

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
CENTRO DE CIÊNCIAS BIOLÓGICAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA EM SAÚDE

LUCIANA SCHLEDER GONÇALVES KOBUS

APLICAÇÃO DA DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS
PARA IDENTIFICAÇÃO DE USUÁRIOS COM DOENÇAS CARDIOVASCULARES
ELEGÍVEIS PARA PROGRAMAS DE GERENCIAMENTO DE CASO

Dissertação apresentada ao Programa de Pós-graduação em Tecnologia em Saúde da PUCPR como requisito parcial para obtenção do título de Mestre em Tecnologia em Saúde.

Linha de Pesquisa: Sistemas de Informação e Apoio à Decisão em Saúde

Orientadora: Dr^a. Sandra Honorato da Silva

Co-orientador: Dr. João da Silva Dias

Curitiba

2006

LUCIANA SCHLEDER GONÇALVES KOBUS

APLICAÇÃO DA DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS
PARA IDENTIFICAÇÃO DE USUÁRIOS COM DOENÇAS CARDIOVASCULARES
ELEGÍVEIS PARA PROGRAMAS DE GERENCIAMENTO DE CASO

Dissertação apresentada ao Programa de Pós-graduação em Tecnologia em Saúde da PUCPR como requisito parcial para obtenção do título de Mestre em Tecnologia em Saúde.

Linha de Pesquisa: Sistemas de Informação e Apoio à Decisão em Saúde

Orientadora: Dr^a. Sandra Honorato da Silva

Co-orientador: Dr. João da Silva Dias

Curitiba

2006

K75a
2006

Kobus, Luciana Schleder Gonçalves

Aplicação da descoberta de conhecimento em bases de dados para
identificação de usuários com doenças cardiovasculares elegíveis para
programas de gerenciamento de caso / Luciana Schleder Gonçalves
Kobus ;
orientadora, Sandra Honorato da Silva ; co-orientador, João da Silva
Dias.
– 2006.
145 f. : il. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná,
Curitiba, 2006

Inclui bibliografia

1. Sistemas de recuperação de informação – Medicina. 2. Saúde –
Administração. 3. Sistema cardiovascular – Doenças – Banco de dados.
I. Silva, Sandra Honorato da. II. Dias, João da Silva. III. Pontifícia
Universidade
Católica do Paraná. Programa de Pós-Graduação em Tecnologia em
Saúde.
IV. Título.

CDD 21. ed. – 610.285



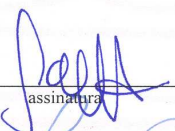
Pontifícia Universidade Católica do Paraná
Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Tecnologia em Saúde

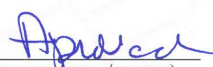
ATA DA SESSÃO PÚBLICA DE DEFESA DE DISSERTAÇÃO DE MESTRADO
DO PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA EM SAÚDE
DA PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ

DEFESA DE DISSERTAÇÃO Nº 029

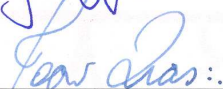
Aos 27 dias do mês de julho de 2006 realizou-se a sessão pública de defesa da dissertação “**Identificação de Usuários com Doenças Cardiovasculares Elegíveis para Programas de Gerenciamento de Caso – Aplicação da Descoberta de Conhecimento em Bases de Dados a uma Organização de Saúde do Tipo Auto-Gestão**”, apresentada por **Luciana Schleder Gonçalves Kobus** como requisito parcial para a obtenção do título de Mestre em Tecnologia em Saúde, – Área de Concentração – Informática em Saúde, perante uma Banca Examinadora composta pelos seguintes membros:


Profª. Drª. Sandra Honorato da Silva,
PUCPR (Presidente)


assinatura


parecer (aprov/ reprov.)

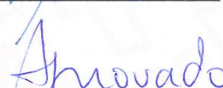
Prof. Dr. João da Silva Dias,
PUCPR (Co-orientador)


João da Silva Dias

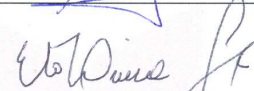

Aprovado

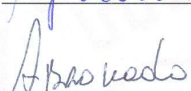
Prof. Dr. Edson Emílio Scalabrin,
(PUCPR)


Edson Emílio Scalabrin


Aprovado

Prof. Dr. Emilton Lima Junior,
(PUCPR)


Emilton Lima Junior


Aprovado

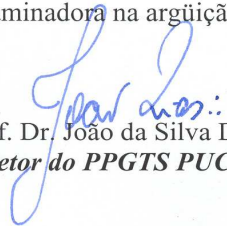
Profª. Drª. Débora Ribeiro Carvalho,
(UTP)


Débora Ribeiro Carvalho


Aprovado

Conforme as normas regimentais do PPGTS e da PUCPR, o trabalho apresentado foi considerado aprovado (aprovado/reprovado), segundo avaliação da maioria dos membros desta Banca Examinadora. Este resultado está condicionado ao cumprimento integral das solicitações da Banca Examinadora registradas no Livro de Defesas do Programa. Deve cumprir em 90 dias todas as adequações solicitadas pela Banca Examinadora na arguição e nos volumes da Dissertação.

Prof. Dr. João da Silva Dias,
Diretor do PPGTS PUCPR





DEDICATÓRIAS

Aos meus pais, Orlando Gonçalves e Lillian Schleder Gonçalves, por terem proporcionado mais do que a educação formal necessária para que eu chegasse até aqui;

Ao meu filho Alexandre;

Ao Christian.

AGRADECIMENTOS

À professora Sandra Honorato da Silva, pela orientação, pelas ricas discussões, pelo exemplo, pela oportunidade de crescimento profissional e pessoal;

Ao professor Edson Scalabrin, pela disponibilidade do seu tempo para ricas discussões, pelo interesse, pelas relevantes contribuições para o rumo deste trabalho;

À professora Déborah Carvalho, pelo seu papel fundamental e apoio imprescindível para o desenvolvimento deste trabalho, pela disponibilidade incondicional, pelas horas disponibilizadas para orientação técnica, pelo constante apoio, pelo seu exemplo;

Ao professor Emílton Lima Junior, pelo incentivo desde a idéia primária até a fase final deste trabalho;

Ao colega Altair Santos pelas discussões e apoio técnicos;

À PUCPR e aos funcionários do PPGTS, em especial à Mariele e Erlí;

AO PPGTS, na pessoa do professor João da Silva Dias, pelo curso, pelas discussões elevadas, e pela bolsa de estudos, sem a qual não poderia ter finalizado este projeto;

Aos meus familiares, que estiveram sempre ao meu lado, dando apoio emocional, técnico, logístico - meus grandes incentivadores em todos os sentidos;

À Maria Clodete Azzolin, por ter ajudado a organizar meus pensamentos, contribuindo para a finalização deste projeto;

Ao Christian, marido, companheiro, amigo, principal incentivador e colaborador deste e de outros projetos. Sem o seu amor, carinho, paciência ..., o desenvolvimento deste projeto não teria o mesmo êxito.

RESUMO

O modelo assistencial predominante nos sistemas de saúde público e suplementar brasileiros é fortemente baseado em práticas curativistas. Este fato gera alto grau de complexidade dos procedimentos, alto custo da assistência à saúde, falhas no atendimento às necessidades de promoção de saúde e prevenção de doenças de usuários, e dificuldades no seu acesso a serviços de saúde. Estes aspectos conflitantes desse modelo interferem no gerenciamento de organizações de saúde. Frente a essa realidade, emergem novas tendências de reorganização da assistência à saúde. A informatização na área da saúde deve ser considerada como um processo que auxilia tanto no desenvolvimento de estratégias voltadas à gestão organizacional, como ao gerenciamento pró-ativo da assistência, contribuindo para que ações de promoção à saúde e prevenção de doenças possam ser implementadas e colaborem para a reversão de uma assistência excessivamente curativista. Esta pesquisa retrospectiva, exploratória-descritiva, de abordagem quantitativa, teve como objetivo propor um modelo para identificação de padrões em uma base de dados de uma organização do tipo auto-gestão, que possibilitem a indicação de usuários com doenças cardiovasculares, elegíveis para o ingresso em programas de gerenciamento de casos. A metodologia adotada foi a descoberta de conhecimentos em bases de dados, que incluiu as etapas de pré-processamento, mineração de dados e pós-processamento; e a técnica de descoberta de associações. Os resultados evidenciaram que usuários incluídos na base de dados da pesquisa apresentaram perfil de elevada morbidade relacionada a doenças cardiovasculares. Os usuários a serem incluídos num programa de gerenciamento de casos de doenças cardiovasculares, são aqueles do sexo masculino, eventos que se relacionam com aqueles de alta complexidade e custo, como cateterismo cardíaco; angioplastia de vaso único; implante de stent; cintilografia do miocárdio; valores referenciais de cateterismo cardíaco; e, valores referenciais de revascularização do miocárdio que se submeteram a procedimentos indicativos de diabetes, e que são. Evidenciou-se a importância de registros de saúde fidedignos e da constante atualização de bases de dados, de modo a tornar possível a prática assistencial baseada em evidências, bem como pesquisas com resultados confiáveis. A metodologia proporcionou alcance dos objetivos propostos, e de uma maneira geral, regras de associação consideradas como relevantes por especialistas apontaram para o desenvolvimento de estratégias preventivas, embora tenha sido evidenciado que usuários buscam assistência numa fase avançada de sua doença coronariana. Essas estratégias podem ser desenvolvidas a partir da estratificação dos riscos da população para estas doenças, e programas estabelecidos de acordo com as necessidades de cada subgrupo. As ferramentas de tecnologia da informação apresentam grande potencial de contribuição para um melhor gerenciamento das organizações e de práticas de saúde, pelo apoio conferido pelo uso de bases de dados necessárias à tomada de decisão clínica e administrativa. Sistemas de informação de organizações de saúde devem ser desenvolvidos a partir da filosofia do conhecimento compartilhado na resolução de problemas, e sob a ótica multiprofissional.

Palavras Chaves: descoberta de conhecimento em bases de dados, gerenciamento de casos, organização de saúde, doenças cardiovasculares.

ABSTRACT

The public and the supplementary Brazilian health system model is strongly based on a health practice focused on curing. This fact leads to a high degree of complexity of procedures, to high costs of health care, to failure of the attendance of the real health clients needs of promotion of health and prevention of diseases, and difficulties on their access to health services. These conflictive aspects of that model interfere on the management of health organizations. In face of this evidence, new tendencies of reorganizing health care arise. The informatics in health area must be considered as a process that contributes to the development of strategies of organizational management, as well as to the proactive health management. It also contributes to the implementation of health promotion and disease prevention actions, which collaborate for the reversion of a health practice excessively focused on curing. This retrospective, descriptive and exploratory research uses a quantitative approach, and aims to propose a model to identify patterns on a health organization database, that might indicate the clients with cardiovascular diseases eligible to case management programs. The methodology adopted was the knowledge discovery in databases, that included the pre-processing, data-mining and post-processing phases, with the technique of discovery of association rules. The results showed that clients registered on the research database had a profile of high morbidity related to cardiovascular diseases. The clients to be enrolled in the case management program to cardiovascular diseases are those who were submitted to procedures that indicate diabetes, and those of male sex, such events related to the ones of high complexity and high cost, such as heart catheterization; angioplasty of single vase; stent implantation; myocardial cintilography; referential values of heart catheterization; and, referential values of myocardial revascularization. The importance of the reliability of health registration and constant database actualization was evidenced in the research, in order to become possible the health practice based on evidences, as well as the occurrence of researches with trusty results. The methodology provided the achievement of the proposed objectives, and the rules considered as relevant by the specialists, in general, pointed out to the development of prevention strategies, although the results had showed that the clients searched for health care when they were already in an advanced phase of their coronary decease. Those strategies can be developed by the population risk stratification for those diseases, and the programs can be developed according to the needs of each group. The informatics technology tools show a great potential to contribute to a better management of health organizations and their practices, by the support conferred by databases which are necessary to the clinical and the administrative decision making process. Health organizations' information systems must be developed considering the philosophy of sharing knowledge, in order to the solution of the problems according to the multiprofessional point of view.

Key Words: knowledge discovery in databases, case management, health organizations, cardiovascular diseases.

LISTA DE FIGURAS

Figura 1 - Proporção de usuários em risco para alto custo e o respectivo consumo de recursos da organização pagadora de serviços de saúde	36
Figura 2 - Principais Componentes do KDD e DM e seus relacionamentos	45
Figura 3 - Representação esquemática das tarefas de Mineração de Dados	53
Figura 4 - Etapas Metodológicas	66
Figura 5 - Resultados alcançados com a presente pesquisa.	124
Figura 6. Representação esquemática das considerações finais.	133

LISTA DE GRÁFICOS

Gráfico 1 – Distribuição de usuários por níveis progressivos de risco e custo total. Health Partners, Minnesota-USA, 2001	38
Gráfico 2 - Relação entre faixa etária, complexidade e quantidade de procedimentos. Instituto Curitiba de Saúde, 2002-2004	89
Gráfico 3 - Relação entre complexidade dos procedimentos, sua utilização pelos usuários e o custo a eles relacionados. Instituto Curitiba de Saúde, 2002- 2004	90
Gráfico 4 - Relação entre o custo, a complexidade dos procedimentos, e a faixa etária. Instituto Curitiba de Saúde, 2002-2004	91
Gráfico 5 - Relação entre o custo dos procedimentos e a faixa etária. Instituto Curitiba de Saúde, 2001-2005	93

LISTA DE QUADROS

Quadro 1 - Exemplos dos modelos de aprendizagem Indutivo e Dedutivo	41
Quadro 2 - Demonstrativo das descrições de indicadores de saúde por área temática	43
Quadro 3 - Descritivo das Técnicas de Mineração de Dados, respectivas características e ferramentas	55
Quadro 4 - Problemas e Técnicas de DM e a tarefa de KDD definida para a sua resolução	67
Quadro 5 - Códigos do CID-10 utilizados para seleção de usuários para a pesquisa	69
Quadro 6 - Descritivo da seleção dos casos, número de registros de atendimento relacionado aos usuários, Instituto Curitiba de Saúde, Curitiba, 2001-2005	79
Quadro 7 – Demonstrativo das variáveis contidas na base de dados inicial e sua disposição no arquivo fornecido pelo administrador da base de dados do ICS	80
Quadro 8 – Demonstrativo das variáveis definidas por junção orientada	81
Quadro 9 - Descritivo da classificação dos procedimentos em ALTO, MÉDIO e BAIXO custos	81
Quadro 10 - Demonstrativo da formatação apresentando o histórico de procedimentos de cada usuário em linhas distintas	82
Quadro 11 - Demonstrativo das associações entre evento e custo	83
Quadro 12 - Demonstrativo das relações dos procedimentos com os seus respectivos códigos do CID-10	85
Quadro 13 – Demonstrativo dos dados que relacionam os procedimentos e o número de suas ocorrências na base de dados da pesquisa	86
Quadro 14 – Demonstrativo dos dados que relacionam usuários que apresentavam em seus registros pelo menos um procedimentos com código que iniciasse com o dígito 40	88

Quadro 15 – Demonstrativo das regras geradas pelo algoritmo Apriori quando aplicado ao arquivo que relacionava os usuários aos seus procedimentos	96
Quadro 16 - Demonstrativo das regras geradas pelo algoritmo Apriori quando aplicado às variáveis provenientes dos casos que continham pelo menos um procedimento que iniciava com o dígito 40	96
Quadro 17 - Demonstrativo de exemplos de regras geradas pelo Apriori	98
Quadro 18 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Cateterismo Cardíaco	100
Quadro 19 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Angioplastia de Vaso Único	101
Quadro 20 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Implante de Stent	103
Quadro 21 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Cintilografia do Miocárdio	104
Quadro 22 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Valores Referenciais de Cateterismo Cardíaco	106
Quadro 23 - Demonstrativo das regras relevantes para o evento significativo Valores Referenciais de Revascularização do Miocárdio	106
Quadro 24 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Consulta de Emergência	107
Quadro 25 - Demonstrativo das regras relevantes para o evento significativo Consulta de Emergência após análise subjetiva e sob a perspectiva da confiança	108
Quadro 26 - Demonstrativo das regras relevantes do evento significativo Valores Referenciais Revascularização do Miocárdio após análise subjetiva e sob a perspectiva da confiança	113
Quadro 27 - Demonstrativo da regra relevante para o evento significativo Valores Referenciais Cateterismo Cardíaco após análise subjetiva e sob a perspectiva da confiança	115

Quadro 28 - Demonstrativo da regra relevante para o evento significativo 'Cintilografia do Miocárdio' após análise subjetiva e sob a perspectiva da confiança	178
Quadro 29 - Demonstrativo da regra relevante para o evento significativo Implante de Stent após análise subjetiva e sob a perspectiva da confiança	119
Quadro 30 - Demonstrativo da regra relevante para o evento significativo Angioplastia Vaso Único após análise subjetiva e sob a perspectiva da confiança	121
Quadro 31 - Demonstrativo das regras relevantes para o evento significativo Cateterismo Cardíaco após análise subjetiva e sob a perspectiva da confiança	123

LISTA DE TABELAS

Tabela 1 – Resultados da análise subjetiva e sob a perspectiva da confiança para o evento significativo ‘Consulta de Emergência’	108
Tabela 2 – Resultados da análise subjetiva e sob a perspectiva da confiança do evento significativo ‘Valores Referenciais Revascularização do Miocárdio	112
Tabela 3 – Resultados da análise subjetiva sob a perspectiva da confiança do evento significativo ‘Valores Referenciais Cateterismo Cardíaco’	115
Tabela 4 – Resultados da análise subjetiva e sob a perspectiva da confiança do evento significativo ‘Cintilografia do Miocárdio’	117
Tabela 5 – Resultados da análise subjetiva e sob a perspectiva da confiança para o evento significativo ‘Implante de Stent’	119
Tabela 6 – Resultados da análise subjetiva e sob a perspectiva da confiança do evento significativo ‘Angioplastia Vaso Único’	121
Tabela 7 – Resultados da análise objetiva e subjetiva do evento significativo ‘Cateterismo Cardíaco’	123

SUMÁRIO

LISTA DE FIGURAS	viii
LISTA DE GRÁFICOS	ix
LISTA DE QUADROS	x
LISTA DE TABELAS	xiii
1 INTRODUÇÃO	16
1.1 Objetivos	20
1.1.1 <i>Objetivo Geral</i>	20
1.1.2 <i>Objetivos Específicos</i>	20
1.2 Justificativa	20
1.3 Organização do Trabalho	23
2 REFERENCIAL TEÓRICO	25
2.1 Considerações Preliminares	25
2.2 A Informação para Gestão da Saúde e Tomada de Decisão	28
2.3 Gerenciamento de Caso	33
2.4 Dados e Informações de Bases de Dados da Saúde	39
2.5 O Processo de Descoberta de Conhecimento	44
2.5.1 <i>Fases do Processo de Descoberta de Conhecimento</i>	46
2.5.1.1 <i>Pré-processamento</i>	47
2.5.1.2 <i>Mineração de Dados</i>	52
2.5.1.3 <i>Pós-Processamento</i>	60
3 METODOLOGIA	64
3.1 População do Estudo	65
3.2 Etapas Metodológicas	66
3.2.1 <i>Definição do Problema e da Técnica de Mineração de Dados</i>	66
3.2.2 <i>Pré-Processamento</i>	68
3.2.3 <i>Seleção de Dados</i>	68
3.2.4 <i>Limpeza dos Dados</i>	70
3.2.5 <i>Construção de Variáveis</i>	70
3.2.6 <i>Formatação dos Dados</i>	71

3.3	Extração de Informações	72
3.4	Mineração de Dados	72
3.5	Pós-Processamento	72
4	RESULTADOS E DISCUSSÕES	77
4.1	Resultados do Pré-Processamento	77
4.2	Resultados da Extração de Informações	89
4.3	Resultados da Mineração dos Dados	95
4.4	Resultados do Pós-Processamento	97
4.4.1	<i>Análise Subjetiva</i>	99
4.4.2	<i>Análise sob a Perspectiva da Confiança</i>	108
5	CONCLUSÕES	126
6	CONSIDERAÇÕES FINAIS	128
	REFERÊNCIAS	134
	ANEXOS	140

1 INTRODUÇÃO

*“A mente que se abre a uma nova idéia
jamais voltará a seu tamanho original”*

- Albert Einstein

O século XXI tem sido referido como “a Era da Informação e do Conhecimento” (BRANCO, 2001; ANGELONI, 2003), devido às implicações destes conceitos nos mais variados contextos da sociedade. A sua associação à tecnologia tornou-se uma necessidade, bem como a inserção desta na vida pessoal, profissional e empresarial.

Na área da saúde, porém, o uso eficiente dos dados e informações como componentes para a tomada de decisão operacional e gerencial ainda é incipiente, se comparada a outros setores (CARVALHO, 1998). Atualmente, percebe-se uma maior conscientização e crescente tendência na utilização de sistemas informatizados para o gerenciamento de dados relativos a atendimentos de saúde de uma determinada população.

No Brasil, a implantação da informatização do Sistema Único de Saúde (SUS), em 1988, deu início à valorização das informações em saúde como elemento estratégico na elaboração de indicadores epidemiológicos e gerenciais. Mais recentemente, a importância da informação na saúde passou a ser relacionada ao planejamento das estratégias a serem utilizadas no setor de assistência à saúde (BRANCO, 2001), tornando-se o mais valioso *commodity*¹ na área da saúde (BRESNAHAN, 2003). Ademais, o cenário atual do sistema de saúde brasileiro requer que a informatização auxilie tanto no desenvolvimento de estratégias voltadas à sua melhor gestão, como para uma possível reestruturação do modelo de assistência à saúde.

¹ *Commodity* – mercadoria, produto de consumo.

Este modelo está fortemente baseado em uma prática curativista², o que leva a um alto grau de complexidade e custos nos procedimentos, ao não atendimento das reais necessidades de saúde dos indivíduos e conseqüente falta de qualidade assistencial. Aliado a todas essas condições acresce-se ainda a dificuldade de incrementar o acesso da população aos serviços de saúde. Estes aspectos conflituosos do modelo de assistência interferem potencialmente nas decisões relativas ao gerenciamento das questões relacionadas à prestação de serviços de saúde (QUEIROZ; BARBOSA, 2003).

Gonzalez *et al.* (2003, p. 3) refletem sobre a necessidade de buscar formas mais eficientes de oferta e utilização dos recursos na área da Saúde frente à “crise fiscal dos Estados, a crise financeira dos sistemas de saúde, a natureza singular dos bens e serviços de saúde, a consciência sobre a escassez de recursos”, e também sobre o fato de a demanda por serviços de saúde ser praticamente infinita.

Diante dessa realidade, emergem novas tendências de reorganização da assistência à saúde. Entre elas, o movimento pela adequação dos custos, pela valorização da evidência no exercício médico, o desenvolvimento de formas de atendimento extra-hospitalares, a redução do número de leitos hospitalares e a busca de tecnologias de alta capacidade resolutiva ambulatorial (QUEIROZ; BARBOSA, 2003).

O desafio de se desenvolver novos modelos de gestão para a saúde está na conciliação da gestão de recursos finitos para uma demanda crescente dos serviços. Além disso, o aperfeiçoamento dos processos operacionais das organizações pagadoras de serviços de saúde precisa ser alcançado, de forma que os dados referentes aos atendimentos dos usuários possam ser acompanhados e prontamente analisados pelos gestores, permitindo um melhor acompanhamento e gestão dos recursos disponíveis aos usuários (MIRANDA, 2004).

Para tanto, a informatização representa uma ferramenta útil no gerenciamento de dados em saúde; disponibilizando-os à administração da organização de saúde,

² Prática curativista: prática assistencial que reflete uma reação à presença da doença. Ou seja, os usuários já devem apresentar a doença instalada para que a ação terapêutica aconteça. O contrário seria a prática assistencial baseada na prevenção de doenças e promoção da saúde.

seja ela prestadora ou pagadora³ de serviços de saúde, gerando informações estratégicas que permitam o conhecimento da realidade organizacional.

A informação é essencial para a tomada de decisão, mas nem sempre o gestor de saúde dispõe das informações necessárias para decidir com segurança e consistência. Portanto, faz-se necessário um sistema de informações que disponibilize os dados relevantes, tanto nos aspectos gerenciais quanto clínicos e financeiros, que agilize a recuperação dos dados, concorrendo para um aumento da produtividade, qualidade e competitividade no mercado (FITZSIMMONS, 2000).

Observa-se grande quantidade e variedade dos dados que se encontram armazenados em bases de dados informatizadas, porém, torna-se relevante considerar seu registro e armazenamento adequados, e agilização na recuperação dos mesmos. Isso se justifica, pois, freqüentemente, tratam-se de dados históricos, representando transações ou situações que ocorreram no passado, podendo ser úteis na predição de situações futuras (ORALLO *et al.*, 2005).

A partir do uso dessa informação, faz-se necessário desenvolver estratégias de descoberta do conhecimento inerente à base de dados. Esta descoberta faz parte de um processo, o qual se concebe como sendo o mais desejável produto final da computação (SOUZA, 2000).

As ciências indutivas vêm se desenvolvendo, desde a metade do século XX, a fim de auxiliar o homem na recuperação de dados, na elaboração de informações e na descoberta de conhecimentos essenciais à tomada de decisão futura, a partir da identificação de padrões ou associações (REZENDE, 2003).

As organizações prestadoras de serviços de saúde poderiam prover um melhor atendimento à saúde de sua clientela e seguir melhores práticas de negócio se expandissem os dados de identificação, admissão ou de exame físico coletados, para a obtenção de um perfil mais detalhado do seu usuário e, conseqüentemente, o conhecimento das suas necessidades assistenciais. Além da coleta dos dados, é fundamental que as organizações prestadoras e pagadoras de serviços de saúde dispensem esforços no sentido de orientar e incentivar o registro adequado destes

³ Organização ou fonte pagadora – pode ser o Sistema Único de Saúde, financiado pelo Estado; o Sistema Suplementar de Assistência, por meio de operadoras de planos de saúde, seguradoras, e auto-gestões; e autofinanciamento do indivíduo (QUEIROZ ; BARBOSA, 2003).

dados, de maneira a contribuir para a construção de uma base com dados confiáveis acerca da realidade.

Já as organizações pagadoras desses serviços deveriam implementar sistemas de informação que possam trabalhar em tempo real com múltiplas bases de dados, a fim de descobrir conhecimento estratégico sobre a utilização do sistema pelos usuários, cruzando atendimentos entre prestadores, entre outras variáveis.

Desta maneira, a análise epidemiológica das bases de dados permitiria um mapeamento da situação de saúde, a produção de estatísticas de acordo com a incidência e prevalência de doenças e riscos à saúde atual e futura dos usuários, conseqüentemente, embasando o planejamento de atividades de prevenção⁴ e promoção⁵ à saúde, e priorizando investimentos (OZBOLT, 1999; MIRANDA, 2004; UNIDAS, 2005).

A partir dessas estatísticas, os gestores poderiam tomar decisões relacionadas aos serviços a serem prestados, de acordo com as reais necessidades de atendimento à saúde de seus usuários e, ainda, promover ações preventivas e pró-ativas para populações identificadas como em risco para alto custo.

Todas essas afirmações corroboram para a crescente motivação de gestores para a utilização de técnicas mais sofisticadas para melhor aproveitar as bases de dados. Esta motivação está relacionada à existência de tecnologias poderosas para coleta, armazenamento e gerenciamento de grandes quantidades de dados. Estes dados podem estar relacionados a tendências e padrões⁶ que podem ser usados para o melhoramento dos processos de tomada de decisão clínica e administrativa (REZENDE, 2003).

Com base no exposto, a questão norteadora do presente trabalho pode ser sintetizada na seguinte pergunta: que padrões podem ser encontrados em uma base

⁴ Prevenção – conceito clássico de Leavell e Clark (1965), divide-se em três níveis: prevenção primária, que impede que o processo de doença se instale; prevenção secundária, que interrompe o processo de doença antes que se torne sintomática; e prevenção terciária, que limita as conseqüências físicas e sociais de doenças sintomáticas (Leavell ; Clark *apud* JEKEL *et al.*, 2005).

⁵ Promoção – ações que contribuem para a prevenção de diversas doenças. Mudanças não médicas, como no estilo de vida, na nutrição e ambiente. Podem ser aplicadas a doenças infecciosas e crônicas (JEKEL *et al.*, 2005).

⁶ Padrões - modelo, exemplo, protótipo, arquétipo. Aquilo que serve de norma ou base para avaliação de qualidade ou quantidade (FERREIRA, 1988). É um tipo ou modelo de declaração sobre o conjunto de dados que está sendo analisado (CARVALHO, 2005).

de dados de uma organização pagadora de serviços de saúde, que possibilitariam a identificação, por meio de técnicas computacionais sofisticadas, de usuários com doenças cardiovasculares, elegíveis para ingresso em programas de gerenciamento de casos?

1.1 Objetivos

1.1.1 Objetivo Geral

Propor um modelo de identificação de padrões que contribua para a indicação de usuários com doenças cardiovasculares elegíveis para ingresso em programas de gerenciamento de casos.

1.1.2 Objetivos Específicos

Identificar as variáveis administrativas⁷ relacionadas aos usuários registrados na base de dados, no período de evolução analisado;

Verificar, por meio de uma ferramenta computacional, a existência de associação entre as variáveis encontradas em uma base de dados de uma operadora de saúde;

Identificar, a partir dos registros da base de dados, eventos alertas⁸ que estejam associados probabilisticamente a eventos de alta complexidade ou alto custo.

1.2 Justificativa

O cenário atual de assistência à saúde caracteriza-se por mudanças progressivas em relação a: envelhecimento da população, queda nas taxas de natalidade e fertilidade, crescente urbanização, aumento da violência, complexidade

⁷ Variáveis Administrativas – são todos os eventos aos quais os usuários podem ser submetidos para fins diagnósticos ou terapêuticos, incluem: dados de identificação relativos à idade, ao sexo, e procedimentos médicos.

⁸ Eventos alerta - são procedimentos que estão relacionados a eventos de alta complexidade ou alto custo, que indicam que este paciente precisa ser incluído em programas de gerenciamento de casos.

do perfil nosológico⁹ das populações, aumento da freqüência das doenças crônico-degenerativas, aumento do desemprego e distorções da distribuição de renda (QUEIROZ; BARBOSA, 2003).

As doenças cardiovasculares são a primeira causa de óbitos, e a segunda causa de morbidade no Brasil (DATASUS, 2006). O Infarto Agudo do Miocárdio (IAM) é exemplo de doença cardiovascular que pode ser prevenida, e quando se instala requer que o indivíduo receba serviços de alta complexidade, com alto custo financeiro e pessoal em termos de sofrimento e afastamento de suas atividades diárias.

Nesse contexto, torna-se relevante a adoção de ferramentas que possibilitem às organizações de saúde, prestadoras ou pagadoras de serviços de saúde, identificarem as necessidades potenciais de assistência à saúde dos usuários e atuarem precocemente. Desta maneira, ações preventivas do IAM podem substituir, quando possível, ações curativas invasivas, de alto risco e de alto custo, compatibilizando qualidade e custo, garantindo eficiência, equidade¹⁰ e eficácia na assistência à sua clientela.

Organizações pagadoras de serviços de saúde, ao utilizarem em suas atividades de auditoria, somente ferramentas não informatizadas para a identificação dos usuários em risco para alto custo¹¹, culminam por identificá-los quando sua assistência já implica em elevado custo e, conseqüentemente, quando estes já sofreram uma significativa deterioração da sua qualidade de vida (STATUSONE, 2004; MIRANDA, 2004).

É evidente que esses processos de auditoria não identificam todos os casos que poderiam ser indicados para outras modalidades de assistência, como: gerenciamento de doenças crônicas; mapeamento e gerenciamento de riscos; programas para terceira idade e gestantes; atendimento domiciliar (*Home Care*); assistência farmacêutica; programas de apoio a dependentes químicos (UNIDAS, 2005). Isso ocorre pelo fato de os auditores atuarem, na maioria das vezes, no

⁹ Perfil Nosológico – *nosológico* é adjetivo de *nosologia*, que significa “estudo das moléstias” (FERREIRA, 1988).

¹⁰ Equidade é a igualdade da assistência à saúde, sem preconceitos ou privilégios de qualquer espécie.

¹¹ Usuário de Alto Custo – neste trabalho considera-se o usuário de alto custo, como aquele cujas despesas com atendimentos de saúde são maiores que a média dos custos de outros usuários com as mesmas patologias.

ambiente hospitalar, em que não se monitoriza aqueles usuários que utilizam os recursos do sistema exteriores a este ambiente.

Daí se extrai, o benefício que os sistemas de saúde públicos ou privados poderiam obter com a utilização de ferramentas de descoberta de conhecimento em bases de dados (KDD). Tais ferramentas poderiam ser capazes de detectar precocemente usuários considerados de risco para desenvolverem determinadas doenças, e padrões de utilização do sistema de saúde. O monitoramento destes subsidiaria o planejamento da assistência necessária na hora certa, de maneira eficiente, numa tentativa de se evitar o surgimento ou cronificação de doenças (CARLSON, 2003).

O descobrimento de conhecimento a partir de bases de dados¹² favorece o surgimento de novos e inusitados modelos de associações de dados, a partir de uma ampla magnitude de dados, condição impossível de ocorrer num tratamento manual de uma enorme quantidade de dados (ORALLO *et al*, 2005), como ocorre nas organizações de saúde.

Por sua vez, o gerenciamento de caso constitui-se numa oportunidade de melhorar a condição de saúde de indivíduos, e pode ser aperfeiçoado mediante técnicas computacionais, como por exemplo, as de inteligência artificial. Tais técnicas tornam possível o desenvolvimento de estratégias antes impossíveis, tais como, a de identificar numa determinada população, candidatos que ainda não são, mas cedo ou tarde poderão se tornar, usuários de alto risco e alto custo para o sistema de saúde (MCCAIN, 2001, GILLESPIE, 2004).

Apesar da produção científica brasileira sobre a utilização de ferramentas computacionais na gestão da saúde na área da saúde ser ainda incipiente, tem-se observado ênfase na sua utilização nas políticas públicas, na academia e no sistema suplementar de assistência.

Rothman e Wagner (2003) comentam que estudos que contemplam a evidência da efetividade de programas de gerenciamento são escassos, embora haja evidências crescentes referentes à eficiência de programas que desenvolvem gerenciamento clínico a partir de protocolos de apoio ao autogerenciamento e monitoramento constante.

Esta pesquisa vem contribuir com a ampliação de conhecimento relacionado

¹² Do inglês, *knowledge discovery from databases* – KDD.

à temática do gerenciamento de casos, como sugerem Rothman e Wagner (2003), além de evidenciar o benefício da aproximação das ferramentas de tecnologia da informação para as questões relacionadas à gestão em saúde.

1.3 Organização do Trabalho

Esta dissertação, além desta Introdução, contempla o Capítulo 2 referente ao referencial teórico, onde são abordadas o modelo assistencial vigente, as mudanças de paradigma pelo qual este modelo tende a passar, bem como da inserção da tecnologia da informação no apoio a decisões clínicas e administrativas.

Ainda, são definidos e discutidos os conceitos de dado, informação e conhecimento e a hierarquia existente entre eles. Além disso, é abordada a importância do entendimento destes conceitos para o melhor aproveitamento do potencial de aplicabilidade de tecnologias da informação, de modo que os sistemas de informação em saúde possam realmente contribuir para o alcance de objetivos organizacionais. Adicionalmente, são descritas as características dos métodos indutivo e dedutivo de aprendizado.

Dando continuidade ao referencial teórico pesquisado, também é discutido o papel fundamental da informação e do conhecimento para a gestão e tomada de decisões em organizações de saúde. São discutidas algumas definições acerca do processo decisório, exemplos de tomadas de decisão na saúde e questões a serem observadas no desenvolvimento de bases de dados e sistemas de informação para a gestão em saúde, incluindo requisitos de conteúdo para o processo de informação em saúde.

São apresentados os tópicos relativos a programas de gerenciamento de casos: definição, histórico, metas, fases e população mais beneficiada. Também são abordados os recursos tecnológicos disponíveis para que iniciativas como as de gerenciamento de casos tenham sucesso, contribuindo para o alcance de objetivos institucionais e da comunidade em geral.

É ainda descrito o processo de descoberta de conhecimento em bases de dados (KDD), e uma etapa relevante dele, que é o processo de mineração de dados (MD). Definições, metodologia, descrição de técnicas e aplicabilidade são abordados, além de um resumo das principais técnicas de mineração de dados, e

dos algoritmos de escolha para cada técnica escolhida para resolução do problema de mineração de dados.

O Capítulo 3 apresenta a metodologia utilizada nesta pesquisa para atingir o objetivo de identificação de padrões que possam indicar usuários com doenças cardiovasculares elegíveis para programas de gerenciamento de casos. De acordo com o apresentado no referencial teórico, são descritas as etapas metodológicas, incluindo o tipo de estudo, população e aspectos éticos observados durante a execução da pesquisa.

No Capítulo 4 são apresentados os resultados advindos da aplicação da metodologia de descoberta de conhecimento em bases de dados, e também dos experimentos realizados em cada uma das suas etapas metodológicas, com as discussões pertinentes, incluindo sugestões para trabalhos futuros. No Capítulo 5 são apresentadas as conclusões referentes aos resultados obtidos, frente aos objetivos propostos e referenciais teórico pesquisados.

Por fim, no Capítulo 6 são apresentadas as considerações finais, indicando as limitações e contribuições da presente pesquisa, além de reflexões referentes à necessidade de mudança do modelo assistencial vigente.

