of the species of origin of raw and cooked meat products using oligonucleotide probes. *Food Chem.*, v. 60, p. 437-442, 1997.

HUR, K.; KIM, J.; CHOI, S.; CHOI, E.; CARP, R.I.; KIM, Y. The pathogenic mechanisms of prion diseases. *Mech. Ageing Dev.*, v. 123, p. 1637-1647, 2002.

HUSS, W. *Microscopy and quality control in the manufacture of animals feed*. Hohenheim: Roche, 1975. 63 p.

ICH (International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use). *Validation of analytical procedures: methodology*. Geneva: ICH/IFPMA, 1996. 8 p.

ILAC (International Laboratory Accreditation Conference). Disponível em: <<http://www.ilac.org>>. Acesso em: 30 de agosto de 2002.

INMAN, E.L.; FRISCHMANN, J.K.; JIMENEZ, P.J.; WINKEL, G.D.; PERSINGER, M.L.; RUTHEFORD, B.S. General method validation guidelines for pharmaceutical samples. *J. Chromatogr. Sci.*, v. 25, p. 252-256, 1987.

INMETRO (Instituto Nacional de Metrologia, Normalização e Qualidade Industrial). *DOQ-CGCRE-008. Orientações sobre validação de métodos de ensaios químicos*. Rio de Janeiro: INMETRO, 2003. 35 p.

ISO (International Standards Organization). *ISO Guide 30. Terms and definitions used in connection with reference materials*. Geneva: ISO, 1992. 8 p.

ISO (International Standards Organization). *International vocabulary of basic and general terms in metrology - VIM*. Geneva: ISO, 1993. 59 p. (a)

ISO (International Standards Organization). *ISO 3534-1. Statistics - Vocabulary and symbols - Part 1: Probability and general statistical terms*. Geneva: ISO, 1993. 53 p. (b)

ISO (International Standards Organization). *ISO 5725-1. Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions*. Geneva: ISO, 1994. 17 p. (a)

ISO (International Standards Organization). *ISO 5725-2. Accuracy (trueness and precision) of measurement methods and results - Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method.* Geneva: ISO, 1994. 42 p. (b)

ISO (International Standards Organization). *ISO 5725-3. Accuracy (trueness and precision) of measurement methods and results - Part 3: Intermediate measures of the precision of a standard measurement method.* Geneva: ISO, 1994. 25 p. (c)

ISO (International Standards Organization). *ISO 5725-4. Accuracy (trueness and precision) of measurement methods and results - Part 4: Basic methods for the determination of the trueness of a standard measurement method.* Geneva: ISO, 1994. 23 p. (d)

ISO (International Standards Organization). *ISO 5725-6. Accuracy (trueness and precision) of measurement methods and results - Part 6: Use in practice of accuracy values.* Geneva: ISO, 1994. 41 p. (e)

ISO (International Standards Organization). *ISO Guide 98 Guide to the expression of uncertainty in measurement (ISO GUM).* Geneva: ISO, 1995. 105 p.

ISO (International Standards Organization). *ISO 11095. Linear calibration using reference materials.* Geneva: ISO, 1996. 29 p.

ISO (International Standards Organization). *ISO/IEC Guide 43-1. Proficiency testing by interlaboratory comparisons - Part 1: Development and operation of proficiency testing schemes.* Geneva: ISO, 1997. 16 p. (a)

ISO (International Standards Organization). *ISO 11843-1. Capability of detection - Part 1: Terms and definitions.* Geneva: ISO, 1997. 10 p. (b)

ISO (International Standards Organization). *ISO 5725-5. Accuracy (trueness and precision) of measurement methods and results - Part 5: Alternative methods for the determination of the precision of a standard measurement method.* Geneva: ISO, 1998. 56 p.

