

CAPÍTULO 3

MODELO DE CALIBRAÇÃO

REGRESSÃO LINEAR VS. REGRESSÃO LINEAR PONDERADA

3.1. FUNDAMENTOS

O êxito de um estudo farmacocinético depende em grande medida da qualidade do método analítico utilizado. Por isso, a validação do método analítico constitui um pré-requisito essencial, capaz de garantir a exactidão e a reprodutibilidade dos dados obtidos. Por conseguinte, qualquer que seja a metodologia analítica, a correcta interpretação da *curva de calibração* afigura-se fundamental, uma vez que a qualidade dos dados obtidos se encontra fortemente dependente da curva de calibração que os gera.

Normalmente, as concentrações de analito são avaliadas utilizando os resultados da *curva de calibração*, definida como a “relação entre a resposta do instrumento e concentrações conhecidas de analito” (Guidance for Industry, 2001). Muito embora, alguns procedimentos analíticos possam necessitar de um modelo de calibração não-linear, a regressão linear é claramente a abordagem mais adoptada, tomando muitas vezes a designação de *recta de calibração*. Neste caso, embora seja ainda comum a utilização de duas ou mais rectas de calibração para estudar diferentes intervalos de concentração, a utilização de um único intervalo de concentrações, amplo e dinâmico, é de grande utilidade na rotina laboratorial. Geralmente, a utilização do detector de UV ou de fluorescência permite um incremento linear ao longo de um amplo intervalo de concentrações.

No entanto, em grande parte dos casos, pelo facto de estarmos perante intervalos de concentração amplos – normalmente, mais de uma ordem de grandeza – é de esperar que a variância em cada ponto da recta seja diferente (Johnson *et al.*, 1988), pelo que, a condição de homogeneidade de variância, designada por *homoscedasticidade*, não é cumprida. Consequentemente, as concentrações mais elevadas tendem a influenciar (ou “pesar”) a recta

de calibração, resultando na perda da exactidão, especialmente nos seus pontos mais baixos (Johnson *et al.*, 1988; Karnes *et al.*, 1991).

A utilização do modelo de Regressão Linear Ponderada (RLP) é a forma mais simples e efectiva de harmonizar as discrepâncias entre as variâncias dos diferentes pontos que constituem a recta (Garden *et al.*, 1980; Johnson *et al.*, 1988; Thompson, 1988; Karnes *et al.*, 1991; Miller, 1991; Mulholland e Hibbert, 1997).

Assim, mais do que entrar em considerações estatísticas complexas por mero rigor académico, pensámos ser de extrema importância a sistematização de conceitos cuja aplicabilidade na rotina analítica é notória. Para tal, foi objectivo deste capítulo o estudo do modelo de regressão da recta de calibração que melhor ajustaria os dados obtidos durante a validação da técnica cromatográfica (HPLC) desenvolvida para a quantificação de lamotrigina em plasma humano (Capítulo 2).

3.1.1. Fundamentos da regressão linear simples vs. ponderada

O objectivo da regressão linear é encontrar um modelo determinístico capaz de estimar os valores assumidos pela variável dependente (y) quando as variáveis independentes (x_n) são conhecidas ou fixas. Este modelo define o tipo de relação existente entre as variáveis, sendo o *Método dos Mínimos Quadrados* a metodologia mais utilizada para determinar os seus parâmetros. A expressão a minimizar é a seguinte:

$$SS = \sum (y_{obs,i} - y_{prev,i})^2 \quad \text{(Equação 3.1)}$$

No modelo de regressão linear a relação entre as variáveis é descrita por uma linha recta, matematicamente expressa por:

$$y = a + bx \quad \text{(Equação 3.2)}$$

sendo y , a variável dependente; x , a variável independente; a , a ordenada na origem; e b , o declive da recta.

Geralmente, o modelo de regressão linear assume que (1) os valores da variável independente x são fixos, (2) os valores da variável x são medidos sem erro associado, (3) para cada valor de x existe uma subpopulação de valores de y com distribuição normal, (4) os valores de y são estatisticamente independentes, (5) as médias dos valores das subpopulações de y caem sobre uma linha recta, e (6) as variâncias das subpopulações de y são iguais (Daniel, 1987).

Porém, é muito comum encontrar situações em que o desvio-padrão das subpopulações de y se altera com o valor de x , o que se descreve como uma situação de *heteroscedasticidade*. Em muitos casos, o desvio-padrão aumenta proporcionalmente com a concentração, o que conduz a um coeficiente de variação constante. Não obstante poderem ser observados comportamentos diferentes, tendo em consideração que estes podem ser função do sinal ou da concentração ou ainda de outros factores, a ocorrência mais comum de heteroscedasticidade é um aumento de variância em função da concentração (Garden *et al.*, 1980; Johnson *et al.*, 1988; Thompson, 1988; Karnes *et al.*, 1991; Miller, 1991; Pateman, 1995).

Geralmente, um método analítico utiliza um número de padrões de matriz idêntica na qual o analito se encontra em concentrações conhecidas e diferentes entre si. Para cada concentração o aparelho fornece uma resposta que está sujeita a um erro, pelo que medições repetidas originarão uma distribuição de frequências. Esta noção encontra-se ilustrada na Figura 3.1, onde é possível observar as distribuições das medidas para cada concentração. Na Figura 3.1a é assumido que a dispersão das medidas é igual para cada padrão, ou seja, é independente da concentração de analito, e pode ser caracterizada por uma única medida de variância, o que se designa por *homoscedasticidade*. Na Figura 3.1b o desvio-padrão da medida aumenta progressivamente com a concentração de analito, o que se denomina por *heteroscedasticidade*.

