

**Fundação Oswaldo Cruz
Escola Nacional de Saúde Pública
Centro de Pesquisas Aggeu Magalhães**

Doutorado Interinstitucional em Saúde Pública

Sistema de Informações Geográficas para
o Programa de Saúde da Família

José Luiz Portugal

Recife, 2003

JOSÉ LUIZ PORTUGAL

**Sistema de Informações Geográficas para o Programa de
Saúde da Família**

Tese apresentada ao Curso de **Doutorado em Saúde Pública** da **Escola Nacional de Saúde Pública – Fundação Oswaldo Cruz**, como parte dos requisitos à obtenção do título de Doutor em Saúde Pública.

Orientadora: Dra. Marília Sá Carvalho

Recife

2003

Catálogo na fonte: Biblioteca do Centro de Pesquisas Aggeu Magalhães

528.88:55

P853s Portugal, José Luiz.

Sistema de Informações Geográficas para o Programa de Saúde da Família / José Luiz Portugal. – Recife: J. L. Portugal, 2003.

126 p.:il., tabs.

Tese (Doutorado em Saúde Pública) – Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, 2003.

Orientadora: Marília Sá Carvalho.

1.Sistema de informações geográficas 2. Programa de saúde da família 3. Cuidados primários de saúde. 4. Métodos e procedimentos estatísticos. 5. Comportamento espacial. 6. Análise estatística. I.Carvalho, Marília Sá. II. Título.

CDU: 528.88:55

Aos meus filhos Carla e Otávio, pelo que a mim representam.

Apresentação

Nas duas últimas décadas a saúde pública tem se reaproximado das categorias espaciais, de forma a incorporar em seu arcabouço teórico e analítico novas variáveis, categorias, métodos e técnicas. As técnicas de Cartografia Digital, de Sistemas de Informação Geográficas (SIG) e de Análise Estatística Espacial se inserem nesse contexto, gerando grandes potencialidades, tanto para estudos acadêmicos quanto para utilizações no planejamento e na gestão dos serviços.

No ano de 1999, aconteceu no Recife uma Oficina de Trabalho que discutiu uma *Proposta de Implantação de um Sistema de Informações Geográficas para Apoio ao Programa de Agentes Comunitários de Saúde/Programa de Saúde da Família*. Essa temática inspirou a presente tese, porque envolvia diversas áreas do conhecimento científico-tecnológico e tinha como pano de fundo uma aplicação prática de significativa complexidade: a necessidade de um Sistema de Informações Geográficas para apoiar um programa de saúde, implantado nacionalmente, com elevado grau de descentralização.

Dentro desse contexto procede-se a um estudo detalhado das informações pertinentes aos protocolos e rotinas de trabalho de todos os envolvidos com o PSF, de modo que se definam as necessidades do Programa. Detalham-se conceitos de SIG, com a finalidade de selecionar a estrutura de dados mais conveniente e as operações de análise, espaciais e não espaciais, passíveis de execução. Identifica-se que o “Processo Pontual Marcado” é uma opção para mapear áreas de risco das doenças cobertas pelo PSF. Formula-se um projeto lógico e este é implementado na forma de um “Piloto”, numa área rural. Conclui-se que o SIG desenvolvido representa um avanço importante em relação ao atual sistema de informação utilizado pelo PSF, constituindo-se em um instrumento facilitador da consolidação da atenção básica no País.

Durante o doutorado, muitas pessoas e instituições estiveram presentes, contribuindo de forma efetiva ao meu aprendizado. A todos os meus agradecimentos, em especial:

- Ao Centro de Pesquisas Aggeu Magalhães, à Secretaria Estadual de Saúde/PE, à Secretaria de Saúde de Caruaru, ao Departamento de

Engenharia de Cartográfica da UFPE, à Equipe de Saúde da Família do Juá, que contribuíram para a execução dos trabalhos de campo;

- À Wayner Vieira de Souza e Tiago Maria Lapa, grandes amigos que me iniciaram na Saúde Pública e muito contribuíram para que esse trabalho tivesse sucesso;
- À Alexandre Bezerra de Carvalho, que então Diretor do CPqAM, autorizou meu ingresso na “Turma Recife” de Doutorado da ENSP e com quem tenho grande identidade;
- À Dra. Marília Sá Carvalho, pela sabedoria, paciência e estímulo a mim dedicados, como professora e orientadora dessa tese;
- Aos colegas de turma, pelo convívio harmônico, em especial à Maria Luiza e Ana Brito, com quem estive próximo em trabalhos de grupo;
- À Luci, exemplo de dedicação e competência, pelo convívio e estímulo ao curso;
- À Mazé pelo auxílio nas inúmeras revisões dessa tese e também por motivos óbvios.

Resumo

Esta tese tem por objetivo criar um Sistema de Informações Geográficas (SIG) para o Programa de Saúde da Família (PSF) que se constitui um instrumento de auxílio ao planejamento, à prestação e à avaliação dos serviços previstos pelo Programa. Para tanto, estudaram-se os protocolos e as rotinas de trabalho de todos os comprometidos com o PSF. Verifica-se que ele é composto por três níveis funcionais, que são o operacional, o gerencial e o político, os quais deveriam ter por suporte a informação, através do Sistema de Informação da Atenção Básica (SIAB). Identifica-se que o SIAB tem restrições que podem ser minimizadas ou mesmo eliminadas com a implantação de um novo sistema composto por informações espaciais e não espaciais, o que é permitido em ambiente de SIG. O indivíduo, associado às coordenadas pontuais de sua residência, deve ser considerado como a unidade básica desse novo sistema, por permitir agregação com todos os níveis hierárquicos superiores, sem perda de informações. Conceitos de SIG, sob a ótica do PSF, são detalhados com ênfase nas estruturas de armazenamento, de organização dos dados e de seus relacionamentos. Observa-se que determinadas operações de análise espacial, importantes para a identificação de áreas de risco, não são disponibilizadas por programas dedicados a SIG. Assim, discute-se alguns métodos estatísticos para o mapeamento de doenças, com ênfase nos que se baseiam em padrões de pontos, mais especificamente o chamado processo pontual marcado. Formula-se, a partir desse estágio, o modelo conceitual do Sistema, na forma de Diagrama Entidade-Relacionamento, com o respectivo Dicionário de Dados. Parte desse modelo é implementado em computador, a partir de dados coletados na USF do Juá, município de Caruaru-PE, detalhando-se um estudo de caso da hipertensão arterial. O sistema é explorado com saídas gráficas na forma de mapas e descritivas na forma de tabelas. Terminada a fase de exploração, conclui-se por sua viabilidade, ressaltando-se os seguintes aspectos: *(i)* ainda não existe um programa de computador único, capaz de executar todas as operações de análise espacial previstas nesse SIG; *(ii)* o município deve ser o local de implantação do sistema como um todo; *(iii)* as Unidades de Saúde da Família (USF) devem ser responsáveis pela sua carga, diretamente em ambiente computacional de banco de dados; *(iv)* o fluxo dos dados das USF para o município deve ser periódico e sistemático, preferencialmente por Internet; *(v)* a obtenção das coordenadas das residências deve se processar a partir de mapas pré-existentes ou, na sua ausência, a partir de imagens de satélite com resolução espacial da ordem de 1 metro; e *(vi)* o

processo pontual marcado é viável para detecção, identificação ou localização, a partir do nível de município, das doenças que o PSF trata, desde que sejam definidas, *a priori*, por profissionais da área de saúde, as covariáveis influentes nesses eventos. O SIG desenvolvido representa um avanço importante em relação ao atual sistema de informação utilizado pelo PSF, constituindo-se em um instrumento facilitador da consolidação da atenção básica no País.

Palavras Chave: Sistema de Informações Geográficas; Programa de Saúde da Família; Sistema de Informações da Atenção Básica; Estatística Espacial; Análise Espacial.

Abstract

The proposal of this thesis is the development of a Geographical Information System (GIS) for the Program of -Family Health (PSF), in order to support the planning, implementation and evaluation of its activities. Those encompasses three functional levels: operational, executive and political, and that all of them rely on information provided by the Basic Care Information System (SIAB). It has been identified some restrictions to the SIAB that can be minimized or even eliminated by the implementation of this scheme, as it comprises both spatial and non-spatial information, in a GIS's Environment. , The point coordinates of the residence of each person assisted, is the basic unit of the new system, and they can be aggregated in any devised way by the upper levels of the health system. GIS concepts are reviewed, subordinated to the main PSF's practices, with emphasis on the structures of data storage, organization and interaction. Certain of spatial analysis operations, relevant for determining differences in risks areas, are not available within GIS's programs. Statistics methods more sophisticated, with emphasis on spatial point patterns are presented. The conceptual model of the system formulated is presented in an entity-relation diagram form, followed by the respective Data Dictionary . The model was implemented as a case study, based on data of the Health-of-Family Unit (USF) of Juá, district of Caruaru, State of Pernambuco. The occurrence of Arterial Hypertension was studied in more detail to test possibilities of analysis, exploring maps, tables, graphic and a more complex statistical model This exercise allows the conclusion that the system is feasible, and the following recommendations are of utmost importance: (i) currently, all the spatial analysis specified in the GIS can not be carried out by the same software; (ii) the municipality should be the site for the implementation of the system, as a whole. (iii) the USF should be responsible for loading the information into the database. (iv) the flow of data from the USF to the municipality must be periodic, preferentially running through an Internet connection (v) geographic coordinates of residences should be obtained from pre-existing maps or, if not available, from satellite images, with resolution of one meter; (vi) the "marked point process" is valid for selecting risk areas, starting at the municipality level and comprising all diseases encompassed by the PSF, since the covariables impacting such events are defined *a priori* by health professionals; (vii) technical training and support at municipal and state levels are vital for implementing the GIS for the PACS/PSF.