- ISO (International Standards Organization). *ISO 11843-2. Capability of detection - Part 2: Methodology in the linear calibration case*. Geneva: ISO, 2000. 24 p.
- ISO (International Standards Organization). *ISO 16140. Microbiology of food and feeding stuffs - Protocol for the validation of alternative methods*. Geneva: ISO, 2003. 74 p.
- ISO (International Standard Organization). *ISO 9000. Quality management systems - Fundamentals and vocabulary*. Geneva: ISO, 2005. 30 p. (a)
- ISO (International Standard Organization). *ISO/IEC 17025. General requirements for the competence of testing and calibration laboratories*. Geneva: ISO, 2005. 28 p. (b)
- JEMMI, T.; SCHLOSSER, H. Tierartbestimmung aus mariniertem und erhitztem Fleisch mittels isoelektrischer Fokussierung. *Fleischwirtsch.*, v. 73, p. 600-602, 1993.
- JENKE, D.R. Chromatographic method validation: a review of current practices and procedures. I. General concepts and guidelines. *J. Liq. Chrom. & Rel. Technol.*, v. 19, p. 719-736, 1996.
- JOHNSTON, J. *Econometric methods*. New York: McGraw-Hill Book Company, 1984. p. 287-342.
- JOINER, R.; VIDAL, F.; MARKS, H. A new powdered agent for flour maturing. *Cereal Ind.*, v. 40, p. 539-553, 1963.
- KELLY, P.C. Outlier detection in collaborative studies. *J. Assoc. Off. Anal. Chem.*, v. 73, p. 58-64, 1990.
- KIMBERLIN, R.H. Bovine spongiform encephalopathy: an appraisal of the current epidemic in the United Kingdom. *Intervirology*, v. 35, p. 208-218, 1993.
- KING, B. In: EUROLAB SYMPOSIUM QUALITY MANAGEMENT AND ASSURANCE IN TESTING LABORATORIES, 1, 1992, Strabourg, France. *Anais...* Strabourg: Eurolab Organization for testing in Europe, 1992. p. 68-72.

- KING, B. *In-house method validation. A guide for chemical laboratories. LGC Limited.* 2003.
- KUMAR, L.; TOOTHILL, J.R.; HO, K. Determination of nitrofurantoin residues in poultry muscle tissues and eggs by liquid chromatography. *J. AOAC Int.*, v. 77, p. 591-595, 1994.
- KUTTATHARMMAKUL, S.; MASSART, D.L.; SMEYERS-VERBEKE, J. Comparison of alternative measurement methods. *Anal. Chim. Acta*, v. 391, p. 203-225, 1999.
- LEARY, J.J.; MESSICK, E.B. Constrained calibration curves: a novel application of Lagrange multipliers in analytical chemistry. *Anal. Chem.*, v. 57, p. 956-957, 1985.
- LEITNER, A.; ZÖLLNER, P.; LINDNER, W. Determination of metabolites of nitrofurantoin antibiotics in animal tissue by high-performance liquid chromatography-tandem mass spectrometry. *J Chromatogr. A*, v. 939, p. 49-58, 2001.
- LEVENE, H. Robust tests for equality of variances. In: OLKIN, I.; GHURYE, S.G.; HOEFFDING, W.; MADOW, W.G.; MANN, H.B. (Ed.) *Contributions to probability and statistics*. Stanford: Stanford University Press, 1960. p. 278-292.
- LITTLE, T.M.; HILLS, F.J. *Statistical methods in agricultural research*. Davis: University of California, 1975. 242 p.
- MAC NEIL, J.D.; PATTERSON, J.; MARTZ, V. Validation of analytical methods – Providing your method is “fit for purpose”. In: FAJGELJ, A.; AMBRUS, A. (Ed.) *Principles and practices of method validation*. Cambridge: Royal Society of Chemistry, 2000. p. 100-107.
- MARDIA, K.V. Tests of univariate and multivariate normality. In: KRISHNAIAH, P.R. (Ed.) *Handbook of statistics: analysis of variance*. Amsterdam: North Holland, 1980. p. 279-320.
- MARK, H. Application of an improved procedure for testing the linearity of analytical methods to pharmaceutical analysis. *J. Pharm. Biomed. Anal.*, v. 33, p. 7-20, 2003.