2 REFERENCIAL TEÓRICO

*“O que sabemos é uma gota,
o que ignoramos é um oceano”
- Isaac Newton*

Neste capítulo serão abordados assuntos que fundamentam a necessidade da implementação de ferramentas computacionais na gestão em saúde, tanto para auxiliar na tomada de decisão clínica como na administrativa. Enfatiza-se a importância da compreensão dos conceitos de *dado*, *informação* e *conhecimento* a fim de evidenciar a aplicabilidade da tecnologia de informação ao gerenciamento em saúde. São apresentadas definições de gerenciamento de casos, seu histórico, metas, fases e população mais beneficiada, bem como as informações básicas que devem contemplar as bases de dados na saúde. O processo de descoberta de conhecimento em bases de dados é abordado como uma metodologia que gera vantagens competitivas à organização que os implementa nas suas atividades cotidianas, principalmente no apoio a decisões em todos os níveis organizacionais.

2.1 Considerações Preliminares

As organizações, em geral, estão cedendo cada vez mais espaço a novas formas de gestão. O setor da saúde, especificamente, encontra-se frente a um divisor de águas, no sentido de abandonar um paradigma basicamente assistencialista para adotar um paradigma de gestão técnica com foco na prestação de serviços de qualidade ao cliente e na sustentabilidade financeira.

Nesta perspectiva, as organizações de saúde, prestadoras ou pagadoras de serviços de saúde, deverão estar aptas a ir ao encontro das necessidades sociais para promover e recuperar a saúde de indivíduos e da comunidade, e ainda gerar seu lucro de maneira razoável, evitando o desperdício (OZBOLT,1999).

Estas organizações vêm passando por um período de fortes pressões geradas tanto pelo aumento dos seus custos operacionais, bem como pelas injustiças sociais

com relação ao acesso a sistemas de saúde (QUINTO NETO, 2000). A visão adotada por estas organizações, por estar fortemente baseada na ética exclusiva de mercado, em detrimento da ética social, acarreta-lhes desprestígio por uma conseqüente falta de resposta às necessidades das comunidades (QUINTO NETO, 2000).

A área da saúde difere das demais por atuar em um contrato social no qual é esperado que se promova saúde. Entretanto, é fundamental que esta atuação ocorra dentro de princípios de negócio; de outra forma, os recursos necessários para alcançar a sua missão estarão perdidos. É neste aspecto que se constata que existe certa tensão no mercado de assistência à saúde, pois existe uma força que impele as organizações para o cumprimento da missão de promover a saúde, e outra que demanda a conservação de recursos e o alcance de lucro com este negócio (OZBOLT, 1999).

Considerando todos estes aspectos, pode-se afirmar que as ferramentas de tecnologia da informação apresentam um grande potencial de contribuição para um melhor gerenciamento dos serviços de saúde, pelo apoio conferido pelo uso de bases de conhecimentos necessárias à tomada de decisão clínica e administrativa (PAHO,1999).

Apesar da utilização ainda incipiente de ferramentas de tecnologia da informação na área da saúde, se comparada a outros setores de serviços e produtos. Contudo, são notáveis os investimentos que vêm sendo aplicados atualmente para a aquisição de sistemas de informação que gerenciem de maneira eficiente a enorme quantidade de dados que advêm dos atendimentos à saúde de uma população.

Entretanto, tal como preconiza o *Institute of Medicine* (IOM), é fundamental que se obtenha informação de qualidade sobre os usuários e o cuidado por eles recebidos, a fim de que os resultados possam ser melhorados – tanto em relação ao cuidado recebido pelo usuário, traduzido em qualidade de vida, quanto ao gerenciamento de recursos da fonte pagadora (METZGER, 2004). A informação de qualidade garante que, ao se aplicar ferramentas computacionais para análise de bases de dados, os resultados sejam confiáveis e fidedignos, podendo ser utilizados para a tomada de decisões clínicas e administrativas.

Novas maneiras de se melhorar a qualidade da assistência em saúde, com controle de custos, vêm sendo almejadas pelas organizações de saúde,

impulsionadas pela crescente demanda de escolha e flexibilidade dos usuários, bem como pela preocupação com a inflação da saúde (LEGROW, 2004). Nota-se existir um crescente interesse em uma abordagem mais sistemática no gerenciamento de usuários com doenças crônicas. Primariamente, isto se deve ao fato de haver maior compreensão das vantagens em prestar assistência que possa prevenir ou adiar o desenvolvimento de doenças crônicas, e complicações a elas relacionadas, o que leva conseqüentemente a maiores despesas com medicamentos, internações, ausência no trabalho, entre outras (METZGER, 2004).

Nesse sentido, auto-gestões¹³ têm investido em ações preventivas realizadas em conjunto com as áreas de gestão de pessoas e medicina ocupacional, com o objetivo de melhorar a qualidade de vida dos seus colaboradores (UNIDAS, 2005), tendo em vista o aumento da expectativa de vida e o crescente envelhecimento dos brasileiros.

Abordagens voltadas à qualidade de vida são de grande importância para o planejamento e avaliação do impacto de ações sociais, de serviços de saúde e bem estar para indivíduos, famílias ou grupos populacionais. Estes, de acordo com sua idade, sexo, e momento da sua existência, têm necessidades variadas e estão expostos a riscos que lhes são característicos. A avaliação destas ações deve considerar “três elementos básicos: a qualidade de vida, o direito e o acesso aos bens e serviços coletivos, e o espaço social como espaço histórico social” (VERMELHO *et al.*, 2004, p.33).

Nesta perspectiva, e focando as organizações pagadoras de serviços de saúde, é fundamental que armazenem e sejam capazes de recuperar registros relacionados à identificação de usuários e a sua utilização dos serviços, entre outros. Conseqüentemente, poderão detectar usuários que possam se beneficiar de estratégias de promoção e prevenção em saúde. Tais estratégias, baseadas em dados e informações confiáveis, promovem a gestão dos recursos destas organizações, numa tentativa de melhorar o acesso e atendimento da população nos serviços de saúde, e garantir a saúde financeira organizacional.

¹³ Auto-Gestão – modalidade de organização pagadora de serviços de saúde, sem fins lucrativos, que se destina a empregados ativos, aposentados, pensionistas ou associados e seus dependentes, participantes de empresas, associações, sindicatos, fundações, caixas de assistência ou entidades de classes profissionais. É a forma na qual as próprias empresas administram programas de saúde para seus empregados (UNIDAS, 2005).

2.2 Informação para Gestão da Saúde e Tomada de Decisão

A informação é essencial para a tomada de decisão, mas nem sempre o gestor de saúde dispõe das informações necessárias para decidir com segurança e consistência. Portanto, ele necessita de um sistema de informações que contenha as informações relevantes, tanto nos aspectos gerenciais, como em relação aos aspectos clínicos, e que lhes possibilite geração de renda, vantagens pelo uso de base de dados, intensificação da produtividade e competitividade no mercado (FITZSIMMONS, 2000).

Sendo assim, faz-se necessário uma priorização do uso da informação em organizações de saúde, pois se trata de uma atividade-meio que contribui significativamente para a gestão, implementação e desenvolvimento da atividade-fim¹⁴ dessas organizações. Esta priorização depende, primordialmente, da conscientização da importância da coleta e de registro adequados pelos profissionais responsáveis por estas atividades-fim (CARVALHO, 1998).

De acordo com Évora (1995), a tecnologia computacional pode contribuir para aprimoramento da qualidade da assistência à saúde, uma vez que pode facilitar o planejamento, a tomada de decisão, a comunicação, o controle gerencial e as mudanças na estrutura organizacional. A autora complementa que os sistemas de informação podem também auxiliar na estruturação, operacionalização, controle e avaliação do desempenho do serviço.

A administração de organizações está intimamente ligada à tomada de decisão, que nada mais é do que o processo de escolha de uma alternativa dentro de um conjunto de possíveis alternativas, a fim de solucionar um determinado problema (MEGGINSON; MOSLEY; PIETRI, 1998). Entretanto, estes mesmos autores e ainda Lacombe e Heilborn (2003) comentam sobre o fato do processo decisório perpassar por uma chance de identificação e análise de uma oportunidade de crescimento organizacional, mediante a superação de metas e alcance de objetivos estabelecidos (STONER; FREEMAN, 2003). Tal superação pode ser conseqüência da utilização não só de informações estratégicas, mas também do

¹⁴ Atividades-fim em organizações de saúde – são as atividades que cumprem efetivamente as finalidades destas organizações, ou seja, a prestação de serviços de assistência à saúde, prestados por diversos profissionais. Em organizações pagadoras destes serviços, as atividades-fim são as atividades de auditoria dos processos de pagamento de serviços de saúde.

conhecimento organizacional, por meio de instrumentos de descoberta de conhecimento em bases de dados.

Autores como Sigulem *et al.* (1998), Carvalho (1998) e Amaral (2002) consideram que o principal objetivo da tecnologia da informação na área da saúde é colocar à disposição dos profissionais as informações relevantes, onde e quando elas forem necessárias. Esses autores, e ainda Bresnahan (1997), ressaltam que é a contínua preocupação com o bem-estar do usuário e a conseqüente necessidade de um processo de decisão ideal que diferenciam o setor da saúde dos outros setores nos quais a informação é essencial, a saber, o sistema bancário e de aviação. Parafraçando Branco (2001), informação em saúde

[...] é aquela produzida com o objetivo de identificar problemas individuais e coletivos do quadro sanitário de uma população, propiciando elementos para a análise da situação encontrada e subsidiando a busca de possíveis alternativas de encaminhamento. Envolve tanto aquelas relativas ao processo de saúde/ doença, quanto as de caráter administrativo, todas essenciais para a tomada de decisão (BRANCO, 2001, p.165).

A informação é a matéria prima para a tomada de decisão, que pode ser definida como a conversão das informações em ação. O processo decisório ocorre independente do nível hierárquico da pessoa que toma a decisão. Por exemplo, quando o médico coleta os dados do usuário para formular seu diagnóstico, seu cérebro processa e analisa os dados coletados na anamnese, transformando-os em informação para que tome a decisão acerca do melhor plano de tratamento para aquele caso. Por sua vez, o auditor de saúde, na avaliação de uma solicitação de um procedimento médico, necessita de informações relacionadas à história clínica do usuário em questão, para avaliar a necessidade e pertinência de tal procedimento.

Alguns autores utilizam o termo “tomador de decisões” em vez de “gerentes”, pois muitos profissionais, independentemente da atividade específica exercida, tomam decisões, principalmente aqueles que desempenham atividades administrativas ou gerenciais (MONTANA, 1999).

Os gerentes necessitam de dados e informações relativos à dinâmica organizacional para poderem planejar, organizar, dirigir, controlar e avaliar ações e serviços da sua organização, a fim de alcançar suas metas de acordo com os recursos disponíveis. Embora, não exista obrigatoriedade na informatização do

processo gerencial, o papel desta é o de organizar as informações relevantes a fim de otimizar o processo decisório (BRANCO, 2001; AMARAL, 2002).

Uma vez definido que o gerente ou gestor é a pessoa envolvida no processo decisório seja ele qual for, pode-se acreditar que nem sempre o gestor está munido com as informações que necessita para tomar suas decisões de maneira segura e consistente. É por este motivo que, às vezes, muitos gerentes tomam decisões de maneira intuitiva ou aleatória, distanciando-os dos objetivos e resultados a serem alcançados pelos funcionários e pela organização, propriamente dita.

Assim, um aspecto importante no desenvolvimento de um sistema de informações em saúde é a diferenciação das informações relevantes das irrelevantes, de acordo com a visão do usuário desse sistema, que atua nos níveis, operacional, gerencial, ou estratégico.

Todavia, deve-se considerar que se vive numa sociedade dependente da velocidade com que as informações aparecem e desaparecem. Com o advento da Internet, que disponibiliza informações de maneira fácil, muito vem sendo discutido sobre a relevância dessas informações na rede mundial de computadores, pois o acesso às informações, embora fácil, pode tornar-se complexo, devido à intensa seleção que deve ser feita. O mesmo acontece dentro das organizações, tornando-se fundamental o estabelecimento de objetivos, ou seja, aonde se quer chegar com os dados coletados em um sistema de informação.

Bresnahan (1997) lembra que muitos sistemas de informação na área da saúde tendem a focalizar sua atenção em operações financeiras, a fim de garantir o retorno de investimentos. Mais recentemente, observou-se que esta mesma tecnologia poderia ser utilizada para propósitos clínicos. Além disso, a autora afirma que organizações contemporâneas e de vanguarda estão partindo para além do lado financeiro das informações, utilizando a tecnologia da informação como um projeto coeso, o qual trabalha com coleta de dados, equipamentos médicos e outros formulários, não mais como projetos isolados.

Amaral (2002), citando Van Bammel, complementa a idéia de integração das informações estratégicas da organização de saúde, escrevendo que

[...] um dos grandes desafios para a comunidade científica no campo da Informática aplicada à Saúde é a implementação de soluções de grande impacto funcional, aliado a um baixo custo operacional, que possa contribuir

para a melhoria da qualidade do setor, facilitar o acesso e expandir serviços ofertados aos usuários, e prover dados organizados e rapidamente disponíveis para os profissionais da saúde, sejam médicos, pesquisadores, epidemiologistas, estatísticos, administradores, etc (AMARAL, 2002, p.226).

Segundo Fitzsimmons (2000, p.88), “a função estratégica da informação como parte integrante das estratégias de gestão de serviços pode ser organizada em quatro categorias: criação de barreiras à entrada de competidores, geração de renda, vantagens pelo uso de bases de dados e intensificação da produtividade”.

Neste sentido, Carvalho (1998) contribui com o parecer de que as soluções da tecnologia da informação contribuem substancialmente no suporte para alguns dos processos administrativos habituais das organizações de saúde, a saber: transmissão e processamento das faturas, avaliação da utilização dos serviços, compra de suprimentos e administração de estoques, pagamento de contas, gerenciamento das finanças, negociação de contratos e gerenciamento da qualidade. A partir destas soluções de tecnologia da informação, o gestor implementa processos de comunicação, logística e apoio à decisão.

Além da contribuição das soluções da tecnologia da informação para os processos administrativos, destaca-se a possibilidade de que dados clínicos advindos da assistência direta aos usuários sejam armazenados e recuperados posteriormente. Assim, é possível construir um histórico de atendimentos que contribua para o conhecimento da evolução de cada usuário, com resultados de exames laboratoriais e de imagem, e evoluções clínicas do atendimento prestado pelos diversos profissionais da saúde, de modo a auxiliar a tomada de decisão futura.

Adicionalmente, é importante salientar que informações administrativas e clínicas correspondem ao corpo de conhecimentos necessário para a tomada de decisão no que tange a utilização de ferramentas de gestão em sistemas de saúde. Um exemplo é o gerenciamento de casos, ferramenta de gestão que, para se desenvolver com eficácia e eficiência exige informações administrativas como: o uso pelo usuário do sistema de saúde em números de consultas e internações hospitalares; e informações clínicas como as patologias às quais os usuários são acometidos, patologias secundárias, resultados de exames que os colocam em situações limítrofes de uma outra patologia, entre outras.

2.3 Gerenciamento de Caso

Sabendo-se que os recursos necessários para a assistência em saúde são finitos, e infinita é a demanda por eles, a solução para este dilema versa entre o aumento desses recursos; a diminuição da demanda, ou pelo menos a sua utilização adequada, com aumento da eficiência dos atendimentos ou tratamentos. Embora, se percebam tentativas nas três frentes de resolução deste problema, tem havido uma maior ênfase na diminuição da demanda e no aumento da eficiência, mediante estratégias de contenção de custos na assistência gerada (JEKEL *et al.*, 2005).

O gerenciamento de casos, utilizada inicialmente nos Estados Unidos, é uma metodologia caracterizada pela atenção gerenciada (MENDES, 1998) à saúde de usuários de um sistema de saúde.

Trata-se de uma forma particular de acompanhamento concorrente e prospectivo de usuários que necessitam de acompanhamento e intervenções de saúde. Iniciou-se nos anos 40, juntamente com a prática do pagamento por procedimento médico; consolidou-se na década de 60, quando vigorava o sistema de financiamento por *capitation*¹⁵; fortaleceu-se nos anos 70, com as organizações de manutenção da saúde (HMOs) e, consolidou-se na última década, com os sistemas integrados de prestação de serviços de saúde. No Brasil, esta prática ainda é incipiente, porém muito se tem discutido acerca de suas vantagens e desvantagens para a realidade brasileira (MENDES, 1998).

O gerenciamento de caso constitui-se numa ferramenta de gestão desenvolvida com o objetivo de organizar formas de intervenção que possam influenciar no risco de indivíduos expostos virem a desenvolver doenças ou de antecipar as suas complicações. Pressupõe

[...] a identificação de casos, em geral de natureza crônica, com potencial para geração de demanda por serviços, e a introdução de atividade gerencial com vistas a adequar a aplicação de recursos para a otimização da assistência no sentido de obter uma relação entre o custo e qualidade mais favorável (MIRANDA, 2003, p.16).

Em outras palavras, o gerenciamento de casos visa coordenar serviços de saúde para indivíduos vulneráveis, ou seja, aqueles que necessitam utilizar um sistema de saúde complexo, de acordo com suas necessidades. Tal coordenação se

¹⁵ *Capitation* – forma de financiamento da assistência à saúde caracterizada pelo atendimento sendo pago per capita, antecipadamente e independente do número de procedimentos efetivamente executados (UNIDAS, 2005).

faz necessária devido ao fato de o sistema de saúde apresentar-se fragmentado, o que demanda que haja uma preocupação com a sistematização da prestação de serviços a fim de que os usuários recebam aqueles serviços pertinentes às suas necessidades; que a prestação de serviços de saúde seja integrada entre os diversos prestadores e ao longo do tempo; e, que sejam prestados de acordo com os objetivos tanto do cliente como do sistema de saúde ao qual ele pertence (NCCC, 2000).

No gerenciamento de casos, uma equipe ou profissional previamente indicada, responsabiliza-se pela determinação e gestão da assistência à saúde de usuários. Por conseguinte, monitora os atendimentos prestados aos usuários pelos quais são responsáveis, dentro de um sistema de saúde, público ou privado (MENDES, 1998).

Esta atividade de gerenciamento compreende o diagnóstico das reais necessidades dos usuários e seus familiares e a proposta, na medida do possível, de melhores práticas para atendê-las (MIRANDA, 2003).

O gerenciamento de casos é considerado um modelo de assistência de difícil implementação, principalmente no sistema público de saúde, por envolver gestão contínua dos atendimentos à saúde do usuário em diferentes serviços de assistência à saúde: hospitais, ambulatórios especializados, atenção domiciliar, entre outros, demandaria a integração de todas as informações geradas pelos atendimentos (MENDES, 1998).

A qualidade da assistência e a eficiência do custo são alguns dos objetivos do gerenciamento de casos, cujas estratégias para alcançá-los devem estar baseadas em valores flexíveis, inovadores e humanos. Neste contexto, considera-se que o enfermeiro é o profissional mais indicado para exercer a função de gestor de casos devido ao seu conhecimento clínico, habilidade para cuidar de maneira integralizada, e por conhecer o processo de trabalho de outros profissionais da área da saúde (GONZALEZ *et al.*, 2003).

A meta do gerenciamento de casos é a busca de benefícios para o usuário e sua família, bem como para os prestadores, e para os pagadores de assistência de saúde.

De acordo com o *National Chronic Care Consortium* (NCCC, 2000, p.1), os benefícios do gerenciamento de casos para os usuários incluem: melhor condição de saúde e funcional; informação e acesso aos serviços necessários de maneira

contínua; maior envolvimento nas decisões relacionadas à assistência à sua saúde; e, serviços custo-efetivos num ambiente menos restritivo. Dentre os seus benefícios para a família estão: a assistência de um profissional gestor de casos experiente; acesso à informação e orientação fundamentais para a tomada de decisões relativas ao processo de saúde-doença; encaminhamentos a serviços que se tornem necessários; e, apoio emocional.

Os benefícios para os prestadores de serviços de saúde são: a coordenação de serviços de saúde entre diversos provedores; o encaminhamento de usuários para outros serviços que se tornem necessários de maneira continuada; o controle da qualidade dos serviços; e, a oferta de dados para a discussão de alternativas de assistência à saúde.

Por sua vez, para as organizações pagadoras de serviços de saúde, os benefícios são: a satisfação dos clientes; a substituição monitorada de serviços de maior custo por aqueles de menor custo; e, ações que evitem a custosa permanência prolongada em internações hospitalares.

Tais metas podem ser alcançadas por meio de (a) busca pela qualidade da assistência à saúde, visando que os serviços sejam apropriados e benéficos para a população assistida; (b) gestão do tempo de permanência do usuário internado no hospital; (c) utilização controlada de recursos de apoio mediante sistemas de gerenciamento de casos que utilizem protocolos, sistemas de apoio à decisão e diretrizes; (d) controle de custos pelo gerenciamento dos casos a fim de controlar as despesas com atendimentos à saúde, bem como assegurar resultados eficazes (MAY, SCHRAEDER E BRITT *apud* GONZALEZ, 2003).

Embora cada caso seja único e possa demandar variações de conduta, o processo de gerenciamento de casos pode ser classificado em cinco fases: busca ativa; *screening*; avaliação abrangente; plano de cuidados; encaminhamento aos serviços; monitorização; e, reavaliação (NCCC, 2000).

Um programa de gerenciamento de caso pode ter sua abordagem voltada a usuários internados. Neste caso, as condutas são de: encaminhamento de profissionais e procedimentos; controle da qualidade; e, de orientações para alta hospitalar. Quando os usuários são aqueles que já se encontram em suas comunidades, as ações geralmente são conduzidas a longo prazo, incluindo orientações aos cuidadores acerca de recursos da própria comunidade, encaminhamentos a serviços ambulatoriais, comparecimento à consultas, mudança

de hábitos de vida, incentivo à adesão e cumprimento do plano terapêutico proposto, entre outras.

Já, as ações realizadas pelas organizações pagadoras de serviços de saúde podem estar relacionadas à prevenção de eventos indesejáveis, como uma utilização repentina e não esperada de serviços devido a um agravamento na saúde do indivíduo que o leve à internação, procedimento de alto custo, e conseqüente modificação na sua condição de vida (NCCC, 2000).

A *American Hospital Association* (1992) descreve cinco modelos de gerenciamento de casos, de acordo com a estrutura e com os objetivos a serem alcançados mediante sua aplicação: (a) o gerenciamento de casos no nível primário, quando o profissional neste nível de assistência à saúde é responsável pelo gerenciamento de todos os cuidados à saúde; (b) gerenciamento de casos “médico”, o qual prevê a monitorização de cuidados médicos a usuários com doenças graves; (c) gerenciamento de casos “social”, a partir do gerenciamento de recursos sociais e econômicos para doentes não agudos, que necessitam de acompanhamento para se evitar uma ocorrência mais grave; (d) gerenciamento de casos médico-social, numa combinação de esforços médicos e sociais para resgatar recursos de saúde, sociais e econômicos; e, (e) gerenciamento de casos “vocacional”, para orientação de pessoas com deficiências na sua inserção na sociedade e no mercado de trabalho.

A população que mais se beneficia com a estratégia de gerenciamento de casos é aquela caracterizada por usuários que apresentam condições de saúde complexas e que necessitam assistência específica e prolongada, incorrendo em alto custo para o sistema de saúde. Dentre eles, os doentes mentais crônicos, os deficientes por problemas de desenvolvimento ou lesões adquiridas, os usuários portadores do vírus da imunodeficiência humana, os usuários de drogas, os idosos, as populações de risco (GONZALEZ *et al.*, 2003), entre outros, a exemplo daqueles que apresentam comorbidades¹⁶. Estes se encontram freqüentemente deteriorados clinicamente, enfrentando desafios sociais e psicológicos.

Como apresentado na figura 1, esses usuários ditos de alto risco correspondem à cerca de 0,5 a 1% dos usuários de um sistema de saúde,

¹⁶ Comorbidade – existência de doenças outras, além de determinada doença objeto de estudo em um indivíduo (Biblioteca Virtual em Saúde., 2006).

consumindo, no entanto, cerca de 30% dos recursos disponíveis (KNABEL; LOUWERS, 2004; STATUSONE, 2004; FORMAN, 2004; CROOKS, 2005).

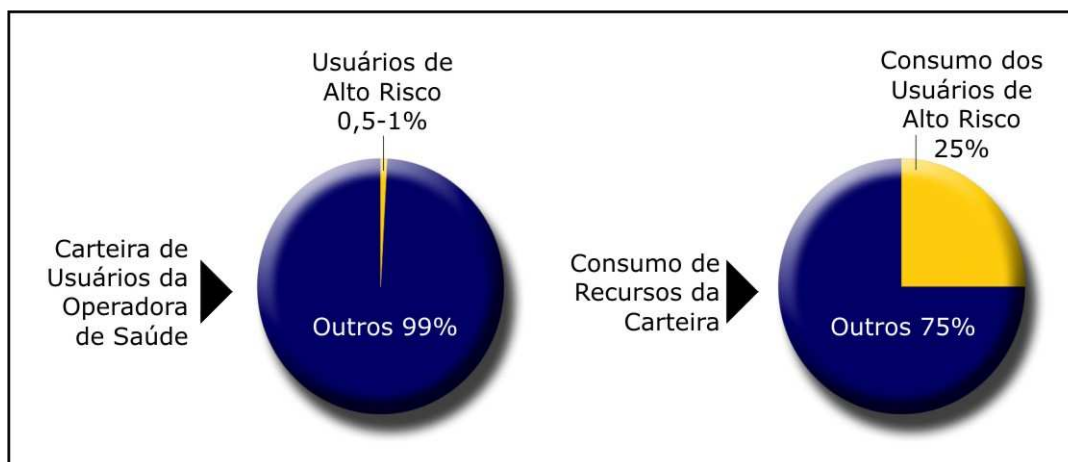


Figura 1 - Proporção de usuários em risco para alto custo e o respectivo consumo de recursos da organização pagadora de serviços de saúde

Fonte: Statusone (2004).

Os usuários, cuja assistência gera altos custos para os sistemas de saúde, são aqueles que apresentam os perfis mais complexos do ponto de vista clínico e psicossocial. Aproximadamente 45% destes apresentam cinco ou mais diagnósticos para descrever a sua condição crônica, podendo, cada um ser o foco de um programa específico de gerenciamento (FORMAN, 2004).

Outra característica desta população é que seu risco pode variar de 'risco moderado' para 'alto risco iminente', devido ao fato de apresentarem várias comorbidades que podem exacerbar uma ou outra condição, e agravar seu estado de saúde. Os problemas de saúde que esses indivíduos apresentam podem retirá-los do convívio com amigos, família e trabalho. Frequentemente, encontram-se num ciclo de isolamento, depressão subclínica e deterioração das condições de saúde (FORMAN, 2004).

Adicionalmente, considera-se que os indivíduos que se encontram em 'alto risco' para desenvolver problemas crônicos de saúde respondem por 70 a 90% das despesas da organização pagadora de serviços de saúde com assistência, e correspondem de 1 a 2% de uma população pela qual esta é responsável (BLAIR, 2004; CROOKS, 2005). De uma maneira geral, os usuários crônicos estão associados a cerca de 20 a 30% das admissões nos hospitais e clínicas, e de 30 a 45% das diárias de internamentos (STATUSONE, 2004).

Com relação a usuários considerados de alto risco e de alto custo, os gestores do sistema de saúde devem se preocupar com a prestação efetiva de serviços, que são previsivelmente necessários, e com a coordenação dos serviços, quando o imprevisível ocorre. Nestes casos, programas de gerenciamento de casos desempenham um papel fundamental, pois tratam de alocar recursos médicos de maneira eficiente e eliminar despesas provenientes daqueles que não produzem resultados satisfatórios (CROOKS, 2005).

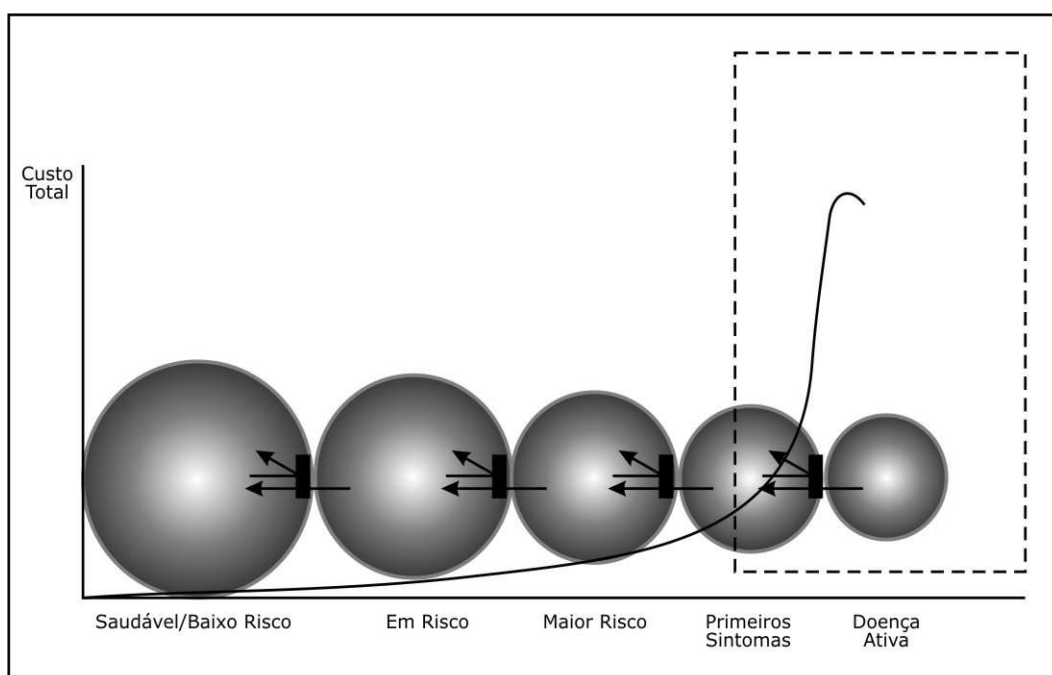
Uma ferramenta computacional para apoiar as decisões referentes a usuários indicados para um programa de gerenciamento de caso pode utilizar recursos de inteligência artificial, tecnologia não convencional a fim de incluir e monitorar todos os usuários do sistema de assistência à saúde em níveis de risco.

Desta maneira, usuários poderão ser avaliados precocemente quanto ao seu risco de desenvolver determinada condição no processo saúde-doença, de acordo com diretrizes previamente estabelecidas em consenso. Tal fato possibilita que o gerenciamento da assistência à saúde proporcione o que cada usuário necessita, dentro de uma abordagem de utilização eficiente dos recursos do sistema de saúde. Todavia, o importante não é somente o gerenciamento da excessiva utilização dos recursos, mas também o da sub-utilização destes na assistência à saúde àqueles usuários que deles necessitam (ISHAM; KRAMER, 2003).

A proposta da identificação de riscos é prover o serviço de saúde adequado, no ambiente adequado, no tempo adequado, pelo preço adequado (NCCC, 2000). Isto pode ser exemplificado pelo gráfico 1, que apresenta os riscos e a distribuição de usuários por níveis progressivos de riscos e custos totais, do modelo de saúde do *Health Partners*, plano de saúde sem fins lucrativos, responsável por 67.000 vidas no Minnesota (USA). Verifica-se que este modelo tem por objetivo mover os usuários para níveis mais baixos de risco.

Neste modelo, iniciativas específicas de gerenciamento da população foram traçadas para cada área da saúde considerada como prioridade para ser gerenciada, tanto em termos de manejo da doença, ou da promoção da saúde. Os usuários do *Health Partners* foram categorizados em segmentos distintos, de acordo com seu nível de risco. Os autores atentam para a identificação e gerenciamento dos usuários de maior risco, considerando que esta população geralmente apresenta uma ou mais condições crônicas associadas (ISHAM ; KRAMER, 2003).

Gráfico 1 – Distribuição de usuários por níveis progressivos de risco e custo total. Health Partners, Minnesota-USA, 2001.



Fonte: Isham e Kramer (2003).

Para que se possa realmente adotar uma atitude pró-ativa de prestação de assistência à saúde, faz-se indispensável a existência de dados que possam levar à informação e conhecimento real dos usuários do sistema de saúde. Só então, estratégias de promoção à saúde e prevenção de doenças poderão ser delineadas já nos primeiros níveis de risco, quando o usuário é classificado como 'saudável' ou 'sem risco', assim que ele se torna membro do sistema. É importante considerar que sem um sistema de informação projetado para tal, dificilmente essas estratégias podem ser colocadas em prática.

Um sistema de informação que vá ao encontro das necessidades da organização é imprescindível, porém, não suficiente para a utilização, não somente dos dados relativos aos processos de trabalho inerentes às práticas desempenhadas nas organizações, mas principalmente das informações e conhecimento que podem ser abstraídos a partir destes dados.

Ferramentas de mineração de dados encontram sua aplicabilidade nos sistemas da saúde tanto no auxílio à redução de custos como no aperfeiçoamento da prestação de serviços de saúde, mediante a identificação estatística dos tratamentos mais efetivos como, por exemplo, através da análise dos resultados

destes tratamentos. Mais ainda, podem desempenhar um papel fundamental na identificação de indivíduos que apresentam um risco estatístico ou probabilidade para certos agravos, os quais podem receber assistência à sua saúde antes que desenvolvam características patológicas de alto custo e morbi-mortalidade (BRESNAHAN, 2003).

2.4 Dados e Informações de Bases de Dados da Saúde

Para um aproveitamento eficiente do potencial que a tecnologia de informação pode proporcionar, faz-se necessária a adequada compreensão dos termos *dado*, *informação* e *conhecimento*. Estes elementos se inter-relacionam e definem a profundidade das ações em tecnologia da informação para o alcance dos objetivos de indivíduos, grupos e culturas.

Sendo assim, *dado* é qualquer indicador bruto, quantificável, sem capacidade de, por si só, levar à compreensão de uma situação ou realidade. Dados podem ser captados e armazenados por recursos de tecnologia da informação, mas não exclusivamente (GOLDSCHMIDT; PASSOS, 2005). São geralmente gerados e utilizados no ambiente operacional, e seu registro, seleção e recuperação feitos a partir de uma base de dados, ou de outras formas de documentos (REZENDE, 2003).

Pelo processamento de dados, a fim de encontrar significado ao que se observa, tem-se a *informação*, que nada mais é do que o dado contextualizado (BRANCO, 2001; REZENDE, 2003).

O *conhecimento* pode ser entendido como o conjunto de padrões que podem ser formulados a partir do relacionamento entre os dados e entre as informações (GOLDSCHMIDT; PASSOS, 2005); também, como informação armazenada ou os modelos usados por pessoas ou máquinas na interpretação, predição e resposta apropriadas ao mundo exterior (ARARIBOIA, 1988). Distingue-se da informação por já ter sido submetido à interpretação, categorização, aplicação e revisão, que possibilitaram a sua utilização no cotidiano das pessoas de forma contínua, e

aplicação em tomadas de decisões, com eficácia¹⁷ e/ ou eficiência¹⁸ (CARVALHO, 1999).

Como exemplo de dado, pode-se citar o número 40080021 na base de dados deste estudo. Como informação, este número representa um código que está relacionado ao procedimento médico de cateterismo cardíaco de câmaras esquerdas; Portanto, só passou a existir a informação quando foi agregado valor ao dado. Como conhecimento, este número trata de um código de procedimento de alto custo.

A transformação dos dados em informação pode ocorrer manual ou eletronicamente, por meio do seu armazenamento, processamento e análise.

A comparação de uma informação com outra, e das suas combinações com muitas outras ligações úteis, e ainda com significado, faz parte do processo de geração de conhecimento. São essas combinações de estruturas de dados com procedimentos interpretativos que levam a um comportamento conhecido que, por sua vez, facilita a tomada de decisão, que nada mais é do que o uso explícito de conhecimento (REZENDE, 2003).

O conhecimento pode ser gerado de diversas maneiras. De acordo com o problema a ser solucionado, é necessária a aquisição de determinados conhecimentos, advindos de diferentes fontes de informação. O homem, por meio das características do seu pensamento, aborda de maneira diferenciada as inúmeras demandas diárias por decisões, seja utilizando da lógica, da cognição, de processos associativos, ou do raciocínio lateral. A tentativa de transferir estes processos da mente humana para o computador, desenvolvendo sistemas inteligentes que simulam o processo humano de decisão, configura-se na área de pesquisa denominada inteligência artificial (REZENDE, 2003). Entretanto, a análise dos dados, a sua correlação com a realidade de cada organização e a tomada de decisão que deve seguir à descoberta da informação, dependem do fator humano (SIGULEM *et al.*, 1998).

No entanto, o aprendizado de dados, informações e conhecimentos não deve

¹⁷ Eficácia, neste contexto, significa “usar o conhecimento com uma taxa de sucesso que seja aceitável, de forma consistente, sem conflitos” (CARVALHO, 1999, p.).

¹⁸ Eficiência, neste contexto, significa “não somente ter condições de resolver o problema, mas também resolvê-lo de forma rápida e adequada, deduzindo as soluções mais prováveis e determinando as informações mais relevantes” (CARVALHO, 1999, p.).

se limitar à percepção destes fenômenos a fim de serem implementados apenas momentaneamente, mas também conduzir à tomada de decisão futura (CARVALHO, 1999). Desta forma, este aprendizado pode ocorrer mediante algumas estratégias como: a implantação direta do conhecimento ou aprendizado por hábito; o aprendizado por instrução; por dedução; por analogia ou por indução (CASTIÑERA *apud* CARVALHO, 1999).

