

ERNESTO JOSUÉ SCHMITT

PREDIÇÃO DE DOENÇAS EM COMUNIDADES

Um estudo de caso em Odontologia.

Dissertação de Mestrado apresentado ao Programa de Pós Graduação em Tecnologia em Saúde da Pontifícia Universidade Católica do Paraná, como parte dos requisitos para a obtenção do título de Mestre em Tecnologia em Saúde.

Curitiba
2005

ERNESTO JOSUÉ SCHMITT

PREDIÇÃO DE DOENÇAS EM COMUNIDADES

Um estudo de caso em Odontologia.

Dissertação de Mestrado apresentado ao Programa de Pós Graduação em Tecnologia em Saúde da Pontifícia Universidade Católica do Paraná, como parte dos requisitos para a obtenção do título de Mestre em Tecnologia em Saúde.

Área de Concentração:
Informática em Saúde

Orientador:
Prof. Dr. João da Silva Dias

Co-Orientador
Prof. Dr. Samuel Jorge Moysés

Curitiba
2005

S355p
2005 Schmitt, Ernesto Josué
Predição de doenças em comunidades : um estudo de caso em odontologia
/ Ernesto Josué Schmitt ; orientador, João da Silva Dias ; co-orientador,
Samuel Jorge Moysés. – 2005.
77 p. : il. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná,
Curitiba, 2005
Bibliografia: p. 70-77

1. Cáries dentárias. 2. Previsão. 3. Epidemiologia. 4. Inteligência artificial. I.
Dias, João da Silva. II. Moysés, Samuel Jorge. III. Pontifícia Universidade
Católica do Paraná. Programa de Pós-Graduação em Tecnologia em Saúde.
III. Título.

CDD 20. ed. – 617.67



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 008

Aos 31 dias do mês de março de 2005 realizou-se a sessão pública de defesa da dissertação **“PREDIÇÃO DE DOENÇAS EM COMUNIDADES – Um Estudo de Caso em Odontologia”**, apresentada por **Ernesto Josué Schmitt**, como requisito parcial para a obtenção do título de **Mestre em Tecnologia em Saúde**, perante uma Banca Examinadora composta pelos seguintes membros:

Prof. Dr. João da Silva Dias
 PUCPR (Presidente)

João Dias
 assinatura

Aprovado
 parecer (aprov/ reprov.)

Prof. Dr. Edson E Scalabrin
 PUCPR

Edson Scalabrin

Aprovado

Prof. Dr. Paulo Sávio A de Goes
 UPE

Paulo Sávio A de Goes

Aprovado

Prof. Dr. Samuel Jorge Moysés
 PUCPR

Samuel Jorge Moysés

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.

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



“ ... para ir arrojadamente onde ninguém jamais foi ...”

(Gene Rodenberry)

DEDICATÓRIA

Aos meus pais, Ernesto e Nelzi, pelo exemplo e apoio nos momentos difíceis.

A minha irmã Nelzi, pelo carinho e lealdade.

Aos meus avós Achilles, Rachel, José e Catarina, pelo exemplo deixado nas gerações.

AGRADECIMENTOS

Ao meu orientador, Prof. João da Silva Dias, com quem muito aprendi; e quem respeito não só pela competência profissional, mas também como pessoa.

Ao meu co-orientador, Prof. Samuel Jorge Moysés, a quem desde há muito tempo devo respeito, amizade e lealdade.

À Profa. Simone Tetu Moysés e ao Prof. Léo Kriger, pelo apoio e incentivo profissional.

Ao Prof. Monir Tacla, diretor do Curso de Odontologia da PUCPR, pessoa pela qual a admiração e respeito, vem desde os anos de Graduação, pelo apoio e incentivo.

Ao amigo Claudio Jose Beltrão, colega durante o curso de Mestrado, hoje amigo e companheiro, pelo apoio.

A Pro Reitoria Comunitária e de Extensão, Prof. Adilson de Moraes Seixas, pela confiança e apoio à pesquisa.

A todos os colegas e amigos que fizeram ou fazem parte desta jornada, e que sempre estiveram ao lado.

Principalmente, a DEUS, por possibilitar novos momentos de descoberta e desafios, e mostrar os caminhos a seguir para alcançar êxito na caminhada.

SUMÁRIO

RESUMO.....	IX
ABSTRACT	X
LISTA DE TABELAS	XI
LISTA DE GRÁFICOS	XII
LISTA DE FIGURAS.....	XIII
1. INTRODUÇÃO	1
1.1 OBJETIVOS	4
1.1.1 <i>Objetivo geral</i>	4
1.1.2 <i>Objetivos específicos</i>	4
1.2 JUSTIFICATIVA	4
1.3 PROPOSTA.....	5
2 - REVISÃO DA LITERATURA	7
2.1. QUADRO ATUAL DA PREVALÊNCIA DA DOENÇA CÁRIE NO BRASIL	7
2.2. DOENÇA CÁRIE.....	7
2.2.1. <i>Deteção e monitoramento das prevalências e incidências das doenças.</i>	Erro!
<i>Indicador não definido.</i>	
2.3 CONSIDERAÇÕES IMPORTANTES SOBRE <i>DATAMINING</i>	11
2.3.1 <i>Exemplos de uso de datamining</i>	15
2.3.2 <i>Fases para implantação do datamining</i>	15
2.3.3 <i>Formas de utilização da Mineração de Dados</i>	16
2.5 SOLUÇÕES PARA A PREDIÇÃO DE DOENÇAS E ANÁLISE DE RISCO	17
3. MATERIAIS E MÉTODOS.....	20
3.1. PADRONIZAÇÃO DO EXAME	21
3.2. FICHA DE EXAME ODONTOLÓGICO	22
3.3 SOFTWARES UTILIZADOS	22
3.3.1 <i>Pacote Estatístico EPIINFO</i>	23
3.3.2 <i>Shell See5</i>	23

3.4. DISTRIBUIÇÃO DA POPULAÇÃO.....	23
3.5 SOLUÇÃO PROPOSTA.....	24
3.6 MODELAGEM DOS DADOS	26
4. RESULTADOS	31
4.1 RESULTADOS GERADOS PELO SEE5	31
4.2 ESTUDO EPIDEMIOLÓGICO DA POPULAÇÃO	35
4.2.1 Caracterização da Amostra.....	35
4.2.2 Índice CPO e ceo.....	37
5. DISCUSSÃO E CONCLUSÃO	40
5.1 DISCUSSÃO.....	40
5.2 CONCLUSÃO	42
5.3 TRABALHOS FUTUROS	44
6 - ANEXOS	45
ANEXO I – FICHA DE COLETA DE DADOS	45
ANEXO II – TELA DE ENTRADA DE DADOS EPI INFO	46
ANEXO III – RELATÓRIO COMPLETO SEE5	47
ANEXO IV – RESUMO DAS REGRAS GERADAS (SEE5) E ANÁLISE ESTATÍSTICA (EPIINFO).....	61
7. REFERÊNCIAS BIBLOGRÁFICAS	66
8. BIBLIOGRAFIA	69

RESUMO

A determinação do risco em adoecer é uma das mais antigas buscas pelos profissionais de saúde e pesquisadores da área de epidemiologia. Com esta predição é possível encaminhar os recursos humanos e financeiros de maneira a tornar o serviço público de saúde mais eficiente e menos oneroso. Neste contexto procura-se determinar um modelo de estudo onde seja possível determinar o risco individual à doença cárie dentro de uma comunidade, levando em consideração determinantes biológicos e sociais. O modelo está baseado na análise dos dados populacionais através da utilização do shell de DataMining See5 com a criação de regras de associação. A validação das mesmas se deu por análise estatística, utilizando o software EPIINFO, e também validado através do estudo Duplo-Cego. Com a análise obteve-se um grau de concordância de 87,5% entre os especialistas e o sistema preditivo do See5, criando-se deste modo um modelo preditivo eficiente e rápido para a utilização em sistemas de saúde. Além disto percebeu-se a grande diferença existente entre os índices de saúde bucal de Guaraqueçaba, onde foram estudadas 471 pessoas, quando comparado com o Brasil e com a meta da OMS, justificando a aplicação de um modelo preditivo baseado em regras.

Palavras Chave: Inteligência Artificial, Epidemiologia, DataMining, Cárie Dental

ABSTRACT

The determination of the risk in getting sick is a search of the oldest for the professionals of health and researchers of the epidemiology area. With this prediction it's possible to direct the human resources and financial in way to turn the more efficient and less onerous public service of health. In this context it tries to determine a study model where is possible to know the individual risk of dental caries inside a community, taking into account biological and social determinant. The model is based on the analysis of the population data with the use of Datamining See5 shell with the creation of association rules. The validation of the same ones felt for statistical analysis, using the software EPIINFO, and validated through the Double-blind study; with the analysis it was obtained a degree of agreement of 87,5% between the specialists and the system of See5, growing up a model predictive efficient and fast for the use in systems of health. Besides it was noticed the great existent difference among the indexes of oral health of Guaraqueçaba, where was studied 471 cases, when compared with Brazil and with the goal of OMS, justifying like this, the application of a model preditivo based on rules.

Keywords: Artificial Intelligence, Epidemiology, DataMining, Dental Caries.

LISTA DE TABELAS

Tabela 1 – Quadro Comparativo do Nº. de Habitantes por Município e Estado – Área Urbana e Rural – 1996/2000.	23
Tabela 2 – Cálculo Amostral em Populações para Levantamentos de Prevalência.	25
Tabela 3 – Cadastro de Individuo:	27
Tabela 4 - Cadastro de Exame	28
Tabela 5 – Atributos a serem utilizados no banco de dados	29
Tabela 6 – Domínio dos Atributos.	30
Tabela 7 – Tabela comparativa do duplo cego	35
Tabela 8 – Distribuição da amostra em Grupos étnicos	36
Tabela 9 – Distribuição da amostra em Grupos étnicos – área Urbana	36
Tabela 10 – Distribuição da amostra em Grupos étnicos – área ilha peças	36
Tabela 11 – Distribuição da amostra em Grupos étnicos – área ilha superagui	37
Tabela 12 – Distribuição da amostra em Grupos étnicos – área rural	37
Tabela 13 – Distribuição da amostra estratificada por sexo	37
Tabela 14 – Distribuição do tipo de cárie da amostra	38
Tabela 15 – distribuição da cárie estratificada por idade e sexo	38

LISTA DE GRÁFICOS

Gráfico 1 – Representação da variação do fator de confiança.....	32
Gráfico 2 – Distribuição das regras geradas pelo see5.....	33
Gráfico 3 – Significância estatística das regras geradas pelo see5	33
Gráfico 4 – Comparação do CPO médio 12 anos	38
Gráfico 5 – Distribuição do tipo de cárie na população estudada.....	39

LISTA DE FIGURAS

Figura 1 - Representação esquemática dos fatores concorrentes para o desenvolvimento da cárie.	7
Figura 2 - Exemplo de Ciclo Virtuoso de DataMining.	14
Figura 3 – A implementação do DM a partir do ciclo virtuoso.....	14
Figura 4 – Demonstrativo das fases para implementação do Datamining.....	16
Figura 5 – Mapa do Município de Guaraqueçaba e adjacências.....	24
Figura 7 - Imagem da tela do módulo de predição do see5.	34

1. INTRODUÇÃO

Dentro do histórico de Políticas de Saúde no Brasil observa-se uma tendência à abordagem curativa em detrimento da preventiva. Este fato não é isolado em regiões do País, mas observável em todo o território Nacional, em todos os períodos históricos. Observa-se também que, especificamente, na Odontologia, não se praticava uma Odontologia Integral como é pregado atualmente (PINTO, 1997; BRATTHAL, 1992; KRIGER, 2002).

O histórico das Políticas de Saúde no Brasil se inicia com a Revolução Industrial e, em conseqüência, com os progressos médicos que se verificaram no século passado, gerando um enfoque predominantemente tecnológico das questões de saúde, o que resultou em uma concepção mecanicista do poder médico, em detrimento do humano e social. Em decorrência disto, a medicina passou a trilhar o caminho da produção em larga escala de equipamentos altamente sofisticados, medicamentos e insumos. A transferência desses conceitos e práticas de saúde para países capitalistas periféricos resultou em uma política economicista, traduzida pela abundância de recursos para um número muito limitado de indivíduos, em contraste com uma grande parcela da população sem acesso algum aos serviços de saúde (PINTO, 1997; CHAVES, 1986).

Os países em desenvolvimento procuram enfrentar essas crises por meio da “racionalização” de seus serviços de saúde, baseada em modelos de atenção abrangentes. No Brasil, no início do período Republicano, a falta de um modelo sanitário para o País deixava as cidades à mercê das epidemias. Neste período histórico, a cidade do Rio de Janeiro apresentava um quadro sanitário caótico, caracterizado pela presença de diversas doenças graves que acometiam a população, como a varíola, a malária, a febre amarela e posteriormente, a peste, o que acabou gerando sérias conseqüências tanto para a saúde coletiva quanto para outros setores, como o do comércio exterior, visto que os navios estrangeiros não mais queriam atracar no porto do Rio de Janeiro em função da situação sanitária existente na cidade. É então nomeado Oswaldo Cruz como Diretor do Departamento Federal de Saúde Pública, que se propôs a erradicar a epidemia de febre amarela na cidade do Rio de Janeiro (modelo

conhecido como *Campanhista*). Posteriormente, a organização da assistência médica sofreu mudanças com a Reforma Carlos Chagas, nos primeiros anos do século XX. Era um período de crescimento industrial notável, com uma agricultura baseada no conceito da monocultura do café. Porém, as condições de vida do trabalhador eram de extrema dificuldade e exploração. Existia um quadro político potencialmente revolucionário, influenciado pelos ideais anarco-sindicalistas e os ideais comunistas, trazidos pelos trabalhadores da Europa. Estes fatos influenciaram decisivamente as medidas governamentais adotadas no campo da Previdência Social após 1920.

Dentro deste contexto histórico, é realizada então a Reforma Carlos Chagas em 1920, que, dentre outras providências, criou o *Departamento Nacional de Saúde Pública* – D.N.S.P, e inicia a intervenção do Estado na prestação de serviços médico-assistenciais. Neste momento, havia a dicotomia entre a prática Curativa e Preventiva, que se arrastou até 1988, com o surgimento do Sistema Único de Saúde – SUS, onde é preconizada a extinção desta prática dicotômica.

No período conhecido como ESTADO NOVO, pretendeu-se estender a previdência social e seus benefícios a todas as categorias do operariado urbano organizado. Desta forma, as antigas CAPs são substituídas pelos INSTITUTOS DE APOSENTADORIA E PENSÕES (IAP). Nestes institutos os trabalhadores eram organizados por categoria profissional (marítimos, comerciários, bancários) e não por empresa. Em 1933, foi criado o primeiro Instituto de Aposentadoria e Pensões: o dos Marítimos (IAPM). Os IAPs foram criados de acordo com a capacidade de organização, mobilização e importância da categoria profissional em questão. Além de servir como importante mecanismo de controle social, os IAPs tinham, até meados da década de 50, papel fundamental no desenvolvimento econômico deste período, como “instrumento de captação de poupança forçada”, através de seu regime de capitalização.

Até o final dos anos 50, a assistência médica previdenciária não era importante. Os técnicos do setor a consideravam secundária no sistema previdenciário brasileiro, e os segurados não faziam dela parte importante de suas reivindicações.

Com o SUS, passa-se a valorizar a Promoção de Saúde, que tem uma participação eficaz, evitando a instalação de novos processos de doença na população; desta forma reduzindo, desta forma os gastos com as necessidades curativas desta população (MOYSÉS, 1997; KRIGER, 1997).

Para a estruturação adequada dos serviços de saúde é necessária a distribuição da atenção em saúde em **Níveis de Atenção**, a saber: *Atenção Primária (Básica)*, *Secundária* e *Terciária*. Com a implementação destes níveis e a reorganização da Atenção Básica pelo Ministério da Saúde, processo que vem ocorrendo há quase uma década, surge então a necessidade de também se reorganizar a Atenção Básica em Odontologia. A principal estratégia do governo ao enfatizar e garantir prioridade ao atendimento básico em saúde se consolida em 1998, com a introdução do Piso de Atenção Básica (PAB), criado através da Portaria GM/MS n.1882, de 18/12/97. O PAB visa garantir aos municípios um valor per capita repassado pelo governo federal, destinado exclusivamente ao custeio de um elenco predeterminado de procedimentos e ações (consultas médicas, odontológicas, vigilância em saúde, promoção de saúde etc.). Esse mecanismo dissocia a produção do faturamento, modelo que, baseado em faturamento e número de atendimentos, perpetuava um modelo de pós-pagamento, incentivando a lógica do aumento crescente de procedimentos e um superfaturamento dos serviços prestados. Procedimentos com maior custo usualmente predominavam sobre os de menor custo, atuando como um desestímulo à adoção de medidas preventivas e de promoção à saúde. Isto posto, é possível perceber que a Atenção Básica é o nível de atenção que pretende atender o maior número de usuários do SUS, senão a todos, uma vez que para ter acesso aos outros níveis de complexidade, passa necessariamente pela atenção básica (BODSTEIN, 2002).

Neste contexto, pode ser percebido que um modelo preditivo é necessidade eminente e solicitada por gestores para que viabilizem uma melhor organização das ações e do serviço, e em conseqüência uma melhor atuação da equipe de saúde bucal. Tem-se portanto a discussão da **ANÁLISE DE RISCO E PREDIÇÃO DE DOENÇAS EM COMUNIDADES** como um tema de grande extensão, pois basta observar a vasta gama de doenças, dentre as quais citam-se as doenças sexualmente transmissíveis, as

relacionadas com a higiene pessoal e ambiental, além da diabete e hipertensão – todas com uma epidemiologia bem definida, hoje presentes na população (ROUQUAYROL, 1999).