The GIS can add meaningful knowledge to the information system currently in use by the PSF, helping to consolidate the *basic care policy* in the country.

Key Words: Geographical Information Systems; Health-of-Family Program; Basic-Care Information System; Spatial Statistics; Spatial Analysis.

Lista de Abreviaturas

ACS	Agente Comunitário de Saúde
CAE	Completa Aleatoriedade Espacial
DER	Diagrama Entidade-Relacionamento
ESF	Equipe de Saúde da Família
FIOCRUZ	Fundação Oswaldo Cruz
FUNASA	Fundação Nacional de Saúde
GPS	Sistema de Posicionamento Global
IBGE	Instituto Brasileiro de Geografia e Estatística
MS	Ministério da Saúde
OMS	Organização Mundial de Saúde
OPAS	Organização Panamericana de Saúde
PACS	Programa de Agentes Comunitários de Saúde
PSF	Programa de Saúde da Família
RIPSA	Rede Interagencial de Informações para a Saúde
SIAB	Sistema de Informações da Atenção Básica
SIG	Sistema de Informações Geográficas
SIH	Sistema de Informações Hospitalares
SIM	Sistema de Informações de Mortalidade
SINAM	Sistema de Informações de Agravos de Notificação
SINASC	Sistema de Informações de Nascidos Vivos
SGBD	Sistema Gerenciador de Banco de Dados
USF	Unidade de Saúde da Família
UTM	Universal Transverso de Mercator

Lista de Figuras

Figura		Página
1	Pirâmide organizacional do PACS/PSF	3
2	Pirâmide organizacional do PACS/PSF sob a ótica da informação	4
3	Setor censitário do Juá	11
4	Estrutura organizacional da ESF	18
5	Estrutura do SIAB	20
6	Estrutura do sistema proposto	21
7	Estrutura hierárquica	23
8	Estrutura em rede	24
9	Estrutura relacional	25
10	Estrutura orientada a objetos	26
11	Sistema de referência plano	30
12	Atributos de residência	31
13	Produto cartesiano $A \times B$	32
14	Produto cartesiano $A \times P$	32
15	Estrutura de dados <i>raster</i>	34
16	Associação cor-valor	35
17	Operação pontual	35
18	Combinação de camadas	36
19	Filtro passa altas	37
20	Estrutura vetorial	38
21	Topologia arco-nó	39
22	Área de influência	41
23	União	42
24	Identidade	42
25	Interseção	43
26	Combinação de feições	43
27	Extração	44
28	Elo implícito	45
29	Elo explícito	46
30	Arquivos de dados	47
31	Relação	48
32	Relação A	48
33	Relações	49
34	União	50
35	Interseção	50
36	Diferença	50
37	Produto cartesiano	51
38	Projeção	51
39	Restrição	52
40	Junção	52
41	Divisão ($T \leftarrow A [y \div z] B$)	52
42	Ligação entidades espaciais – atributos	53
43	Distribuição de residências	55
44	Superfície dos nascidos vivos, Recife, PE, 1995-1999	56
45	Nascidos vivos por bairro, Recife, PE, 1995-1999	57
46	Fluxo entre regiões de residência e de óbitos pós-neonatal, Rio de Janeiro, 1995	58
47	Variação da densidade demográfica, segundo grade de 5 metros	60

48	Variação da densidade demográfica, segundo grade de 0,5 metros	61
49	Interpretação geométrica do estimador kernel	62
50	Variação da densidade demográfica, segundo raio de busca de 20 metros	63
51	Diagrama entidade-relacionamento	70
52	Área piloto	79
53	Diagrama entidade-relacionamento implementado	81
54	Método Stop and Go	86
55	Resultados dos levantamentos espaciais	86
56	Estrutura do banco de dados	87
57	Formulário	88
58	Esquema do modelo georrelacional	89
59	Distribuição da população por faixa etária e sexo, Juá, 2001	91
60	Distribuição da população por faixa etária e renda, Juá, 2001	91
61	Distribuição dos hipertensos por faixa etária e sexo, Juá, 2001	92
62	Distribuição da população por faixa etária e escolaridade, Juá, 2001	92
63	Totalização de crianças, por agente, com idade inferior a 9 anos e alfabetizadas, Juá, 2001	93
64	Mapas da componente não paramétrica $g(s)$ e da superfície de p -valores	103
65	Aglomerado populacional de parte da região do Juá	111
66	Obstáculos para os sinais de satélites	112
67	Imagem IKONOS, Salvador – BA, 2002	113
68	Fluxo das informações	115

Lista de Tabelas

Tabela		Página
1	Distribuições de freqüência das variáveis disponíveis, Juá, 2001	90
2	Resultados da regressão com todas as variáveis	100
3	Resultados da regressão com todas as variáveis e interações	101
4	Resultados da regressão incluindo-se a variável espaço	102
5	Resultados obtidos para diversos valores de largura de banda	102

SUMÁRIO**Página**

Apresentação	iii
Resumo	v
Abstract	vii
Lista de Abreviaturas	ix
Lista de Figuras	x
Lista de Tabelas	xi
1. Introdução	1
1.1. Objetivo	6
1.1.1. Objetivo geral	6
1.1.2. Objetivos específicos	6
2. Programa de Saúde da Família	8
2.1. Entidades componentes do PSF	9
2.2. Equipe de saúde da família	12
2.2.1. Diagnóstico da realidade	13
2.2.2. Programação das atividades	17
2.2.3. Acompanhamento e avaliação	18
2.3. Crítica ao SIAB	20
3. Sistema de Informações Geográficas	23
3.1. Escolha da estrutura mais adequada	27
3.2. Dado geográfico	29
3.3. Estruturas de armazenamento e análise dos dados gráficos	33
3.3.1. Estrutura de dados <i>raster</i>	33
3.3.1.1. Análise espacial na estrutura <i>raster</i>	35
3.3.2. Estrutura de dados vetorial	37
3.3.2.1 Análise espacial na estrutura vetorial	40
3.4. Armazenamento dos atributos em meio magnético	44
3.4.1. Modelo de banco de dados relacional	47
3.5. Ligação dos dados gráficos aos atributos	53
4. Métodos Estatísticos para mapeamento de doenças	55
4.1. Padrões de pontos	58
4.2. Exploração do padrão de pontos	60
4.3. Processo pontual marcado	63
5. Proposta de um modelo conceitual de SIG para o PSF	68
5.1. Projeto lógico	70
5.2. Dicionário de dados	72
5.2.1. Dicionário das informações espaciais	72
5.2.2. Dicionário das informações não espaciais	74
6. Implementação do projeto lógico	79
6.1. Caracterização da área e população de estudo	79
6.2. Projeto lógico implementado	80
6.2.1 Dicionário de dados	81
6.3. Carga do banco de dados	85
6.3.1. Trabalho de campo	85

6.3.2. Validação dos trabalhos de campo	88
6.4. Exploração do sistema	89
6.4.1. Análise espacial em instâncias superiores	93
7. Estudo de caso: a hipertensão arterial no Juá	99
8. Considerações finais e recomendações	108
9. Conclusão	118
10. Referências Bibliográficas	121
Anexos	127

1. INTRODUÇÃO

A política de saúde no Brasil sofreu significativa mudança a partir da década de 1990, quando a atenção básica passa a ser uma área de concentração de esforços, com a disponibilização de incentivos federais, calculados e transferidos em base *per-capita* (MS, 2002). Dentro dessa ótica, dois programas interligados têm-se destacado: o Programa de Agentes Comunitários (PACS) e o Programa de Saúde da Família (PSF). Em âmbito nacional, o PACS foi criado em 1992, começando-se, assim, o enfoque da família como unidade de ação programática de saúde e a introdução da noção de área de cobertura por família (MS, 2002; Viana & Dal Poz, 1998).

Com início formal em 1993 e vinculado à Fundação Nacional de Saúde (FUNASA), o PSF representa uma possibilidade de reestruturação da atenção primária, a partir de um conjunto de ações conjugadas e em sintonia com os princípios de territorialização, intersetorialidade, descentralização, co-responsabilidade e priorização de grupos populacionais com maior risco de adoecer ou morrer (Trad & Bastos, 1998).

A expansão desses programas tem sido tal que, em dezembro de 2000, o PSF passa a existir em 3.090 municípios, com 10.473 equipes de saúde da família, cobrindo 22% da população brasileira. O PACS, por sua vez, é operacionalizado em 4.610 municípios, com 145.299 agentes, cobrindo 51% da população do país (Souza et al., 2000).

