

CARACTERIZAÇÃO DAS INTERNAÇÕES POR INFARTO AGUDO DO  
MIOCÁRDIO NO ESTADO DO RIO DE JANEIRO POR MEIO DE ANÁLISE DE  
CORRESPONDÊNCIA MÚLTIPLA E DISTÂNCIA DE TOLERÂNCIA

Felipe Machado Huguenin

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Biomédica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Biomédica.

Orientadores: Antonio Fernando Catelli  
Infantosi.

Renan Moritz Varnier Rodrigues  
de Almeida.

Rio de Janeiro  
Outubro de 2011

CARACTERIZAÇÃO DAS INTERNAÇÕES POR INFARTO AGUDO DO  
MIOCÁRDIO NO ESTADO DO RIO DE JANEIRO POR MEIO DE ANÁLISE DE  
CORRESPONDÊNCIA MÚLTIPLA E DISTÂNCIA DE TOLERÂNCIA

Felipe Machado Huguenin

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO  
LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA  
(COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE  
DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE  
EM CIÊNCIAS EM ENGENHARIA BIOMÉDICA.

Examinada por:

---

Prof. Antonio Fernando Catelli Infantsi, Ph.D.

---

Prof. Roberto Macoto Ichinose, D.Sc.

---

Prof<sup>ª</sup>. Rejane Sobrino Pinheiro, D.Sc

---

Prof<sup>ª</sup>. Maria Stella de Castro Lobo, D.Sc.

RIO DE JANEIRO, RJ - BRASIL

OUTUBRO DE 2011

Huguenin, Felipe Machado

Caracterização das internações por infarto agudo do miocárdio no Estado do Rio de Janeiro por meio de Análise de Correspondência Múltipla e Distância de Tolerância / Felipe Machado Huguenin – Rio de Janeiro: UFRJ/COPPE, 2011.

XII, 73 p.: il.; 29,7 cm.

Orientadores: Antonio Fernando Catelli Infantosi.  
Renan Moritz Varnier Rodrigues de Almeida.

Dissertação (mestrado) – UFRJ/ COPPE/ Programa de Engenharia Biomédica, 2011.

Referências Bibliográficas: p. 65-71.

1. Acessibilidade em saúde. 2. Infarto Agudo do Miocárdio 3. Análise de Correspondência. 4. Distância de Tolerância I. Infantosi, Antônio Fernando Catelli *et al.*. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Biomédica. III. Título.

Este trabalho é dedicado aos meus pais,  
Miguel e Maria Teresa, minhas jóias raras.  
Dedicado, ainda, à minha irmã, Érica, ao  
meu cunhado, Raphael e a minha tão  
amada sobrinha, Maria Clara.

## **Agradecimentos:**

À Deus em sua Trindade Santa, acima de tudo, pelo dom da vida, sempre presente em todos os momentos da minha existência, cujo amor, bondade e compreensão tornaram tudo possível.

Aos meus orientadores, Professor Dr. Renan Moritz Varnier Rodrigues de Almeida e Professor Dr. Antonio Fernando Catelli Infantosi, que me estimularam desde o início quando tudo parecia muito mais difícil, e que se tornaram mais do que simples professores, mas incentivadores na busca por novos conhecimentos. Agradeço pelos ensinamentos e pela experiência passada, que transcenderam os limites acadêmicos, mas formaram para a vida.

Às professoras Rosimary Almeida e Rejane Pinheiro. Esta por disponibilizar o banco de dados fruto de suas pesquisas no Instituto de Estudos em Saúde Coletiva (UFRJ), que serviu de base para este trabalho e pelos conhecimentos passados. Aquela por cumprir com maestria a sua função de mestre, mostrando-me como ser um professor melhor, e acima de tudo por estar sempre com a mão amiga estendida para auxiliar.

A todos os professores do Programa de Engenharia Biomédica da COPPE/UFRJ, pelos conhecimentos repassados e pela dedicação, o que mantém o curso como um dos melhores do Brasil.

Aos integrantes do Laboratório de Engenharia de Sistemas de Saúde (LESS) e dos demais laboratórios do PEB, que, com certeza, oportunizaram a conclusão deste trabalho. Em especial a Aline Nascimento e ao João Carlos Gama, que por meio das conversas, possibilitaram o surgimento e o aprimoramento de idéias, e a Nárrima, Fernanda, Bruno e Francisco, pelas conversas e conselhos, meu sincero muito obrigado.

Aos funcionários administrativos, eficientes nas suas tarefas, mas acima de tudo pessoas amigas com paciência inesgotável para resolverem nossos problemas.

Aos meus familiares e amigos, pelo apoio, incentivo e exemplos de vida.

Aos amigos do Centro Educacional de Santíssimo, do Centro de Estudos de Jovens e Adultos da Secretaria de Estado de Educação do Rio de Janeiro e do Laboratório de Biomecânica e Comportamento Motor da Universidade do Estado do Rio de Janeiro, sempre compreensivos e interessados nos meus estudos.

Finalmente, citando-os mais uma vez, agradeço aos meus pais, pela compreensão, apoio e carinho, que tornam a caminhada da vida mais tranquila e feliz. À minha irmã, Érica e ao meu cunhado, Raphael, pela amizade e companheirismo.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

CARACTERIZAÇÃO DAS INTERNAÇÕES POR INFARTO AGUDO DO  
MIOCÁRDIO NO ESTADO DO RIO DE JANEIRO POR MEIO DE ANÁLISE DE  
CORRESPONDÊNCIA MÚLTIPLA E DISTÂNCIA DE TOLERÂNCIA

Felipe Machado Huguenin

Outubro/2011

Orientadores: Antonio Fernando Catelli Infantosi.

Renan Moritz Varnier Rodrigues de Almeida.

Programa: Engenharia Biomédica

Doenças coronarianas, como o Infarto Agudo do Miocárdio (IAM), têm um grande impacto financeiro sobre o sistema de saúde, constituindo-se em cerca de 5% dos gastos com internação. O presente trabalho tem como objetivo caracterizar as internações dos pacientes do Sistema Único de Saúde por IAM, no Estado do Rio de Janeiro, por meio da identificação de agrupamentos sugeridos por uma Análise de Agrupamentos tradicional e por uma Análise de Correspondência Múltipla (ACM) associada a um índice heurístico de proximidade (Distância de Tolerância), e tendo como variável de maior interesse o gasto hospitalar. Investigou-se a existência de associação entre as categorias das variáveis, e, como decorrência, pôde-se associar maiores gastos com a utilização de CTI e com o uso de tecnologias de moderada complexidade, e menores com uso de tecnologias de baixa complexidade e não utilização de CTI. Por outro lado, a utilização de tecnologias de alta complexidade apresentou-se isolada, sem associação aparente com CTI ou com as demais variáveis. Também foi detectada associação entre a categoria “menores gastos” e as seguintes categorias: não-deslocamento do paciente, sexo feminino, idade entre 56 e 75 anos, óbito até trinta dias e óbito até um ano. A metodologia proposta mostrou-se uma ferramenta útil para caracterizar as internações, podendo contribuir para o processo de tomada de decisão por parte dos gestores em saúde e ratificando a importância da utilização de técnicas exploratórias para avaliação de serviços de saúde e análise de dados epidemiológicos.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

CHARACTERIZATION OF ADMISSIONS FOR ACUTE MYOCARDIAL  
INFARCTION IN RIO DE JANEIRO STATE USING MULTIPLE  
CORRESPONDENCE ANALYSIS AND TOLERANCE DISTANCE

Felipe Machado Huguenin

October/2011

Advisors: Antonio Fernando Catelli Infantosi.  
Renan Moritz Varnier Rodrigues de Almeida.

Program: Biomedical Engineering

Coronary heart diseases, like Acute Myocardial Infarction (AMI) have a large financial impact on the Brazilian Unified Health System (SUS), constituting about 5% of hospital admissions. This study aims to characterize the admissions of Brazilian Health System patients with AMI in Rio de Janeiro State, through the identification of clusters suggested by a traditional grouping analysis and a Multiple Correspondence Analysis (MCA) associated with a heuristic index for similar categories (Tolerance Distance); the variable of greatest interest is “hospital spending”. Concerning the association between variable categories, “higher costs” are associated with “use of ICU (Intensive Care Unit)” and “use of technologies of moderate complexity”, and “lower costs” were also associated with “use of technologies of low complexity” and “not using ICU”. On the other hand, cases of “use of technologies of high complexity” were non-clustered, with no apparent association with “use of ICU” or other categories. Other identified associations were “lower costs” and “no patient travelling”, “female”, “age between 56 and 75 years”, “death within thirty days” and “death within one year”. The proposed methodology proved to be a useful tool to characterize the hospital admissions profile and may contribute to decision making by managers and professionals inside the Brazilian Health System, confirming the importance of the use of exploratory techniques for evaluation of health services and epidemiological data analysis.

## SUMÁRIO

1	Introdução.....	1
1.1	Objetivo .....	3
1.2	Estrutura do trabalho .....	3
2	Infarto Agudo do Miocárdio.....	5
2.1	Epidemiologia .....	5
2.2	Fisiopatologia.....	12
2.3	Diagnóstico, condição de acesso e uso de tecnologias.....	15
3	Técnicas Multivariadas e aplicações ao IAM.....	17
3.1	Análise de Correspondência Múltipla .....	18
3.1.1	Distância de Tolerância.....	28
3.2	Cluster Hierárquico .....	30
3.3	Aplicação de Técnicas Multivariadas ao IAM.....	31
4	Materiais e Métodos.....	39
4.1	Base de Dados do IAM.....	39
4.1.1	Seleção das variáveis .....	39
4.2	Aplicação da ACM.....	45
5	Resultados.....	47
6	Discussão.....	55
7	Conclusão .....	63
8	Referências Bibliográficas.....	65
9	Anexo I.....	72

## LISTA DE FIGURAS

Figura 2.1 – Evolução da mortalidade proporcional segundo causas, Brasil.	7
Figura 2. 2 – Mortalidade proporcional segundo causas selecionadas, Brasil.	7
Figura 2. 3 – Volume de internações e óbitos por IAM, 1998 e 2008, RJ.	11
Figura 2. 4 – Gastos em reais com internações por IAM, 1998 e 2008, RJ.	11
Figura 2.5 – Diagrama do coração humano.	12
Figura 2.6 – Irrigação sanguínea do coração – circulação coronariana.	13
Figura 3.1 – Mapa de Correspondência gerado pelo programa R 2.11.1.	24
Figura 3.2 – Matriz indicadora ou Matriz disjunta.	25
Figura 3.3 – Matriz padronizada <b>G</b> .	26
Figura 4.1 – Fluxograma da seleção das variáveis.	41
Figura 5.1 – Mapa de Correspondência gerado por meio da ACM.	48
Figura 5.2 – Ampliação do Mapa de Correspondência.	49
Figura 5.3 – Gráfico de barras mostrando as associações encontradas.	52
Figura 5.4 – Dendrograma.	53
Figura 5.5 – Gráfico de análise do comportamento do nível de fusão.	54

## LISTA DE TABELAS

Tabela 2.1 – Óbitos segundo o capítulo CID 10, Brasil, 1998 e 2008.	8
Tabela 3.1 – Tabela de contingência.	19
Tabela 3.2 – Probabilidades marginais de linhas e colunas.	19
Tabela 3.3 – Valores esperados em cada casela.	20
Tabela 3.4 – Perfis de linha.	21
Tabela 3.5 – Perfis de coluna.	21
Tabela 3.6 – Frequências relativas.	22
Tabela 4.1 – Frequências absolutas e relativas das variáveis.	44
Tabela 5.1 – Associações encontradas para a variável desfecho.	51

## LISTA DE ABREVIATURAS E SIGLAS

AC	Análise de Correspondência
ACM	Análise de Correspondência Múltipla
AIH	Autorizações de Internação Hospitalar
AnC	Angioplastia Coronariana
BBS	Bulletin Board System
BVS	Biblioteca Virtual em Saúde
CEP	Comitê de Ética e Pesquisa
COPPE	Instituto Alberto Luiz Coimbra de Pós-graduação e Pesquisa de Engenharia
CTI	Centro de Tratamento Intensivo
DCNT	Doenças Crônicas Não-Transmissíveis
DIC	Doenças Isquêmicas do Coração
DIM	Dimensão
DIP	Doenças Infecto-parasitárias
DVS	Decomposição em Valores Singulares
EEAN	Escola de Enfermagem Anna Nery
IAM	Infarto Agudo do Miocárdio
IC	Intervalo de Confiança
InC	Insuficiência Coronariana
OMS	Organização Mundial de Saúde
OR	Odds Ratio
PEB	Programa de Engenharia Biomédica
ROC	Receiver Operator Characteristic
RVM	Revascularização Miocárdica
SIH	Sistema de Informações Hospitalares
SIM	Sistema de Informações sobre Mortalidade
SUS	Sistema Único de Saúde
UFRJ	Universidade Federal do Rio de Janeiro

# 1 INTRODUÇÃO

A partir da década de 1960, o cenário epidemiológico mundial revela uma diminuição das doenças infecto-parasitárias (DIP) e um crescimento das Doenças Crônicas Não-Transmissíveis (DCNT), das quais as doenças do aparelho circulatório merecem destaque devido à sua grande magnitude (OMS, 2011). Em 2004, segundo estimativas da Organização Mundial da Saúde (OMS), 17,1 milhões de indivíduos morreram devido a doenças cardiovasculares, sendo, aproximadamente, 7,2 milhões por doença cardíaca coronariana. Atualmente, estima-se que cerca de 20 milhões de indivíduos sofram de doenças do aparelho cardiovascular em todo mundo, dos quais, aproximadamente, 12 milhões são vítimas fatais de Infarto Agudo do Miocárdio (IAM) (OMS, 2011).

De modo similar ao que vem ocorrendo mundialmente, o Brasil tem revelado uma alteração no perfil de morbi-mortalidade da sua população. Enquanto em 1930 as doenças do aparelho circulatório representavam somente 12% das mortes da população, em 2003 já respondiam por cerca de 31% dos óbitos (MALTA *et al.*, 2006). Mais recentemente (2008), houve 318 mil mortes (152 mil mulheres e 166 mil homens) por doenças do aparelho circulatório no país, das quais quase 75 mil por ocorrência de IAM (MINISTÉRIO DA SAÚDE, 2011).

Também em 2008 foram registradas cerca de um milhão e cem mil internações devido a estas doenças, com gastos totais em torno de um bilhão e seiscentos milhões de reais (MINISTÉRIO DA SAÚDE, 2011). Assim, as doenças coronarianas têm um grande impacto financeiro sobre o sistema de saúde, constituindo-se em cerca de 5% dos gastos com internação (OLIVEIRA *et al.*, 2006). Dado esse alto custo, é de especial importância identificar variáveis que possam caracterizar ou prever esses gastos.

Ao mesmo tempo, a facilidade de acesso aos bancos de dados de utilização de serviços de saúde e a necessidade de justificar o uso crescente de recursos, como sempre, escassos, têm promovido um aumento no número de estudos sobre alocação de recursos em saúde no Brasil (RAMIARINA *et al.*, 2006). A melhor distribuição destes recursos permitiria atender às premissas do Sistema Único de Saúde (SUS): universalidade (saúde é direito de todos), equidade (diminuição das desigualdades sociais) e integralidade (as ações de saúde devem ser combinadas e voltadas para a prevenção, promoção e reabilitação dos pacientes).

Neste contexto, várias técnicas estatísticas têm sido propostas para análise de dados e auxílio à tomada de decisão na área de saúde (ESCOSTEGUY *et al.*, 2002; MELO *et al.*, 2004; ESCOSTEGUY *et al.*, 2005). Uma destas, a Análise de Correspondência Múltipla (ACM), que consiste basicamente na representação da posição relativa de variáveis em um Espaço Euclidiano, possui grande potencial para esse tipo de análise, no entanto, sua utilização ainda não é tão empregada. Esta técnica possibilita a visualização de relações entre as categorias de variáveis, sugerindo associações com base na distância entre elas (quanto mais próximas, maior a probabilidade de estarem associadas), sem necessidade de suposições como de gaussianidade e linearidade dos dados (LE ROUX e ROUANET, 2004; GREENACRE, 2007). Assim, a ACM pode ser empregada não somente como técnica exploratória, mas também com vistas a sugerir associações entre categorias de variáveis.

No entanto, como a avaliação de proximidade entre as categorias no “mapa de correspondência” (o espaço Euclidiano onde as variáveis são representadas), é, em geral, subjetiva, é necessário o desenvolvimento de índices que permitam uma avaliação mais objetiva das associações na ACM. Um exemplo é o índice heurístico de proximidade, baseado em uma distância Euclidiana ponderada, denominado *Distância*

de *Tolerância* (COSTA et al., 2008; ALMEIDA et al., 2009), o qual demanda ainda validação e investigação adicional, de forma a garantir e estender a sua aplicabilidade.

## 1.1 Objetivo

Caracterizar as internações dos pacientes do SUS por IAM, em municípios do Estado do Rio de Janeiro, por meio da identificação de agrupamentos sugeridos pela ACM e por um índice heurístico de proximidade – a *Distância de Tolerância*. Deste modo, espera-se obter informação que venha a contribuir com o uso mais eficiente dos recursos relativos às doenças do aparelho circulatório, principalmente do IAM, bem como auxiliar na tomada de decisão por parte dos gestores em saúde.

O trabalho possui como objetivos específicos:

1. Averiguar e identificar, por meio da metodologia proposta, as possíveis categorias de variáveis em associação, utilizando-se como variável desfecho: *gastos com paciente internado por IAM*;
2. Comparar os resultados do método da *Distância de Tolerância* (COSTA et al., 2008; ALMEIDA et al., 2009) com os obtidos por meio da técnica de *Cluster Hierárquico*;
3. Desenvolver e validar, por meio da técnica de *bootstrap*, o método da *Distância de Tolerância* (COSTA et al., 2008; ALMEIDA et al., 2009), proposto para facilitar a identificação de associações entre as categorias de variáveis estudadas.

## 1.2 Estrutura do trabalho

O presente trabalho é descrito em tópicos, com a introdução ao IAM no Capítulo 2, ressaltando a sua epidemiologia, fisiopatologia e diagnóstico. A introdução às técnicas de análise multivariadas é abordada no Capítulo 3, tanto em relação à Análise

de Correspondência (Simples e Múltipla) quanto à Distância de Tolerância e quanto a Análise de Agrupamento (*Cluster* Hierárquico). O Capítulo 4 apresenta a descrição completa dos dados (critérios de inclusão e exclusão e pré-processamento) e os procedimentos utilizados para o desenvolvimento da pesquisa, além da descrição do programa utilizado para as análises estatísticas. No Capítulo 5 são apresentados os resultados da Análise de Correspondência Múltipla em conjunto com a Distância de Tolerância e da Análise de Agrupamentos Hierárquicos. O Capítulo 6 apresenta a discussão dos resultados, comparando-os aos estudos expostos no Capítulo 3. Além disto, são apresentadas as conclusões e sugestões para os futuros trabalhos nesta linha de pesquisa.