Os dois métodos mais significativos de aprendizagem são a indução e a dedução. O primeiro é a maneira mais utilizada pela mente para produzir conhecimento novo. Uma vez que são fornecidos exemplos com objetos pertencentes à uma classe pré-definida, e outros com objetos a ela não pertencentes, o aprendiz¹⁹ verifica os conjuntos de objetos e reconhece similaridades, a fim de encontrar uma descrição da classe que englobe os exemplos pertencentes (positivos) e não pertencentes (negativos). Já no método dedutivo, as deduções são feitas a partir de uma premissa maior para uma menor, e também por meio de generalizações e especializações, até que se alcance o critério previamente exigido (CARVALHO, 1999). Exemplos das duas estratégias de aprendizado podem ser visualizados no Quadro 1.

Quadro 1 - Exemplos dos modelos de aprendizagem Indutivo e Dedutivo

Modelo Indutivo	Modelo Dedutivo
– enunciados específicos com possibilidade de generalização	– parte de uma premissa maior para uma premissa menor
Microalbuminúria é exame para detectar diabetes melitus.	O procedimento de Revascularização do Miocárdio é considerado de alto risco no tratamento de doenças cardiovasculares.
Diabetes melitos predispõe à doença cardiovascular.	Homens que foram submetidos ao exame de hemoglobina glicosilada foram também submetidos à Revascularização do Miocárdio.
Logo, usuários que realizam o exame de Microalbuminúria podem estar predispostos para doenças cardiovasculares.	Logo, homens que foram submetidos ao exame de hemoglobina glicosilada podem ser de alto risco para doenças cardiovasculares.

Tais modelos de aprendizagem podem influenciar sobremaneira o modo como seres humanos e máquinas desempenham a análise dos dados para gerar informações e conhecimento para tomada de decisão. Os algoritmos computacionais

¹⁹ Aprendiz – é o agente que aprende. Pode ser um indivíduo ou um sistema computacional (CARVALHO, 1999).

que trabalham como mineração de dados possuem na sua concepção estes modelos de aprendizagem.

Por sua vez, em relação aos dados básicos que devem constar nas bases de dados de sistemas de informação de organizações de serviços de saúde, a Organização Panamericana de Saúde (PAHO) preconiza aqueles relacionados à estrutura, processo e resultados advindos da prestação de cuidados em saúde e de atividades de prevenção (PAHO, 1999).

Em relação aos dados necessários, planejados, coletados ou produzidos na saúde, alguns fatores devem ser considerados: (a) questões específicas da organização de saúde; (b) utilização operacional da informação pelos responsáveis pela tomada de decisão clínica, educacional e gerencial; (c) opções quanto ao uso de dados novos a serem gerados ou o uso dos dados já *existentes*, tanto de fontes **externas** como de fontes **externas** ao sistema; (d) nível de agregação das informações; e (e) expectativas com relação à qualidade e adequação das informações resultantes, diretamente ligadas à viabilidade da coleta e produção adequada da informação (PAHO, 1999).

Por sua vez, a Organização Panamericana de Saúde (PAHO, 1999) preconiza que os dados individuais dos usuários devem versar sobre três aspectos, coletados de cada usuário: o de identificação, o administrativo, e o clínico.

Os dados de identificação devem estar reunidos sob um código único de identificação, para que possam ser identificados o sexo, a idade, o nome e o endereço do usuário. Os dados administrativos devem contemplar o tipo de cobertura contratual e o número do seu contrato. Dentre os procedimentos e normas, devem ser estabelecidos os protocolos aplicáveis a cada usuário, de acordo com o seu perfil, demanda²⁰ e necessidade²¹ de serviços de saúde.

Dentre os dados de interesse clínico, devem-se registrar as doenças pré-existent, o risco, medidas antropométricas, diagnósticos, testes de desempenho

²⁰ Demanda por serviços de saúde – possui uma definição médica e econômica. A médica traduz-se na quantidade de cuidados médicos que as pessoas usariam se não houvesse barreiras para o acesso a esses cuidados (custo, conveniência, medo, localização dos prestadores de serviço). A definição econômica é a quantidade de cuidados que é comprada a determinado preço (JEKEL *et al.*, 2005).

²¹ Necessidade de serviços de saúde – geralmente considerada como um julgamento profissional (JEKEL *et al.*, 2005).

clínico e deficiências. Devem também ser registrados, dados referentes ao processo de assistência à saúde, como consultas, atendimentos hospitalares, procedimentos como exames, tratamentos, achados clínicos, e as percepções dos usuários acerca do seu problema de saúde. Ademais, devem ser contemplados os recursos utilizados nos procedimentos clínicos e a demanda por outros recursos de apoio como fisioterapia, fonoaudiologia, órteses, entre outros (PAHO, 1999).

Dados sociais, como renda salarial, ocupação, estrutura familiar e dados étnicos são considerados importantes, bem como o grau de conhecimento do usuário acerca do seu problema de saúde, satisfação com a assistência recebida, opiniões passadas e atuais levantadas em pesquisas sobre as percepções e experiências com os prestadores de saúde (PAHO, 1999).

A partir do levantamento desses dados dos usuários, é possível construir indicadores que podem ser categorizados de acordo com: a área temática da informação à qual estão relacionados; a dimensão do sistema de saúde a ser avaliado; ou à situação de saúde que se pretende representar (PAHO, 1999). Alguns indicadores comumente utilizados na área da saúde podem ser visualizados no Quadro 2, que os apresenta por área temática e suas descrições.

Quadro 2 - Demonstrativo das descrições de indicadores de saúde por área temática

ÁREA TEMÁTICA	DESCRIÇÃO DOS INDICADORES
Relacionados à condição de saúde.	Moradia, nível socioeconômico e educacional, saneamento básico, coleta de lixo, contato com animais.
Relacionados às necessidades de saúde.	Privação socioeconômica, situações de saúde-doença e de risco.
Indicadores de promoção da saúde e prestação de cuidados.	Estrutura (recursos e prestadores), o processo de assistência, e resultados.
Indicadores de acesso dos usuários, cobertura e utilização do sistema de saúde.	Acesso, cobertura de recursos, de atividades e sua utilização. Alguns podem ser utilizados para análise da equidade.
Indicadores de avaliação da eficiência.	Utilizados para analisar o quão eficientemente os recursos são alocados entre áreas ou grupos populacionais, indicadores para análises de custo-benefício, desempenho ou produtividade dos recursos utilizados (devem ser comparados com padrões de comparação).
Indicadores de avaliação da efetividade.	Relacionados aos resultados das intervenções em saúde e seu impacto no status de saúde global de grupos populacionais.
Indicadores de qualidade da educação e assistência em saúde.	Relacionados à estrutura e ao processo de promoção e assistência à saúde, pela perspectiva técnica e científica.
Indicadores de satisfação dos usuários e dos profissionais de saúde.	Envolve condições físicas e humanas do ambiente e do processo de assistência à saúde.
Indicadores de participação social.	Grau de envolvimento de grupos da comunidade e outros parceiros.
Indicadores de aderência a programas de saúde específicos.	Relacionados aos serviços preventivos e curativos específicos.

Fonte: (PAHO, 1999).

O Anexo 1 apresenta uma tabela com elementos essenciais que devem estar presentes em bases de dados de sistemas de informação em saúde, elaborada pela Organização Panamericana de Saúde (PAHO, 1999). Acredita-se que, a partir da coleta adequada dos dados propostos nessa tabela, sistemas de informação poderão proporcionar as informações e conhecimento adequados da realidade de um sistema de saúde, contribuindo para o processo decisório em todos os níveis organizacionais.

Contudo, é importante salientar que é fundamental que os dados sejam trabalhados e cheguem ao gestor de saúde, numa performance passível de utilização nos processos decisório que permeiam suas atividades.

2.5 Processo de Descoberta de Conhecimento

A descoberta de conhecimento em bases de dados é uma área de pesquisa, que vem crescendo nos últimos anos devido à tecnologia inerente que possibilita a manipulação de grandes bases de dados, que podem conter padrões e tendências. Tal tecnologia, capaz de gerar conhecimento novo, poderia otimizar as decisões nos diversos tipos de negócios, trazendo benefícios à atividade operacional prática, social e econômica (REZENDE, 2005). Ao longo deste desenvolvimento serão apresentadas e discutidas algumas tecnologias utilizadas nos processos de descoberta de conhecimento em bases de dados.

Este processo pode ser definido como o processo realizado a partir dos dados de identificação não trivial de padrões válidos, novos, potencialmente úteis e compreensíveis (FAYYAD *et al.*, 1996).

O processo de mineração de dados é, por vezes, confundido com a descoberta de conhecimento em bases de dados, mas aquele é uma parte deste, o qual engloba uma série de etapas operacionais (GOLDSCHMIDT; PASSOS, 2005). O uso da descoberta de conhecimento em bases de dados se preocupa com: o entendimento e o preparo dos dados; a verificação; e, a aplicação do conhecimento descoberto. Já na mineração de dados a preocupação reside na extração dos padrões, propriamente dita (CIOS; KURGAN, 2002). A Figura 2 ilustra o processo de descoberta de conhecimento e de mineração de dados.

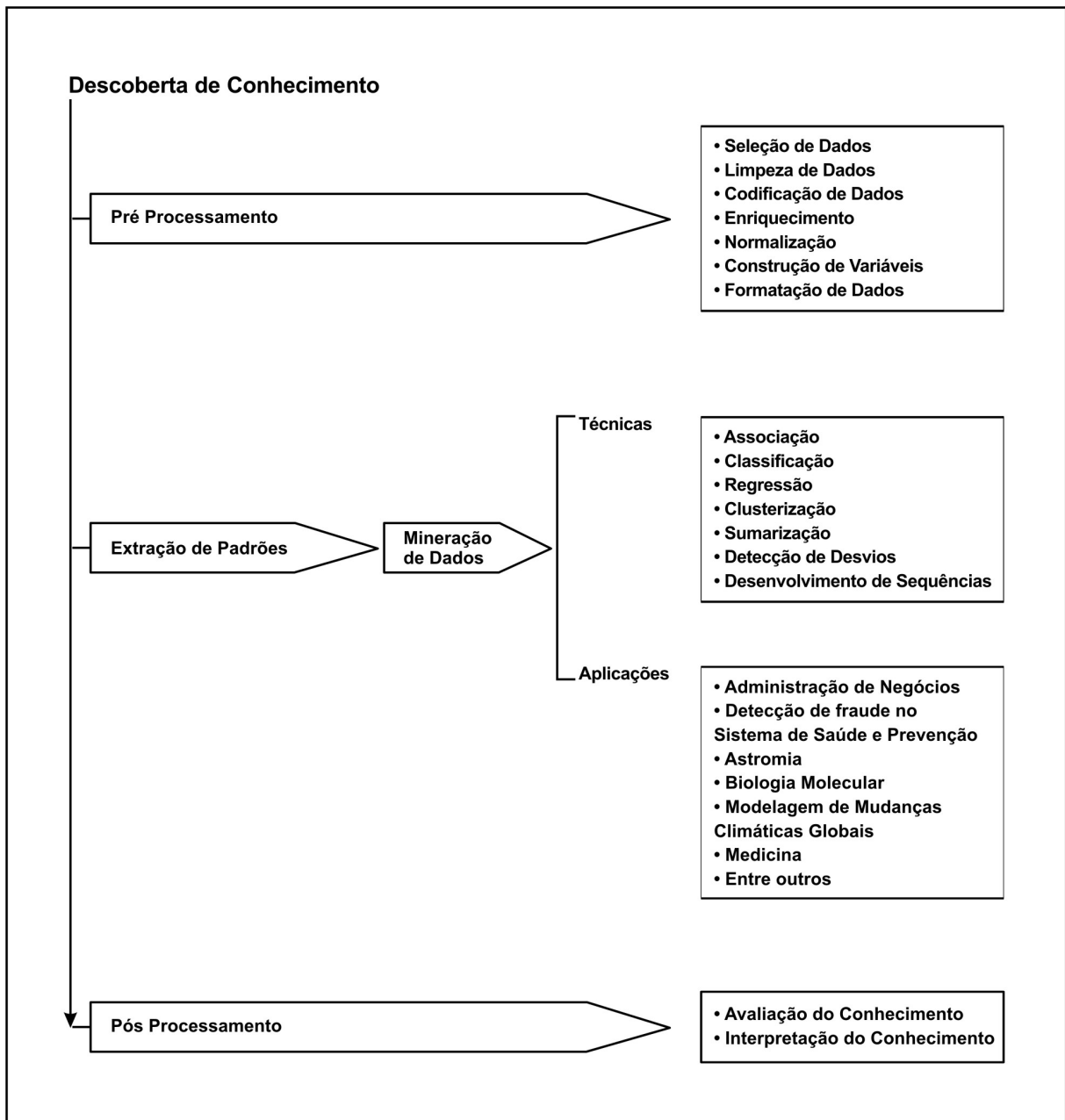


Figura 2 - Principais Componentes do KDD e DM e seus relacionamentos

Fonte: Adaptado de LEE *et al.* (2000).

A mineração de dados advém da estatística clássica, da inteligência artificial e do aprendizado de máquina²². Estas três técnicas se unem no estudo de dados para descobrir tendências e padrões úteis (OLIVEIRA, 2001; COUTINHO, 2006).

²² Aprendizado de máquina – “uma área da Inteligência Artificial cujo objetivo é o desenvolvimento de técnicas computacionais sobre o aprendizado bem como a construção de sistemas capazes de adquirir conhecimento de forma automática” (Rezende, 2003, p.89).

Entretanto, a mineração de dados “é o processo de extrair informação válida, *previamente desconhecida* e de máxima abrangência a partir de grandes bases de dados, usando-as para efetuar decisões cruciais” (COUTINHO, 2006, p.1). Já a estatística clássica, além de ser a base das técnicas de mineração de dados, desempenha um papel fundamental no pós-processamento, fase que deve acontecer após a extração dos padrões na mineração de dados. A estatística oferece uma análise objetiva e ao mesmo tempo mais ampla dos dados abstraídos a partir das técnicas de mineração de dados. Sem ela, os dados geralmente são probabilísticos, e refletem a realidade da base de dados na qual foram realizadas as técnicas de mineração, sem grande poder de generalização. Assim, a mineração de dados permite extrair informações novas.

A eficácia do processo de mineração de dados decorre da integração dos atores envolvidos neste processo, que devem estar presentes durante toda sua execução, a saber: (a) o especialista do domínio, o qual deve ser detentor de amplo conhecimento acerca do domínio da aplicação e fornecer apoio para a execução do processo; (b) o analista, o qual deve conhecer intimamente o processo de mineração de dados e extração de conhecimento; e (c) o usuário final, que é aquele que utiliza o conhecimento extraído para a tomada de decisões (REZENDE, 2005), no caso o profissional de saúde gestor.

2.5.1 Fases do Processo de Descoberta de Conhecimento

O processo de descoberta de conhecimento em bases de dados constitui-se de um conjunto de atividades realizadas continuamente e que compartilham o conhecimento advindo das bases de dados (CARVALHO, 1999). Pode ser subdividido em três etapas, a saber: o pré-processamento, a mineração de dados e o pós-processamento.

A primeira fase deste processo inicia-se com a definição do problema a ser solucionado. Os conceitos abordados nesta fase podem auxiliar o analista na escolha da melhor técnica de mineração a fim de alcançar os objetivos traçados, bem como no ajuste de parâmetros do processo durante a fase de extração de padrões, ou mineração de dados, propriamente dita (REZENDE, 2005).

2.5.1.1 Pré-processamento

O sucesso da tarefa de mineração de dados depende do esforço empregado na fase de pré-processamento dos dados. Isso pode ser particularmente necessário em projetos na área da saúde, nos quais pode-se verificar que os prestadores de serviços de saúde se preocupam com o preenchimento de códigos que garantem os pagamentos, em detrimento de outros dados que dizem respeito aos usuários, considerados sem importância para fins financeiros (GILLESPIE, 2004).

Durante a fase de pré-processamento são realizadas funções de captação, organização, tratamento e preparação dos dados para a mineração dos mesmos. Estas funções estão relacionadas à possível correção de dados errados até o ajuste da sua formatação. Dentre as possíveis técnicas de pré-processamento, têm-se a seleção de dados, limpeza, codificação, enriquecimento, normalização, construção de variáveis, partição do conjunto de dados e formatação dos dados (GOLDSCHMIDT; PASSOS, 2005).

a) Seleção dos Dados

Trata-se da identificação das informações que devem ser consideradas para o processo de descoberta de conhecimento em bases de dados. Muitas vezes, para que a manipulação dos dados não interfira nas atividades operacionais da organização, são feitas cópias ou são utilizadas estruturas de *Data Warehouse*²³.

Muitas operações podem ser realizadas com a finalidade de selecionar os dados, entre elas, as operações de: junção direta; junção orientada; redução de dados horizontal e vertical; eliminação direta de variáveis; e, análise de componentes principais e redução de valores.

Goldschmidt e Passos (2005) definem que a função de seleção de dados chamada de **junção direta** é realizada

[...] quando todas as variáveis e registros da base de dados são incluídos numa nova tabela a ser trabalhada, sem uma análise crítica quanto a quais variáveis e quais casos podem realmente contribuir para o processo de

²³ Data Warehouse – conjunto de dados baseado em assuntos, integrado, não volátil, variável em relação ao tempo, e destinado a auxiliar em decisões de negócios. Diferem das bases de dados operacionais no sentido que estas podem ser definidas como base de dados clássicos, que contêm informações detalhadas a respeito do negócio em nível transacional. Naqueles, os dados estão voltados a fornecer informações para os níveis gerencial e estratégico das empresas (GOLDSCHMIDT; PASSOS, 2005).

descoberta de conhecimento em bases de dados (GOLDSCHMIDT; PASSOS, 2005, p.26).

Já a seleção de dados por **junção orientada** é uma técnica que permite definir variáveis que realmente apresentam potencial para responder à questão norteadora, aos objetivos do estudo e ao problema de mineração de dados (GOLDSCHMIDT; PASSOS, 2005).

A **redução horizontal de dados** é caracterizada pela escolha de casos mediante amostragem aleatória, eliminação direta de casos, segmentação da base de dados e agregação de informações, esta com o objetivo de “reunir os dados de forma a reduzir o conjunto de dados original” (GOLDSCHMIDT; PASSOS, 2005, p. 28).

Já a **redução vertical de dados** é considerada uma importante técnica de pré-processamento, também chamada de redução de dimensão. Seu objetivo é descobrir um conjunto mínimo de atributos de maneira que a informação original seja mantida. Em casos de classificação, existem duas abordagens possíveis de serem realizadas, a independente de modelo ou *filter*, e a dependente de modelo, ou *wrapper*. Na primeira, a seleção de atributos é realizada sem ser considerado o algoritmo de mineração de dados que será aplicado; na segunda, o algoritmo de mineração de dados é aplicado a cada conjunto de atributos a fim de se avaliar os resultados gerados (GOLDSCHMIDT; PASSOS, 2005).

Ainda para a escolha dos atributos, tanto se a escolha tiver sido pela abordagem *filter* como pela *wrapper*, existem três estratégias clássicas e simples para norteá-la, a saber: (a) seleção seqüencial para frente, ou *forward selection*; (b) seleção seqüencial para trás, ou *backward selection*; e, (c) combinação das estratégias de seleção seqüencial para trás e para frente (GOLDSCHMIDT; PASSOS, 2005).

Os autores afirmam que existem algoritmos, que também podem ser aplicados para selecionar atributos em problemas de classificação, e então, a partir dos dados, construir árvores de decisão, de maneira que os atributos que não apareçam na árvore não são considerados como importantes para a resolução do problema (GOLDSCHMIDT; PASSOS, 2005).

A **eliminação direta de atributos** pode ser realizada a partir de duas heurísticas²⁴: aqueles que são atributos identificadores e aqueles que apresentam valores constantes na base de dados (GOLDSCHMIDT; PASSOS, 2005).

A **análise de componentes principais** é realizada a partir da normalização dos dados, para que os atributos pertençam a uma mesma faixa de valores; e, da sua ordenação em ordem decrescente de variância. Deste modo, o conjunto de dados pode ser reduzido a partir dos componentes que apresentam menor variância (GOLDSCHMIDT; PASSOS, 2005).

A técnica de **redução de valores** pode estar relacionada a valores nominais ou contínuos, e apresenta-se como uma opção à técnica de eliminação de atributos, pela redução vertical de dados. Visa à redução do número de valores distintos em determinados atributos, o que pode melhorar o desempenho de alguns algoritmos de mineração de dados (GOLDSCHMIDT; PASSOS, 2005). Dentre as operações, pode-se citar a clusterização.

b) Limpeza

Quando se trabalha com bases de dados reais, é comum deparar-se com dados incompletos, ruidosos ou inconsistentes. Dados incompletos são aqueles cujas tabelas apresentam dados ausentes sobre alguma variável, ou sem detalhes. Dados ruidosos são aqueles cujos valores são diferentes do padrão normal esperado. Dados inconsistentes apresentam discrepância semântica entre si (GOLDSCHMIDT; PASSOS, 2005).

Assim, se estas ocorrências são comuns numa base de dados, é fundamental que o analista dispense tempo para trabalhar nesta fase do processo de descoberta de conhecimento. A qualidade dos modelos de conhecimento depende diretamente da qualidade dos dados a partir dos quais ele é abstraído (GOLDSCHMIDT; PASSOS, 2005).

A fase de limpeza tem como objetivo primordial a correção da base de dados, a partir da verificação da consistência dos dados, correções de possíveis erros, preenchimento ou eliminação de valores redundantes ou desconhecidos, e eliminação de valores que estão fora do intervalo aceitável. Desta forma, procura-se

²⁴ Heurística – é o “conhecimento de regras práticas que auxiliam a reduzir o esforço de busca por soluções” (REZENDE, 2003, p.16).

diminuir o número de consultas a ser realizado pelo algoritmo de mineração de dados, a fim de potencializar o seu desempenho e otimizar os resultados (GOLDSCHMIDT; PASSOS, 2005).

Outras formas de atenuar o problema das informações ausentes são: o preenchimento manual de valores; o preenchimento com valores globais constantes; o preenchimento com medidas estatísticas; e, o preenchimento com métodos de mineração. Para a correção de inconsistências podem ser realizadas a **exclusão de casos** e a **correção de erros**. Na exclusão de casos são eliminados os casos e as variáveis cujos valores divergem daqueles esperados para a situação. Já a correção de erros pode ser realizada tanto manualmente como de maneira automática, substituindo os valores errados por meio de comandos de atualização de dados em bases de dados relacionais (GOLDSCHMIDT; PASSOS, 2005).

c) Codificação

Também chamada de transformação, a codificação trata da apresentação dos dados para que seja possível a sua entrada no algoritmo de mineração de dados (GOLDSCHMIDT; PASSOS, 2005).

Dentre as formas de codificar os dados, encontram-se os tipos: categórica-numérica; e numérica-categórica. Na primeira, variáveis categóricas são codificadas em numéricas. Na segunda, as variáveis numéricas são codificadas em categóricas, sendo possível também que sejam adotados intervalos para a codificação, o que se chama de normalização (GOLDSCHMIDT; PASSOS, 2005).

No entanto, outras transformações possíveis são: resumo, quando se faz necessário o agrupamento dos dados para formar resumos; e de tipo, quando, por exemplo, a variável data deve ser transformada em outro tipo de representação para ser aplicado ao algoritmo de maneira adequada (GOLDSCHMIDT; PASSOS, 2005).

d) Enriquecimento

Durante esta fase, é possível tentar enriquecer a base de dados sob a qual se trabalha para melhorar a qualidade dos dados a fim de fornecer novos elementos que contribuam com a descoberta de conhecimento. Isso pode ocorrer por meio de pesquisa junto às fontes de dados, bem como por meio da utilização de bases de dados externas, sempre com o auxílio do especialista no domínio de aplicação (GOLDSCHMIDT; PASSOS, 2005).

Exemplo de enriquecimento em utilizações de bases de dados na saúde seria a agregação de dados dos Correios com os códigos do CEP, somados à utilização de dados de empresas de água e esgoto a fim de identificar a disponibilidade de água tratada e esgoto nos domicílios dos usuários pesquisados.

e) Normalização

Nesta fase, ocorre o ajuste de intervalos de valores a fim de não produzir viés na aplicação de alguns algoritmos de mineração adotados, de modo que algumas variáveis apresentem escalas de valores maiores que outras. É possível ajustar os valores das variáveis de maneira que pertençam a intervalos menores (GOLDSCHMIDT; PASSOS, 2005).

Os tipos de normalização são: a normalização linear, a normalização por desvio padrão; a normalização pela soma dos elementos; a normalização pelo valor máximo dos elementos; e, a normalização por escala decimal.

f) Construção de Variáveis

Trata-se da construção de novas variáveis a partir daqueles já existentes na base de dados (GOLDSCHMIDT; PASSOS, 2005). Por exemplo, cria-se a variável “IDADE” a partir do variável “DATA_NASC” (data de nascimento) e da variável “DATA_ATEND” (data de atendimento), a fim de determinar a idade do usuário no momento do registro do atendimento.

g) Partição do Conjunto de Dados

Faz-se necessária a divisão do conjunto de dados quando o algoritmo de mineração de dados trabalha com um conjunto de treinamento e um conjunto de testes (GOLDSCHMIDT; PASSOS, 2005).

h) Formatação dos dados

Apesar de não ser citada por Goldschmidt e Passos (2005), a fase de formatação dos dados é imprescindível a fim de que o algoritmo de mineração de dados possa ler e identificar os padrões entre os dados. A partir da definição do problema de pesquisa e, conseqüentemente da técnica a ser utilizada, cada algoritmo demanda que os dados estejam em uma determinada formatação, de maneira a tornar possível sua aplicação.

2.5.1.2 Mineração de Dados

A mineração de dados consiste na busca efetiva de conhecimentos novos e úteis numa base de dados. Para tanto, são aplicados algoritmos, que são fundamentados em técnicas, e atuam nesta busca a partir de determinados paradigmas e exploram os dados. Nesta fase do processo é realizada a extração de padrões (REZENDE, 2003).

Na área da saúde, considera-se que a aplicação de técnicas de mineração de dados pode permitir aos prestadores de saúde garantir uma assistência de qualidade planejada, otimizar custos e salvar vidas (SIAU, 2003). São aplicados algoritmos que fazem a busca efetiva por conhecimentos implícitos e úteis a serem utilizados na resolução de um problema já estabelecido.

A mineração de dados não faz hipóteses sobre os dados, pelo contrário, procura relacionamentos e correlações escondidos entre os dados (GILLESPIE, 2004). Juntamente com os especialistas no domínio da aplicação, devem ser estabelecidos metas, objetivos e restrições da aplicação dos algoritmos. A participação destes deve acontecer desde a definição do problema e em todo o processo, provendo o conhecimento específico necessário aos analistas.

O que determina a técnica de mineração de dados a ser utilizada é o tipo de **tarefa de KDD**²⁵ a ser realizada a partir do problema a ser resolvido e a linguagem de representação dos padrões a serem encontrados (REZENDE, 2003).

Entre as técnicas de mineração de dados, têm-se: a descoberta de associação; a classificação; a regressão; a clusterização; a sumarização; e, a descoberta de seqüências, as quais podem ser adaptadas ou compostas, originando novas tarefas (GOLDSCHMIDT; PASSOS, 2005). As tarefas voltadas à extração de padrões podem ser agrupadas em atividades preditivas ou descritivas, como mostra a Figura 3.

²⁵ Tarefa de KDD – equivale a uma operação de KDD que pertença à etapa de Mineração de Dados, podendo ser primárias ou compostas (GOLDSCHMIDT; PASSOS, 2005).

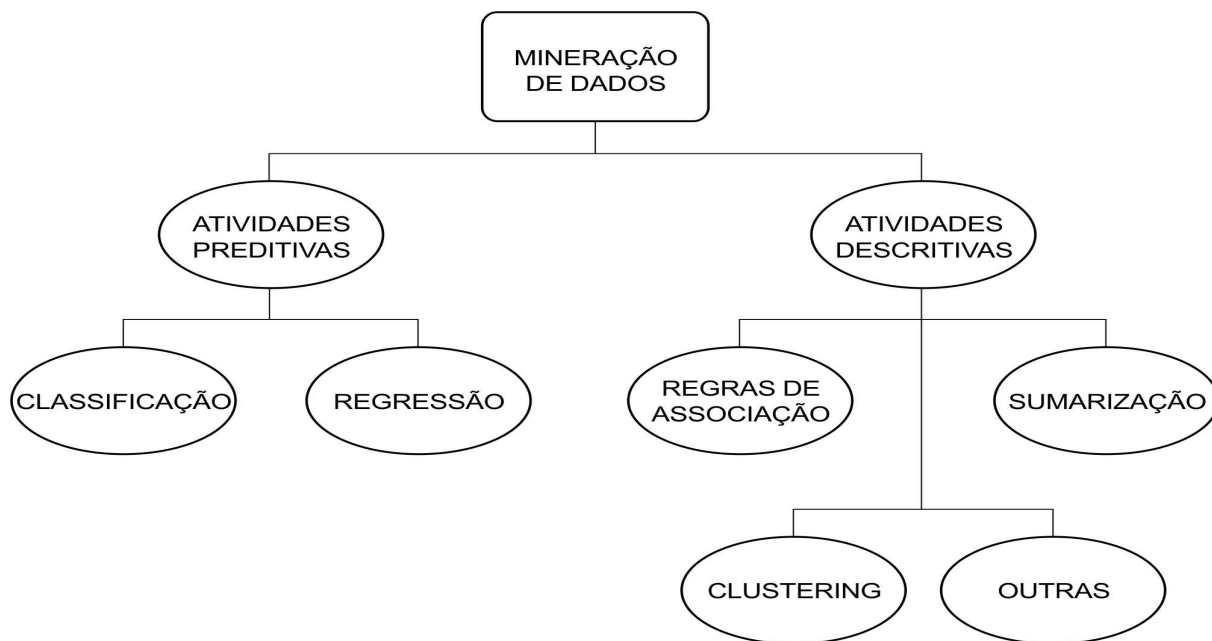


Figura 3 - Representação esquemática das tarefas de Mineração de Dados

Fonte: Rezende (2005).

De acordo com Rezende (2003), a mineração de dados preditiva tem como objetivo a generalização de exemplos, com classe conhecida, a fim de reconhecer a classe de um novo exemplo. As tarefas de mineração que lidam com atividades preditivas são: a classificação, para atributos categóricos; e a regressão, para atributos numéricos.

Um algoritmo de predição poderia ser utilizado com a base de dados desta pesquisa para que, a partir de dados históricos de utilização de procedimentos ou conjuntos de procedimentos, relacionados ao seu custo e utilização, se pudesse aprender com os dados históricos para predizer casos semelhantes no futuro. Para tanto, o pré-processamento da base de dados deve ser voltado para a predição. Ainda, as classes deveriam ser de conhecimento do analista e dos especialistas envolvidos na pesquisa.

Já a mineração de dados descritiva tem como objetivo a descoberta de comportamentos intrínsecos ao conjunto de dados, considerando que estes dados não apresentam classe específica. As tarefas de mineração que lidam com atividades descritivas são clusterização, regras de associação e sumarização, as quais serão explicitadas adiante.

A estratificação dos usuários em de baixo, médio e de alto custo, de acordo com os procedimentos realizados, pode ser um exemplo de mineração de dados descritiva que poderia ser aplicada na base de dados desta pesquisa, por meio da clusterização.

Tendo em vista o exposto, fica claro que não existe um único algoritmo que seja utilizado para todas as tarefas de descoberta de conhecimento em bases de dados. Assim, às vezes se faz necessária a utilização de dois ou mais algoritmos para o alcance dos resultados esperados para a resolução de um determinado problema. O Quadro 3 apresenta as técnicas de mineração de dados, suas principais características no sentido das funções possíveis de serem executadas para a resolução de um determinado problema, e as ferramentas referentes à cada técnica, de acordo com o referencial de Goldschmidt e Passos (2005), e Rezende (2005).

Como pode se observar no Quadro 3, quando se pretende buscar itens freqüentes e que ainda ocorram simultaneamente em transações de bases de dados, a técnica de mineração a ser utilizada é a descoberta de associações. Esta técnica seria a adequada para o alcance de um dos objetivos deste trabalho, a identificação dos eventos que, associados, indicam um usuário elegível para programas de gerenciamento de caso. Assim, existem inúmeras ferramentas que apresentam esta técnica nos seus aplicativos, tais como o Apriori, *Clementine* e *Intelligent Miner*.

A classificação é a técnica de mineração de dados a ser utilizada quando se pretende descobrir uma função que mapeie um conjunto de registros em um conjunto de rótulos pré-definidos - as classes. Quando aplicada a novos registros, prevê a classe em que tais registros se enquadram, por meio dos aplicativos WEKA, *Pollyanalyst* e *Xpert/RuleMiner*, por exemplo, que utilizam de técnicas como redes neurais, algoritmos genéticos e lógica indutiva.

Quadro 3 - Descritivo das Técnicas de Mineração de Dados, respectivas características e ferramentas

Técnicas de DM	Características	Tecnologia – Aplicativos
Descoberta de Associação	Busca por itens que freqüentemente ocorram de forma simultânea em transações na base de dados.	Aplicativos: Apriori, GSP, DHP Pollyanalyst, Magnum Opus, Xpert/Rule Miner, Data Mite, Oracle 9i Data Mining, Mineset, WEKA, Intelligent Miner, Clementine
Classificação	Descobre uma função que mapeie um conjunto de registros em um conjunto de rótulos pré-definidos (classes). Assim, esta função pode ser aplicada a novos registros de forma a prever a classe em que tais registros se enquadram.	Tecnologia: Redes Neurais, Algoritmos Genéticos, Lógica Indutiva Aplicativos: Pollyanalyst, Xpert/Rule Miner, Microsoft Data Analyzer 2002, Oracle 9i Data Mining, Darwin, Mineset, WEKA, Intelligent Miner, See5, Clementine, Data-Miner Software Kit, MLC++
Regressão	Mapeia os registros de uma base de dados, mas apenas para atributos numéricos	Tecnologia: Estatística, Redes Neurais Aplicativos: Pollyanalyst, Darwin, Mineset, WEKA, Cubist, Data-Miner Software Kit, MLC++
Clusterização	Separa os registros de uma base de dados em subconjuntos (<i>clusters</i>), de maneira que os elementos de um <i>cluster</i> compartilhem de propriedades comuns a outros <i>clusters</i> . Seu objetivo é maximizar a similaridade intracluster e minimizar a similaridade intercluster.	Aplicativos: K-Means, K-Modes, K-Prototypes, Kohonen Pollyanalyst, Xpert/Rule Miner, Microsoft Data Analyzer 2002, Darwin, Intelligent Miner Mineset, Clementine, MLC++
Sumarização	Procura identificar e indicar características comuns em bases de dados.	Tecnologia: Lógica Indutiva, Algoritmos Genéticos Aplicativos: Pollyanalyst, Intelligent Miner
Detecção de Desvios	Identifica registros da base de dados cujas características não atendam aos padrões considerados normais ao contexto (<i>outliers</i>)	Tecnologia: Estatística

Fonte: Goldschmidt e Passos (2005), e Rezende (2005).

Quando se torna necessário separar os registros de uma base de dados em subconjuntos ou *clusters*, de maneira que os elementos de um *cluster* compartilhem

de propriedades comuns a outros *clusters*, a técnica de mineração de dados a ser escolhida é a **clusterização**, por meio dos aplicativos *K-Means*, *K-Modes*, *K-Prototypes*, entre outros.

Por meio da lógica indutiva e dos algoritmos genéticos, quando se tem o objetivo de identificar e indicar características comuns em bases de dados está se utilizando a técnica de mineração de dados de sumarização. Aplicativos que executam essa tarefa são o *Pollyanalyst*, *Intelligent Miner*.

Já a detecção de desvios, por meio da estatística, procura identificar registros da base de dados cujas características não atendam aos padrões considerados normais ao contexto, os chamados *outliers*.

A **descoberta de associação** relaciona-se à busca de regras de associação²⁶ entre dados, de modo a descobrir, numa base de dados, conjuntos de dados que ocorram simultânea e freqüentemente. Primeiramente definida por Agrawal *et al.* em 1993, descrevia um experimento no qual se trabalhava com uma tabela *basket*, cuja característica é a de que a quantidade de dados nesta tabela não é limitada pelo número de variáveis da relação (AGRAWAL *et al.*, 1993). A tabela utilizada descrevia transações de compra de produtos em um estabelecimento, e cada transação de compra relacionava-se a todos os produtos adquiridos por um cliente durante um período estabelecido. Esta técnica de mineração de dados tem sua aplicação muito difundida na área de marketing, na qual são exploradas as associações entre produtos vendidos (OLIVEIRA, 2001).

A representação matemática das regras de associação é na forma $Y \leftarrow X$; sendo $[X]$ e $[Y]$ conjuntos de dados, tais que $[X] \cap [Y] = \emptyset$ (GOLDSCHMIDT; PASSOS, 2005). Carvalho (2005) comenta que o significado de regras como esta é que os conjuntos de dados $[X]$ e $[Y]$ ocorrem freqüentemente juntos num mesmo registro. As regras também podem ser lidas na forma “SE X; ENTÃO Y”.