O PACS/PSF tem por característica um alto grau de descentralização, e cobertura em âmbito nacional, respeitando as especificidades regionais. Sem dúvida, o sucesso desses programas depende em grande parte de instrumentos práticos, inovadores, dinâmicos e flexíveis que apoiem seu processo de implantação, gerenciamento e tomada de decisão (Tasca et al., 1993). Dentro dessa linha de raciocínio, é estruturado em 1988 o Sistema de Informações da Atenção Básica (SIAB).

O SIAB tem por objetivo acompanhar a implementação e monitorar os resultados do PACS e PSF (MS, 2002). A unidade privilegiada de geração e coleta das informações do sistema é a família/domicílio. O processamento das informações no sistema consiste basicamente na agregação dos dados entre os diversos níveis operacionais (microárea,

área, segmento, município, regional de saúde, estado) e na respectiva construção de indicadores. As saídas são relatórios de dados e indicadores agregados, voltados para o acompanhamento e avaliação do desempenho das equipes de saúde da família e dos agentes de saúde (MS, 1998).

Uma crítica a esse sistema é que através dele é impossível realizar estudos a nível individual, dado que as informações, apesar de coletadas nesse nível só são registradas no SIAB de forma agregada. Questiona-se também o modo de agregação dos dados, uma vez que da maneira como está estruturado, perde-se a conexão entre os diversos níveis operacionais, devido à ausência de uma chave de ligação (elo) entre eles. Essa concepção do SIAB gera dificuldades para o desenvolvimento de diversas atividades, entre as quais citam-se:

- Utilização das informações para vigilância epidemiológica (busca ativa e investigação de casos);
- Vinculação entre as informações do SIAB com outros sistemas de informação em saúde e com bases de dados demográficos e sócio-econômicos, ressaltando-se aqui que a integração entre as bases de dados (Painho et al, 1999) só é possível se estiverem relacionadas às mesmas unidades de referência (domicílio, microárea, área, etc), caso contrário, a acessibilidade, a qualidade e a atualização das informações será prejudicada (Barcellos, 1996);
- Georreferenciamento de informações sobre eventos de saúde, o que compromete o uso de Sistemas de Informações Geográficas (SIG) e de análise espacial de dados, como a identificação do padrão de ocorrência espacial da morbi-mortalidade, a identificação de conglomerados de casos numa área específica; etc. (FIOCRUZ, 1999).

Levantados esses questionamentos, convém identificar como o PACS/PSF insere-se no âmbito do Ministério da Saúde, do ponto de vista da informação. Considere-se que uma organização pública ou privada é geralmente composta por 3 níveis funcionais: nível operacional, responsável pela produção do serviço; nível gerencial, responsável pelo controle dos recursos empregados no nível operacional; nível político, responsável pela definição de leis, normas e programas que garantam o

sucesso da organização (Huxhold & Levinsohn, 1995). Adaptando-se esse conceito para o PACS/PSF, pode-se representá-lo graficamente, na forma de uma pirâmide, conforme mostrado na Figura 1.

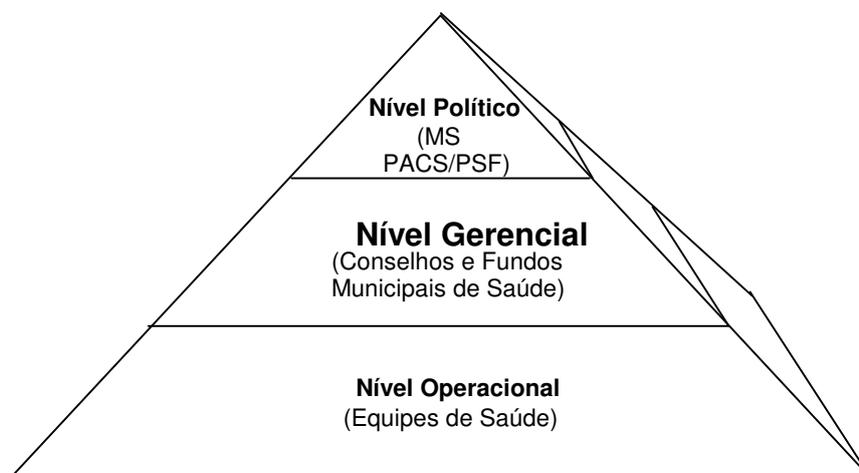


Figura 1 – Pirâmide organizacional do PACS/PSF

Na base da pirâmide, o nível operacional, é onde as políticas e o gerenciamento do programa se materializam em ações de atenção básica à saúde, através das equipes de saúde da família. Na parte central da pirâmide, o nível gerencial, é onde as políticas são materializadas em planos de ação e onde os recursos do PACS/PSF são controlados, através das secretarias de saúde estaduais e municipais, no último caso pelo Conselho e pelo Fundo Municipal de Saúde. Na parte superior da pirâmide, o nível político, é onde os orçamentos são aprovados, normas e ordens são emitidas, adaptações e outras políticas de atenção básica são definidas pelo Ministério da Saúde.

Cada um dos níveis funcionais tem por suporte a informação, por meio de uma integração horizontal e vertical de dados. A integração horizontal caracteriza-se pela combinação de informações dentro de um mesmo nível e a vertical pela agregação das informações (Tobler, 1989) de um nível que fluem para os níveis superiores. Assim, determinadas informações coletadas no nível operacional são empregadas no nível político para iniciar planos e programas. Estes, por sua vez, são enviados ao nível gerencial que os transformam em ações para serem colocadas em prática pelo nível operacional, formando-se assim um sistema que se retroalimenta, graficamente representado na Figura 2.

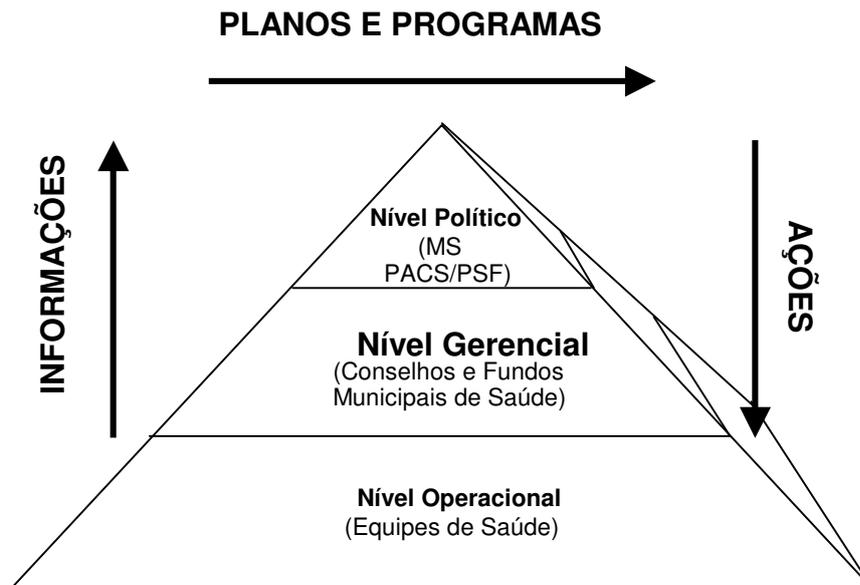


Figura 2 - Pirâmide organizacional do PACS/PSF sob a ótica da informação

Cabe observar que a coleta da informação acontece no nível operacional, ou seja, nas Equipes de Saúde da Família. Especial atenção deve ser dada a esse aspecto, uma vez que uma coleta imprecisa ou incompleta pode comprometer todo o sistema de informação e, conseqüentemente, o programa.

Outra consideração a ser destacada é que a agregação horizontal e vertical dos dados do PACS/PSF pode se processar levando-se em consideração a posição geográfica dos eventos que o programa trata, (Starfield, 2002). Isso é permitido por sistemas de informação especialmente desenvolvidos para esse fim, denominados Sistemas de Informações Geográficas (SIG).

Na década de 1990, houve grande disseminação do geoprocessamento, definido como o conjunto de tecnologias de coleta, tratamento, manipulação e apresentação de informações espaciais (Pina, 1998). Neste campo inserem-se os Sistemas de Informações Geográficas (SIG).

O SIG é um sistema de informações, desenvolvido sobre um programa de computador específico, que utiliza um Sistema Gerenciador de Banco de Dados (SGBD), permitindo tratar dados das mais diversas fontes, processar, armazenar, recuperar e analisar os mesmos, com o objetivo de

gerar informações e apresentar resultados em formato passível de ser compreendido pelo usuário, oferecendo um ambiente integrado (Strauch, 1998).

O principal fator de sucesso de um SIG é a aplicação que se deseja do sistema (Painho, 1999). É essa que vai identificar a parcela do mundo real a ser representada, tanto gráfica quanto descritivamente e também a estrutura dos dados e suas relações. Para que isso se processe, recomenda-se que as 3 fases a seguir sejam respeitadas (Sá, 2001; Runbaugh et al. 1994; Silva, 1991):

1. Abstração do mundo real;
2. Definição do modelo conceitual;
3. Elaboração do modelo físico.

A abstração do mundo real se caracteriza pelo estudo do fenômeno bem como dos dados que o explicam e deve ser sempre acompanhada por profissionais ligados ao dito fenômeno.