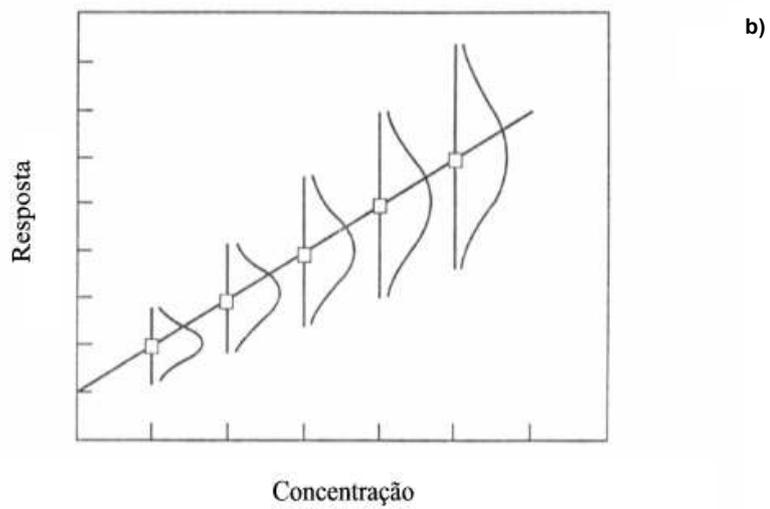
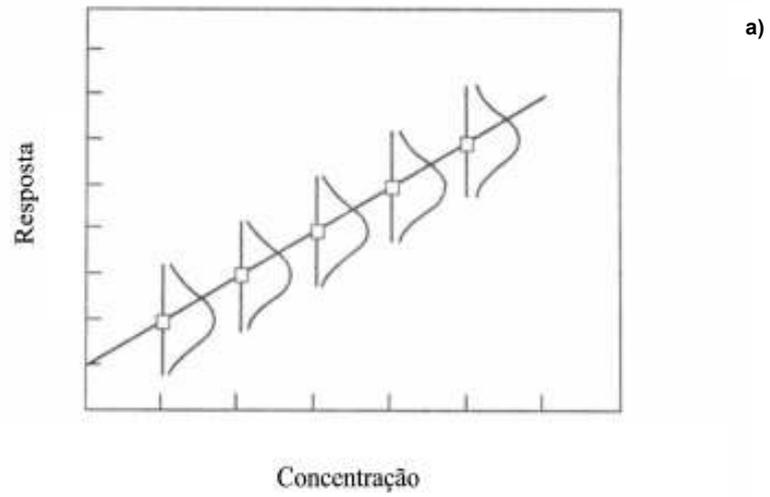


Figura 3.1. Modelos de regressão linear: homoscedasticidade^{a)} - igual desvio-padrão para todas as sub-populações de y ; heteroscedasticidade^{b)}; aumento do desvio-padrão em função da concentração.

Assim, quando as concentrações mais elevadas tendem a influenciar (“pesar”) mais a recta de regressão do que os pequenos desvios associados às concentrações mais baixas, resultando na diminuição da exactidão (Johnson *et al.*, 1988; Karnes *et al.*, 1991), deve-se recorrer ao modelo de regressão linear ponderada (Garden *et al.*, 1980; Johnson *et al.*, 1988; Miller e Miller, 1988; Thompson, 1988; Karnes *et al.*, 1991; Miller, 1991).

A expressão a ser minimizada toma a seguinte forma (Johnson *et al.*, 1988):

$$SS = \sum \frac{(y_{obs,i} - y_{prev,i})^2}{\sigma_i^2} \quad \text{(Equação 3.3)}$$

em que σ_i^2 representa a variância associada à resposta dos padrões de calibração.

Tomando o objectivo da RLP em conta, o factor de ponderação mais apropriado, W_i , corresponderá ao inverso da variância:

$$W_i = \frac{1}{\sigma_i^2} \quad \text{(Equação 3.4)}$$

No entanto, este factor de ponderação é considerado pouco prático, tendo em atenção que seriam necessárias várias determinações de cada padrão de calibração para se calcular o valor de variância, pelo que, outras aproximações de variância têm sido apontadas, baseadas quer na variável independente x (concentração), quer na variável dependente y (resposta) (Johnson *et al.*, 1988; Karnes *et al.*, 1991).

3.1.2. Teste de homoscedasticidade

O modelo de regressão linear é claramente o modelo mais utilizado nos métodos analíticos, em que a variável independente é a concentração e a variável dependente a resposta. Tal como já foi referido, este tipo de regressão assume que a medida de x não está associada a

erros e que os resíduos calculados são independentes, normalmente distribuídos e possuem igual variância, condição que se designa por homoscedasticidade. Contudo, na maioria das análises bioanalíticas, a variância não é constante em todos os pontos da recta de calibração e, por conseguinte, a condição de homoscedasticidade não é cumprida.

Assim sendo, é importante verificar a assunção de homoscedasticidade em qualquer análise de regressão linear. Esta avaliação é feita por análise gráfica da distribuição dos resíduos em função da concentração (Garden *et al.*, 1980; Johnson *et al.*, 1988; Karnes *et al.*, 1991; Szabo *et al.*, 1994; Mulholland e Hibbert, 1997) e através da aplicação do *F*-teste (Cardone *et al.*, 1990; 1990; Hartmann *et al.*, 1995).