## **2 INFARTO AGUDO DO MIOCÁRDIO**

Este capítulo aborda o Infarto Agudo do Miocárdio. De início, é exposto o processo de transição epidemiológica (a modificação do perfil de morbi-mortalidade de uma população), apresentando o cenário epidemiológico relacionado ao IAM, tanto no mundo, quanto no Brasil e, em especial, no Estado do Rio de Janeiro. A seguir, serão apresentados os principais conceitos fisiopatológicos e de diagnóstico de infarto. Além disso, a condição de acesso e o uso de tecnologias relacionadas a esta doença também serão abordadas.

### **2.1 Epidemiologia**

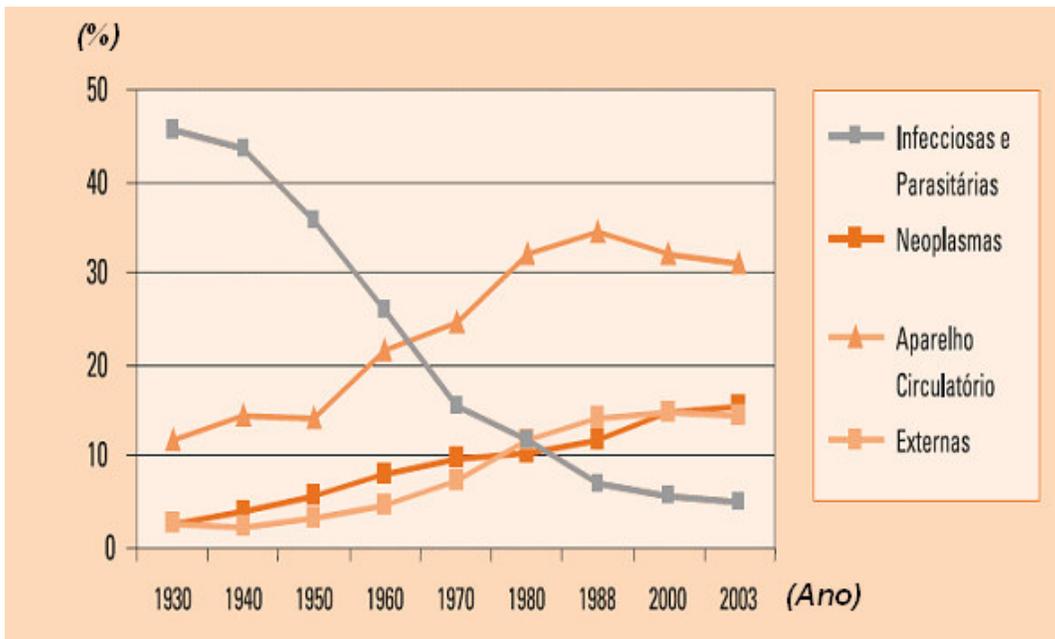
Em termos epidemiológicos, a partir da década de 1960 observam-se dois processos de transição que produziram importantes mudanças no perfil das doenças da população mundial. Um processo é denominado “transição demográfica”, com uma significativa diminuição das taxas de fecundidade, natalidade e aumento progressivo na expectativa de vida, caracterizando um crescente aumento do quantitativo de idosos em relação aos demais grupos etários; enquanto que o segundo, denominado “transição epidemiológica”, é representado pelo perfil de morbi-mortalidade da população, que vem sofrendo alterações (MINISTERIO DA SAÚDE, 2005).

O cenário epidemiológico mundial revela uma diminuição das DIP e um crescimento das DCNT. Estas são doenças multifatoriais, com interação entre vários fatores, caracterizadas por: “história natural prolongada, longo curso sem sintomas e período clínico lento, prolongado e permanente; manifestação clínica com períodos de remissão ou exacerbação; lesões celulares irreversíveis; e evolução para graus variados de incapacidade ou morte” (ROUQUAYROL & ALMEIDA FILHO, 2003).

Neste contexto, entre as DCNT, as doenças do aparelho circulatório têm a maior magnitude (OMS, 2011). Atualmente, estima-se que cerca de 20 milhões de indivíduos sofram de doenças do aparelho cardiovascular em todo mundo, dos quais, aproximadamente, 12 milhões são vítimas fatais de IAM (OMS, 2011). Baseado nos atuais padrões, a OMS estima cerca de 23 milhões de óbitos por doenças cardiovasculares em 2030 (OMS, 2011).

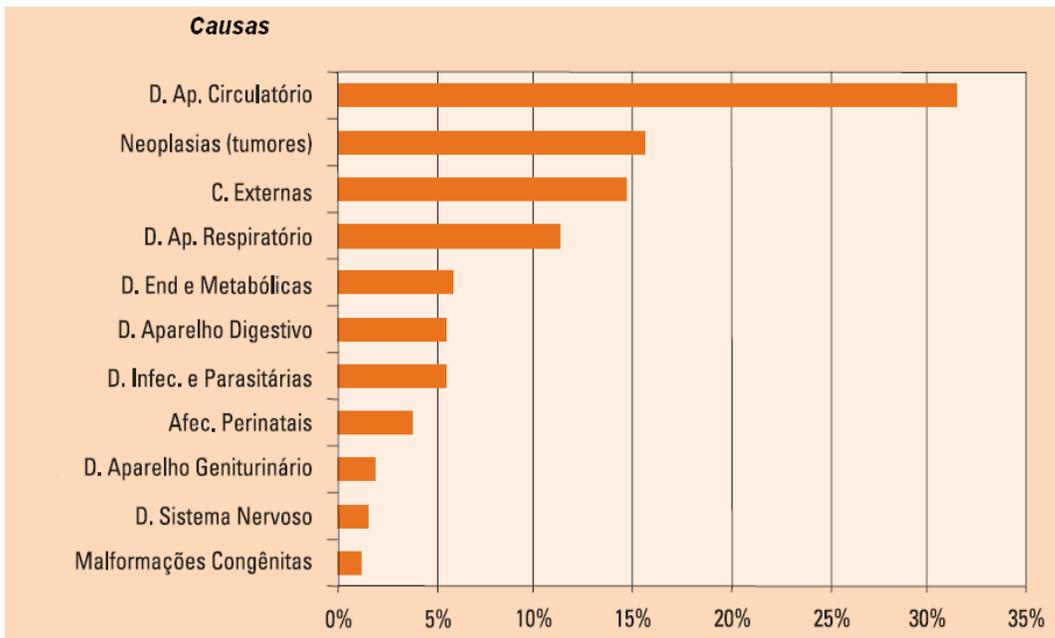
Seguindo a tendência mundial, o cenário epidemiológico brasileiro revela uma diminuição das DIP e um crescimento das DCNT, representado, principalmente, pelo aumento das doenças do aparelho circulatório. A Figura 2.1 apresenta a evolução da mortalidade segundo causas no Brasil de 1930 a 2003 e a Tabela 2.1 mostra o panorama dos óbitos, em números absolutos, segundo o capítulo CID-10 no Brasil nos anos de 1998 e 2008 (MINISTÉRIO DA SAÚDE, 2011).

Assim, pode-se dizer que as doenças cardiovasculares representam quase um terço das mortes no Brasil (Figura 2.2) (MINISTÉRIO DA SAÚDE, 2004). No entanto, o processo de transição epidemiológica brasileiro não pode ser visto de forma homogênea, como se ocorresse da mesma forma em todas as regiões (SCHRAMM *et al.*, 2004). Devido às diferenças culturais, históricas e sociais, as situações epidemiológicas de diferentes regiões tornam-se contrastantes, apresentando um modelo polarizado, com fases de transição distintas no mesmo país. Regiões de menor desenvolvimento socioeconômico podem experimentar a convivência com doenças infecto-parasitárias (DIP), enquanto regiões mais desenvolvidas podem apresentar um aumento das doenças crônicas não-transmissíveis (DCNT).



Fonte: Sistema de Informações sobre Mortalidade (SIM) – Ministério da Saúde – 2004.

**Figura 2.1** – Evolução da mortalidade proporcional segundo causas, Brasil.



Fonte: Sistema de Informações sobre Mortalidade (SIM) – Ministério da Saúde – 2004.

**Figura 2.2** – Mortalidade proporcional segundo causas selecionadas – Brasil 2003.

**Tabela 2.1** – Óbitos segundo o capítulo CID-10 – Brasil 1998 e 2008.

<i>Capítulo CID-10</i>	<i>1998</i>	<i>2008</i>
I. Algumas doenças infecciosas e parasitárias	48.792	47.295
II. Neoplasias (tumores)	110.799	167.677
III. Doenças do sangue	4.173	5.825
IV. Doenças endócrinas; nutricionais e metabólicas	39.841	64.631
V. Transtornos mentais e comportamentais	5.478	11.852
VI. Doenças do sistema nervoso	10.872	21.609
VII. Doenças do olho e anexos	21	39
VIII. Doenças do ouvido e da apófise mastóide	117	125
<b>IX. Doenças do aparelho circulatório</b>	<b>256.511</b>	<b>317.797</b>
X. Doenças do aparelho respiratório	91.983	104.989
XI. Doenças do aparelho digestivo	40.739	55.272
XII. Doenças da pele e do tecido subcutâneo	1.447	2.642
XIII. Doenças do sistema osteomuscular e tecido conjuntivo	2.076	4.094
XIV. Doenças do aparelho geniturinário	12.878	19.790
XV. Gravidez, parto e puerpério	1.997	1.691
XVI. Algumas afec. originadas no período perinatal	36.221	26.080
XVII. Malformação congênita; anomalias cromossômicas	9.331	10.502
XVIII. Sint. sinais e achad. anorm. Ex. clín. e laborat.	140.929	79.161
XX. Causas externas de morbidade e mortalidade	117.690	135.936

Fonte: Sistema de Informações sobre Mortalidade (SIM) – Ministério da Saúde – 2011.

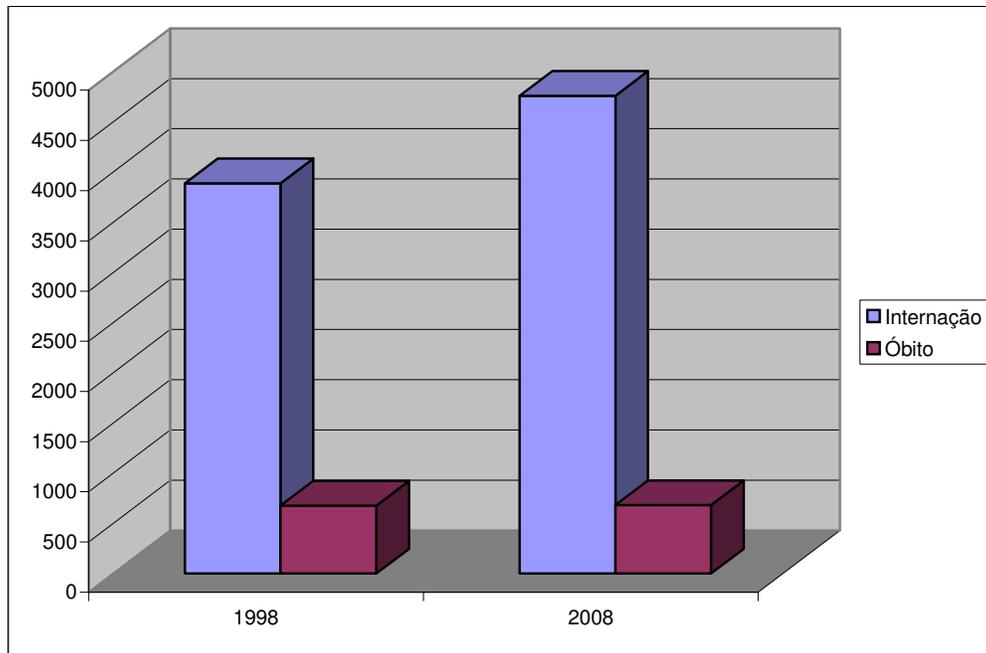
Segundo o MINISTÉRIO DA SAÚDE (2011), registraram-se cerca de 255 mil óbitos ocasionados por doenças cardiovasculares em 1998. Uma década depois, em 2008, ocorreram, aproximadamente, 318 mil mortes (152 mil mulheres e 166 mil homens), por doenças do aparelho circulatório, das quais, aproximadamente, 75 mil por ocorrência de infarto. Este aumento mostra-se proporcional ao crescimento demográfico, revelando uma possível estabilização do processo, provavelmente com a ajuda dos conhecimentos acumulados nos últimos anos, não só na terapêutica como também na prevenção da doença.

Ao mesmo passo que os óbitos aumentaram, as internações hospitalares por doenças do aparelho circulatório também cresceram. Em 1998 foram registradas um pouco mais de um milhão de internações devido a estas doenças, enquanto que, em 2008, este quantitativo ampliou para cerca de um milhão e cem mil internações. Estima-se que o total de gastos que era de aproximadamente 575 milhões de reais em 1998 elevou-se para cerca de 1 bilhão e 600 milhões de reais em 2008 (MINISTÉRIO DA SAÚDE/SIH/SUS, 2011). Crescimento ainda maior foi registrado com relação às internações por IAM. Em 1998, ocorreram quase 35 mil internações (5.717 óbitos), enquanto que, em 2008, perto de 62 mil foram registradas (8.510 óbitos). Os gastos, neste período, passaram de, aproximadamente, 27 milhões de reais para 150 milhões de reais (MINISTÉRIO DA SAÚDE/SIH/SUS, 2011). Esse acréscimo nos gastos talvez não decorra somente do aumento no número de internações, mas sim também da incorporação de novas tecnologias na terapêutica cardiovascular, as quais são em geral mais dispendiosas e complexas (RIBEIRO *et al.*, 2005).

No Estado do Rio de Janeiro, o panorama epidemiológico não difere do nacional. Segundo o MINISTÉRIO DA SAÚDE (2011), em 1998 ocorreram 34.314

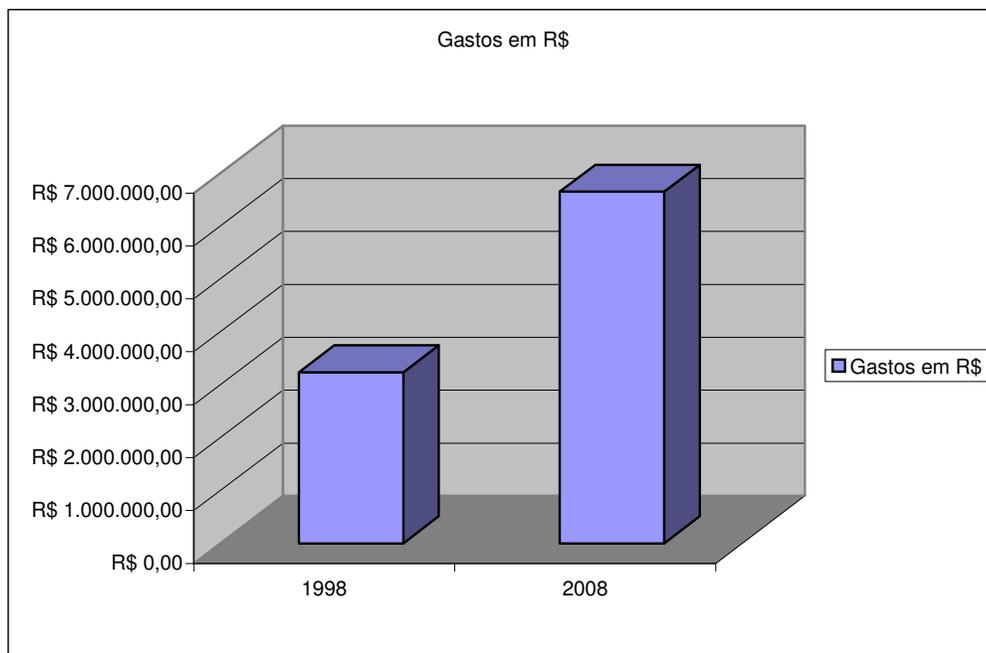
óbitos por doenças do aparelho circulatório, e, em 2008, este quantitativo ampliou-se para 37.080 óbitos. Ao mesmo tempo em que os óbitos totais cresceram no Estado, as internações diminuíram e os gastos aumentaram. Em 1998, foram registradas pouco mais de 90 mil internações (9.808 óbitos), enquanto que em 2008, perto de 76 mil internações foram notificadas (6.608 óbitos). Os gastos passaram de, aproximadamente, 54 milhões de reais para 115 milhões de reais (MINISTÉRIO DA SAÚDE/SIH/SUS, 2011).

Ao passo que as internações por doenças do aparelho circulatório diminuíram no Estado do Rio de Janeiro, especificamente nos casos de internação por IAM, esse quantitativo aumentou. Em 1998, ocorreram 3.888 internações, sendo que 678 pacientes foram a óbito, enquanto que, em 2008, 4.759 internações foram registradas, das quais 683 óbitos. Juntamente com este aumento, ampliaram-se também os gastos hospitalares, que passaram de, aproximadamente, 3,2 milhões de reais para 6,6 milhões de reais, no mesmo período de tempo (MINISTÉRIO DA SAÚDE/SIH/SUS, 2011). Observa-se, a partir destes dados, que os aumentos no número de pacientes internados e nos gastos não foram proporcionais. Enquanto as internações cresceram cerca de 22%, o aumento dos gastos está em torno de 100%. Isto indica que esse aumento não deve estar somente relacionado ao crescimento quantitativo das internações, mas também à incorporação de novas tecnologias. A Figura 2.3 mostra o volume de internações e óbitos por IAM e a Figura 2.4 apresenta os gastos em reais com internações por IAM, ambas relacionadas ao Estado do Rio de Janeiro nos anos de 1998 e 2008.



Fonte: Adaptado de Ministério da Saúde – Sistema de Informações Hospitalares do SUS, 2011.

**Figura 2.3** – Volume de internações e óbitos por IAM nos anos de 1998 e 2008, RJ.

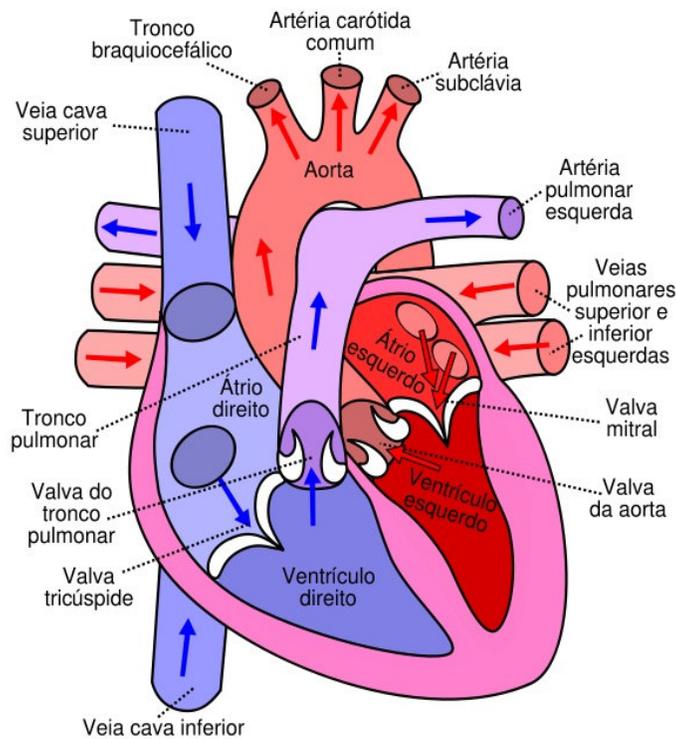


Fonte: Adaptado de Ministério da Saúde – Sistema de Informações Hospitalares do SUS, 2011.