A definição do modelo conceitual se caracteriza pela representação lógica da abstração. Essa atividade é materializada pela especificação de diagramas que identificam a estrutura e fluxo dos dados, bem como o modo como esses se relacionam (Setzer, 1989).

A elaboração do modelo físico se caracteriza pela implementação em computador do que foi definido no modelo lógico, envolvendo inclusive a escolha do programa adequado.

Dadas as possibilidades dos Sistemas de Informações Geográficas, julga-se oportuno seu emprego como instrumento de aperfeiçoamento para o PACS/PSF e para o SIAB, por apresentar potencial para auxiliar o planejamento, a prestação e a avaliação dos serviços à população adscrita, para melhorar as bases do conhecimento nas quais se encontra o progresso da prestação dos serviços, além de possibilitar a inclusão de aspectos ambientais na análise da saúde (Scholten, 1991). Mais especificamente, a criação de um SIG para o PACS/PSF pode contribuir para (FIOCRUZ, 1999):

- Subsidiar a definição da população/território adscrita a uma microárea, área ou unidade de saúde da família;
- Identificar a cobertura geográfica do PACS/PSF;

- Identificar estratos populacionais expostos a maiores riscos ambientais, sócio-econômicos e de morbi-mortalidade;
- Identificar/localizar eventos de saúde no espaço geográfico;
- Validar a coleta de dados e indicadores construídos pela equipe de saúde da família através da comparação com outras bases de dados;
- Visualizar indicadores operacionais e de resultados do programa através da construção de mapas temáticos;
- Integrar diferentes fontes de informação em uma mesma base geográfica;
- Utilizar técnicas de análise espacial para identificar dependência espacial no padrão de ocorrência dos eventos de saúde;
- Identificar fontes ambientais de risco para a saúde, na área de abrangência do programa.

Convém observar que essas contribuições são determinadas na fase da **abstração do mundo real**.

Definido o que se espera do sistema, identificam-se como críticas e complexas as especificações das estruturas de armazenamento, de organização dos dados e seus relacionamentos, ou seja, a **elaboração do modelo conceitual**. Na realidade essa deve ser transparente para o usuário, porém fundamental para que se possa implementar em computador o ambiente organizacional do PACS/PSF, sob a ótica da informação.

Face à relevância dessa temática, definem-se a seguir os objetivos da presente tese.

1.1 OBJETIVOS

1.1.1 OBJETIVO GERAL

Propor um Sistema de Informações Geográficas para apoio ao planejamento, operacionalização e avaliação das atividades dos Programas de Agentes Comunitários e de Saúde da Família.

1.1.2 OBJETIVOS ESPECÍFICOS

- Contextualizar o Programa de Saúde da Família sob a ótica de sistemas de informação;
- Descrever os conceitos de Sistemas de Informações Geográficas com ênfase nas estruturas de armazenamento, organização dos dados e seus relacionamentos, visando o uso pelo PACS/PSF;
- Elaborar um modelo conceitual de SIG tomando-se por unidade básica a família/domicílio;
- Implementar em computador o modelo conceitual, na forma de um projeto piloto, em área previamente estabelecida;
- Testar o sistema, por meio de determinadas operações de consulta, análise espacial e visualização de dados.

2. PROGRAMA DE SAÚDE DA FAMÍLIA

A Saúde da Família é uma estratégia que prioriza as ações de promoção, proteção e recuperação da saúde dos indivíduos e da família, do recém-nascido ao idoso, sadios ou doentes, de forma integral e contínua. Tem por objetivo a reorganização da prática assistencial em novas bases e critérios, em substituição ao modelo tradicional de assistência, orientado para a cura de doenças. A atenção está centrada na família, entendida e percebida a partir do seu ambiente físico e social, o que vem possibilitando às equipes de Saúde da Família uma compreensão ampliada do processo saúde/doença e da necessidade de intervenções que vão além de práticas curativas (MS, 2000).

Dentro desse contexto, Programa de Saúde da Família foi idealizado para aproximar os serviços de saúde da população, para cumprir o princípio constitucional do Estado de garantir ao cidadão seu direito de receber atenção integral à saúde - com prioridade para as atividades preventivas, mas sem prejuízo dos serviços assistenciais - e para permitir que os responsáveis pela oferta dos serviços de saúde, os gestores do SUS, aprofundem o conhecimento sobre aqueles a quem devem servir (MS, 2000).

O PSF está estruturado a partir da Unidade de Saúde da Família - uma unidade pública de saúde, com equipe multiprofissional que assume a responsabilização por uma determinada população, a ela vinculada, onde desenvolve ações de promoção da saúde e de prevenção, tratamento e reabilitação de agravos.

A Unidade de Saúde da Família atua com base nos seguintes princípios:

- **Caráter substitutivo:** corresponde à substituição das práticas convencionais de assistência por um novo processo de trabalho, cujo eixo está centrado na vigilância à saúde;
- **Integralidade e Hierarquização:** A Unidade de Saúde da Família está inserida no primeiro nível de ações e serviços do sistema local de saúde, denominado atenção básica. Deve estar vinculada à rede de serviços de forma que se garanta atenção integral aos indivíduos e famílias e seja assegurado a referência e contra-referência para os diversos níveis do sistema, sempre que for

requerido maior complexidade tecnológica para a resolução de situações ou problemas identificados na atenção básica.

- Territorialização e adscrição da clientela: trabalha com território de abrangência definido e é responsável pelo cadastramento e acompanhamento da população adscrita a esta área.
- Equipe multiprofissional: A equipe de Saúde da Família é composta minimamente por um médico generalista ou médico de família, um enfermeiro, um auxiliar de enfermagem e de quatro a seis agentes comunitários de saúde.

Definidas as linhas gerais do PSF, esse capítulo destina-se a contextualizá-lo sob a ótica de sistemas informação, de modo a atender as necessidades de todas as pessoas envolvidas nos níveis operacionais e gerencias do programa.

Para cumprir esse objetivo, se faz necessário um estudo criterioso das informações pertinentes aos protocolos e rotinas de trabalho de todos os comprometidos com o PSF. Esse estudo tem por base a observação da realidade numa Unidade de Saúde e a consulta aos Cadernos de Atenção Básica (MS,2000), por ser esse o documento que melhor detalha o Programa.

2.1 – ENTIDADES COMPONENTES DO PSF

As *entidades*¹ que compõe o PSF, que são a família/domicílio, a microárea, a área, o segmento territorial, o município, todas essas identificadas univocamente por um número² e a unidade de saúde da família.

A **Família** é o conjunto de pessoas ligadas por laços de parentesco, dependência doméstica ou normas de convivência que residem na mesma unidade domiciliar. Inclui empregado (a) doméstico (a) que reside no domicílio, pensionistas e agregados (MS, 1988).

O **Domicílio** designa o local de moradia estruturalmente separado e independente, constituído por um ou mais cômodos. A separação fica

¹ *Entidade* sob a ótica de sistemas de informação, corresponde a todo objeto que tem existência própria, quando considerado no contexto das atividades de uma organização e é descrita por seus atributos (Furtado et al, 1989).

² Ver Anexo 1 – Ficha A (Frente)

caracterizada quando o local de moradia é limitado por paredes (muros ou cercas, entre outros) e coberto por um teto que permita que seus moradores se isolem e cujos residentes arcam com parte ou todas as despesas de alimentação ou moradia. Considera-se *independente* o local de moradia que tem acesso direto e que permite a entrada e a saída de seus moradores sem a passagem por local de moradia de outras pessoas. Em casas de cômodos (cortiços), considera-se como um domicílio cada unidade residencial. Também são considerados domicílios os prédios em construção, as embarcações, carroças, vagões, tendas, grutas e outros locais que estejam servindo de moradia para a família (MS, 1988).

A **Microárea** é o espaço geográfico delimitado, onde residem cerca de 400 a 750 pessoas e correspondem à área de atuação de um agente comunitário de saúde - ACS (MS, 1988). Essa entidade tem restrições de ordem prática para ser implementada, na medida que prevê delimitação do espaço. Numa região urbana o problema pode ser minimizado, tomando-se por referência o arruamento, desde que perfeitamente definido. O problema cresce em complexidade no interior de favelas e regiões rurais, dado que nessas situações não existem referências concretas para limites e as pessoas distribuem-se em aglomerados populacionais. Uma solução possível para definição de microárea é sua representação por meio de croquis³. Caso essa opção seja empregada, as operações de análise espacial ficam restritas a métodos que não envolvam geometria.

A **Área** é o conjunto de microáreas contíguas sob a responsabilidade de uma equipe de saúde da família, onde residem em torno de 2400 a 4500 pessoas (MS, 1988). Essa entidade apresenta as mesmas restrições da anterior.