- MARK, H. WORKMAN, J. Linearity in calibration: the Durbin-Watson statistic: a discussion of how DW can be a useful tool when different statistical approaches show different sensitivities to particular departures from the ideal. *Spectroscopy*, v. 20, p. 34-38, 2005. (a)
- MARK, H. WORKMAN, J. Linearity in calibration: the importance of non-linearity. *Spectroscopy*, v. 20, p. 56-59, 2005. (b)
- MAROTO, A.; BOQUÉ, R.; RIU, J.; RIUS, F.X. Estimation of measurement uncertainty by using regression techniques and spiked samples. *Anal. Chim. Acta*, v. 446, p. 133-145, 2001.
- MAROTO, A.; RIU, J.; BOQUÉ, R.; RIUS, F.X. Estimating uncertainties of analytical results using information from the validation process. *Anal. Chim. Acta*, v. 391, p. 173-185, 1999.
- MATSUNAGA, T.; CHIKUNI, K.; TANABE, R.; MUROYA, S.; SHIBATA, K.; YAMADA, J.; SHINMURA, Y. A quick and simple method for the identification of meat species and meat products by PCR assay. *Meat Sci.*, v. 51, p. 143-148, 1999.
- MC KELLAR, Q.A. Ecotoxicology and residues of anthelmintic compounds. *Vet. Parasitol.*, v. 72, p. 413-435, 1997.
- MCSHEEHY, S.; SZPUNAR, J.; MORABITO, R.; QUEVAUVILLER, P. The speciation of arsenic in biological tissues and the certification of reference materials for quality control. *Trends Anal. Chem.*, v. 22, p. 191-209, 2003.
- MELOUN, M.; MILITKÝ, J. Detection of single influential points in OLS regression model building. *Anal. Chim. Acta*, v. 439, p. 169-191, 2001.
- MEYER, P.C.; ZUND, R.E. *Statistical methods in analytical chemistry*. New York: John Wiley & Sons, 1993. p. 81-134.
- MEYER, R.; CANDRIAN, U.; LÜTHY, J. Detection of pork in heated meat products by polymerase chain reaction. *J. AOAC Int.*, v. 77, p. 617-622, 1994.



- MEYER, R.; HÖFELEIN, C.; LÜTHY, J.; CANDRIAN, U. Polymerase chain reaction-restriction fragment length polymorphism analysis: a simple method for species identification in food. *J. AOAC Int.*, v. 78, p. 1542-1551, 1995.
- MILLER, J.C.; MILLER, J.N. *Statistics for analytical chemistry*. New York: Ellis Horwood Limited. 1993. p. 101-141.
- MINDAK, W.L.; DOLAN, S.P. Determination of arsenic and selenium in food using a microwave digestion-dry ash preparation and flow injection hydride generation atomic absorption spectrometry. *J. Food Compos. Anal.*, v. 12, p. 111-122, 1999.
- MOMCILOVIC, D.; RASOOLY, A. Detection and analysis of animal materials in food and feed. *J. Food Prot.*, v. 63, p. 1602-1609, 2000.
- MONTGOMERY, D.C.; RUNGER, G.C. *Applied statistics and probability for engineers*. New York: John Wiley & Sons, 1994. p. 471-624.
- MORALES, P.; GARCÍA, T.; GONZÁLES, I.; MARTÍN, R.; SANZ, B.; HERNÁNDEZ, P.E. Monoclonal antibody detection of porcine meat. *J. Food Prot.*, v. 57, p. 146-149, 1994.
- MULHOLLAND, M.; HIBBERT, D.B. Linearity and the limitations of least squares calibration. *J. Chromatogr. A*, v. 762, p. 73-82, 1997.
- NATA (National Association of Testing Authorities - Australia). *Technical note 17. Format and content of test methods and procedures for validation and verification of chemical test methods*. Sydney: NATA, 1997. 8 p.
- NÄYKKI, T.; PERÄMÄKI, P.; KUJALA, J.; MIKKONEN, A. Optimization of a flow injection hydride generation atomic absorption spectrometric method for the determination of arsenic, antimony and selenium in iron chloride/sulfate-based water treatment chemical. *Anal. Chim. Acta*, v. 439, p. 229-238, 2001.
- NIEDERER, M.; BOLLHALDER, R. Identification of species specific central nervous tissue by gas chromatography-mass spectrometry (GC-MS) a possible method for supervision of meat products and cosmetics. *Mitt. Gebiete Lebensm. Hyg.*, v. 92, p. 133-144, 2001.