3.1.2.1. Análise de resíduos

A análise gráfica dos resíduos *versus* concentração permite-nos observar uma banda horizontal de resíduos aleatoriamente distribuídos em redor do eixo dos xx, em caso de homoscedasticidade (Figura 3.2a); por outro lado, um aumento gradual da variância em função da concentração repercutir-se-á na distribuição dos resíduos que tomam uma forma característica de leque ou cone (Figura 3.2b), a ocorrência mais comum de heteroscedasticidade.

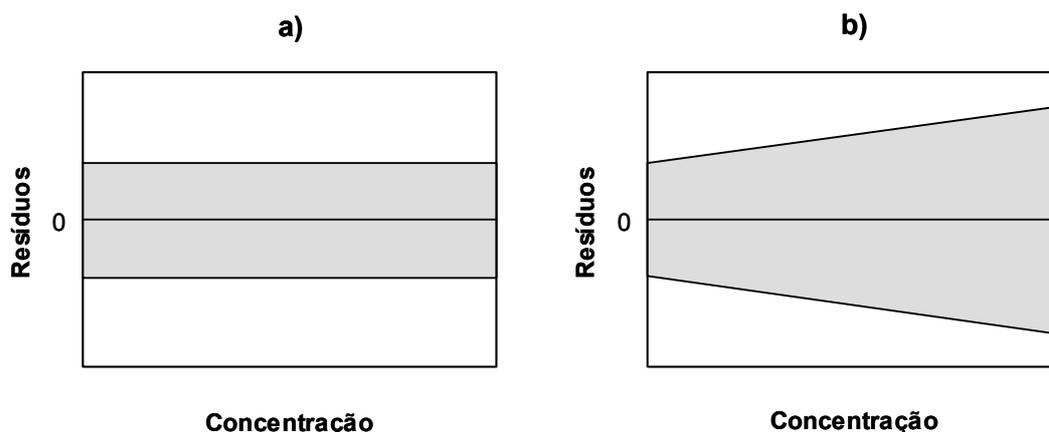


Figura 3.2. Esquemas representativos da distribuição dos resíduos vs. concentração (zona sombreada) observada em caso de variância constante^{a)} e quando se verifica o aumento da variância em função da concentração^{b)}.

3.1.2.2. Análise estatística

A variância das concentrações é analisada nos limites da curva de calibração, tendo em consideração que:

$$F_{\text{exp}} = \frac{S_1^2}{S_2^2} \quad (\text{Equação 3.5})$$

$$F_{\text{tab}}(f_1, f_2; 0,99)$$

sendo que S_1 representa a variância da resposta do padrão de calibração mais elevado e S_2 a variância do padrão de calibração mais baixo.

Os valores experimentais F_{exp} são expressos pela razão entre as variâncias obtidas para os padrões de concentração nos limites da curva de calibração, e os valores tabelados F_{tab} para um nível de confiança a 99% e $f_1 = f_2 = (n - 1)$ graus de liberdade (ISO, 1990).

Por sua vez, a variância da amostra é dada pela seguinte equação:

$$S_i^2 = \frac{\sum (y_i - \bar{y}_i)^2}{n_i - 1} \quad (\text{Equação 3.6})$$

Mantendo-se a variância constante ao longo da recta de calibração, os resíduos devem distribuir-se aleatoriamente em redor do eixo dos xx, e o valor do F_{exp} deve ser inferior ao valor tabelado.

3.1.3. Escolha do modelo de regressão

Perante a evidência de que os dados são heteroscedásticos, a fase seguinte passa pela escolha de um modelo de regressão mais adequado à sua natureza. Como já se referiu, nesta situação, o factor de ponderação mais adequado é o inverso da variância, porém, esta

abordagem não se coaduna com a rotina laboratorial, pois requer várias determinações por ponto, o que torna o processo impraticável, dada a necessidade efectiva de preparar rectas de calibração diariamente ou sempre que o método é utilizado. Por conseguinte, o recurso a outras aproximações deve ser equacionado, nomeadamente, a utilização de factores de ponderação empíricos, denominados como “aproximação de variância”, tais como, $1/x^{1/2}$, $1/x$, $1/x^2$, $1/y^{1/2}$, $1/y$ e $1/y^2$.

O factor de ponderação mais adequado é escolhido com base no cálculo do Erro Relativo percentual (%ER), o qual compara a concentração obtida através da equação resultante da regressão (C_{exp}), com a concentração teórica ou nominal (C_{nom}):

$$\%ER = \frac{C_{exp} - C_{nom}}{C_{nom}} \cdot 100 \quad \text{(Equação 3.7)}$$

O erro relativo percentual, avaliado quer pela análise gráfica dos seus valores em função da concentração, quer pela soma de todos os valores para o total da gama de concentrações, é um indicador sensível da qualidade do ajuste na avaliação da regressão linear ponderada (Johnson *et al.*, 1988).

O melhor factor de ponderação será aquele que for capaz de originar valores de %ER aleatoriamente distribuídos em redor do eixo do x , em forma de uma banda horizontal estreita, e que simultaneamente, apresente a menor soma de valores de %ER ao longo do intervalo de concentrações.