1.1 OBJETIVOS

1.1.1 Objetivo geral

O presente trabalho tem por objetivo propor um modelo que seja capaz de diagnosticar o aparecimento e distribuição da cárie dental em populações, viabilizando a melhoria do modelo assistencial, ou seja, da atenção em saúde.

1.1.2 Objetivos específicos

- a. Avaliar a relação existente entre o modelo assistencial e o perfil epidemiológico da população, visando definir o grau da necessidade de um modelo baseado em predição de doença.
- b. Obtenção de regras de associação que tornem viável a avaliação dos fatores concorrentes para o aparecimento da cárie dental em populações.

1.2 JUSTIFICATIVA

Existe hoje uma maior tendência de aplicações de uma Política de Saúde em Odontologia, calcada na Odontologia de Promoção de Saúde, sem esquecer em hipótese alguma das necessidades já instaladas na população que está em atendimento. Promoção, prevenção e cura devem, juntos, construir a atenção básica. Com a implantação de um modelo preditivo para a cárie dental, a melhora da qualidade da atenção em saúde bucal é justificativa suficiente para este estudo, pois poderá melhor organizar o serviço de odontologia e, desta forma, os índices de saúde bucal tenderão a melhoras significativas.

O uso da ferramenta de Inteligência Artificial - o *Datamining*, é o suporte que o estudo necessitava. Somada à crescente afinidade das pesquisas em Saúde Coletiva e Tecnologia em Saúde, cria-se o contexto ideal para a viabilização e discussão da hipótese de que é possível saber quem de determinada comunidade irá adoecer de cárie dental.

1.3 PROPOSTA

Para a aplicação desta metodologia se torna imprescindível a necessidade de se conhecer a Epidemiologia de cada ambiente a ser trabalhado. Neste caso específico, o ambiente é a comunidade de Guaraqueçaba – PR. Guaraqueçaba é uma pequena cidade em número de habitantes e não em extensão territorial. Situada no litoral norte do Paraná, na divisa com São Paulo, possui cerca de 8.300 habitantes, divididos na área de continente e na região das ilhas. Este município possui características especiais por se localizar em uma Área de Proteção Ambiental (APA), que acaba por restringir o crescimento populacional. O trabalho com esta comunidade se justifica por se tratar de um grupo bastante estancado em termos epidemiológicos e que, portanto, possui uma variação populacional quase nula, o que dará respaldo aos dados obtidos. A PUCPR já desenvolve ações dentro da comunidade através do Programa de Ação Comunitária e Ambiental - Proação, o que facilita a tarefa de busca de dados, necessidades básicas e de logística.

A população hoje assistida na área urbana é de cerca de 2.400 habitantes da sede do município. Porém, a pretensão é de se trabalhar com uma amostra representativa desta população, com representantes da área urbana, rural e ilhéus.

Este trabalho poderá iniciar um novo momento dentro da concepção de Sistemas de Apoio à Tomada de Decisão, ou Sistemas Baseados em Conhecimento, para a Odontologia, podendo ser de aplicação imediata e ampla, com restrições que podem ser diminuídas ainda mais quando se verifica que o software a ser implementado poderá trabalhar com uma base de dados de casos anteriores da doença em questão na comunidade ou em comunidades correlatas. Isto só é possível quando se considera que as doenças a serem estudadas estão relacionadas com fatores

inerentes a hospedeiros e agentes causadores (ROUQUAYROL, 1999; MALTZ,1997; BEZERRA, 1997; NEWBRUM, 1988; BRATTHAL, 2001).

A tendência maior dentro da Informática em Saúde são os Sistemas de Apoio à Tomada de Decisão e, em consequência disto, são os que mais geram polêmica, pelo fato de muitos técnicos se sentirem ameaçados com a presença da tecnologia. Este modelo busca automatizar o processo de predição com a presença dos técnicos para atuar nos focos potenciais de desenvolvimento de doença (micro-áreas de risco) apontados pelo sistema. Hoje este processo é feito de forma muito lenta, utilizando-se o recurso de busca ativa e análise de vários fatores. Normalmente, em uma cidade como Curitiba, em uma área de abrangência para cada unidade básica que conta com cerca de 16.000 habitantes existem dois ou no máximo três técnicos de formação superior, além dos auxiliares para tal pesquisa.

2 - REVISÃO DA LITERATURA

Devido à interação dos temas a serem abordados na pesquisa, esta revisão de literatura, será apresentada em três partes distintas, mas que devem trabalhar de forma integrada: 1) Entendimento da doença cárie; 2) Entendimento da cárie no contexto Brasil e 3) Estudo da Inteligência Artificial - Datamining.

2.1 DOENÇA CÁRIE

Para definir meios de controle da doença cárie é necessário o entendimento da evolução da doença. A cárie é uma doença multifatorial, tradicionalmente definida pelo modelo da História Natural das Doenças, onde se estabelece a necessidade da atuação em conjunto de quatro fatores, a saber:

- a. **Hospedeiro**, a estrutura morfológica dos dentes, alterações de formação do esmalte e hábitos de higiene;
- b. **Microbiota**, que varia dependendo da superfície do dente em estudo;
- c. **Substrato ou Dieta**, para que ocorra um desenvolvimento ou crescimento microbiano;
- d. **Tempo**, fator de extrema importância, uma vez que, com o aumento do tempo de permanência do substrato associado ao tempo de maturação da placa bacteriana, há o favorecimento do desenvolvimento da lesão.

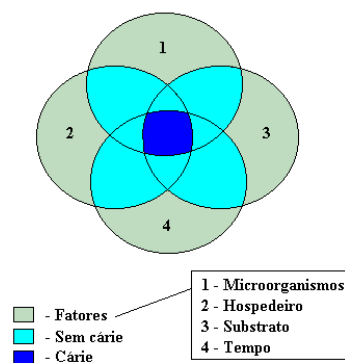


FIGURA 1 - REPRESENTAÇÃO ESQUEMÁTICA DOS FATORES CONCORRENTES PARA O DESENVOLVIMENTO DA CÁRIE.

Ao considerar apenas os determinantes biológicos da cárie dental, facilmente se comete um equívoco, pois, conforme vários autores defendem, a cárie dental possui também determinantes sociais que devem ser percebidos na perspectiva coletiva obtida com o uso do método epidemiológico. A identificação de grupos de risco e a construção de novos modelos explicativos para a doença englobam determinantes sociais e psicológicos, até então ignorados pela Odontologia.

Além das variáveis clássicas como sexo, idade e distribuição geográfica, a abordagem da doença cárie através de um ordenamento social pode ser organizada em três grupos distintos de variáveis: as socioeconômicas, as culturais e as comportamentais. As variáveis socioeconômicas, segundo o mesmo autor, acabam definindo maior ou menor acesso a meios preventivos, à educação e a uma dieta cariogênica. As variáveis culturais também assumem um papel importante na determinação da doença, pois elas vão definir os padrões de alimentação, o consumo de açúcar e também a adoção de métodos preventivos. Já as variáveis comportamentais estão ligadas ao comportamento individual; entram nesse grupo de variáveis o auto-cuidado com a saúde e hábitos de higiene (FREITAS, 2001).

Trabalhos recentes indicam como os hábitos de vida de uma família podem modificar as condições de saúde de uma pessoa, como: “escolaridade materna abaixo ou igual a 8 anos, renda familiar menor que 6 salários mínimos, não freqüentar pré-escola e consumo de doces pelo menos uma vez ao dia aos 6 anos de idade foram fatores de risco à cárie” (PERES, 2003). Suas conclusões ainda explicitam que, além de métodos de controle de doença específicos, dever-se-ia utilizar ferramentas de intervenção relacionadas com os fatores de risco social. Em reforço a isto, cita-se recente estudo em Minas Gerais, que comprovou com análises multivariadas a associação entre condições socioeconômicas desfavoráveis e piores índices de Saúde Bucal (LUCAS, 2005).

Evidencia-se, desta forma, a necessidade de um modelo de dados com abrangência suficiente para dar suporte e interpretar as variáveis biológicas e sociais

das doenças. Desta forma se obtêm-se ferramentas de intervenção focadas nas reais necessidades da comunidade em estudo, descobrindo assim, novas e importantes relações na história natural das doenças.

2.2. QUADRO ATUAL DA PREVALÊNCIA DA DOENÇA CÁRIE NO BRASIL

O Brasil, desde há muito tempo, vem sendo considerado o “País dos desdentados”, apesar de ter um estoque muito grande de cirurgiões-dentistas, um dos maiores do mundo (Ministério da Saúde – Brasil, 1986).

Este paradigma pode e já está sendo quebrado. Com a aplicação pelas autoridades de um modelo assistencial mais comprometido com a saúde e a conscientização da população, houve quedas sensíveis na incidência de doenças bucais, em especial a cárie, em cidades como Curitiba, onde a meta da OMS para o ano 2010 é um CPO-D¹ igual ou menor a 1,0 (GOMES, 2004).

O trabalho hoje realizado e preconizado pelo Ministério da Saúde não é marcado por uma dicotomia entre a Odontologia Curativa e a Odontologia Preventiva, e sim por uma interdisciplinaridade marcante para que seja alcançado um atendimento calcado em uma Odontologia Integral (MOYSÉS, 1997; KRIGER, 1997; BARATIERI, 1995), como podemos observar ao estudar o programa Brasil Sorridente², que hoje está sendo implantado no Brasil pelo Ministério da Saúde.

Uma grande preocupação dos Cirurgiões-Dentistas que se dedicam à Saúde Coletiva em especial é a diminuição a valores extremamente pequenos dos índices de saúde bucal. Neste contexto o CPOD e ceod, preconizados pela Organização Mundial de Saúde (BLINKHORN, 2001; GOMES, 2004), a prevalência de doenças bucais, tendo como um foco especial de atenção a doença cárie, que é em boa parte responsável pelo edentulismo de grande parte da população, segundo o MS, na faixa etária de 65 a 74 anos, o número de dentes perdidos dos brasileiros é de 25,83 em média, o que representa quase 93% do total do índice (MINISTÉRIO DA SAÚDE – BRASIL, 2004).

¹ - Índice preconizado pela Organização Mundial de Saúde, que mede os dentes **C**ariados, **P**erdidos e **O**bturados.

² - Programa de Atenção Básica em Odontologia – MS Brasil.

Em estudos anteriores da cárie dental no Brasil, os pesquisadores afirmavam: “Aos 07 anos, cada criança tem 2,2 dentes cariados, sendo que nesta idade 21% tiveram acesso à Odontologia; aos 12 anos o CPO-D é 6,6 e este índice coloca o Brasil entre os países com maior prevalência de cárie; em consequência de um grande número de extrações, as necessidades protéticas da população são elevadas” (MINISTÉRIO DA SAÚDE – BRASIL, 1986); após 18 anos de controle da cárie com a aplicação de um modelo assistencial mais adequado à realidade brasileira, o Brasil encontra-se com o seguinte perfil epidemiológico: “Quase 27% das crianças de 18 a 36 meses apresentam pelo menos um dente decíduo com experiência de cárie dentária, sendo que a proporção chega a quase 60% das crianças de 5 anos de idade. Quanto à cárie dentária na dentição permanente, quase 70% das crianças brasileiras de 12 anos e cerca de 90% dos adolescentes de 15 a 19 apresentam pelo menos um dente permanente com experiência de cárie dentária e CPO-D médio igual a 2,78. De uma forma geral, são valores considerados altos, particularmente quando são levadas em conta as faixas etárias de 18 a 36 meses e a idade de 5 anos” (MINISTÉRIO DA SAÚDE – BRASIL, 2004).

Considerando estes índices obtidos pelo Ministério da Saúde, é fácil constatar a necessidade de uma atuação coletiva e de visão integral. Uma vez que a doença cárie deixa as suas seqüelas, as restaurações e os dentes perdidos continuam a inchar o índice CPOD, já que este, além de conter os dentes cariados, também possui na sua composição os dentes já restaurados e os perdidos por cárie. “Destaca-se o fato que o componente perdido é responsável por cerca de 66% do índice no grupo de 35 a 44 anos e quase 93% no grupo de 65 a 74 anos” (MINISTÉRIO DA SAÚDE – BRASIL, 2004).

Com estas considerações se faz necessária uma busca nos fatores de desenvolvimento e controle da doença cárie para que se obtenha sucesso em um programa, demonstrado pela queda dos índices epidemiológicos e pela proximidade das metas estabelecidas pela OMS.

2.3 CONSIDERAÇÕES IMPORTANTES SOBRE *DATAMINING*

Datamining ou Mineração de Dados é um termo criado pela comunidade de pesquisadores Trata-se de um conjunto de técnicas reunidas da Estatística e Inteligência Artificial, com o objetivo de descobrir “conhecimento novo”, que possa estar escondido em grandes massas de dados (CARVALHO, 2001).

Muitas das ferramentas utilizadas neste processo são conhecidas e aplicadas há muito no meio gerador de tecnologia. Porém, esta ferramenta está se popularizando por diversos fatores, dentre os quais podemos citar o crescimento do número de computadores nas empresas prestadoras de serviços públicos e privados e, conseqüentemente, o aumento significativo no volume e pertinência de informações armazenadas em forma digital (CARVALHO, 2001).

O Datamining pode ser considerado como o processo de *Machine Learning* – *ML*, e para tal é necessária a definição deste processo no contexto no qual a pesquisa ocorreu. Para aplicar métodos de *ML*, tanto para a extração de conhecimento como para a construção de classificadores, é preciso primeiramente definir o paradigma e a estratégia que será aplicada. Enquanto o paradigma diz respeito à forma como o conhecimento adquirido será representado, a estratégia é a forma como o conhecimento é propriamente adquirido. Diversos paradigmas e estratégias de *ML* foram propostos. Entre os paradigmas, podemos citar: conexionista, genético, estatístico, instanced-based e simbólico (MONARD, 1997).

Para que se possa criar uma relação pertinente entre o prestador de serviços e o tomador dos serviços, é necessária uma análise criteriosa do banco de dados denominado de *data warehouse*³, enquanto a capacidade de aprendizado é dada pelas técnicas de **Inteligência Artificial**: Desde o início dos anos 60, a comunidade de Inteligência Artificial – IA tem pesquisado sistemas que seriam capazes de aprender e de induzir um modelo do processo de decisão de um especialista, com base em um conjunto de exemplos - **o modelo Estatístico** – onde os algoritmos usados em IA são

³ Coleção de dados orientada por assuntos, integrada, variante no tempo, não volátil, que tem por objetivo dar suporte aos processos de tomada de decisão [IHMO92].

adequados para descobrir regras e modelos em conjuntos de dados artificiais e relativamente pequenos. A premissa feita é a de que todas as informações necessárias estão disponíveis. Toda a informação utilizada pelo especialista para tomar uma decisão é armazenada na base de dados. Nestas condições, o algoritmo de indução gera modelos que fazem predições corretas para cada exemplo no conjunto de dados e modela corretamente o processo de decisão do especialista. Conjuntos de dados reais não contêm toda a informação necessária para se tomar decisões corretas. Isto não significa que alguns dados têm ruído ou que algumas variáveis são desconhecidas, significa que algumas informações relevantes simplesmente não estão disponíveis. Neste momento a estatística pode entrar. Datamining utiliza inteligência artificial combinada à estatística para gerar bons modelos, mesmo que toda a informação necessária não esteja presente. Testes estatísticos são necessários para validar a qualidade do modelo: o Datamining.

Durante o processo de busca, a qualidade de muitos modelos tem de ser validada. Informações estatísticas sobre os dados são necessárias para avaliar o quão correto é o modelo gerado. Desta forma, datamining tipicamente envolve o envio de muitas consultas aos bancos de dados da empresa, resultando em carga de trabalho pesada para toda a base e para a rede. Técnicas de datamining avançadas utilizam métodos de otimização para reduzir a interação com as bases de dados. Por exemplo, o resultado de uma consulta é temporariamente armazenado, de forma que consultas subseqüentes a informações similares podem ser atendidas sem que a base de dados seja acessada.

A mineração de dados consiste basicamente em descobrir relações entre situações, classificar usuários, prever acesso, problemas, doenças, dentre outras. Ao considerar isto, verifica-se que, com tantas atividades agrupadas sob o mesmo teto, há uma grande possibilidade de confusão sobre o tema. Este processamento e classificação pode facilitar a compreensão. Michel Berry (<http://www.data-miners.com>) propõe a classificação de sistemas de datamining em três dimensões:

- Natureza da tarefa

- Classificação: associação um item a uma classe, de opções pré-definidas;
 - Segmentação: é uma tarefa de classificação sem classes pré-definidas; consiste, então em agrupar registros semelhantes e separar os diferentes;
 - Predição: preencher o valor de um registro, baseado em atributos anteriores;
 - Descrição: percepções diferentes das apresentadas em tabelas e/ou relatórios.
- Objetivo a ser buscado
 - Produção de um modelo: é o mais utilizado para as tarefas de predição, produzindo um *score*, que possa ser utilizado em dados ainda brutos;
 - Produção de informação: normalmente utilizada para as tarefas descritivas, buscando uma nova perspectiva ao analisar os dados.
 - Grau de estruturação dos dados

Com esta classificação agrupam-se sistemas parecidos, para assim aplicar algoritmos e procedimentos similares. Desta forma, a definição para Mineração de Dados é “*o processo de exploração e análise de grandes quantidades de dados com o objetivo de descobrir padrões ou regras que permitam uma melhor compreensão da informação contida nos mesmos*” (<http://www.integra.com.br>).

Após a definição de mineração de dados e o surgimento de um ambiente favorável à implementação, é necessário se definir qual o processo ou metodologia específico a ser utilizado para este projeto. Os processos de mineração de dados normalmente utilizam o “*ciclo virtuoso de datamining*”, onde algoritmos e a informação obtida se beneficiam de processos anteriores do mesmo modelo, incorporando as respostas obtidas em futuros processamentos.