Um exemplo desta regra, considerando uma determinada carteira de usuários de uma operadora de saúde, pode ser:

Ecocardiograma \leftarrow Cateterismo Cardíaco (11%, 78%)

²⁶ A descoberta de associação também pode ser chamada de regra de associação ou regras associativas (GOLDSCHMIDT; PASSOS, 2005).

Nesta regra, 11% dos usuários sofreram um procedimento de cateterismo cardíaco, sendo que destes, 78% também foram submetidos a um ecocardiograma.

A frequência de uma associação reflete o total de vezes que o conjunto de dados ocorre em relação ao número total de casos existentes na base de dados. Esta associação é considerada freqüente caso seja maior que o **suporte mínimo** (frequência mínima) estipulado pelo analista. O suporte permite identificar as associações que ocorrem em uma quantidade significativa em relação às demais. O valor de 11% seria o suporte do exemplo acima.

É importante salientar que, neste trabalho, a leitura não foi realizada como proposto por Agrawal (1993), por se tratar de uma leitura mais restritiva das regras abstraídas pelo Apriori. Borgelt (1996) declara que, para Agrawal (1993), o suporte de uma regra pode ser definido como o número relativo de casos nos quais a regra ocorre²⁷, e já na sua concepção do algoritmo Apriori, considerou-se a definição de que o suporte de uma regra é igual ao número relativo de casos no qual esta regra é aplicável (BORGELT, 1996, p.2).

Outro valor estipulado para cada aplicação é o da **confiança mínima**. Esta medida assegura o quanto a ocorrência do antecedente da regra determina a ocorrência do conseqüente da regra. Para o exemplo dado acima, a confiança seria de 78%.

Estes parâmetros de suporte e confiança, respectivamente, restringem a quantidade de regras que serão geradas e descrevem a sua qualidade (OLIVEIRA, 2001).

De acordo com Goldschmidt e Passos:

A tarefa de Descoberta de Associações pode ser definida formalmente como a busca por regras de associação freqüentes e válidas em um base de dados, a partir da especificação dos parâmetros de confiança e suporte mínimos (GOLDSCHMIDT; PASSOS, 2005, p.62).

Já com relação a **tarefas de classificação**, esses mesmos autores postulam, sucintamente, que são tarefas que procuram “descobrir uma função que mapeia um conjunto de rótulos categóricos predefinidos, chamados de classes. Uma vez descoberta, pode ser aplicada a novos registros de forma a prever a classe em que tais registros se enquadram” (GOLDSCHMIDT; PASSOS, 2005, p. 13).

²⁷ Considerando a regra $A \text{ e } B \rightarrow C$, o suporte deve ser lido da seguinte maneira: a presença do item C segue a presença dos itens A e B.

Esta técnica é utilizada, segundo Carvalho (1999, p.31), para resolver problemas do tipo “descobrir regras para a previsão das classes de novos exemplos do domínio”, por meio da indução, sendo que esta predição está relacionada aos valores que assumem certas variáveis dependentes.

A autora lembra ainda que, ao utilizar dados da experiência passada para fazer as descobertas é fundamental que exista um conjunto destes exemplos, disponível para que um novo exemplo ou conjunto de exemplos seja classificado. Para tanto, os dois conjuntos de dados devem apresentar a variável dependente nas diversas classes (CARVALHO, 1999).

Considerando a área da saúde, Blair (2004) comenta ser interessante a capacidade de se separar e identificar indivíduos numa população, que em poucos anos a partir de hoje desenvolverão uma doença crônica. A autora refere ser importante salientar, também, que a intervenção mediante programas de gerenciamento pode influenciar o comportamento do indivíduo de maneira a reduzir a incidência de doenças, antes que elas se tornem crônicas e de alto custo.

Hinkle *apud* Carlson (2003, p.4) afirma que o trabalho pesado do gerenciamento de casos tem início depois de realizado o trabalho computacional, mas que este se trata de “uma ferramenta superior que melhor identifica usuários para programas de gerenciamento da saúde apropriados”.

Organizações de saúde nos Estados Unidos utilizam abordagens preditivas nos seus programas computacionais como base para seus programas de gerenciamento. Por meio dos dados demográficos e de solicitações de pagamento, quantificam o impacto das doenças, avaliam a utilização relativa de seus recursos de assistência à saúde e as tendências de utilização dos serviços a fim de prever futuras demandas por serviços (KNABEL; LOUWERS, 2004).

Técnicas de inteligência artificial, como regras, redes bayesianas ou redes neurais utilizam conjuntos de dados administrativos (consultas, compras em farmácias, dados de identificação) que associados com avaliações de risco em saúde, prevêm o risco de uma determinada pessoa desenvolver condições crônicas que futuramente irão onerar o sistema de saúde (BLAIR, 2004).

Embora seja relevante o fato de a tecnologia auxiliar na manipulação e avaliação de dados a fim de predizer o custo futuro da saúde de indivíduos, é igualmente importante estar ciente das questões éticas envolvidas neste processo, a fim de não penalizar os próprios usuários por suas condições de saúde (BLAIR,

2004). Neste sentido, é fundamental estabelecer normas para segurança por escrito, implementá-las e monitorá-las junto aos usuários do sistema de informação, tais como profissionais da saúde e pessoal administrativo. Deve haver garantia do controle do acesso²⁸, da integridade²⁹, da proteção física³⁰ e da confidencialidade³¹ dos dados e informações dos usuários (PAHO, 1999).

Infelizmente, por mais sofisticada que seja a técnica de inteligência artificial utilizada numa ferramenta computacional, ela por si só não é capaz de gerar mudança de comportamento. Para tanto, é necessário que se incluam protocolos ou diretrizes de intervenção e monitoração desses casos, para que os profissionais responsáveis pela orientação dos usuários o façam da maneira mais custo-efetiva possível, primando pela qualidade do atendimento ao usuário identificado como portador de algum grau de risco (BLAIR, 2004).

Uma ferramenta que trabalha com predição constrói um modelo por meio do exame dos relacionamentos históricos entre os dados conhecidos e a variável de interesse. Uma vez conhecidos os relacionamentos históricos e o modelo construído, a ferramenta é aplicada e utilizada para prever o comportamento da variável (COUSINS *et al.*, 2002)

Os classificadores, segundo Brieman *et al.* (1984), devem seguir dois propósitos: o primeiro seria o de predição de um valor e o segundo, da compreensão da relação entre os atributos previsores e a classe. Para tanto, é fundamental que o algoritmo não apenas classifique, mas que também explicita o conhecimento de forma inteligível. A fim de avaliar a qualidade das regras geradas são utilizados como critérios “a precisão preditiva, compreensibilidade e o grau de interesse do conhecimento descoberto” (CARVALHO, 2005, p.8).

Os Métodos de KDD, mais particularmente os métodos de mineração de dados, são definidos por Goldschmidt e Passos (2005) como:

[...] operadores definidos a partir de precondições e efeitos. A precondição é um predicado que estabelece um requisito a ser cumprido antes da execução do método de KDD. Já o efeito é também um predicado que se

²⁸ Controle do Acesso – a prevenção de entradas não autorizadas nos recursos informacionais (PAHO, 1999)

²⁹ Controle da Integridade – a prevenção de modificações não autorizadas das informações (PAHO, 1999)

³⁰ Proteção Física – a proteção dos dados ou dos equipamentos de processamento dos dados contra acidentes intencionais ou propositalis (PAHO, 1999)

³¹ Confidencialidade – a prevenção da revelação não autorizada das informações (PAHO, 1999)

refere a uma situação gerada após a aplicação do método (GOLDSCHMIDT; PASSOS 2005, p.83).

Entre os métodos de KDD pode-se citar: redes neurais; algoritmos genéticos; métodos baseados em instâncias; métodos estatísticos; métodos específicos; métodos baseados em indução de árvores de decisão; e, métodos baseados em lógica nebulosa (GOLDSCHMIDT; PASSOS, 2005).

Cada um dos métodos apresenta pontos fracos e fortes. A escolha por um ou outro método depende do tipo dos dados disponíveis e do objetivo da pesquisa. Alguns autores optam por aplicar todas as técnicas nos mesmos dados e assim verificar qual delas apresenta o melhor desempenho (CARLSON, 2003).

O algoritmo Apriori é um exemplo referenciado como um método específico, voltado para a descoberta de associações. Este algoritmo pode ser decomposto em duas etapas iterativas. A primeira é responsável por encontrar todos os conjuntos de itens freqüentes, ou seja, que satisfazem a condição de suporte mínimo. A segunda, parte da identificação dos itens freqüentes para a geração de regras de associação que satisfazem a condição de confiança mínima (GOLDSCHMIDT; PASSOS, 2005). Já o algoritmo C4.5³² é um exemplo de um método de indução de árvores de decisão.

Cousins *et al.* (2002) citam diversos artigos nos quais são realizados estudos na área da saúde utilizando métodos de mineração de dados, tais como os modelos de séries temporais, árvores de classificação, regressão linear e não linear, e redes neurais.

2.5.1.3 Pós-Processamento

Na fase de Pós-Processamento, o conhecimento extraído é analisado, interpretado e avaliado, de acordo com a sua utilidade na resolução do problema (REZENDE, 2005), por meio de simplificações, transformações de modelos de conhecimentos, organização e apresentação dos resultados (GOLDSCHMIDT; PASSOS, 2005).

Collazos *et al.* (2002) lembram que existem três tipos de resultados provenientes do processo de descoberta de conhecimento em bases de dados, a

³² C4.5 – algoritmo de classificação

saber: (a) confirmação do óbvio, quando os resultados anteriores e posteriores ao processo são os mesmos; (b) novo conhecimento, quando a partir da base de dados são descobertas informações não óbvias por meio de uma simples observação dos dados; e, (c) nenhuma relação significativa achada, porque não se encontrou nada de novo, ou porque as variáveis foram insuficientes.

Na etapa de pós-processamento ocorre a disponibilização dos conjuntos de padrões extraídos aos usuários finais. Após análise e processamento destes padrões, eles se tornam disponíveis para sua inclusão num sistema inteligente, ou para sua incorporação como apoio no processo decisório do próprio homem (REZENDE, 2003).

Muitas vezes, os algoritmos de mineração de dados descobrem um número excessivo de regras e padrões. Entretanto, é importante que o analista mostre ao usuário somente regras ou padrões relevantes à sua prática, para que o processo de descoberta de conhecimento em bases de dados seja reconhecido como uma boa ferramenta de apoio, e não complique ainda mais o processo decisório dos 'tomadores de decisão', inviabilizando a interpretação de padrões que poderiam ser interessantes (CARVALHO, 2005).

Para tanto, é fundamental a compreensão do conceito de medida de interesse a fim de que o processo de descoberta de conhecimento em bases de dados obtenha sucesso. Primeiramente, as medidas de interesse podem ser utilizadas após a mineração de dados para organizar ou filtrar os padrões descobertos de acordo com o interesse nesses padrões. Ainda, elas podem ser utilizadas para guiar ou restringir o espaço da busca, eliminando conjuntos de padrões que não satisfaçam condições predeterminadas.

As medidas de interesse associadas aos modelos de conhecimento em mineração de dados são objetivas e subjetivas. As objetivas são baseadas nas estruturas dos padrões descobertos e estatísticas a eles relacionados, como taxa de erro, consistência e suporte. Já as subjetivas referem-se às avaliações dos especialistas no domínio da aplicação em relação aos dados e modelos de conhecimentos gerados, como: surpresa; contradições; geradoras de alternativas; e de ações estratégicas (GOLDSCHMIDT ; PASSOS, 2005; REZENDE, 2003).

O usuário final precisa compreender o que foi gerado pelo algoritmo, ou seja, as regras devem ser de fácil interpretação. Algumas vezes, o processo de extração de padrões pode ser repetido, ajustando-se os parâmetros ou melhorando o pré-

processamento a fim de se obter melhores resultados para os usuários finais (REZENDE, 2003).

Pela simplificação e transformação do modelo de conhecimento, é possível torná-lo menos complexo, porém sem perder informações relevantes; e também converter sua forma de representação, a fim de facilitar a análise. Neste caso, árvores de decisão podem ser convertidas em regras, ou vice versa. No caso da simplificação e transformação do modelo de conhecimento, principalmente no caso de grandes conjuntos de regras, o estabelecimento de limites de precisão da regra pode auxiliar na detecção e exclusão das regras que não satisfaçam estes limites. Ainda, poderão ser realizadas exclusões de atributos, e no caso de algoritmos de classificação, poda³³ de árvores de decisão (REZENDE, 2003).

Fica evidente que o passo final da descoberta de conhecimento em base de dados não é somente a extração de informações a partir de tarefas de mineração de dados. Faz-se necessária a avaliação destas informações, principalmente o grau de contribuição que o processo traz ao conhecimento já estabelecido por parte dos especialistas. Ainda, caso o conhecimento gerado pelo algoritmo seja considerado de pouca utilidade para os especialistas e não responda aos objetivos propostos, é possível repetir o processo de extração de informações, melhorando a escolha dos dados ou ajustando os parâmetros dos algoritmos.

Em bases de dados da área da saúde podem ser encontrados muitos registros de atendimentos, relacionamentos construídos através do tempo, muitos diagnósticos e interações. Tal complexidade tende a aumentar, elevando o grau de aplicabilidade de técnicas de descoberta de conhecimento em bases de dados e mineração de dados. Esta situação pode colocar as aplicações desta tecnologia na área da saúde no mesmo patamar de utilização nas áreas da indústria que mais as utilizam atualmente, como a bancária, entre outras (BRESNAHAN, 1997).

O desenvolvimento deste capítulo sugere que modelos de saúde, que remuneram serviços e profissionais por procedimento realizado, necessitam ser

³³ A poda é uma técnica de lidar com ruído e *overfitting*. O ruído ocorre quando se trabalha com dados reais, que são imperfeitos. Um exemplo de ruído é quando exemplos tem os mesmos valores de atributos, mas com classes diferentes. O *overfitting* ocorre quando, a partir dos exemplos disponíveis, a hipótese seja muito específica para o conjunto de treinamento utilizado, ajustando-se em excesso à este conjunto. Nestes casos, a poda permite o aprendizado de uma hipótese bem genérica a partir do conjunto de treinamento, melhorando o desempenho para exemplos não vistos (REZENDE, 2003).

revisados. Tais modelos fazem com que sejam gerados cada vez mais procedimentos, dificultando o gerenciamento eficiente dos recursos financeiros sem, necessariamente garantir, a qualidade da assistência aos usuários.

Neste sentido, o gerenciamento de casos se constitui numa alternativa relevante para o aperfeiçoamento da gestão dos recursos e da qualidade da assistência aos usuários. Entretanto, esta ferramenta de gestão deve estar em consonância com o modelo de saúde e apoiada pela tecnologia da informação a fim de que seus benefícios sejam amplamente aproveitados.

3 METODOLOGIA

“Cavalheiros, o que pode ser mais fácil que colocar esse ovo em pé, algo que vocês disseram ser impossível? É a coisa mais simples do mundo. Qualquer um poderia fazer - depois de saber como”
- Colombo

A questão norteadora deste trabalho, conforme descrito no capítulo introdutório, está ligada à descoberta de padrões na base de dados de uma organização pagadora de serviços de saúde, que possibilitariam a identificação de usuários com doenças cardiovasculares, elegíveis para programas de gerenciamento de casos. Assim, tornar-se-ia possível o gerenciamento da assistência à saúde prestada de acordo com a necessidade que o usuário apresenta, de maneira a melhorar a sua qualidade de vida, e utilizar os recursos de financiamento, eficientemente.

Para tanto, foi desenvolvido um estudo quantitativo, retrospectivo, de cunho exploratório descritivo. A pesquisa quantitativa tem como objetivo a análise e compreensão de informações e fatos traduzidos em números, a fim de quantificá-los, utilizando-se de recursos e técnicas da estatística (SILVA; MENEZES, 2001).

Um estudo retrospectivo caracteriza-se pela análise retroativa de casos, com a análise das conseqüências de fatos passados. Um aspecto importante deste tipo de estudo é que como os fatos analisados não foram observados diretamente, sua quantificação depende das informações disponíveis serem suficientes e confiáveis. Muitas vezes, os dados encontrados são incompletos ou inconsistentes, tal como nas bases de dados da saúde, em que diversos profissionais têm acesso à documentação, porém trabalham com diferentes padrões ou critérios de registro (SILVA, 1999).

Já estudos exploratórios descritivos têm como característica observar, descrever e explorar aspectos de uma situação (POLIT; HUNGLER, 1995).

É importante ressaltar que todos os dados relacionados à identificação dos usuários foram omitidos. Após as primeiras buscas na base de dados, os números

de identificação dos usuários para o sistema foram substituídos por números em ordem crescente, a fim de manter a privacidade dos indivíduos.

3.1 População do Estudo

A população do estudo foi constituída de uma base de dados composta de dados de usuários do Instituto Curitiba de Saúde (ICS), que é uma operadora de saúde do tipo auto-gestão, responsável pelos atendimentos à saúde dos funcionários da Prefeitura Municipal de Curitiba, e seus dependentes. Inicialmente, foram analisados os dados relativos ao período de 2001 a 2005, que totalizavam 55.814 usuários e 1.168.983 registros. Contudo, pelo fato de os dados relativos aos anos de 2001 e 2005 apresentarem inconsistências, estes foram desprezados.

Considerando que o perfil epidemiológico do ICS está em consonância com o da região de Curitiba, e à necessidade operacional de restrição dos dados da base original, foram definidos critérios de seleção dos dados.

O primeiro foi que os usuários tivessem idade igual ou acima de 40 anos. Nesta operação a amostra inicial foi reduzida para 700.801 registros referentes a 24.086 usuários. O segundo critério foi que estes usuários deveriam ter em seus registros de atendimentos, pelo menos um registro relacionado com o grupo das Doenças do Aparelho Cardiovascular, ou seja, que iniciassem com a letra I do CID-10³⁴. Assim, o conjunto de dados foi reduzido para 401.041 registros, referentes a 8.457 usuários.

Após autorização do ICS e do Comitê de Ética em Pesquisa da Pontifícia Universidade Católica do Paraná, por meio do registro nº 924, foi iniciado o processo de coleta dos dados contidos na base de dados informatizados daquela instituição. A aprovação do Comitê de Ética encontra-se no Anexo 2. Foram realizadas buscas na referida base a fim de selecionar os registros, tendo em vista os critérios de inclusão e exclusão.

³⁴ CID-10 – Código Internacional das Doenças, na sua 10ª revisão (OMS, 1993).

3.2 Etapas Metodológicas

As etapas metodológicas foram desenvolvidas segundo o referencial de Goldschmidt e Passos (2005) e adaptadas do modelo proposto por Souza (2000). A Figura 5 apresenta uma representação esquemática das etapas metodológicas desenvolvidas para o alcance dos objetivos deste estudo.

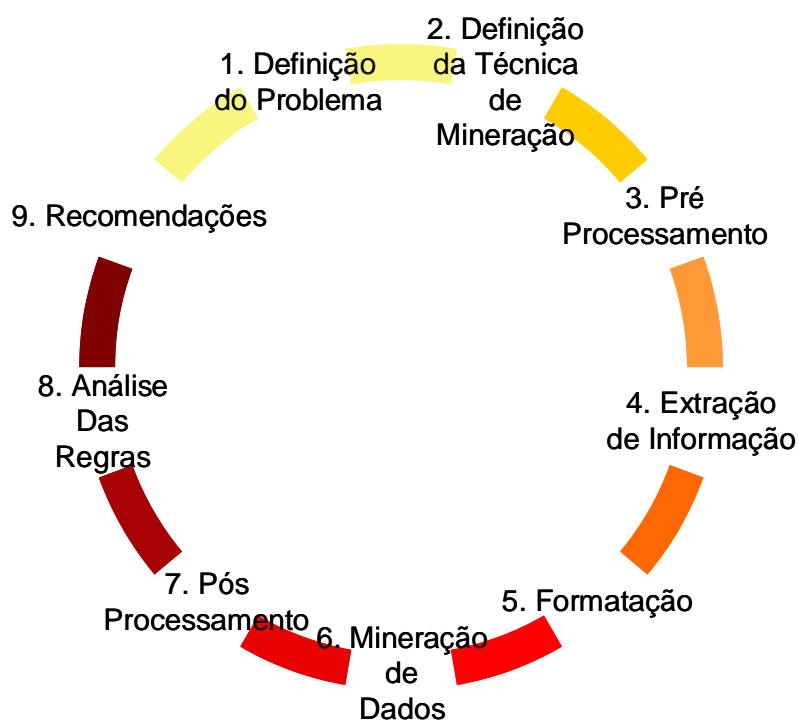


Figura 4 - Etapas Metodológicas

Fonte: Goldschmidt; Passos (2005; adaptado de Souza (2000)

3.2.1 Definição do Problema e da Técnica de Mineração de Dados

Em consonância com os objetivos específicos deste trabalho, foram abstraídos quatro problemas passíveis de serem resolvidos por algoritmos de mineração de dados.

Para alguns outros problemas definidos no início da pesquisa, também poderiam ter sido utilizados algoritmos para sua resolução. Entretanto, esses problemas foram manualmente trabalhados. A estratificação dos usuários em baixo, médio e de alto custo e a classificação dos procedimentos como de baixa, média e

alta complexidade são exemplos desses problemas que foram tratados manualmente. Para tanto, foi utilizado o conhecimento de especialistas da área da saúde.

Pelo fato de este trabalho ter a característica da multidisciplinariedade, a escolha do especialista não foi uma tarefa fácil. Foram definidos como critérios para escolha do especialista a experiência assistencial com usuários portadores de doenças cardiovasculares; a experiência em atividades de gestão em operadoras ou prestadoras de serviços de saúde; que apresentasse como titulação acadêmica o grau de doutor; e que tivesse disponibilidade para participar de reuniões a fim de discutir os dados e resultados advindos do processo metodológico. Obedecendo aos critérios estabelecidos, foram encontrados dois especialistas, um da área da medicina e outro da área da enfermagem.

O Quadro 4 apresenta os problemas e as técnicas de mineração de dados definidas para a resolução de cada um deles.

Quadro 4 - Problemas e Técnicas de DM e a tarefa de KDD definida para a sua resolução

Problema	Tipo de Tarefa de KDD
1. Descobrir a existência de associação entre procedimentos que geram um padrão indicativo de alto custo e de alta complexidade, e aprender com esta associação para detectar casos semelhantes no futuro.	Regras de Associação, cuja aplicação foi feita por meio do algoritmo Apriori (BORGELT, 1996)
2. Estratificar os usuários em baixo, médio e alto custo, de acordo com os procedimentos realizados.	Clusterização
3. Constatar a evolução dos procedimentos ou conjuntos de procedimentos, relacionado ao seu custo e aprender com esta constatação para prever casos semelhantes no futuro.	Classificação
4. Constatar a evolução dos atendimentos de usuários, relacionado à complexidade para aprender com esta constatação e prever casos semelhantes no futuro.	Classificação

Para trabalhar um problema em consonância com os objetivos desta pesquisa, foi escolhido o Problema 1, ou seja, descobrir a existência de associação entre procedimentos que geram um padrão indicativo de alto custo e de alta

complexidade e aprender com esta associação para detectar casos semelhantes no futuro. O problema foi resolvido mediante a aplicação do algoritmo Apriori, por meio de descoberta de regras de associação.

Os demais problemas definidos poderiam contribuir para a descoberta de outros conhecimentos inerentes à base de dados, e também para a identificação de usuários para programas de gerenciamento de casos. Entretanto, a fim de focar o trabalho e alinhá-lo com os objetivos traçados, dentro das condições técnicas e do tempo disponível para a realização da pesquisa, a aplicação do processo de descoberta de conhecimento em bases de dados foi restrita à resolução desse primeiro problema.

3.2.2 Pré-Processamento

O pré-processamento compreende a execução de funções de seleção, limpeza, construção e formatação dos dados. Antes mesmo de se executar estas funções, foram realizadas diversos experimentos com os dados, até mesmo para um melhor conhecimento acerca da base utilizada. Desta forma, buscou-se definir as variáveis que seriam utilizadas, considerando a real importância de cada uma delas no processo de descoberta de conhecimento em bases de dados, a melhor maneira de apresentação de cada uma delas, além de definições relacionadas ao domínio do problema.

A seguir cada uma das funções de pré-processamento está descrita da maneira como foi realizada, de acordo com a necessidade do algoritmo de associação, uma vez que esta tarefa foi definida para a resolução do problema.

É importante salientar que a escolha da técnica de associação se deu a partir da definição do problema: descobrir a existência de associação entre procedimentos que geram um padrão indicativo de alto custo e de alta complexidade, e aprender com esta associação para detectar casos semelhantes no futuro.

3.2.3 Seleção de Dados

Foi solicitado ao administrador da base de dados, um conjunto de dados referente às transações operacionais dos seus prestadores de serviços de saúde.

Os dados fornecidos foram selecionados através da função de seleção de dados de **junção direta**. A partir desta base de dados, os dados foram selecionados

considerando o objetivo proposto para o estudo e os critérios de seleção dos dados, como já explicitado. O Quadro 5 apresenta os códigos do CID-10 considerados para a seleção.

Quadro 5 - Códigos do CID-10 utilizados para seleção de usuários para a pesquisa

Capítulo IX	Doenças do Aparelho Circulatório
Código	Descrição
I00 – I02	Febre Reumática Aguda
I05- I09	Doença Reumática Crônica do Coração
I10 – I15	Doenças Hipertensivas
I20 – I25	Doenças Isquêmicas do Coração
I26 – I28	Doenças cardíaca pulmonar e da circulação pulmonar
I30 – I52	Outras formas de doença do coração
I60 – I69	Doenças Cerebrovasculares
I70 – I79	Doenças das artérias, arteríolas e capilares
I80 – I89	Doenças das veias, vasos linfáticos e dos gânglios linfáticos, não classificados em outra parte.
I95 – I99	Outros transtornos do aparelho circulatório e os não especificados
I32	Pericardite
I39	Endocardite
I41	Miocardite
I52	Outros transtornos cardíacos em doença classificada em outra parte
I68	Transtornos cerebrovasculares em doenças classificadas em outra parte
I79	Transtornos das artérias, das arteríolas e capilares em doenças classificadas em outra parte
I98	Outros transtornos do aparelho circulatório em doenças classificadas em outra parte

Fonte: CID-10 - Código Internacional de Doenças, 10ª revisão.

O próximo passo adotado foi a seleção de dados por **junção orientada**. Em seguida, foram realizadas todas as demais funções de pré-processamento, até que a base de dados contivesse os dados prontos para serem formatados e então submetidos ao algoritmo de mineração de dados.

Por meio da **eliminação direta** foram eliminados casos cujo código de identificação era o número 99999, reconhecidamente relacionado à identificação dos exames periódicos realizados com os funcionários da Prefeitura Municipal de Curitiba.

Inicialmente, quando se cogitava trabalhar com os valores pagos aos diversos prestadores, consideravam-se três variáveis que faziam referência destes. A primeira era referente ao valor pago pela Prefeitura Municipal de Curitiba; a

segunda, ao valor pago pelo ICS; a terceira, a co-participação do usuário, se houvesse. Por meio da função de **agregação de informações**, como parte da redução horizontal de dados, estes valores foram reunidos e somados, obtendo-se o valor total pago para cada procedimento realizado, para cada caso.

Por meio da **eliminação direta de variáveis** foi excluída a variável NR_CART (números da carteira dos usuários), referente à identificação dos usuários, a fim de preservar o anonimato dos mesmos.

3.2.4 Limpeza dos Dados

Ao manipular a base de dados percebeu-se que na grande maioria dos casos havia variáveis com preenchimento incompleto, tais como os códigos de especialidades, código do conselho profissional, grau de escolaridade, profissão, entre outros.

Tais variáveis foram excluídas da base a ser trabalhada, entretanto, esta não se mostrou uma etapa complicada de ser realizada, uma vez que a falta de preenchimento global das variáveis já indicou quais delas deveriam ser excluídas, para não prejudicar a qualidade dos padrões a serem gerados pelo algoritmo de mineração de dados.

Foram eliminados os casos cujos valores de CID-10 não estavam contidos no rol de códigos, por exemplo, os CID-10 de valor “ics”, o qual obviamente não existe na tabela de códigos CID-10, e aqueles cujo identificador não apresentava um valor válido.

3.2.5 Construção de Variáveis

Foi criada a variável “IDADE” a partir da variável “DATA_NASC” (data de nascimento) e da variável “DATA_ATEND” (data de atendimento), para determinar a idade do usuário no momento do registro do seu primeiro atendimento. Para este estudo, a idade do usuário é aquela que ele possuía no momento do registro do seu primeiro atendimento, por mais que o usuário tenha mais de um atendimento ao longo dos quatro anos de registro, e a sua idade varie com o passar do tempo.

Ademais, como ficou definido que as variáveis contidas na base de dados que se relacionavam aos valores dos procedimentos não seriam utilizadas devido a inúmeras inconsistências, foi criada uma nova tabela com dados provenientes da Tabela AMB-Ciefas. Esta tabela apresenta um rol de procedimentos das diversas

áreas médicas, divididos em capítulos que são as grandes áreas da Medicina, e grupos que são os diferentes tipos de procedimentos. Além disso, cada procedimento está relacionado com seu valor.

Cada procedimento foi classificado como sendo de baixo, médio ou alto custo. Para cada conjunto de procedimentos do mesmo grupo foi computada a média dos valores dos procedimentos do grupo e realizada a sua comparação com o valor do procedimento a ser classificado. O procedimento seria classificado como de baixo custo, caso seu valor se encontrasse abaixo da média dos procedimentos do grupo em questão. Seria considerado de médio custo, caso seu valor fosse igual; e de alto custo, se o valor estivesse acima da média do seu grupo.

Acreditou-se ser mais prudente não utilizar as variáveis relacionadas ao custo presentes na base de dados original, pelo fato de apresentarem inúmeras inconsistências e para evitar resultados também inconsistentes. Como se trata de uma base de dados real, os resultados do processo de mineração de dados poderiam causar impacto irreal, positivo ou negativo, o que poderia influenciar de maneira errônea no processo de tomada de decisão.

3.2.6 Formatação dos Dados

Com base na definição do problema selecionado para mineração de dados, fato que determinou o tipo de técnica de mineração de dados a ser utilizada, conseqüentemente foi também determinada a formatação apropriada para o algoritmo específico para a técnica.

Arquivos no formato texto foram gerados, os quais continham tabelas que mostravam, respectivamente, o relacionamento entre os procedimentos médicos e o número de vezes que eles ocorriam para cada usuário; entre os procedimentos médicos e os CID-10; e entre os procedimentos médicos e seus valores, de acordo com a classificação em de baixo, médio e de alto custo, realizadas a partir da tabela AMB-Ciefas 1996.

Adicionalmente, os dados foram transformados de SYBASE para formato *.txt, de maneira que cada registro de linha representasse um caso, com todos os procedimentos dispostos seqüencialmente. Conseqüentemente, os arquivos com os dados nesta formatação puderam ser submetidos ao algoritmo Apriori.

3.3 Extração de Informações

A etapa de extração de informações ocorreu a partir da manipulação do conjunto de dados gerados após a etapa de pré-processamento, pois os dados tinham que estar selecionados, limpos e corrigidos.

A própria etapa de pré-processamento dos dados contribuiu significativamente para o melhor conhecimento da base de dados: tanto das suas potencialidades, como das suas fragilidades.

Adicionalmente, foram realizados experimentos estatísticos com os dados da base, por meio do *software* Microsoft Excel®, os quais geraram estatísticas acerca dos usuários, procedimentos e custos durante o início do ano de 2002 até o final de 2004.

3.4 Mineração de Dados

Para execução da tarefa de descoberta de regras de associação foi escolhido o algoritmo Apriori, com o intuito de descobrir os padrões de ocorrência dos procedimentos médicos entre os diversos usuários com registros na base de dados.

Trata-se de um aplicativo de fácil manipulação. Os parâmetros de suporte e confiança são ajustados facilmente, antes da execução do aplicativo.

Os parâmetros *default* do algoritmo Apriori são de 10 para suporte e de 50 para confiança. Foram realizadas tentativas, mal sucedidas, de aumentar os valores destes parâmetros, de modo que as regras obtidas tivessem maior representatividade. Conseqüentemente, não foi possível identificar os procedimentos relacionados a eventos cardiovasculares com valores de suporte mais altos. Desta maneira, foram adotados os valores de suporte e confiança *default* do aplicativo, a fim de identificar todos os casos possíveis, deixando a restrição das regras para o pós-processamento, por meio da análise subjetiva e objetiva dos resultados.

3.5 Pós-Processamento

O arquivo de saída do Apriori apresentou as regras no formato

$$Y \leftarrow X (\text{suporte}\%, \text{confiança}\%),$$

Onde: X é o antecedente da regra, e Y, o conseqüente. A regra neste formato torna difícil sua compreensão pelos especialistas da área da saúde. Para resolver esta questão, as regras foram passadas para o aplicativo Microsoft Excel®, onde foram substituídos os códigos dos eventos por suas respectivas descrições.

Ainda, devido à grande quantidade de regras geradas (112.720 kB), foi necessário definir com os especialistas, e também de maneira a manter a coerência com os objetivos do trabalho, quais regras estavam alinhadas aos objetivos.

Dentre todas as regras geradas, foram selecionadas para análise aquelas que apresentavam, no conseqüente da regra, eventos reconhecidamente de alto custo, tais como: internação em UTI; consulta de emergência³⁵; revascularização do miocárdio; implante de marcapasso; implante de *stent*; cateterismo cardíaco; angioplastia de vaso único; eletrofisiologia cardíaca; cintilografia do miocárdio; valores referenciais de revascularização do miocárdio e valores referenciais de cateterismo cardíaco³⁶. Estes eventos foram classificados pelos especialistas como eventos significativos.

O fato de se ter focado a análise nos eventos do conseqüente da regra (Y) justifica-se pelo próprio objetivo do trabalho, que é a identificação de padrões que indiquem usuários com doenças cardiovasculares elegíveis para programas de gerenciamento de casos. Por conseguinte, todos os procedimentos do antecedente da regra (X) formam conjuntos de procedimentos que, associados ou não, são importantes para esta identificação. E o conseqüente (Y), indica o evento final gerado pela associação.

Só então, as regras foram classificadas subjetivamente em **relevantes**, **irrelevantes** e **insignificantes**, com a contribuição da análise de especialistas em gestão em saúde e também em cardiologia.

Para esta pesquisa, uma regra foi classificada como **relevante** quando a associação entre os eventos agregou valor à análise, e apontou evento(s) alerta(s)

³⁵ A consulta emergencial, se analisada isoladamente, não está relacionada a alto custo. Entretanto, consultas de emergência são consideradas eventos não programados, freqüentemente relacionados à instabilidade clínica do usuário. Sendo assim, freqüentemente são desencadeadores de novos eventos (procedimentos de baixo, médio e alto custo, internações, entre outros).

³⁶ Importante salientar que eventos que possuem a descrição de valores referenciais são referentes a códigos criados no sistema relacionados a “pacotes” negociados com os prestadores. Como a base possui dados de 2001 a 2005, existem dois tipos de códigos: o do procedimento propriamente dito, cujo valor é somente do procedimento; e o que se refere a valores referenciais, criado posteriormente, cujo valor se refere ao procedimento, taxas hospitalares, materiais, medicamentos e honorários médicos. Por isso, ambos os códigos foram mantidos neste estudo.

para a indicação de usuário elegível para programas de gerenciamento de casos. Uma regra **irrelevante** foi classificada quando a associação entre os eventos não agregou valor à análise, e não apontou evento(s) alerta(s) para a indicação de usuário elegível para esses. Por sua vez, uma regra foi considerada **insignificante** quando a associação entre os eventos refletiu eventos realizados rotineiramente e, portanto, indicou uma associação óbvia.

Em seguida, as regras foram filtradas de acordo com a sua classificação, e somente as regras relevantes foram analisadas sob a perspectiva da confiança, pelo analista.

Considerando como exemplo as seguintes estruturas de regras:

1. $A \leftarrow B$ (17,1%, 53,8%)
2. $A \leftarrow B; C$ (13,2%, 50,0%)

Exemplos como a segunda regra foram considerados **redundantes**, em relação exemplos semelhantes à primeira regra, pois apresentou os mesmos elementos da primeira regra, e ainda mais outros elementos, o que fez com que diminuíssem os valores de suporte e a confiança.

Como exemplo prático do que foi exposto acima, tem-se os seguintes exemplos:

1. CONS_EMERG \leftarrow ANGIOPL_VASO_UNICO (18.4%, 57.1%)
2. CONS_EMERG \leftarrow ANGIOPL_VASO_UNICO CATETERISMO_CARD CONS_AMB (11.8%, 55.6%)

Na primeira regra, a leitura é feita da seguinte maneira: ‘dentre os usuários da amostra, 18,4% têm probabilidade de apresentar o procedimento angioplastia de vaso único em seus registros; destes, 57,1% têm probabilidade de apresentar uma consulta de emergência associada, em seus registros’.

Para a segunda regra, a leitura seria: ‘dentre os usuários da amostra, 11,8% têm probabilidade de apresentar uma consulta ambulatorial associada ao procedimento de cateterismo cardíaco, associados ao procedimento de angioplastia de vaso único em seus registros; destes, 55,6% têm probabilidade de apresentar uma consulta de emergência associada em seus registros’.