**Figura 2.4** – Gastos em reais com internações por IAM nos anos de 1998 e 2008.

## 2.2 Fisiopatologia

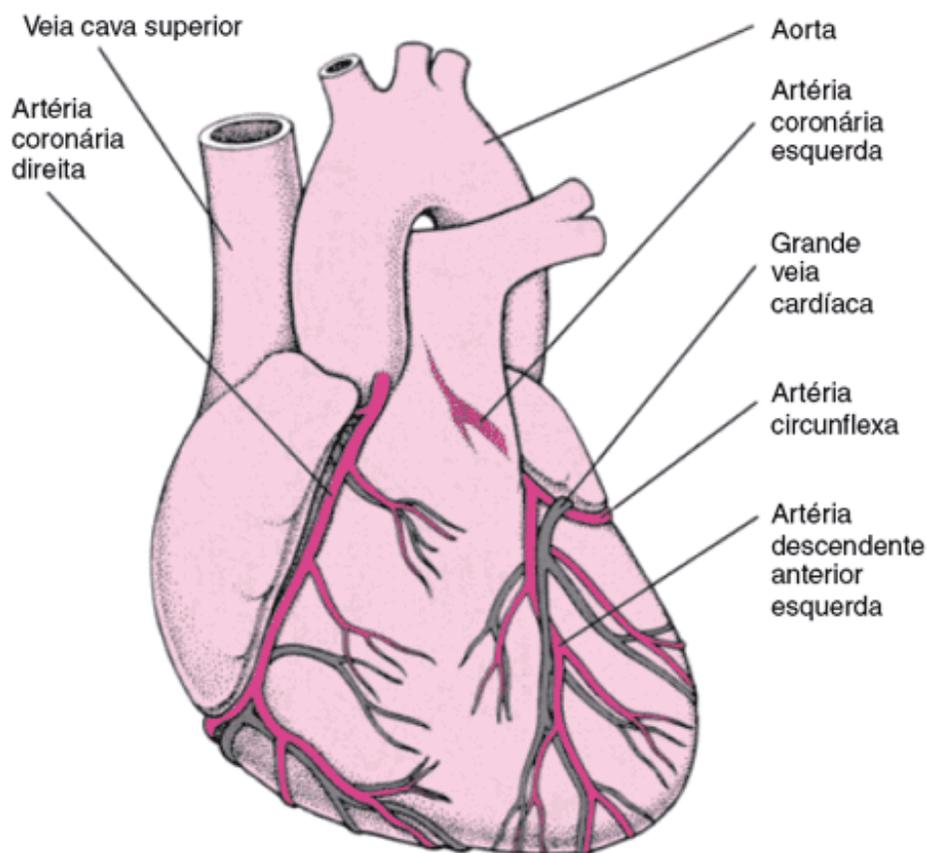
A cardiopatia isquêmica é a causa de morte mais comum na sociedade ocidental (Guyton e Hall 2006). Parte destas mortes ocorre abruptamente, como consequência da oclusão coronária aguda ou da fibrilação do coração, enquanto outras ocorrem lentamente, ao longo do período de semanas ou anos, em consequência do enfraquecimento progressivo do mecanismo de propulsão do sangue. A figura 2.5 apresenta de forma esquemática o coração humano.



Fonte: Wikicommons – Wikimedia.org – 2011.

**Figura 2.5** – Diagrama esquemático do coração humano. A circulação sanguínea pode ser dividida em dois grandes circuitos: *circulação pulmonar*, o sangue sai do ventrículo direito, passa pela artéria pulmonar, chega aos pulmões, retorna pelas veias pulmonares e chega ao átrio esquerdo; *circulação sistêmica*, o sangue sai do ventrículo esquerdo, passa pela artéria aorta, chega aos outros sistemas fisiológicos, retorna pelas veias cava e chega ao átrio direito.

A insuficiência coronária aguda ocorre quando o miocárdio fica privado de oxigênio por um extenso período e consegue restabelecer oxigenação. No entanto, quando o músculo cardíaco fica em isquemia intensa e contínua, causando lesões irreversíveis das células miocárdicas, apresenta-se o quadro clínico de IAM. Lesões teciduais irreversíveis originadas ao miocárdio ocorrem quando uma artéria coronária principal ou um de seus ramos sofre oclusão (Figura 2.6), que pode ser desencadeada por um coágulo sanguíneo, ou por lesões ateroscleróticas, ou, ainda, por uma partícula da placa ateromatosa que se desprende do seu local de origem e bloqueia uma ramificação.



Fonte: MSDONLINE – Manual Merck saúde para família, 2011.

**Figura 2.5** – Irrigação sanguínea do coração – circulação coronariana.

Com a prática de exercícios extenuantes ou quando se é acometido por uma anemia profunda e súbita, o miocárdio necessita de uma maior quantidade de oxigênio para manter as suas funções fisiológicas. Caso o aporte de oxigênio ao miocárdio fique acentuadamente comprometido, mesmo com as artérias coronárias preservadas, ou seja, com funcionamento fisiológico normal, pode ocorrer o infarto.

Com o IAM, três áreas do tecido lesado tornam-se aparentes, sendo bem distintas umas das outras. A primeira é a área da necrose, na qual as células do miocárdio encontram-se irreversivelmente lesadas. Ao redor destas células, encontra-se a área de tecido não-funcional, devido à incapacidade de contração e de condução de impulso nervoso, mas que pode ser recuperada se a oxigenação for adequadamente restabelecida. Mais ao redor, encontra-se a terceira área, que ainda mantém contração, porém fracamente, na qual a recuperação do tecido isquêmico é mais fácil (GUYTON e HALL, 2006)

Quanto à espessura do IAM, pode-se classificá-lo em: transmural, quando a lesão ocorre em todas as camadas do músculo cardíaco, possivelmente afetando a função propulsora do coração; ou, subendocárdico, quando só atinge as camadas internas do miocárdio, não comprometendo os movimentos cardíacos.

Além destas classificações, o IAM pode ser classificado quanto à localização da lesão. Denomina-se: infarto anterior, quando a obstrução encontra-se na artéria descendente anterior; infarto anterolateral, quando ocorre a obstrução da artéria diagonal; infarto inferior, quando a obstrução ocorre na coronária direita ou nos ramos da artéria circunflexa.

Quanto à sintomatologia, nem todos os pacientes têm dor no peito de forma intensa e aguda. Os sintomas do infarto podem variar de pessoa para pessoa, e entre os sintomas mais comuns, pode-se citar: angina peitoral ou desconforto torácico, dor

concentrada diretamente abaixo do esterno, com características de pressão ou aperto, que pode ser irradiada para o braço e para o pescoço; e sensação de desconforto nos ombros, costas, mandíbula ou no estômago. A dor pode ser acompanhada por náuseas, vômitos, sudorese excessiva, tonturas e desmaios. Palidez da pele, inquietação, palpitações e dispnéias também podem ocorrer (WERF *et al.*, 2003; BERTRAND *et al.*, 2002).

### **2.3 Diagnóstico, condição de acesso e uso de tecnologias**

Um tempo considerável pode decorrer do início da dor até a procura por socorro (primeiro momento), e, posteriormente, da procura por socorro até o deslocamento ao hospital mais próximo (segundo momento). Essa demora é nociva, pois a maioria das mortes relacionadas ao IAM ocorrem nas primeiras horas de manifestação da doença (TIMERMAN e FEITOSA, 2003). O primeiro momento encontra-se associado ao reconhecimento da dor e à procura imediata pelos serviços de emergência. O segundo está diretamente relacionado à assistência de saúde, que deve ser acessível e rápida, além de contar com recursos tecnológicos e humanos treinados para o diagnóstico do infarto.

Até o ano de 2000, seguindo recomendação da OMS, o IAM era diagnosticado com base na tríade: quadro clínico, alterações eletrocardiográficas e dosagem enzimática. No entanto, com o desenvolvimento de marcadores mais específicos para a detecção da lesão do músculo cardíaco, ocorreu a revisão dos critérios para diagnóstico do infarto (MYOCARDIAL INFARCTION REDEFINED, 2000). Entre as modificações, a utilização das troponinas como marcadores de necrose tecidual foi a mais significativa (NEWBY *et al.*, 2002).

Entretanto, mesmo com os novos critérios de diagnóstico, nem sempre o paciente pode ser prontamente diagnosticado e internado. O acesso aos serviços de

saúde, juntamente com a assistência ao paciente tem papel de destaque neste contexto, pois, como dito, o tempo decorrido entre o início dos sintomas e o atendimento ao indivíduo é fundamental para sua sobrevivência (MELO, 2007).

Contudo, sendo o IAM um evento agudo que sempre requer internação hospitalar, a caracterização das internações dos pacientes enfartados, obtendo-se o perfil das internações e os fatores associados, depende de um grande número de variáveis. Desta forma, torna-se necessária a utilização de métodos estatísticos que analisem simultaneamente múltiplas medidas (Análise Multivariada), permitindo assim uma visão global das variáveis analisadas, e, conseqüentemente uma visualização ampla do fenômeno.

### 3 TÉCNICAS MULTIVARIADAS E APLICAÇÕES AO IAM

Com a complexidade crescente dos problemas tratados pela ciência, muitas são as dificuldades em transformar dados em informações úteis. Cada vez mais, se faz necessária a utilização de ferramentas estatísticas que traduzam dados em informação. Entre estas ferramentas destaca-se a mencionada Análise Multivariada, que, diferentemente das ferramentas que analisam as variáveis de maneira isolada (estatística univariada) refere-se a métodos que utilizam simultaneamente variáveis para explicar um fenômeno (MOITA NETO, 2004; 2006). Assim, pode-se definir a Análise Multivariada como o conjunto de métodos e técnicas estatísticas que possibilitam analisar ao mesmo tempo várias dimensões (multidimensional) das variáveis em estudo, visando a estabelecer relações entre as mesmas (HAIR *et al.*, 2009).

Algumas técnicas de análise multivariada se destacam pela importância em apoiar o processo de decisão na área de saúde, tanto para um pesquisador quanto para um gestor. Neste capítulo buscou-se descrever os conceitos teóricos fundamentais ao desenvolvimento do estudo. O primeiro tópico apresenta a técnica de Análise de Correspondência Múltipla, principal técnica utilizada, sua origem histórica, conceitos e interpretações. Além disso, a Distância de Tolerância (índice heurístico de proximidade) é definida e explicada. O segundo tópico aborda um breve relato da técnica de Análise de Agrupamento (*Cluster* Hierárquico), a qual é, posteriormente, comparada com a ACM + Distância de Tolerância, para auxiliar no desenvolvimento desta. No terceiro tópico são expostos trabalhos na área de Saúde que utilizaram alguma das técnicas Multivariadas com aplicação ao IAM (Estado da Arte).

### 3.1 Análise de Correspondência Múltipla

A Análise de Correspondência Simples (AC), ao que tudo indica, tem sua origem na década de 30, com o trabalho de Hirschfield intitulado “*A Connection Between Correlation and Contingency*”, no qual a relação entre linhas e colunas de uma tabela de contingência é formulada algebricamente. No entanto, sua estruturação como análise de dados sistemática iniciou-se somente na década de 60, a partir dos estudos de um grupo de franceses liderados por *Jean Paul Benzécri*. A princípio a técnica foi denominada “*Analyse Factorielle des Correspondances*”, com o intuito de exprimir o sistema de associações entre os elementos de dois conjuntos. Posteriormente, o nome da técnica foi reduzido para “*Analyse des Correspondances*”, empregado atualmente (GREENACRE e HASTIE, 1987; LE ROUX e ROUANET, 2004; GREENACRE e NENADIC, 2007).

Na AC ocorre a conversão de uma matriz de dados não-negativos em uma representação gráfica, na qual linhas e colunas da matriz são expressas em um número menor de dimensões. Dessa maneira, a AC permite visualizar as relações e semelhanças existentes entre o conjunto de linhas e entre o conjunto de colunas de uma tabela de contingência. Para isso, as categorias das variáveis analisadas são representadas por meio de pontos em um diagrama de espalhamento, em um plano Euclidiano, de tal forma que suas posições relativas indicam níveis de similaridade ou associação. Sendo assim, a AC permite maior facilidade na apresentação dos resultados numéricos, facilitando a interpretação dos dados e gerando hipóteses que podem ser, posteriormente, testadas (GREENACRE, 1992).

Para facilitar a compreensão do método da AC, os dados utilizados por GREENACRE e HASTIE (1987) serão apresentados abaixo. A Tabela 3.1 mostra uma tabela de contingência I (linhas) x J (colunas), com as variáveis: grupos educacionais, 5

categorias, do menor grau de instrução para o maior, nas linhas; e leitura predileta, 3 categorias nas colunas. O total de indivíduos é de 312.

Um conceito fundamental para o cálculo do valor esperado é o da probabilidade marginal, assim designada por estar à margem da tabela de contingência. Esta probabilidade é calculada pelo valor total da linha ou da coluna dividido pelo total de indivíduos (N). Assim, a probabilidade marginal associada à primeira linha, ou seja, ao grupo educacional E1, é  $14/312=0,0449$ . A Tabela 3.2 mostra as probabilidades marginais de linhas e colunas.

**Tabela 3.1** – Tabela de contingência (Greenacre & Hastie, 1987). Grupos educacionais (E1, E2, E3, E4 e E5) e tipo predileto de leitura (C1, C2 e C3).

	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>Total</b>
<b>E1</b>	5	7	2	14
<b>E2</b>	18	46	20	84
<b>E3</b>	19	29	39	87
<b>E4</b>	12	40	49	101
<b>E5</b>	3	7	16	26
<b>Total</b>	57	129	126	N=312

**Tabela 3.2** – Tabela com as probabilidades marginais de linhas e colunas.

	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>Total</b>
<b>E1</b>	5	7	2	$14/312=0,0449$
<b>E2</b>	18	46	20	0,2692
<b>E3</b>	19	29	39	0,2788
<b>E4</b>	12	40	49	0,3237
<b>E5</b>	3	7	16	0,0833
<b>Total</b>	0,1827	0,4135	0,4038	N=312

Com as probabilidades marginais de linhas e colunas, pode-se calcular o valor esperado para cada casela, dado pelo produto das probabilidades marginais associadas à célula e o total de indivíduos. Estes valores estão na Tabela 3.3.

**Tabela 3.3** – Tabela com os valores esperados em cada casela.

	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>Total</b>
<b>E1</b>	0,0449.0,1827.312 = 2,6	5,8	5,7	14
<b>E2</b>	15,3	34,7	33,9	84
<b>E3</b>	15,9	36,0	35,1	87
<b>E4</b>	18,5	41,8	40,8	101
<b>E5</b>	4,8	10,8	10,5	26
<b>Total</b>	57	129	126	<i>N</i> =312

Uma das formas de averiguar a dependência entre as variáveis é pela análise dos resíduos padronizados, denominados Resíduos de Pearson (3.1), dados por:

$$resíduos = \sum \frac{(O_{ij} - E_{ij})}{\sqrt{E_{ij}}} \quad (3.1)$$

onde *O* e *E* são os valores observado e esperado, respectivamente.

Dessa forma, cada célula *ij* da tabela possui um resíduo padronizado, e, quanto mais ele se afasta do seu valor esperado, formado por duas categorias de variáveis distintas, mais a célula contribui para a independência entre as variáveis.

Mais do que analisar a dependência entre as variáveis, a AC tem a finalidade de decompor os dados e analisar as interrelações entre as categorias das variáveis. A partir da transformação dos valores observados em Perfis de Linha e Perfis de Coluna

(Tabelas 3.4 e 3.5, respectivamente), apresentam-se graficamente as categorias que possuem maior ou menor prevalência, ou seja, as categorias que mais ou menos contribuem para a representação.

**Tabela 3.4** – Perfis de Linha.

	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>Total</b>
<b>E1</b>	0,3571	0,5000	0,1429	1
<b>E2</b>	0,2143	0,5476	0,2381	1
<b>E3</b>	0,2184	0,3333	0,4483	1
<b>E4</b>	0,1188	0,3960	0,4851	1
<b>E5</b>	0,1154	0,2692	0,6154	1
<b>Média</b>	0,1827	0,4135	0,4038	

**Tabela 3.5** – Perfis de Coluna.

	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>Média</b>
<b>E1</b>	0,0877	0,0543	0,0159	0,0449
<b>E2</b>	0,3158	0,3566	0,1587	0,2692
<b>E3</b>	0,3333	0,2248	0,3095	0,2788
<b>E4</b>	0,2105	0,3101	0,3889	0,3237
<b>E5</b>	0,0526	0,0543	0,1270	0,0833
<b>Total</b>	1	1	1	

A proporção de E1 na população, denominada Massa do Ponto, é  $14/312=0,0449$ , o que representa o peso de E1 nesta população, ou seja, a medida de importância entre cada categoria de variável. A coluna **Média** da Tabela 3.5 apresenta as demais massas das categorias da variável grupo educacional. Estes valores são entendidos como um ponto no espaço Euclidiano multidimensional, dado pelas coordenadas do vetor  $I=[0,0449; 0,2692; 0,2788; 0,3237; 0,0833]$ . Analogamente, em

relação à leitura favorita existe o vetor  $c=[0,1827; 0,4135; 0,4038]$ . O vetor  $l$  representa o Centróide de Linha e o vetor  $c$  o Centróide de Coluna. O Centróide é a média das distâncias dos pontos ponderada pelas massas associadas a estes. Também chamado de centro de gravidade, funciona como se fosse o ponto de equilíbrio da matriz de dados. Assim, quanto mais a massa se aproxima da unidade, maior será a sua importância e mais próxima do centróide estará a categoria correspondente à variável.

Outra forma de obter-se estas informações é por meio da tabela de frequências relativas, com a qual obtém-se a proporção de cada combinação das categorias de variáveis e também do total das categorias em relação ao total de unidades classificadas. Para isto, todas as células foram divididas pelo total de indivíduos, neste caso, 312. A Tabela 3.6 apresenta as frequências relativas.