O alto potencial da mineração de dados só pode ser alcançado quando o ciclo é corretamente implementado: identificação de oportunidades, elaboração de modelo

para determinar as regras, elaboração de ações e medição dos resultados destas ações que, por sua vez, possibilitam a identificação de novas oportunidades e assim sucessivamente, conforme a figura 2.

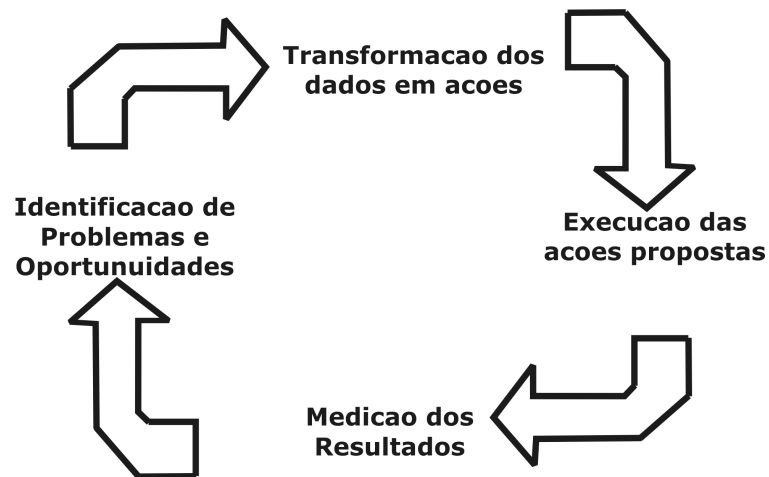


FIGURA 2 - EXEMPLO DE CICLO VIRTUOSO DE DATAMINING.

O processo de implementação da mineração de dados pode ser caracterizado pela figura 3:

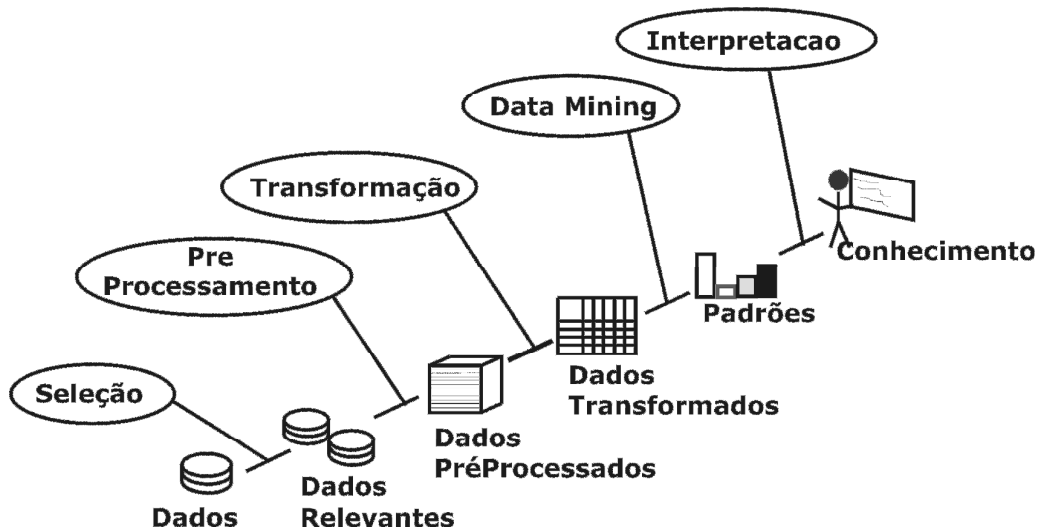


FIGURA 3 – A IMPLEMENTAÇÃO DO DM A PARTIR DO CICLO VIRTUOSO.

Fica evidente que não basta ter a base de dados, é necessário saber o que se quer extrair da base para gerar o conhecimento. As técnicas do datamining se tornam utilizadas para exploração de dados por vários motivos, dentre os quais:

1. Por necessitar de um grande volume de dados, isto é necessário para calibrar os algoritmos e extrair dos dados conclusões confiáveis.
2. Padronização dos dados com a utilização do datawarehouse, já que estas técnicas necessitam de dados limpos, padronizados e organizados.

2.3.1 Exemplos de uso de datamining

Na área médica, é possível a criação de grandes bancos de dados sobre a sintomatologia, exames complementares, tratamento e curso destas doenças nos pacientes. A mineração de dados, neste caso, pode dar novas perspectivas nas relações entre as doenças e perfis profissionais, sócioeconômicos, hábitos pessoais e local de moradia. Essas relações são utilizadas para melhor entendimento das doenças e seus tratamentos.

2.3.2 Fases para implantação do datamining

A técnica de *datamining* não se resume apenas em analisar grande volume de dados e novas relações. Para se obter resultados compatíveis com a expectativa, deve-se ter um planejamento adequado, no qual se podem destacar as seguintes fases, conforme a figura 3:

1. Identificação de um problema ou definição de um objetivo a ser alcançado (nesta pesquisa, o desenvolvimento/aplicação de um sistema capaz de diagnosticar com eficiência o aparecimento e distribuição de doenças em populações);
2. Descoberta de novas relações por técnicas de *datamining*;
3. Análise da pertinência das novas relações descobertas;
4. Uso racional das novas relações descobertas;

5. Avaliação dos resultados e
6. Redefinição de problemas ou objetivos.

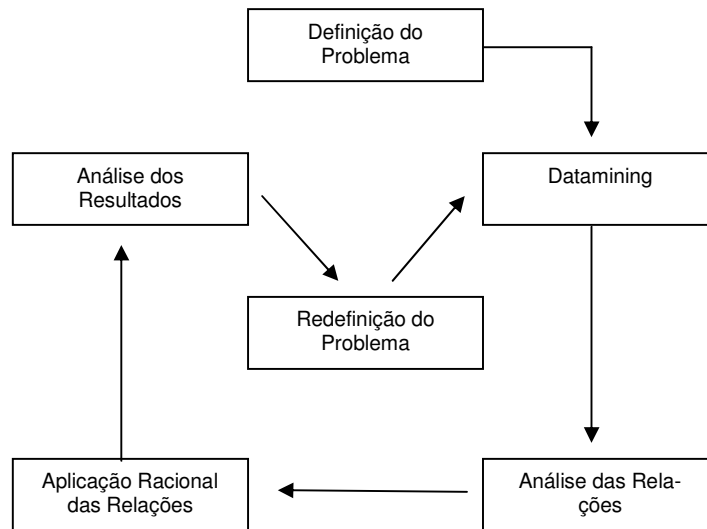


FIGURA 4 – DEMONSTRATIVO DAS FASES PARA IMPLEMENTAÇÃO DO DATAMINING.

2.3.3 Formas de utilização da Mineração de Dados

Em função do tipo do problema que esteja sendo estudado e relacionando-o com o conhecimento acumulado sobre este problema, define-se qual a forma de aplicação da *mineração de dados*, a saber:

1. Descoberta não supervisionada de relações, utilizada freqüentemente quando não se conhece as relações e as implicações do fenômeno em estudo. Esta técnica gera várias relações, que precisam ser estudadas para separar as que realmente são pertinentes ao estudo. Pequeno nível de conhecimento específico da área para a implantação da mineração de dados.
2. Teste de hipóteses, utilizada quando se conhece algumas relações e se deseja conhecer novas características, fazendo uma confirmação ou refutação de uma hipótese. Razoável nível de conhecimento específico da área para a implantação da mineração de dados.

3. Modelagem matemática de dados, utilizada quando se tem um maior conhecimento da área e da relação que se deseja estudar. Alto nível de conhecimento específico da área para a implantação do *datamining*.

2.5 SOLUÇÕES PARA A PREDIÇÃO DE DOENÇAS E ANÁLISE DE RISCO

Na literatura especializada, são raras as referências que tratam da odontologia e do trabalho com grande número de pessoas, visando melhorar o Planejamento Estratégico e a implantação de Serviços de Saúde; as soluções propostas são sempre em termos individuais e visando soluções para pacientes que se encontram com algum tipo de patologia já instalada.

Quando se busca referenciais para a predição, deve-se citar FERRO, 2001, que afirma e comprova ser possível a implantação de Inteligência Artificial e técnicas de descobrimento de conhecimento em áreas de pesquisa clínicas, como na utilização em Processamento de Sêmen, que consiste de dados reais de uma clínica de reprodução humana, e em estudos de cardiopatias (conjunto de dados *Heart Disease*). Ambos os trabalhos utilizaram estudos bibliográficos e consultas com especialistas para a compreensão do domínio e a relevância do estudo em questão. Para o estudo das cardiopatias, foi utilizado o banco de dados da *heart disease*; já no outro estudo (processamento de sêmem), foi construído um banco de dados específico para as pesquisas. No estudo das cardiopatias foi aplicado o método de Aprendizado de Máquina Indutivo Simbólico Supervisionado utilizando o C 4.5 que é um algoritmo que pode ser usado para auxiliar nessas tarefas, sendo algumas vezes apenas uma decisão de sim ou não, concebida por um fornecedor de crédito ou algo um pouco mais complexo como o diagnóstico de uma doença. Muitas vezes essas classificações seguem um padrão, por isso muitos especialistas fazem a classificação de um caso olhando para modelos anteriores cuja classe é conhecida. O C 4.5 é capaz de aprender, olhando para um conjunto desses casos, como eles são classificados e a partir daí fazer uma predição para novos casos. Ou seja, o C 4.5 gera um classificador que é capaz de agir como um especialista, classificando os casos desconhecidos -

para indução de árvores de decisão, pelo fato de buscar novos conhecimentos interessantes, e a construção de classificadores para apoio à tomada de decisão.

A análise final foi realizada aplicando os algoritmos citados acima e, juntamente com o especialista, os resultados gerados pelas árvores e as regras eram analisados usando-se a técnica de estudos epidemiológicos denominada *Duplo Cego*, na qual se descreve atributos presentes na árvore e o especialista emite sua opinião sobre em qual classe aquela regra deve se encaixar. Isso era feito sem que ele soubesse qual era a classe atribuída àquela regra pelo indutor.

Para a análise do banco de sêmem foi utilizado o datamining aplicando o mesmo método de aprendizado de máquina, com o algoritmo C 4.5. Como relata o autor, nos dois trabalhos, foi possível extrair um conhecimento do banco de dados da área médica de onde, de forma manual, não seria possível obter tal relação.

Para Eric Bramir, diretor da Pediatria do Bradem Tumor Research Program at Children's Memorial Hospital in Chicago (http://www.healthcare-informatics.com/issues/2004/04_04/hagland.htm), o datamining possibilita uma revolução na pesquisa médica, como pode ser comprovado na sua pesquisa que estudava os genes do câncer infantil, na qual cerca de 30.000 a 40.000 mil genes do genoma humano eram passíveis de estudo. Com a utilização de um software baseado no SPSS⁴, pode-se isolar melhor as diferenças entre as crianças e aplicar melhores doses de quimioterapia e radioterapia. Bremer frisa que a próxima fronteira é a pesquisa médica. Cita ainda a Oracle, que está trabalhando na implantação de um Protocolo de Datamining – bastante complexo – para ser utilizado no Projeto Genoma. Estão incluídas no projeto instituições como National Institutes of Health, Bethesda, Md. Com datamining, cita o pesquisador, "nós podemos descobrir centenas, talvez milhares de variáveis, então poderemos classificar e organizar em categorias". O potencial para o trato com o paciente e para a saúde coletiva tem encorajado poucas organizações pioneiras a desenvolverem seus próprios softwares. Médicos pesquisadores do Hospital de Crianças de Boston desenvolveram um sistema analítico chamado de AEGIS

⁴ - Statistical Package for the Social Sciences.

(Sistema Integrado Geotemporal e Epidemiológico Automatizado). Este sistema busca e analisa os dados a partir de vários departamentos de emergência, procurando por padrões, conforme explica Kenneth D. Mandl, M.D., professor da Harvard Medical School.

Dados relativos a informações e exames médicos de pacientes com trombose eram analisados pelo software SPSS Clementine datamining. A habilidade para prever o começo e o prognóstico da trombose é fundamental na intervenção da doença, pois padrões de sintomas e exames de laboratório podem indicar uma pré-trombose para condição de trombose ativa. A pesquisadora Susan Jensen, da SPSS (2001) utilizou modelos de predição, regras de associação e descoberta de sucessão para investigar estes padrões. A trombose é uma complicação potencial de doenças de colágeno; para entender os mecanismos responsáveis e o começo de trombose, foram empreendidas análises para quantificar um perfil dos resultados de exame de laboratório que indicam presença de trombose e a sucessão de a) sintomas b) resultado de exame c) tipos de doenças de colágeno que podem prever ou podem se associar com o começo de trombose. Durante o processo, o autor cita sua dificuldade de julgar a validade dos modelos devido à falta de conhecimento de domínio. Devemos considerar, dentre as conclusões alcançadas pela pesquisadora, como a mais significativa a que relata que nos dados de treinamento havia uma precisão de predição global de 88%, contra aproximadamente 82% de precisão de prever presença de trombose com exames de laboratório.

Porém, o que se discute neste caso é a possibilidade de implantação de um modelo de predição capaz de auxiliar na tomada de decisão do clínico. No julgamento de alguns casos, o especialista pode não ter a necessidade de tal sistema, porém o médico vinculado ao serviço de atendimento, considerado como generalista passa a ter esta necessidade durante a sua rotina clínica, considerando-se a falta de especialistas em áreas distantes, custo de formação e dificuldade de transporte e exames complementares.

3. MATERIAIS E MÉTODOS

Testes preditivos podem auxiliar na detecção de doenças e a interpretação depende basicamente dos seus valores de sensibilidade e especificidade. A predição positiva (valor preditivo positivo) se associa basicamente à especificidade: testes específicos, por terem poucos falsos positivos, levam a altos valores preditivos positivos. Auxiliam a diagnosticar a doença quando não estão alterados. A predição negativa (valor preditivo negativo) se associa à sensibilidade: testes sensíveis, por terem poucos falsos negativos, geralmente têm altos valores preditivos negativos. Auxiliam a afastar a presença da doença quando estão normais.

O teste ideal para predição da Cárie Dental deveria ser de simples realização, não invasivo, de fácil reprodução e com altos valores de sensibilidade e especificidade (CHAVES, 1986; OMS, 1999). A cárie é uma doença que deixa suas seqüelas, sendo então importante um índice que forneça uma idéia da situação atual e da evolução da doença e dos tratamentos realizados na população (CHAVES, 1986; OMS, 1999). Segundo JOHNSON in CHAVES, “um índice útil para a indicação do estado de saúde bucal, em relação com uma doença ou condição determinada, deve preencher os seguintes critérios: Pertinência, Confiança e Significado”. Desta forma tem-se o entendimento claro para a utilização de um índice como o CPO D, que representa Dentes Cariados/ Perdidos/ Obturados (Restaurados). Este índice foi escolhido para representar a situação atual da doença cárie na população, pois mostra claramente quantos dentes o indivíduo/população possui remanescentes na boca, hígidos e cariados, além de ser preconizado pela OMS como índice-base para comparações entre estudos internacionais de saúde bucal (CHAVES, 1986; OMS, 1991; OMS, 1999).

Faz-se importante lembrar que a queda do CPOD, quando este se encontra entre os valores Alto/Médio, é extremamente rápida; quando está entre Médio/Baixo, a queda passa a ser mais difícil, ou seja, mais lenta. Contudo, a queda do CPOD no patamar médio pode ser potencializada quando aplicadas medidas coerentes aos grupos de risco da doença (BRATTHAL, 1992; PITT, 1992; BRATTHAL, 2001). Esta

consideração é importante para que se possa montar o modelo de análise dos dados e a análise correta das relações geradas pelo software, quando da mineração no banco de dados.

Este trabalho se relaciona com fatores de risco (social e biológico) da cárie dental a todos os habitantes de casas pré-selecionadas randomicamente. Todos os residentes da habitação sorteados terão suas fichas cadastradas e acompanhadas, uma vez que a resposta que o sistema tem por objetivo gerar, não tem foco em uma faixa etária, mas sim em um grupo populacional.

Com a análise estatística destes dados que serão apresentados a seguir tentar-se-á demonstrar que a Promoção de Saúde é condição imprescindível para o controle da doença cárie. A capacidade de discernimento e reconhecimento de novas relações do processo saúde-doença a partir da análise correta do banco de dados com o Datamining, utilizando o shell See5 e com a validação estatística do EPIINFO⁵, torna o processo de datamining fundamental para o modelo proposto, além de demonstrar que sistemas computacionais podem ser referência para o planejamento estratégico em saúde, quando associados à epidemiologia e à prática baseada em evidências.

3.1. PADRONIZAÇÃO DO EXAME

Conforme a padronização adotada pela Organização Mundial de Saúde para Levantamentos Epidemiológicos os exames serão realizados de acordo com os seguintes critérios(OMS, 1991; OMS, 1999):

1. Dente Hígido: é considerado como hígido o dente que não apresentar evidências de cáries, tratadas ou não. Os estágios iniciais do desenvolvimento de cárie não são considerados pela dificuldade em detectá-los em exames clínicos, mas também serão considerados como hígidos os dentes que apresentarem os fatores: manchas esbranquiçadas; descoloração ou manchas rugosas; sulcos e fissuras do esmalte, manchados, que prendam o explorador, mas não apresentam uma base amolecida, esmalte socavado⁶ ou amolecimento das

⁵ Pacote de estatística desenvolvido para aplicação em Saúde. Distribuído pela OMS.

⁶ Socavado – dente com cárie, onde o dente encontra-se escavado por baixo do esmalte.

paredes; áreas escuras, brilhantes, duras e fissuradas do esmalte de um dente com fluorose moderada ou severa.