- PATRA, M.; BHOWMIK, N.; BANDOPADHYAY, B.; SHARMA, A. Comparison of mercury, lead and arsenic with respect to genotoxic effects on plant systems and the development of genetic tolerance. *Environ. Exp. Bot.*, v. 52, p. 199-223, 2004.
- PECORELLI, I.; BIBI, R.; FIORONI, L.; GALARINI, R. Validation of a confirmatory method for the determination of sulphonamides in muscle according to the European Union regulation 2002/657/EC. *J. Chromatogr. A.*, v. 1032, p. 23-29, 2004.
- PEREIRA, A.S.; DONATO, L.J.; DE NUCCI, G. Implications of the use of semicarbazide as a metabolic target of nitrofurazone contamination in coated products. *Food Addit. Contam.*, v. 21, p. 63-69, 2004.
- PERETZ, D.; WILLIAMSON, R.A.; MATSUNAGA, Y.; SERBAN, H.; PINILLA, C.; BASTIDAS, R.B.; ROZENSHTEYN, R.; JAMES, T.L.; HOUGHTEN, R.A.; COHEN, F.E.; PRUSINER, S.B.; BURTON, D.R. A conformational transition at the N terminus of the prion protein features in formation of the scrapie isoform. *J. Mol. Biol.*, v. 273, p. 614-622, 1997.
- PIERCE, F.D.; BROWN, H.R. Inorganic interference study of automated arsenic and selenium determination with atomic absorption spectrometry. *Anal. Chem.*, v. 48, p. 693-695, 1976.
- POLLMEIER, M.; MAIER, S.; MORIARTY, K.; DEMONTIGNY, P. High-performance liquid chromatography assay for the determination of a semisynthetic avermectin analog (eprinomectin) in bovine milk at parts per billion levels - method development and validation. *J. Chromatogr. B*, v. 772, p. 99-105, 2002.
- PRUSINER, S.B.; MCKINLEY, M.P.; GROTH, D.F.; BOWMAN, K.A.; MOCK, N.I.; COCHRAN, S.P.; MASIARZ, F.R. Scrapie agent contains a hydrophobic protein. *Proc. Natl. Acad. Sci. USA*, v. 78, p. 6675-6679, 1981.
- RIPLEY, B.D.; THOMPSON, M. Regression techniques for the detection of analytical bias. *Analyst*, v. 112, p. 377-383, 1987.

RIU, J.; RIUS, F.X. Univariate regression models with errors in both axes. *J. Chemom.*, v. 9, p. 343-362, 1995.

ROUDAUT, B. Multiresidue method for the determination of avermectin and moxidectin residues in the liver using HPLC with fluorescence detection. *Analyst*, v. 123, p. 2541-2544, 1998.

RSC (Royal Society of Chemistry). *Analytical Methods Committee Technical Brief: What should be done with results below the detection limit? Mentioning the unmentionable.* RSC, 2001. 2 p. (a)

RSC (Royal Society of Chemistry). *Analytical Methods Committee Technical Brief: Robust statistics: a method of coping with outliers.* RSC, 2001. 2 p. (b)

RSC (Royal Society of Chemistry). *Analytical Methods Committee Technical Brief: Terminology - the key to understanding the analytical science. Part 1: accuracy, precision and uncertainty.* RSC, 2003. 2 p.

RSC (Royal Society of Chemistry). *Analytical Methods Committee Technical Brief: The amazing Horwitz function.* RSC, 2004. 2 p.

