

variáveis. A este submodelo foi aplicada a estatística de Wilks, donde concluiu-se pela não rejeição da sua admissibilidade. No modeloICC.29 o número de preditores contínuos é numeroso. Nesta situação, a tarefa de identificar preditores cuja contribuição para o modelo beneficiasse com a utilização de transformações é difícil. Foi efectuada a análise à linearidade dos preditores contínuos, técnica essa utilizada para verificar a existência de relações de linearidade entre a variável resposta e os preditores contínuos do modelo de RL. Caso tal relação não se venha a revelar linear nas variáveis em causa, poder-se-á proceder a várias transformações (Harrell, 2005). A referida análise deu origem ao modelo modeloICC.30. No campo da bioestatística alguns tipos de interações têm-se revelado importantes na previsão de resultados (Harrell, 2005). Tendo em conta as interações entre variáveis foi construído o modeloICC.interacção. A Tabela 3 compara todos os modelos propostos até ao momento para as várias funções de ligação estudadas. Na ligação Aranda Ordaz, considerou-se  $\alpha \in [0; 1.5]$ . O parâmetro  $\alpha$  adoptado foi o que minimizou o desvio residual.

## 2.5 Teste de Hosmer e Lemeshow

Hosmer e Lemeshow (2000) propuseram um teste que tem como objectivo verificar a qualidade de ajustamento no modelo de RL. O presente teste consiste no cálculo da estatística  $\chi^2$  de Pearson. No cálculo desta estatística serão considerados  $g$  grupos. A criação de  $g$  grupos tem como base as probabilidades estimadas pelo modelo de RL para cada uma das  $n$  observações em estudo. Os grupos criados serão dez ( $g = 10$ ) com igual dimensão. A escolha de 10 grupos é justificada pela determinação dos 10 decis associados às probabilidades estimadas. A hipótese nula para o teste de Hosmer e Lemeshow é formulada no seguinte termo:  $H_0$ : O modelo de RL em estudo está correctamente ajustado. A estatística de teste é construída com base numa medida global de ajuste entre a proporção de frequências observadas em cada um dos grupos ( $\bar{y}_g$ ) e a proporção de frequências estimadas pelo modelo em cada um dos grupos ( $\bar{\pi}_g$ ). Essa medida é dada por:

$$C = \sum_{g=1}^{10} n_g \frac{(\bar{y}_g - \bar{\pi}_g)^2}{\bar{\pi}_g(1 - \bar{\pi}_g)}. \quad (5)$$

Se  $H_0$  for verdadeira, devem registar-se pequenas diferenças entre  $\bar{y}_g$  e  $\bar{\pi}_g$  e, consequentemente,  $C$  deve tomar valores baixos. Pelo contrário, um valor de  $C$  elevado constitui um indicio de que há um desajuste entre as proporções referidas. O teste de Hosmer e Lemeshow foi aplicado a todos os MLGs propostos, com função ligação Logit. Constatou-se que não é rejeitada a admissibilidade de um correcto ajuste em todos os modelos. Na Figura 3 é visível que as proporções observadas e estimadas são praticamente coincidentes, situação essa retratada pela sobreposição total ou parcial dos círculos para cada um dos grupos.