3.1.4. Equação de regressão linear ponderada

Os parâmetros do modelo (a e b) da equação de regressão linear ponderada podem agora ser determinados. A conversão das equações da regressão linear simples nas suas contrapartes ponderadas é facilmente realizada por multiplicação de um factor por cada somatório da equação, e alterando n por $\sum w_i$ (Levie, 1986; Johnson *et al.*, 1988).

Assim, sabendo que os valores dos parâmetros a e b para a regressão linear simples (RLS) podem ser obtidos pelas seguintes fórmulas:

$$b = \frac{n \cdot \sum x_i y_i - \sum x_i \cdot \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2} \quad (\text{Equação 3.8})$$

$$a = \frac{\sum x_i^2 \cdot \sum y_i - \sum x_i \cdot \sum x_i y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2} \quad (\text{Equação 3.9})$$

os valores dos parâmetros a e b para a RLP são obtidos pelas seguintes fórmulas modificadas:

$$b = \frac{\sum w_i \cdot \sum w_i x_i y_i - \sum w_i x_i \cdot \sum w_i y_i}{\sum w_i \cdot \sum w_i x_i^2 - (\sum w_i x_i)^2} \quad (\text{Equação 3.10})$$

$$a = \frac{\sum w_i x_i^2 \cdot \sum w_i y_i - \sum w_i x_i \cdot \sum w_i x_i y_i}{\sum w_i \cdot \sum w_i x_i^2 - (\sum w_i x_i)^2} \quad (\text{Equação 3.11})$$

onde (x_i, y_i) representam o par de dados i do n total de pares de dados e w_i representa o factor de ponderação.

O grau de dependência estabelecido entre as duas variáveis, expresso pelo coeficiente de correlação (r), é obtido pela fórmula seguinte:

$$r = \frac{\sum w_i \cdot \sum w_i x_i y_i - \sum w_i x_i \cdot \sum w_i y_i}{\sqrt{\sum w_i \cdot \sum w_i x_i^2 - (\sum w_i x_i)^2} \cdot \sqrt{\sum w_i \cdot \sum w_i y_i^2 - (\sum w_i y_i)^2}} \quad (\text{Equação 3.12})$$

As fórmulas acima apresentadas para a RLP podem ser inseridas numa folha de cálculo (Microsoft Excel®) de modo a proceder à determinação dos parâmetros de regressão das rectas de calibração, utilizando, para tal, o factor de ponderação desejado. Este tipo de estudo pode também ser realizado através de um programa informático que inclua nas suas funções a

regressão linear ponderada, com a possibilidade de seleccionar o factor de ponderação pretendido na regressão. O programa estatístico SPSS® (release 11.5 for Windows; 2002) possui essas potencialidades e permite ainda obter dados relativos à qualidade do ajuste dessa recta.

Todo o processo de análise de dados, desde a observação dos resíduos vs. concentração até ao cálculo da equação que melhor exprime a natureza dos dados pode ser esquematicamente traduzido pela Figura 3.3.