2. Dente Cariado: Quando uma lesão em sulco, fissura ou superfície lisa apresentar tecido amolecido não base, descoloração do esmalte ou de parede ou ainda possuir restauração provisória.
3. Dente Obturado e Cariado: Quando possuir uma ou mais restaurações e ao mesmo tempo uma ou mais áreas cariadas.
4. Dente Obturado sem Cárie: Quando uma ou mais restaurações definitivas estiverem presentes sem cárie.
5. Dente Perdido Devido a Cárie: Elemento da dentição permanente/decídua foi extraído por cárie (na dentição decídua aplica-se este código apenas se a faixa etária não justifique a sua exfoliação).
6. Dente Excluído: Quando por qualquer razão o dente não possa ser examinado.

3.2. FICHA DE EXAME ODONTOLÓGICO

A ficha de levantamento de dados da OMS traz dados de relevância para o estudo e é utilizada como padrão em levantamentos realizados no mundo (OMS, 1991; OMS, 1999).

Nesta ficha padronizada mínima ainda pode ser incluído algum dado que seja relevante no diagnóstico e manutenção dos cadastros. Esta ficha pode ser avaliada no anexo I.

3.3 SOFTWARES UTILIZADOS

Para o desenvolvimento deste trabalho, de forma a tornar a pesquisa de fácil reprodução, foram utilizados dois softwares que podem ser encontrados para download, a saber:

3.3.1 Pacote Estatístico EPIINFO

O EPIINFO é um pacote desenvolvido pelo Centro para o Controle de Doenças e Prevenção (CDC), que é uma instituição ligada à OMS. O software destina-se ao uso de pesquisadores e gestores de serviços de saúde no tratamento estatístico dos dados que são obtidos em levantamentos epidemiológicos, ou para a geração de índices para a avaliação dos serviços prestados. O link para download deste software é: <http://www.cdc.gov/epiinfo/index.htm>

3.3.2 Shell See5

Foi utilizada a versão demo do Shell See5, que utiliza o algoritmo de datamining capaz de processar a base de dados e devolver ao pesquisador a listagem de regras que são então avaliadas.

É importante salientar que este mesmo shell possui uma ferramenta que, após o teste ser rodado, pode ser iniciada, processando outros casos conforme o aprendizado obtido na exploração do banco de dados. Com esta facilidade, foi possível, de forma rápida, trabalhar o estudo duplo de forma adequada.

3.4. DISTRIBUIÇÃO DA POPULAÇÃO

A população de Guaraqueçaba se distribui dentro da área geográfica (figura 5) do município (IBGE, 2001) conforme:

TABELA 1 – QUADRO COMPARATIVO DO Nº. DE HABITANTES POR MUNICÍPIO E ESTADO – ÁREA URBANA E RURAL – 1996/2000.

	POPULAÇÃO RESIDENTE					
	EM AGOSTO DE 1996	EM 01.08.2000				
		TOTAL	HOMENS	MULHERES	URBANA	RURAL
PARANÁ	9.003.804	9.558.126	4.735.030	4.823.096	7.782.005	1.776.121
GUARAQUEÇABA	8.035	8.288	4.414	3.874	2.582	5.706
GUARATUBA	31.685	27.242	13.853	13.389	23.140	4.102
CURITIBA	1.476.253	1.586.898	760.677	826.221	1.586.898	0

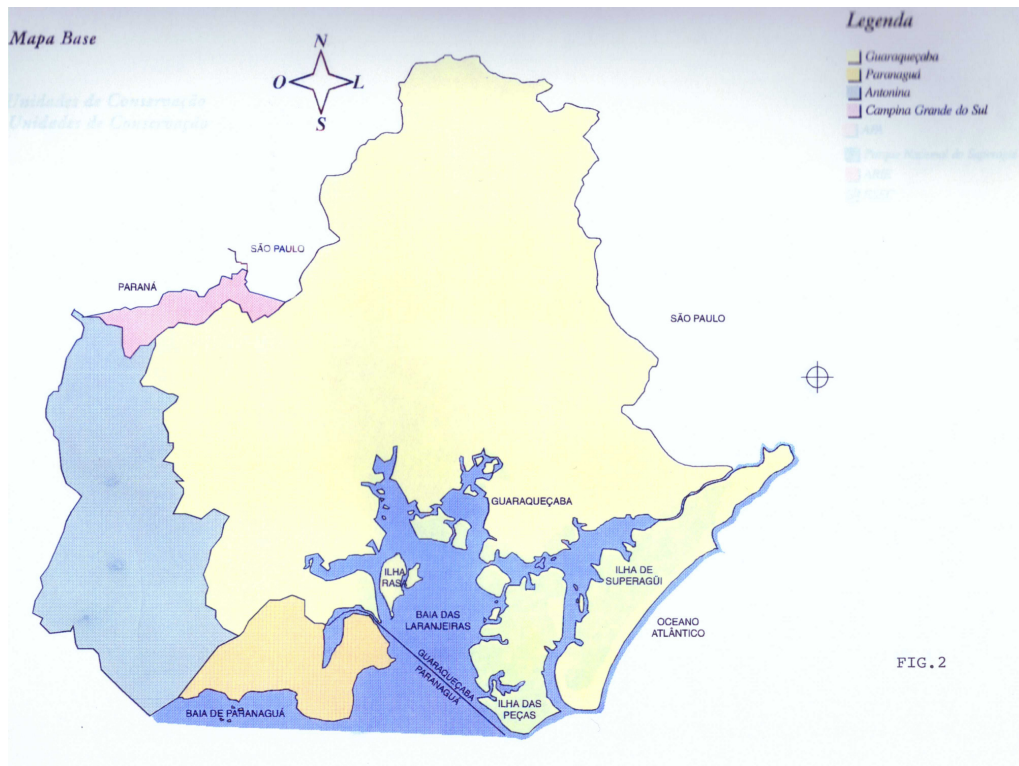


FIGURA 5 – MAPA DO MUNICÍPIO DE GUARAQUEÇABA E ADJACÊNCIAS

Para o estudo proposto, que é um estudo de prevalência, não haverá interferência sobre o normal desenvolvimento da comunidade relacionada, onde para um nível de confiança de 95%, tem-se uma amostra de 60 pessoas, de um universo total de 2400 possíveis examinados (Tabela 2). A obtenção destes dados se deu a partir da análise estatística e epidemiológica da população, segundo o módulo Statcalc do sw EPIINFO.

3.5 SOLUÇÃO PROPOSTA

Como solução para o objetivo primeiro deste trabalho, que é a predição da doença, sugere-se:

1. O uso do DataMining, para determinar a prevalência de associações pertinentes ao desenvolvimento da cárie dental na população.
2. Análise estatística inicial dos dados obtidos com o freeware EPIINFO.

3. Com a análise de diversos exames coletados na mesma população, determinar probabilisticamente quais indivíduos podem adoecer.
4. Utilizar o freeware EPIINFO para validar as regras de associação de DataMining geradas pelo software See5, com testes de representação estatística.
5. Com um teste duplo cego, verificar a pertinência dos diagnósticos encontrados.

Tabela 2 – Cálculo Amostral em Populações para Levantamentos de Prevalência.

Pesquisa em População ou Estudos Descritivos Usando uma Amostra Aleatória	
Tamanho da População (Population Size)	2400
Frequência esperada da doença (Expected frequency)	20%
Pior condição de erro (Worst Acceptable)	10%
Nível de Confiança (Confidence Level)	Tamanho Amostral (Sample Size)
80,0%	26
90,0%	43
95,0%	60
99,0%	102
99,9%	162
99,99%	220

Tem-se a predição⁷ como termo de eleição, pois tem uma relação maior com o objeto de estudo, uma vez que o termo prognóstico⁸ guarda em si a doença já instalada, contrapondo-se, ao objetivo de evitar o aparecimento de doença na comunidade.

Com estas soluções, os gestores da saúde podem alterar o seu planejamento estratégico e redefinir os métodos de atuação, usando de maneira mais racional o recurso recebido, mudando assim o perfil epidemiológico da população assistida.

⁷ - [Do lat. *praedictione*.] Substantivo feminino.. Ato ou efeito de predizer

⁸ - [Do gr. *prognostikón*, pelo lat. *prognosticu*.] Substantivo masculino.. Conjetura sobre o desenvolvimento

3.6 MODELAGEM DOS DADOS

Para que o estudo alcance o objetivo proposto, é necessário que os dados a serem cadastrados e que irão compor a base de dados para o estudo tenham pertinência entre eles e o objeto de estudo.

Os algoritmos de *DataMining* necessitam de dados que guardem uma relação com o objeto de estudo, **mas que não representem por eles mesmos o objeto**. Com isto, pode-se determinar que os dados que compõem a base de estudo não sejam todos os registrados na ficha de exame (anexo I), justificando-se, assim, o uso de alguns atributos, determinados pela epidemiologia conhecida da doença cárie.

A seguir são listados os dados que compõem o escopo da pesquisa, que se encontram divididos no cadastro do indivíduo (Tabela 3) e no cadastro de exame (Tabela 4).

TABELA 3 – Cadastro de Indivíduo:

Campo	Tipo	Função
Código*	Numérico AutoPreenchimento	Localização do indivíduo na base de dados
Nome	Alfanumérico	Localização do indivíduo na base de dados
Nome da Mãe	Alfanumérico	Diferencial na localização do indivíduo na base de dados
Data Nascimento	Numérico	Sócio demográfico
Naturalidade	Alfanumérico	Sócio demográfico
Endereço	Alfanumérico	Sócio demográfico
Localidade	Alfanumérico	Sócio demográfico

A criação de duas tabelas de atributos se explica pela necessidade da separação dos determinantes sociais (Tabela 3) dos determinantes biológicos (Tabela 6) da doença em estudo.

A partir do modelo de cadastro mostrado nas tabelas 3 e 4, pode-se então extrair os dados pertinentes ao universo da pesquisa, como é mostrado na tabela 5, que representam os atributos que compõe a multifatoriedade da doença cárie, no indivíduo ou em coletivo.

Com a definição dos atributos a serem utilizados para a determinação da presença ou não da doença, é necessária a definição do domínio de cada atributo, ou seja, o tipo de dado que será suportado para cada atributo, possibilitando o correto funcionamento do algoritmo e a utilização das regras estatísticas pelo EPIINFO. Este domínio está estabelecido na Tabela 6 e foi construído conforme a realidade da comunidade em estudo, considerando-se desde dados pessoais até índices epidemiológicos para a percepção do excesso de uso de flúor e seus suplementos.

Tabela 4 - Cadastro de Exame

Campo	Tipo	Função
Código*	Numérico	Localização do indivíduo na base de dados
Nome	Alfanumérico Autopreenchimento	Confirmação do indivíduo na base de dados
Data do exame	Alfanumérico	Dado epidemiológico temporal
Dieta de risco	Alfanumérico	Dado epidemiológico para análise de risco atual e fator de predição
Atividade de doença	Alfanumérico	Dado epidemiológico para análise de risco atual e fator de predição
Presença de placa bacteriana	Alfanumérico	Dado epidemiológico para análise de risco atual e fator de predição
Tipo de cárie	Alfanumérico	Dado epidemiológico para análise de risco atual e fator de predição
Suplemento de flúor	Alfanumérico	Dado epidemiológico para análise de risco atual e fator de predição
Gengivite	Alfanumérico	Dado epidemiológico para análise de risco atual e fator de predição
Tipo de fluorose	Alfanumérico	Dado epidemiológico para análise de risco atual
Renda familiar	Alfanumérico	Dado epidemiológico para análise de risco e fator de predição
CPO/ceo	Alfanumérico	Dado epidemiológico para análise de risco e principal fator de predição

TABELA 5 – ATRIBUTOS A SEREM UTILIZADOS NO BANCO DE DADOS

Atributo	Função
Idade	Dado epidemiológico para análise de risco atual e fator de predição
Sexo	Dado epidemiológico para análise de risco atual e fator de predição
Grupo Étnico	Dado epidemiológico para análise de risco atual e fator de predição
Dieta de risco	Dado epidemiológico para análise de risco atual e fator de predição
Presença de placa bacteriana	Dado epidemiológico para análise de risco atual e fator de predição
Suplemento de flúor	Dado epidemiológico para análise de risco atual e fator de predição
Gengivite	Dado epidemiológico para análise de risco atual e fator de predição
Tipo de fluorose	Dado epidemiológico para análise de risco atual e fator de predição
Renda familiar	Dado epidemiológico para análise de risco e fator de predição

TABELA 6 – DOMÍNIO DOS ATRIBUTOS.

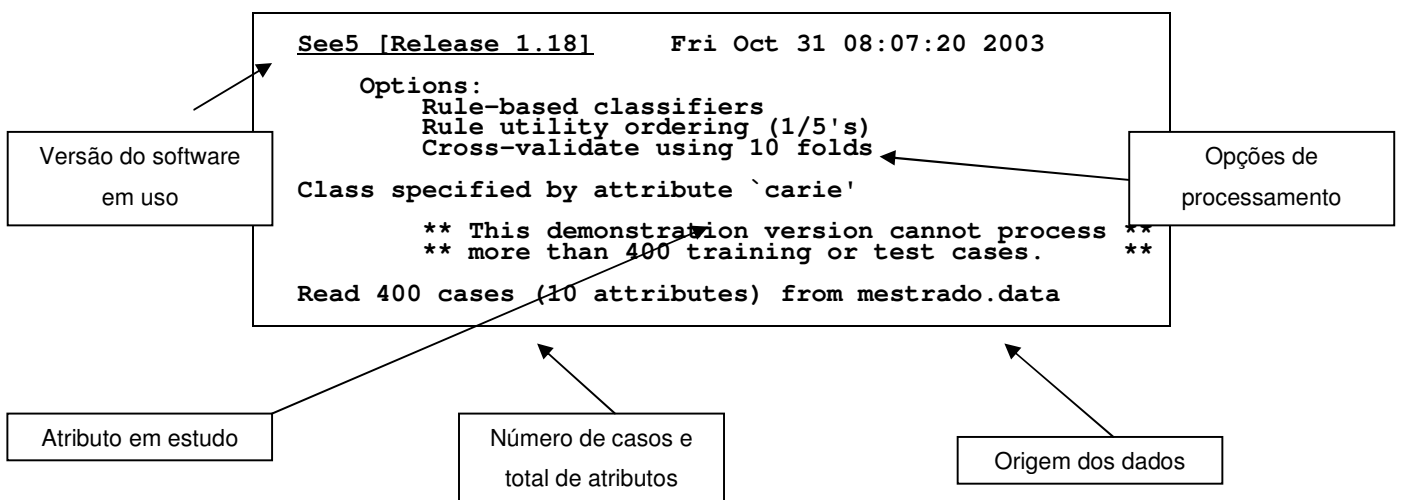
Atributo	Domínio
Idade	Contínuo
Sexo	Masculino/Feminino
Grupo Étnico	Branco/Negro/Mulato/Índio
Dieta de risco	Sim ou não
Presença de placa bacteriana - IHO	(sem placa) 0, 1, 2, 3 (muita placa)
Suplemento de flúor	Água de abastecimento Creme dental Bochechos Sem suplemento
Gengivite	Sim ou não
Tipo de fluorose	Normal Muito Leve Leve Moderada Severa
Renda familiar	Contínuo Número de salários mínimos

4. RESULTADOS

4.1 RESULTADOS GERADOS PELO See5

Os resultados gerados pelo See5 são emitidos de forma a apresentar o(s) atributo(s) que compõe a regra e a confiança desta regra, em relação ao atributo que se pretende consultar (a cárie dental, neste trabalho). Demonstra-se, para um melhor entendimento, no Quadro 1, como se compõe o cabeçalho deste relatório.

QUADRO 1 – CABEÇALHO DO RELATÓRIO DO See5



Na Tabela 9 são compiladas todas as regras geradas pelo See5, para facilitar a interpretação e posterior julgamento das regras.

Conforme o objetivo, procura-se detectar quais indivíduos expostos a fatores de riscos irão desenvolver a doença. Portanto nesta tabela encontram-se:

1. Os atributos e seus domínios, definidos nas Tabelas 7 e 8, relacionados com cada regra;
2. Os valores dos domínios de cada atributo para cada regra;
3. O resultado da regra, ou seja, a que tipo de cárie (zero – sem cárie, aguda, crônica ou tratada) se espera encontrar no paciente ao qual a regra se aplica;
4. E finalmente, a confiança da regra (Conf.), dada pela equação:

$$\text{Confiança}(\text{conf.}) = \frac{\text{número_de_registros_X_e_Y}}{\text{número_de_registros_X}}$$

onde X e Y são conjuntos de itens, $X \cap Y \neq \emptyset$.

O Gráfico 1 tem o objetivo de representar o fator de confiança de cada regra e a comparação deste fator entre as regras geradas pelo *shell*, onde se pode observar que o menor Fator de Confiança (Conf) é 0,467; o Gráfico 2 representa o resultado obtido com a aplicação da regra em relação aos estados de predição referentes à Cárie Dental.

Para que fosse possível obter a análise estatística de significância (p), foi utilizado o EPIINFO e a análise é demonstrada no Gráfico 3, incluindo aí a linha de corte da amostra (0,4).