Percebe-se que na segunda regra, a associação presente no antecedente da regra (consulta ambulatorial; cateterismo cardíaco; angioplastia de vaso único) é muito mais específica do que a associação da primeira regra (angioplastia de vaso único). Inclusive, as taxas de suporte e confiança diminuíram, o que significa que a segunda regra apresenta associações encontradas em menor proporção (de 18,4% para 11,8%), além de menor força de determinação do conseqüente em relação ao antecedente da regra (de 57,1% para 55,6%).

O objetivo da análise sob a perspectiva da confiança foi diminuir o conjunto de regras a serem reapresentadas para os especialistas, a fim de otimizar sua análise e discussão. Entretanto, tendo em vista as peculiaridades da área da saúde em relação à estruturação do conhecimento médico; as particularidades e especificidade da instalação do processo saúde-doença em diferentes indivíduos; e ainda, considerando o objetivo de identificar usuários elegíveis para programas de gerenciamento de casos, decidiu-se apresentar também as regras consideradas redundantes pelo analista, de maneira que os especialistas pudessem contribuir com a sua análise destas regras.

Acreditou-se que, na fase de análise sob a perspectiva da confiança, a opinião dos especialistas poderia validar os resultados obtidos a partir deste modelo proposto, e então contribuir para a identificação de conhecimento específico relevante que poderia influenciar no entendimento das regras do negócio.

Sendo assim, com a ajuda dos especialistas, as regras relevantes foram novamente avaliadas, considerando sua taxa de confiança, e também a possibilidade de, entre estas regras, ainda existirem regras redundantes ou realmente relevantes.

O conjunto de regras relevantes foi separado em três grandes grupos, para facilitar sua visualização. No primeiro grupo havia regras com taxa de confiança acima de 66% de confiança. Este valor foi definido junto com os especialistas, pois se refere a dois terços da probabilidade de ocorrência das regras. No segundo grupo, as regras apresentavam taxa de confiança de 51% a 65%. No terceiro grupo, foram alocadas as regras com taxa de confiança abaixo de 50%.

Uma regra com confiança próxima de 50% significa que a probabilidade da sua ocorrência é aleatória, a mesma daquela que se encontra quando se joga uma moeda para cima, ou seja, 50% de chance de ocorrer, e 50% de chance de não ocorrer.

Entretanto, na perspectiva do especialista, dependendo dos elementos componentes da regra, a probabilidade desta ocorrer em 50% pode ser muito elevada para um evento que deveria ser exceção, por exemplo. Desta maneira, a ocorrência desta regra, mesmo com uma taxa de confiança baixa, próxima ou mesmo abaixo de 50%, torna-se relevante.

Assim, foi abstraído um conjunto de regras realmente relevantes, tanto em relação ao suporte como à confiança.

Vale destacar que nesta pesquisa não foi adotado o pós-processamento dos padrões descobertos sob as perspectivas: objetiva e subjetiva, discutidas por Freitas (1999). Essa forma de pós-processamento será proposta em trabalhos futuros.

Exemplos das regras geradas, bem como suas análises subjetivas e sob a perspectiva da confiança, são apresentados no Capítulo 4, referente aos resultados e discussões.

4 RESULTADOS E DISCUSSÕES

Conhecimento é poder.

- Francis Bacon

Considerando os objetivos traçados para esta pesquisa e a metodologia adotada, neste capítulo são apresentados e discutidos os resultados dos experimentos realizados em cada etapa metodológica do processo descoberta de conhecimento na base de dados estudada: do pré-processamento, da extração de informações, da mineração de dados e do pós-processamento.

4.1 Resultados do Pré-Processamento

Antes de apresentar os resultados e discussões relativas à fase de pré-processamento é importante salientar o quão intensiva foi a fase de preparação dos dados para a realização da pesquisa.

Diversos experimentos foram realizados com a base entregue pelo administrador da base de dados do ICS. Um primeiro complicador foi a quantidade de registros nela contidos. A disposição dos registros na base dificultou a realização de uma análise mais profunda dos dados. Diversas tentativas no intuito de contribuir para se criar o melhor conjunto de dados, e conseqüentemente, os melhores resultados possíveis, foram realizadas. Entre outras, pode-se citar a mudança da formatação das variáveis e as composições e associações de eventos que poderiam acontecer, concomitantemente.

Mesmo após terem sido definidos os dados a serem trabalhados e a formatação a que eles deveriam ser submetidos, quando se pensava ter obtido os resultados, era necessário novo pré-processamento, nova aplicação de filtros, novas justificativas, a fim de que os resultados pudessem ser analisados à luz dos objetivos da pesquisa.

A própria literatura revela que a manipulação de bases de dados na saúde é árdua, devido a inconsistências encontradas, dados incompletos (COLLAZOS *et al.*,

2002), e também pelo viés financeiro no qual estão baseados muitos dos registros dos atendimentos em saúde nas operadoras de planos de saúde. Isso ocorre em detrimento de outros dados que poderiam contribuir para a compreensão do fenômeno saúde-doença na população atendida.

Entretanto, ficou evidente a grande importância desta etapa operacional, tanto para o profundo conhecimento de inconsistências, potencialidades e fragilidades da base de dados, como para desencadear com sucesso as etapas seguintes do processo de descoberta de conhecimento na base de dados.

A análise do tempo despendido nesta fase, voltado principalmente para o preparo da base de dados para a aplicação do algoritmo, conduziu a reflexões acerca da exeqüibilidade de iniciativas de trabalhos como este fora do ambiente acadêmico.

Primeiramente, considerando as necessidades de pré-processamento da base, foram necessárias diversas horas de trabalho em conjunto com os especialistas na área da saúde e na área de análise de sistemas. Além deste trabalho operacional, inúmeras discussões foram conduzidas acerca das possibilidades de estudos e resultados esperados, sendo que muitos experimentos foram realizados e não tiveram êxito.

No sentido de otimizar esta fase do processo, é fundamental que as bases de dados na saúde sejam concebidas de maneira multiprofissional e, além disso, que desde a sua concepção sejam consideradas as informações possíveis de serem extraídas a partir dos seus dados. Assim, tornar-se-á possível que metodologias como esta sejam exeqüíveis não somente no ambiente acadêmico, mas também nas organizações de saúde.

Iniciando a descrição das operações realizadas na etapa de pré-processamento, por meio da **seleção direta de casos**, foi restringido tanto o número de registros de atendimento, bem como o número de usuários, a fim de possibilitar uma análise focalizada para tornar exeqüível a realização da presente pesquisa. O Quadro 6 apresenta um resumo das operações de seleção realizadas, bem como os resultados das seleções em termos de número de registros de atendimentos e número de usuários (ou casos) restantes na base a ser manipulada.

Quadro 6 - Descritivo da seleção dos casos, número de registros de atendimento relacionado aos usuários, Instituto Curitiba de Saúde, Curitiba, 2001-2005

Seleção	nº de registros	nº de usuários
Junção Direta	1.168.983	55.814
Idade >= 40 anos	700.801	24.086
CID -10 que inicia com a letra I	401.041	8.457
Casos com pelo menos um atendimento com código que se inicia com o dígito 40	2.987	77

A proposta de seleção dos casos que apresentavam, em seus registros, pelo menos um atendimento com código que se iniciasse com o dígito 40 (procedimentos cardiovasculares de alto custo) ocorreu devido ao fato de casos complexos, de risco ou de alto custo responderem por cerca de 0,5 a 1% de uma população atendida (KNABEL; LOUWERS, 2004; STATUSONE, 2004; FORMAN, 2004; CROOKS, 2005).

Numa primeira análise, o fato de se ter diminuído consideravelmente o número de usuários e de registros, pode causar questionamentos relativos à representatividade da população selecionada, em relação à quantidade de usuários e registros na base de dados inicial.

Entretanto, é importante considerar que a desigual distribuição das diversas patologias numa população é uma característica de bases de dados da área da saúde. É fundamental considerar que os indivíduos, isoladamente, ou as comunidades, estão mais ou menos expostas a multicausalidade do processo saúde-doença, tais como a: fatores biológicos e comportamentais; fatores ambientais; fatores imunológicos; fatores nutricionais; fatores genéticos; fatores sociais, espirituais e relacionados ao acesso a sistemas de saúde (JEKEL, 2005).

Como resultado da operação de seleção de **junção direta**, foi gerado um arquivo *.mdb (Microsoft Access), com variáveis referentes a cada um dos 1.168.983 registros relativos aos atendimentos do total de 55.814 usuários, com tamanho de cerca de 2GB, como mostra o Quadro 7.

Quadro 7 – Demonstrativo das variáveis contidas na base de dados inicial e sua disposição no arquivo fornecido pelo administrador da base de dados do ICS

CART	SEXO	NASC	CID PRINCL	DESCRICAO	DT_PROCE DIMENTO	NR_PRO CESSO	EXAME	TOTAL	MATMED		
1	F	18/02/1921	G45	ACID VASC CEREBR ISQUEMICOS TRANS SINDR COR	09/07/2004	1026909	36010014	R\$ 713,30	R\$ 154,00		
			I10	HIPERTENSAO ESSENCIAL	01/04/2003	666421	20010028	R\$ 136,56			
							20010133	R\$ 136,56			
			M819	OSTEOPOROSE NE	25/03/2003	640613	32080115	R\$ 135,00			
			R51	CEFALEIA	14/02/2004	904014	36010014	R\$ 464,80			
					904015	36010189	R\$ 2.846,25	R\$ 375,95			
2	M	15/08/1929	E11	DIABETES MELLITUS NAO-INSULINO-DEPENDENTE	04/12/2002	578019	28010493	R\$ 58,80			
									28010507	R\$ 58,80	
									28010540	R\$ 58,80	
									28010973	R\$ 176,40	

Fonte: ICS, 2001-2005.

A etapa seguinte constituiu-se na escolha das variáveis com grande potencial de resposta aos problemas da pesquisa (Quadro 8). Tal seleção foi realizada pela operação de **junção orientada**.

Quadro 8 – Demonstrativo das variáveis definidas por junção orientada

Variável	Descrição
NR_CART	Identificação do usuário
SEXO	Sexo
DT_NASC	Data de nascimento
DT_PROCEDIMENTO	Data do procedimento
EXAME	Procedimento realizado

Numa tentativa de construir uma nova variável que indicasse o custo dos procedimentos realizados pelos usuários, foi proposta uma classificação dos procedimentos em BAIXO, MEDIO, e de ALTO custo.

No Quadro 9 estão apresentados alguns exemplos dos procedimentos da tabela AMB-Ciefas, com seus respectivos grupos e a classificação de acordo com seu custo, a fim de exemplificar o processo realizado.

Quadro 9 - Descritivo da classificação dos procedimentos em ALTO, MÉDIO e BAIXO custos

Código do Procedimento	Grupo	Custo
21010188	21	BAIXO
21010196	21	ALTO
21010200	21	ALTO
21010218	21	ALTO
21010226	21	ALTO
21010234	21	ALTO
21010242	21	ALTO
21010250	21	BAIXO
22010017	22	BAIXO
22010025	22	BAIXO
22010033	22	BAIXO
22010041	22	BAIXO

No Quadro 10 pode ser observada a criação do atributo IDADE, a partir da data de nascimento por ocasião do primeiro procedimento registrado na base de dados, bem como a formatação da base de dados com o histórico de cada usuário representado em uma linha, de modo a possibilitar aplicação do algoritmo Apriori.

Quadro 10 - Demonstrativo da formatação apresentando o histórico de procedimentos de cada usuário em linhas distintas

Variáveis	
Identificação – Sexo – Idade	Eventos
1 F 82	36010014 20010028 20010133 32080115 36010014 36010189
2 M 72	28010493 28010507 28010540 28010973 28011023 28011210 28011279 28011392 28130367 28010493 28010507 28010973 28011023 28011392 28130367 28010493 28010507 28010973 28011023 28011392 28010175 28010493 28010507 28010973 28011023 28011392 28010175 28010493 28010507 28010973 28011023 28011392 28010493 28010507 28010973 28011023 28011392 28010493 28010507 28010973 28011023 28011392 28061624 28010493 28010507 28010973 28011023 28011392 28010493 28010507 28010973 28011023 28011392 20010176 20010184 32020066 32040040 32040067
3 F 69	28010493 28010507 28010540 28010876 28010973 28011139 28011392 28040490 28130367 28010493 28010507 28011023 28011392 28010493 28010507 28011023 28011392 23010037 20010052 10014 28010493 28010507 28010540 28011210 28011279 28011392 28040481 28010493 28010507 28011392 21010021 32020066 32020040 32020066 28010540 28010973 28040481 28040490 28040732 28010493 28010507 28011023 28011392 28010493 28010507 28011392 10014 10014 95010173 50010190 32040075
4 M 56	10022 20010010 87020211 32050054 10022 28010558 28010566 87020211 28061624 32050054 10014 23010037 23020148 87050234 21010021 20010010 20010028
5 F 74	50010123 50010190 50010212 33010358 32040075 32030037 32050054 28010973 28040139 28040481 10022 34010068 32050054 20010133 10022

Nesta formatação, foram registrados o código de identificação do usuário, o sexo, a idade e todos os procedimentos realizados pelo usuário, em ordem cronológica.

Na etapa de construção de variáveis, foram gerados arquivos contendo relações entre os procedimentos e o valor a eles atribuído, a fim de analisar seu custo. O Quadro 11 apresenta exemplos que associam o procedimento ao seu custo.

Quadro 11 - Demonstrativo das associações entre evento e custo

Variáveis			
Identificação – Sexo – Idade	Evento/Custo		
1 F 82	20010028_MEDIO 36010014_ALTO	20010133_MEDIO 36010189_ALTO	32080115_ALTO
2 M 72	20010176_MEDIO 28010493_BAIXO 28010973_BAIXO 28011279_BAIXO 28130367_BAIXO 32040067_MEDIO	20010184_MEDIO 28010507_BAIXO 28011023_BAIXO 28011392_BAIXO 32020066_MEDIO	28010175_BAIXO 28010540_BAIXO 28011210_BAIXO 28061624_MEDIO 32040040_MEDIO
3 F 69	10014_BAIXO 23010037_MEDIO 28010540_BAIXO 28011023_BAIXO 28011279_BAIXO 28040490_BAIXO 32020040_MEDIO 50010190_MEDIO	20010052_MEDIO 28010493_BAIXO 28010876_BAIXO 28011139_BAIXO 28011392_BAIXO 28040732_BAIXO 32020066_MEDIO 95010173_ALTO	21010021_BAIXO 28010507_BAIXO 28010973_BAIXO 28011210_BAIXO 28040481_BAIXO 28130367_BAIXO 32040075_MEDIO
4 M 56	10014_BAIXO 20010028_MEDIO 23020148_BAIXO 28061624_MEDIO 87050234_MEDIO	10022_BAIXO 21010021_BAIXO 28010558_BAIXO 32050054_MEDIO	20010010_MEDIO 23010037_MEDIO 28010566_BAIXO 87020211_MEDIO
5 F 74	10022_BAIXO 28040139_BAIXO 32040075_MEDIO 34010068_ALTO 50010212_MEDIO	20010133_MEDIO 28040481_BAIXO 32050054_MEDIO 50010123_MEDIO	28010973_BAIXO 32030037_MEDIO 33010358_MEDIO 50010190_MEDIO

O principal atributo escolhido para ser estudado nesta base de dados foi o atributo ‘procedimento’, nesta pesquisa chamado de ‘eventos’. Tal escolha fundamenta-se no fato de que o outro atributo que poderia também ser analisado, os códigos do CID-10, poderia causar um viés na pesquisa.

Neste sentido, Forman (2004) salienta que, para fins de detecção de usuários com indicação para programas de gerenciamento de casos, os códigos de

diagnósticos de doenças não devem ser considerados como um atributo-meta. O autor comenta que usuários de alto risco geralmente são acometidos de muitas patologias, ou seja, apresentam comorbidades associadas, e caso seja realizada uma busca apenas por determinadas doenças, poderão ser perdidos os dados referentes a outras patologias.

Jekel *et al.* (2005) comentam que a qualidade do diagnóstico médico vem se modificando com o passar dos anos e relacionam este fato com as freqüentes alterações nas versões das classificações de doenças. Por sua vez, Gillespie comenta sobre a dificuldade em sensibilizar os profissionais da saúde, sobretudo os médicos, para o preenchimento correto de códigos relacionados à sua prática clínica, e não somente em relação àqueles relacionados às operações financeiras.

Corroborando com esta declaração, na prática de atividades de auditoria de contas médico-hospitalares é muito comum os auditores se depararem com códigos do CID-10 que indicam investigação diagnóstica ou doenças inespecíficas³⁷, ou seja, não indicam realmente a doença que o usuário apresenta. Muitas vezes, este código permanece nas solicitações de liberação de procedimentos, o que prejudica sobremaneira a análise posterior das liberações, e também a análise do histórico do usuário, se forem considerados somente estes códigos para tais análises.

Entretanto, considerou-se importante gerar um arquivo que mostrasse a relação entre os procedimentos e os seus respectivos códigos do CID-10. Os dados advindos deste arquivo poderiam contribuir para a análise dos casos quanto as comorbidades apresentadas, conseqüentemente auxiliando na tomada de decisão clínica e administrativa. Como os procedimentos estariam todos relacionados em ordem cronológica e para cada caso, poder-se-ia verificar as múltiplas comorbidades às quais são acometidos os usuários do estudo.

O Quadro 12 apresenta as relações dos procedimentos com os seus respectivos códigos do CID10.

³⁷ Exemplo de códigos do CID-10 que indicam: (a) investigação diagnóstica: Z00; (b) doenças inespecíficas: I95 a I99 - Outros transtornos do aparelho circulatório e os não especificados.

Quadro 12 - Demonstrativo das relações dos procedimentos com os seus respectivos códigos do CID-10

Variáveis	
Identificação – Sexo – Idade	Eventos
1 F 82	36010014_G45 20010028_I10 20010133_I10 32080115_M819 36010014_R51 36010189_R51
2 M 72	28010493_E11 28010507_E11 28010540_E11 28010973_E11 28011023_E11 28011210_E11 28011279_E11 28011392_E11 28130367_E11 28010493_E11 28010507_E11 28010973_E11 28011023_E11 28011392_E11 28130367_E11 28010493_E11 28010507_E11 28010973_E11 28011023_E11 28011392_E11 28010175_E11 28010493_E11 28010507_E11 28010973_E11 28011023_E11 28011392_E11 28010175_I10 28010493_I10 28010507_I10 28010973_I10 28011023_I10 28011392_I10 28010493_I10 28010507_I10 28010973_I10 28011023_I10 28011392_I10 28010493_I119 28010507_I119 28010973_I119 28011023_I119 28011392_I119 28061624_I119 28010493_E11 28010507_E11 28010973_E11 28011023_E11 28011392_E11 28010493_E11 28010507_E11 28010973_E11 28011023_E11 28011392_E11 20010176_I672 20010184_I672 32020066_M190 32040040_M190 32040067_M190
3 F 69	28010493_E781 28010507_E781 28010540_E781 28010876_E781 28010973_E781 28011139_E781 28011392_E781 28040490_E781 28130367_E781 28010493_E782 28010507_E782 28011023_E782 28011392_E782 28010493_E782 28010507_E782 28011023_E782 28011392_E782 23010037_E782 20010052_I10 10014_I10 28010493_I11 28010507_I11 28010540_I11 28011210_I11 28011279_I11 28011392_I11 28040481_I11 28010493_I11 28010507_I11 28011392_I11 21010021_L859 32020066_M150 32020040_M549 32020066_M549 28010540_E11 28010973_E11 28040481_E11 28040490_E11 28040732_E11 28010493_E11 28010507_E11 28011023_E11 28011392_E11 28010493_E782 28010507_E782 28011392_E782 10014_H261 10014_H261 95010173_H261 50010190_H261 32040075_M170
4 M 56	10022_I10 20010010_I10 87020211_I10 32050054_J111 10022_K30 28010558_R074 28010566_R074 87020211_R10 28061624_Z125 32050054_Z136 10014_H521 23010037_K29 23020148_K29 87050234_K29 21010021_K29 20010010_Z136 20010028_Z136
5 F 74	50010123_H25 50010190_H25 50010212_H25 33010358_I83 32040075_M704 32030037_R52 32050054_R52 28010973_Z000 28040139_Z000 28040481_Z000 10022_G43 34010068_G91 32050054_M545 20010133_R94 10022_Z00

Com o objetivo de conhecer o número de procedimentos realizados para cada caso, foi gerado um arquivo com todos os usuários, todos os procedimentos

realizados por cada um deles, porém associados a um contador. Nesse caso, cada registro de procedimento foi considerado singular, e os procedimentos não se apresentaram na ordem cronológica de acontecimento. Esta operação diminuiu sobremaneira o número de registros, mantendo, porém, os dados originais, facilitando a análise dos mesmos.

O Quadro 13 apresenta os dados que relacionam os procedimentos e o número de vezes em que eles ocorreram na base de dados.

Quadro 13 – Demonstrativo dos dados que relacionam os procedimentos e o número de suas ocorrências na base de dados da pesquisa

Variáveis			
Identificação – Sexo – Idade	Eventos		
1 F 82	20010133_1 36010189_1	32080115_1	36010014_2
2 M 72	20010184_1 28010507_9 28011023_9 28011392_9 32020066_1	28010175_2 28010540_1 28011210_1 28061624_1 32040040_1	28010493_9 28010973_9 28011279_1 28130367_2 32040067_1
3 F 69	20010052_1 28010493_7 28010876_1 28011139_1 28011392_7 28040732_1 32020066_2 95010173_1	21010021_1 28010507_7 28010973_2 28011210_1 28040481_2 28130367_1 32040075_1	23010037_1 28010540_3 28011023_3 28011279_1 28040490_2 32020040_1 50010190_1
4 M 56	10022_2 21010021_1 28010558_1 32050054_2	20010010_2 23010037_1 28010566_1 87020211_2	20010028_1 23020148_1 28061624_1 87050234_1
5 F 74	20010133_1 28040481_1 32050054_2 50010123_1	28010973_1 32030037_1 33010358_1 50010190_1	28040139_1 32040075_1 34010068_1 50010212_1

Após todas as etapas de pré-processamento, alguns experimentos foram realizados, utilizando o algoritmo Apriori a fim de verificar, a partir dos resultados gerados nestes primeiros experimentos, se a base de dados apresentava potencial

para produzir resultados que respondessem à questão norteadora do trabalho e aos objetivos traçados, bem como para que a pesquisadora conhecesse os dados.

A partir da análise das regras geradas nestes primeiros experimentos, foi constatado que a base possuía potencial para permitir o alcance dos objetivos. Entretanto, não haviam sido geradas regras que contivessem os códigos de procedimentos relacionados às doenças cardiovasculares.

Conseqüentemente, foi realizada mais uma etapa de pré-processamento. As regras foram submetidas a um filtro, buscando-se identificar apenas os usuários que apresentavam, dentre seus procedimentos, pelo menos um que apresentasse códigos que se iniciavam com o dígito “40”, pois esses são os que indicam procedimentos cardiológicos invasivos. O Anexo 3 apresenta os códigos da tabela de procedimentos do ICS, baseada na tabela AMB-Ciefas 1996, que se relacionam a procedimentos cardiológicos diagnósticos e terapêuticos.

Com a realização desta etapa de pré-processamento, o número de usuários ficou significativamente reduzido, porém, representativo dos usuários que efetivamente tiveram no seu histórico, eventos relacionados a procedimentos cardiovasculares. O Quadro 14 apresenta os usuários e os procedimentos realizados, sendo que todos apresentavam em seus registros, pelo menos um evento cujo código iniciasse com o dígito 40.

Durante a execução das etapas metodológicas foi definido que era preciso identificar os usuários que apresentavam, em seu histórico de registros de eventos em saúde, códigos relacionados a condições de alto custo e alta complexidade.

A partir de então, foi possível examinar, mediante associações, quais destes eventos, ou conjuntos de eventos que, quando detectados precocemente por um sistema informatizado, poderiam atuar como alertas para uma possível indicação deste usuário para um programa de gerenciamento de caso. Os eventos e suas associações foram identificados pelo algoritmo Apriori.

Quadro 14 – Demonstrativo dos dados que relacionam usuários que apresentavam em seus registros pelo menos um procedimentos com código que iniciasse com o dígito 40

Variáveis	
Identificação – Sexo – Idade	Eventos
35 F 70	10014 10022 20010 20010010 20010133 22010130 28010175 28010493 28010507 28010540 28010558 28010566 28010973 28011210 28011279 28011392 28030141 28040481 28040490 28040732 28040791 28050703 28050720 28060504 28130367 28999991 32020066 32040091 32040105 32050038 33010358 40040178 40080404 40080420 50010123 51010020 51010046 51010054 51010127 87010046 87020068 87020300 87020335 87030039 87050307 91000025 96010010 96010029
299 M 62	10014 10022 10200 10766 20010 20010010 20010028 20010133 20010176 20010184 22010130 27999998 28010507 28010973 28011023 28011392 28011511 28061624 28130367 28999991 32050038 32050054 34010068 34010084 39030016 39030164 39040267 40040100 50010131 51010380 87010046 87020203 87020300 91000033 91000068 95010378 95010750 96010029
308 M 73	20010133 28010540 28010973 28011104 28011210 40050068 87010305 87020300 87030047 87030063 87040107
484 F 54	10014 10200 20010 20010010 20010028 20010133 20010150 25040030 27020010 27030059 27040046 27040062 27040151 27040194 27040216 27040232 27040240 27040283 27040348 27040356 27040364 27040372 27040399 27040410 27040437 27040453 28010329 28010540 28010558 28010566 28010965 28010973 28011120 28011210 28011279 28011414 28040139 28040414 28040422 28040481 28040562 28040732 28040791 32050038 32050054 32120010 39030016 39040267 40020045 40040054 40040100 40040135 40040143 40060012 40080200 87010020 87010305 87020068 87020076 87020211 87020300 87020319 87030012 87030020 87030055 87030063 87030071 87040107 87059991 95010750 96010029
703 M 77	10022 10200 20010 20010010 20010052 20010133 28010540 28010973 28011104 28011210 28999991 32050038 32050054 40050076 40050084 87010046 87020025 87020300 87030020 87040018 92020046 96010029

4.2 Resultados da Extração de Informações

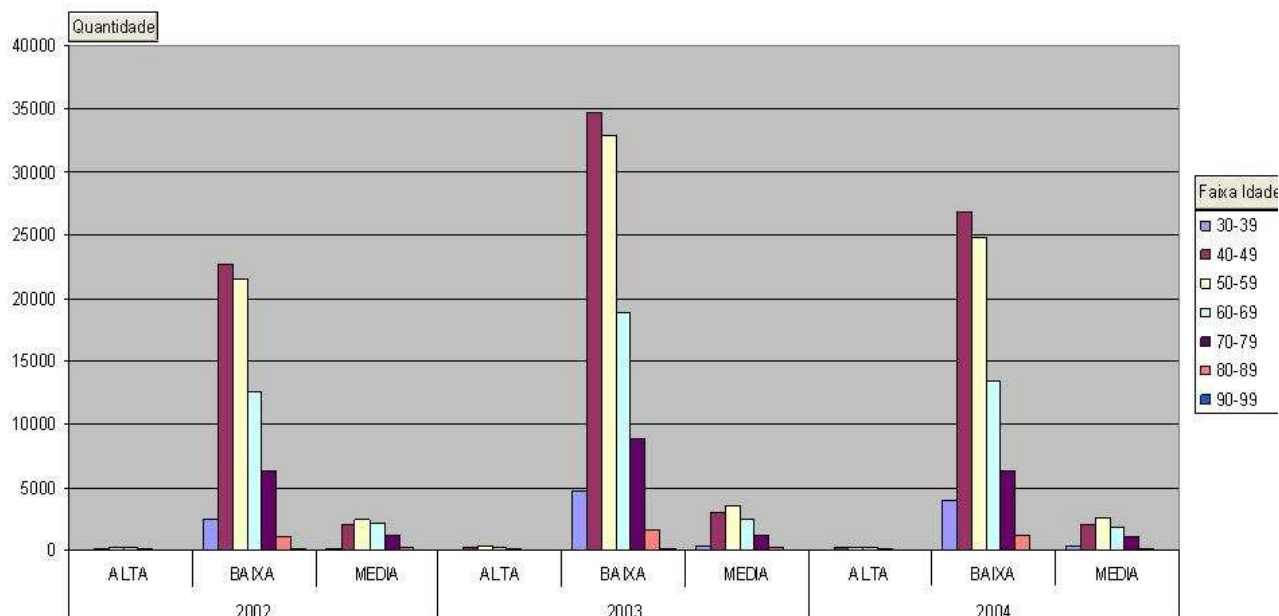
Esta análise refere-se às estatísticas que puderam ser realizadas com a base de dados da pesquisa. Foi realizada uma busca, considerando os códigos do CID-10 e o custo a eles relacionados no ano de 2004. Sendo assim, obteve-se que os códigos associados a maior custo, encontrados com maior frequência, foram àqueles relacionados no Capítulo IX do CID-10, a saber, das Doenças do Aparelho Circulatório.

Os dados do ICS revelam que em 2004, a faixa etária com maior incidência das doenças relacionadas neste capítulo do CID-10 foi a de 50 a 59 anos, com cerca de 11.790 atendimentos realizados, gerando um custo de R\$ 2.371.665,87 neste ano.

No intuito de corroborar com o conhecimento acerca da base de dados desta pesquisa foram construídos gráficos a partir dos dados, os quais evidenciaram condições interessantes acerca da faixa etária, do grau de complexidade, da utilização dos serviços, do custo dos serviços, e suas inter-relações.

O Gráfico 2 apresenta a correlação entre faixa etária, complexidade e a quantidade dos procedimentos aos quais os usuários foram submetidos.

Gráfico 2 - Relação entre faixa etária, complexidade e quantidade de procedimentos. Instituto Curitiba de Saúde, 2002-2004

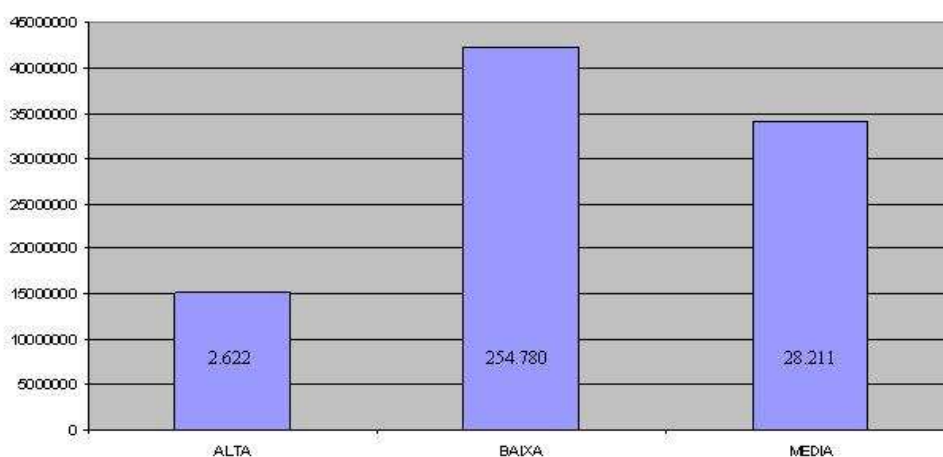


Percebe-se que a faixa etária que mais se destaca na utilização dos procedimentos cardiovasculares, nos três anos, foi a dos 40-49 anos, seguida dos 50-59 anos, e dos 60-69 anos. Assim, pode-se afirmar que predomina, nesta base de dados, a utilização de serviços de saúde por uma população de faixa etária mais avançada.

O conhecimento da faixa etária predominante é uma informação importante acerca da população de um sistema de saúde, uma vez que cada faixa etária apresenta sua predisposição a fatores de risco para diferentes agravos em saúde, estilo de vida característicos, entre outros. O conhecimento deste indicador pode influenciar diversos aspectos do gerenciamento em saúde, entre eles o cálculo atuarial³⁸, a negociação de valores referenciais para os procedimentos específicos desta faixa etária, e programas de promoção à saúde e prevenção de doenças específicas. Observa-se ainda que a utilização do sistema de saúde está voltada a procedimentos de baixa complexidade, em todas as faixas etárias, nos três anos analisados nesta pesquisa.

O Gráfico 3 apresenta a relação entre o custo dos procedimentos, com seus valores em reais, e a sua complexidade.

Gráfico 3 - Relação entre complexidade dos procedimentos, sua utilização pelos usuários e o custo a eles relacionados. Instituto Curitiba de Saúde, 2002-2004



³⁸ Cálculo atuarial – cálculo que contempla, além dos custos passados e o custeio das despesas totais de administração, a previsão da evolução futura dos custos num ambiente de inflação da saúde, o envelhecimento da população beneficiária, o risco de aumento de custos, de utilização (sinistralidade) ou (VENDRAMINI, 2006, p.).

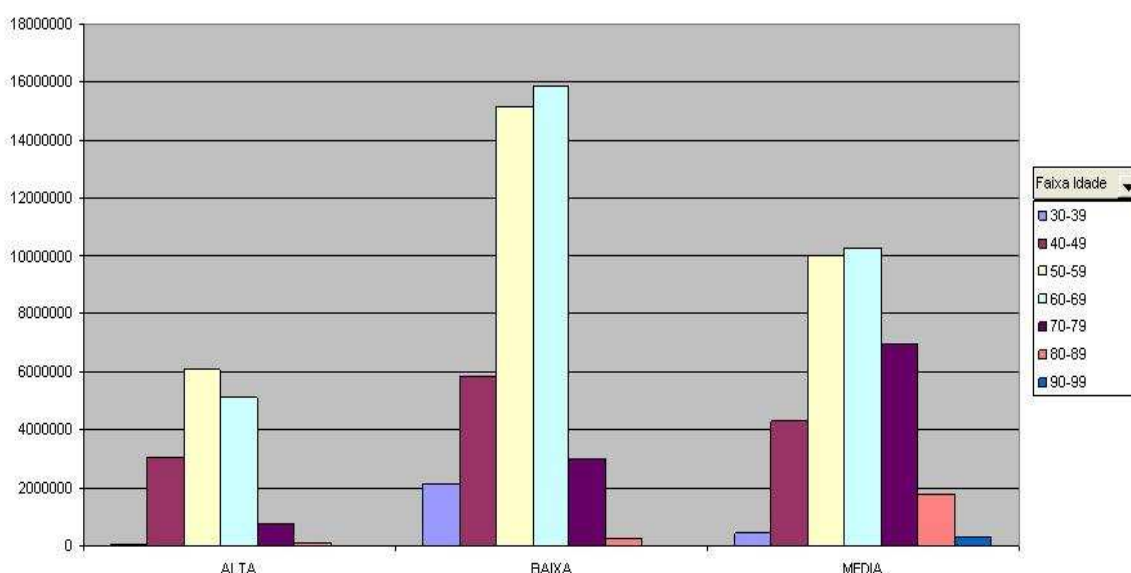
Percebe-se que os procedimentos de baixa complexidade estão relacionados a um maior custo total, durante os três anos pesquisados, seguidos dos procedimentos de média complexidade, e só então dos de alta complexidade, tal como aponta também o Gráfico 2.

A relação entre o alto custo e os procedimentos de baixa complexidade pode estar relacionada com atividades de auditoria, dentro das operadoras de saúde, voltadas aos procedimentos de alta complexidade, em detrimento dos procedimentos de baixa complexidade.

O que se observa é que a prática de atividades de auditoria, que pertencem à função administrativa de controle de custos, vem sendo realizada de maneira eficiente pelas organizações pagadoras de serviços de saúde, a fim de conter a grande utilização dos procedimentos de alta complexidade. Entretanto, estratégias devem ser implementadas a fim de regular também a utilização dos procedimentos de baixa complexidade.

O Gráfico 4 apresenta a relação entre o custo, a complexidade dos procedimentos e a faixa etária.

Gráfico 4 - Relação entre o custo, a complexidade dos procedimentos, e a faixa etária. Instituto Curitiba de Saúde, 2002-2004



Quando se compara a faixa etária com os custos relacionados à complexidade dos procedimentos, nota-se que, para os procedimentos de alta complexidade, a faixa etária que apresenta os maiores custos é a dos 50-59 anos, seguida dos 60-69

anos. Para os procedimentos de baixa e média complexidade, a faixa etária que apresenta os maiores custos é a dos 60-69 anos.

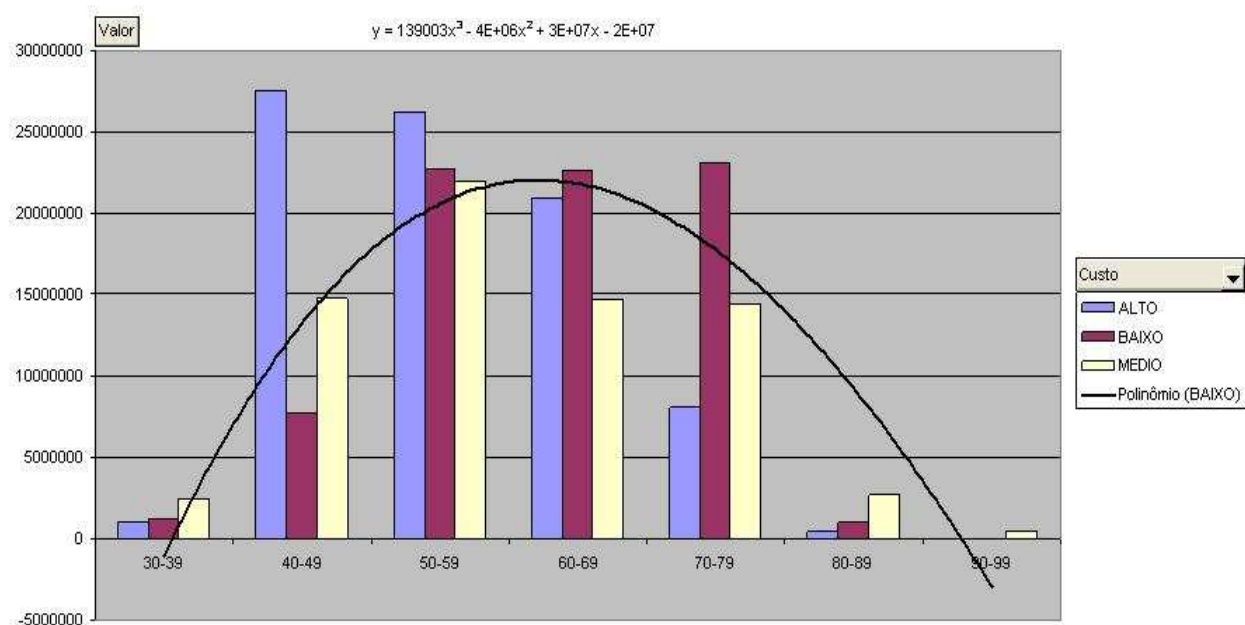
Ampliando esta discussão, o Gráfico 5 apresenta a relação entre o custo dos procedimentos e a faixa etária, além de uma linha de tendência polinomial.