**Tabela 3.6** – Frequências relativas.

	<b>C1</b>	<b>C2</b>	<b>C3</b>	<b>Totais</b>
<b>E1</b>	0,0160	0,0224	0,0064	0,0449
<b>E2</b>	0,0577	0,1474	0,0641	0,2692
<b>E3</b>	0,0609	0,0929	0,1250	0,2788
<b>E4</b>	0,0385	0,1282	0,1571	0,3237
<b>E5</b>	0,0096	0,0224	0,0513	0,0833
<b>Totais</b>	0,1827	0,4135	0,4038	<b>1</b>

Finalmente, para se obter as coordenadas das categorias das duas variáveis, utiliza-se o algoritmo da Decomposição em Valores Singulares (DVS). Este, aplicado a uma matriz retangular (Matriz de dados), a decompõe em três matrizes: vetores singulares relativos às linhas (matriz  $U$ ), vetores singulares relativos às colunas (matriz

$V$ ) e matriz diagonal e quadrada, composta pelos valores singulares associados às linhas e colunas (matriz  $D$ ), dadas por (2) (GREENACRE, 1992).

$$\text{Matriz de dados} = U.D.V^T \quad (3.2)$$

Aplicando-se o algoritmo da DVS à matriz de dados do exemplo utilizado (Tabela 3.6), obtêm-se as matrizes  $U$ ,  $D$  e  $V$ , conforme exposto posteriormente:

$$\begin{bmatrix} 0,0160 & 0,0224 & 0,0064 \\ 0,0577 & 0,1474 & 0,0641 \\ 0,0609 & 0,0929 & 0,1250 \\ 0,0385 & 0,1282 & 0,1571 \\ 0,0096 & 0,0224 & 0,0513 \end{bmatrix} \quad (\text{Matriz de dados})$$

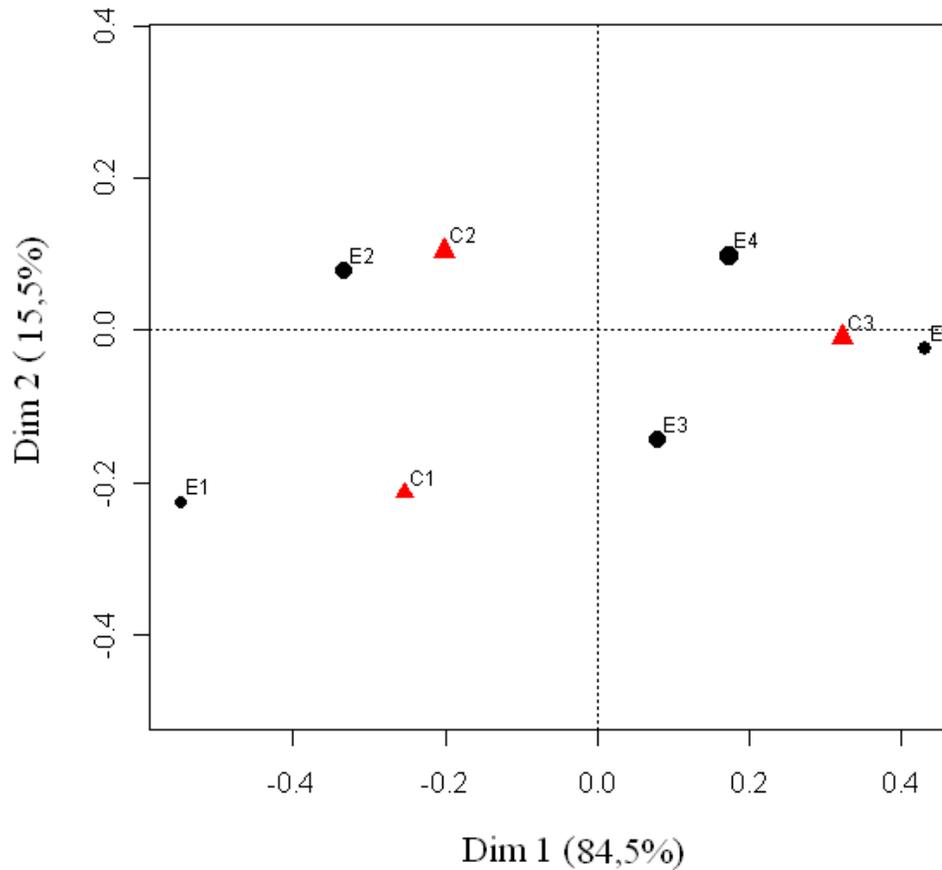
$$\begin{bmatrix} -0,076 & 0,183 & 0,255 \\ -0,512 & 0,818 & -0,070 \\ -0,528 & -0,271 & 0,761 \\ -0,652 & -0,368 & -0,592 \\ -0,167 & -0,298 & 0,004 \end{bmatrix} \quad (\text{Matriz } U)$$

$$\begin{bmatrix} 0,312 & 0 & 0 \\ 0 & 0,072 & 0 \\ 0 & 0 & 0,026 \end{bmatrix} \quad (\text{Matriz } D)$$

$$\begin{bmatrix} -0,287 & 0,234 & 0,929 \\ -0,682 & 0,631 & -0,370 \\ -0,672 & -0,740 & -0,021 \end{bmatrix} \quad (\text{Matriz } V)$$

Dessa forma, a partir de uma tabela de contingência padronizada como uma matriz de I linhas e J colunas, a Matriz  $U$  representa as coordenadas das observações no mapa de correspondência, a Matriz  $D$  é composta pelos autovalores da Matriz de dados

e a Matriz  $V$  representa os autovetores desta mesma matriz. Com isto, é possível representar geometricamente os dados no Espaço Euclidiano, tornando mais fácil estabelecer as relações existentes entre as variáveis estudadas ao observar a visualização. A Figura 3.1 representa o mapa de correspondência do exemplo utilizado.



**Figura 3.1** – Mapa de Correspondência gerado pelo Programa R 2.11.1 com os dados de GREENACRE e HASTIE (1987).

O percentual total da variação explicada pelo mapa de correspondência é uma medida de acurácia da solução obtida, e corresponde ao percentual entre parênteses das dimensões 1 e 2 da Figura 3.1. A variação explicada, mais conhecida com Inércia, é a

quantificação da variação total de cada nuvem de pontos. A dimensão 1 (Dim 1) contém a maior inércia, ou seja, é responsável pelo maior percentual da variação explicada. A dimensão 2 (Dim 2) é o vetor perpendicular ao primeiro eixo (Dim 1), também tendo a maior inércia possível.

A ACM é a aplicação da Análise de Correspondência Simples a uma base de dados categóricos multidimensional, os quais são dispostos em uma matriz disjunta, denominada Matriz Indicadora, e não mais em uma tabela de contingência. Neste caso, os pacientes ou indivíduos são representados nas linhas ( $n$ ), e as categorias das variáveis nas colunas ( $p$ ) (GREENACRE, 2007). A Figura 3.2 mostra uma Matriz Indicadora simplificada, ou seja, com a quantidade de colunas reduzida, a partir dos dados desta dissertação.

Pacientes	Variáveis						Total
	Gastos (GAST)		Tipo de hospital (HOSP)		Deslocamento (DESL)		
	inferiores a R\$ 905,00	superiores a R\$ 905,00	Contratado	Próprio	Sem	Com	
1	1	0	1	0	1	0	3
2	0	1	0	1	1	0	3
...	...	...	...	...	...	...	...
4435	0	1	1	0	0	1	3
Total	3212	1223	1276	3159	3781	654	13305

**Figura 3.2** – Matriz indicadora ou Matriz disjunta.

Por meio da padronização de cada elemento desta matriz ( $z_{ij}$ ) e pelos correspondentes perfis de linha ( $r_i$ ) e coluna ( $c_j$ ), relativos às respectivas massas, obtém-

se a matriz padronizada **G** (Figura 3.3), por meio da expressão posterior (3.3):

$$g_{ij} = \frac{z_{ij}}{\sqrt{r_i \cdot c_j}} \text{ para } 0 \leq i \leq n \text{ e } 0 \leq j \leq p. \quad (3.3)$$

Pacientes	Variáveis					
	Gastos (GAST)		Tipo de hospital (HOSP)		Deslocamento (DESL)	
	inferiores a R\$ 905,00	superiores a R\$ 905,00	Contratado	Próprio	Sem	Com
1	0,01019	0,00000	0,01616	0,00000	0,00939	0,00000
2	0,00000	0,01651	0,00000	0,01027	0,00939	0,00000
...	...	...	...	...	...	...
4435	0,00000	0,01651	0,00000	0,01027	0,00000	0,02258
Total	1	1	1	1	1	1

**Figura 3.3** – Matriz padronizada **G**.

A partir da fatoração da matriz **G** por meio do algoritmo da Decomposição por Valores Singulares (DVS), descrito no tópico sobre AC, obtém-se (3.4):

$$G = U \cdot D \cdot V^T \quad (3.4)$$

onde **U** e **V** são matrizes de vetores singulares, o primeiro relativo às linhas, e o segundo, às colunas. **D** é a matriz diagonal e quadrada, composta pelos valores singulares associados a cada uma das linhas e colunas. Desta forma, as coordenadas

principais das linhas e colunas são calculadas pelo produto das coordenadas padrões e seus respectivos valores singulares (GREENACRE, 1992).

Como na ACM os valores singulares das categorias-pontos não exprimem a variância, ou seja, a inércia. Com vistas a garantir que a inércia calculada pela DVS seja compatível com a variabilidade dos dados, Benzécri sugeriu um ajuste (GREENACRE, 2007), expresso por (3.5):

$$\tilde{\lambda}_j^2 = \left( \frac{Q}{Q-1} \right)^2 \left( \lambda_j^2 - \frac{1}{Q} \right)^2 \quad (3.5)$$

onde  $Q$  é o número de variáveis ativas,  $\lambda_j$  os valores singulares referentes à Matriz Indicadora e  $\tilde{\lambda}_j$  o valor ajustado.

Os passos para a obtenção do mapa de correspondência a partir de uma Matriz Indicadora, com  $1 < i < n$  categorias nas linhas e  $1 < j < p$  categorias nas colunas, são os seguintes:

1. Obtenção dos vetores  $r$  e  $c$ , relativos aos perfis das linhas e colunas, respectivamente, dados por:

$$r_i = \sum_{j=1}^p z_{ij} \quad \text{e} \quad c_j = \sum_{i=1}^n z_{ij} ;$$

2. Padronização da Matriz Indicadora em seus valores relativos, ou seja, obtenção da matriz padronizada  $G$ , por meio da expressão:

$$g_{ij} = \frac{z_{ij}}{\sqrt{r_i \cdot c_j}} \quad \text{para } 0 \leq i \leq n \text{ e } 0 \leq j \leq p. ;$$

3. Fatoração da Matriz  $\mathbf{G}$  por meio do algoritmo da Decomposição em Valores Singulares (DVS):

$$G = U \cdot D \cdot V^T ;$$

4. Ajuste da inércia, para garantir que esta seja compatível com a variabilidade dos dados, utilizando a fórmula sugerida por Benzécri:

$$\tilde{\lambda}_j^2 = \left( \frac{Q}{Q-1} \right)^2 \left( \lambda_j^2 - \frac{1}{Q} \right)^2 ;$$

5. Obtenção do Mapa de Correspondência;

### 3.1.1 Distância de Tolerância

Como mencionado, a interpretação usual do Mapa de Correspondência, resultante da ACM, baseia-se na proximidade espacial entre as categorias das diferentes variáveis. No entanto, esta proximidade depende, entre outras, da inércia das dimensões consideradas, da escala do gráfico e, mesmo, da experiência do usuário. Uma alternativa para a avaliação da associação entre as categorias é a utilização de um índice heurístico de associação, denominado *Distância de Tolerância*, expresso por (3.6) (COSTA et al, 2008; ALMEIDA et al, 2009):

$$\text{Distância de tolerância} = \sqrt{(X - Y) \cdot \Lambda \cdot (X - Y)^T} \quad (3.6)$$

onde  $\mathbf{X}$  e  $\mathbf{Y}$  são as coordenadas principais das categorias;  $\mathbf{X} = (x_1, x_2, \dots, x_s)$ ,  $\mathbf{Y} = (y_1, y_2, \dots, y_s)$ , sendo que  $\underline{s}$  é a máxima dimensão obtida, ou por meio de um *scree plot* ou pré-determinada com base na variação explicada pelos eixos.  $\Lambda$  é a matriz diagonal,

cujos elementos  $\alpha_j$  são dados por (3.7), e, logo, a distância de tolerância é uma distância Euclidiana ponderada, na qual os pesos são dados pela razão entre as inércias ajustadas e a inércia da primeira dimensão.

$$\alpha_j = \frac{\tilde{\lambda}_j^2}{\tilde{\lambda}_1^2} \text{ para } 0 \leq j \leq p. \quad (3.7)$$

Quando existe necessidade de estabelecimento de associações entre variáveis explicativas e a variável resposta, como em uma regressão, a variável resposta pode ser inserida no gráfico da análise de correspondência como variável suplementar, por meio de fórmulas de transição (GREENACRE, 2007). Deste modo, esta variável não influencia os cálculos das coordenadas principais e nem na variabilidade. A distância de tolerância fornece uma interpretação intuitiva de proximidade entre as categorias, possibilitando inferências posteriores sobre as relações entre as variáveis. Quando comparada com as distâncias Euclidianas normal e estatística, apresenta um menor limiar para associação.

### 3.1.1.1 O método *Bootstrap*

O *bootstrap*, terminologia introduzida por Efron em 1979, é basicamente uma técnica de reamostragem (amostragem com reposição), que permite, a partir da amostra observada, fazer suposições sobre a população em estudo (DAVINSON e HINKLEY, 1997). A amostragem é feita, com reposição, da distribuição da qual os dados são obtidos, se esta é conhecida (*bootstrap* paramétrico) ou da amostra original (*bootstrap* não-paramétrico).

Na técnica *bootstrap* as observações são escolhidas de forma aleatória e as estimativas recalculadas. Normalmente, essa metodologia é utilizada para calcular o

intervalo de confiança de parâmetros e calcular os valores-p, determinando assim regiões de confiança (LEBART, 2007). No entanto, no presente trabalho, o mesmo será empregado para determinar distâncias de tolerância “espelhadas” no espaço Euclidiano original. Desta forma, as associações encontradas no espaço original poderão ser comparadas com aquelas dadas pelos “espaços espelhados”.

Como em toda análise estatística, a verificação da robustez quanto ao resultado encontrado é desejável, principalmente ao ser introduzido um novo método. Assim, a técnica *bootstrap* será empregada para a validação dos resultados sugeridos pela metodologia da *Distância de Tolerância*.

### **3.2 Cluster Hierárquico**

A análise de agrupamentos (*cluster analysis*) é um conjunto de técnicas estatísticas multivariadas, cujo objetivo principal é classificar indivíduos ou categorias de variáveis (objetos) em grupos com base nas características que eles possuem. Esta análise reúne os objetos em relação a um conjunto de características selecionadas, de modo que cada um deles é considerado similar a outros no agrupamento. Os grupos resultantes devem ter elevada homogeneidade interna (dentro do grupo) e elevada heterogeneidade externa (entre os grupos) (MINGOTI, 2005).

O *Cluster* Hierárquico tem sido há muito tempo a metodologia mais popular de agrupamento, envolvendo uma série de  $n - 1$  decisões ( $n$  número de observações) que combinam observações em uma estrutura hierárquica ou tipo árvore. Existem dois tipos de procedimentos hierárquicos: aglomerativo, com o qual cada objeto começa como seu próprio agrupamento; e, divisivo, com o qual todas as observações iniciam em um agregado e são sucessivamente divididas até que cada uma forma um agregado unitário (HAIR *et al.*, 2009).

O procedimento para a escolha do número de grupos para definir a partição do conjunto de dados analisados tem grande importância. Uma das formas de avaliar isto é por meio da análise do comportamento do “nível de fusão”. À medida que a similaridade entre os grupos agregados vai decrescendo e, portanto, a distância entre os agrupamentos aumentando, pode-se apresentar um gráfico do número de grupos *versus* o nível de distância do agrupamento (fusão) em cada estágio do processo. A ocorrência de “pontos de salto” indica o momento ideal de parada do algoritmo, ou seja, o número de grupos final. Quando o segmento de reta no gráfico se estabiliza, não apresentando variações significativas, obtém-se a partição ótima (MINGOTI, 2005). Depois de estabelecido o número de agrupamentos, a forma gráfica mais utilizada para representar o resultado final dos diversos agrupamentos é o dendrograma, no qual a escala vertical indica o nível de similaridade entre os objetos e no eixo horizontal são marcados os objetos utilizados.

### **3.3 Aplicação de Técnicas Multivariadas ao IAM**

OLIVEIRA *et al.* (2006) publicaram um artigo sobre a letalidade das Doenças Isquêmicas do Coração (DIC) no estado do Rio de Janeiro no período entre 1999 e 2003. Utilizaram dados derivados das AIH pagas pelo SUS, tendo como objetivo estimar as taxas de letalidade por DIC aguda e crônica e por revascularização miocárdica (RVM) e angioplastia coronariana (AnC) no estado do Rio de Janeiro. Os autores lembram que as AIH, embora tenham sido criadas como mecanismo de pagamento da assistência hospitalar prestada, mostraram-se instrumentos adequados para avaliar os dados relacionados às internações nas unidades de saúde, servindo como fonte de informação útil para a tomada de decisões gerenciais.

O estudo constatou que, do total de 68.375 internações ocorridas por DIC no período estudado, 34% dos pacientes tiveram o diagnóstico de IAM, dos quais 1,4%

foram revascularizados cirurgicamente e 3,4% foram submetidos à angioplastia coronariana. As letalidades por angina, IAM, outras DIC agudas e DIC crônicas foram de 2,8%, 16,2%, 2,9% e 3,9%, respectivamente. Mesmo sem identificar os hospitais, observou-se que existem diferenças entre as unidades prestadoras de serviços ao SUS, por meio das informações de letalidade. De modo geral, concluiu-se que a RVM e AnC no tratamento das DIC no estado do Rio de Janeiro vêm aumentando progressivamente e que as taxas de letalidade foram elevadas tanto nos Procedimentos de Alta Complexidade como no tratamento clínico de IAM, esta, por sua vez, semelhante à existente quando não se utilizam trombolíticos (16,7%).

Três estudos tiveram como objetivo analisar a pertinência do uso do Sistema de Informações Hospitalares (SIH) do SUS, porém cada um deles se propôs a fazer isto com objetivos específicos bem distintos. O primeiro, realizado por ESCOSTEGUY *et al.* (2002), analisou 1936 formulários de AIH do SIH/SUS registrados com diagnóstico principal de IAM na cidade do Rio de Janeiro em 1997, e também uma amostra aleatória de 391 prontuários médicos estratificada por hospitais. Após estes dois passos, analisou-se a concordância entre o diagnóstico principal registrado na AIH e o registro do prontuário médico, por meio do coeficiente de correlação intraclasse *Kappa*. O estudo constatou que a qualidade do registro de IAM notificado na AIH foi satisfatória, tendo, segundo critérios estabelecidos pela literatura específica, um percentual de confirmação alta (cerca de 91,7%, IC95% = 88,3-94,2). De modo geral, a precisão das variáveis, tanto demográficas (sexo, faixa etária), de processo (procedimentos, intervenções) e de resultado (óbito, motivo da saída) foi satisfatória. A maior limitação enfrentada foi o elevado sub-registro do diagnósticos secundários na AIH. Assim, considerando-se a ampla disponibilidade da base de dados; disponível pelo DATASUS via *internet* e em CD-ROM, concluiu-se que a utilização do SIH, como instrumento

auxiliar, é pertinente na avaliação da qualidade da assistência ao infarto, ao menos no município do Rio de Janeiro.