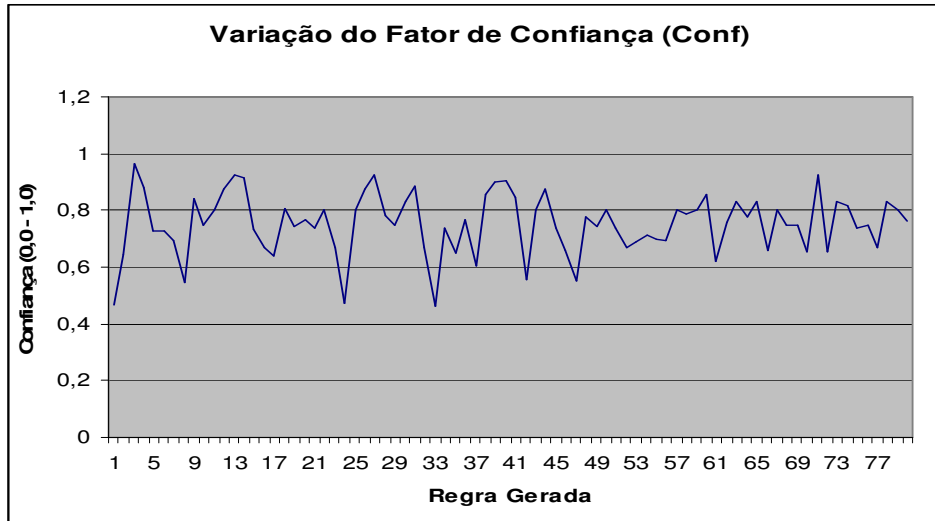


GRÁFICO 1 – REPRESENTAÇÃO DA VARIAÇÃO DO FATOR DE CONFIANÇA

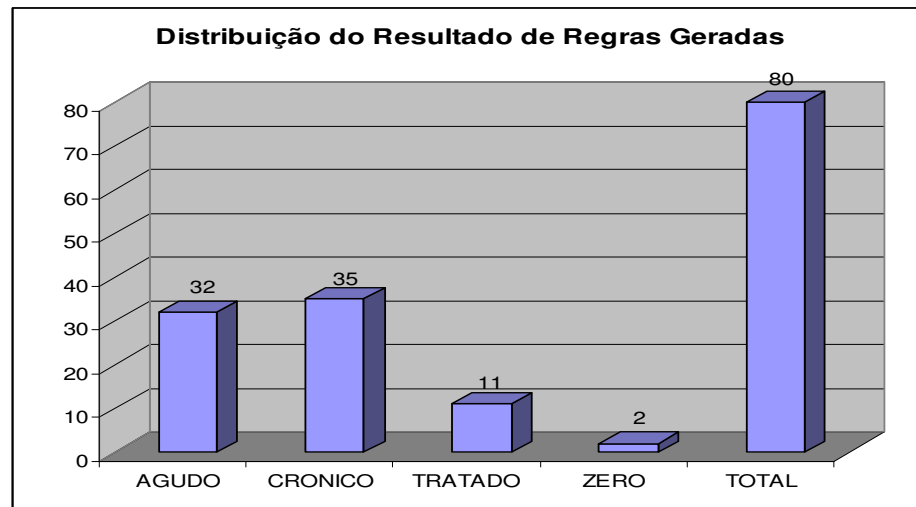


GRÁFICO 2 – DISTRIBUIÇÃO DAS REGRAS GERADAS PELO SEE5

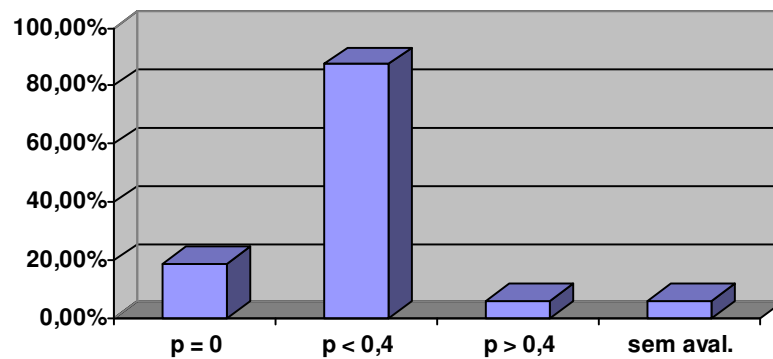


GRÁFICO 3 – SIGNIFICÂNCIA ESTATÍSTICA DAS REGRAS GERADAS PELO See5

A análise da pertinência das regras frente à aplicação do conhecimento do especialista tem também como objetivo a validação do sistema. Os testes foram realizados aplicando-se a metodologia do DUPLO CEGO, como descrito na pesquisa de FERRO, 2001.

O estudo do DUPLO CEGO foi organizado segundo os critérios apresentados abaixo, sendo o resultado resumido na Tabela 7.

1. Seleção de 50 fichas com os dados coletados dos prontuários;

2. Seleção de 02 profissionais cirurgiões-dentistas para analisarem as fichas e produzirem um diagnóstico, levando em consideração a experiência clínica de cada um. Os profissionais são caracterizados como: um deles com experiência na região de Guaraqueçaba, e no atendimento a ilhéus, com cerca de 05 anos de formado; o outro é o próprio pesquisador;
 - a. Processamento destas fichas através do módulo de predição do See5;
3. Considerou-se como terceiro especialista o próprio sistema preditivo, pelo fato do mesmo ter sido treinado, considerando-se o sucesso das técnicas de algoritmos indutivos e do C 4.5, referenciadas anteriormente.

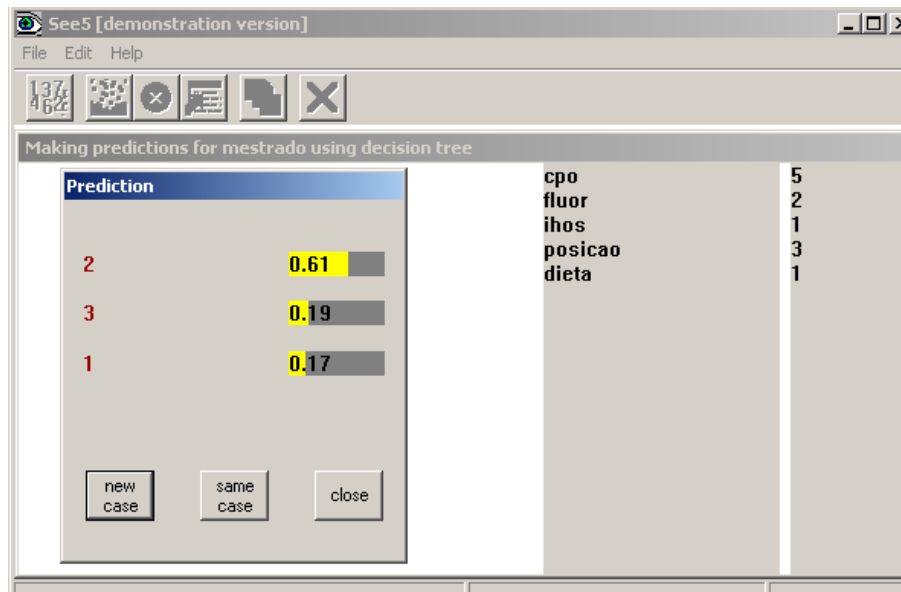


FIGURA 6 - IMAGEM DA TELA DO MÓDULO DE PREDIÇÃO DO SEE5.

TABELA 7 – TABELA COMPARATIVA DO DUPLO CEGO

	n	%	n	%
03 especialistas têm a mesma opinião	20	41,67%		
Especialista + See5 têm a mesma opinião	17	35,42%	42	87,5%
Especialista tem opinião igual à segunda escolha do See5	5	10,41%		
02 especialistas têm a mesma opinião	6	12,5%	6	12,5%
Total de casos	48	100%	48	100%

4.2 ESTUDO EPIDEMIOLÓGICO DA POPULAÇÃO

O presente estudo ainda possibilitou conhecer o perfil epidemiológico da população de Guaraqueçaba. De todas as variáveis levantadas, algumas se destacam quando relacionadas à doença cárie. Para este trabalho, são apresentados a distribuição populacional da amostra (caracterização da amostra), o índice CPO e ceo, fluorose e acesso a flúor.

4.2.1 Caracterização da Amostra

A região de Guaraqueçaba é dividida em:

1. Área Urbana
2. Área Rural
 - a. Ilhas
 - b. Continente

Com uma população de pessoas com baixa renda, sem perspectivas de emprego, ou de crescimento profissional, a própria curva de crescimento populacional de Guaraqueçaba encontra-se pouco crescente. Os indivíduos do município foram distribuídos em quatro grupos, a saber: (1) Brancos, (2) Negros, (3) Mulatos, (4) Indígenas (tabela 8).

TABELA 8 – DISTRIBUIÇÃO DA AMOSTRA EM GRUPOS ÉTNICOS

Grupo Étnico	Freqüência	Percentagem	Percentagem acumulada	
Branco	366	77,7%	77,7%	
Negro	17	3,6%	81,3%	
Mulato	71	15,1%	96,4%	
Indígena	17	3,6%	100,0%	
Total	471	100,0%	100,0%	

Como a pesquisa ocorreu levando em conta os diferentes locais de moradia da população, as Tabelas de 9 a 12, demonstram como a amostra desta população está dividida dentro do município de Guaraqueçaba. Deve-se ressaltar, ainda, que a idade média da população durante a pesquisa foi de 17,06 anos, com a distribuição por sexo sendo caracterizada pela Tabela 13.

TABELA 9 – DISTRIBUIÇÃO DA AMOSTRA EM GRUPOS ÉTNICOS – ÁREA URBANA

Grupo Étnico	Freqüência	Percentagem	Percentagem acumulada	
Branco	167	77,3%	77,3%	
Negro	6	2,8%	80,1%	
Mulato	33	15,3%	95,4%	
Indígena	10	4,6%	100,0%	
Total	216	100,0%	100,0%	

TABELA 10– DISTRIBUIÇÃO DA AMOSTRA EM GRUPOS ÉTNICOS – ÁREA ILHA PEÇAS

Grupo Étnico	Freqüência	Percentagem	Percentagem acumulada	
Branco	1	100,0%	100,0%	
Negro	0	0,0%	100,0%	
Mulato	0	0,0%	100,0%	
Indígena	0	0,0%	100,0%	
Total	1	100,0%	100,0%	

TABELA 11 – DISTRIBUIÇÃO DA AMOSTRA EM GRUPOS ÉTNICOS – ÁREA ILHA SUPERAGUI

Grupo Étnico	Frequência	Percentagem	Percentagem acumulada	
Branco	190	78,8%	78,8%	
Negro	7	2,9%	81,7%	
Mulato	37	15,4%	97,1%	
Indígena	7	2,9%	100,0%	
Total	241	100,0%	100,0%	

TABELA 12– DISTRIBUIÇÃO DA AMOSTRA EM GRUPOS ÉTNICOS – ÁREA RURAL

Grupo Étnico	Frequência	Percentagem	Percentagem acumulada	
Branco	8	61,5%	61,5%	
Negro	4	30,8%	92,3%	
Mulato	1	7,7%	100,0%	
Indígena	0	0,0%	100,0%	
Total	13	100,0%	100,0%	

TABELA 13 – DISTRIBUIÇÃO DA AMOSTRA ESTRATIFICADA POR SEXO

Sexo	Frequência	Percentagem	Percentagem acumulada	
Masculino	218	46,3%	46,3%	
Feminino	253	53,7%	100,0%	
Total	471	100,0%	100,0%	

4.2.2 Índice CPO-D e ceo-d

Como o índice caracterizado anteriormente, estes dois índices a medem média de doença cárie e sua distribuição na população. Os resultados foram estratificados nas Tabelas 14 a 15 e apresentam: (1) a média CPO-D e ceo-d da amostra, (2) a distribuição do tipo de cárie dentro da amostra.

No estudo do índice se percebeu-se que crianças com zero a seis anos têm, em média, 5,44 dentes cariados, perdidos ou obturados; as crianças com idade entre seis e doze anos têm um índice igual a 6,65; já aos 12 anos, a população apresenta um índice CPO-D de 7,03 (ver Tabela 15). É importante ainda caracterizar a amostra e compará-la com dados nacionais. O Gráfico 4 cumpre este papel e o Gráfico 5 caracteriza melhor a distribuição do tipo de cárie dentro da amostra estudada.

TABELA 14 – DISTRIBUIÇÃO DO TIPO DE CÁRIE DA AMOSTRA

Tipo de Cárie	Freqüência	Percentagem	Percentagem acumulada
Zero Cárie	28	5,9%	5,9%
Aguda	218	46,3%	52,2%
Crônica	174	36,9%	89,2%
Tratada	51	10,8%	100,0%
Total	471	100,0%	100,0%

TABELA 15 – DISTRIBUIÇÃO DA CÁRIE ESTRATIFICADA POR IDADE E SEXO

	< 6 anos	6 a 12 anos	12 anos	Amostra Geral
Masculino	5,25	7,25	7,62	9,34
Feminino	5,67	6,05	7,53	10,30
Média	5,44	6,65	7,03	9,86
N	45	234	28	471

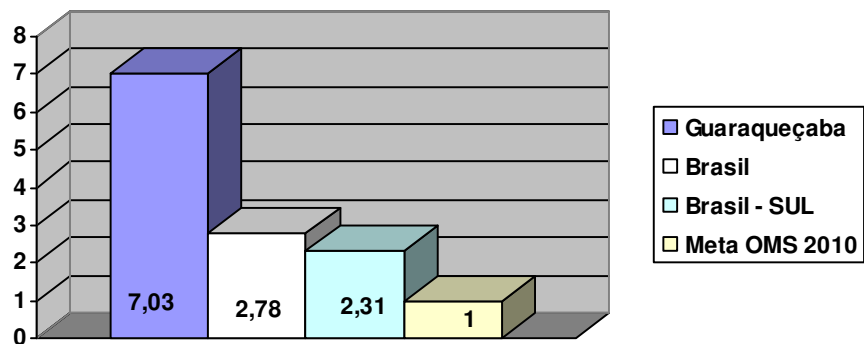


GRÁFICO 4 – COMPARAÇÃO DO CPO MÉDIO 12 ANOS

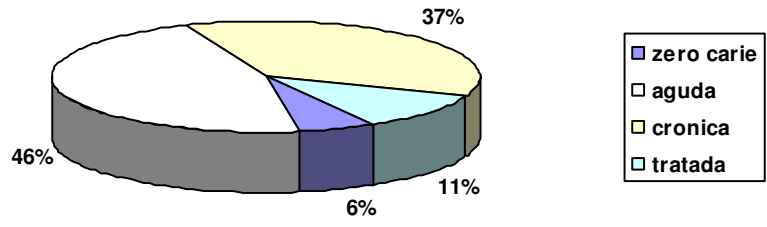


GRÁFICO 5 – DISTRIBUIÇÃO DO TIPO DE CÁRIE NA POPULAÇÃO ESTUDADA

5. DISCUSSÃO E CONCLUSÃO

5.1 DISCUSSÃO

O resultado deste trabalho é a análise das regras geradas pelo See5, considerando então a possível utilização deste modelo nos Sistemas de Atenção à Saúde em Odontologia para a melhora da Gestão e Organização dos Serviços, sendo, portanto uma melhoria na qualidade dos indicadores em Saúde Bucal da comunidade. A predição obtida está representada pelo Resultado da Regra, que pode ser: 0 – Zero cárie, 1 – Cárie Aguda, 2 – Cárie Crônica e 3 – Cárie Tratada (anexo IV).

No anexo IV, que representa as regras comparam-se as variáveis que foram estudadas, os atributos que geraram as regras de associação e a significância obtida pela análise pelo EPIINFO. Na coluna denominada *Resultado da Regra*, está representada qual a situação do indivíduo exposto às variáveis em relação à doença cárie, a saber: Zero: sem exposição e/ou experiência de cárie; Aguda: passando por um surto da doença (é necessária a intervenção do profissional de forma rápida e resolutiva); Crônica: paciente já teve a experiência de cárie, porém agora está com a mesma controlada, podendo apresentar cavidades abertas ou restaurações provisórias; Tratada: o paciente já recebeu tratamento adequado e deve fazer parte de um grupo de controle da doença. Já a coluna *CONF.* representa o fator de confiança da regra gerada, sendo que quanto mais próximo de 1,0 maior o fator de confiança para a regra. No Gráfico 1 verifica-se que o valor médio CONF para as regras geradas é de 0,753, com uma variância de 0,0116 e desvio padrão de 0,1078, representando a pertinência das regras geradas pelo sistema.

O Gráfico 2 demonstra a distribuição das regras em relação à exposição do paciente (risco do paciente) relacionada à doença cárie, conforme discutido anteriormente. Na análise do gráfico e das regras fica evidente a maior necessidade em políticas que retirem o paciente do grupo de ATIVIDADE DE DOENÇA – Cárie Ativa, incluindo-o no grupo de Crônicos ou Tratada. Esta análise se torna importante porque quanto mais regras houver que favoreçam a doença aguda maior é a necessidade e heterogeneidade da população em relação à implantação de políticas capazes de

controlar o aparecimento e incremento da doença. Esta uma análise que leva em consideração o impacto clínico da doença na vida da população.

Já a validação estatística das regras listadas no Anexo IV e geradas pelo See5 se dá pela análise da pertinência das regras frente à estatística pela aplicação da ferramenta EPIINFO sobre cada regra, resultando no Anexo IV, que utiliza o módulo Análise do software.

Esta tabela apresenta como resultado da análise o valor de significância estatística de cada regra frente a toda a base de dados que foi analisada. Ao contrário do See5, que gera as regras, a intenção agora é de verificar, estatisticamente quais regras podem efetivamente fazer parte de um conjunto válido de regras para predição.

Com esta análise percebeu-se que 18,75% das regras apresentaram uma alta significância estatística, 87,5% (71 regras de um total de 80 geradas pelo sistema) apresentaram um $p < 0,4$, sendo que apenas 6,25% apresentaram um $p \geq 0,4$, e 6,25% das regras não puderam ser avaliadas. Somadas, elas representam 12,5% equivalendo a nove regras. Esta distribuição é demonstrada no Gráfico 3.