Percebe-se que os procedimentos de alto custo são mais utilizados pelos usuários pertencentes às faixas etárias dos 40-49 anos e dos 50-59 anos. Os procedimentos de baixo custo são mais utilizados pelos usuários pertencentes às faixas etárias dos 60-69 anos e dos 70-79 anos. Por sua vez, os procedimentos de médio custo são mais utilizados pelos usuários pertencentes às faixas etárias dos 30-39 anos, 80-89 anos e 90-99 anos.

A linha de tendência apresenta seu pico entre as faixas etárias de 50-59 anos e 60-69 anos. Isto significa que, para a população desta base de dados, a faixa etária que vai dos 50 aos 69 anos é a que apresenta maiores custos relacionados a procedimentos de baixa, média e alta complexidades.

Assim pode-se pensar que a população-alvo de programas de gerenciamento de casos deve ser aquela que se encontra numa faixa etária anterior a esta, ou seja, anterior aos 50 anos, a fim de atuar de maneira pró-ativa, realizando detecção precoce, evitando o desenvolvimento da doença. Desta maneira, usuários e operadora de saúde podem ser beneficiados; esta, com a possibilidade de gerir técnica e eficientemente seus custos, aqueles, com melhora da sua condição de saúde e qualidade de vida.

Gráfico 5 - Relação entre o custo dos procedimentos e a faixa etária. Instituto Curitiba de Saúde, 2001-2005



É fundamental salientar que as projeções para 2020 apontam o Brasil como a sexta população mundial de idosos. Sendo assim, os profissionais da saúde devem estar cada vez mais familiarizados com as modificações estruturais, funcionais, e hemodinâmicas relacionadas ao envelhecimento. Ainda, com a finalidade de melhor conduzir o tratamento destes idosos, devem basear suas práticas em informações obtidas por meio de evidências, a fim de estabelecer condutas padronizadas para esta faixa etária, que visem não só o tratamento, mas também a prevenção dos principais agravos característicos da terceira idade (LORGA, 2002).

Adicionalmente, a discussão acerca da crucial importância da mudança de paradigma assistencial também merece atenção especial. Os dados apresentados nos gráficos evidenciam que o atual sistema de assistência à saúde está equivocado, por focar na doença, e não na promoção da saúde.

Miranda (2004) salienta que, de acordo com a Organização Mundial da Saúde, dentre os fatores que fazem com que um indivíduo passe dos 65 anos, 53% estão ligados ao estilo de vida; 20% ao meio ambiente; 17% à herança genética; e, apenas 10% estão ligados à assistência médica, foco do sistema de saúde público e suplementar no Brasil. Sendo assim, parece óbvio que as estratégias preventivas

deveriam ser mais exploradas, entretanto, para algumas operadoras de saúde, prevenção é custo (MIRANDA, 2004, p.5).

Este mesmo autor considera que, numa tentativa de amenizar as questões relativas à grande utilização dos sistemas de saúde, que os onera a ponto de perceber-se a limitação do acesso da população aos seus serviços. Ainda que estratégias de 'marketing positivo' devem ser a chave para a mudança do paradigma assistencial curativista, para um outro baseado na promoção da saúde e prevenção de doenças. Salaria que:

[...] o consumidor de planos de saúde é "bombardeado" com um marketing focado na doença. É a rede abundante de médicos, de clínicas e de hospitais, todos dotados com equipamentos de alta tecnologia. Claro que ao comprar um plano de saúde ele irá à busca disto, porque é exatamente isto que ele foi induzido a consumir pela propaganda. Em qualquer segmento econômico sabemos que o consumidor é muito mais dependente da oferta do que indutor desta. São as empresas, de modo geral, que determinam o quê o consumidor deve comprar e não o contrário. Está na hora, portanto, do consumidor passar a ser induzido a contratar planos de saúde e não planos de doença, como até agora fazem. É simples, mas com certeza este é o caminho para o verdadeiro gerenciamento da saúde (MIRANDA, 2004, p.5).

Goldsmith, Hutchson e Hurley (2004) comentam sobre os estudos publicados em 1986 por Louise Russell, que tratavam da avaliação dos benefícios de saúde e econômicos das intervenções preventivas, e testava a validação da hipótese de que a prevenção leva à economia. Ao final dos seus estudos, concluiu que a prevenção raramente está ligada a menores custos, sendo que os fatores que mais influenciam o impacto econômico da prevenção são: o tamanho da população que recebe as intervenções preventivas; a frequência com a qual as intervenções preventivas são repetidas; a magnitude e incerteza dos riscos das intervenções preventivas; a valorização das intervenções preventivas pela população que as recebe; como estes valores podem ser comparados a outros resultados, e o tempo que decorre dos investimentos até que os benefícios possam ser evidenciados.

Estes mesmos autores salientam que a prevenção e a cura possuem diferentes objetivos, e servem para diferentes fins. A prevenção tem por finalidade a manutenção da saúde, enquanto a cura tem como objetivo o alívio da dor e sofrimento. Pelo fato de serem meios para diferentes fins, não necessariamente devem competir por recursos (GOLDSMITH; HUTCHSON; HURLEY, 2004).

Deve-se também considerar que é impossível substituir o atual sistema focado na doença no seu papel de atender a casos de urgência e a toda situação

que envolva risco de vida, ou tratamento de doenças já instaladas no organismo dos indivíduos. No entanto, é preciso questionar as técnicas de diagnóstico e tratamento, e estabelecer ações de promoção de saúde, que contribuem para diminuir o número de indivíduos doentes, porém, deve-se ter clareza de que quando se caracteriza a necessidade de tratamento de doenças, este sistema focado na doença é necessário (MIRANDA, 2004).

É neste sentido que ações de gerenciamento em saúde devem considerar casos específicos, por meio da identificação de usuários com risco potencial. Assim pode-se obter um perfil destes usuários de modo a estabelecer a necessidade de recursos para a sua assistência, desenvolver um programa de assistência monitorizado que poderá auxiliar a reduzir seu risco para doenças e hospitalização (ROTHMAN; WAGNER, 2003), contribuindo para a gestão dos recursos do sistema de saúde.

4.3 Resultados da Mineração dos Dados

O algoritmo Apriori foi aplicado à base de dados preparada para a pesquisa. O Quadro 15 apresenta alguns exemplos das regras geradas.

Tomando-se como exemplo a primeira regra gerada, a sua leitura é feita da seguinte maneira: 'dos casos desta amostra, 11,1% têm probabilidade de apresentar o código de procedimento 32080115 (densitometria óssea duo energética em coluna vertebral) no seu histórico; destes, 81,6% têm probabilidade de serem do sexo feminino'.

Uma vez que neste experimento o número de regras geradas foi muito grande, a análise foi dificultada. Mesmo assim, percebeu-se que as regras geradas não associavam casos que incluíam eventos cardíacos associados com outros eventos cardíacos de alto custo e de alta complexidade.

Quadro 15 – Demonstrativo das regras geradas pelo algoritmo Apriori quando aplicado ao arquivo que relacionava os usuários aos seus procedimentos

Conseqüente da Regra	Antecedente da Regra	Suporte e Confiança
F ←	32080115	(11.1%, 81.6%)
87020300 ←	87020319	(16.2%, 81.3%)
28130367 ←	28100093 28040481	(10.2%, 83.6%)
28130367 ←	28100093 28010973	(10.4%, 83.9%)
28010540 ←	28100093 28040481	(10.2%, 86.3%)
28010540 ←	28100093 28010973	(10.4%, 87.9%)
28010973 ←	28100093 28040481	(10.2%, 94.6%)
28040481 ←	28100093 28010973	(10.4%, 92.2%)
10022 ←	28100093 28040481	(10.2%, 80.4%)
10014 ←	28100093 28040481	(10.2%, 81.0%)
10022 ←	28100093 28010973	(10.4%, 80.1%)

Como já explicitado anteriormente, decidiu-se, então aplicar o algoritmo Apriori na seleção dos 76 casos, que tinham pelo menos um evento com o código de dígito inicial 40. O algoritmo gerou um arquivo texto com 23.443 regras, com 113MB. Um exemplo das regras geradas para esta seleção dos dados pode ser verificado no Quadro 16.

Quadro 16 - Demonstrativo das regras geradas pelo algoritmo Apriori quando aplicado às variáveis provenientes dos casos que continham pelo menos um procedimento que iniciava com o dígito 40

Conseqüente da Regra	Antecedente da Regra	Suporte e Confiança
40060012 ←	27040399	(10.5%, 50.0%)
40060012 ←	27040062	(10.5%, 50.0%)
40060012 ←	87030063	(10.5%, 50.0%)
40040143 ←	87030063	(10.5%, 50.0%)
23020148 ←	23010037	(10.5%, 62.5%)
27020061 ←	27040399	(10.5%, 62.5%)
27020061 ←	27040062	(10.5%, 62.5%)
27020061 ←	28011120	(10.5%, 50.0%)

Tomando-se como exemplo a primeira regra gerada, a sua leitura é feita da seguinte maneira: 'dos casos desta amostra, 10,5% têm probabilidade de apresentar o código de procedimento 27040399 (TGP ou ALP por componente hemoterápico) no seu histórico; destes, 50,0% têm probabilidade de apresentarem o código 40060012 (perfusionista) associado.

Após esta fase, foi iniciada a fase de Pós-Processamento.

4.4 Resultados do Pós-Processamento

A fim de tornar as regras geradas pelo algoritmo Apriori compreensíveis aos profissionais da saúde, estas foram transferidas para o *software* Microsoft Excel®, considerando que, primariamente, encontravam-se no formato $Y \leftarrow X$ (suporte%, confiança%).

Primeiramente, os códigos que constavam na Tabela ICS de procedimentos foram substituídos pelas suas descrições, ou abreviações. O Quadro 17 apresenta exemplos dessas regras.

Num segundo momento, foram definidos junto com os especialistas os eventos significativos, extraídos das próprias regras geradas pelo algoritmo Apriori.

Salienta-se que estes eventos encontravam-se no conseqüente da regra. Considerando-se a forma $Y \leftarrow X$; os eventos significativos são representados pelo Y. São eventos significativos: cateterismo cardíaco; angioplastia de vaso único; implante de *stent*; cintilografia do miocárdio; valores referenciais de cateterismo cardíaco; valores referenciais de revascularização do miocárdio; e, consulta de emergência.

A etapa de pós-processamento, nesta pesquisa, foi dividida em duas fases: a etapa de análise subjetiva e a etapa de análise sob a perspectiva da confiança.

Quadro 17 - Demonstrativo de exemplos de regras geradas pelo Apriori

EVENTO SIGNIFICATIVO	EVENTO	EVENTO	EVENTO	SUPORTE	CONFIANÇA
INTERNACAO	← REVASC_MIOCARDIO			(11.8%,	55.6%)
INTERNACAO	← VR_REVASC_MIOCARDIO	CONS_AMB		(11.8%,	55.6%)
INTERNACAO	← REVASC_MIOCARDIO	CAT_ART_RADIAL		(10.5%,	50.0%)
INTERNACAO	← REVASC_MIOCARDIO	UREIA		(10.5%,	50.0%)
INTERNACAO	← REVASC_MIOCARDIO	POTASSIO		(10.5%,	50.0%)
INTERNACAO	← REVASC_MIOCARDIO	HEMOGRAMA_COMPL		(11.8%,	66.7%)
INTERNACAO	← REVASC_MIOCARDIO	CREATININA		(11.8%,	66.7%)
INTERNACAO	← REVASC_MIOCARDIO	GLICOSE		(11.8%,	77.8%)
INTERNACAO	← VR_REVASC_MIOCARDIO	ROTINA_URINA	CONS_AMB	(13.2%,	70.0%)
INTERNACAO	← VR_REVASC_MIOCARDIO	TRIGLICERIDEOS	CONS_AMB	(11.8%,	55.6%)
INTERNACAO	← VR_REVASC_MIOCARDIO	COLESTEROL_TOTAL	CONS_AMB	(13.2%,	50.0%)
INTERNACAO	← VR_REVASC_MIOCARDIO	GLICOSE	CONS_AMB	(11.8%,	77.8%)
HEMOTERAPIA	← REVASC_MIOCARDIO			(14.5%,	63.6%)
HEMOTERAPIA	← REVASC_MIOCARDIO	CAT_ART_RADIAL		(11.8%,	77.8%)
HEMOTERAPIA	← REVASC_MIOCARDIO	UREIA		(14.5%,	63.6%)
HEMOTERAPIA	← REVASC_MIOCARDIO	POTASSIO		(11.8%,	88.9%)
HEMOTERAPIA	← REVASC_MIOCARDIO	HEMOGRAMA_COMPL		(15.8%,	66.7%)
CREATININA	← REVASC_MIOCARDIO	CREATININA		(11.8%,	66.7%)
CREATININA	← REVASC_MIOCARDIO	GLICOSE		(15.8%,	50.0%)
CREATININA	← REVASC_MIOCARDIO	UREIA		(11.8%,	66.7%)
CREATININA	← REVASC_MIOCARDIO	POTASSIO		(15.8%,	50.0%)
CREATININA	← REVASC_MIOCARDIO	HEMOGRAMA_COMPL		(11.8%,	77.8%)
CREATININA	← REVASC_MIOCARDIO	CREATININA		(15.8%,	58.3%)
CREATININA	← REVASC_MIOCARDIO	GLICOSE		(11.8%,	66.7%)
CREATININA	← EXTRA_CORPOREA	DISSECCAO_VEIA		(15.8%,	50.0%)
CREATININA	← EXTRA_CORPOREA	CAT_ART_RADIAL		(11.8%,	66.7%)
CREATININA	← EXTRA_CORPOREA	CREATININA		(15.8%,	50.0%)
CREATININA	← EXTRA_CORPOREA	TX_REG_INTERN		(11.8%,	66.7%)
CREATININA	← EXTRA_CORPOREA	GLICOSE		(15.8%,	50.0%)
CREATININA	← EXTRA_CORPOREA	CONS_AMB		(11.8%,	77.8%)
CREATININA	← REVASC_MIOCARDIO			(17.1%,	53.8%)

4.4.1 Análise Subjetiva

Para cada evento significativo foi elaborado um quadro demonstrativo das associações com outros eventos, o suporte e a confiança.

As regras relevantes, após análise subjetiva, para o evento significativo 'Cateterismo Cardíaco' (Quadro 18) apresentaram suporte que variou de 10,5% até 18,4%, e confiança que variou de 70,0% a 88,9%. Tomando como exemplo a primeira regra considerada relevante pelos especialistas, a leitura deve ser feita da seguinte maneira: 'para esta amostra, 18,4% apresentavam nos seus históricos o procedimento de angioplastia de vaso único; destes, 78,6% estavam associados ao procedimento de cateterismo cardíaco'.

As regras para o evento significativo 'Angioplastia de Vaso Único' (Quadro 19) apresentaram suporte de 14,5%, e confiança de 90,9%. Tomando como exemplo a primeira regra considerada relevante pelos especialistas, a leitura deve ser feita da seguinte maneira: 'para esta amostra, 14,5% apresentaram nos seus históricos os procedimentos de cateterismo cardíaco associado a implante de *stent*; destes, 90,9% estavam associados ao procedimento de angioplastia de vaso único'.

Quadro 18 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Cateterismo Cardíaco

(continua)

EVENTO SIGNIFICATIVO		EVENTO	EVENTO	EVENTO	SUPORTE	CONF.
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO			(18.4%,	78.6%)
CATETERISMO_CARD	←	IMPLANTE_STENT			(18.4%,	78.6%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT		(17.1%,	76.9%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM		(14.5%,	72.7%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	M		(13.2%,	80.0%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	CONS_EMERG		(10.5%,	75.0%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	CONS_AMB		(14.5%,	81.8%)
CATETERISMO_CARD	←	IMPLANTE_STENT	TESTE_ERGOM		(14.5%,	72.7%)
CATETERISMO_CARD	←	IMPLANTE_STENT	M		(11.8%,	77.8%)
CATETERISMO_CARD	←	IMPLANTE_STENT	ECOCARDIO_BIDIM_DOPPLER		(10.5%,	87.5%)
CATETERISMO_CARD	←	IMPLANTE_STENT	CONS_EMERG		(10.5%,	75.0%)
CATETERISMO_CARD	←	IMPLANTE_STENT	CONS_AMB		(14.5%,	81.8%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	TESTE_ERGOM	(13.2%,	70.0%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	M	(11.8%,	77.8%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	CONS_AMB	(13.2%,	80.0%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	M	(11.8%,	77.8%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	CONS_AMB	(13.2%,	80.0%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	M	CONS_AMB	(10.5%,	87.5%)
CATETERISMO_CARD	←	IMPLANTE_STENT	TESTE_ERGOM	TX_SALA_OBS	(10.5%,	87.5%)
CATETERISMO_CARD	←	IMPLANTE_STENT	TESTE_ERGOM	M	(10.5%,	75.0%)
CATETERISMO_CARD	←	IMPLANTE_STENT	TESTE_ERGOM	ECG	(10.5%,	75.0%)
CATETERISMO_CARD	←	IMPLANTE_STENT	TESTE_ERGOM	CONS_AMB	(13.2%,	80.0%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	M	(11.8%,	77.8%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	INTERNACAO	(10.5%,	75.0%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	CONS_AMB	(13.2%,	80.0%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	M	(11.8%,	77.8%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	ECG	(10.5%,	75.0%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	CONS_AMB	(13.2%,	80.0%)
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	M	CONS_AMB	(10.5%,	87.5%)

Quadro 18 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Cateterismo Cardíaco (conclusão)

EVENTO SIGNIFICATIVO		EVENTO	EVENTO	EVENTO	SUPORTE	CONF.
CATETERISMO_CARD	←	IMPLANTE_STENT	TESTE_ERGOM	TX_SALA_OBS	(10.5%,	87.5%)
CATETERISMO_CARD	←	IMPLANTE_STENT	TESTE_ERGOM	M	(10.5%,	75.0%)
CATETERISMO_CARD	←	IMPLANTE_STENT	TESTE_ERGOM	ECG	(10.5%,	75.0%)
CATETERISMO_CARD	←	IMPLANTE_STENT	TESTE_ERGOM	CONS_AMB	(13.2%,	80.0%)
CATETERISMO_CARD	←	IMPLANTE_STENT	TX_SALA_OBS	CONS_AMB	(11.8%,	88.9%)

Quadro 19 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Angioplastia de Vaso Único

EVENTO SIGNIFICATIVO		EVENTO	EVENTO	SUPORTE	CONF.
ANGIOPL_VASO_UNICO	←	IMPLANTE_STENT	CATETERISMO_CARD	(14.5%,	90.9%)
ANGIOPL_VASO_UNICO	←	IMPLANTE_STENT	CONS_AMB	(14.5%,	90.9%)
ANGIOPL_VASO_UNICO	←	IMPLANTE_STENT	TESTE_ERGOM	(14.5%,	90.9%)
ANGIOPL_VASO_UNICO	←	IMPLANTE_STENT	CATETERISMO_CARD	(14.5%,	90.9%)
ANGIOPL_VASO_UNICO	←	IMPLANTE_STENT	CONS_AMB	(14.5%,	90.9%)

As regras relevantes, após análise subjetiva, para o evento significativo 'Implante de *Stent*' (Quadro 20) apresentam suporte que variou de 10,5% até 19,7%, e confiança que variou de 50,0% a 100,0%. Tomando como exemplo a primeira regra considerada relevante pelos especialistas, a leitura deve ser feita da seguinte maneira: 'para esta amostra, 14,5% apresentaram nos seus históricos o procedimento de angioplastia de vaso único associada ao procedimento de teste ergométrico; destes, 90,9% estavam associados ao procedimento de implante de *stent*'.

As regras relevantes, após análise subjetiva, para o evento significativo 'Cintilografia do Miocárdio' (Quadro 21) apresentam suporte que variou de 10,5% até 13,2%, e confiança de 50,0%. Tomando como exemplo a primeira regra considerada relevante pelos especialistas, a leitura deve ser feita da seguinte maneira: 'para esta amostra, 14,5% apresentaram nos seus históricos o procedimento de Cateterismo Cardíaco associado a teste ergométrico e implante de *stent*, destes, 50,0% estavam associados ao procedimento cintilografia do miocárdio'.

Quadro 20 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Implante de Stent

(continua)

EVENTO SIGNIFICATIVO	EVENTO	EVENTO	EVENTO	SUPORTE	CONF.
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	TESTE_ERGOM		(14.5%,	90.9%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	TX_SALA_OBS		(11.8%,	100.0%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	M		(13.2%,	90.0%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	CATETERISMO_CARD		(14.5%,	90.9%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	CONS_EMERG		(10.5%,	87.5%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	CONS_AMB		(14.5%,	90.9%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	TESTE_ERGOM		(14.5%,	90.9%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	M		(13.2%,	90.0%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	M	(11.8%,	88.9%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	CATETERISMO_CARD	(10.5%,	87.5%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	TX_SALA_OBS	CATETERISMO_CARD	(10.5%,	100.0%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	TX_SALA_OBS	CONS_AMB	(10.5%,	100.0%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	M	CATETERISMO_CARD	(10.5%,	87.5%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	M	CONS_AMB	(10.5%,	87.5%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	CATETERISMO_CARD	CONS_AMB	(11.8%,	88.9%)
IMPLANTE_STENT ←	MAPEAMENTO_RETINA_MONO	CATETERISMO_CARD	CONS_AMB	(10.5%,	50.0%)
IMPLANTE_STENT ←	CINTILOGR_MIOC_PERFUS_REP_STRESS	TESTE_ERGOM	CONS_AMB	(11.8%,	55.6%)
IMPLANTE_STENT ←	TESTE_ERGOM	TX_SALA_OBS	CATETERISMO_CARD	(14.5%,	63.6%)
IMPLANTE_STENT ←	TESTE_ERGOM	TX_SALA_OBS	CONS_AMB	(19.7%,	53.3%)
IMPLANTE_STENT ←	TESTE_ERGOM	M	ECOCARDIO_BIDIM_DOPPLER	(13.2%,	50.0%)
IMPLANTE_STENT ←	TESTE_ERGOM	M	CATETERISMO_CARD	(11.8%,	66.7%)
IMPLANTE_STENT ←	TESTE_ERGOM	M	CONS_AMB	(18.4%,	50.0%)
IMPLANTE_STENT ←	TESTE_ERGOM	ECOCARDIO_BIDIM_DOPPLER	CATETERISMO_CARD	(14.5%,	54.5%)
IMPLANTE_STENT ←	TX_SALA_OBS	ECOCARDIO_BIDIM_DOPPLER	CATETERISMO_CARD	(14.5%,	54.5%)
IMPLANTE_STENT ←	TX_SALA_OBS	ECOCARDIO_BIDIM_DOPPLER	CONS_AMB	(15.8%,	50.0%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	M	(11.8%,	88.9%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	CATETERISMO_CARD	(10.5%,	87.5%)
IMPLANTE_STENT ←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	CONS_AMB	(13.2%,	90.0%)

Quadro 20 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Implante de *Stent* (conclusão)

EVENTO SIGNIFICATIVO	EVENTO	EVENTO	EVENTO	SUPORTE	CONF.
IMPLANTE_STENT ← ANGIOPL_VASO_UNICO		TX_SALA_OBS	CONS_AMB	(10.5%,	100.0%)
IMPLANTE_STENT ← ANGIOPL_VASO_UNICO		M	CATETERISMO_CARD	(10.5%,	87.5%)
IMPLANTE_STENT ← ANGIOPL_VASO_UNICO		M	CONS_AMB	(10.5%,	87.5%)
IMPLANTE_STENT ← ANGIOPL_VASO_UNICO		CATETERISMO_CARD	CONS_AMB	(11.8%,	88.9%)
IMPLANTE_STENT ← MAPEAMENTO_RETINA_MONO		CATETERISMO_CARD	CONS_AMB	(10.5%,	50.0%)
IMPLANTE_STENT ← CINTILOGR_MIOC_PERFUS_REP_STRESS		TESTE_ERGOM	CONS_AMB	(11.8%,	55.6%)
IMPLANTE_STENT ← TESTE_ERGOM		TX_SALA_OBS	CATETERISMO_CARD	(14.5%,	63.6%)
IMPLANTE_STENT ← TESTE_ERGOM		M	ECOCARDIO_BIDIM_DOPPLER	(13.2%,	50.0%)
IMPLANTE_STENT ← TESTE_ERGOM		M	CATETERISMO_CARD	(11.8%,	66.7%)
IMPLANTE_STENT ← TESTE_ERGOM		M	CONS_AMB	(18.4%,	50.0%)
IMPLANTE_STENT ← TESTE_ERGOM		ECOCARDIO_BIDIM_DOPPLER	CATETERISMO_CARD	(14.5%,	54.5%)
IMPLANTE_STENT ← TX_SALA_OBS		ECOCARDIO_BIDIM_DOPPLER	CATETERISMO_CARD	(14.5%,	54.5%)
IMPLANTE_STENT ← TX_SALA_OBS		ECOCARDIO_BIDIM_DOPPLER	CONS_AMB	(15.8%,	50.0%)

Quadro 21 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Cintilografia do Miocárdio

EVENTO SIGNIFICATIVO	EVENTO	EVENTO	EVENTO	SUPORTE	CONF.
CINTILOGR_MIOC_PERFUS_REP_STRESS ←	IMPLANTE_STENT	TESTE_ERGOM	CATETERISMO_CARD	(10.5%,	50.0%)
CINTILOGR_MIOC_PERFUS_REP_STRESS ←	IMPLANTE_STENT	TESTE_ERGOM	CONS_AMB	(13.2%,	50.0%)
CINTILOGR_MIOC_PERFUS_REP_STRESS ←	IMPLANTE_STENT	TESTE_ERGOM	TX_SALA_OBS	(10.5%,	50.0%)
CINTILOGR_MIOC_PERFUS_REP_STRESS ←	IMPLANTE_STENT	TESTE_ERGOM	CATETERISMO_CARD	(10.5%,	50.0%)
CINTILOGR_MIOC_PERFUS_REP_STRESS ←	IMPLANTE_STENT	TESTE_ERGOM	CONS_AMB	(13.2%,	50.0%)

As regras relevantes, após análise subjetiva, para o evento significativo 'Valores Referenciais de Cateterismo Cardíaco' (Quadro 22) apresentaram suporte de 10,5% e confiança de 50,0%. Tomando como exemplo a primeira regra considerada relevante pelos especialistas, a leitura deve ser feita da seguinte maneira: 'para esta amostra, 10,5% apresentaram nos seus históricos o procedimento de consulta ambulatorial associada ao procedimento hemoglobina glicosilada, de mapeamento de retina e de microalbuminúria; destes, 50,0% estavam associados ao procedimento de valores referenciais de cateterismo cardíaco'.

As regras relevantes, após análise subjetiva, para o evento significativo 'Valores Referenciais de Revascularização do Miocárdio' (Quadro 23) apresentaram suporte que variou de 10,5% até 11,8%, e confiança que variou de 50,0% a 100,0%. Tomando como exemplo a primeira regra considerada relevante pelos especialistas, a leitura deve ser feita da seguinte maneira: 'para esta amostra, 14,5% apresentaram nos seus históricos o sexo masculino associado ao procedimento de hemoglobina glicosilada; destes, 100,0% estavam associados ao procedimento de valores referenciais de revascularização do miocárdio'.

As regras relevantes, após análise subjetiva, para o evento significativo 'Consulta de Emergência' (Quadro 24) apresentaram suporte que variou de 10,5% até 18,4%, e confiança que variou de 50,0% a 75,0%. Tomando como exemplo a primeira regra considerada relevante pelos especialistas, sua leitura é feita da seguinte maneira: 'para esta amostra, 10,5% apresentaram nos seus históricos o procedimento de revascularização do miocárdio; destes, 75,0% estavam associados a uma consulta de emergência'.

Quadro 22 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Valores Referenciais de Cateterismo Cardíaco

EVENTO SIGNIFICATIVO	EVENTO	EVENTO	EVENTO	EVENTO	SUORTE	CONF.
VR_CATETERISMO_CARD ←	MICROALBUMINURIA	MAPEAMENTO_RETINA	HB_GLICOSILADA	CONS_AMB	(10.5%,	50.0%)
VR_CATETERISMO_CARD ←	MAPEAMENTO_RETINA	HB_GLICOSILADA	TESTE_ERGOM	GLICOSE	(10.5%,	50.0%)
VR_CATETERISMO_CARD ←	MAPEAMENTO_RETINA	HB_GLICOSILADA	TESTE_ERGOM	CONS_AMB	(10.5%,	50.0%)
VR_CATETERISMO_CARD ←	TX_PROC_AMB	TESTE_ERGOM	ECOCARDIO_BIDIM_DOPPLER	CONS_AMB	(10.5%,	50.0%)

Quadro 23 - Demonstrativo das regras relevantes para o evento significativo Valores Referenciais de Revascularização do Miocárdio

EVENTO SIGNIFICATIVO	EVENTO	EVENTO	EVENTO	SUORTE	CONF.
VR_REVASC_MIOCARDIO ←	HB_GLICOSILADA	M		(11.8%,	100.0%)
VR_REVASC_MIOCARDIO ←	MICROALBUMINURIA	MAPEAMENTO_RETINA	HB_GLICOSILADA	(10.5%,	50.0%)
VR_REVASC_MIOCARDIO ←	MAPEAMENTO_RETINA	HB_GLICOSILADA	TESTE_ERGOM	(10.5%,	75.0%)
VR_REVASC_MIOCARDIO ←	HB_GLICOSILADA	TESTE_ERGOM	ECOCARDIO_BIDIM_DOPPLER	(10.5%,	62.5%)

Quadro 24 - Demonstrativo das regras relevantes após análise subjetiva para o evento significativo Consulta de Emergência

EVENTO SIGNIFICATIVO		EVENTO	EVENTO	EVENTO	SUPORTE	CONF.
CONS_EMERG	←	REVASC_MIOCARDIO			(10.5%,	75.0%)
CONS_EMERG	←	VR_REVASC_MIOCARDIO	COLESTEROL_HDL		(11.8%,	66.7%)
CONS_EMERG	←	VR_REVASC_MIOCARDIO	CONS_AMB		(13.2%,	50.0%)
CONS_EMERG	←	ELETROFISIOLOGIA_CARD			(10.5%,	75.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO			(18.4%,	57.1%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT		(17.1%,	53.8%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM		(14.5%,	54.5%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	CATETERISMO_CARD		(14.5%,	54.5%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	CONS_AMB		(14.5%,	54.5%)
CONS_EMERG	←	IMPLANTE_STENT	TESTE_ERGOM		(14.5%,	54.5%)
CONS_EMERG	←	IMPLANTE_STENT	ECOCARDIO_BIDIM_DOPPLER		(10.5%,	75.0%)
CONS_EMERG	←	IMPLANTE_STENT	CATETERISMO_CARD		(14.5%,	54.5%)
CONS_EMERG	←	IMPLANTE_STENT	GLICOSE		(14.5%,	72.7%)
CONS_EMERG	←	IMPLANTE_STENT	CONS_AMB		(14.5%,	54.5%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	TESTE_ERGOM	(13.2%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	CATETERISMO_CARD	(13.2%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	CONS_AMB	(13.2%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	CATETERISMO_CARD	(10.5%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	CONS_AMB	(13.2%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	TX_SALA_OBS	CATETERISMO_CARD	(10.5%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	CATETERISMO_CARD	CONS_AMB	(11.8%,	55.6%)
CONS_EMERG	←	IMPLANTE_STENT	TESTE_ERGOM	CATETERISMO_CARD	(10.5%,	50.0%)
CONS_EMERG	←	IMPLANTE_STENT	TESTE_ERGOM	CONS_AMB	(13.2%,	50.0%)
CONS_EMERG	←	IMPLANTE_STENT	TX_SALA_OBS	CATETERISMO_CARD	(11.8%,	55.6%)
CONS_EMERG	←	IMPLANTE_STENT	CATETERISMO_CARD	CONS_AMB	(11.8%,	55.6%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	TESTE_ERGOM	(13.2%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	CATETERISMO_CARD	(13.2%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	IMPLANTE_STENT	CONS_AMB	(13.2%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	CATETERISMO_CARD	(10.5%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	TESTE_ERGOM	CONS_AMB	(13.2%,	50.0%)
CONS_EMERG	←	ANGIOPL_VASO_UNICO	CATETERISMO_CARD	CONS_AMB	(11.8%,	55.6%)

4.4.2 Análise sob a Perspectiva da Confiança

Após a análise subjetiva dos especialistas, as regras foram analisadas sob a perspectiva da confiança, ainda com auxílio dos especialistas, como já explicitado o capítulo 3.

A Tabela 1 apresenta os resultados provenientes da análise subjetiva e sob a perspectiva da confiança do evento significativo ‘Consulta de Emergência’, e o Quadro 25 as regras relevantes após análise subjetiva e sob a perspectiva da confiança.

Tabela 1 – Resultados da análise subjetiva e sob a perspectiva da confiança para o evento significativo ‘Consulta de Emergência’

	IRRELEV	INSIGNIF	RELEV	RED	TOTAL
ANÁLISE SUBJETIVA	162	74	24		260
ANÁLISE CONFIANÇA	162		8	90	260

Na análise subjetiva, das 260 regras geradas relacionadas ao evento significativo ‘Consulta de Emergência’, 162 foram classificadas como irrelevantes, 74 como insignificantes, e 24 como relevantes. Na análise sob a perspectiva da confiança, das 24 regras consideradas relevantes na análise subjetiva, apenas 8 foram confirmadas como relevantes, as demais regras primariamente relevantes foram classificadas como redundantes (90), juntamente com as insignificantes.

Quadro 25 - Demonstrativo das regras relevantes para o evento significativo Consulta de Emergência após análise subjetiva e sob a perspectiva da confiança

EVENTO SIGNIFICATIVO	EVENTO	EVENTO	EVENTO	SUP.	CONF.
CONS_EMERG	← REVASC_MIOCARDIO			(10.5%,	75.0%)
CONS_EMERG	← ELETROFISIOLOGIA_CARD			(10.5%,	75.0%)
CONS_EMERG	← IMPLANTE_STENT	ECOCARDIO_BIDIM_DOPPLER		(10.5%,	75.0%)
CONS_EMERG	← IMPLANTE_STENT	GLICOSE		(14.5%,	72.7%)
CONS_EMERG	← ANGIOPL_VASO_UNICO	IMPLANTE_STENT	TESTE_ERGOM	(13.2%,	50.0%)
CONS_EMERG	← ANGIOPL_VASO_UNICO	IMPLANTE_STENT	CATETERISMO_CARD	(13.2%,	50.0%)
CONS_EMERG	← ANGIOPL_VASO_UNICO	TESTE_ERGOM	CATETERISMO_CARD	(10.5%,	50.0%)
CONS_EMERG	← IMPLANTE_STENT	TESTE_ERGOM	CATETERISMO_CARD	(10.5%,	50.0%)

A seguir, são apresentadas as leituras das regras relevantes após a análise subjetiva e sob a perspectiva da confiança, para o evento significativo 'Consulta de Emergência', considerando a amostra desta pesquisa.

Regra 1 → 10,5% dos usuários apresentaram no seu histórico o procedimento revascularização do miocárdio; destes, 75% têm probabilidade de apresentar associação com uma consulta de emergência.

Regra 2 → 10,5% dos usuários apresentaram no seu histórico o procedimento eletrofisiologia cardíaca; destes, 75% têm probabilidade de apresentar associação com uma consulta de emergência.

Regra 3 → 10,5% dos usuários apresentaram no seu histórico o procedimento ecocardiograma bidimensional com Doppler associado ao procedimento de implante de *stent*; destes, 75% têm probabilidade de apresentar associação com uma consulta de emergência.

Regra 4 → 10,5% dos usuários apresentaram no seu histórico o procedimento de mensuração de glicose associado ao procedimento de implante de *stent*; destes, 75% têm probabilidade de apresentar associação com uma consulta de emergência.

Regra 5 → 13,2% dos usuários apresentaram no seu histórico o procedimento teste ergométrico associado ao procedimento de implante de *stent* e à angioplastia de vaso único; destes, 50% têm probabilidade de apresentar associação com uma consulta de emergência.