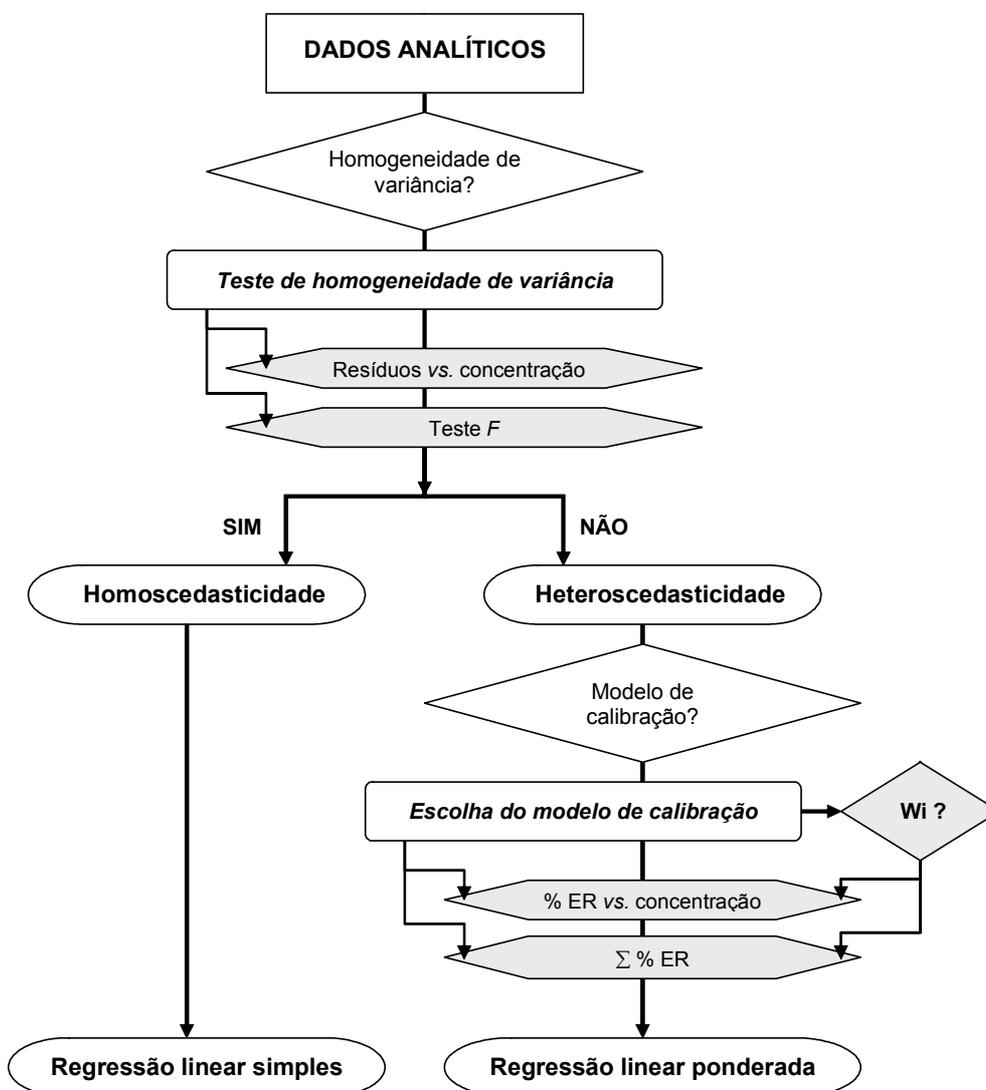


Figura 3.3. Representação esquemática da escolha do modelo de regressão (ER=Erro relativo; Wi=Ponderação).

3.2. RESULTADOS

Para exemplificar este processo de escolha do modelo de regressão utilizámos o conjunto de dados gerados no ensaio intra-dia (Tabela 2.1) durante a validação da técnica de determinação de lamotrigina em plasma humano (Capítulo 2).

3.2.1. Teste de homoscedasticidade

A distribuição dos resíduos obtidos para o conjunto de dados do ensaio intra-dia da validação da técnica de lamotrigina pode ser analisada através da Figura 3.4. Esta figura demonstra uma aumento gradual da magnitude do erro em função da concentração, não sendo possível observar uma banda horizontal de erro aleatoriamente distribuído em redor do eixo das abcissas, indicativa de homoscedasticidade.

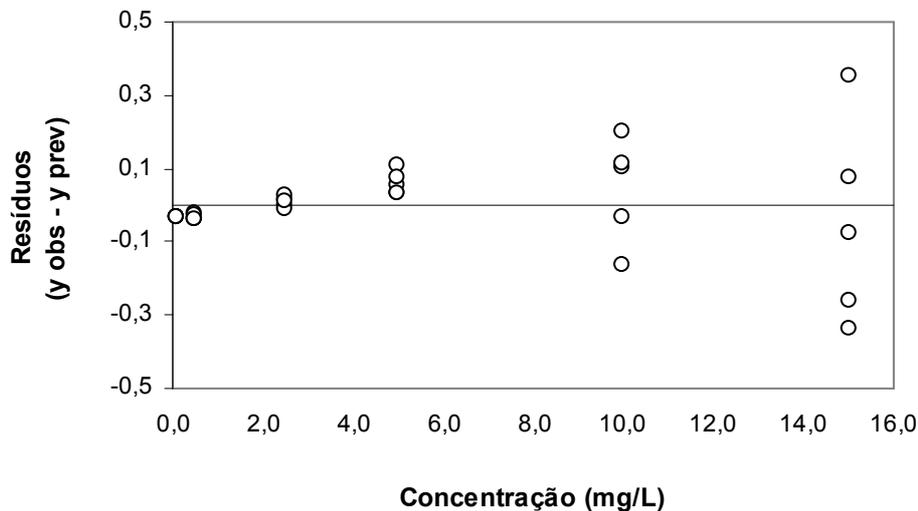


Figura 3.4. Distribuição dos resíduos vs. concentração obtida no ensaio intra-dia da validação da técnica analítica de determinação de lamotrigina em plasma humano (prev=Previsto; obs=Observado).

O teste F (Tabela 3.1) relevou uma diferença significativa de variâncias entre os dois extremos da recta de calibração quando o valor experimental ($F_{\text{exp}} = 1,45 \times 10^5$) foi comparado com o valor tabelado ($F_{\text{tab}} = 15,98$).

Os resultados indicam que a assunção de homoscedasticidade não foi cumprida; os dados são heteroscedásticos.

Tabela 3.1. Teste de homogeneidade de variâncias (s^2): Teste F .

Concentração (mg/L)	A	A _{PI}	Resposta (A/ A _{PI})	s ²
0,1	75381	3373583	0,022	5,25x10 ⁻⁷
	62834	2744443	0,023	
	68050	2863618	0,024	
	73212	3190782	0,023	
	75930	3144582	0,024	
15,0	12209191	3244845	3,763	7,60x10 ⁻²
	12953689	2911392	4,449	
	11682576	2903904	4,023	
	10940196	2852877	3,835	
	11147470	2672128	4,172	

$$F_{\text{exp}} = \frac{s^2_{15,0}}{s^2_{0,1}} = \frac{7,60 \times 10^{-2}}{5,25 \times 10^{-7}} = 1,45 \times 10^5$$

$$F_{\text{tab}} (4; 4; 0,99) = 15,98$$

A=Área do pico de lamotrigina; A_{PI}=Área do pico do padrão interno.

3.2.2. Escolha do modelo de regressão

A distribuição do erro relativo percentual (%ER) em função da concentração encontra-se representada na Figura 3.5, para o modelo de regressão linear simples (Modelo 1) e para os modelos de regressão ponderada (Modelos 2-7). O Modelo 1 originou uma subestimação das concentrações próximas do extremo inferior da curva de calibração, principalmente no limite de quantificação (LQ). Os Modelo 4 e 7 apresentaram os melhores resultados nesses pontos da recta.