No segundo estudo, MELO *et al.* (2004) tinham como objetivo específico avaliar a qualidade da informação sobre óbito por IAM utilizando o SIH/SUS e o Sistema de Informações sobre Mortalidade (SIM). Em uma primeira etapa, foram analisados os dados sobre mortalidade hospitalar por IAM no município do Rio de Janeiro no ano de 2000 referentes a essas bases de dados. Na segunda etapa, ocorreu o levantamento dos dados dos prontuários médicos de dois hospitais do SUS da cidade do Rio de Janeiro. Neste caso também foi analisada a concordância entre os dados das declarações de óbito e da AIH, com base nos registros dos prontuários médicos, pelo *Kappa* de Cohen e pelo coeficiente de correlação intraclasse. O estudo verificou que o total de óbitos hospitalares por IAM registrados no SIM/SUS era expressivamente maior que no SIH/SUS. O SIM registrou 4007 óbitos em que o infarto foi a causa básica, dos quais, 2405 ocorreram em hospital. Em contra-partida, o SIH/SUS registrou 2896 internações com diagnóstico principal de IAM, sendo notificados 438 óbitos. Esta discrepância foi explicada a partir de três fontes: ausência na emissão de AIHs, notificação de outro diagnóstico principal no SIH/SUS e sub-notificação do óbito na AIH.

Nesse segundo estudo, a proporção de confirmação do diagnóstico de IAM na AIH foi mais baixa do que a observada por ESCOSTEGUY *et al.* (2002) no município do Rio de Janeiro, em 1997. No entanto, se fossem considerados apenas os casos de óbito, as proporções seriam similares. Além disto, os dois estudos confirmaram o elevado sub-registro do diagnóstico secundário na AIH. Contudo, Melo *et al.* concluíram que em relação à validade, verificou-se um alto grau de sensibilidade no SIH/SUS e no SIM e uma confiabilidade satisfatória para grande parte das variáveis analisadas, corroborando os resultados de ESCOSTEGUY *et al.* (2002). Entretanto,

acrescentaram que é necessário implementar medidas voltadas para a melhoria da qualidade do SIH/SUS, tais como a padronização de critérios para emissão de AIH e o treinamento das equipes responsáveis pelos registros.

O terceiro artigo (ESCOSTEGUY *et al.*, 2005), desenvolvido pela mesma equipe do primeiro, teve como objetivo avaliar o banco AIH em relação ao prontuário médico na análise de fatores associados à variação do risco de óbito hospitalar no IAM, no município do Rio de Janeiro, em 1997. O universo do estudo foi definido com base no diagnóstico principal encontrado na AIH, sendo utilizados 1936 registros. A partir destes registros, realizou-se um estudo seccional de uma amostra aleatória de 391 prontuários médicos, estratificada por hospitais. Os fatores associados com o óbito hospitalar por infarto foram estudados por meio de modelagem logística. Os modelos foram obtidos diretamente a partir das AIH e dos prontuários médicos. Ao final, foram construídas curvas ROC (*Receiver Operator Characteristic*) para os modelos estudados, permitindo sua comparação.

O estudo encontrou uma confirmação diagnóstica do IAM igual a 91,7%. O modelo logístico derivado dos prontuários médicos teve um melhor ajuste, obtendo uma concordância de 90,1%. Embora o modelo desenvolvido com as AIH não tivesse um ajuste tão bom quando comparado ao dos prontuários, com concordância de 70,6%, o seu desempenho não foi modificado significativamente. Além disto, a confirmação do diagnóstico de IAM e a precisão das variáveis da AIH foram satisfatórias, o que também foi concluído nos dois artigos anteriores. Os autores destacam que um melhor preenchimento das autorizações poderia melhorar ainda mais o ajuste do modelo final.

Como visto, o infarto é uma patologia que acomete um número elevado de indivíduos, sendo estudado em todo o mundo. Na Espanha, o estudo realizado por SENDRA *et al.* (2005) analisou a mortalidade hospitalar por IAM e os fatores

associados a ela, tendo em vista que as enfermidades cardiovasculares (e, dentro deste grupo, as doenças esquêmicas do coração) são a primeira causa de morte na Espanha.

A mortalidade hospitalar por episódios de IAM foi estudada por meio da base de dados da comunidade de Madri no ano de 2001, baseando-se em variáveis sócio-demográficas e hospitalares, envolvendo procedimentos realizados, fatores de risco e co-morbidades associadas. As técnicas estatísticas descritivas e a regressão logística foram utilizadas para analisar os dados. Constatou-se que dos 5306 registros de infarto do estudo, 71% são homens e a idade média dos pacientes é de 68 anos. 29% dos pacientes fizeram revascularização com colocação de *stent* coronariano. A mortalidade averiguada foi de 10,8%. Na análise multivariada, o aumento da idade, a presença de arritmias, a insuficiência cardíaca e a insuficiência renal são associados a uma maior mortalidade.

Esse estudo concluiu que a mortalidade hospitalar por IAM, considerando-se uma base de dados administrativa, é similar à detectada por estudos clínicos. Isto confirma o que OLIVEIRA *et al.* (2006) apontou, no Brasil, sobre bases administrativas, ou seja, elas servem como fonte de informação útil para a tomada de decisões gerenciais. O trabalho também afirma que a mortalidade é associada a variáveis mencionadas na literatura, mas que novos estudos são necessários para confirmar a associação com outras variáveis.

No Brasil, o estudo realizado por EVANGELISTA *et al.* (2008) se propôs a investigar os fatores associados ao óbito hospitalar nas internações por IAM e insuficiência coronariana (InC), e se a via de internação pela Central de Internação da Secretaria Municipal de Saúde de Belo Horizonte estava associada ao óbito hospitalar após ajuste por fatores relevantes. Os dados foram obtidos das AIH e dos laudos da

última internação realizada com hipótese diagnóstica de IAM ou InC. Para identificar fatores de risco para o óbito hospitalar realizou-se uma análise multivariada.

O estudo detectou que não houve associação entre a via de acesso à internação e risco de óbito hospitalar por essas causas. Por outro lado, a análise multivariada demonstrou maior risco de óbito para pacientes com sessenta anos de idade ou mais (*odds ratio* – OR = 2,9), hipóteses diagnóstica de IAM (OR = 3,0), uso de Unidade de Terapia Intensiva (UTI) (OR = 1,6), sexo feminino (OR=1,4), especialidade cirúrgica (OR = 1,9) e hospital público (OR = 3,5). Um ponto de destaque foi, nas internações por IAM, o maior risco de morte de pacientes internados no fim de semana (OR = 1,7). Concluiu-se que, para avaliar a assistência prestada, são necessárias investigações que levem em consideração mais fatores relacionados aos hospitais, aos pacientes e ao processo da assistência, subsidiando assim propostas que garantam maior equidade e qualidade da assistência pública.

Entre os artigos utilizados na revisão, o que mais se aproxima da presente dissertação é o produzido por SOARES *et al.* (2006), devido à utilização de dados do mesmo ano e retirados do mesmo sistema de informação. Esse estudo avaliou a influência dos principais fatores prognósticos disponíveis no SIH/SUS na sobrevivência dos pacientes internados por IAM nos serviços de saúde do SUS, no estado do Rio de Janeiro, no ano de 2002. Foi realizado o relacionamento probabilístico de bases de dados das AIH de 2002 e das declarações de óbitos de 2002 e 2003. Entre os óbitos, 56,7% eram homens, porém a letalidade foi maior entre as mulheres, 27,7%, em todas as idades. Mais de 50% das internações concentraram-se na região metropolitana, enquanto as internações das demais regiões dividiram-se homoganeamente. A reinternação por IAM foi de 4,1%. A permanência hospitalar média foi de 9,5 dias. Cerca de 29,1% dos pacientes fizeram uso da UTI, com maior utilização entre homens

(65,3%) do que entre mulheres (34,7%) e em indivíduos mais idosos. Não houve efeito estatisticamente significativo em relação à utilização de UTI e prevenção do óbito. A AnC e a RVM foram realizadas em 2,8% e 1,3% dos pacientes, respectivamente, com proporção menor entre as mulheres. A letalidade em trinta dias foi de 13,6% e a letalidade em um ano foi de 21,1%, com pior evolução em mulheres nos dois períodos e em indivíduos mais idosos. Cerca de 50% dos pacientes que foram a óbito em um ano morreram em até 9 dias a partir da data de internação. Apenas 25% dos pacientes sobreviveram até 68,5 dias, o que demonstrou que a maioria dos óbitos ocorreu precocemente. O estudo concluiu que a letalidade foi maior em indivíduos mais idosos, em ambos os sexos. A proporção de casos de IAM foi maior entre o sexo masculino, porém a letalidade foi maior entre as mulheres. A utilização de UTI, em ambos os sexos, encontrou-se muito abaixo do preconizado e foi identificada uma baixa realização de AnC e RVM principalmente entre as mulheres. Além disto, o tempo médio de permanência e a mortalidade hospitalar em 30 dias puderam sugerir a prática de altas precoces e baixa qualidade de assistência, evidenciadas pelo alto número de óbitos hospitalares pouco tempo após a internação por IAM.

Tomando-se como base os artigos supracitados, pode-se observar que a utilização das AIH, fornecidas por meio do Sistema de Informações Hospitalares do SUS (SIH/SUS), é frequente para avaliar os dados relacionados às internações nas unidades de saúde. Embora criadas como forma de pagamento de serviços prestados, elas servem como fonte de informação útil, tanto para tomada de decisão quanto para o cálculo de diversas estimativas. No entanto, alguns estudos chamam a atenção para o precário preenchimento das autorizações, o que prejudica a utilização dos dados. Ainda que a confirmação diagnóstica seja alta perante os prontuários médicos, os registros das AIH ainda contam com possíveis erros de digitação e as falhas de informação.

Relativamente aos trabalhos internacionais, mesmo com a utilização de bases secundárias pode-se perceber uma diferença grande entre os cuidados aos pacientes enfartados no Brasil e em outros países. Por exemplo, um dos estudos mostra que em Madri na Espanha, 29% dos pacientes internados por IAM realizaram revascularização e que a mortalidade foi, aproximadamente, igual a 10,8% (SENDRA *et al.*, 2005). Entretanto, no Estado do Rio de Janeiro, apenas 1,3% dos pacientes internados realizaram revascularização e o percentual de óbito hospitalar foi de, aproximadamente, 28% (SOARES *et al.*, 2006).

Relativamente aos métodos empregados, um estudo utilizou as taxas de letalidade por DIC aguda e crônica e por revascularização miocárdica e angioplastia coronariana, para observar as diferenças entre as unidades prestadoras de serviços ao SUS (OLIVEIRA *et al.*, 2006). Dois artigos utilizaram o coeficiente de correlação intraclasse *Kappa* para comparar, principalmente, as informações da base de dados das AIH do SIH/SUS e os prontuários médicos (ESCOSTEGUY *et al.*, 2002; MELO *et al.*, 2004). Com o mesmo objetivo dos dois artigos anteriores, ESCOSTEGUY *et al.* (2005) fizeram uso da modelagem logística, com posterior construção de curvas *ROC*, permitindo a comparação dos modelos. Dois estudos utilizaram a análise multivariada, apresentando os fatores associados ao IAM (SENDRA *et al.*, 2005; EVANGELISTA *et al.*, 2008), e, finalmente, SOARES *et al.* (2006) empregaram o relacionamento probabilístico de base de dados para avaliar a influência dos principais fatores prognósticos na sobrevida dos pacientes internados por IAM.

## **4 MATERIAIS E MÉTODOS**

### **4.1 Base de Dados do IAM**

O presente trabalho utilizou dados que se encontram disponíveis por meio de uma colaboração entre o Programa de Engenharia Biomédica do Instituto Alberto Luiz Coimbra de Pós-graduação e Pesquisa de Engenharia da Universidade Federal do Rio de Janeiro (PEB/COPPE/UFRJ) e o Instituto de Estudos em Saúde Coletiva da mesma Universidade (IESC/UFRJ).

Primeiramente, foram selecionados, do SIH-SUS as internações com o diagnóstico principal de IAM (Classificação Internacional de Doenças 10ª Revisão: I21 ou I22), realizadas em hospitais vinculados ao SUS, no Estado do Rio de Janeiro, no ano de 2002. Os registros duplicados e de menores de dezoito anos foram excluídos. A seguir, os registros de internação restantes foram relacionados aos do Sistema de Informação sobre Mortalidade (SIM) dos anos de 2002 e 2003, utilizando-se o *software* RecLink e metodologia proposta por COELI e CAMARGO Jr. (2002).

Desta forma, o estudo tem base populacional e desenho transversal, e sua população é constituída de indivíduos adultos de ambos os sexos, com idades entre 18 e 100 anos. Assim, a partir das AIH, a base de dados original possui 4435 registros com 48 variáveis (ANEXO I).

#### **4.1.1 Seleção das variáveis**

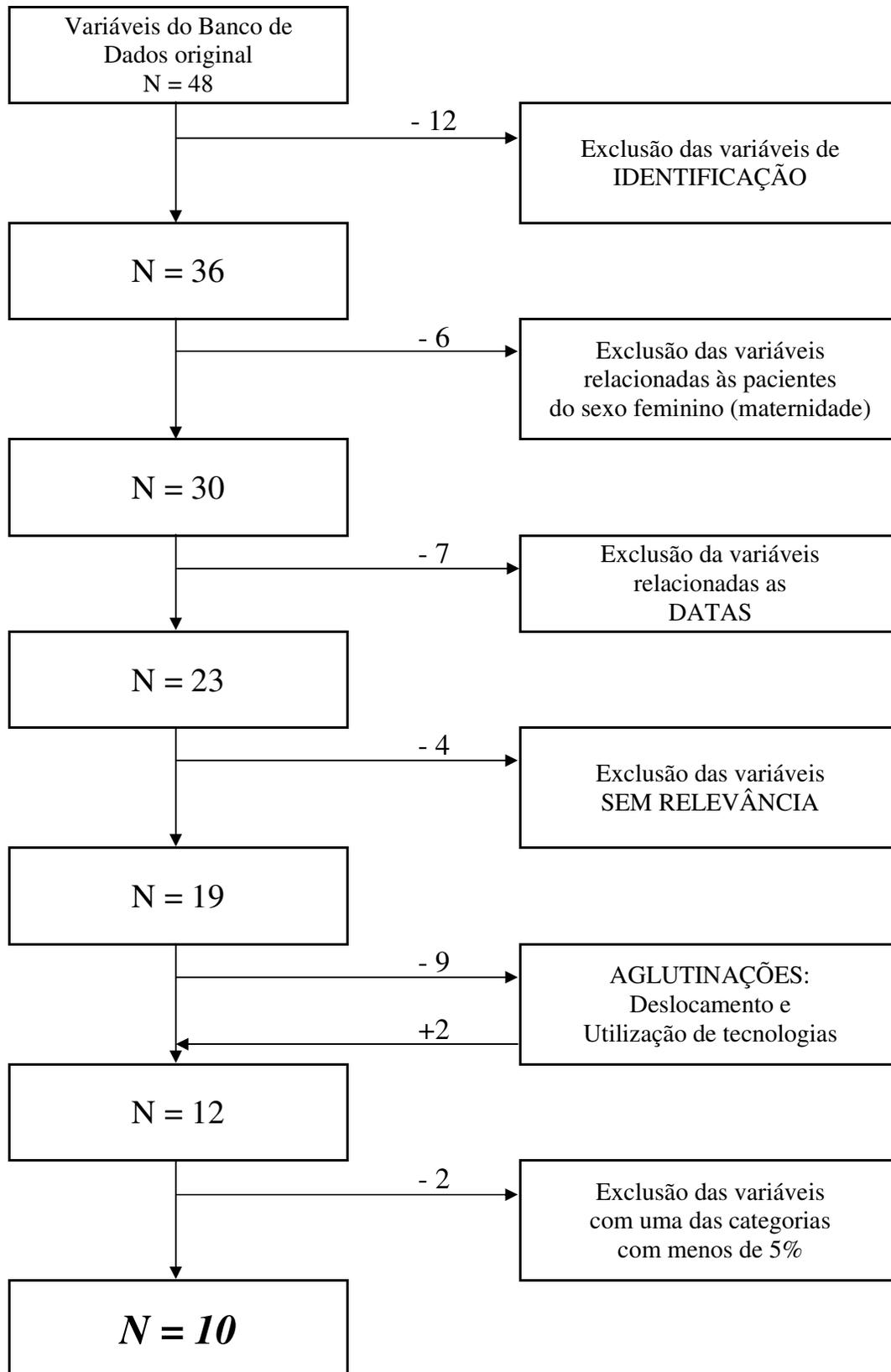
Tendo em vista que as variáveis registradas são formadas por meio de um banco de dados administrativo (AIH), foi realizada uma seleção das variáveis para a formação da base de dados do estudo, sendo que inicialmente foram excluídas as variáveis de identificação, tanto do paciente quanto dos hospitais e suas dependências (exclusão de 12 variáveis). Deste modo, as variáveis relacionadas aos pacientes não permitem

identificá-los, garantindo sua privacidade. Desta forma, o estudo teve aprovação do Comitê de Ética e Pesquisa (CEP) da Escola de Enfermagem Anna Nery da Universidade Federal do Rio de Janeiro (EEAN/UFRJ).

Como o banco de dados original possui uma seção somente relacionada às pacientes do sexo feminino, na qual encontram-se variáveis referentes à maternidade e afins (tais como: número de filhos e realização de cirurgia de laqueadura), foram excluídas seis variáveis por não serem de interesse para o estudo. Além destas, foram retiradas as variáveis relacionadas às datas (sete variáveis) e as sem relevância para o estudo, tais como: IP-descritivo e cobrança (quatro variáveis). As variáveis *local de atendimento* e *origem do paciente* foram aglutinadas formando uma variável, denominada de *Deslocamento intermunicipal* (duas variáveis tornaram-se uma). Similarmente, as variáveis *Revascularização, Angioplastia, Transplantes, Órteses e Próteses, Diagnoses e Terapias, Transfusão de sangue, Tomografia e Ressonância* foram aglutinadas, formando a variável *Utilização de tecnologias* (sete variáveis formaram uma). Após estes procedimentos, apenas as variáveis que possuíam categorias com percentual maior que 5% de variabilidade foram mantidas, resultando em 10 variáveis. A Figura 4.1 sumariza o procedimento de seleção das variáveis.

Todos os registros de internação por IAM no ano de 2002 foram armazenados e as variáveis selecionadas, conforme descrito, foram categorizadas da seguinte forma:

- 1) *Tipo de hospital*: hospital **contratado** ou hospital **próprio**. O hospital contratado refere-se às instituições que prestam serviços ao SUS (filantrópicas ou não). Por outro lado, o hospital próprio faz menção aos institutos que são administrados diretamente por órgãos públicos, tanto em âmbito Federal, Estadual ou Municipal, fazendo parte também os Hospitais Universitários, independente da esfera administrativa.



**Figura 4.1** – Fluxograma da seleção das variáveis.

- 2) *Deslocamento intermunicipal*: **sem** deslocamento ou **com** deslocamento. Utilizaram-se duas variáveis para compor a variável deslocamento intermunicipal: município de origem, local onde reside o paciente; e, município de atendimento do paciente, local onde os primeiros socorros foram prestados;
- 3) *Região de atendimento*: **metropolitana** ou **não-metropolitana**. Originalmente, o banco contava com o registro das sete regiões administrativas do Estado do Rio de Janeiro. Estas foram aglutinadas, formando apenas duas regiões. A região metropolitana refere-se ao município do Rio de Janeiro, capital do Estado, e a não-metropolitana corresponde aos demais municípios do Rio de Janeiro.
- 4) *Idade*: **15 a 55 anos**; **56 a 75 anos** ou **maior de 76 anos** (inclusive) Variável originalmente contínua, foi categorizada de acordo com a literatura científica brasileira (SOARES *et al.*, 2006), formando três categorias.
- 5) *Sexo*: **masculino** ou **feminino**. Variável binária, relacionada ao gênero do paciente.
- 6) *Utilização de tecnologias*: **alta**; **moderada** ou **baixa** complexidade. Esta variável foi construída a partir de sete variáveis contidas no banco de dados: revascularização; angioplastia; transplantes; órteses e próteses; diagnoses e terapias; transfusões de sangue; tomografia e ressonância. Considerou-se a utilização de tecnologias de alta complexidade os casos em que os pacientes passaram por uma revascularização, angioplastia e/ou por um transplante, e a utilização de tecnologias de moderada complexidade os casos em que os paciente foram submetidos à colocação de órtese ou prótese e/ou submetidos a transfusões de sangue. A utilização de tecnologias de baixa complexidade

incluiu os pacientes submetidos a diagnoses e terapias e/ou tomografias e ressonâncias.