Regra 6 → 13,2% dos usuários apresentaram no seu histórico o procedimento cateterismo cardíaco associado ao procedimento de implante de *stent* e à angioplastia de vaso único; destes, 50% têm probabilidade de apresentar associação com uma consulta de emergência.

Regra 7 → 10,5% dos usuários apresentaram no seu histórico o procedimento de cateterismo cardíaco associado ao procedimento de teste ergométrico e a angioplastia de vaso único, 50% têm probabilidade de apresentar associação com uma consulta de emergência.

Regra 8 → 10,5% dos usuários apresentaram no seu histórico o procedimento cateterismo cardíaco associado ao procedimento teste ergométrico e ao implante de *stent*; destes, 50% têm probabilidade de apresentar associação com uma consulta de emergência.

A elevada associação de procedimentos cardiovasculares com consultas de emergência pode estar relacionada ou à condição grave do usuário, ou à falta de monitoração do usuário que foi submetido a esse tipo de procedimento, uma vez que poucos procedimentos ou exames são solicitados de maneira freqüente para o diagnóstico de problemas cardiovasculares, como mostram as regras apresentadas pelo algoritmo.

O fato de um evento de 'consulta de emergência' estar associado a um evento relacionado a um procedimento cardíaco evidencia a importância de se estabelecer um protocolo de acompanhamento a usuários submetidos a procedimentos cardiovasculares. O acompanhamento ambulatorial pós-procedimento cardíaco deve ser uma prioridade, porque traz benefícios não só para o usuário, mas também para a gestão de organizações pagadoras e prestadoras de serviços de saúde.

Além disso, a adoção de programas de gerenciamento de casos no nível ambulatorial pode auxiliar no monitoramento dos usuários e incorrer em economia de cerca de 15 a 30% em internações e consultas de emergência (NCCC, 2000). Pode contribuir também pelo acompanhamento da equipe multidisciplinar de gerenciamento, que detém o conhecimento sobre usuário como um todo. Esta estratégia deve ser desenvolvida a partir da sensibilização dos médicos para consultas mais amplas, ou seja, da necessidade de anamnese do indivíduo como um todo, além da disponibilização de uma rede de apoio para que possa referenciar usuários com potencial para algum risco, e assim monitorá-los.

É importante salientar que as consultas ambulatoriais apresentam a característica de serem previamente agendadas pelo usuário com o profissional de sua escolha, ou que o vem acompanhando. Ou seja, o usuário marca previamente a consulta com o médico que o conhece, seja por acompanhamento clínico ou por ter realizado o procedimento cirúrgico. De outra maneira, na consulta de emergência, o usuário se depara com um profissional que não o conhece. Neste último caso, existe uma grande chance de que este usuário seja submetido a diversos procedimentos diagnósticos ou mesmo seja internado.

Com a diminuição do tempo de permanência nos hospitais, o gerenciamento de casos ambulatorial é visto como uma ferramenta que pode auxiliar a evitar os

dispendiosos episódios de cuidados críticos, e também garantir a continuidade do processo assistencial.

No final da década de 90, uma das grandes empresas de seguro saúde norte americana apresentou uma economia de cerca de oito milhões de dólares, referente a um trabalho extensivo de gerenciamento de casos junto a dezesseis clínicas ambulatoriais do sul da Califórnia. Relatam a utilização de um instrumento de mensuração de risco, bem como dados farmacêuticos, entre outros, que levaram à identificação dos pacientes que estavam em risco de declínio do seu estado de saúde devido à grande demanda por hospitalização ou consultas de emergência. A partir de então, trabalharam para intervir no processo saúde-doença destes pacientes de forma pró-ativa, antes do que normalmente aconteceria (NCCC, 2000).

A coordenação da assistência prestada pelos médicos, nos diversos ambientes de assistência à saúde que o usuário frequenta para fins de definição diagnóstica, é de crítica importância em se tratando da assistência a usuários com doenças crônicas. Os usuários valorizam muito o fato de terem um médico ou equipe que esteja ciente do “todo” do seu problema, considerando o fato que também a sua família ou cuidador pode entrar em contato com este profissional ou equipe. Um sistema em que se possa desenvolver tal coordenação, ou gerenciamento, está associado a melhores resultados gerenciais e satisfação por parte do usuário (ROTHMAN; WAGNER, 2003).

É fundamental que as organizações pagadoras de serviços de saúde trabalhem em parceria com os prestadores, a fim de organizar o acompanhamento de usuários que apresentem condições de saúde e doença que estejam relacionadas a um maior risco de internamento e procedimentos de alto custo, como os programas de gerenciamento de casos. Além disso, parcerias podem ser desenvolvidas com outros profissionais da saúde que possam se encarregar da monitorização destes usuários, a fim de garantir a continuidade da assistência, identificando situações que possam desencadear problemas de saúde, tais como ansiedade, conflitos familiares, segurança residencial, necessidade de transporte ou equipamentos (NCCC, 2000).

Entretanto, sabe-se que por trás da prática desorganizada de assistência à saúde que se verifica atualmente existe um modelo de atenção à saúde ultrapassado, curativista, que induz a realização de vários procedimentos, com

pouca resolubilidade e qualidade. Para que ações de gerenciamento de caso sejam exeqüíveis e resolutivas, é realmente necessário que haja uma mudança do atual paradigma assistencial, rumo a uma prática pró-ativa, de promoção à saúde e prevenção de agravos, para o bem dos indivíduos que necessitam de cuidados de saúde, como para o próprio sistema de saúde (NCCC, 2000; ROTHMAN; WAGNER, 2003; MIRANDA, 2004).

A literatura indica que algumas tentativas vêm sendo realizadas a fim de expandir a assistência à saúde prestada somente em ambientes reconhecidamente destinados a este tipo de serviço. Como exemplo, já no início da década de 90 em Chicago, organizações de assistência à saúde trabalhavam com equipes multidisciplinares em programas de gerenciamento de casos, estendendo-os para além das suas portas. Procuravam desenvolver ações integradas com recursos da comunidade, e assim conseguiram trabalhar com promoção à saúde, prevenção de agravos, contribuindo para a detecção precoce de problemas como hipertensão, depressão e diabetes (NCCC, 2000).

A Tabela 2 apresenta os resultados provenientes da análise subjetiva e sob a perspectiva da confiança do evento significativo 'Valores Referenciais de Revascularização do Miocárdio'. Por sua vez, o Quadro 26 apresenta as regras relevantes após as etapas de análise subjetiva e sob a perspectiva da confiança.

Tabela 2 – Resultados da análise subjetiva e sob a perspectiva da confiança do evento significativo 'Valores Referenciais Revascularização do Miocárdio'

	IRRELEV	INSIGNIF	RELEV	RED	TOTAL
ANÁLISE SUBJETIVA	481	153	5		639
ANÁLISE CONFIANÇA	481		4	154	639

Na análise subjetiva, das 639 regras geradas relacionadas ao evento significativo 'Valores Referenciais Revascularização do Miocárdio', 481 foram classificadas como irrelevantes, 153 como insignificantes, e 5 como relevantes. Na análise sob a perspectiva da confiança, das 5 regras consideradas relevantes na análise subjetiva, 4 foram confirmadas como relevantes, as demais regras primariamente relevantes foram classificadas como redundantes (154), juntamente com as insignificantes.

Quadro 26 - Demonstrativo das regras relevantes do evento significativo Valores Referenciais Revascularização do Miocárdio após análise subjetiva e sob a perspectiva da confiança

EVENTO SIGNIFICATIVO	EVENTO	EVENTO	EVENTO	SUP.	CONF.
VR_REVASC_MIOCARDIO	← CINTILOGR_MIOC_	HB_GLICOSILADA		(11.8%,	100.0%)
VR_REVASC_MIOCARDIO	← HB_GLICOSILADA	M		(11.8%,	100.0%)
VR_REVASC_MIOCARDIO	← MAPEAMENTO_RETINA	HB_GLICOSILADA	TESTE_ERGOM	(10.5%,	75.0%)
VR_REVASC_MIOCARDIO	← MICROALBUMINURIA	MAPEAMENTO_RETINA	HB_GLICOSILADA	(10.5%,	50.0%)

A seguir, são apresentadas as leituras das regras relevantes após as análises subjetiva e sob a perspectiva da confiança, para o evento significativo 'Valores Referenciais de Revascularização do Miocárdio', considerando a amostra deste estudo.

Regra 1 → 11,8% dos usuários apresentaram no seu histórico o procedimento de mensuração da hemoglobina glicosilada associado ao procedimento de cintilografia de miocárdio; destes, 100% têm probabilidade de apresentar associação com valores referenciais de revascularização do miocárdio.

Regra 2 → 11,8% dos usuários apresentaram no seu histórico o fato de serem do sexo masculino associado ao procedimento de mensuração da hemoglobina glicosilada; destes, 100% têm probabilidade de apresentar associação com valores referenciais de revascularização do miocárdio.

Regra 3 → 10,5% dos usuários apresentaram no seu histórico o teste ergométrico associado ao procedimento de mensuração da hemoglobina glicosilada e ao mapeamento de retina; destes, 75% têm probabilidade de apresentar associação com valores referenciais de revascularização do miocárdio.

Regra 4 → 10,5% dos usuários apresentaram no seu histórico o procedimento de mensuração da hemoglobina glicosilada associada ao mapeamento de retina e ao de mensuração da microalbuminúria; destes, 50% têm probabilidade de apresentar associação com valores referenciais de revascularização do miocárdio.

É importante salientar que estas regras foram consideradas relevantes pelos especialistas porque demonstram a associação entre exames que estão relacionados à presença de diabetes (mensuração da hemoglobina glicosilada, o mapeamento da retina e a mensuração da microalbuminúria) e o procedimento de

revascularização do miocárdio. Outro aspecto a ser considerado é que, de acordo com os especialistas, estes exames não são solicitados como exames de avaliação periódica (*check ups*) para a população em geral, somente para aqueles indivíduos que já apresentam fatores de risco para diabetes ou doença coronariana já diagnosticadas.

A segunda regra, especificamente, foi considerada relevante, pois indica relação entre o sexo masculino, a mensuração da hemoglobina glicosilada e o procedimento de revascularização do miocárdio. Segundo os especialistas, esta é uma informação importante no sentido de voltar ações preventivas e de educação à saúde para a população do sexo masculino.

Embora o fato de que os homens apresentam, de uma maneira geral, maior probabilidade de doença coronariana já seja conhecido pelos especialistas, os experimentos nesta base de dados apontam uma probabilidade de 100% da sua ocorrência quando estes homens apresentam indicativo de diabetes, fato este que por si só já denota relevância.

Pelo fato da doença coronariana ser reconhecidamente uma patologia que acomete primariamente homens brancos, deixou de ser pesquisada e mesmo monitorada em mulheres. Entretanto, em 1999, 512.904 mulheres norte-americanas morreram devido à doença coronariana, mais que o câncer de mama e outros tipos de câncer. Atualmente, sabe-se que as mulheres tendem a apresentar mais complicações referentes à doença coronariana (AMERICAN HEART ASSOCIATION, 2001).

Estudos recentes, dentre eles o de Kosuge *et al.* (2006) mostram que as taxas de doença coronariana e a sua morbi-mortalidade em mulheres vêm aumentando, o que sugere que esta população também deve ser investigada. O fato do algoritmo não ter detectado casos do sexo feminino foi surpreendente à despeito do crescente número de mulheres com doença coronariana.

Gottschall e Leite (2002) apontam o sexo masculino como um fator de baixo risco para a angioplastia enquanto o sexo feminino representa um fator de alto risco, ainda mais quando associado à idade avançada (maior de 70 anos), que apresente doença de múltiplos vasos, com diabete melitos, história de insuficiência cardíaca, fração de ejeção ventricular menor de 40%, lesões equivalentes à doença de tronco

de coronária esquerda, inadequada terapêutica anti-plaquetária, presença de angina instável, angioplastia terapêutica trombolítica, estenose maior de 90% e lesões de alto risco. Estes mesmos autores comentam também sobre o fator preditivo de oclusão coronariana aguda durante angioplastia, assim como de maior mortalidade após angioplastias realizadas a usuários do sexo feminino, associada também a outros fatores, (GOTTSCHALL; LEITE, 2002).

Sendo assim, estudos futuros deverão incluir a investigação sobre as associações de eventos cardiovasculares com o sexo feminino.

A Tabela 3 apresenta os resultados provenientes da análise subjetiva e sob a perspectiva da confiança do evento significativo ‘Valores Referenciais Cateterismo Cardíaco’. Já o Quadro 27 apresenta as regras relevantes após as etapas objetiva e subjetiva de análise.

Tabela 3 – Resultados da análise subjetiva sob a perspectiva da confiança do evento significativo ‘Valores Referenciais Cateterismo Cardíaco’

	IRRELEV	INSIGNIF	RELEV	RED	TOTAL
ANÁLISE SUBJETIVA	1772	0	5		1777
ANÁLISE CONFIANÇA	1772		1	4	1177

Na análise subjetiva, das 1777 regras geradas relacionadas ao evento significativo ‘Valores Referenciais Cateterismo Cardíaco’, 1172 foram classificadas como irrelevantes, nenhuma como insignificante, e 5 como relevantes. Na análise sob a perspectiva da confiança, das 5 regras consideradas relevantes na análise subjetiva, apenas uma foi confirmada como relevante, e as demais regras primariamente relevantes (4) foram classificadas como redundantes, juntamente com as insignificantes.

Quadro 27 - Demonstrativo da regra relevante para o evento significativo Valores Referenciais Cateterismo Cardíaco após análise subjetiva e sob a perspectiva da confiança

EVENTO SIGNIFICATIVO	EVENTO	EVENTO	EVENTO	EVENTO	SUP.	CONF.
VR_CATETERISMO_CARD	← MICROALBUMINURIA	MAPEAMENTO_RETINA	HB_GLICOSILADA	GLICOSE	(10.5%,	50.0%)

A regra relevante para o evento significativo 'Valores Referenciais Cateterismo Cardíaco', após análise subjetiva e sob a perspectiva da confiança, deve ser lida da seguinte maneira: 'para esta amostra, 10,5% dos usuários apresentaram no seu histórico o procedimento de mensuração da glicose associado ao procedimento de mensuração da hemoglobina glicosilada, ao mapeamento da retina e à mensuração de microalbuminúria; destes, 50% têm probabilidade de apresentar associação com valores referenciais de cateterismo cardíaco'.

Embora esta regra apresente uma taxa de confiança baixa (50%), foi considerada relevante pelos especialistas pelo fato de associar diversos procedimentos relacionados ao diabetes (mensuração da glicose, da hemoglobina glicosilada, mapeamento da retina e a mensuração de microalbuminúria), que tendem a apresentar associação com um procedimento de alto custo como o cateterismo cardíaco.

Além disso, outro aspecto interessante desta regra é que procedimentos de menor complexidade mostraram-se associados a um procedimento de alta complexidade. Este fato vai ao encontro do objetivo desta pesquisa, pois aponta que ações pró-ativas de prevenção podem ser desenvolvidas a fim de evitar um evento mais complexo, tanto em termos financeiros, como de risco que o indivíduo está exposto, que afetam a sua qualidade de vida e incorrem em sacrifícios pessoais.

A associação apontada por essa regra reflete a necessidade de monitorização de indivíduos que apresentem este padrão em seu histórico de saúde, por meio de programas de gerenciamento de casos, por exemplo.

Há forte associação de diabete com a doença coronariana, inclusive a literatura refere maior chance de óbito para usuários portadores daquela doença (ESCOSTEGUY *et al.*, 2005; ROTHMAN; WAGNER, 2003; BRAUNWALD *et al.*, 2000). Em alguns estudos, a doença cardiovascular foi listada como a causa da morte em 65 a 75% dos portadores de diabete (BRAUNWALD *et al.*, 2000; KOSIBOROD *et al.*, 2005). A hiperglicemia predispõe à dislipidemia, à agregação plaquetária e à função alterada das hemáceas, o que pode levar a formação de trombos (SMELTZER; BARE, 2003).

Jensen *et al.* (2000) citam estudos que salientam que para populações de indivíduos diabéticos, assim como para a população em geral, a microalbuminúria, independente de outros fatores, precede à doença cardiovascular aterosclerótica.

Estes mesmos autores, num estudo com pacientes hipertensos, descobriram que a microalbuminúria estava presente em 28% dos pacientes que subseqüentemente desenvolveram doença cardíaca isquêmica, comparados com apenas 8% dos pacientes do grupo controle, concluindo que a microalbuminúria é o determinante mais forte de doenças cardíacas isquêmicas entre pacientes já hipertensos (JENSEN *et al.*, 2000).

Seria interessante incluir em estudos futuros uma ferramenta que pudesse, após identificar padrões de procedimentos que levem a um outro procedimento mais complexo, voltar aos casos da base para identificar os usuários que possuam esses padrões encontrados no seu histórico. Desta maneira, seria possível analisar outras variáveis que poderiam ser interessantes, além de possibilitar, na prática, o gerenciamento de casos de forma individualizada.

A Tabela 4 apresenta os resultados provenientes da análise subjetiva e sob a perspectiva da confiança do evento significativo 'Cintilografia do Miocárdio', e o Quadro 28 apresenta as regras relevantes após as etapas objetiva e subjetiva de análise.

Tabela 4 – Resultados da análise subjetiva e sob a perspectiva da confiança do evento significativo 'Cintilografia do Miocárdio'

	IRRELEV	INSIGNIF	RELEV	RED	TOTAL
ANÁLISE SUBJETIVA	56	3	3		62
ANÁLISE CONFIANÇA	56		1	5	62

Na análise subjetiva, das 62 regras geradas relacionadas ao evento significativo 'Cintilografia do Miocárdio', 56 foram classificadas como irrelevantes, 3 como insignificantes, e 3 como relevantes. Na análise sob a perspectiva da confiança, das 3 regras consideradas relevantes na análise subjetiva, apenas uma foi confirmada como relevante, as demais regras primariamente relevantes foram classificadas como redundantes (5), juntamente com as insignificantes.

Quadro 28 - Demonstrativo da regra relevante para o evento significativo ‘Cintilografia do Miocárdio’ após análise subjetiva e sob a perspectiva da confiança

EVENTO SIGNIFICATIVO	EVENTO	EVENTO	EVENTO	SUP.	CONF.
CINTILOGR_MIOC	← IMPLANTE_STENT	TESTE_ERGOM	CATETERISMO_CARD	(10.5%,	50.0%)

Após as análises subjetiva e sob a perspectiva da confiança do evento significativo ‘Cintilografia do Miocárdio’, a regra classificada como relevante para o evento significativo ‘Cintilografia do Miocárdio’ pode ser lida da seguinte maneira: ‘para esta amostra, 10,5% dos usuários apresentaram no seu histórico o procedimento de cateterismo cardíaco associado a teste ergométrico e implante de *stent*; destes, 50% têm probabilidade de apresentar associação com o procedimento de cintilografia do miocárdio perfusão em repouso e stress’.

Considerando que a cintilografia do miocárdio é o procedimento de escolha para monitorização dos resultados dos procedimentos de cateterismo cardíaco e de implante de *stent*, a regra foi considerada relevante, uma vez que este procedimento deveria ser mais freqüente do que foi encontrado nesta pesquisa.

É importante salientar que os princípios do gerenciamento de casos contemplam a gestão dos recursos, de maneira que os usuários do sistema de saúde obtenham a assistência de saúde adequada às suas necessidades, e evitar a não utilização ou utilização inadequada destes recursos.

Faz-se necessário utilizar o conhecimento advindo da medicina baseada em evidências a fim de garantir que condutas que visem o monitoramento adequado do paciente sejam aplicadas na prática, fato que contribui para a qualidade de vida e de assistência do paciente, e conseqüentemente à saúde financeira da operadora de saúde. Evita-se, desta maneira, que os usuários cheguem num estágio avançado da doença pelo fato de não ter acesso a serviços que seriam indispensáveis ao seu diagnóstico e tratamento (CARLSON, 2003)

A Tabela 5 apresenta os resultados provenientes da análise subjetiva e sob a perspectiva da confiança do evento significativo ‘Implante de *Stent*’, e o Quadro 27 apresenta as regras relevantes após as etapas objetiva e subjetiva de análise.

Tabela 5 – Resultados da análise subjetiva e sob a perspectiva da confiança para o evento significativo ‘Implante de *Stent*’

	IRRELEV	INSIGNIF	RELEV	RED	TOTAL
ANÁLISE SUBJETIVA	699	3	32		734
ANÁLISE CONFIANÇA	699		1	34	734

Na análise subjetiva, das 734 regras geradas relacionadas ao evento significativo ‘Implante de *Stent*’, 699 foram classificadas como irrelevantes, 3 como insignificantes, e 32 como relevantes. Na análise sob a perspectiva da confiança, das 32 regras consideradas relevantes na análise subjetiva, apenas uma foi confirmada como relevante, as demais regras primariamente relevantes foram classificadas como redundantes (34), juntamente com as insignificantes.

Quadro 29 - Demonstrativo da regra relevante para o evento significativo Implante de Stent após análise subjetiva e sob a perspectiva da confiança

EVENTO SIGNIFICATIVO	EVENTO	SUP.	CONF.
IMPLANTE_STENT	← ANGIOPL_VASO_UNICO	(18.4%,	92.9%)

Após a análise subjetiva e sob a perspectiva da confiança, a regra classificada como relevante para o evento significativo ‘Implante de *Stent*’ pode ser lida da seguinte maneira: ‘para esta amostra, 18,4% dos usuários apresentaram no seu histórico o procedimento de angioplastia; destes, 92,9% têm probabilidade de apresentar associação com o procedimento de implante de *stent*’.

O que chamou a atenção dos especialistas foi o alto índice de ocorrência de implantes de *stents* (92,9%), na base de dados estudada. Alguns autores, entre eles Braunwald *et al.* e (2001) e Smith *et al.* (2001), comentam que cerca de 50 a 80% de todos os procedimentos cardíacos invasivos estão associados ao implante de pelo menos um *stent*.

Os procedimentos de angioplastia podem ser realizados somente com um cateter balão, ou com a implantação de *stents*, que são tubos fenestrados de aço purificado que podem ser expandidos pelo balão e prover um aumento do calibre

das artérias coronárias de dentro para fora de forma sustentada (GOTTSCHALL; LEITE, 2002).

Nos dias atuais, com o aperfeiçoamento do material utilizado nesse procedimento, praticamente toda lesão uni-arterial pode ser acessível ao cateter balão, com exceção de lesões em vasos com calibre inferior a dois milímetros. Ainda não existem balões compatíveis com esse calibre, que está associado a um alto índice de oclusão aguda e re-estenose na dilatação de vasos finos periféricos ou ramos. O emprego de aterectomia rotacional em alguns casos, e de *stents* apropriados em outros, auxiliam a diminuir dramaticamente as complicações e a re-estenose (GOTTSCHALL; LEITE, 2002).

A literatura indica que o uso de *stents* coronarianos vem reduzindo de maneira significativa as complicações isquêmicas pós-angioplastia por balão, quando implantados após dissecação ou oclusão aguda, ou recolhimento elástico agudo. Entre as principais complicações pode-se citar óbito, infarto e a cirurgia de revascularização do miocárdio (GOTTSCHALL; LEITE, 2002).

Sabe-se que a implantação dos *stents* reduz o aparecimento de re-estenose, quando seus resultados são comparados com utilização do balão. Além disso, quando associados a drogas imunossupressoras, como a rapamicina, demonstram resultados impressionantes referentes ao desaparecimento definitivo da re-estenose. Otimizam resultados do cateter balão no caso de discretas dissecações ou de estenose residual menor que 20% após angioplastia, no caso de lesões de tronco de coronária esquerda sem possibilidade cirúrgica, bem como no caso de lesões ostiais ou de bifurcações, estenoses em ponte de safena com alto potencial de complicações e de re-estenose, como oclusões totais após recanalização, e no IAM (GOTTSCHALL; LEITE, 2002).

Considerando o exposto, acredita-se que o alto índice de implantes de *stent* associados a procedimentos de angioplastia retrata que os pacientes são diagnosticados numa fase avançada da doença coronariana, ou que os procedimentos são pouco auditados e, por isso, ocorrem numa grande freqüência. Entretanto, ao visualizar o Gráfico 3 (na página 76) percebe-se que os procedimentos de alta complexidade, categoria na qual se incluem os procedimentos de angioplastia e implante de *stent*, de maneira geral não são tão freqüentes,

corroborando com a idéia de que os usuários que compõem esta base de dados realmente apresentam condições graves de doença coronariana, fato que justifica a grande utilização deste procedimento.

A Tabela 6 apresenta os resultados provenientes da análise subjetiva e sob a perspectiva da confiança do evento significativo ‘Angioplastia Vaso Único’, e o Quadro 30 apresenta as regras relevantes após as etapas objetiva e subjetiva de análise.

Tabela 6 – Resultados da análise subjetiva e sob a perspectiva da confiança do evento significativo ‘Angioplastia Vaso Único’

	IRRELEV	INSIGNIF	RELEV	RED	TOTAL
ANÁLISE SUBJETIVA	309	28	54		391
ANÁLISE CONFIANÇA	309		1	81	392

Na análise subjetiva, das 392 regras geradas relacionadas ao evento significativo ‘Angioplastia Vaso Único’, 309 foram classificadas como irrelevantes, 28 como insignificantes, e 55 como relevantes. Na análise sob a perspectiva da confiança, das 55 regras consideradas relevantes na análise subjetiva, apenas 1 foi confirmada como relevante, sendo as demais regras relevantes classificadas como redundantes (80), juntamente com as insignificantes.

Quadro 30 - Demonstrativo da regra relevante para o evento significativo Angioplastia Vaso Único após análise subjetiva e sob a perspectiva da confiança

EVENTO SIGNIFICATIVO	EVENTO	SUP.	CONF.
ANGIOPL_VASO_UNICO	← IMPLANTE_STENT	(18.4%,	92.9%)

Após a análise subjetiva e sob a perspectiva da confiança, as regras classificadas como relevantes para o evento significativo ‘Angioplastia Vaso Único’, têm a sua leitura da seguinte maneira, para a primeira regra, ‘para esta amostra, 18,4% dos usuários apresentaram no seu histórico o procedimento implante de

stent; destes, 92,9% têm probabilidade de apresentar associação com o procedimento angioplastia de vaso único’.

De acordo com os especialistas, esta regra foi considerada como relevante pela sua taxa de confiança elevada (92,9%).

Nota-se que as taxas de suporte e de confiança desta regra são as mesmas da primeira regra do Quadro 27, o qual apresenta as regras relevantes para o evento significativo ‘Implante de *Stent*’. A seguir são apresentadas a regra do quadro 30 e a primeira regra do quadro 29.

IMPLANTE_STENT	←	ANGIOPL_VASO_UNICO	(18,4%	92,9%)
ANGIOPL_VASO_UNICO	←	IMPLANTE_STENT	(18,4%,	92,9%)

Uma vez que este é um achado surpreendente, a fim de se evitar quaisquer questionamentos com relação à fidedignidade dos resultados do algoritmo, os dados foram checados com as saídas originais, demonstrando que as regras fornecidas para os especialistas estavam de acordo com as saídas originais do algoritmo.

Assim, o fato das duas regras terem a mesma taxa de suporte e confiança indica que os procedimentos ocorrem sempre associados. Esta afirmação vem corroborar com a análise dos resultados, que vêm demonstrando o perfil elevado de morbidade relacionada a doenças cardiovasculares dos usuários desta base de dados.

Essas afirmações podem significar que os usuários chegam ao serviço que realiza este tipo de procedimento em condições avançadas de doença coronariana, e que a sua grande maioria precisa de uma angioplastia, após um *stent* já ter sido implantado. Esta interpretação pode ser feita tendo em vista que os eventos ocorrem em dois momentos diferenciados. Por outro lado, esta situação evidencia a relevância e necessidade de implantação de programas preventivos e de promoção à saúde pela organização pagadora objeto desta pesquisa.

A Tabela 7 apresenta os resultados provenientes da análise subjetiva e sob a perspectiva da confiança do evento significativo ‘Cateterismo Cardíaco’. Por sua vez, o Quadro 31 as regras relevantes após as etapas objetiva e subjetiva de análise.

Tabela 7 – Resultados da análise objetiva e subjetiva do evento significativo ‘Cateterismo Cardíaco’

	IRRELEV	INSIGNIF	RELEV	RED	TOTAL
ANÁLISE SUBJETIVA	152	40	34		226
ANÁLISE CONFIANÇA	152		2	72	226

Na análise subjetiva, das 392 regras geradas relacionadas ao evento significativo ‘Cateterismo Cardíaco’, 159 foram classificadas como irrelevantes, 40 como insignificantes, e 34 como relevantes. Na análise sob a perspectiva da confiança, das 34 regras consideradas relevantes na análise subjetiva, apenas 2 foram confirmadas como relevantes, sendo as demais regras relevantes classificadas como redundantes (72), juntamente com as insignificantes.

Quadro 31 - Demonstrativo das regras relevantes para o evento significativo Cateterismo Cardíaco após análise subjetiva e sob a perspectiva da confiança

EVENTO SIGNIFICATIVO		EVENTO	SUP.	CONF.
CATETERISMO_CARD	←	ANGIOPL_VASO_UNICO	(18.4%,	78.6%)
CATETERISMO_CARD	←	IMPLANTE_STENT	(18.4%,	78.6%)

A seguir, são apresentadas as leituras das regras relevantes após análise subjetiva e sob a perspectiva da confiança, para o evento significativo ‘Cateterismo Cardíaco’, considerando a amostra deste estudo.

Regra 1 → 18,4% dos usuários apresentaram no seu histórico o procedimento angioplastia de vaso único; destes, 78,6% têm probabilidade de apresentar associação com o procedimento cateterismo cardíaco.

Regra 2 → 18,4% dos usuários apresentaram no seu histórico o procedimento implante de *stent*; destes, 78,6% têm probabilidade de apresentar associação com o procedimento cateterismo cardíaco.

As duas regras, tanto a que relaciona o procedimento de angioplastia como a que relaciona o procedimento de implante de *stent* ao procedimento de cateterismo cardíaco, apresentam as mesmas taxas de suporte e confiança. Isso reflete que, nesta base de dados, estes procedimentos (angioplastia e implante de *stent*)

ocorrem na mesma freqüência, ou seja, ocorrem associados ao procedimento de cateterismo cardíaco.

Esta afirmação contribui com a análise de que os usuários cujos registros encontram-se nesta base apresentam condições de saúde complexas, sendo que os usuários geralmente são submetidos a procedimentos de angioplastia, implante de *stent*, e também a cateterismo cardíaco (em 78,6% dos casos), seja para avaliação do diâmetro do vaso lesado ou à adequada implantação do *stent*.

Um resumo dos resultados obtidos com a presente pesquisa, tendo em vista as etapas metodológicas desenvolvidas e os objetivos traçados pode ser visualizado na Figura 5

RESULTADOS

Usuários a serem indicados para programas de gerenciamento de casos cardiovasculares	Variáveis administrativas	Eventos Alertas
<p>Com mais de 40 anos, que apresentarem em seu histórico procedimentos indicativos de diabetes</p> <p>Microalbuminúria Hemoglobina Glicosilada Mapeamento de Retina</p> <p>e que forem do sexo masculino.</p>	<p>Idade, códigos de procedimentos dos usuários, custo, e complexidade dos eventos</p>	<p>Cateterismo Cardíaco Angioplastia de vaso único Implante de Stent Cintilografia do Miocárdio VR Cateterismo Cardíaco VR Revascularização Miocárdio Consultas de Emergência Hemoglobina Glicosilada Mapeamento de Retina Microalbuminúria</p>

Figura 5 - Resultados alcançados com a presente pesquisa.

Salienta-se que todos os dados estavam armazenados na base de dados, no entanto, para que estas informações pudessem ser estudadas, houve necessidade de um grande trabalho de pré-processamento, condição muitas vezes incapaz de ser realizada na prática de organizações de saúde. Em decorrência disto, não dispõem

de informação e conhecimento para gerenciar proativamente seus procedimentos assistenciais.

5 CONCLUSÕES

*“Estamos afogados em informação, mas
Morrendo de fome por conhecimento”
- John Naisbett*

O desenvolvimento desta pesquisa teve como questão norteadora a indagação acerca dos padrões que poderiam ser encontrados em uma base de dados de uma organização pagadora de serviços de saúde que possibilitariam a identificação de usuários com doenças cardiovasculares elegíveis para programas de gerenciamento de casos por meio de técnicas computacionais sofisticadas.

A metodologia de descoberta de conhecimentos em bases de dados, utilizando a técnica de descoberta de associações, tornou possível o alcance do objetivo geral proposto nesta pesquisa. A própria metodologia originou um modelo para identificar os padrões, em uma base de dados, em forma de regras.

Após análise realizada por especialistas de maneira subjetiva e sob a perspectiva da confiança, esses padrões indicaram que os usuários a serem incluídos em um programa de gerenciamento de casos são aqueles que apresentam procedimentos indicativos de diabetes e que são do sexo masculino. Assim, podem ser evitados procedimentos cardiológicos de cateterismo cardíaco, angioplastia com implante de *stent*, ou mesmo de revascularização do miocárdio.

As variáveis administrativas identificadas nos registros dos usuários na base de dados, no período de evolução analisado, foram: a idade; os códigos de procedimentos dos usuários; o custo; e, a complexidade dos procedimentos. Entretanto, variáveis de interesse clínico, social e de utilização do sistema de saúde, não incluídas na base de dados estudada, poderiam contribuir significativamente para uma análise mais ampla por parte dos especialistas.

A partir da identificação dos eventos significativos e, conseqüentemente, dos eventos que os geravam, associados ou isoladamente, foi possível identificar alguns

eventos alertas que possibilitariam a identificação de usuários com potencial para doença cardiovascular, elegíveis para programas de gerenciamento de casos.

Numa perspectiva de se poder visualizar, numa tela de computador com ajuda de uma linha do tempo em que se encontrassem todos os procedimentos solicitados e liberados para o usuário, os eventos alertas seriam aqueles que disparariam um alerta para o gerente de casos, avisando que este usuário, de acordo com o seu histórico de procedimentos, apresentaria determinada probabilidade de desenvolver um determinado evento significativo, de alto custo ou alta complexidade, e por isso deveria ser encaminhado para o programa de gerenciamento de casos específico para a sua situação de saúde-doença.

Nesta pesquisa, os eventos alertas identificados foram, primeiramente, os eventos classificados pelos especialistas como significativos. Eventos como cateterismo cardíaco; angioplastia de vaso único; implante de *stent*; cintilografia do miocárdio; valores referenciais de cateterismo cardíaco; e, valores referenciais de revascularização do miocárdio, são associados à assistência de alta complexidade e de alto custo. Desta maneira, caso ocorram no histórico de um determinado usuário com CID do grupo das doenças cardiovasculares, devem ser considerados como alerta.

Já as consultas de emergência, igualmente consideradas como eventos alertas, devem ser monitoradas, pois podem estar relacionadas à condição grave de saúde do usuário, ou a acompanhamento médico ambulatorial insuficiente.

Além dos eventos significativos, foram classificados como eventos alertas os procedimentos de mensuração da hemoglobina glicosilada, mapeamento de retina e microalbuminúria, quando encontrados no histórico de pacientes com CID do grupo de doenças cardiovasculares.

Assim, usuários maiores de 40 anos, do sexo masculino, e que apresentem em seu histórico os procedimentos acima citados devem ser investigados e deve ser analisada a possibilidade de incluí-los num programa de gerenciamento de casos.

6 CONSIDERAÇÕES FINAIS

*“O impossível de hoje será o possível de amanhã
se fizermos o possível de hoje”*

- Paulo Freire

Com o desenvolvimento desta pesquisa, procurou-se aliar o conhecimento das ciências exatas, a Computação e a Informática, com o conhecimento das ciências humanas, a Administração, ambas aplicadas às ciências da saúde.

Independentemente dos resultados que poderiam advir desta tentativa, ao final deste processo evidencia-se a grande contribuição da associação de uma ou mais ciências, em prol de uma outra, ainda mais quando se trata de um assunto com aplicação prática, e de conseqüências importantes como o acesso dos indivíduos a serviços de saúde de qualidade.

Esta contribuição pode ser compreendida primariamente em termos da troca de informações e conhecimentos entre profissionais de diversas formações acadêmicas. Somando-se a essa contribuição, encontra-se a ampliação da visão de mundo, das diferentes formas de raciocínio, e do entendimento das contribuições de diferentes formas de pensamento na resolução de um problema, seja ele qual for, desde que as pessoas tenham interesse e disponibilidade para ouvir e aprender.

A interdisciplinariedade é, cada vez mais, uma realidade dentro de organizações de saúde. O recente entendimento da necessidade da informatização dos seus processos de trabalho promove o encontro entre profissionais de outras áreas do conhecimento com os da área da saúde, compartilhando conhecimentos e contribuindo para a otimização do processo de tomada de decisão referente a situações antes vivenciadas apenas pelos profissionais da área da saúde.

É a partir dessa filosofia de compartilhar conhecimentos para que problemas sejam resolvidos sob a ótica de diversos profissionais que devem ser desenvolvidos os sistemas de informação de prestadores e operadoras de saúde.

Embora iniciativas como o curso de Pós-Graduação em Tecnologia em Saúde da PUCPR contribuam para o intercâmbio de profissionais da saúde com ferramentas computacionais, e de profissionais da área tecnológica com problemas da área da saúde, é fundamental a compreensão de que ainda é preciso que todos os profissionais façam parte de uma equipe, que vise resolver problemas de tecnologia da informação na área da saúde. Só assim pode ser mantida a expertise de cada disciplina a fim de que possa haver intersecção das áreas, sem detrimento do conhecimento específico inerente a cada uma delas.