RSC (Royal Society of Chemistry). *Analytical Methods Committee Technical Brief: Is my calibration linear?.* RSC, 2005. 2 p. (a)

RSC (Royal Society of Chemistry). *Analytical Methods Committee Recommendation: The estimation and use of recovery factors.* RSC, 2005. 2 p. (b)

RUPASINGHE, T.; CARDWELL, T.J.; CATTRALL, R.W.; POTTER, I.D.; KOLEV, S.D. Determination of arsenic by pervaporation-flow injection hydride generation and permanganate spectrophotometric detection. *Anal. Chim. Acta*, v. 510, p. 225-230, 2004.

RUPP, H.S.; MUNNS, R.K.; LONG, A.R. Simultaneous determination of nitrofurazone and furazolidone in shrimp (*Penaeus vannamei*) muscle tissue by liquid chromatography with UV detection. *J. AOAC Int.*, v. 76, p. 1235-1239, 1993.

- RUTLEDGE, D.N.; BARROS, A.S. Durbin-Watson statistic as a morphological estimator of information content. *Anal. Chim. Acta*, v. 454, p. 277-295, 2002.
- RUZICKA, J.; HANSEN, E.H. *Flow Injection Analysis*. New York: Wiley, 1988. 498 p.
- RUZICKA, J.; HANSEN, E.H. Flow injection analysis - where are we heading? *Trends Anal. Chem.*, v. 17, p. 69-73, 1998.
- RYAN, T.A.; JOINER, B.L. *Normal probability plots and tests for normality*. The State College: Pennsylvania State University, 1976. 15 p.
- SALISBURY, C.D.C. Modified method for the determination of ivermectin residues in animal tissues. *J. AOAC Int.*, v. 76, p. 1149-1151, 1993.
- SAMS, R. Chemical assay of avermectins by high performance liquid chromatography with fluorescence detection. *Vet. Parasitol.*, v. 48, p. 59-66, 1993.
- SCHENCK, F.J.; LAGMAN, L.H. Multiresidue determination of abamectin, doramectin, ivermectin, and moxidectin in milk using liquid chromatography and fluorescence detection. *J. AOAC Int.*, v. 82, p. 1340-1344, 1999.
- SCHWARTZ, L.M. Effect of constraints on precision of calibration analyses. *Anal. Chem.*, v. 58, p. 246-250, 1986.
- SCOTT, M.R.; WILL, R.; IRONSIDE, J.; NGUYEN, H.B.; TREMBLAY, P.; DEARMOND, S.J.; PRUSINER, S.B. Compelling transgenetic evidence for transmission of bovine spongiform encephalopathy prions to humans. *Proc. Natl. Acad. Sci. USA*, v. 96, p. 15137-15142, 1999.
- SHABIR, G.A. Validation of high-performance liquid chromatography methods for the pharmaceutical analysis. Understanding the differences and similarities between validation requirements of the US Food and Drug Administration, the US Pharmacopeia and the International Conference on Harmonization. *J. Chromatogr. A*, v. 987, p. 57-66, 2003.

- SNEDECOR, G.W.; COCHRAN, W.G. *Statistical methods*. Ames: Iowa State University, 1989. 503 p.
- SOUZA, S.V.C.; SILVA, G.; DINIZ, M.H.G.M.; SANTOS, E.V.; LIMA, J.A.; TEODORO, J.C. Determinação de resíduos de avermectinas em fígado bovino por cromatografia líquida de alta eficiência. *Ciênc. Tecnol. Aliment.*, v. 23, p. 54-58, 2003.
- STADLER, R.H.; MOTTIER, P.; GUY, P.; GREMAUD, E.; VARGA, N.; LALLJIE, S.; SCHULTZ, C.; WHITAKER, R.; KINTSCHER, J.; DUDLER, V.; CASTLE, L.; READ, W.A. Semicarbazide is a minor thermal decomposition product of azodicarbonamide used in the gaskets of certain food jars. *Analyst*, v. 129, p. 276-281, 2004.
- STRONG, F.C. Regression line that starts at the origin. *Anal. Chem.*, v. 51, p. 298-299, 1979.
- TAVERNIERS, I.; DE LOOSE, M.; VAN BOCKSTAELE, E. Trends in quality in the analytical laboratory. II. Analytical method validation and quality assurance. *Trends Anal. Chem.*, v. 23, p. 535-552, 2004.
- THOMPSON, M. Recent trends in inter-laboratory precision at ppb and sub-ppb concentrations in relation to fitness for purpose criteria in proficiency testing. *Analyst*, v. 125, p. 385-386, 2000.
- THOMPSON, M.; ELLISON, S.L.R.; FAJGELJ, A.; WILLETTS, P.; WOOD, R. Harmonized guidelines for the use of recovery information in analytical measurement. *Pure Appl. Chem.*, v. 71, p. 337-348, 1999.
- THOMPSON, M.; ELLISON, S.L.R.; WOOD, R. Harmonized guidelines for single-laboratory validation of methods of analysis. *Pure Appl. Chem.*, v. 74, p. 835-855, 2002.
- THOMPSON, M.; ELLISON, S.L.R.; WOOD, R. The international harmonized protocol for the proficiency testing of analytical chemistry laboratories. *Pure Appl. Chem.*, v. 78, p. 145-196, 2006.