O **Segmento Territorial** é o conjunto de áreas contíguas que pode corresponder à delimitação de um distrito sanitário, de uma zona de informação do IBGE ou a outro nível de agregação importante para o planejamento e avaliação em saúde no município. *É a divisão territorial utilizada para a análise espacial dos dados em um determinado município* (MS, 1988). Essa entidade apresenta 2 problemas, o primeiro trata de delimitação conforme já comentado para microárea e área. O segundo é também crítico, na medida que transfere para o município a identificação

³ Ver Capítulo 3

do nível de agregação a ser empregado. Isso acontecendo, não existe a possibilidade de sistematizar a informação, dado que determinado município pode empregar um nível que não seja o empregado por outro.

O **Município** é a circunscrição administrativa autônoma do estado, governada por um prefeito e uma câmara de vereadores.

A **Unidade de Saúde da Família (USF)** é considerada como a porta de entrada e primeiro nível de atenção, devendo estar integrada em uma rede de serviços dos diversos níveis de complexidade, estabelecendo um sistema de referência e contra-referência que garanta resolutividade e possibilite o acompanhamento dos pacientes (MS, 2000).

Dado o que foi levantado, observa-se que a geometria das entidades é praticamente desconsiderada para o PSF. É da opinião desse pesquisador que tal fato tem sua razão de ser porque extrapola a competência dos envolvidos no Programa.

A primeira vista esse problema poderia ser resolvido empregando-se mapas já confeccionados, porém esses são raros no país para uma escala compatível com microárea, área e segmento territorial. Deve ser ressaltado, que até mesmo a malha de setores censitários disponibilizada pelo IBGE é, em grande parte, confeccionada como croquis, conforme mostrado na Figura 3.

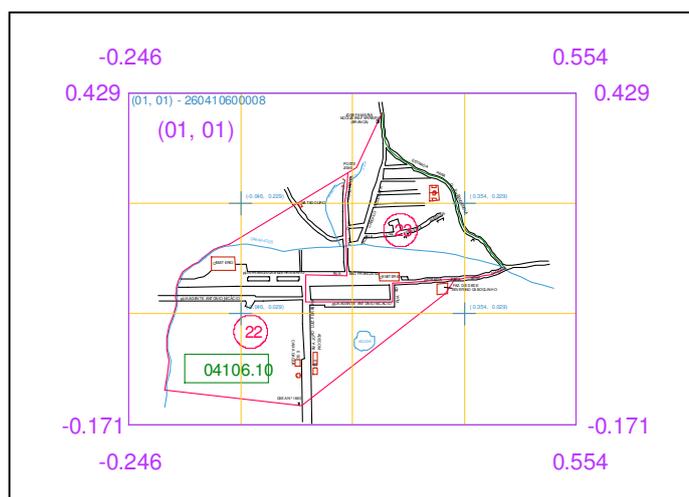


Figura 3 – Setor Censitário do Juá (Fonte IBGE – 2000)

Na figura acima, observa-se que a malha de coordenadas não tem compromisso com a forma nem com a posição real do setor.

Faz-se necessário esclarecer, que a atribuição do IBGE é a precisão das informações descritivas dos setores e não de sua geometria. Desse modo é perfeitamente aceitável, por exemplo, que esses sejam definidos simplesmente pelos arruamentos que os envolvam, quer estes arruamentos tenham posição definida dentro de um sistema de referência previamente especificado, que não.

A falta de compromisso com a geometria dificulta, ou até mesmo impede, que algumas atribuições da Equipe de Saúde da Família sejam implementadas e também inviabilizam o emprego de determinados métodos de análise espacial. Para sanar esses problemas, sugere-se que sejam levantadas as posições espaciais dos indivíduos, localizados pontualmente nos seus domicílios, respeitando um sistema de referência único. Essa solução contorna os problemas de limites e também permite que agregações se processem em todos os níveis hierárquicos do PSF, ou seja, agregando-se família/domicílio chega-se a microárea, agregando-se microárea chega-se à área e assim sucessivamente.

2.2. EQUIPE DE SAÚDE DA FAMÍLIA

A Equipe de Saúde da Família (ESF) está inserida no nível operacional do Programa e é responsável pela coleta e armazenamento das informações mais desagregadas. De modo geral é recomendado que cada ESF assista de 600 a 1000 famílias, com limite máximo de 4500 habitantes. Cada **Agente Comunitário de Saúde (ACS)** da equipe cobre entre 200 a 250 famílias. Tais valores podem variar, função das realidades geográfica, econômica e sócio-política da área, levando-se sempre em conta a densidade populacional e a facilidade de acesso à Unidade.

Recomenda-se que a ESF seja composta, no mínimo, por um médico, um enfermeiro, um auxiliar de enfermagem e entre quatro a seis Agentes Comunitários de Saúde, cujas atribuições básicas são:

- Conhecer a realidade das famílias pelas quais são responsáveis, com ênfase nas suas características sócio-econômicas, psicoculturais, demográficas e epidemiológicas;
- Identificar os problemas de saúde mais comuns e situações de risco aos quais a população está exposta;

- Programar atividades e reestruturar o processo de trabalho;
- Executar, de acordo com a qualificação de cada profissional, os procedimentos de vigilância à saúde e de vigilância epidemiológica, nos diversos ciclos da vida;
- Atuar no controle de doenças transmissíveis como a tuberculose, a hanseníase, as DSTs e AIDS, de doenças infecto-contagiosas em geral, das doenças crônico-degenerativas e de doenças relacionadas ao trabalho e meio ambiente;
- Promover, através da educação continuada, a qualidade de vida e contribuir para que o meio ambiente torne-se mais saudável; entre outras.

Para que as atividades da ESF respondam às necessidades de saúde da população de uma área é necessário realizar o planejamento local das ações. Este é composto pelo **diagnóstico da realidade**, pela **programação das atividades** e por um **acompanhamento e avaliação** permanentes.

2.2.1. DIAGNÓSTICO DA REALIDADE

O diagnóstico da realidade visa a identificação dos problemas existentes, suas causas e riscos e escolha de prioridades, sendo executado a partir da coleta e da análise dos dados.

A coleta dos dados ocorre através de entrevistas com os membros da família, associada a informações obtidas pela observação da realidade e por questionamentos a outras instituições que atuam na área. Todas as famílias são cadastradas pelo ACS com base em roteiro pré-estabelecido para as entrevistas. Os dados referem-se a aspectos demográficos, sócio-econômicos, sócio-culturais, do meio ambiente e sanitário, são armazenados em mídia mecânica (fichas em papel) e incluem o seguinte conjunto de informações, que não representam um limite máximo de conteúdo:

- **Demográficos**
 - Endereço
 - Idade
 - Sexo

- Origem dos membros da família (nascimento e movimentação geográfica)
- **Sócio-econômicos**
 - Renda
 - Condições de trabalho, ocupação, estudo dos membros da família
 - Condições de moradia (tipo de habitação, número de cômodos ou peças, energia elétrica e saneamento básico – abastecimento, tratamento e armazenamento da água, destino dos dejetos e do lixo)
 - Escolaridade
 - Comunicação (telefone, televisão, rádio e outros)
 - Serviços de saúde que normalmente acessam
 - Transporte que mais utiliza
- **Sócio-culturais**
 - Estrutura familiar (composição, papéis, hierarquia)
 - Religião
 - Participação em grupos comunitários
- **Meio Ambiente**
 - Condição das águas (rios, lagoas e mar)
 - Condição da terra (solo, vegetação e uso de agrotóxicos)
 - Condição do ar (poluição)
 - Condição do clima (umidade e temperatura)
- **Mortalidade e Morbidade**
 - Mortalidade (óbitos na família, grau de parentesco, residentes com a família, causa referida e data)
 - Morbidade e condições de risco (gestação, alcoolismo, fumo, dependência de outras drogas, deficiência física e mental, hipertensão arterial, diabetes, distúrbios mentais, hanseníase, epilepsia, tuberculose, malária, doença de chagas e outras condições prevalentes na área).

É pertinente observar, que a quantidade de informações que deveria ser coletada é significativamente extensa e rica em conteúdo, contemplando dados referentes ao indivíduo, à família e ao ambiente. Destaca-se que os dados referentes ao indivíduo podem ser trabalhados em dois níveis, o ecológico, caso sejam agregados, e o individual propriamente dito. Os

dados da família/domicílio e ambiente também podem sofrer agregações sucessivas. Estas considerações são importantes para o delineamento de estudos epidemiológicos baseados nos dados do PSF.

Sobre os dados de ambiente é oportuno observar que além de serem de difícil obtenção, não são especificadas suas variáveis determinantes. Citando-se, por exemplo, o caso de condições das águas, seria necessário que fossem definidos, a priori, os parâmetros e respectivas unidades de medida capaz de permitir-lhes classificação ou distinção. Segundo opinião desse pesquisador essa tarefa não é da alçada da ESF. Vale entretanto ressaltar, que caso tais informações sejam disponibilizadas por órgãos competentes, devidamente georreferenciadas, poderão ser empregadas em análise espacial, por meio de combinação de planos de informação⁴.

Sobre o dado ao nível de indivíduo, verifica-se uma desagregação ainda maior que a prevista anteriormente que era família/domicílio. Desse modo, pode-se dizer que um conjunto de indivíduos compõe uma família e segue-se a hierarquia já definida. Cabe aqui a observação que esse nível de informação pode ser agregado com outros sistemas de informação em saúde, como por exemplo o SIM, SINAN, SINASC, SIH, entre outros. Isso é possível a partir da criação de elos⁴, dado que todos operam com a mesma unidade que é o indivíduo.