Os sistemas de informação na área da saúde devem efetivamente ser desenvolvidos em acordo com os objetivos organizacionais, as regras do negócio, os processos de trabalho dos diversos profissionais que atuam na organização, além de possibilitar a recuperação dos dados que podem levar à informação estratégica e conhecimento.

Para que o resultado das práticas desenvolvidas na organização de saúde seja o conhecimento acerca dos processos pelos quais se alcançam serviços de qualidade e o bem estar do usuário, é fundamental que os profissionais envolvidos no registro dos dados no sistema de informações estejam cientes do seu papel essencial neste processo.

Sendo assim, a educação em serviço, visando à valorização do registro correto nos sistemas de informação, deve ser implementada em todas as organizações de saúde. Além disso, o apoio gerencial é fundamental para que ações de educação em serviço sejam desenvolvidas, bem como para que as bases de dados sejam adequadamente atualizadas. O registro correto dos dados e a atualização das bases de dados influenciam sobremaneira o resultado de pesquisas em bases de dados e a construção do conhecimento organizacional.

A construção de bases de dados em saúde deve considerar com veemência questões éticas e legais envolvidas na manipulação de dados de saúde de populações, de modo que os usuários não sejam penalizados pelas suas condições de saúde-doença. Estratégias devem ser desenvolvidas no sentido de garantir a confidencialidade, controle do acesso, da integridade e proteção física dos dados e informações. Os dados dos usuários devem ser coletados quando do seu ingresso no sistema de saúde, incluindo dados de identificação, administrativos e clínicos,

considerando a tênue linha que os diferencia com relação à importância para estudos de custo-benefício, entre tantos outros possíveis de serem realizados em uma organização de saúde.

É de suma importância, tanto para as atividades práticas que envolvem a manipulação de dados (seu registro, armazenamento e recuperação), como para as atividades de pesquisa em bases de dados, que os registros sejam confiáveis e fidedignos à realidade. Desta maneira, esforços devem ser envidados no sentido de manter as bases de dados atualizadas.

As etapas metodológicas do processo de descoberta de conhecimento em bases de dados mostram-se exequíveis em bases de dados da saúde, porém é essencial valorizar a etapa de pré-processamento, a fim de deixar a base de dados o mais preparada possível para os experimentos. Similarmente, a fase de pós-processamento é fundamental na transformação das informações extraídas a partir da mineração dos dados, por meio de análises subjetivas e objetivas, de modo a decodificar o conhecimento para quem o utiliza para tomar decisões.

O auxílio de especialistas que trabalhem conjunta e efetivamente na análise, e a compreensão e discussão das regras encontradas pelos algoritmos de mineração é significativo para o êxito desses processos, pois só estes poderão comparar e avaliar a real contribuição de técnicas mais elaboradas de descoberta de padrões em dados à realidade.

A metodologia proporcionou o alcance dos objetivos propostos, e as regras consideradas como relevantes pelos especialistas, de uma maneira geral, apontam para que estratégias preventivas sejam desenvolvidas para a população desta base de dados. Embora tenha sido evidenciado que os usuários já apresentavam doenças coronarianas em fases avançadas, tais estratégias poderiam ser desenvolvidas a partir da estratificação dos riscos da população para estas doenças, e programas estabelecidos de acordo com as necessidades de cada subgrupo.

Com relação aos resultados da pesquisa com esta base de dados, demonstram que estratégias de auditoria devem estar voltadas não só ao controle dos procedimentos de alto custo e alta complexidade, mas também à monitorização dos procedimentos de baixa e média complexidade, pois estão fortemente associados a altos custos devido à alta utilização.

Programas de gerenciamento de casos podem ser ajustados aos objetivos organizacionais, de modo a prestar assistência precoce e preventiva a indivíduos com algum grau de risco para desenvolver as patologias mais frequentes nesta população. Entretanto, por se tratar de uma ferramenta de gestão que subentende o investimento de recursos para atividades preventivas, no nosso país este tipo de estratégia ainda é incipiente, principalmente por não existirem estudos confiáveis que demonstrem o retorno do investimento neste tipo de programa.

O gerenciamento de casos parece ser uma ferramenta de muita utilidade no auxílio à gestão da saúde de populações, porém deve encontrar base no modelo assistencial e também contar com a contribuição dos profissionais envolvidos. Por conseguinte, para que tais ferramentas de gestão obtenham êxito que beneficie, tanto aos usuários como às operadoras de saúde, é necessário que o modelo de assistência à saúde seja revisto, pois seus princípios e fundamentos permeiam todas as ações em saúde.

A presente pesquisa apresentou como limitações: (a) a base de dados original, a qual apresenta dados incompletos, omissos, digitados de maneira errada, que além de demandaram uma extensa fase de pré-processamento, limitaram o tamanho da base de dados da pesquisa; (b) as dificuldades técnicas encontradas, relativas ao tamanho da base de dados e aos equipamentos disponíveis, que incorreram em retrabalho, desperdício do tempo, limitando os experimentos que poderiam ser realizados.

Acredita-se que essas limitações incentivam a busca pela sua superação, e nesse sentido, novos estudos sejam realizados a fim de tentar resolver com outros algoritmos os problemas identificados e não trabalhados neste estudo. Outros problemas poderão ser detectados a partir do enriquecimento da base de dados com dados clínicos, de moradia, educacionais, entre outros.

Dando continuidade a esta linha de pesquisa, com relação ao pós-processamento, as regras deverão ser analisadas de maneira objetiva, utilizando-se da estatística, porém ainda com auxílio dos especialistas para validar os experimentos.

Adicionalmente, pretende-se aplicar o conhecimento descoberto nesta pesquisa na base de dados inicial. Desta maneira, poder-se-ia identificar os

pacientes com os eventos significativos no seu histórico, indicando à operadora de saúde os pacientes que deveriam ser incluídos num programa de gerenciamento de casos, voltado à problemas cardiovasculares. Ademais, a mesma metodologia poderia ser utilizada para outros grupos de patologias.

De maneira geral, esta pesquisa pode ser visualizada como uma contribuição para um relacionamento mais estreito entre os profissionais da área da saúde com os da área tecnológica, e de ambos com organizações de saúde. Procurou-se evidenciar intersecções entre estas áreas, o esforço para que se observasse o rigor metodológico, e a responsabilidade com os resultados encontrados, tendo em vista as limitações inerentes à base de dados.

Além disso, em decorrência da pesquisa ter trabalhado com uma base de dados real, que exigiu uma limpeza rigorosa destes dados, é que se insiste na afirmação que projetos de desenvolvimento de sistemas de informação em saúde devam, necessariamente, unir profissionais das áreas da saúde aos da tecnologia da informação. Desta maneira, poderão ser minimizadas as limitações das bases de dados, fato que influencia diretamente os resultados de estudos como este. Caso contrário, projetos como este, que demandam um trabalho árduo, detalhista, esforço computacional, intelectual e apoio de especialistas com disponibilidade e interesse no projeto, continuarão a acontecer apenas em ambiente acadêmico, sem transferir para a comunidade em geral os benefícios que deles podem advir.

A Figura 6 apresenta esquematicamente as considerações finais da presente pesquisa. A pirâmide apresenta propositalmente na sua base o trabalho multidisciplinar na concepção das bases de dados na área da saúde. Considera-se que este tipo de trabalho em conjunto é a base para o alcance do objetivo final, que se encontra no topo da pirâmide: a mudança do modelo assistencial no sistema de saúde suplementar brasileiro, de essencialmente curativista para efetivamente baseado em prevenção, nos seus níveis primário, secundário e terciário.



Figura 6. Representação esquemática das considerações finais.

Entretanto, somente a concepção das bases de dados de maneira multidisciplinar não garante esta transformação. Para tanto, estratégias de educação em serviço são fundamentais para a conscientização de todos os profissionais que trabalham com registros de dados de saúde acerca da importância destes dados serem fidedignos, além de para a conscientização da necessidade de mudança do modelo assistencial.

Acredita-se que educando e envolvendo um maior número de profissionais que trabalham em organizações de saúde, tanto no gerenciamento como na assistência direta aos clientes, é que metodologias como a do processo de descoberta de conhecimento em bases de dados poderá ser desenvolvida não só em ambiente acadêmico, mas também nas organizações de saúde, possibilitando que o conhecimento gerado por este processo possa ser efetivamente utilizado para embasar as práticas assistenciais, as decisões gerenciais e estratégicas nestas organizações.

REFERÊNCIAS

AGRAWAL, R. *et al.* Mining association rules between sets of items in large databases. In: THE 1993 AC SIGMOD CONFERENCE, 1993. Washington, DC. **Proceedings of the 1993 ac sigmod conference.** Washington, DC., 1993.

ALMEIDA, C. **O mercado privado de serviços de saúde no Brasil: panorama atual e tendências da assistência médica suplementar.** Textos para Discussão: IPEA. Rio de Janeiro, 1998.

AMARAL, M. B. Tecnologia da informação e gestão em saúde. **O mundo da Saúde.** São Paulo, ano 25, v. 26, n. 2, abr.-jun. 2002.

American Heart Association. **Heart and stroke statistical update.** Dallas, TX: American Heart Association: 2001.

AMERICAN HOSPITAL ASSOCIATION. Case Management, an aid to quality and continuity of care. In: AMERICAN HOSPITAL ASSOCIATION. **Case Management & Social Work Practice.** New York: Longman Publishing Group, 1992.

ANGELONI, M. T. Elementos intervenientes na tomada de decisão. **Ci. Inf. Brasília.** V. 32, n.1, p. 17-22, jan-abr, 2003.

ARARIBOIA, G. **Inteligência artificial.** Rio de Janeiro: LTC,1988.

BLAIR, R. Tech and tenderness. **Health management Technology.** V. 25, n. 5. May, 2004.

BORGELT, C. Apriori: find association rules/ hyperedges with apriori algoitiithm. Copyright 1996 – 2001.

BRANCO, M.A.F. Informação em saúde como elemento estratégico para gestão. IN: BRASIL. Ministério da Saúde. **Gestão municipal da saúde, textos básicos.** Rio de Janeiro: Brasil, Ministério da Saúde, p. 163-169, 2001.

BRAUNWALD, E., *et al.* ACC/AHA guidelines for the management of patients with unstable angina and non-ST-segment elevation MI: executive summary and recommendations. **Circulation.** v. 102, n.10, p. 1193-1290, 2000.

BRESNAHAN, J. Data Mining: a delicate operation. **CIO Magazine.** June 15, 1997. Disponível em <www.cio.com/archive/061597/mining_content.html?> Acesso em 15 jan. 2003.

BRIEMAN, L; FRIEDMAN, J.H.; OLHSEN, R.A.; STONE, C.J. **Classification and regression trees.** Monterey, Ca: Wadsworth and Brooks, 1984.

CARLSON, B. Predictive modeling, sharp lens on near future. **Managed Care**. Julho, 2003. Disponível em:
<www.managedcaremag.com/archives/0307/0307.predictive.html> Acesso em 10 ago 2004.

CARVALHO, A. O. Tecnologias da informação na gestão da saúde. **Revista de Administração Pública**. v. 32, n. 2, mar – abr, 1998.

CARVALHO, D.R. **Mineração de Dados através de indução de regras e algoritmos genéticos**. Dissertação de Mestrado.. Programa de Pós-Graduação em Informática Aplicada. Pontifícia Universidade Católica do Paraná. Curitiba, 1999.

_____. **Árvore de decisão/ algoritmo genético para tratar o problema de pequenos disjuntos em classificação de dados**. Tese de Doutorado. Programas de Pós-Graduação em Computação de Alto Desempenho/Sistemas Computacionais do Programa de Engenharia Civil da Universidade Federal do Rio de Janeiro. Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2005.

CIOS, K. J.; KURGAN, L. A. **Trends in Data Mining and Knowledge Discovery**. Disponível em <<http://citeseer.ist.tsu.edu/cios05trends.html>> Acesso em 01 dez 2005.

COLLAZOS, K. S.; ROISENBERG, M.; BARRETO, J. M. Dificuldades na Aplicação de KDD em Medicina. II Workshop de Informática aplicada à Saúde – CBComp (Anais) 2002.

COUSINS, M. S.; SHICKLE, L. M.; BANDER, J. A. Introduction to predictive modeling. **Disease Management**. v. 5, n.3, 2002. Disponível em:
<www.choosehmc.com/press/11-00-02-V5_N3_2002_Introduction_to_predictive_modeling.pdf> Acesso em 10 ago 2004.

COUTINHO, F.V. **Data mining**. Disponível em:
<<http://www.dwbrasil.com.br/html/dmining.html>> Acesso em 21 jun 2006.

CROOKS, P. Managing high-risk, high cost patients: the Southern California Kaiser permanente experience in the Medicare ESRD Demonstration Project. **The Permanente Journal**. v.9, n.2, 2005.

DATASUS. **Informações de saúde**. Disponível em <www.datasus.gov.br> Acesso em 01 fev 2006.

ESCOSTEGUY, C. C.; PORTELA, M. C.; MEDRONHO, R. A.; VASCONCELLOS, M. T. L. AIH versus prontuário médico no estudo de risco de óbito hospitalar no infarto agudo do miocárdio no município do Rio de Janeiro, Brasil. **Cad. Saúde Pública**. n. 21, v.4, p. 1065-1076, jul-ago, 2005.

ÉVORA, Y. D. M. **Processo de informatização na enfermagem**. São Paulo: EPU, 1995.

FAYYAD, U. PIATESKY-SHAPIO, G., SMYTH, P. **Advances in knowledge discovery and data mining**. Boston: MIT Press, 1996.

FITZSIMMONS, J. A. **Administração de serviços**: operações, estratégia e tecnologia de informação. 2. ed. Porto Alegre: Bookman, 2000.

FORMAN, S. Targeting the Highest-Risk Population to Complement Disease Management. **Health Management Technology**. v. 25, n. 7, jul, 2005.

GILLESPIE, G. Data Mining: solving car, cost capers. **Health Data Management**. Nov, 2004. Disponível em <www.healthdatamanagement.com> Acesso em 01 nov 2005.

GOLDSCHMIDT, R.; PASSOS, E. **Mineração de Dados**: um guia prático. Rio de Janeiro: Elsevier, 2005.

GOLDSMITH, L. S.; HUTCHISON, B.; HURLEY, J. **Economic evaluation across the four faces of prevention: a canadian perspective**. McMaster University, Centre of Health Economics and Policy Analyss, 2004.

GONZALEZ, R. I. C.; CASARIN, S. N. A.; CALIRI, M. H. L.; SASSAKI, C. M.; MONROE, A. A.; VILLA, T. C. S. Gerenciamento de caso: um novo enfoque de cuidado à saúde. **Rev. Latino-Am. Enfermagem**. v. 11, n. 2. Ribeirão Preto mar./abr. 2003.

GOTTSCHALL, C. A. M.; LEITE, R. S. Angioplastia coronariana transluminal percutânea: indicações e complicações. IN: INSTITUTO DE CARDIOLOGIA DO RIO GRANDE DO SUL. **Manual de Condutas**. Porto Alegre: Fundação Universitária de Cardiologia, 2002. p.199-215.

FERREIRA, A. B. H. **Dicionário Aurélio Básico da Língua Portuguesa**. São Paulo: Nova Fronteira, 1988.

ISHAM, G.J.; KRAMER, K.K. Identifying and managing high risk members. **Group Practice Journal**. v. 52, n. 5, May, 2003.

JEKEL, J. F.; KATZ, D. L.; ELMORE, J. G. **Epidemiologia, Bioestatística e Medicina Preventiva**. 2. ed. Porto Alegre: Artmed, 2005.

KNABEL, T; LOUWERS, J. Intervenability: another measure of health risk. **Health Management Technology**. v.25, n.7, julho, 2004.

KOSUGE, M. *et al.* Sex differences in early mortality of patients undregoin primary *stentíng* for acute myocardial infarctation. **Circulation Journal**. v.70, p. 217-221, mar, 2006.

LACOMBE, R. J. M; HEILBORN, G. L. J. **Administração**: princípios e tendências. São Paulo: Saraiva, 2003.

LEE, I. N.; LIAO, S. C.; EMBRECHTS, M. Data mining techniques applied to medical information. **Med. Inform.** v.25, n.2, 81-102, 2000.

LEGROW, G. S. Predictive modeling: no crystal ball but an important tool.

Healthplan. jan/fev, 2004. Disponível em:

<<http://www.ahip.org/content/default.aspx?bc=31|130|136|259|262>> Acesso em 21 set 2004.

LORGA, A. *et al.* I Diretrizes do Grupo de Estudos em Cardiogeriatría da Sociedade Brasileira de Cardiologia. **Arq. Bras. Cardiol.**, 2002, vol.79 suppl.1, p.1-46. ISSN 0066-782X.

MCCAIN, J. Predictive modeling holds promise of earlier identification, treatment.

Managed Care Magazine. Disponível em

<www.managedcaremag.com/archives/0109/0109.predictive.html> Publicado em set. 2001. Acesso em 04 jan 2005.

MEGGINSON, L. C.; MOSLEY, D. C; PIETRI, P. H. J. **Administração:** conceitos e aplicações. 4 ed. São Paulo: Harba, 1998.

MENDES, E.V. A reengenharia do sistema de serviços de saúde no nível local: a gestão da atenção à saúde. In: MENDES, E.V. **A organização da saúde no nível local.** São Paulo: HUCITEC; 1998. p. 57-86.

METZGER, J. **Using computerized registries in chronic disease care.** California HealthCare Foundation, 2004.

MIRANDA, C.R.M. **Gerenciamento de custos em planos de assistência à saúde.**

Disponível em

<[http://www.ans.gov.br/portal/upload/biblioteca/TT_AS_20_ClaudioMiranda_Gerenciamento de Custo.pdf](http://www.ans.gov.br/portal/upload/biblioteca/TT_AS_20_ClaudioMiranda_Gerenciamento%20de%20Custo.pdf)> Publicado em nov. 2003. Acesso em 12 mar 2004.

MONTANA, P.J. **Administração.** São Paulo: Saraiva, 1999.

NCCC - NATIONAL CHRONIC CARE CONSORTIUM. **Case management:** methods and issues. Bloomington, MN, USA: NCCC, 2000. Disponível em <www.ncccconline.org> Acesso em 01 nov 2005.

OLIVEIRA, I. C. **Aplicação de Data Mining na Busca de um Modelo de Prevenção da Mortalidade Infantil.** Dissertação de Mestrado. Programa de Pós Graduação em Engenharia de Produção e Sistemas. Universidade Federal de Santa Catarina. Florianópolis, 2001.

ORALLO, J.H.; LIZANDRA, M.C.J.; COLLADO, N. M.; ARANDA, C.M. **Extracción y Visualización de Conocimiento de Bases de Datos Médicas.** Disponível em

<www.dsic.upv.es/~jorallo/escrits/Kdd-ACTA.pdf> Acesso em 26 set 2005.

OZBOLT, J.G. Personalized health care ad business success: can informatics bring us to the promised land? **Journal of the American Medical Informatics Association.** v.6, n.15, Sep./Oct., 368-373.1999.

PAHO - PAN AMERICAN HEALTH ORGANIZATION. **Setting up healthcare services information systems**: a guide for requirement analysis, application, specification, and procurement. Washington, DC: PAHO, 1999.

POLIT, Denise F., HUNGLER, Bernadette P. **Fundamentos de pesquisa em enfermagem**. 3. ed. Porto Alegre: Artes Médicas, 1995. 391 p.

ROGOSKI, R. R. Managed Care's Challenges. **Health Management Technology**. v. 24, n. 3. março, 2003.

QUEIROZ, A. C. S.; BARBOSA, A. P. Racionalidade e incorporação de tecnologia em saúde: a experiência de um hospital de alta complexidade em São Paulo. **RAE-eletrônica**. V. 2, n. 1, jan-jun, 2003.

QUINTO NETO, A. **Processo de acreditação**: a busca da qualidade nas organizações de saúde. Porto Alegre: Dacasa, 2000.

REZENDE, S. O. **Sistemas Inteligentes**: fundamentos e aplicações. Barueri, SP: Manole, 2005.

_____; PUGLIESI, J.B., MELANDA, E.A.; DE PAULA, M.F. Mineração de Dados. IN: REZENDE, S. O. **Sistemas Inteligentes**: fundamentos e aplicações. Cap. 12, 308-335. Barueri, SP: Manole, 2005.

ROTHMAN, A. A.; WAGNER, E. H. Chronic illness management: what is the role of primary care? **Ann Intern Med**. v.138. n. 3, feb, 2003.

SIAU, K. Health care informatics. **IEEE Transactions on Information Technology in Biomedicine**. v. 7, n.1, mar., 2003.

SILVA, J. B. **O Projeto de Pesquisa**: passo a passo. Disponível em: <www.meac.ufc.br/public/projpesq/pesquisa.htm#cap3> Fortaleza, CE, 1999. Acesso em 02 mai 2006.

SILVA, E. L.; MENEZES, E. M **Metodologia da pesquisa e elaboração de dissertação**. Florianópolis: Laboratório de Ensino à Distância da UFSC, 2001.

SIGULEM, D. *et al.* **Sistemas de apoio à decisão em medicina**. Disponível em: www.virtual.epm.br/material/tis/curr-med/sad_tml/sistema.htm Acesso em 12 dez. 2002.

SMELTZER, S.; BARE, B. **Brunner & Suddarth's. textbook of medical surgical nursing**. 10th edition. Lippincot Williams, 2003.

SMITH, SC; *et al.* ACC/AHA Guidelines for percutaneous coronary intervention (revision of the 1993 PTCA guidelines): executive summary. **Journal of the American College of Cardiology**. v. 37, n.8, p. 2215-2238, 2001.

SOUZA, O.R.M. **Mineração de Dados de um plano de saúde para obter regras de associação**. Dissertação de Mestrado. Florianópolis, 2000. 115p. Programa de

Pós Graduação em Engenharia da Produção - Universidade Federal de Santa Catarina.

STATUSONE. **Proactive care**: cost versus care. Disponível em <www.statusone.com> Acesso em 04 nov 2004.

STONER, J. A. F.; FREEMAN, R. E. **Administração**. 5. ed. Rio de Janeiro: LTC, 1999.

UNIDAS. **Autogestão no Brasil**: 1890-2005 história da organização e consolidação do setor. São Paulo: Unidas, 2005.

VENDRAMINI, L. F. **Precificação de planos de assistência à saúde**: uma visão integrada. Disponível em <rnti.fesppr.br/include/getdoc.php?id=100&article=27&mode=pdf> Acesso em 18 jun 2006.

VERMELHO, L. L.; COSTA, A. J. L; KALE, P. L. Indicadores de Saúde. IN: MEDRONHO, R. A. **Epidemiologia**. São Paulo: Atheneu, p. 33-55, 2004.

ANEXOS

ANEXO 1

Tabela 1A. Dados essenciais para sistemas de informação em saúde.

1	Identificador Pessoal Único
2	Data de Nascimento
3	Sexo
4	Raça e Etnia
5	Residência
6	Estado Civil
7	Moradia – (com quem mora, tipo de residência: própria, asilos, outras instituições)
8	Estado de Saúde Referido
9	Estado Funcional – grau de dependência nas atividades da vida diária
10	Escolaridade
11	Relação de parentesco com o titular
12	Ocupação Atual ou a mais recente
13	Tipo de Atendimento (hospitalar: emergência ou internamento; ambulatorial)
14	Data da Admissão (internamento)
15	Data da Alta (internamento)
16	Data do Atendimento (ambulatorial ou consulta)
17	Identificação do Prestador
18	Tipo do Prestador ou do Atendimento
19	Identificação do Profissional de Saúde
20	Endereço do Prestador ou Local do Atendimento
21	Identificação do Médico Assistente (internamento)
22	Identificação do Médico que prestou o atendimento principal
23	Especialidade do Profissional de Saúde
24	Diagnóstico Principal (internamento)
25	Diagnóstico Primário (responsável pela maior parte dos procedimentos do internamento)
26	Outros Diagnósticos (internamento)
27	Qualificador para Outros Diagnósticos
28	Razão pela qual o usuário procurou o serviço de saúde
29	Diagnóstico responsável pelos serviços prestados (ambulatorial)
30	Outros Diagnósticos (ambulatorial)
31	Causa de Lesão Externa
32	Peso do Recém Nascido
33	Procedimento Principal (internamento)
34	Outros procedimentos (internamento)
35	Datas dos procedimentos (internamento)
36	Procedimentos e Serviços (ambulatorial)
37	Medicações Prescritas
38	Tipo de Alta (internamento)
39	Tipo de Alta (ambulatorial)
40	Expectativa de Pagamento pelo Usuário
41	Lesão relacionada ao Trabalho
42	Valor Total da Conta Médica

Fonte: (PAHO, 1999)

ANEXO 2

Autorização do Comitê de Ética da PUCPR e do Instituto Curitiba de Saúde (ICS) para utilização da base de dados para pesquisa.



Pontifícia Universidade Católica do Paraná
Pró-Reitoria de Pesquisa e Pós-Graduação

Curitiba, 08 de fevereiro de 2006
Of. 006/06/CEP-PUCPR

Ref. "Ferramenta de Modelagem Preditiva para identificação de pacientes elegíveis para programas de gerenciamento de casos".

Prezado (a) Pesquisador (es),

Venho por meio deste informar a Vossa Senhoria que o Comitê de Ética em Pesquisa da PUCPR, no dia 08 de fevereiro do corrente ano aprovou o Projeto Intitulado "Ferramenta de Modelagem Preditiva para identificação de pacientes elegíveis para programas de gerenciamento de casos", pertencente ao Grupo III, sob o registro no CEP n° 924, e será encaminhado a CONEP para o devido cadastro. Lembro ao senhor (a) pesquisador (a) que é obrigatório encaminhar relatório anual parcial e relatório final a este CEP.

Atenciosamente,

Profª M. Sc Ana Cristina Miguez Ribeiro
Coordenadora do Comitê de Ética em Pesquisa - PUCPR

Ilma Sra.
Luciana Scheleder gonçalves Kubus

ANEXO 3

Quadro 3A - Códigos de procedimentos cardiológicos e suas descrições
(continua)

CÓDIGO	DESCRIÇÃO
40010015	COMISSUROTOMIA VALVAR
40010031	TROCA VALVAR UNICA
40010058	PLASTIA VALVAR
40010066	AMPLIACAO DO ANEL VALVAR
40020010	ANEURISMECTOMIA DE VENTRICULO ESQUERDO
40020045	REVASCULARIZACAO DO MIOCARDIO
40030016	CANAL ARTERIAL PERSISTENTE-CORRECAO CIRURGICA
40030024	COARCTACAO DA AORTA-CORRECAO CIRURGICA
40030032	CONFECACAO DE BANDAGEM DA ARTERIA PULMONAR
40030040	CORRECAO CIRURGICA DA COMUNICACAO INTERATRIAL
40030059	CORRECAO CIRURGICA DE COMUNICAO INTERVENTRICULAR
40030067	AMPLIACAO (ANEL VALVAR, GRANDES VASOS, ATRIO, VENTRICULO)
40030075	RESSECCAO (INFUNDIBULO, SEPTO, MEMBRANAS, BANDAS)
40030083	TRANSPOSICOES (VASOS, CAMARAS)
40030091	REDIRECIONAMENTO DO FLUXO SANGUINEO (C/ANASTOMOSE DIRETA, C/
40040011	ANEURISMA AORTA-TORACICA - CORRECAO CIRURGICA
40040020	ANEURISMA TORACO-ABDOMINAL-CORRECAO CIRURGICA
40040038	CARDIOTOMIA(FERIMENTO,CORPO ESTRANHO,EXPLORACAO)
40040046	COLOCACAO DE BALAO INTRAORTICO
40040054	DRENAGEM DO PERICARDIO
40040062	PERICARDIOCENTESE
40040070	PERICARDIECTOMIA
40040089	TRANSPLANTE CARDIACO
40040097	BIOPSIA DO MIOCARDIO
40040100	INSTALACAO DO CIRCUITO DE CIRCULACAO EXTRA CORPOREA CONVENC
40040119	INSTALACAO DE CIRCUITO DE CIRCULACAO EXTRA-CORPOREA EM CRIAN
40040127	HIPTERMIA PROFUNDA COM OU SEM PARADA CIRCULATORIA TOTAL
40040135	INSTALACAO DE MARCA PASSO EPIMIOCARDIO TEMPORARIO
40040143	INSTALACAO DE CATETERES INTRACAVITARIOS PARA MONITORIZACAO H
40040151	INSTALACAO DE CIRCUITO PARA ASSISTENCIA MECANICA CIRCULATORI
40040160	MANUTENCAO DE CIRCUITO PARA ASSISTENCIA MECANICA CIRCULATORI
40040178	CORRECAO CIRURGICA DAS ARRITMIAS
40040186	EMBOLECTOMIA PULMONAR
40040194	RETIRADA DE ORGAOS PARA TRANSPLANTE
40050025	RECOLOCACAO DO ELETRODO(MARCA PASSO)
40050033	TROCA DE GERADOR(MARCA PASSO)
40050041	IMPLANTE DE DESFIBRILADOR INTERNO, PLACAS E ELETRODOS
40050050	IMPLANTE DE ELETRODO ATRIAL
40050068	IMPLANTE DE ELETRODO VENTRICULAR
40050076	IMPLANTE DE ELETRODO ATRIAL E VENTRICULAR
40050084	IMPLANTE DE GERADOR
40060012	PERFUSIONISTA
40080013	CATETERISMO DE CAMARAS CARDIACAS DIREITAS
40080021	CATETERISMO DE CAMARAS CARDIACAS ESQUERDAS
40080030	CATETERISMO DE CAMARAS CARDIACAS DIREITAS E ESQUERDAS

Quadro 3A - Códigos de procedimentos cardiológicos e suas descrições
(continuação)

CÓDIGO	DESCRIÇÃO
40080048	CATETERISMO DE CAMARAS CARDIACAS DIREITAS E ESQUERDAS POR TE
40080056	CATETERISMO DE CAMARAS CARDIACAS DIREITAS E ESQUERDAS COM TE
40080064	AVALIACAO HEMODINAMICA DA REATIVIDADE VASCULAR PULMONAR
40080072	CATETERISMO DE CAMARA CARDIACA DIREITA COM CINEANGIOGRAFIA
40080080	CATETERISMO DE CAMARA CARDIACA ESQUERDA COM CINEANGIOGRAFIA
40080099	CATETERISMO DE CAMARAS CARDIACAS DIREITA E ESQUERDA COM CINE
40080102	CATETERISMO DE CAMARAS CARDIACAS DIREITA E ESQUERDA COM CINEA
40080110	CATETERISMO DE CAMARAS CARDIACAS DIREITA E ESQUERDA COM CINE
40080129	CATETERISMO DE CAMARAS CARDIACAS DIREITA COM CINEANGIOGRAFIA
40080137	CATETERISMO DE CAMARAS ESQUERDAS COM CINEANGIOGRAFIA E ANGIO
40080145	CATETERISMO DE CAMARAS DIREITAS E ESQUERDAS COM CINEANGIO/ G
40080153	CATETERISMO DE CAMARAS CARDIACAS DIREITA E ESQUERDA, CINEAN/
40080161	AVALIACAO HEMODINAMICA DA REATIVIDADE VASCULAR PULMONAR COM
40080170	DIAGNOSTICO ANATOMICO DAS CONEXOES SISTEMICO PULMONARES
40080188	INDIVIDUALIZACAO DAS REPERCUSOES HEMODINAMICAS NAS FISTULAS
40080196	ESTUDO HEMODINAMICO EM CANDIDATOS A TRANSPLANTES CARDIACOS O
40080200	CATETERISMO DE CAMARAS ESQUERDAS COM ESTUDO CINECORONARIOGRA
40080218	CATETERISMO DE CAMARAS DIREITA E ESQUERDA COM ESTUDO CINECOR
40080226	CATETERISMO DE CAMARAS CARDIACA DIREITA E ESQUERDA COM ESTUD
40080234	CATETERISMO DE CAMARAS CARDIACA DIREITA E ESQUERDA POR TECNI
40080242	CATETERISMO DE CAMARAS CARDIACA DIREITA E ESQUERDA COM ESTUD
40080250	CATETERISMO DA CAMARA CARDIACA ESQUERDA, CINECORONARIOGRAFIA
40080269	CATETERISMO DA CAMARA ESQUERDA E CINECORONARIOGRAFIA COM EST
40080277	CATETERISMO ESQUERDO E ESTUDO CINEANGIOGRAFICO DA AORTA E SE
40080285	CATETERISMO DE CAMARA CARDIACA ESQUERDA E/OU ESQUERDA E DIRE
40080293	ESTUDO DE METABOLISMO MIOCARDIO COM CATETERISMO DO SEIO VENO
40080307	ESTUDO DE METABOLISMO MIOCARDIO COM CATETERISMO DO SEIO VENO
40080315	BIOPSIA ENDOMIOCARDICA PERCUTANEA
40080323	BIOPSIA ENDOMIOCARDICA PERCUTANEA COM CATETERISMO DE CAMARAS
40080331	BIOPSIA ENDOMIOCARDICA PERCUTANEA COM CATETERISMO DE CAMARAS
40080340	ANGIOGRAFIA POR SUBTRACAO DIGITAL COM INJECAO VENOSA
40080358	ANGIOGRAFIA POR SUBTRACAO DIGITAL COM INJECAO ARTERIAL
40080366	ANGIOGRAFIA POR SUBTRACAO DIGITAL DA AORTA E OU SEUS RAMOS
40080374	ANGIOGRAFIA POR SUBTRACAO DIGITAL DE CAVIDADES CARDIACAS
40080382	ANGIOGRAFIA POR SUBTRACAO DIGITAL DAS ARTERIA CORONARIAS
40080390	ESTUDO ELETROFISIOLOGICO INTRACAVITARIO DO SISTEMA DE CONDUC
40080404	ESTUDO ELETROFISIOLOGICO INTRACAVITARIO DO SISTEMA DE CONDUC
40080412	MAPEAMENTO DE FEIXES ANOMALOS E FOCOS ECTOPICOS POR ELETROFI
40080420	MAPEAMENTO DE FEIXES ANOMALOS E FOCOS ECTOPICOS POR ELETROFI
40080439	ESTUDO ULTRA SONOGRAFICO TRANSLUMINAL CORONARIO OU INTRACAVI
40090019	ANGIOPLASTIA TRANSLUMINAL CORONARIA DE VASO UNICO
40090027	ANGIOPLASTIA TRANSLUMINAL DE VASOS MULTIPLOS
40090035	ANGIOPLASTIA TRANSLUMINAL CORONARIA COM USO SIMULTANEO DE BA
40090043	ANGIOPLASTIA TRANSLUMINAL CORONARIA COM SUPORTE POR CIRCULAC
40090051	ATERECTOMIA CORONARIA TRANSLUMINAL PERCUTANEA

Quadro 3A - Códigos de procedimentos cardiológicos e suas descrições
(conclusão)

CÓDIGO	DESCRIÇÃO
40090060	IMPLANTE DE ENDOPROTESE INTRACORONARIA
40090078	TROMBOLISE INTRACORONARIA
40090086	RECANALIZACAO MECANICA POR ANGIOPLASTIA TRANSLUMINAL CORONAR
40090094	VALVOTOMIA PERCUTANEA POR VIA VENOSA
40090108	VALVOTOMIA PERCUTANEA POR VIA ARTERIAL
40090116	VALVOTOMIA PERCUTANEA POR VIA TRANSEPTAL
40090124	AORTOPLASTIA TRANSLUMINAL PERCUTANEA
40090132	ANGIOPLASTIA TRANSLUMINAL PERCUTANEA NAS ESTENOSES PERIFERIC
40090140	ANGIOPLASTIA TRANSLUMINAL PERCUTANEA EM CONEXOES SISTEMICO-P
40090159	ANGIOPLASTIA TRANSLUMINAL PERCUTANEA EM ESTENOSES VENOSAS CE
40090167	ANGIOPLASTIA TRANSLUMINAL PERCUTANEA NAS RAMIFICACOES DA AOR
40090175	IMPLANTE TRANSLUMINAL PERCUTANEO DE ENDOPROTESE INTRAVASCULA
40090183	EMBOLOTERAPIA NAS FISTULAS A.V. PULMONARES
40090191	EMBOLOTERAPIA NAS CONEXOES SISTEMICO-PULMONARES
40090205	OCLUSAO PERCUTANEA DE P.C.A.
40090213	OCLUSAO PERCUTANEA DOS DEFEITOS SEPTAIS INTRACARDIACOS
40090221	RETIRADA PERCUTANEA DE CORPOS ESTRANHOS NO SISTEMA CIRCULATO
40090230	RETIRADA PERCUTANEA DE EMBOLOS PULMONARES
40090248	OCLUSAO DE VEIA CAVA INFERIOR COM FILTROS
40090256	IMPLANTE PERCUTANEO DE BALAO INTRA-AORTICO
40090264	IMPLANTE DE ELETRODO DE MARCA PASSO TEMPORARIO
40090272	ATRIOSSEPTOSTOMIA POR CATETER BALAO OU POR LAMINA
40090280	PERICARDIOCENTESE APOS CATETERISMO COMPLICADO
40090299	PERICARDIOSTOMIA COM UTILIZACAO DE CATETER BALAO
40090302	CONTROLE DE HEMORRAGIAS POR CATETER BALAO
40090329	ABLACAO DE CIRCUITO ARRITMOG. POR CATETER DE RADIOFREQ