- 7) *Tempo de internação: até 7 dias; de 8 a 30 dias* ou *mais de trinta dias*. Variável originalmente contínua, foi categorizada de acordo com a literatura científica brasileira, formando três categorias (ESCOSTEGUY, 2005).
- 8) *Utilização de CTI: não* ou *sim*. Esta variável indica a utilização do Centro de Tratamento Intensivo.
- 9) *Óbito: não; até 30 dias* ou *em 1 ano*. Variável originalmente contínua, foi categorizada de acordo com a literatura científica brasileira, formando 3 categorias (ESCOSTEGUY, 2005).
- 10) *Gastos: inferiores a R\$ 905,00* (inclusive) ou *superiores a R\$ 905,00*. O valor foi obtido a partir do Sistema de Informações Hospitalares (SIH/SUS), e representa a média de gastos por pacientes infartados no Estado do Rio de Janeiro no ano de 2002.

Para a estruturação da Matriz Indicadora, foram utilizadas 10 variáveis. O percentual de óbitos identificado na população estudada foi de aproximadamente 29%, ou seja, dos 4435 registros de pacientes internados por IAM, 1272 foram a óbito. Os demais percentuais das variáveis com suas respectivas categorias no conjunto amostral são mostrados na Tabela 4.1. Os números encontrados após cada categoria (0, 1 ou 2) são utilizados como codificação para adequá-las à aplicação das técnicas multivariadas.

**Tabela 4.1** – Frequências absolutas e relativas das variáveis.

<b>Variáveis / Categorias</b>		<b>Frequência absoluta</b>	<b>Frequência Relativa (%)</b>
<i>Tipo de hospital (HOSP)</i>	<i>Contratado (0)</i>	<b>1276</b>	<b>29</b>
	<i>Próprio (1)</i>	<b>3159</b>	<b>71</b>
<i>Deslocamento intermunicipal (DESL)</i>	<i>Sem deslocamento (0)</i>	<b>3781</b>	<b>85</b>
	<i>Com deslocamento (1)</i>	<b>654</b>	<b>15</b>
<i>Região de atendimento (REG)</i>	<i>Não-metropolitana (0)</i>	<b>1374</b>	<b>31</b>
	<i>Metropolitana (1)</i>	<b>3061</b>	<b>69</b>
<i>Idade (ID)</i>	<i>15 a 55 anos (0)</i>	<b>1591</b>	<b>36</b>
	<i>56 a 75 anos (1)</i>	<b>2256</b>	<b>51</b>
	<i>maior de 76 anos (2)</i>	<b>588</b>	<b>13</b>
<i>Sexo (SEXO)</i>	<i>Feminino (0)</i>	<b>1669</b>	<b>38</b>
	<i>Masculino (1)</i>	<b>2766</b>	<b>62</b>
<i>Utilização de tecnologias (TEC)</i>	<i>Baixa complexidade (0)</i>	<b>3025</b>	<b>68</b>
	<i>Moderada complexidade (1)</i>	<b>1219</b>	<b>27</b>
	<i>Alta complexidade (2)</i>	<b>191</b>	<b>5</b>
<i>Tempo de internação (INT)</i>	<i>Até 7 dias (0)</i>	<b>2393</b>	<b>54</b>
	<i>8 a 30 dias (1)</i>	<b>1864</b>	<b>42</b>
	<i>Mais de 30 dias (2)</i>	<b>178</b>	<b>4</b>
<i>Utilização de CTI (CTI)</i>	<i>Não (0)</i>	<b>3145</b>	<b>71</b>
	<i>Sim (1)</i>	<b>1290</b>	<b>29</b>
<i>Óbito (OB)</i>	<i>Não (0)</i>	<b>3163</b>	<b>71</b>
	<i>Até 30 dias (1)</i>	<b>722</b>	<b>16</b>
	<i>Em um ano (2)</i>	<b>550</b>	<b>13</b>
<i>Gastos (GAST)</i>	<i>Inferiores a R\$ 905,00 (0)</i>	<b>3212</b>	<b>72</b>
	<i>Superiores a R\$ 905,00 (1)</i>	<b>1223</b>	<b>28</b>

## 4.2 Aplicação da ACM

Na aplicação da ACM utilizou-se como variável desfecho o gasto hospitalar com o paciente internado por IAM (saída da ACM), sendo as demais consideradas entradas, ou seja, possivelmente associadas a este gasto. Primeiramente, ocorreu a padronização da Matriz Indicadora (4435 linhas, na qual se encontram os pacientes enfartados; 25 colunas, nas quais estão as variáveis, com suas respectivas categorias) em valores relativos (equação 3.3) obtendo-se a Matriz Padronizada  $\mathbf{G}$  (análoga à exemplificada pela Figura 3.2).

Com a utilização do algoritmo da Decomposição em Valores Singulares (equação 3.4) realizou-se a fatoração da Matriz  $\mathbf{G}$ , formando-se três novas matrizes (Matriz  $\mathbf{U}$ , que representa as coordenadas das observações no mapa de correspondência; Matriz  $\mathbf{D}$ , composta pelos autovalores da Matriz de dados e Matriz  $\mathbf{V}$ , que representa os autovetores desta mesma matriz). Posteriormente, a inércia foi ajustada de acordo com a equação 3.5, a fim de garantir que a mesma seja compatível com a variabilidade dos dados.

Desta forma, foi possível representar geometricamente os dados no Espaço Euclidiano por meio do Mapa de Correspondência, tendo sido a máxima dimensão  $s$  escolhida por meio de um critério gráfico muito utilizado na Análise Fatorial (*scree plot*) para a variação explicada pelos eixos (COSTA *at al.*, 2008). Posteriormente, a Distância de Tolerância foi calculada (equação 3.6), de modo que o quinto percentil (5%) de todas as distâncias calculadas foi considerado como o limiar de associação entre as categorias das variáveis, as quais foram agrupadas duas-a-duas. Duas categorias foram consideradas associadas quando suas distâncias encontravam-se abaixo desse limiar.

O método de simulação *bootstrap* (construção de distribuições amostrais por reamostragem, tendo por base a idéia de tratar a amostra como se fosse a população que deu origem aos dados), foi utilizado para verificar se as associações encontradas por intermédio da Distância de Tolerância eram estáveis. Assim, aproximadamente, metade dos registros foi selecionada aleatoriamente ( $n=2200$ ), repetindo-se a ACM e o cálculo da Distância de Tolerância. Este procedimento foi repetido 1000 vezes, e a quantidade de agrupamentos comuns foi registrada, sendo aqueles que apresentaram mais de 75% das ocorrências considerados estáveis.

A partir da Análise de Agrupamento (*Cluster Hierárquico*) construiu-se um dendrograma, representando o agrupamento de variáveis semelhantes. Para tal, no eixo horizontal são colocadas as amostras, e, no eixo vertical, o índice de similaridade. O número de agrupamentos foi determinado com base na análise do comportamento de fusão (seção 3.2). Ao final, os agrupamentos formados por cada metodologia foram comparados.

Para todo o tratamento estatístico do banco de dados foi utilizado o Programa estatístico R, versão 2.11.1 (R DEVELOPMENT CORE TEAM, 2004), com a utilização do pacote *ca*. Este programa é fruto de uma colaboração internacional, sendo de distribuição gratuita e proporcionando análises estatísticas abrangentes e de fácil utilização.

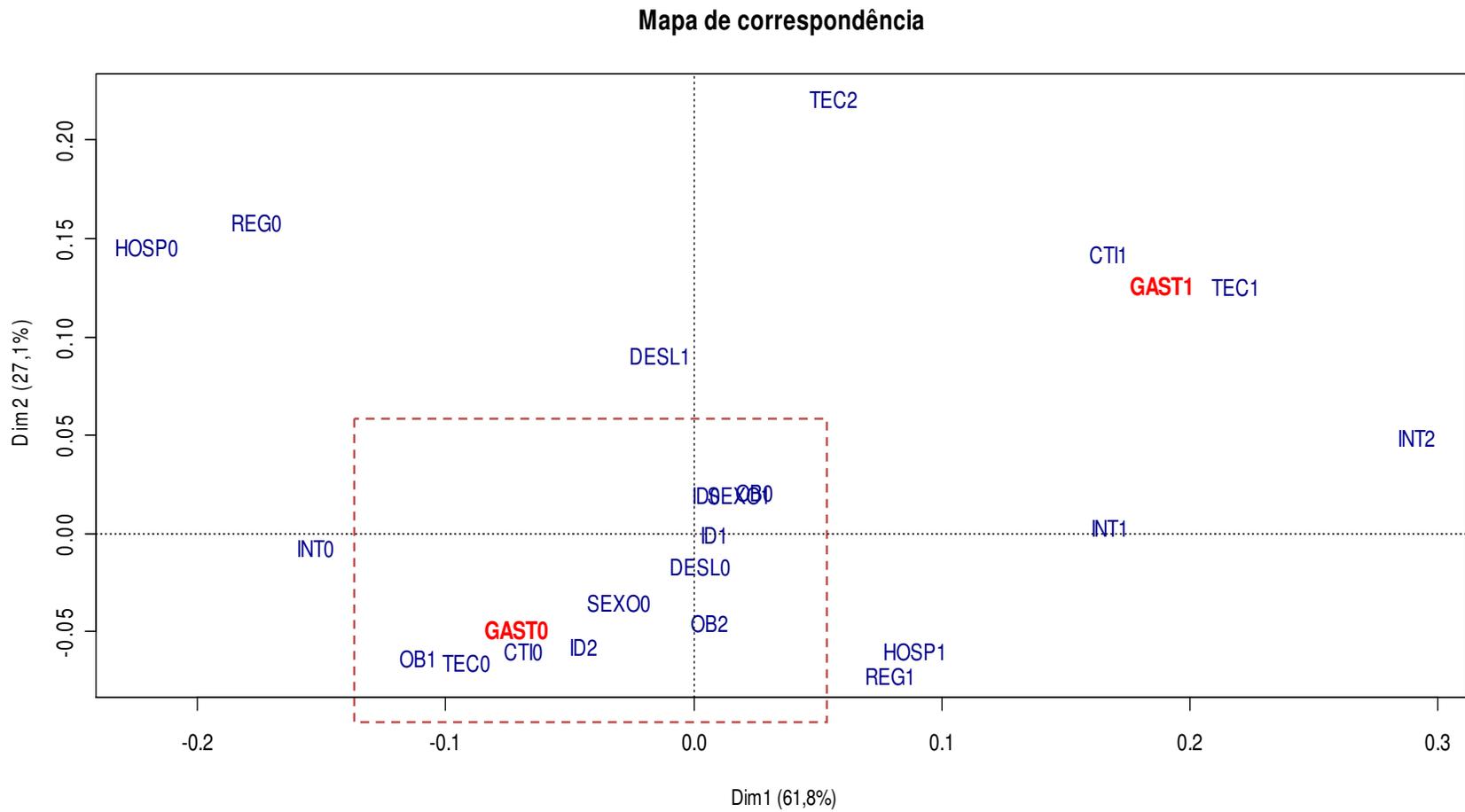
## 5 RESULTADOS

A aplicação da ACM na amostra resultou no Mapa de Correspondência da Figura 5.1. A Dimensão 1 explica, aproximadamente, 61,8% da variabilidade dos dados e está mais relacionada à estrutura do atendimento hospitalar. Por outro lado, a Dimensão 2 explica, aproximadamente, 27,1% da variabilidade e encontra-se mais associada aos processos aplicados aos pacientes enfartados. Assim, a inércia total da ACM, levando-se em consideração as duas primeiras dimensões, foi de aproximadamente 0,89.

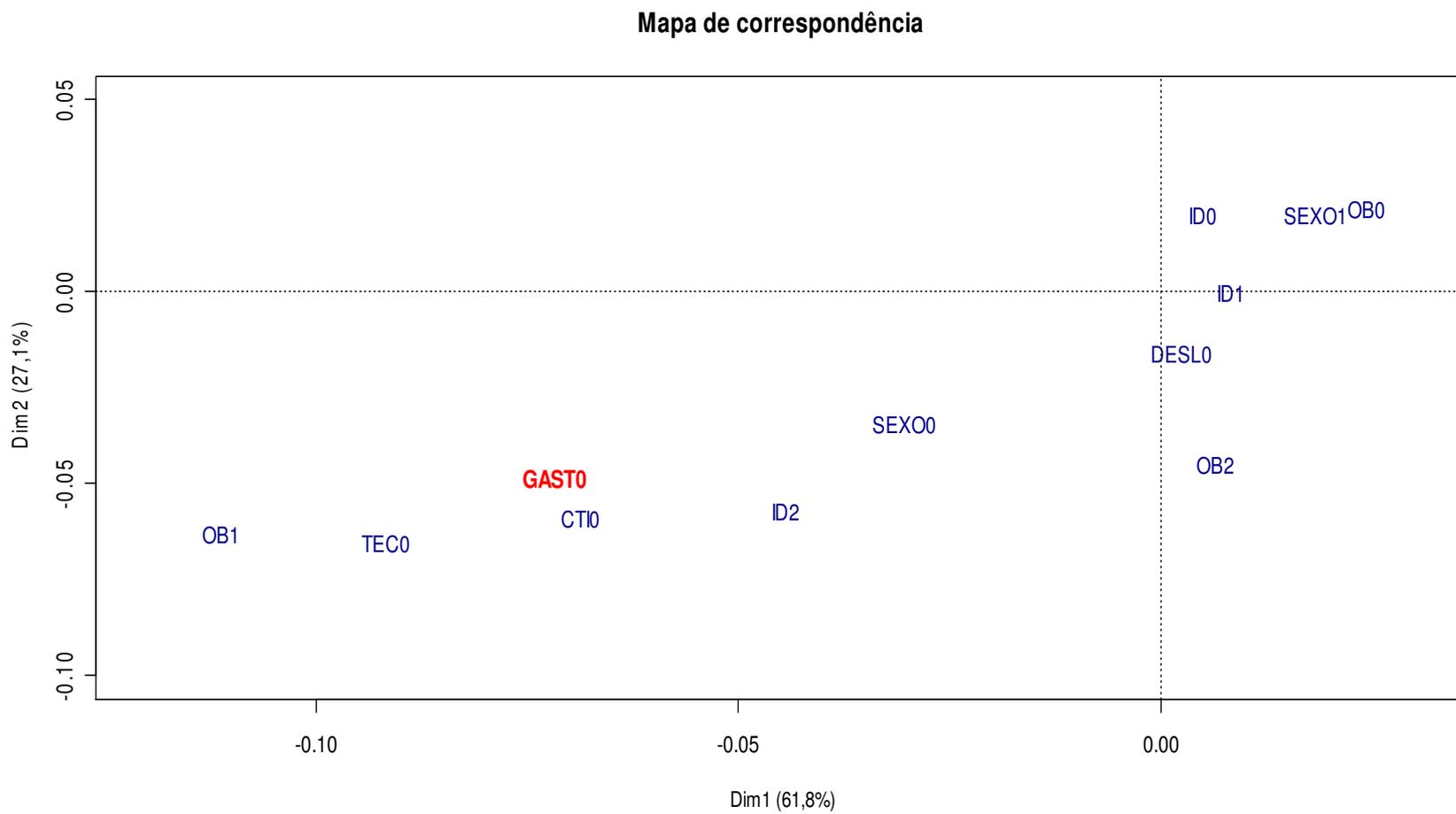
Com vistas a facilitar a visualização das categorias das variáveis, a região próxima ao centróide é mostrada ampliada na Figura 5.2. Tomando-se como referência a variável desfecho, o Mapa de Correspondência sugere associação entre o *gasto inferior a R\$ 905,00* (GAST0) e *óbito até 30 dias* (OB1), *utilização de tecnologias de baixa complexidade* (TEC0), *não utilização de CTI* (CTI0), *idade maior que 76 anos* (ID2) e *sexo feminino* (SEXO0). No entanto, em relação ao *gasto superior a R\$ 905,00* o Mapa sugere associação com: *utilização de CTI* (CTI1) e *utilização de tecnologias de moderada complexidade* (TEC1).

Considerando-se a Distância de Tolerância de 0,024 (calculada até a segunda dimensão) e os nove grupos formados dois-a-dois que incluem a variável desfecho, obteve-se sete associações entre categorias de variáveis com *gasto inferior a R\$905,00*, a saber, *sem deslocamento* (DESL0), *idade entre 56 e 75 anos* (ID2), *sexo feminino* (SEXO0), *utilização de tecnologias de baixa complexidade* (TEC0), *óbito em até 30 dias* (OB1), *óbito em um ano* (OB2) e *não utilização de CTI* (CTI0). Duas outras associações, porém com *gasto superior a R\$ 905,00* foram identificadas: *utilização de tecnologias de moderada complexidade* (TEC1) e *utilização de CTI* (CTI1).





**Figura 5.1** – Mapa de Correspondência gerado por meio da ACM. A área retangular demarcada pela linha tracejada corresponde a Figura 5.2.