A importância da avaliação final com profissionais odontólogos como o preconizado pelo ensaio duplo cego é para que estes possam cumprir sua função, tornando-se imperativo que o duplo cego seja bem planejado, com as hipóteses a serem testadas descritas a *priori*. Sempre que possível, deve-se garantir que os indivíduos envolvidos não tenham conhecimento do grupo que esteja recebendo a intervenção ou o controle – duplo cego (Newell, 1992). Assim feito, profissionais odontólogos analisaram as fichas dos pacientes e consideraram as mesmas variáveis analisadas pelo sistema, havendo coincidência de diagnóstico em 87,5% dos casos. Ambos os profissionais diferem da opinião do sistema em 12,5% dos casos. Chama a atenção o fato destes valores serem exatamente os mesmos das pertinências estatísticas das regras encontradas pelo EPIINFO.

Como pode ser percebido nos resultados da pesquisa, a comunidade de Guaraqueçaba possui índices de saúde bucal (doença cárie) muito aquém dos resultados obtidos pelo SB Brasil⁹ de outras cidades e, principalmente, das metas da

⁹ - Levantamento epidemiológico de saúde bucal realizado pelo MS no Brasil e com dados divulgados em 2004.

OMS para 2010. Fato que deve ser levado em consideração são as condições de vida da população. Guaraqueçaba está entre os cinco maiores bolsões de pobreza do País e isto influencia não somente na pobreza financeira como também na pobreza de saúde, educação e cultura. Ao se comparar, no Gráfico 4, a média das cáries entre a população de Guaraqueçaba e o restante do Brasil, fica evidente tal diferença. Diferença esta que pode ser modificada com a mudança do perfil de gestão municipal, utilizando-se de meios mais democráticos de acesso a serviços de saúde. Principalmente na Atenção Básica, a criação e gestão de programas especiais para comunidades da área rural (continente e ilha) são medidas simples e eficazes para tal transformação.

Existe ainda um cenário ainda mais preocupante na região. Quando se estuda a distribuição do tipo de cárie na comunidade, percebe-se uma alta prevalência de cáries agudas (46,3%) e apenas 5,9% de pessoas livres de experiência de cárie (Gráfico 5). No estudo SB Brasil de 2003 do Ministério da Saúde pode-se perceber que, na faixa etária de 12 anos há em média 31% das crianças sem experiência de cárie, em contraponto ao estudo realizado, aos 12 anos 3,6% das crianças de Guaraqueçaba estavam livres de cárie.

5.2 CONCLUSÃO

O estudo dos dados, com o auxílio do software EPIINFO, possibilitou criar os gráficos¹⁰ de forma a melhor analisar os dados obtidos pela pesquisa. Fica evidenciado que, com análise estatística das regras e o estudo duplo cego, o Método da Análise de Banco de Dados pelo See5 (DataMining) é **pertinente** para a predição de doenças, principalmente quando em conjunto com a aplicação metodológica do duplo cego, que confirma com 87,5% de igualdade as opiniões dos especialistas com o See5.

Acredita-se que o estudo cumpriu os objetivos propostos, principalmente porque foi criado um novo modelo metodológico capaz de predizer doenças em comunidades.

¹⁰ - Ver gráficos no Capítulo 4 – Materiais e Métodos.

Assim torna-se possível de agora em diante, saber qual habitante e em que casa vai adoecer. Considerações importantes são pertinentes ainda nesta discussão, em relação ao modelo estudado:

1. O modelo leva em consideração as características sócio-econômicas, demográficas e biológicas do indivíduo;
2. O modelo é, portanto, único para cada comunidade em estudo;
3. Como é possível alterar as variáveis da análise, é possível, também, aplicar o modelo estudado para as doenças de que se conheçam os fatores etiológicos (predizer doenças e risco) e para doenças de que não se conheçam estes fatores (o modelo pode definir quais são os indicadores de risco, baseado no banco de dados da comunidade). Desta forma, possibilitam-se o planejamento estratégico e a organização de forma adequada dos serviços de saúde.

É, importante salientar que o estudo nunca planejou excluir o profissional de saúde deste processo de diagnóstico, mas sim incluí-lo de forma pró-ativa em um modelo onde este profissional terá uma resposta mais efetiva do sistema. A comunidade, por outro lado, receberá uma atenção em saúde de melhor qualidade, uma vez que os recursos serão aplicados onde realmente são necessários, nos fatores pertinentes a cada morador. É importante ressaltar ainda que existe a possibilidade da retro-alimentação do sistema, quando da implementação de um software capaz de fazer esta gestão, podendo ainda incluir estes dados epidemiológicos, cruzados com outros dados do usuário disponível no banco de dados do gestor municipal.

Com a análise dos dados epidemiológicos, evidencia-se que a comunidade ainda tem acesso restrito a uma odontologia não calcada nos princípios da promoção de saúde e que o modelo assistencial existente no município deve ser revisto. Deve-se esquecer o modelo dicotômico da prática curativa e preventiva que sempre está em evidência em detrimento ao controle da doença. Com o modelo de predição de doença, as equipes de Saúde Bucal podem atuar de forma a evitar o aparecimento da doença, tratando os pacientes dos fatores pelos quais irão adoecer. Neste contexto políticas públicas de programas de bochechos, água fluoretada são de importância ímpar e se

constituem uma forma barata e democrática de acesso a serviços de controle de doença cárie.

Os profissionais da odontologia devem estar menos preocupados em tratar cárie e sim tratar a doença cárie, entendendo o ser humano como um indivíduo inserido em um grupo familiar, com anseios e condições diversas e adversas que compõem o seu contexto de vida e este, por sua vez, dita suas condições de saúde. Mais que um sistema de diagnóstico, pretendeu-se discutir a contextualização do indivíduo na família e na comunidade e qual entendimento poderia ser extraído quando analisadas tais condições frente às suas necessidades no processo saúde-doença.

Fica, portanto, explícita a necessidade de implantação de um software para a predição da doença integrando, Prontuário Eletrônico do Paciente e bases de dados do Sistema de Informação da Atenção Básica (SIAB) – que passa a ser o próximo objeto de estudo.

5.3 TRABALHOS FUTUROS

Ao considerar este trabalho como o início de uma proposta, é necessário dar continuidade ao estudo após os resultados positivos desta fase inicial. A experiência obtida durante a execução deste trabalho permite que sejam considerados alguns trabalhos futuros que podem contribuir ainda mais para o modelo estudado, com os expostos abaixo:

- Aumentar a base de dados para que o algoritmo do See5 possa gerar regras mais refinadas, evitando algum viés.
- Aplicação do modelo em outras doenças (multifatoriais) para concluir a generalidade do modelo.
- Desenvolvimento de um Prontuário Eletrônico do Paciente em Odontologia (PEPO), baseado na estratégia do PSF, podendo ainda ser usado na territorialização, predição da doença e acompanhamento da evolução dos indicadores de saúde bucal da comunidade.

ANEXO II – TELA DE ENTRADA DE DADOS EPI INFO

Enter

File Edit Options Help

1 Page

Save data
Mark record as deleted
End

New

Record
473 New Record
<< < > >>

MustEnter , Range:0 - 99

I Do Not Love You
Microsoft PowerPoint
C I A F - Profissional 2.30
Controle e tratamento ...

27 minuto(s) restante(s) Cancelar

Iniciar Proação Epi Info 2002 Epi Info 2002 Enter 12:44

Dados Pessoais e de Localização

Idade (anos) Sexo
(1-Masc/2-Fem) Grupo étnico Posição geográfica
1-Bra/2-Neg/3-Mul/4-Ind 1-Urb/2-Pes/3-Sup/4-Rur

Perfil de Acesso

Acesso Flúor Tipo de Cárie IHO-S Dieta de
0-nenhum/1-Água (0-Zero/1-Ág/2-Cr/3-Trat) Risco

2-Creme/3-Ambos

INDICADORES EPIDEMIOLÓGICOS

CPO/CEO Fluorose Detectada

PPGIA-PUCPR
PROGRAMA DE PÓS GRADUAÇÃO EM INFORMÁTICA APLICADA
" INFORMÁTICA EM SAÚDE "

Aluno: ERNESTO JOSUÉ SCHMITT
Orientador: PROF. JOÃO DA SILVA DIAS

ANEXO III – RELATÓRIO COMPLETO SEE5

See5 [Release 1.18] Fri Oct 31 08:07:20 2003

Options:
 Rule-based classifiers
 Rule utility ordering (1/5's)
 Cross-validate using 10 folds

Class specified by attribute `carie`

** This demonstration version cannot process **
 ** more than 400 training or test cases. **

Read 400 cases (10 attributes) from mestrado.data

[Fold 0]

Rules:

Rule 1: (336/179, lift 1.1)
 cpo > 0
 -> class 2 [0.467]

Rule 2: (107/38, lift 1.8)
 cpo > 0
 ihos > 1
 idade <= 14
 -> class 1 [0.642]

Rule 3: (24, lift 13.8)
 cpo <= 0
 -> class 0 [0.962]

Rule 4: (23/2, lift 6.9)
 cpo > 27
 ihos <= 0
 -> class 3 [0.880]

Rule 5: (57/15, lift 2.0)
 cpo > 0
 fluor = 0
 ihos > 0
 -> class 1 [0.729]

Rule 6: (9/2, lift 1.7)
 dieta = 0
 ihos > 1
 posicao = 1
 -> class 2 [0.727]

Rule 7: (34/10, lift 1.6)
 cpo > 0
 fluor = 2
 ihos <= 0
 idade <= 31
 -> class 2 [0.694]

Rule 8: (124/56, lift 1.5)
 cpo > 6
 cpo <= 13
 -> class 1 [0.548]

Rule 9: (23/3, lift 1.9)
 cpo <= 27
 fluor = 2
 idade > 31
 -> class 2 [0.840]

Rule 10: (18/4, lift 1.7)
 cpo > 6
 cpo <= 27
 fluor = 2
 dieta = 1
 ihos <= 1
 posicao = 3

```

-> class 2 [0.750]
Rule 11: (3, lift 6.3)
  fluor = 2
  ihos > 2
  idade > 14
  idade <= 21
-> class 3 [0.800]
Rule 12: (6, lift 2.0)
  cpo <= 27
  ihos > 2
  sexo = 2
  idade > 21
  idade <= 31
-> class 2 [0.875]
Rule 13: (11, lift 2.1)
  cpo > 0
  cpo <= 5
  fluor = 2
  dieta = 1
  posicao = 1
  idade <= 10
-> class 2 [0.923]
Rule 14: (10, lift 2.5)
  posicao = 4
-> class 1 [0.917]
Rule 15: (28/7, lift 2.0)
  dieta = 1
  ihos > 1
  ihos <= 2
  sexo = 2
  idade <= 31
-> class 1 [0.733]
Rule 16: (4/1, lift 5.2)
  cpo > 0
  cpo <= 2
  fluor = 2
  ihos > 0
  ihos <= 1
  posicao = 1
-> class 3 [0.667]

```

Default class: 2

Evaluation on hold-out data (40 cases):

```

          Rules
-----
      No      Errors
      16     13(32.5%)  <<

```

[Fold 1]

Rules:

```

Rule 1: (335/179, lift 1.1)
  cpo > 0
-> class 2 [0.466]
Rule 2: (120/43, lift 1.7)
  cpo > 0
  ihos > 1
  idade <= 19
-> class 1 [0.639]
Rule 3: (25/1, lift 13.3)
  cpo <= 0
-> class 0 [0.926]
Rule 4: (29/5, lift 6.3)
  ihos <= 0
  idade > 19
-> class 3 [0.806]

```

```

Rule 5: (58/14, lift 2.0)
  cpo > 0
  fluor = 0
  ihos > 0
  -> class 1 [0.750]

Rule 6: (41/10, lift 1.7)
  cpo > 0
  fluor = 2
  posicao = 1
  grupo = 1
  idade <= 10
  -> class 2 [0.744]

Rule 7: (11/2, lift 1.8)
  cpo > 0
  cpo <= 4
  dieta = 0
  ihos > 1
  -> class 2 [0.769]

Rule 8: (51/13, lift 2.0)
  cpo > 7
  ihos > 0
  posicao = 3
  idade <= 19
  -> class 1 [0.736]

Rule 9: (3, lift 1.8)
  fluor = 0
  dieta = 0
  ihos > 1
  -> class 2 [0.800]

Rule 10: (25/8, lift 1.5)
  cpo <= 27
  fluor = 2
  dieta = 1
  idade > 14
  -> class 2 [0.667]

Rule 11: (17/9, lift 3.7)
  cpo > 0
  dieta = 1
  ihos <= 1
  posicao = 3
  sexo = 2
  -> class 3 [0.474]

Rule 12: (13/2, lift 1.8)
  cpo > 7
  fluor = 2
  dieta = 1
  ihos <= 1
  idade <= 19
  -> class 2 [0.800]

Rule 13: (6, lift 2.4)
  ihos > 1
  posicao = 1
  grupo = 1
  sexo = 2
  idade > 8
  idade <= 10
  -> class 1 [0.875]

Rule 14: (11, lift 2.5)
  posicao = 4
  -> class 1 [0.923]

Rule 15: (53/11, lift 2.1)
  cpo > 0
  cpo <= 27
  fluor = 0
  dieta = 1
  -> class 1 [0.782]

Rule 16: (10/2, lift 1.7)
  fluor = 2
  ihos > 1
  grupo = 3

```

```

idade > 10
-> class 2 [0.750]
Rule 17: (4, lift 2.3)
  ihos > 1
  grupo = 3
  idade > 12
  idade <= 19
-> class 1 [0.833]
Rule 18: (24/2, lift 2.0)
  cpo > 17
  cpo <= 27
  fluor = 2
  idade > 19
-> class 2 [0.885]
Rule 19: (4/1, lift 5.2)
  cpo > 14
  cpo <= 17
  fluor = 2
  idade > 19
-> class 3 [0.667]
Rule 20: (204/110, lift 1.3)
  dieta = 1
-> class 1 [0.461]
Default class: 2
Evaluation on hold-out data (40 cases):

```

Rules		
No	Errors	
20	15 (37.5%)	<<

[Fold 2]

```

Rules:
Rule 1: (335/179, lift 1.1)
  cpo > 0
-> class 2 [0.466]
Rule 2: (55/14, lift 2.0)
  cpo > 0
  fluor = 0
  ihos > 0
-> class 1 [0.737]
Rule 3: (25/1, lift 13.3)
  cpo <= 0
-> class 0 [0.926]
Rule 4: (24/2, lift 6.9)
  cpo > 27
  ihos <= 0
-> class 3 [0.885]
Rule 5: (52/18, lift 1.8)
  cpo > 0
  ihos > 1
  posicao = 3
  idade <= 10
-> class 1 [0.648]
Rule 6: (11/2, lift 2.1)
  cpo > 0
  ihos > 2
  sexo = 1
  idade > 10
  idade <= 31
-> class 1 [0.769]
Rule 7: (41/16, lift 1.6)
  ihos > 1
  ihos <= 2
  sexo = 2

```



```

idade <= 31
-> class 1 [0.605]

Rule 8: (5, lift 2.3)
cpo > 6
ihos <= 1
posicao = 3
sexo = 2
idade > 14
idade <= 19
-> class 1 [0.857]

Rule 9: (8, lift 2.1)
fluor = 2
ihos > 1
posicao = 1
idade <= 8
-> class 2 [0.900]

Rule 10: (29/2, lift 2.5)
cpo > 7
cpo <= 27
fluor = 0
-> class 1 [0.903]

Rule 11: (11/1, lift 2.3)
ihos > 1
posicao = 1
idade > 10
idade <= 14
-> class 1 [0.846]

Rule 12: (7/3, lift 4.3)
cpo > 0
cpo <= 2
fluor = 2
ihos <= 1
posicao = 1
-> class 3 [0.556]

Rule 13: (3, lift 6.3)
cpo > 0
cpo <= 6
dieta = 1
ihos <= 1
posicao = 3
sexo = 2
idade <= 15
-> class 3 [0.800]

Rule 14: (9, lift 2.5)
posicao = 4
-> class 1 [0.909]

Rule 15: (7/1, lift 2.1)
ihos > 1
posicao = 1
sexo = 2
idade > 8
idade <= 10
-> class 1 [0.778]

Rule 16: (6, lift 2.0)
cpo <= 6
ihos <= 1
idade > 15
-> class 2 [0.875]

```

Default class: 2

Evaluation on hold-out data (40 cases):

Rules		
No	Errors	
16	14 (35.0%)	<<

[Fold 3]

Rules:

- Rule 1: (335/179, lift 1.1)
 cpo > 0
 -> class 2 [0.466]
- Rule 2: (82/21, lift 2.0)
 cpo > 4
 ihos > 1
 idade <= 14
 -> class 1 [0.738]
- Rule 3: (25/1, lift 13.3)
 cpo <= 0
 -> class 0 [0.926]
- Rule 4: (25/2, lift 7.0)
 cpo > 27
 ihos <= 0
 -> class 3 [0.889]
- Rule 5: (39/13, lift 1.5)
 cpo <= 27
 fluor = 2
 ihos > 1
 idade > 14
 -> class 2 [0.659]
- Rule 6: (131/59, lift 1.5)
 cpo > 7
 cpo <= 27
 ihos > 0
 -> class 1 [0.549]
- Rule 7: (16/3, lift 1.8)
 cpo > 7
 cpo <= 27
 fluor = 2
 dieta = 1
 ihos <= 1
 posicao = 3
 -> class 2 [0.778]
- Rule 8: (60/15, lift 2.0)
 cpo > 0
 cpo <= 27
 fluor = 0
 -> class 1 [0.742]
- Rule 9: (3, lift 2.2)
 cpo > 27
 fluor = 0
 ihos > 0
 -> class 1 [0.800]
- Rule 10: (44/11, lift 1.7)
 cpo > 2
 cpo <= 27
 fluor = 2
 ihos <= 1
 posicao = 1
 -> class 2 [0.739]
- Rule 11: (4/1, lift 5.2)
 cpo > 0
 cpo <= 2
 fluor = 2
 ihos > 0
 ihos <= 1
 posicao = 1
 -> class 3 [0.667]
- Rule 12: (10, lift 2.5)
 posicao = 4
 -> class 1 [0.917]
- Rule 13: (30/9, lift 1.6)
 cpo > 0
 cpo <= 7
 ihos <= 1
 posicao = 3