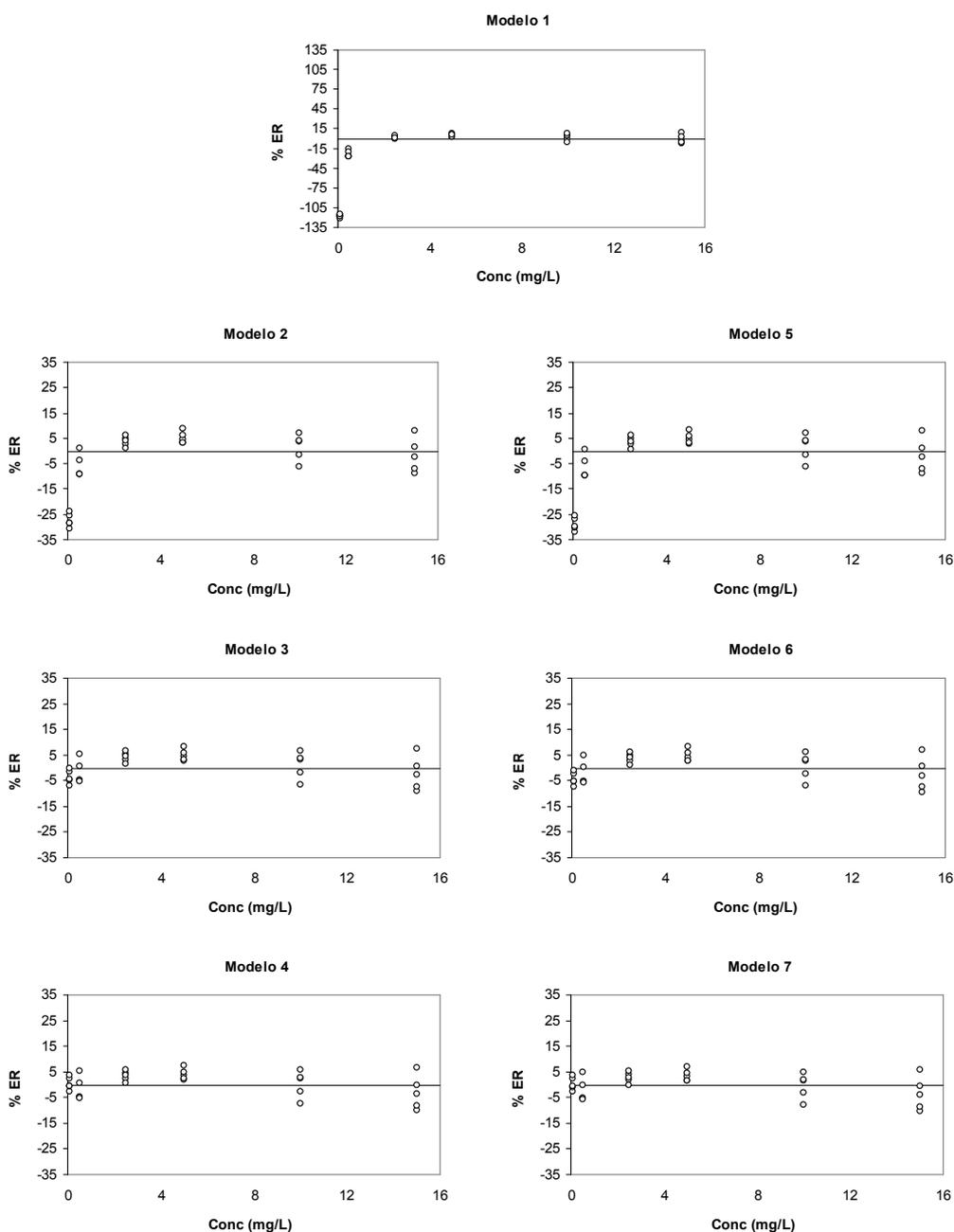


Figura 3.5. Distribuição do erro relativo percentual (%ER) em função da concentração de lamotrigina (Conc.) obtida para o Modelo 1 ($w_i=1$), Modelo 2 ($w_i=1/y^{1/2}$), Modelo 3 ($w_i=1/y$), Modelo 4 ($w_i=1/y^2$), Modelo 5 ($w_i=1/x^{1/2}$), Modelo 6 ($w_i=1/x$) e Modelo 7 ($w_i=1/x^2$).

Os parâmetros de regressão gerados a partir de cada factor de ponderação e o respectivo somatório dos erros relativos percentuais encontram-se resumidos na Tabela 3.2. Estes parâmetros de regressão foram obtidos com base nas equações 3.10, 3.11 e 3.12, introduzidas numa folha de cálculo (Microsoft Excel®, versão 2002). A análise estatística destes resultados foi levada a cabo pelo SPSS® (release 11.5 for Windows; 2002).

Tabela 3.2. Parâmetros de regressão da recta de calibração gerada para cada factor de ponderação (w_i) e respectiva soma dos erros relativos percentuais ($\Sigma \%ER$); $n = 30$.

Modelo	w_i	b	a	r	$\Sigma \%ER$
1	1	0,2713	+0,0281	0,997	788,90
2	$1/y^{1/2}$	0,2743	+0,0033	0,998	264,31
3	$1/y$	0,2762	-0,0034	0,998	133,48
4	$1/y^2$	0,2789	-0,0048	0,998	116,47
5	$1/x^{1/2}$	0,2747	+0,0037	0,998	271,34
6	$1/x$	0,2770	-0,0033	0,998	133,71
7	$1/x^2$	0,2804	-0,0049	0,999	112,38

O modelo 7 aparece associado ao somatório mais baixo, pelo que a aproximação de variância $1/x^2$ apresenta-se como a mais válida de todas as estudadas. Estes resultados estão de acordo com a distribuição gráfica assumida pelos erros (Figura 3.5). Por conseguinte, o Modelo 7 foi escolhido como o modelo de regressão mais adequado ao conjunto de dados estudados.