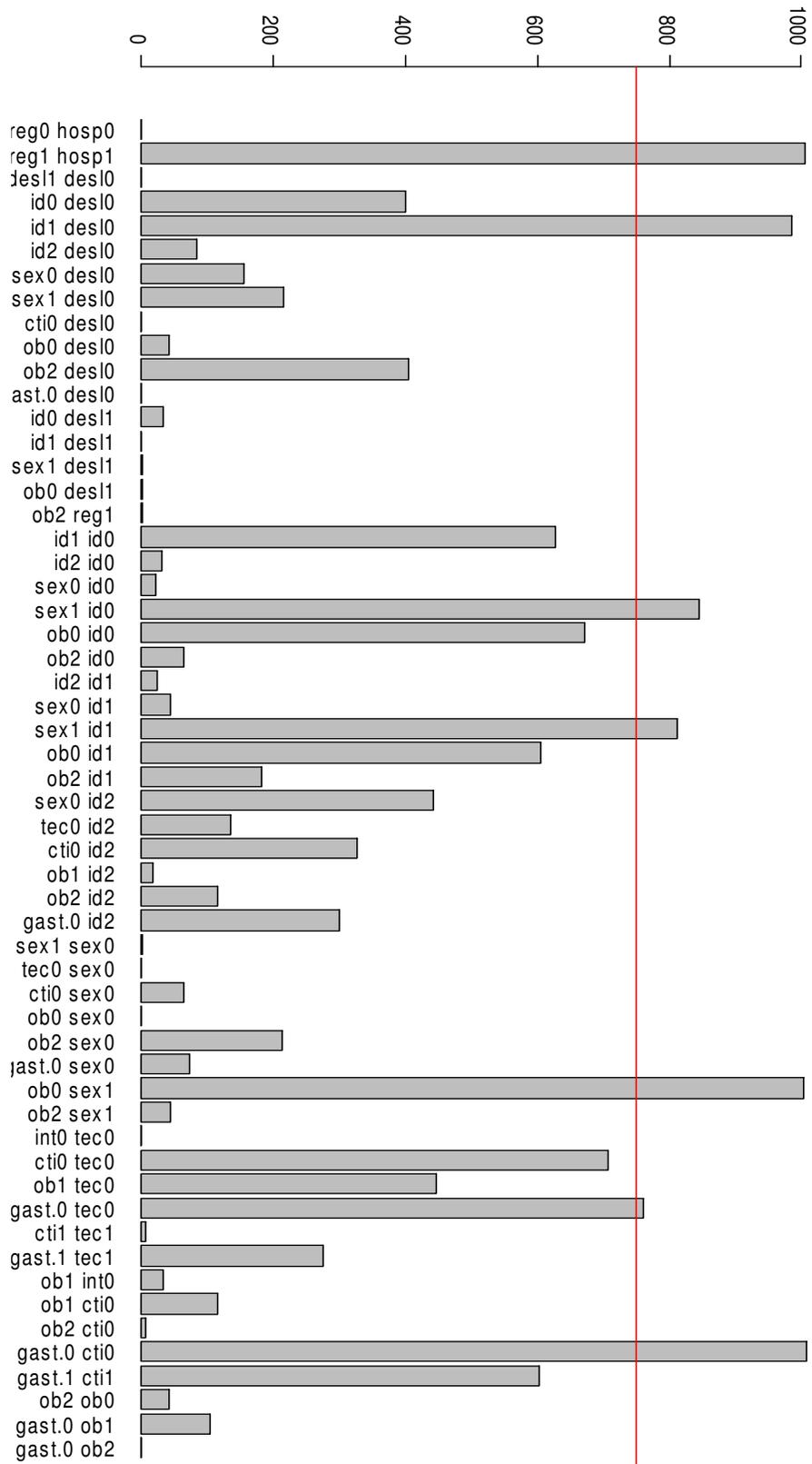
**Figura 5.2** – Ampliação do Mapa de Correspondência apresentado na Figura 5.1.

Após a validação (método *Bootstrap*), somente sete associações puderam ser consideradas estáveis, das quais duas incluíam a variável desfecho. Conforme mostra a Figura 5.3, nas associações encontradas pela ACM juntamente com a Distância de Tolerância as barras representam o número de ocorrências encontrado pela validação, enquanto a linha mostra o limite definido de 75% para estabilidade. A Tabela 5.2 mostra as categorias associadas com a variável desfecho.

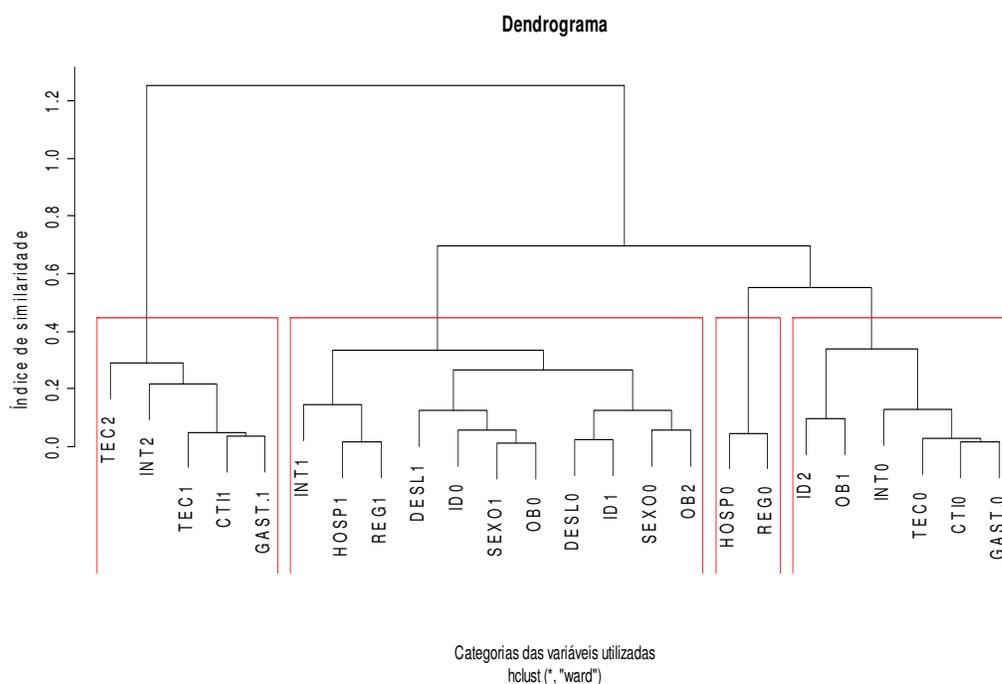
**Tabela 5.1** – Associações para a variável desfecho (*gastos hospitalares com internação*) detectadas pela Análise de Correspondência Múltipla, 4435 pacientes internados em hospitais SUS, Rio de Janeiro, 2002. Associações estáveis por uma análise *bootstrap* indicadas por um asterisco.

<b>Gastos</b>	<b>Categorias</b>
<i>Inferiores a R\$ 905,00 (GAST0)</i>	<i>Sem deslocamento (DESL0)</i>
	<i>Idade entre 56 e 75 anos (ID2)</i>
	<i>Sexo feminino (SEXO0)</i>
	<b><i>Utilização de tecnologias de baixa complexidade (TEC0)*</i></b>
	<i>Óbito em até 30 dias (OB1)</i>
	<i>Óbito em um ano (OB2)</i>
	<b><i>Não utilização de CTI (CTI0)*</i></b>
<i>Superiores a R\$ 905,00 (GAST1)</i>	<i>Utilização de tecnologias de moderada complexidade (TEC1)</i>
	<i>Utilização de CTI (CTI1)</i>

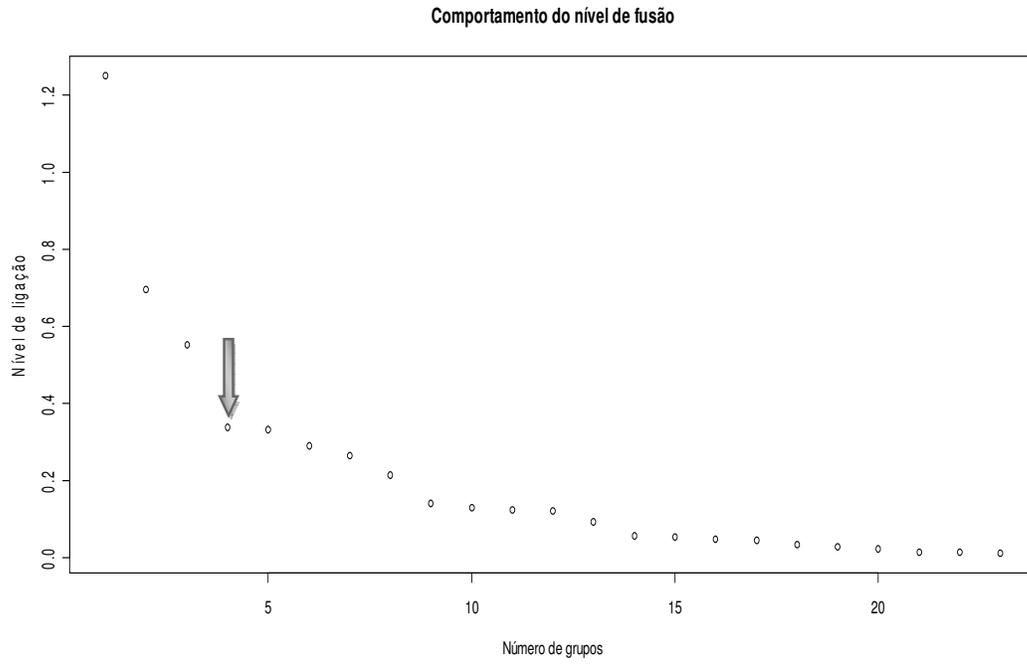
**Figura 5.3** – Gráfico de barras mostrando as associações encontradas pela ACM juntamente com a Distância de Tolerância.



A análise de agrupamentos (*Cluster Hierárquico*) possibilitou a construção do Dendrograma apresentado na Figura 5.4, o qual apresenta na sua parte inferior os agrupamentos sugeridos por esta metodologia. Associados ao *gasto inferior a R\$905,00* encontram-se: *não utilização de CTI (CTI0)*, *utilização de tecnologias de baixa complexidade (TEC0)*, *até 7 dias de internação (INT0)*, *até 30 dias para o óbito (OB1)* e *maior de 76 anos de idade (ID2)*. Quatro associações com *gasto superior a R\$ 905,00* foram sugeridas: *utilização de CTI (CTI1)*, *utilização de tecnologias de moderada complexidade (TEC1)*, *mais de 30 dias de internação (INT2)* e *utilização de tecnologias de alta complexidade (TEC2)*. O critério utilizado na determinação do número de agrupamentos foi a análise de comportamento de fusão (Figura 5.5). O gráfico do critério mostra o número de agrupamentos em função do nível de ligação dos grupos em cada estágio do procedimento e permite indicar o número final destes (vide seção 3.2).



**Figura 5.4** – Dendrograma apresentado a partir dos dados utilizados.



**Figura 5.5** – Gráfico da análise do comportamento do nível de fusão. A seta indica o número de agrupamentos utilizado para a formação do dendrograma.

## 6 DISCUSSÃO

O IAM é um evento agudo, que, como discutido, sempre requer atendimento emergencial seguido de internação. Por sua magnitude, impacto na mortalidade, letalidade, volume de internações e sensibilidade a tecnologias médicas, ele tem sido especialmente apontado como um agravo importante no desenvolvimento de padrões e indicadores para o monitoramento da assistência médica (ESCOSTEGUY *et al.*, 2002; MELO *et al.*, 2007). A proposta principal do presente estudo foi analisar o perfil das internações dos pacientes do SUS por IAM no Estado do Rio de Janeiro, verificando as associações entre as categorias das variáveis escolhidas como relevantes para sua caracterização. “*Gastos*” foi escolhida como variável desfecho por se tratar de uma variável importante sobre a utilização dos recursos relativos às doenças do aparelho circulatório, principalmente do infarto, visto que, atualmente, as doenças coronarianas têm um grande impacto financeiro sobre o sistema de saúde, constituindo cerca de 5% dos gastos com internação (OLIVEIRA *et al.*, 2006).

Uma característica importante do presente trabalho é a representatividade da amostra utilizada, que contém todos os dados das internações registradas no SUS com diagnóstico de IAM no ano de 2002, no Estado do Rio de Janeiro. Por outro lado, isso implicou em uma grande quantidade de informação, dificultando a avaliação e análise da assistência aos pacientes com IAM. Assim, o banco de dados possuía originalmente um elevado número de variáveis (48), algumas das quais, apesar de terem importância para o estudo (como *Reinternação* ou *Diagnóstico secundário*), não puderam ser aproveitadas, devido às subnotificações. As variáveis revascularização miocárdica e angioplastia coronariana representam, normalmente, procedimentos bastante utilizados em pacientes enfartados. Como ambas tinham limitada variabilidade no banco de dados, preferiu-se seu uso na formação de uma nova variável, “*utilização de tecnologias*”,

criada pela combinação destas com as variáveis: transplantes, órteses e próteses, diagnoses e terapias, transfusões de sangue, tomografia e ressonância. Assim, esta variável foi criada, permitindo uma melhor caracterização do perfil de internações. Após as exclusões e aglutinações, das 48 variáveis que formavam o banco de dados original, apenas 10 foram utilizadas para a análise final.

Os estudos com pacientes internados por IAM, baseados nos prontuários médicos ou em dados secundários (da própria AIH) em geral utilizam metodologia multivariada, tal como a modelagem logística reportado por ESCOSTEGUY *et al.* (2005). Esse estudo tinha como objetivo avaliar o desempenho do banco da AIH em relação ao prontuário médico na análise de fatores associados ao risco de óbito hospitalar por IAM. Para tal, os autores desenvolveram dois modelos, um a partir das AIH e outro baseado nos prontuários médicos, sendo o melhor ajuste obtido para o modelo derivado dos prontuários. Embora o modelo proveniente das AIH não tivesse um ajuste tão bom, não foram observadas modificações significativas entre os modelos. No entanto, os autores ressaltaram que um melhor ajuste no modelo das AIH poderia ser conseguido caso as mesmas fossem preenchidas de forma mais adequada.

Outra metodologia já utilizada para a análise do IAM a partir dos dados das AIH e dos prontuários é o coeficiente de correlação intraclass *Kappa* (ESCOSTEGUY *et al.*, 2002; MELO *et al.*, 2004). Os dois estudos tinham o objetivo de analisar a concordância entre o diagnóstico principal registrado na AIH e o registro do prontuário médico, e ambos constataram que a qualidade do registro de IAM notificado na AIH foi satisfatória.

Um aspecto especialmente importante ao lidar-se com bancos de dados desenvolvidos por razões administrativas é o *relacionamento probabilístico* realizado a partir das bases de dados existentes (SOARES *et al.*, 2006). Esse procedimento consiste

em relacionar duas bases de dados, sendo utilizado para identificar indivíduos presentes em ambas. A idéia básica é utilizar a frequência de uma característica para calcular um *score* para cada par hipotético, de forma a caracterizar sua probabilidade de correção. Em SOARES *et al.* (2006), a metodologia foi utilizada para averiguar os óbitos registrados nas AIH de 2002 e os encontrados nas declarações de óbito de 2002 e 2003. No Brasil, outros trabalhos que se utilizaram dessa metodologia ESCOSTEGUY *et al.* (2002; 2005) e MELO *et al.* (2004).

Esses trabalhos, no entanto, não se utilizaram da metodologia da ACM ou de outras metodologias de agrupamento de dados, como o *Cluster* Hierárquico. A utilização da ACM mostrou-se adequada por permitir a análise de um grande número de variáveis de forma visual e simplificada, além de também permitir a análise das associações de variáveis em termos de suas *categorias*. Isso permitiu estabelecer o perfil das interações por IAM no Estado do Estado do Rio de Janeiro, detectando e representando estruturas subjacentes do banco de dados, por meio da redução da sua dimensionalidade. Como discutido, nessa técnica as associações entre categorias de variáveis são representadas graficamente, como “mapas”, de forma diferente das metodologias multivariadas mais clássicas. Assim, o uso da ACM tem pertinência por facilitar a interpretação da estrutura do banco de dados, vista a grande quantidade de informações do mesmo.

O mapa de correspondência gerado a partir da ACM apresentou uma inércia total, considerando-se as duas primeiras dimensões, de 0,89, com a dimensão 1 explicando cerca de 61,8% da variabilidade dos dados e a dimensão 2, aproximadamente, 27,1%. A dimensão 1 pode ser interpretada com representando de forma mais direta a estrutura do atendimento hospitalar, destacando-se as variáveis *Tipo de hospital* e *Região de atendimento*. Por outro lado, a dimensão 2 pode ser

considerada como mais relacionada aos procedimentos adotados no atendimento aos pacientes enfartados, destacando-se a variável *Utilização de tecnologias*. Assim, a partir das variáveis analisadas, que não incluíam informações clínicas detalhadas sobre o paciente, mas sim concentravam-se nas características da prestação do serviço, constatou-se que a estrutura do atendimento hospitalar (*tipo de hospital; região de atendimento*) possui um maior percentual de explicação da variabilidade dos casos.

A ACM, em conjunto com a Distância de Tolerância, sugeriu 56 associações, das quais nove se referem à variável desfecho; sete associadas a *gastos inferiores a R\$ 905,00*, sendo elas: *sem deslocamento, idade entre 56 e 75 anos, sexo feminino, utilização de tecnologias de baixa complexidade, óbito em até trinta dias, óbito em um ano e não utilização de CTI*; e, duas relacionadas a *gastos superiores a R\$ 905,00: utilização de tecnologias de moderada complexidade e utilização de CTI*. Entre as sete, duas foram consideradas estáveis pelo método de validação empregado (*Bootstrap*). Estas associam o *gasto inferior a R\$ 905,00 (GAST0)* à *utilização de tecnologias de baixa complexidade (TEC0)* e à *não utilização de CTI (CTI0)*.

Como mencionado, o IAM é um evento agudo que sempre requer atendimento emergencial seguido de internação. O Mapa de correspondência, juntamente com a Distância de Tolerância, mostrou que, como esperado, os casos que utilizam tecnologias de baixa complexidade encontram-se associados com a não utilização de CTI, e que ambos associam-se a menores gastos. Em contrapartida, os maiores gastos estão mais associados à utilização de tecnologias de moderada complexidade e à utilização de unidade altamente especializada (CTI). Embora a utilização de CTIs envolva o tratamento de alta complexidade, com emprego de equipamentos e tecnologias onerosas, os casos de utilização de tecnologias de alta complexidade (categoria TEC2 na Figura 5.1), encontravam-se “isolados”, sem associação com utilização de CTI

(apesar de localizados no mesmo quadrante da categoria *utilização de tecnologias de moderada complexidade*). Este achado sugere que as facilidades disponíveis na CTI não devem estar sendo adequadamente utilizadas no atendimento destes pacientes e, conseqüentemente, em desfecho negativo, o que se refletiria nos mapas de correspondência em indicar menor associação entre esses casos e a utilização de tecnologias.

A metodologia também sugeriu associação entre os casos que utilizam tecnologias de menor complexidade e os pacientes com idade entre 56 e 75 anos. Os indivíduos nesta faixa etária já possuem circulação colateral (arteriogênese) (McARDLE, 2006), a qual abranda a intensidade do infarto, tornando o IAM menos grave, não necessitando de intervenções de alta complexidade. Por outro lado, pessoas mais novas, por não possuírem um mecanismo auxiliar que auxilie quando da oclusão total ou parcial de um vaso; e as pessoas mais idosas, debilitadas e com a musculatura cardíaca enfraquecida, são, normalmente, vítimas de infartos mais graves, de média e alta complexidade. Todavia, a associação dos casos que *utilizam tecnologias de menor complexidade* com o óbito em até 30 dias ou em um ano e com a não utilização de CTI, apesar de pouco elucidativa, pode estar captando diferenças entre diagnósticos (influenciando diretamente no tipo de tratamento, devido à real gravidade do caso) e possíveis casos de altas precoces.

A metodologia da ACM + Distância de Tolerância não apresentou diferenças significativas relativamente ao método de *Cluster Hierárquico*. Os agrupamentos representados pelo Dendrograma (Figura 5.4), elaborados a partir da análise do “nível de fusão”, que utilizou, posteriormente, quatro grupos (Figura 5.5), são semelhantes aos achados da metodologia da Distância de Tolerância. Nota-se, porém, que os grupos formados pela metodologia de agrupamentos hierárquicos englobam categorias que

aparentemente encontram-se mais distantes. Essas associações não foram detectadas pela Distância de Tolerância, a qual é mais rigorosa, formalizada e sistemática na definição de associações. Além disto, a Distância de Tolerância permite posteriormente processos de validação, como a análise de estabilidade, procedimento que outras metodologias não permitem. Desta forma, a Distância de Tolerância pode ser adotada como um índice objetivo de proximidade entre as categorias das variáveis.