-> class 2 [0.688]

Default class: 2

Evaluation on hold-out data (40 cases):

Rules		
No	Errors	
13	17(42.5%)	<<

[Fold 4]

Rules:

Rule 1: (335/203, lift 1.1)
 cpo > 0
 -> class 1 [0.395]

Rule 2: (42/9, lift 1.8)
 cpo > 2
 cpo <= 27
 fluor = 2
 ihos <= 1
 posicao = 1
 -> class 2 [0.773]

Rule 3: (25/1, lift 13.3)
 cpo <= 0
 -> class 0 [0.926]

Rule 4: (34/10, lift 5.4)
 cpo > 27
 -> class 3 [0.694]

Rule 5: (33/9, lift 1.6)
 cpo <= 27
 fluor = 2
 ihos > 1
 idade > 19
 -> class 2 [0.714]

Rule 6: (28/8, lift 1.6)
 cpo > 0
 cpo <= 6
 ihos <= 1
 posicao = 3
 -> class 2 [0.700]

Rule 7: (53/16, lift 1.6)
 cpo > 0
 fluor = 2
 dieta = 1
 ihos <= 2
 idade <= 10
 -> class 2 [0.691]

Rule 8: (7, lift 2.0)
 cpo > 0
 cpo <= 6
 dieta = 0
 ihos > 2
 -> class 2 [0.889]

Rule 9: (3, lift 1.8)
 fluor = 0
 dieta = 0
 ihos > 1
 -> class 2 [0.800]

Rule 10: (12/2, lift 1.8)
 cpo > 11
 fluor = 2
 dieta = 1
 ihos <= 2
 idade <= 19
 -> class 2 [0.786]

Rule 11: (4/1, lift 5.2)

```

      cpo > 0
      cpo <= 2
      fluor = 2
      ihos > 0
      ihos <= 1
      posicao = 1
      -> class 3 [0.667]
Rule 12: (3, lift 1.8)
      fluor = 2
      ihos > 1
      grupo = 2
      -> class 2 [0.800]
Rule 13: (5, lift 2.0)
      dieta = 0
      ihos > 1
      ihos <= 2
      grupo = 3
      -> class 2 [0.857]
Rule 14: (12, lift 2.5)
      posicao = 4
      -> class 1 [0.929]
Default class: 2
Evaluation on hold-out data (40 cases):

```

Rules		
No	Errors	
14	18 (45.0%)	<<

[Fold 5]

```

Rules:
Rule 1: (334/179, lift 1.1)
      cpo > 0
      -> class 2 [0.464]
Rule 2: (119/45, lift 1.7)
      ihos > 1
      idade <= 19
      -> class 1 [0.620]
Rule 3: (26/1, lift 13.4)
      cpo <= 0
      -> class 0 [0.929]
Rule 4: (23/1, lift 7.2)
      cpo > 27
      ihos <= 0
      -> class 3 [0.920]
Rule 5: (56/13, lift 2.1)
      cpo > 0
      cpo <= 27
      fluor = 0
      dieta = 1
      -> class 1 [0.759]
Rule 6: (10/1, lift 1.9)
      fluor = 2
      ihos > 1
      ihos <= 2
      sexo = 2
      idade <= 10
      -> class 2 [0.833]
Rule 7: (7, lift 2.1)
      cpo > 0
      cpo <= 6
      dieta = 0
      ihos > 2
      -> class 2 [0.889]
Rule 8: (7/1, lift 2.1)

```

```

      cpo > 7
      dieta = 0
      ihos <= 1
      posicao = 3
      idade <= 19
      -> class 1 [0.778]

Rule 9: (4, lift 6.5)
      cpo > 1
      cpo <= 2
      fluor = 2
      ihos > 0
      ihos <= 1
      -> class 3 [0.833]

Rule 10: (3, lift 2.2)
      cpo > 27
      fluor = 0
      ihos > 0
      -> class 1 [0.800]

Rule 11: (39/13, lift 1.5)
      cpo > 0
      fluor = 2
      dieta = 1
      ihos <= 2
      sexo = 1
      -> class 2 [0.659]

Rule 12: (4/1, lift 5.2)
      cpo > 14
      cpo <= 17
      fluor = 2
      idade > 19
      -> class 3 [0.667]

Rule 13: (8/1, lift 1.8)
      cpo > 7
      dieta = 0
      posicao = 1
      -> class 2 [0.800]

Rule 14: (2, lift 5.9)
      fluor = 2
      ihos > 2
      idade > 14
      idade <= 19
      -> class 3 [0.750]

Rule 15: (11, lift 2.5)
      posicao = 4
      -> class 1 [0.923]

Rule 16: (18/4, lift 1.7)
      cpo > 11
      fluor = 2
      ihos <= 2
      idade <= 19
      -> class 2 [0.750]

```

Default class: 2

Evaluation on hold-out data (40 cases):

```

          Rules
-----
      No      Errors
      16      16 (40.0%)  <<

```

[Fold 6]

Rules:

```

Rule 1: (335/180, lift 1.1)
      cpo > 0
      -> class 2 [0.463]

Rule 2: (53/10, lift 2.2)
      cpo > 0

```

```

      cpo <= 27
      fluor = 0
      dieta = 1
      -> class 1 [0.800]
Rule 3: (25/1, lift 13.3)
      cpo <= 0
      -> class 0 [0.926]
Rule 4: (23/2, lift 6.9)
      cpo > 27
      ihos <= 0
      -> class 3 [0.880]
Rule 5: (39/11, lift 1.9)
      cpo > 0
      dieta = 1
      ihos > 2
      idade <= 14
      -> class 1 [0.707]
Rule 6: (24/8, lift 1.8)
      cpo > 6
      dieta = 0
      posicao = 3
      idade <= 16
      -> class 1 [0.654]
Rule 7: (28/4, lift 2.3)
      cpo > 4
      dieta = 1
      ihos > 1
      ihos <= 2
      idade <= 14
      -> class 1 [0.833]
Rule 8: (4/1, lift 5.2)
      cpo > 6
      dieta = 0
      ihos <= 1
      idade > 16
      idade <= 25
      -> class 3 [0.667]
Rule 9: (12, lift 2.5)
      posicao = 4
      -> class 1 [0.929]
Rule 10: (4/1, lift 5.2)
      cpo > 0
      cpo <= 2
      fluor = 2
      ihos > 0
      ihos <= 1
      posicao = 1
      -> class 3 [0.667]
Rule 11: (9/1, lift 2.2)
      cpo > 2
      fluor = 2
      dieta = 0
      ihos > 1
      posicao = 3
      sexo = 1
      idade <= 14
      -> class 1 [0.818]

```

Default class: 2

Evaluation on hold-out data (40 cases):

```

      Rules
-----
      No      Errors
      11     14(35.0%)  <<

```

[Fold 7]

Rules:

```

Rule 1: (335/180, lift 1.1)
      cpo > 0
      -> class 2 [0.463]

Rule 2: (85/22, lift 2.0)
      cpo > 4
      ihos > 1
      idade <= 14
      -> class 1 [0.736]

Rule 3: (25/1, lift 13.3)
      cpo <= 0
      -> class 0 [0.926]

Rule 4: (23/2, lift 6.9)
      cpo > 27
      ihos <= 0
      -> class 3 [0.880]

Rule 5: (55/13, lift 2.0)
      cpo > 0
      cpo <= 27
      fluor = 0
      dieta = 1
      -> class 1 [0.754]

Rule 6: (3, lift 2.2)
      cpo > 27
      fluor = 0
      ihos > 0
      -> class 1 [0.800]

Rule 7: (9, lift 2.5)
      posicao = 4
      -> class 1 [0.909]

```

Default class: 2

Evaluation on hold-out data (40 cases):

```

----- Rules -----
      No      Errors
      7      12(30.0%)  <<

```

[Fold 8]

Rules:

```

Rule 1: (334/202, lift 1.1)
      cpo > 0
      -> class 1 [0.396]

Rule 2: (26/1, lift 12.9)
      cpo <= 0
      -> class 0 [0.929]

Rule 3: (46/13, lift 1.6)
      cpo > 2
      cpo <= 5
      fluor = 2
      -> class 2 [0.708]

Rule 4: (41/9, lift 1.8)
      cpo > 5
      cpo <= 27
      fluor = 2
      ihos > 0
      posicao = 3
      idade > 21
      -> class 2 [0.767]

Rule 5: (37/11, lift 5.4)
      cpo > 27
      -> class 3 [0.692]

Rule 6: (39/12, lift 1.6)
      cpo > 0

```

```

    fluor = 2
    dieta = 1
    ihos <= 2
    idade <= 9
    -> class 2 [0.683]
Rule 7: (27/7, lift 1.7)
    cpo > 5
    fluor = 2
    ihos <= 2
    sexo = 2
    idade <= 14
    -> class 2 [0.724]
Rule 8: (20/4, lift 1.8)
    cpo > 5
    cpo <= 27
    fluor = 2
    ihos <= 0
    -> class 2 [0.773]
Rule 9: (3, lift 2.2)
    cpo > 27
    fluor = 0
    ihos > 0
    -> class 1 [0.800]
Rule 10: (25/5, lift 1.8)
    cpo > 0
    cpo <= 5
    fluor = 2
    sexo = 1
    -> class 2 [0.778]
Rule 11: (3, lift 6.3)
    fluor = 2
    ihos > 2
    idade > 14
    idade <= 21
    -> class 3 [0.800]
Rule 12: (2, lift 1.7)
    fluor = 3
    dieta = 0
    -> class 2 [0.750]
Rule 13: (1, lift 1.5)
    fluor = 1
    -> class 2 [0.667]
Rule 14: (9/4, lift 4.3)
    posicao = 1
    idade > 14
    -> class 3 [0.545]
Default class: 2
Evaluation on hold-out data (40 cases):

```

```

----- Rules -----
      No      Errors
      14      17(42.5%)  <<

```

[Fold 9]

```

Rules:
Rule 1: (334/178, lift 1.1)
    cpo > 0
    -> class 2 [0.467]
Rule 2: (86/23, lift 2.0)
    cpo > 4
    ihos > 1
    idade <= 14
    -> class 1 [0.727]
Rule 3: (26/1, lift 12.9)

```



```

      cpo <= 0
      -> class 0 [0.929]
Rule 4: (22/2, lift 7.0)
      cpo > 27
      ihos <= 1
      -> class 3 [0.875]
Rule 5: (49/10, lift 2.1)
      cpo > 0
      cpo <= 27
      fluor = 0
      dieta = 1
      -> class 1 [0.784]
Rule 6: (29/9, lift 1.8)
      cpo > 5
      fluor = 2
      dieta = 0
      posicao = 3
      idade <= 19
      -> class 1 [0.677]
Rule 7: (14/2, lift 1.9)
      cpo > 4
      dieta = 0
      posicao = 1
      -> class 2 [0.813]
Rule 8: (11/2, lift 2.1)
      cpo > 13
      cpo <= 18
      fluor = 2
      ihos > 1
      posicao = 3
      grupo = 1
      -> class 1 [0.769]
Rule 9: (3, lift 6.4)
      cpo > 0
      cpo <= 5
      dieta = 1
      ihos <= 1
      posicao = 3
      sexo = 2
      idade <= 19
      -> class 3 [0.800]
Rule 10: (12, lift 2.5)
      posicao = 4
      -> class 1 [0.929]
Rule 11: (21/4, lift 1.8)
      cpo <= 13
      idade > 21
      -> class 2 [0.783]
Rule 12: (4/1, lift 5.3)
      cpo > 0
      cpo <= 2
      fluor = 2
      ihos > 0
      ihos <= 1
      posicao = 1
      -> class 3 [0.667]
Rule 13: (4, lift 2.3)
      cpo > 18
      idade <= 23
      -> class 1 [0.833]
Rule 14: (10/2, lift 6.0)
      cpo > 27
      fluor = 2
      sexo = 2
      -> class 3 [0.750]
Rule 15: (3/1, lift 4.8)
      ihos > 1
      posicao = 1
      idade > 14

```

```

-> class 3 [0.600]
Rule 16: (33/6, lift 2.2)
  cpo > 0
  fluor = 0
  ihos > 1
-> class 1 [0.800]
Rule 17: (26/9, lift 1.5)
  cpo > 0
  cpo <= 4
  ihos > 1
-> class 2 [0.643]
Rule 18: (44/10, lift 1.7)
  cpo > 2
  fluor = 2
  ihos <= 1
  posicao = 1
-> class 2 [0.761]

```

Default class: 2

Evaluation on hold-out data (40 cases):

Rules		
No	Errors	
18	15 (37.5%)	<<

[Summary]

Fold	Rules	
	No	Errors
0	16	32.5%
1	20	37.5%
2	16	35.0%
3	13	42.5%
4	14	45.0%
5	16	40.0%
6	11	35.0%
7	7	30.0%
8	14	42.5%
9	18	37.5%
Mean	14.5	37.8%
SE	1.2	1.5%

(a)	(b)	(c)	(d)	<-classified as
27			1	(a): class 0
	87	56	4	(b): class 1
1	59	109	5	(c): class 2
	3	22	26	(d): class 3

Time: 0.8 secs

ANEXO IV – RESUMO DAS REGRAS GERADAS (See5) E ANÁLISE ESTATÍSTICA (EPIINFO).

Regra Número	Regras de Associação Geradas pelo See5										Análise estatística EPIINFO
	Variáveis/Atributos									Resultados	
Idade	Sexo	Grupo Étnico	Local de Moradia	Acesso a Flúor	IHO-S	Dieta de Risco	CPO ou ceo	Fluorose	Resultado da Regra ¹¹	Conf.	
1							> 0		Crônica	0,467	0,0
2	≤14						> 1		Aguda	0,642	1,605
3							≤0		Zero	0,962	0,0
4							≤0		Tratado	0,880	1,148
5					Sem		>0		Aguda	0,729	0,3291
6				Urbana			>1	Não	Crônica	0,727	X
7	≤31				Creme		≤0		Crônica	0,694	0,3134
8									Aguda	0,548	0,1605
9	>31				Creme				Crônica	0,840	0,2231
10				Superagüi	Creme		≤1	Sim	Crônica	0,750	0,0435
11	>14≤21				Creme		>2		Tratado	0,800	0,0022
12	>21≤31	Feminino					>2		Crônica	0,875	0,1573

¹¹ Tipo da cárie a qual o indivíduo tem, ou poderá desenvolver.

13	≤10			Urbana	Creme		Sim	>0≤5		Crônica	0,923	0,2650
14				Rural						Aguda	0,917	0,0
15	≤31	Feminino				>1≤2	Sim			Aguda	0,733	0,1270
16				Urbana	Creme	>0≤1		>0≤2		Tratado	0,667	0,0022
17	≤19					>1		>0		Aguda	0,639	0,1605
18	>19					≤0				Tratado	0,806	0,0824
19	≤10		Branco	Urbana	Creme			>0		Crônica	0,744	0,3530
20						>1	Não	>0≤4		Crônica	0,769	0,4459
21	≤19			Superagüi		>0		>7		Aguda	0,736	0,1573
22					Sem	>1	Não			Crônica	0,800	X
23	>14				Creme		Sim	≤27		Crônica	0,667	0,3679
24		Feminino		Superagüi		≤1	Sim	>0		Tratado	0,474	0,0
25	≤19				Creme	≤1	Sim	>7		Crônica	0,800	0,0833
26	>8≤10	Feminino	Branco	Urbana		>1				Aguda	0,875	0,0
27				Rural						Aguda	0,923	X
28					Sem		Sim	>0≤27		Aguda	0,782	0,0
29	>10		Mulato		Creme	>1				Crônica	0,750	0,0227
30	>19				Creme	>1		≤27		Crônica	0,833	0,0001
31	>19				Creme			>17≤27		Crônica	0,885	0,0114
32	>19				Creme			>14≤17		Tratado	0,667	X
33							Sim			Aguda	0,461	0,0
34					Creme	>0		>0		Aguda	0,737	0,0342
35	≤10			Superagüi		>1		>0		Aguda	0,648	0,0001

36	>10≤31	Masculino				>2		>0		Aguda	0,769	0,0461
37	≤31	Feminino				>1≤2				Aguda	0,605	0,0000
38	>14≤19	Feminino		Superagüi		≤1		>6		Aguda	0,857	0,0389
39	≤8			Urbana	Creme	>1				Crônica	0,900	0,4351
40					Sem			>7≤27		Aguda	0,903	0,0
41	>10≤14			Urbana		>1				Aguda	0,846	0,0006
42				Urbana	Creme	≤1		>0≤2		Tratado	0,556	0,3263
43	≤15	Feminino		Superagüi		≤1	Sim	>0≤6		Tratado	0,800	0,1991
44	>15					≤1		≤6		Crônica	0,875	0,1573
45	≤14					>1		>4		Aguda	0,738	0,2893
46	>14				Creme	>1		≤27		Crônica	0,659	0,1573
47						>0		>7≤27		Aguda	0,549	0,1666
48				Superagüi	Creme	≤1	Sim	>7≤27		Crônica	0,778	0,0785
49					Sem			>0≤27		Aguda	0,742	0,0767
50					Sem	>0		>27		Aguda	0,800	0,0006
51				Urbana	Creme	≤1		>2≤27		Crônica	0,739	0,0
52				Urbana	Creme	>0≤1		>0≤2		Tratado	0,667	0,3263
53				Superagüi		≤1		>0≤7		Crônica	0,688	0,0180
54	>19				Creme	>1		≤27		Crônica	0,714	0,0001
55				Superagüi		≤1		>0≤6		Crônica	0,700	0,0665
56	≤10				Creme	≤2	Sim	>0		Crônica	0,691	0,3903
57					Sem	>1	Não			Crônica	0,800	0,0000
58	≤19				Creme	≤2	Sim	>11		Crônica	0,786	0,0715