Terminada a coleta, a equipe deve proceder à **análise dos dados**, que é composta pela **consolidação dos dados** e pelo **mapeamento da área de atuação**. A descrição e especificação dessas atividades, constantes dos Cadernos de Atenção (MS, 2000), estão relativamente obscuras. Também foi constatado em trabalhos de campo que elas não são implementadas do modo como é proposto. Baseado nesses fatos, o que o que será descrito a seguir tem interpretação desse pesquisador.

É definido que a **consolidação dos dados** tem por objetivo convertê-los em informações passíveis de análise, que são as seguintes:

- Distribuição da população por faixa etária, por sexo e sua distribuição geográfica;
- Condições de moradia, inclusive os aspectos ligados ao saneamento básico;
- Grau de escolaridade e o tipo de ocupação dos membros da área;

⁴ Ver Capítulo 3

- Localização dos recursos hídricos, de elementos poluentes e de outras informações importantes para o conhecimento da realidade local;
- Identificação das lideranças locais, dos grupos religiosos e das formas de organização social da realidade.

Partindo-se do princípio que consolidação é a criação de um indicador, seja totalizado ou relacionado à outra variável, verifica-se que o próprio nome da atividade está em desacordo com os objetivos para ela propostos. Na realidade o que se poderia consolidar e o é, são os dados referentes a:

- População por faixa etária, por sexo, por grau de escolaridade e por tipo de ocupação;
- Condições de moradia e aspectos ligados a meio ambiente.

Também é definido que o **mapeamento da área de atuação** deve ser elaborado a partir da consolidação dos dados e se destina a representação em papel das principais informações geográficas (residências, ruas, estradas, estabelecimentos escolares, religiosos, de saúde, rios, montes, córregos e outras), das principais informações populacionais (grupos específicos a serem priorizados), além da localização de gestantes, idosos, desnutridos, microáreas de risco, etc. Esse mapa deve servir de base para a elaboração do roteiro das visitas, para o direcionamento, a execução e o acompanhamento das atividades na comunidade e para a apresentação de mudanças positivas ou negativas, decorrentes das ações e outros fins.

Ao que tudo indica, tal mapeamento seria o resultado da análise e se assim o for, severas críticas merecem destaque, citando-se as seguintes:

- Confeccionar um mapa somente com a representação dos acidentes naturais e artificiais da área é impeditivo para os integrantes da equipe, por exigir conhecimentos técnicos específicos;
- Identificar num mapa informações sobre populações e microáreas de risco exige significativos conhecimentos de técnicas de análise espacial, que não está ao alcance da ESF, até mesmo por falta de recursos computacionais;
- Utilizar um mapa para elaborar roteiro de visitas não se justifica, dado que as agentes são residentes na área de atuação da equipe e sabem por onde e como se deslocar;

- O mapa como sugerido, por si só, não tem condições de servir de base para o direcionamento, execução e acompanhamento das atividades da comunidade. Ele deve ser entendido, como um instrumento a mais para apoio a decisão.

Sobre o que foi apresentado acerca do diagnóstico da realidade observa-se que a coleta dos dados e seu armazenamento se processam conforme o previsto e que a análise resume-se a totalizações e razões dos dados, sobre a população e condições de moradia.

2.2.2. PROGRAMAÇÃO DAS ATIVIDADES

A programação das atividades tem por base o diagnóstico da realidade, correspondendo ao plano de trabalho da equipe a ser desenvolvido num determinado período de tempo. Os resultados esperados dessa fase devem ter condições de identificar o seguinte (MS, 2000):

- Os problemas priorizados, já tendo sido identificados no momento do diagnóstico da realidade;
- Os objetivos a serem atingidos, de acordo com as necessidades da população e as possibilidades da equipe;
- As atividades a serem realizadas e o detalhamento de sua realização;
- As metas a serem atingidas, que representam os objetivos quantificados;
- Os recursos necessários para a realização das atividades;
- O tempo que será despendido com tais atividades, de acordo com um cronograma de trabalho e os resultados esperados.

A programação deve estar sob permanente avaliação pela ESF, comparando-se os resultados obtidos com os objetivos e metas previamente estabelecidas. É sugerido que essa avaliação seja semestral, desde que não ocorra nenhum fato relevante que exija encurtar esse prazo.

Cabe no momento, sintetizar o que é proposto para o diagnóstico da realidade e para a programação das atividades. Essas duas fases são representadas esquematicamente, conforme mostrado na Figura 4, tomando-se por referência o modelo já descrito no Capítulo 1.

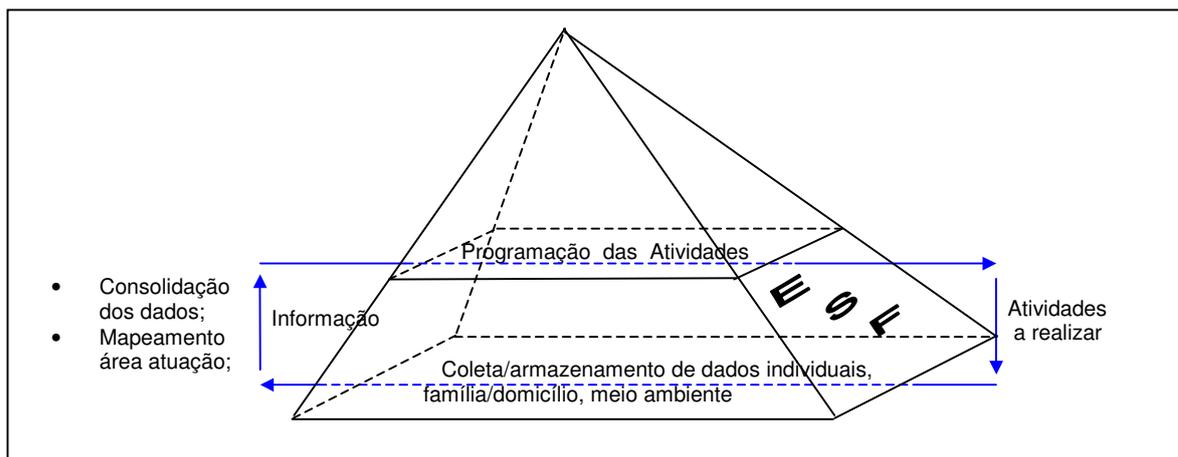


Figura 4 – Estrutura organizacional da ESF

Do ponto de vista teórico, identifica-se uma coerente organização e perfeito funcionamento da ESF, sob a ótica da informação, entretanto, determinadas limitações de ordem prática devem ser consideradas, listando-se a seguir as mais pertinentes:

- A coleta de dados contempla informações sobre o indivíduo, a família/domicílio e meio ambiente, sendo que todos são armazenados em mídia mecânica (fichas em papel), dificultando sobremaneira sua organização, consulta e atualização;
- A consolidação dos dados se processa pela consulta de até 1000 fichas da família/domicílio e 4500 fichas individuais, tornando significativamente complexa e demorada sua execução;
- A confecção do mapeamento da área de atuação, com a riqueza de detalhes e dinamismo sugeridos não é executada.

Dadas essas limitações, verifica-se que a informação disponível é extensa, porém só parte dela é utilizada, impedindo assim uma eficiente programação das atividades.

2.2.3. ACOMPANHAMENTO E AVALIAÇÃO

Os resultados das atividades realizadas pela ACS são avaliadas periodicamente, durante reuniões promovidas pelo instrutor-supervisor, ocasião em que os agentes recebem também novas orientações. A principal fonte de informação para essas avaliações periódicas é o SIAB. Os

relatórios que esse sistema fornece, deveriam permitir conhecer a realidade sócio-sanitária da população acompanhada, avaliar a adequação dos serviços de saúde oferecidos e readequá-los sempre que necessário e por fim, melhorar a qualidade dos serviços de saúde.

Os instrumentos de coleta de dados do sistema (Anexo 1) são as fichas de cadastramento das famílias (Fichas A), de acompanhamento de pacientes (Fichas B e C) e de registro de atividades, procedimentos e notificações (Ficha D). As saídas (Anexo 2) são os relatórios consolidados de situação de saúde e acompanhamento de famílias (SSA), os relatórios consolidados de produção dos serviços, de ocorrência de doenças e/ou de situações consideradas como marcadoras (PMA), ambos com periodicidade mensal e os relatórios consolidados do cadastramento familiar, com periodicidade anual, ou sempre que houver alteração no cadastro familiar. A Figura 5 identifica num diagrama entidade relacionamento, a estrutura desse sistema.