3.2.3. Equação de regressão linear ponderada

A recta de calibração obtida através do Modelo 7, modelo de regressão ponderada, que utiliza o factor de ponderação $1/x^2$, é igual a $y = 0,2804x - 0,0049$, com $r = 0,999$.

A exactidão dos dados, expressa pelo valor de *Bias ou Viés*, avaliada ao longo do intervalo de concentrações estudadas, utilizando os Modelo 1 e 7, encontra-se na Tabela 3.3.

Tabela 3.3. Comparação da exactidão do modelo de regressão linear simples ($w_i=1$) e do modelo de regressão ponderado ($w_i=1/x^2$) durante o ensaio de validação da lamotrigina^a.

Concentração nominal (mg/L)	% BIAS	
	Modelo 1 $w_i = 1$ [$y = 0,2713x - 0,0281$]	Modelo 7 $w_i = 1/x^2$ [$y = 0,2804x - 0,0049$]
0,1	-117,99	+0,28
0,5	-23,38	-2,33
2,5	+1,31	+2,73
5,0	+4,50	+3,46
10,0	+1,64	-0,48
15,0	-1,21	-3,63

^a Os valores representam a média de cinco réplicas.

3.3. DISCUSSÃO

A *curva de calibração* e a natureza da função matemática utilizada no modelo de regressão têm merecido, nos últimos anos, a atenção de alguns grupos de trabalho (Baumann e Watzig, 1995; Mulholland e Hibbert, 1997; Nagaraja *et al.*, 1999; Kirkup e Mulholland, 2004). O reconhecimento das limitações dos modelos de regressão linear e as consequências da sua aplicação a dados provenientes de métodos analíticos, frequentemente caracterizados por heteroscedasticidade, chamaram a atenção para a necessidade de se escolher o modelo de regressão mais adequado durante a validação dos procedimentos analíticos (Baumann e Watzig, 1995; Mulholland e Hibbert, 1997; Nagaraja *et al.*, 1999; Kirkup e Mulholland, 2004). Apesar de ser ainda comum a utilização de duas ou mais rectas de calibração para abranger diferentes intervalos de concentração (Albertioni *et al.*, 1995; Mccrudden e Tett, 1999), não obstante a maior dificuldade encontrada na sua aplicação na rotina laboratorial, a utilização de modelos de regressão linear ponderada tem sido apontada por vários autores como forma de harmonizar as diferenças de variância (Baumann e Watzig, 1995; Nagaraja *et al.*, 1999; Sadray *et al.*, 2003), assistindo-se nos últimos anos a uma crescente aplicação desta abordagem, não só nos métodos cromatográficos (Sparreboom *et al.*, 1999; Sadray *et al.*, 2003; Santoyo e Verma, 2003; Zhou *et al.*, 2003), mas também noutros campos de análise, por exemplo, a *electroforese capilar* (Baumann e Watzig, 1995), ou a *reação em cadeia da polimerase em tempo real*, mais conhecida por “*real-time PCR*” (Burns *et al.*, 2004).

Contrariamente ao que acontece nas análises de produtos farmacêuticos, em que a concentração de analito se encontra definida entre limites conhecidos, os métodos de determinação de substâncias em matrizes biológicas, caracterizam-se por utilizarem, geralmente, intervalos de concentração amplos, pois só assim é possível quantificar substâncias sujeitas a processos dinâmicos como a absorção, distribuição, metabolismo e excreção (Buick *et al.*, 1990). Quando o intervalo de concentrações utilizado para uma dada recta de calibração é extenso, é provável que as variâncias encontradas para cada ponto da recta não sejam iguais, pelo que a assunção de homoscedasticidade, inerente ao método de regressão linear simples, não é cumprida.

Apesar de alguns autores preconizarem a escolha do modelo de regressão durante a pré-validação (Lang e Bolton, 1991; Wieling *et al.*, 1996), a utilização das amostras dos ensaios da validação propriamente dita, com as várias curvas de calibração em vários lotes, permite simular mais adequadamente as condições da rotina (Nagaraja *et al.*, 1999), pelo que o presente estudo do modelo de regressão foi realizado com o conjunto de dados obtidos durante a validação da

técnica analítica (Capítulo 2). Neste estudo, as concentrações utilizadas para a recta de calibração encontram-se compreendidas entre 0,1 e 15,0 mg/L, tendo o teste de homoscedasticidade sido levado a cabo. Se o modelo de regressão linear simples ajustasse bem os dados, os resíduos distribuir-se-iam aleatoriamente numa banda horizontal em redor do eixo das abcissas (Karnes *et al.*, 1991). No entanto, os dados obtidos durante o ensaio intra-dia demonstraram claramente um aumento gradual do valor dos resíduos em função da concentração. O teste *F* veio demonstrar a existência de diferentes variâncias nos extremos da recta, confirmando, por isso, a existência de heteroscedasticidade.