O percentual de óbitos por IAM no estudo é semelhante ao percentual estimado pelo MINISTÉRIO DA SAÚDE (2011) por meio dos dados obtidos do SIM para óbitos hospitalares. Em termos gerais, o perfil de gastos segue o padrão nacional, estando os menores gastos relacionados aos casos que utilizam tecnologias de baixa complexidade e à não utilização de CTI, e vice-versa. Duas questões, porém, merecem destaque: a primeira está relacionada à utilização do CTI, que em ambos os sexos (29%), encontra-se muito abaixo do observado em países desenvolvidos, como a Espanha (SENDRA *et al.*, 2005; ESCOSTEGUY *et al.*, 2006). Essa baixa utilização pode ajudar a explicar, também, a mencionada falta de associação entre os casos de utilização de tecnologias de alta complexidade e a utilização de CTI, observada neste estudo. A segunda refere-se ao fato de que, no mapa de correspondência, a categoria *sexo masculino* (SEXO1) encontra-se próxima da *não óbito*, e que a categoria *sexo feminino* (SEXO0) está próxima das categorias *óbito em até trinta dias* (OB1) e *óbito em um ano* (OB2). Somente com a visualização do mapa, seria possível que ocorresse a sugestão da associação do gênero como óbito. No entanto, uma análise mais detalhada (Tabela 5.1), mostra que a proporção de casos de IAM foi maior entre os homens (*sexo masculino* – SEXO1), porém a letalidade foi maior entre as mulheres (*sexo feminino* – SEXO0), o que ocasiona a proximidade entre o não óbito e o sexo masculino e o óbito e o sexo feminino, relações que vem sendo observadas na literatura científica nos últimos anos

(ANDERSON e PEPINE, 2007). O mapa de correspondência indica ainda o bem conhecido fato de que hospitais públicos encontram-se, em sua maioria, na região metropolitana, e que os hospitais contratados encontram-se principalmente nas regiões não-metropolitanas. Esse modelo de atenção hospitalar acentua a centralização dos serviços nos grandes centros urbanos, com grande poder político historicamente determinado.

A ACM é uma técnica ainda não utilizada para a obtenção do perfil das internações por IAM e sua relação com o uso dos recursos financeiros. Sendo assim, algumas limitações deste estudo devem ser consideradas. A utilização de um banco de dados secundário, ou seja, o uso de informações existentes para investigar outras questões, para as quais o banco não foi originalmente concebido, pode provocar problemas, causados, por exemplo, por subnotificações, por falhas metodológicas, ou até mesmo por erros no preenchimento. No entanto, a consolidação dos dados das AIH dos pacientes enfartados a partir dos dados provenientes do SIM/SUS (COELI e CAMARGO Jr., 2002), permitiu que tais dificuldades fossem minimizadas. Além disto, as bases de dados primárias utilizadas para a formação da base secundária possuem confiabilidade satisfatória (ESCOSTEGUY *et al.*, 2002; 2005; MELO *et al.*, 2004), principalmente no que se refere à utilização desse tipo de banco de dados na avaliação da assistência ao infarto.

Adicionalmente, uma importante limitação no banco das AIH refere-se à insuficiência de dados para distinguir pacientes com níveis de gravidade distintos, impedindo aprofundamento na análise. A utilização dos registros de internações apenas do ano de 2002 e dos registros de mortalidade dos anos de 2002 e 2003 consiste em outro limite do estudo. Além disto, outro obstáculo reside no fato de que informações não investigadas podem ter participação no perfil das internações, como por exemplo:

intervalo de tempo desde o surgimento dos sintomas até a admissão hospitalar, tamanho do infarto e situação hemodinâmica no momento da avaliação inicial. Devido à falta de registros destas variáveis relacionadas aos pacientes enfartados, torna-se difícil chegar-se a conclusões mais específicas e detalhadas acerca dos possíveis desfechos clínicos.

Independentemente das limitações encontradas, o estudo possui relevância científica devido à investigação e utilização de uma metodologia ainda não utilizada para caracterizar as internações por IAM, além de contribuir para o desenvolvimento metodológico da discutida Distância de Tolerância. No âmbito social, a relevância encontra-se no fato do estudo ter gerado informações úteis quanto aos perfis de internação dos pacientes enfartados, permitindo o uso mais eficiente dos recursos relativos à saúde, com o objetivo de melhorar a oferta da assistência, e, conseqüentemente, melhorar o atendimento aos pacientes. Em termos de pesquisas futuras, seria interessante a realização de estudos contemplando a utilização de bases de dados que tenham, originalmente, em sua estrutura, registros sobre o paciente, ou seja, variáveis relacionadas aos aspectos clínicos e físicos dos indivíduos. Também seria importante a utilização destas bases por um período de tempo maior do que o utilizado, o que possibilitaria a obtenção de informações mais precisas, permitindo a utilização mais eficiente dos recursos financeiros pelos gestores.

## 7 CONCLUSÃO

O IAM é uma patologia multicausal, que atinge todos os estratos da sociedade, sendo especialmente preocupante devido à dificuldade ou falta de acesso à assistência da população do país, além da falta de medidas preventivas e de acesso à informação. Logo, é importante caracterizar as interações por esta patologia, verificando as associações entre as categorias das variáveis envolvidas, para que se obtenha uma visão do panorama geral dentro da realidade do país.

A metodologia empregada, Análise de Correspondência Múltipla juntamente com a Distância de Tolerância, indicou que os menores gastos encontram-se mais associados à não utilização do CTI e com os casos de utilização de tecnologias de baixa complexidade. Em contrapartida, os maiores gastos encontram-se mais associados à utilização dos CTI e com a utilização de tecnologias de moderada complexidade, sendo este efeito não constatado para casos de utilização de tecnologias de alta complexidade. Este último aspecto, sugere que as facilidades tecnológicas disponíveis no CTI não estariam sendo adequadamente empregadas no atendimento dos pacientes internados, não impedindo, assim, que estes fossem a óbito. Tal interpretação baseia-se em estudos relatados na literatura que constataram ser a utilização do CTI muito abaixo do observado em países desenvolvidos.

A ACM e a Distância de Tolerância mostraram-se úteis, com potencial para contribuir com o processo de tomada de decisão por parte dos gestores em saúde, no que concerne à utilização mais racional dos recursos. Embora os resultados com esta metodologia proposta não difiram substancialmente daqueles obtidos com método de *Cluster* Hierárquico, a ACM em conjunto com Distância de Tolerância se mostrou mais robusta, conforme demonstrado pela aplicação do *bootstrap*.

Além disto, o estudo ratifica a importância da utilização de técnicas exploratórias para a análise de bancos de dados utilizados em estudos epidemiológicos ou na caracterização de serviços de saúde. Assim, a partir da aplicação de técnicas como a ACM e a Distância de Tolerância, espera-se melhor entender possíveis associações entre variáveis e, deste modo, visualizar relações em grandes bancos de dados, como comum na área de saúde.

## 8 REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, R. M. V. R., INFANTOSI, A. F. C., SUASSUNA, J. H. R., COSTA, J. C. G. D., 2009, “Multiple Correspondence Analysis in predictive logistic modelling: application to a living-donor kidney transplantation data”, *Computer Methods and Programs in Biomedicine*, v. 95, n. 2 (Ago), pp. 116-128.

ANDERSON, R. D., PEPINE, C. J., 2007, “Gender differences in the treatment for Acute Myocardial Infarction: Bias or Biology?”, *Circulation*, v. 115, n. 7 (Fev), pp. 823-826.

BERTRAND, M. E., SIMOONS, M. L., FOX, K. A., WALLENTIN, L. C., HAMM, C. W., McFADDEN, E., DE FEYTER, P. J., SPECCHIA, G., RUZYLLO, W., 2002, “Management of acute coronary syndromes in patient presenting without persistent ST-segment elevation”, *European Heart Journal*, v. 23, n. 23 (Dez), pp. 1809-1840.

COELI, C. M., CAMARGO Jr., K. R., 2002, “Avaliação de diferentes estratégias de blocagem no relacionamento probabilístico de registros”, *Revista Brasileira de Epidemiologia*, v. 5, pp. 185-196.

COSTA, J. C. G. D., ALMEIDA, R. M. V. R., INFANTOSI, A. F. C., SUASSUNA, J. H. R., 2008, “A heuristic index for selecting similar categories in Multiple Correspondence Analysis applied to living donor kidney transplantation”, *Computer Methods and Programs in Biomedicine*, v. 90, n. 3, pp. 217-229.

DAVISON, A. C., HINKLEY, D. V., 1997, *Bootstrap Methods and Their Application*, 1 ed. Cambridge, Cambridge University Press.

ESCOSTEGUY, C. C., PORTELA, M. C., MEDRONHO, R. D. A., DE VASCONCELOS, M. T. L., 2002, “The Brazilian hospital information system and the acute myocardial infarction care”, *Revista de Saúde Pública*, v. 36, pp. 491-499.

ESCOSTEGUY, C. C., PORTELA, M. C., MEDRONHO, R. D. A., DE VASCONCELOS, M. T. L., 2005, “Hospital admissions forms versus medical records to asses risk of in-hospital death from acute myocardial infarction in Rio de Janeiro, Brazil”, *Cadernos de Saúde Pública*, v. 21, n. 4 (Jul-Ago), pp. 1065-1076.

EVANGELISTA, P. A., BARRETO, S. M., GUERRA, H. L., 2008, “Hospital admissions and hospital death associated to ischemic heart diseases at the National Health System (SUS)”, *Arquivos Brasileiros de Cardiologia*, v. 90, n. 2, pp. 130-138.

GREENACRE, M. J., 1992, “Correspondence Analysis in medical research”, *Statistical Methods in Medical Research*, v. 1, pp. 97-117.

GREENACRE, M. J., HASTIE, T., 1987, “The geometric interpretation of correspondence analysis”, *Journal of American Statistical Association*, v. 82, pp. 437-447.

GREENACRE, M., 2007, *Correspondence Analysis in practice*. 2 ed. Barcelona, Chapman & Hall

GREENACRE, M. J., NENADIC, O., 2007, “Correspondence Analysis in R, with two and three-dimensional graphics: the ca package”, *Journal of Statistical Software*, v. 20, n. 3.

GUYTON, A. C., HALL, J. E., 2006, *Tratado de Fisiologia Médica*. 11 ed. Rio de Janeiro, Elsevier.

HAIR, J. R., HAIR, J. F., ANDERSON, R. E., TATHAM, R., BLACK, W., 2009, *Análise multivariada de dados*. 6 ed. São Paulo, Bookman.

LEBART, L., 2007, “Which Bootstrap for Principal Axes Method?”, In: *Selected Contributions in Data Analysis and Classification, Studies in Classification, Data Analysis and Knowledge Organisation Series*, Springer, pp. 581-588.

LE ROUX, B., ROUANET, H., 2004, *Geometric Data Analysis*, 1 ed. Dordrecht, The Netherlands, Kluwer Academic Publishers.

MALTA, D.C., CEZÁRIO, A. C., MOURA, L., MORAIS NETO, O. L., SILVA Jr, J. B., 2006, “A construção da vigilância e prevenção das doenças crônicas não transmissíveis no contexto do Sistema Único de Saúde”, *Epidemiologia e Serviços de Saúde*, v. 15, pp. 47-65.

McARDLE, W. D., KATCH, F. I., KATCH, V. L., 2006, *Fisiologia do exercício: energia, nutrição e desempenho humano*, 2 ed. São Paulo, Manole.

MELO, E. C. P., TRAVASSOS, C. M. R., CARVALHO, M. S., 2004, “Quality of data on myocardial infarction deaths, Brazil”, *Revista de Saúde Pública*, v. 38, n. 3, pp. 385-391.

MELO, E. C. P., TRAVASSOS, C. M. R., CARVALHO, M. S., 2007, “Infarto agudo do miocárdio no Município do Rio de Janeiro: qualidade dos dados, sobrevida e distribuição espacial”, *Epidemiologia e Serviços de Saúde*, v. 16, pp. 121-123.

MINGOTI, S. A., 2005, *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*, 1 ed. Belo Horizonte, Editora UFMG.

MINISTÉRIO DA SAÚDE, 2004, *Saúde Brasil 2004: uma análise da situação da saúde*. Secretaria de Vigilância em Saúde, Brasília.

MINISTÉRIO DA SAÚDE, 2005, *A vigilância, o controle e a prevenção das doenças crônicas não-transmissíveis: DCNT no contexto do Sistema Único de Saúde brasileiro*. Organização Pan-Americana da Saúde, Brasília.

MINISTÉRIO DA SAÚDE, 2011, *Sistema Informações sobre Saúde*, Disponível em: <<http://www2.datasus.gov.br/DATASUS/index.php?area=02>>. Acesso em 22 mar. 2011, 10:23.

MOITA NETO, J. M., 2004, *Estatística Multivariada: uma visão didática-metodológica*, Disponível em: <[http://criticanarede.com/cien\\_estatistica.html](http://criticanarede.com/cien_estatistica.html)>. Acesso em 6 ago. 2010, 15:55.

MOITA NETO, J. M., 2006, *Estatística Multivariada*, 1 ed. Piauí, Fundação de Amparo à Pesquisa do Estado do Piauí.

MSDONLINE – Manual Merck saúde para família, 2011, *Irrigação sanguínea do coração*. Disponível em: <[http://mmspf.msdonline.com.br/pacientes/manual\\_merck/secao\\_03/cap\\_014.html](http://mmspf.msdonline.com.br/pacientes/manual_merck/secao_03/cap_014.html)>. Acesso em: 27 out. 2011, 16:58.

MYOCARDIAL INFARCTION REDEFINED, 2000, “Myocardial infarction redefined: a consensus document of the joint European society of cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction”, *European Heart Journal*, v. 21, n. 18, pp. 1502-1513.

NEWBY, L. K., ALPERT, J. S., OHMAN, E. M., THYGESEN, K., CALIFF, R. M., 2002, “Changing the diagnosis of acute myocardial infarction: implications for practice and clinical investigations”, *American Heart Journal*, v. 144, n. 6 (Dez), pp. 957-980.

OLIVEIRA, G. M. M. D., KLEIN, C. H., SOUZA E SILVA, N. A. D., GODOY, P. H., FONSECA, T. M. P., 2006, “Ischemic heart disease lethality in the State of Rio de Janeiro between 1999 and 2003”, *Arquivos Brasileiros de Cardiologia*, v. 86, pp. 131-137.

ORGANIZAÇÃO MUNDIAL DE SAÚDE (OMS), 2011, *Cardiovascular Diseases (CVDs)*, Fact Sheet n°. 317. Disponível em: <<http://www.who.int/media/centre/factsheets/fs317/en/index.html>>. Acesso em: 01 fev. 2011, 9:45

RAMIARINA, R., ALMEIDA, R. M. V. R., PEREIRA, W., 2006, “Determinação e predição do custo de internação e sua relação com as características dos pacientes”. In: *Anais do Congresso Brasileiro de Engenharia Biomédica*, São Pedro, São Paulo.

RIBEIRO, R. A., MELLO, R. G. B., MELCHIOR, R., DILL, J. C., HOHMANN, C. B., LUCCHESI, A. M., STEIN, R., RIBEIRO, J. P., POLANCZYK, C. A., 2005, “Annual cost of ischemic heart disease in Brazil. Public and private perspective”, *Arquivos Brasileiros de Cardiologia*, v. 85, n. 1 (Jul), pp. 3-8.

ROUQUAYROL, M. Z., ALMEIDA FILHO, N., 2003, *Epidemiologia e Saúde*, 6 ed. Rio de Janeiro, Editora Medsi.

R Development Core Team, R: *A language for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2007, Disponível em <<http://www.R-project.org>>.

SCHRAMM, J. M. A., OLIVEIRA, A. F., LEITE, I. C., VALENTE, J. G., GADELHA, A. M. J., PORTELA, M. C., CAMPOS, M. R., 2004, “Transição epidemiológica e o estudo da carga de doença no Brasil”, *Ciência e Saúde Coletiva*, v. 9, n. 4, pp. 897-908.

SENDRA, J. M., SANTAMERA, A. S., IÑIGOC, J., REGIDORA, E., 2005, “Factores asociados a la mortalidad intrahospitalaria del infarto de miocardio.

Resultados de un estudio observacional”, *Medicina Clinica (Barc)*, v. 125, n. 17, pp. 641-646.

SOARES, C. C. P., PINHEIRO, R. S., ESCOSTEGUY, C. C., 2006, “Survival after acute myocardial infarction in patients admitted in the Brazilian Health Unified System hospitals” *Cadernos de Saúde Coletiva*, v. 14, pp. 411-416.

TIMERMAN, A., FEITOSA, G. A., 2003, *Síndromes coronarianas agudas*, 1 ed. Rio de Janeiro, Atheneu.

WERF, F. V., ARDISSINO, D., BETRIU, A., COKKINOS, D. V., FALK, E., FOX, K. A., JULIAN, D., LENGYEL, M., NEUMANN, F. J., RUZYLLO, W., THYGENSEN, C., UNDERWOOD, S. R., VAHANIAN, A., VERHEUGT, F. W., WIJNS, W., 2003, “Management of acute myocardial infarction in patients presenting with ST-segment elevation. The task force on the management of acute myocardial infarction of the European society of cardiology”, *European Heart Journal*, v. 24, n. 1, pp. 28-66.

WIKICOMMONS – WIKIMEDIA.ORG, 2011, *Diagram of the human heart*. Disponível em: <[http://commons.wikimwdia.org/wiki/File:Diagram\\_of\\_the\\_human\\_heart\\_\(cropped\)\\_pt.svg](http://commons.wikimwdia.org/wiki/File:Diagram_of_the_human_heart_(cropped)_pt.svg)>. Acesso em: 10 mar. 2011, 14:37.

## 9 ANEXO I

Variáveis	Processos
Número do paciente Naturalidade Número da AIH Número do município Número do prontuário CGC Nome do hospital Número do hospital Complemento UF CEP Protocolo AIH	<i>Exclusão das variáveis de IDENTIFICAÇÃO</i>
Nascimento do filho Sexo do filho Bairro de atendimento CEP atendimento Município de atendimento Código do estabelecimento	<i>Exclusão das variáveis relacionadas à MATERNIDADE e AFINS</i>
Data de internação Data de saída Data de nascimento Data de óbito Data de nascimento do filho Data de internação (maternidade) Data de saída (maternidade)	<i>Exclusão das variáveis relacionadas as DATAS</i>
IP-descritivo Descrição do DP Descrição do DS Cobrança	<i>Exclusão das variáveis SEM RELEVÂNCIA</i>

Atendimento  
Origem

Formação da variável DESLOCAMENTO

---

Revascularização  
Angioplastia  
Diagnoses e terapias  
Órteses e próteses  
Transfusão de sangue  
Ressonância  
Transplantes

---

Formação da variável UTILIZAÇÃO DE  
TECNOLOGIAS

Reinternação  
Diagnóstico secundário

---

Exclusão das variáveis com uma das  
categorias com menos de 5%

---