59			Negro		Creme	>1				Crônica	0,800	0,0347
60			Mulato			>1≤2	Não			Crônica	0,857	0,00004
61	≤19					>1				Aguda	0,620	0,0
62					Sem		Sim	>0≤27		Aguda	0,759	0,0002
63	≤10	Feminino			Creme	>1≤2				Crônica	0,833	0,0000
64	≤19			Superagüi		≤1	Não	>7		Aguda	0,778	0,6127
65					Creme	>0≤1		>1≤2		Tratada	0,833	0,0307
66		Masculino			Creme	≤2	Sim	>0		Crônica	0,659	0,0124
67				Urbana			Não	>7		Crônica	0,800	0,0
68	≤19				Creme	≤2		>11		Crônica	0,750	0,0229
69	>14≤19				Creme	>2				Tratada	0,750	0,0289
70	≤14					>2	Sim	>0		Aguda	0,654	0,0008
71								≤0		Zero	0,926	0,0
72	≤16			Superagüi			Não	>6		Aguda	0,654	0,6315
73	>16≤25					≤1	Sim	>4		Aguda	0,833	0,2520
74	≤14	Masculino		Superagüi	Creme	>1	Não	>2		Aguda	0,818	0,413272
75	≤14					>1		>4		Aguda	0,736	0,8204
76					Ambos		Não			Crônica	0,750	0,0302
77					Água					Crônica	0,667	0,1991
78	≤23							>18		Aguda	0,833	0,0
79					Sem	>1		>0		Aguda	0,800	0,0
80				Urbana		≤1				Crônica	0,761	0,0

7. REFERÊNCIAS BIBLIOGRÁFICAS

- Pinto VG. Epidemiologia das Doenças Bucais no Brasil. In: *Promoção de Saúde Bucal*, Anonymous 1997. pp. 27-42.
- Moysés SJ. Conceito de Promoção de Saúde na Construção de Sistemas de Atenção em Saúde Bucal. In: *Promoção de Saúde Bucal*, 1997. pp. 371-408.
- Bratthal D. Caries, views and perspectives *Scand.J.Dent.Res.*, vol. 100(1), pp. 47-51, 1992.
- Chaves MM. *Odontologia Social*, 1986.
- Kruger L, Moysés ST. Filosofia da Promoção de Saúde na Clínica Privada. In: *Promoção de Saúde Bucal*, Anonymous 1997. pp. 409-432.
- Rouquayrol MZ, Filho NA. *Epidemiologia e Saúde*, 1999.
- Maltz M, Carvalho J. Diagnóstico da Doença Cárie. In: *Promoção de Saúde Bucal*, Anonymous 1997. pp. 69-92.
- Bezerra ACB, Toledo OA. Nutrição, Dieta e Cárie. In: *Promoção de Saúde Bucal*, Anonymous 1997. pp. 43-68.
- Newbrum E. *Cariologia*, Sao Paulo: Santos, 1988. pp. -326
- Bratthal D, Pakhomov G. Dental Caries and Prevention. 2001.
- Ministério da Saúde - Brasil. *Levantamento Epidemiológico em Saúde Bucal, Zona Urbana - 1986*, Brasília: Divisão Nacional de Saúde Bucal, 1988.
- Baratieri, LN. *Dentística: Procedimentos Preventivos e Restauradores*, Sao Paulo: Editora Santos; Quintessence Books, 1995. pp. -509
- Blinkhorn A, Attwood D, Gavin G, O'hickey S. Joint Epidemiological Dental Health of 12 years old school children in Dublin and Glasgow *Community Dent Oral Epidemiology*, vol. 20(5), pp. 307-308, Oct, 2001.
- Reddy J., The WHO Oral Health Goals for The Year 2000 in South of Africa *Int.Dent.J.*, vol. 42(3), pp. 150-156, 1992.
- Pitt N. The Scottish Health Boards Dental Epidemiological Programme: Initial Surveys of 05 and 12 Years Old *Br.Dent.J.*, vol. 172(11), pp. 408-413, 1992.

Organização Mundial de Saúde. *Levantamento Epidemiológico Básico em Saúde Bucal - Manual de Instrução*, Genebra:1991.

Organização Mundial de Saúde. *Levantamentos Básicos em Saúde Bucal*, Sao Paulo: Santos, 1999.

Bratthal D. *Dental Caries and Prevention*. 2001.

IBGE, <http://www.ibge.gov.br> . 01-04-2001.

Glover F. Tabu Search and Adaptive memory programming: Advances, applications and challenges, *Interfaces in Computer Science and Oper., Res.*, Kluwer Academic Publ., pp: 1-75. (1996).

Radcliffe NJ. The Algebra of Genetic Algorithms, In *Annals of Mathematics and Artificial Intelligence* 10, pp: 339-384. (1994)

Monard M C, Batista GEAPA, Kawamoto S, Pugliesi, JB. “Uma Introdução ao Aprendizado Simbólico de Máquina por Exemplos”. Notas do ICMC, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, Outubro 1997.

Ferro M, Lee HD. “O Processo de KDD – Knowledge Discovery in Database para Aplicações na Medicina” . SEMINC 2001 p. 57-62

Data Mining Stronger computer tools allow deeper analysis of medical research, patient care and insurance data. http://www.healthcare-informatics.com/issues/2004/04_04/hagland.htm. Mark Hagland, April 2004 Visualizado em 10/11/2004.

Jenses S. Mining Medical Data for Predictive and Sequential Patterns: PKDD 2001 , SPSS UK Ltd.

Bodstein R. Primary care in the agenda of public health sector in Brazil. *Ciênc. saúde coletiva*, 2002, vol.7, no.3, p.401-412. ISSN 1413-8123.

Ministério da Saúde - Brasil. Projeto SB Brasil 2003 -Condições de saúde bucal da população brasileira 2002-2003 Resultados Principais. Brasília – Coordenação Nacional de Saúde Bucal – 2004.

Gomes PR, Costa SC, Cypriano, S et al. Dental caries in Paulínia, São Paulo State, Brazil, and WHO goals for 2000 and 2010. *Cad. Saúde Pública*, May/June 2004, vol.20, no.3, p.866-870. ISSN 0102-311X.

Carvalho LAV – A mineração de dados no marketing, Medicina, Economia, Engenharia e Administração. São Paulo: Érica, 2001.

<http://www.integra.com.br>, visitado em 08.02.05, 15:58 horas.

<http://www.inf.ufpr.br/~aurora/tutoriais/arvoresdecisao/IAcap01.html>, visitado em 08.02.05, 17:11 horas.

Newell DJ. Intention-to-treat analysis: implications for quantitative and qualitative research. *Int J Epidemiol* 1992; 21(5):837-

http://genesis.nce.ufrj.br/dataware/tutorial/cbasicos.html#tg_I3, visitado em 19.06.05, 15:43 horas.

Freitas, S F T. “A História Social da Doença Cárie”, BAURU, EDUSC, 126 pp. 2001.

Peres, M A, Latorre, M A D O, Sheiham, A. e cols - “Determinantes Sociais e Biológicos da Cárie Dentária em crianças de 6 anos de idade: um estudo transversal aninhado numa coorte de nascidos vivos no Sul do Brasil”. *Rev. Bras. Epidemiol.* Vol. 6, Nº 4, 2003

Lucas, S D, Portela M C, Mendonça L L, - “Variações no nível de cárie dentária entre crianças de 5 e 12 anos em Minas Gerais, Brasil” *Cad. Saúde Pública*, Rio de Janeiro, 21(1):55-63, jan-fev, 2005

8. BIBLIOGRAFIA CONSULTADA

Baratieri LN. *Dentística: Procedimentos Preventivos e Restauradores*, Sao Paulo: Editora Santos;Quintessence Books, 1995. pp. -509

Barreto JM. *Inteligência Artificial no Limiar do Século XXI*, Florianópolis/SC: 1999.

Barreto JM. Sistemas Especialistas. In: *Inteligência Artificial No Limiar do Século XXI*, Florianópolis/SC: 1999.pp. 197-207.

Battle DL, Vose MD. Isomorphisms of Genetic Algorithms, *Artificial Intelligence* 60, pp: 155-165, Elsevier. (1993)

Bellazi R, Magni P, Larizza C, De Nicolao G, Riva A, Stefanelli M. Mining Biomedical Time Series By Combining Structural Analysis and Temporal Abstractions AnonymousAnonymous1998.

Benn DK, Dankel DD, Kostewicz SH. Can Low Accuracy Disease Risk Predictor Models Improve Health Care Using Decision Support Systems? AnonymousAnonymous1998.

Berkowics, RJ. Primary infection of infants with *Streptococcus mutans*. Apud. *Odontologia para bebês*. São Paulo: Artes Médicas, 1996.

Bezerra ACB, Toledo OA. Nutrição, Dieta e Cárie. In: *Promoção de Saúde Bucal*, Anonymous1997.pp. 43-68.

Blinkhorn A, Attwood D, Gavin G, and O'hickey S. Joint Epidemiological Dental Health of 12 years old school children in Dublin and Glasgow *Community Dent Oral Epidemiology*, vol. 20(5), pp. 307-308, Oct, 2001.

Bodstein, R. Primary care in the agenda of public health sector in Brazil. *Ciênc. saúde coletiva*, 2002, vol.7, no.3, p.401-412. ISSN 1413-8123.

Bratthal D, Pakhomov G. *Dental Caries and Prevention*. 2001.

Bratthal, D. Caries, views and perspectives *Scand.J.Dent.Res.*, vol. 100(1), pp. 47-51, 1992.

Bratthal D. *Dental Caries and Prevention*. 2001.

Buckles BP, Petry FE. *Genetic algorithms*, Washigton:1992. pp. -109

Cantu-Paz E. A summary of research on Parallel Genetic Algorithms, IlliGAL Report, 95007, University of Illinois at Urbana-Champaign. (1995)

Carvalho LAV – A mineração de dados no marketing, Medicina, Economia, Engenharia e Administração. São Paulo: Érica, 2001.

Caufield PW. Dental caries- a transmissible and infectious disease a revisted: position paper. *Ped. Dent.* , v.19, n.8, p. 491-498, 1997.

Caufield PW et al. Initial acquisition of mutans streptococcus by infants: evidence for a discrete window of infectivity. *J. Dent. Res.* v.72, n.1, p. 37- 45, 1993.

Chaves MM. *Odontologia Social*, 1986.

Chen RJ, Meyer R., Yeckel, J. A genetic algorithm for diversity minimization and its parallel implementation, Proc. of the Fifty Int. Conf. on GA, pp: 163-170, San Mateo, CA, Morgan Kaufmann Publishers. (1993)

Davis L. *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, NY. (1991)

Elhanan G, Cimino, JJ. Automated Integration of a Diagnostic Decision Suport System with a Large Scale Clinical Information System-Apects of Feasibility *American Medical Informatics Association*, vol. 1998.

Fogel LJ. *Evolutionary Programming in Perspective: The Top-Down View - Comp. Intelligence - IEEE Press.* (1966)

Freitas, S F T. “A História Social da Doença Cárie”, BAURU, EDUSC, 126 pp. 2001.

Gallagher RJ, Sengupta, S, Hripcsak G, Barrows RC, Clayton PD. An Investigation into the Strengths and Limitations of Artificial Neural Networks: An Application to an Adult ICU Patient Database *American Medical Informatics Association*, vol. 1998.

Garrafa V. Bioética e Odontologia. In: *Promoção de Saúde Bucal*, Anonymous 1997. pp. 465-475.

Goldberg DE, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA. (1989) .

Gomes PR, Costa SC, Cypriano S et al. Dental caries in Paulínia, São Paulo State, Brazil, and WHO goals for 2000 and 2010. *Cad. Saúde Pública*, May/June 2004, vol.20, no.3, p.866-870. ISSN 0102-311X.

Harmon P, King D. *Sistemas Especialistas*, New York: Editora Campus Ltda, 1988.

Haupt RL, Haupt SE. *Practical genetic algorithms* , 1998.

Holland JH. Adaptation in natural and artificial systems, Univ. of Michigan Press, Ann Arbor (1975)

Hongiman B, Bates DW, Light P. Computerized Data Mining for Adverse Drug Events in Outpatient Setting *American Medical Informatics Association*, vol. 1998.

IBGE, <http://www.ibge.gov.br> . 01-04-2001.

Jensen S. Mining Medical Data for Predictive and Sequential Patterns: PKDD 2001 , SPSS UK Ltd.

Keyes PH. Medidas presentes y futuras para controlar la cárie dentária. Apud. Odontologia para bebês. São Paulo: Artes Médicas, 1996.

Krasse B. Risco de cárie: Guia prático para controle e assessoramento. Trad. J. L. F. Andrade. São Paulo: Quintessence, 1986.

Kruger L, Moysés ST. Filosofia da Promoção de Saúde na Clínica Privada. In: *Promoção de Saúde Bucal*, Anonymous 1997. pp. 409-432.

Krol M. Development of a Decision Support System for Detecting Critical Conditions During Anesthesia *American Medical Informatics Association*, vol. 2001.

Loesche W J. Cárie Dental: Uma infecção tratável. Trad.: TELES, J. C. B. et al. Rio de Janeiro: Cultura Médica, 1993.

Lucas, S D, Portela M C, Mendonça L L, - "Variações no nível de cárie dentária entre crianças de 5 e 12 anos em Minas Gerais, Brasil" *Cad. Saúde Pública*, Rio de Janeiro, 21(1):55-63, jan-fev, 2005

Maltz M., Carvalho J. Diagnóstico da doença cárie. In. *Promoção de saúde bucal - Aboprev*. São Paulo: Artes Médicas, 1997.

Marshall J. Using Neural Networks to Predict Patient Satisfaction *American Medical Informatics Association*, vol. 1998.

Medeiros UV. Avaliação comparativa dos testes de atividade cariogênica. *Rev. Odont. Esp.*, v.1, n.1, p. 75-89, 1995.

Medeiros UV. Controle clínico da doença cárie. In: *Dentística Restauradora: o essencial para o clínico*. São Paulo: Santos, 1998.

Ministério da Saúde - Brasil. *Levantamento Epidemiológico em Saúde Bucal, Zona Urbana - 1986*, Brasília: Divisão Nacional de Saúde Bucal, 1988.

Ministério da Saúde - Brasil. Projeto SB Brasil 2003 - Condições de saúde bucal da população brasileira 2002-2003 Resultados Principais. Brasília – Coordenação Nacional de Saúde Bucal – 2004.

Monard, M. C.; Batista, G. E. A. P. A.; Kawamoto, S.; Pugliesi, J. B. “Uma Introdução ao Aprendizado Simbólico de Máquina por Exemplos”. Notas do ICMC, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, Outubro 1997.

Morita, et al. Cárie Dentária. Apud. Odontologia para o bebê. São Paulo: Artes Médicas, 1996.

Moysés SJ. Conceito de Promoção de Saúde na Construção de Sistemas de Atenção em Saúde Bucal. In: *Promoção de Saúde Bucal*, 1997. pp. 371-408.

Newbrun E. Cariologia. São Paulo: Santos, 1988.

Newell DJ. Intention-to-treat analysis: implications for quantitative and qualitative research. *Int J Epidemiol* 1992; 21(5):837-

Newton G et al. Empirical Bayes Analysis. <http://www.biostat.wisc.edu/geda/ebasummary.html> . 8-2-2001.

Norsys. Belief Networks and Decision Networks. <http://www.norsys.com/belief.html> .

Organização Mundial de Saúde. *Levantamento Epidemiológico Básico em Saúde Bucal - Manual de Instrução*, Genebra:1991.

Organização Mundial de Saúde. *Levantamentos Básicos em Saúde Bucal*, São Paulo: Santos, 1999.

Peres, M A, Latorre, M A D O, Sheiham, A. e cols. - “Determinantes Sociais e Biológicos da Cárie Dentária em crianças de 6 anos de idade: um estudo transversal aninhado numa coorte de nascidos vivos no Sul do Brasil”. *Rev. Bras. Epidemiol.* Vol. 6, Nº 4, 2003

Pinto VG. Epidemiologia das Doenças Bucais no Brasil. In: *Promoção de Saúde Bucal*, Anonymous 1997. pp. 27-42.

Pitt N. The Scottish Health Boards Dental Epidemiological Programme: Initial Surveys of 05 and 12 Years Old *Br.Dent.J.*, vol. 172(11), pp. 408-413, 1992.

Reddy J. The WHO Oral Health Goals for The Year 2000 in South of Africa *Int.Dent.J.*, vol. 42(3), pp. 150-156, 1992.

Rouquayrol, MZ, Filho NA. *Epidemiologia e Saúde*, 1999.

Slavkin HC. First encounters: transmission of infectious oral diseases from mother to child. *J. Am. Dent. Ass.* v.128, n.6, p. 773-8, 1997.

Thylstrup A, Fejerskov, O. *Tratado de Cariologia*. Trad.: WEYNE, S.. Rio de Janeiro: Cultura Médica, 1998.

Walter LRF et al. *Odontologia para o bebê*. SãoPaulo: Artes Médicas, 1996.

Walter LRF, Nakama L. Cárie Dentária. Apud. *Odontologia para o bebê*. SãoPaulo: Artes Médicas, 1996.

Wang PC, Korfhage W. Process scheduling using genetic algorithms, *IEEE Symposium on Parallel and Distributed Processing*, San Antonio, USA, pp: 638-641. (1995)

http://genesis.nce.ufrj.br/dataware/tutorial/cbasicos.html#tg_l3, visitado em 19.06.05, 15:43 horas.