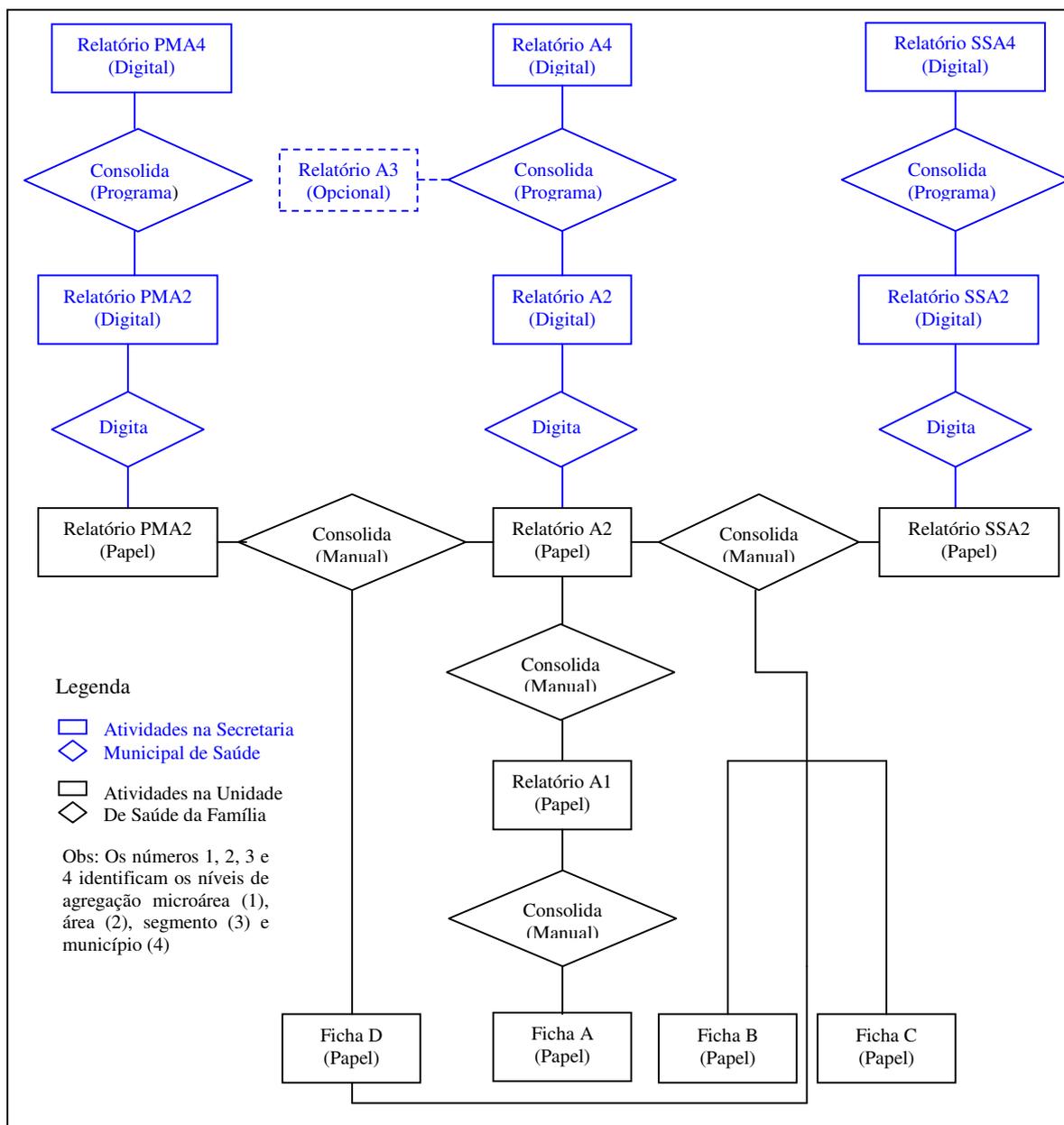


Figura 5 - Estrutura do SIAB

2.3. CRÍTICA AO SIAB

Baseado na estrutura do SIAB merecem destaque as seguintes observações, que são um detalhamento das limitações já levantadas para a fase da programação das atividades:

- Excesso de tempo gasto na USF para consolidação dos relatórios A1, A2, SSA2 e PMA2 e na secretaria de saúde municipal para digitá-las, dada sua periodicidade mensal;

- Disponibilização das informações agregadas por áreas, segmentos (opcional) e municípios em meio digital e por famílias em fichas.
- Ausência da informação espacial, em escala compatível com a família/domicílio, impossibilitando o mapeamento da área de atuação, conforme previsto na fase do diagnóstico da realidade.

Os problemas acima descritos podem ser minimizados ou até mesmo eliminados, desde que as Fichas sejam carregadas em meio digital, diretamente na USF e que na Ficha A sejam acrescentados 2 campos, identificando a posição espacial (coordenadas) dos domicílios. Uma vez efetuadas essas alterações, identifica-se o que se pode chamar de SIAB aprimorado, contextualizado no diagrama mostrado na Figura 6.

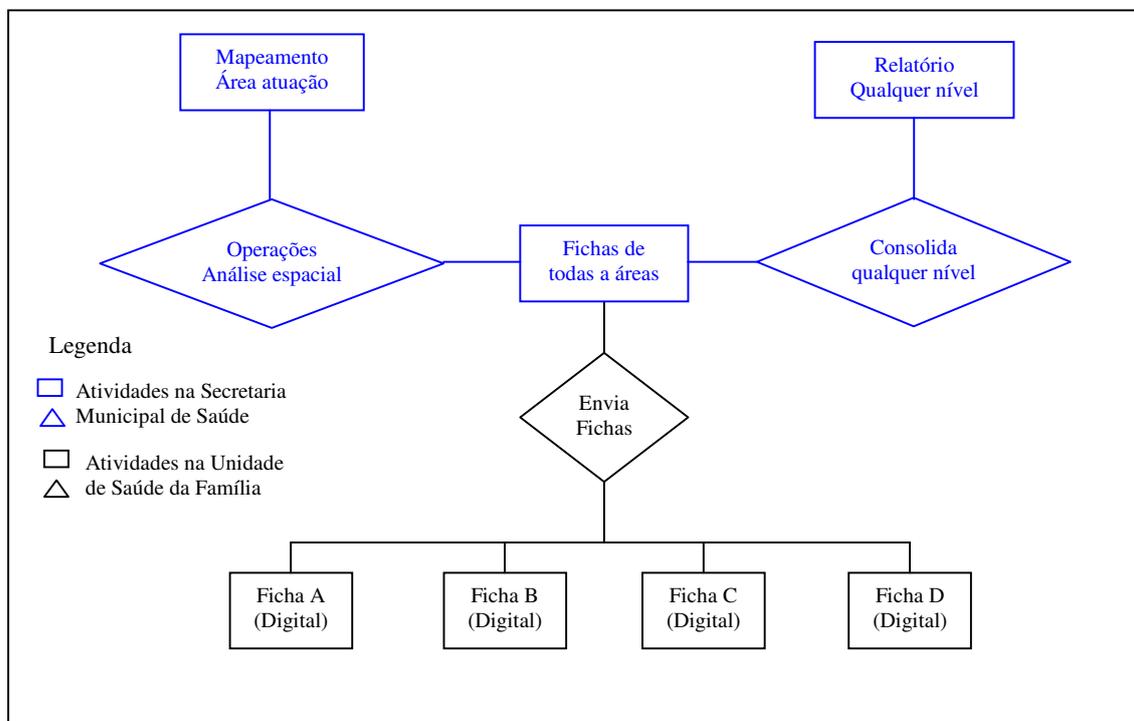


Figura 6 - Estrutura do sistema proposto

Dentro dessa nova ótica que é a de um SIG, são apresentadas as seguintes vantagens em relação ao sistema anterior:

- Adaptação do sistema ao PSF e não do PSF ao sistema;
- Simplificação das tarefas dos encarregados pela carga das informações, com economia de tempo e pessoal;
- Flexibilidade no que diz respeito à capacidade de identificar problemas e recursos relevantes a nível local (Tasca et al., 1993);

- Especificidade e capacidade de discriminação das informações, obtidas através de um enfoque microlocalizado, acerca da distribuição espacial dos próprios problemas e recursos, por sua vez visualizados em mapas e banco de dados informatizados (Tasca et al., 1993);
- Forte capacidade de elaboração de dados (Tasca et al., 1993);
- Gradualidade no desenho e na implementação do sistema (Tasca et al., 1993).

3. SISTEMA DE INFORMAÇÕES GEOGRÁFICAS

Os Sistemas de Informações Geográficas são os que realizam o tratamento computacional dos dados geográficos, armazenando sua geometria e atributos num banco de dados geográficos.

O banco de dados geográfico é controlado por um sistema gerenciador de banco de dados (SGBD), cuja finalidade é armazenar e recuperar as informações de modo eficiente (Date, 1995; Worboys, 1995), além de permitir que tais informações sejam processadas em computador, de forma a espelhar os fenômenos do mundo real em estudo (Burrough, 1998).

Os SGBD empregam diferentes estruturas de organização da informação, chamadas de modelos de dados, sendo esses classificados em **hierárquicos**, **em rede**, **relacional** e **orientado a objetos** (Burrough,1998).

A **estrutura de banco de dados hierárquica** é conveniente quando os dados tem relação de “*um para muitos*” (1:n). Ela assume que cada parte da hierarquia pode ser pesquisada empregando-se uma chave que descreve toda a estrutura. A Figura 7 representa o mapa de uma área, composta de 2 microáreas e sua respectiva codificação.