- THOMPSON, M.; FEARN, T. What exactly is fitness for purpose in analytical measurement? *Analyst*, v. 121, p. 275-278, 1996.
- THOMPSON, M.; LOWTHIAN, P.J. A Horwitz-like function describes precision in a proficiency test. *Analyst*, v. 120, p. 271-272, 1995.
- THOMPSON, M.; WOOD, R. Harmonized guidelines for internal quality control in analytical chemistry laboratories. *Pure Appl. Chem.*, v. 67, p. 649-666, 1995.
- TOTH, B.; SHIMIZU, H. 1-carbamyl-2-phenylhydrazine tumorigenesis in Swiss mice. Morphology of lung adenomas. *J. Natl. Cancer Inst.*, v. 52, p. 241-251, 1974.
- TOTH, B.; SIMIZU, H.; ERICKSON, J. Carbamylhydrazine hydrochloride as a lung and blood vessel tumor inducer in Swiss mice. *Europ. J. Cancer*, v. 11, p. 17-22, 1975.
- TSALEV, D.L. Vapor generation or electrothermal atomic absorption spectrometry? - Both! *Spectrochim. Acta Part B*, v. 55, p. 917-933, 2000.
- TUKEY, J.W. One-degree of freedom for non-additivity. *Biometrics*, v. 5, p. 232-242, 1949.
- TUKEY, J. W. *Exploratory data analysis*. Reading: Addison-Wesley, 1977. 499 p.
- TURNER, M.J.; SCHAEFFER, J.M. Mode of action of ivermectin. In: CAMPBELL, W.C. (Ed.) *Ivermectin and abamectin*. New York: Springer-Verlag, 1989. p. 73-88.
- USDA (United States Department of Agriculture - Food Safety and Inspection Service - Office of Public Health and Science). Determination of arsenic by atomic absorption spectrophotometry. SOP CLG-ARS.03, 3, Washington DC: USDA/FSIS, 2001. 16 p.
- VAHTER, M.; MARAFANTE, E. In vivo methylation and detoxication of arsenic. In: CRAIG, P.J.; GLOCKLING, F. (Ed.) *The biological alkylation of heavy elements*. London: Royal Society of Chemistry, 1988.