A escolha do modelo de regressão foi realizada de acordo com a análise gráfica da distribuição dos ER percentuais e com o seu somatório. O melhor factor de ponderação é aquele que apresenta a menor soma de erros (Johnson *et al.*, 1988). O modelo 7, correspondente ao factor de ponderação $1/x^2$, apresentou os valores mais baixos de somatório de erros, ao mesmo tempo que exibiu os melhores resultados de distribuição gráfica de erros em redor do eixo dos xx na zona do limite de quantificação (Tabela 3.2 e Figura 3.5). Para além disso, à semelhança do que acontece na maioria dos métodos analíticos (Pateman, 1995), o desvio-padrão da resposta foi proporcional à concentração (Figura 3.4). De acordo com estas circunstâncias o factor de ponderação seleccionado foi o $1/x^2$.

Ficou ainda demonstrado que uma larga percentagem de erros no extremo inferior da recta de calibração pode coexistir com valores de coeficiente de correlação (*r*) aceitáveis (Tabela 3.2). Este aspecto é particularmente importante, se pensarmos que alguns analistas continuam a utilizar o coeficiente de correlação (de 0,999 a 1,000) como único critério de aceitação na avaliação do *Método dos Mínimos Quadrados*, preocupação que foi também expressa por Kirkup e Mulholland (2004). De facto, o valor de *r* é apenas um indicador de associação ou correlação entre duas variáveis, e não um indicador de linearidade (Analytical Methods Committee, 1988).

Este estudo revelou um aumento da exactidão para o modelo de regressão baseado no factor de ponderação $1/x^2$ (Modelo 7), aumento esse particularmente evidente no extremo inferior da recta, tal como era esperado, sem que por isso fosse alterada a exactidão existente a nível das porções média e superior da recta de calibração. O valor de *Bias* encontrado para o Modelo 1 foi consideravelmente superior aos limites aceitáveis de 20% (Shah *et al.*, 1992) para os valores mais baixos da recta, especialmente no limite de quantificação (Tabela 3.3). Contrariamente, a regressão linear ponderada, expressa no Modelo 7, foi capaz de apresentar bons resultados de exactidão, originando portanto, uma diminuição do LQ. Esta redução do LQ através da utilização da regressão linear ponderada foi também referida por Szabo *et al.* (1994). Desta forma é possível utilizar um intervalo de concentrações alargado e dinâmico com elevado

grau de exactidão (Johnson *et al.*, 1988; Karnes *et al.*, 1991), originando assim, o aumento da precisão na análise de rotina laboratorial (Garden *et al.*, 1980). Um estudo recente refere que o modelo de regressão linear ponderado permite melhorar a avaliação de parâmetros como a sensibilidade e o limite de detecção (Santoyo e Verma, 2003). A quantificação exacta de baixas concentrações séricas ao longo do tempo assume particular relevância nos estudos farmacocinéticos e farmacodinâmicos, no desenvolvimento e avaliação de medicamentos.

Por fim, uma vez que a regressão linear é o modelo mais utilizado para a obtenção das curvas de calibração, fornecendo um instrumento eficaz de ajuste de dados a um dado modelo, importa referir que é fundamental certificarmo-nos da legitimidade da assunção do modelo matemático utilizado. A regressão linear simples assume que o resíduo surge aleatoriamente e que a magnitude das diferenças entre os pontos de calibração e a linha ajustada são independentes da concentração. Embora a regressão linear ponderada seja naturalmente mais complexa do que a regressão linear simples, esse facto não deve demover os analistas de perseguirem o seu objectivo, ou seja assegurar a fiabilidade dos resultados.

Finalmente, resta salientar que o interesse gerado em torno da *curva de calibração*, e por conseguinte, do *modelo de regressão* deve-se também, em parte, à globalização e formação de grupos económicos, como a União Europeia (UE), que levou as entidades governamentais a instituírem directivas e recomendações, num esforço conjunto de harmonização, sendo notória a preocupação relativamente a este assunto por parte de entidades regulamentares como a EMEA (EMEA, 1995b; EMEA, 1995a; EMEA, 1998) ou a FDA (Guidance for Industry, 2001), ou ainda de outros organismos internacionais como a EURACHEM (1998; 2000). Esta preocupação pela validação dos métodos analíticos não é um problema exclusivo dos estudos farmacocinéticos e estudos de bioequivalência, mas também de outras áreas que dependem da química analítica, algumas das quais com elevado impacto económico, como é o caso do controlo de resíduos de medicamentos em alimentos, dada a importância da determinação dos limites máximos de resíduos (LMRs) e dos intervalos de segurança dos medicamentos veterinários (European Commission, 2003).

Uma das recomendações da *Conferência de Washington* referente ao tema “curva de calibração” foi a de que “a curva deve ser ajustada aos padrões utilizando a função mais simples” (Shah *et al.*, 1992). Alguns autores classificaram esta conclusão de “simplificação perigosa” (Hooper, 1995), não obstante, esse facto levou a que as entidade regulamentares exigissem uma justificação para a utilização de qualquer algoritmo que não seja o modelo linear simples - “a ponderação e a utilização de uma equação de regressão complexa deve ser justificada” (Guidance for Industry, 2001).

O presente trabalho pode contribuir para clarificar as circunstâncias em que é necessário utilizar o modelo de regressão linear ponderada, fornecendo a metodologia para atingir esse fim. Para além disso, a regressão linear ponderada poder-se-á tornar tanto mais importante quanto maior for o grau de exigência por parte de futura legislação da UE (European Commission, 2000; Burns *et al.*, 2004).