

é uma das doenças que afecta o normal funcionamento do coração. Traduz-se pela incapacidade do coração bombear sangue de forma eficiente, impedindo que este chegue a todo o organismo. A ICC é um problema grave de saúde pública que atinge proporções epidémicas. A sua prevalência na Europa e nos EUA é de 2%, enquanto que em Portugal continental esse valor mais que duplica (4.36%) (Ceia *et al.*, 2000). Morrem mais doentes de ICC do que de todos os tumores malignos em conjunto. Em países industrializados, 2% do orçamento da saúde é gasto em pacientes com ICC sendo que 70% desse total em internamentos e reinternamentos (Fonseca, 2007). A existência frequente de outras patologias associadas à ICC, nomeadamente a doença pulmonar, insuficiência renal, diabetes e anemia, dificulta o tratamento farmacológico. A importância que a ICC assume, revelou-se uma motivação para a temática deste estudo bem como o recurso a uma classe de modelos designada por Modelos Lineares Generalizados (MLGs). Esta classe de modelos permite uma maior flexibilidade do que o Modelo Linear Clássico, quando a variável resposta não se limita à distribuição normal. Neste estudo, a variável resposta é de tipo binário e será representada por ICC. O estudo da ICC será efectuado na base de dados *Right Heart Catheterization* criada por Connors *et al.* (1996) e disponível em <http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/DataSets>. Os dados são provenientes dos cinco maiores hospitais universitários norte-americanos no período compreendido entre 1989 e 1994 e contém informação relativa ao primeiro dia de hospitalização de 5220 adultos gravemente doentes que estão a receber tratamento na unidade de cuidados intensivos.

## 2 Metodologia

### 2.1 Abordagem inicial

A idade do paciente é um factor de risco para qualquer doença cardíaca, em particular para a ICC. Fará sentido ter como covariável, a variável idade. A representação gráfica da variável resposta ICC em função da variável idade revela duas particularidades:

1. todos os pontos se dispõem em cima de uma de duas rectas paralelas:  $ICC=0$  (ausência de doença) e  $ICC=1$  (presença da doença). Tal facto prende-se com a própria natureza da variável resposta.
2. a variabilidade associada à variável resposta relativamente à variável idade é grande, o que torna difícil expressar uma relação funcional entre elas.

Por forma a ultrapassar este obstáculo, ou seja redução de alguma variabilidade mantendo a estrutura da relação entre ambas as variáveis, Hosmer e Lemeshow (2000) sugerem a criação de intervalos de classe para a variável independente e a posterior determinação da média dos resultados dentro de cada classe desta variável. A Figura 1 concretiza a sugestão dada por Hosmer e Lemeshow e permite identificar uma relação funcional, semelhante a uma curva sigmóide, entre