- VALENZUELA, A.I.; POPA, D.S.; REDONDO, M.J.; MAÑES J. Comparison of various liquid chromatographic methods for the analysis of avermectins residues in citrus fruits. *J. Chromatogr. A*, v. 918, p. 59-65, 2001.
- VAN DER VOET, H.; VAN RHIJN, J.A.; VAN DE WIEL, H.J. Inter-laboratory, time, and fitness-for-purpose aspects of effective validation. *Anal. Chim. Acta*, v. 391, p. 159-171, 1999.
- VAN LOCO, J.; BEERNAERT, H. An alternative method validation strategy for the European Decision 2002/657/EC. In: PROCEEDINGS OF EUROPEAN FOOD CHEMISTRY, 12, 2003, Brugges, Belgium. *Proceedings...* Brugges: European Food Chemistry, 2003. p. 91-94.
- VAN OVERMEIRE, I.; VAN LOCO, J.; ROOS, P.; CARBONNELLE, S.; GOEYENS, L. Interpretation of CALUX results in view of the EU maximal TEQ level in milk. *Talanta*, v. 63, p. 1241-1247, 2004.
- VAN RHIJN, J.A.; KENNEDY, D.G.; MULDER, P.P.J.; ELBERS, I.J.W.; ZUIDEMA, T.; YOUNG, P.B. Is semicarbazide an unambiguous marker residue for nitrofurazone? In: PROCEEDINGS OF EUROPEAN FOOD CHEMISTRY, 12, 2003, Brugges, Belgium. *Proceedings...* Brugges: European Food Chemistry, 2003.
- VESSMAN, J.; STEFAN, R.I.; VAN STADEN, J.E.; DANZER, K.; LINDNER, W.; BURNS, D.T.; FAJGELJ, A.; MULLER, H. Selectivity in analytical chemistry. *Pure Appl. Chem.*, v. 73, p. 1381-1386, 2001.
- WAIBLINGER, H.U.; WEBER, W.; KLEINERT, T. Differenzierung von Wildfleischproben mittels DNA- und Protein-Analytik. *Lebensmittelchemie*, v. 52, p. 97-99, 1998.
- WEBB, J.L. *Enzyme and metabolic inhibitors*. New York: Academic Press, 1966. p. 595-793.
- WEISBERG, S. *Applied linear regression*. New York: Wiley, 1985. 324 p.

- WELZ, B.; STAUSS, P. Interferences from hydride-forming elements in hydride generation atomic absorption spectrometry with a heated quartz tube atomizer. *Spectrochim. Acta Part B*, v. 48, p. 951-976, 1993.
- WENTWORTH, W.E. Rigorous least squares adjustment. *J. Chem. Educ.*, v. 42, p. 96-103, 1965.
- WHO (World Health Organization). *Report of a WHO consultation on public health issues related to human and animal transmissible spongiform encephalopathies. With the participation of FAO and OIE*. Geneva: WHO, 2-3 april, 1996.
- WILL, R.G.; IRONSIDE, J.W.; ZEIDLER, M.; COUSENS, S.N.; ESTIBEIRO, K.; ALPEROVITCH, A.; POSER, S.; POCCHIARI, M.; HOFMAN, A.; SMITH, P.G. A new variant of Creutzfeldt-Jacob disease in the UK. *Lancet*, v. 347, p. 921-995, 1996. (a)
- WILL, R.G.; ZEIDLER, M.; BROWN, P.; HARRINGTON, M.; LEE, K.H.; KENNEY, K.L. Cerebrospinal-fluid test for new variant Creutzfeldt-Jacob disease. *Lancet*, v. 348, p. 955, 1996. (b)
- WINTERO, A.K.; THOMSEN, P.D.; DAVIES, W. A comparison of DNA-hybridization, immunodiffusion, countercurrent immunoelectrophoresis and isoelectric focussing for detecting the admixture of pork and beef. *Meat Sci.*, v. 27, p. 75-85, 1990.
- WOLF, C.; LÜTHY, J. Quantitative competitive (QC) PCR for quantification of porcine DNA. *Meat Sci.*, v. 57, p. 161-168, 2001.
- WOOD, R. How to validate analytical methods. *Trends Anal. Chem.*, v. 18, p. 624-632, 1999.
- WORKMAN, J.; MARK, H. Comparison of goodness of fit statistics for linear regression, Part I. *Spectroscopy*, v. 19, p. 38-41, 2004.
- WU, Z.; LI, J.; ZHU, L.; LUO, H.; XU, X. Multi-residue analysis of avermectins in swine liver by immunoaffinity extraction and liquid chromatography-mass spectrometry. *J. Chromatogr. B*, v. 755, p. 361-366, 2001.



YODEN, W.J.; STEINER, E.H. *Statistical manual of the Association of Official Analytical Chemists. Statistical techniques for collaborative tests*. Arlington: AOAC, 1975. 88 p.