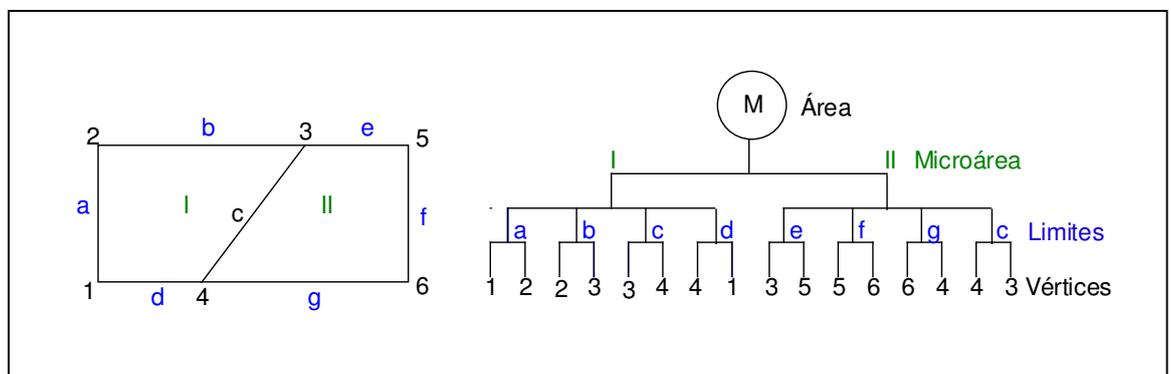


Figura 7 – Estrutura Hierárquica

A estrutura hierárquica tem por vantagens a simplicidade de implementação e a rapidez no acesso aos dados que a compõe. Apresenta como desvantagens a dificuldade de associar seus dados a atributos descritivos e a grande redundância de dados. Esse último aspecto ocorre porque cada vértice é armazenado duas vezes quando definem limites não comuns às microáreas (limites *a,b,e,f,g,d*) e 4 vezes em caso contrário (limite *c*). Outra limitação dessa estrutura é a imposição de que todas as

consultas desejadas devam ser conhecidas *a priori*, de forma a permitir a criação da estrutura. For isso esse modelo não se adequa ao tratamento de dados ambientais.

A **estrutura de banco de dados em rede**, mostrada na Figura 8, é menos rígida que a hierárquica pois permite que se especifique consultas *a posteriori*, uma vez que elas são função das ligações materializadas por ponteiros.

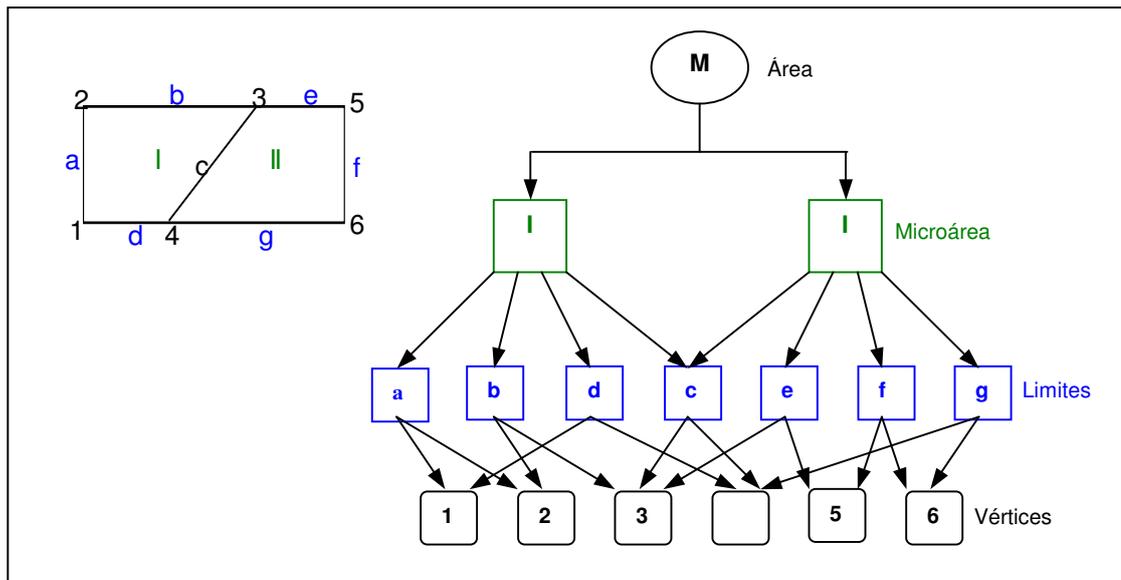


Figura 8 – Estrutura em rede

Observa-se que a redundância é minimizada porque cada vértice é armazenado uma única vez. A generalização cartográfica, exemplificada pela união de microáreas, se processa com relativa simplicidade, bastando eliminar os limites comuns e respectivas ligações. A desvantagem da estrutura é o aumento de informações no banco de dados por conta da grande quantidade de ponteiros.

A **estrutura de banco de dados relacional** não tem hierarquia e não mantém ponteiros. Os dados são armazenados em tabelas bidimensionais, onde cada linha (*registro*) identifica o dado e cada coluna os atributos desse dado (*campo*). As tabelas são conectadas por meio de atributos comuns a elas (*chaves*), conforme mostrado na Figura 9.

Essa estrutura tem por características apresentar grande flexibilidade, além de permitir qualquer tipo de consulta empregando-se lógica booleana e operações matemáticas. A eliminação ou inclusão de dados é facilitada

porque só envolve a eliminação ou inclusão de registros. As operações de conexão entre tabelas se realizam com simplicidade, porém consomem significativo tempo de máquina. Seu emprego é pertinente quando todos os registros têm o mesmo número de campos. Salienta-se, ainda, que tem sido modelo de grande emprego, como ferramenta de comunicação entre o projetista do banco de dados e o usuário final, durante as fases de análise de requisitos e projeto conceitual (Lisboa, 1997).

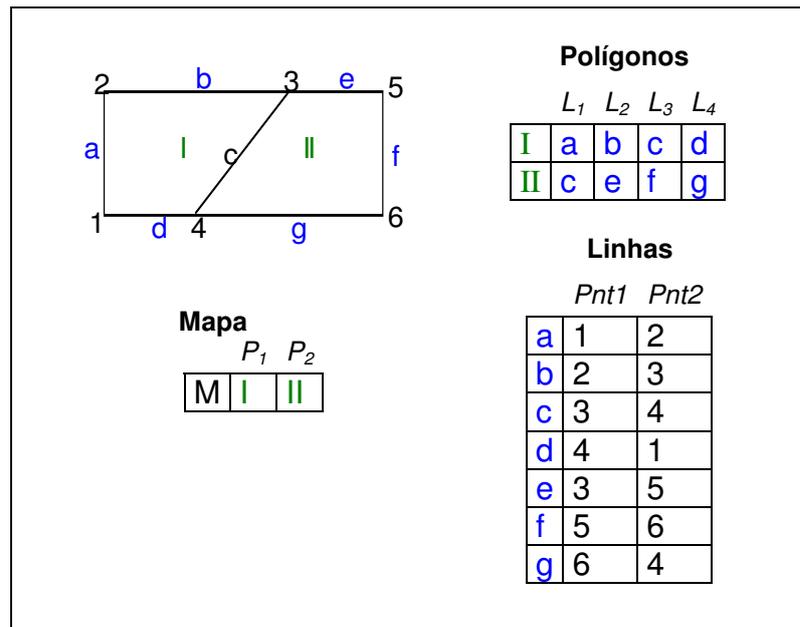


Figura 9 – Estrutura Relacional

Uma extensão dessa estrutura é a chamada **relacional híbrida** (Burrough, 1998), também conhecida por **georrelacional** (Câmara et al, 1996). Nela é permitido gerenciar os dados espaciais separando-os em 2 partes. A primeira trata da representação geométrica e topológica das entidades espaciais e a segunda do manuseio dos atributos descritivos das mesmas. Apresenta as seguintes vantagens:

- Os atributos descritivos não necessitam compor o banco de dados espacial, porém nele podem ser incluídos a qualquer tempo, até mesmo por *Internet*;
- Os atributos descritivos podem ser acessados, eliminados e atualizados, sem necessidade de alterar a geometria e topologia do banco;

- A recuperação dos atributos descritivos pode se processar por intermédio de linguagens padrão, como SQL (*standard query language*), independentemente do SGBD empregado;
- Os atributos descritivos podem ser conectados, por meio de chaves, às suas geometria e topologia.

A **estrutura de dados orientada a objetos** foi originalmente desenvolvida para resolver os problemas de redundância e busca seqüencial que são comuns em bancos relacionais (Burrough, 1998) e também para organizar programas de computador, tomando-se por referência o domínio do problema a ser solucionado (Yourdon & Arguila, 1999). Tal domínio é representado por intermédio de objetos, sendo estes definidos como entidades independentes, assíncronas e concorrentes que têm *estado* (possuem atributos), que encapsulam *serviços* (realizam trabalho) e se *conectam* com outros objetos para executar as funções finais do sistema (Silva, 2002; Yourdon & Coad, 1991).

De acordo com essa abordagem, uma estrutura orientada é representada na Figura 10.

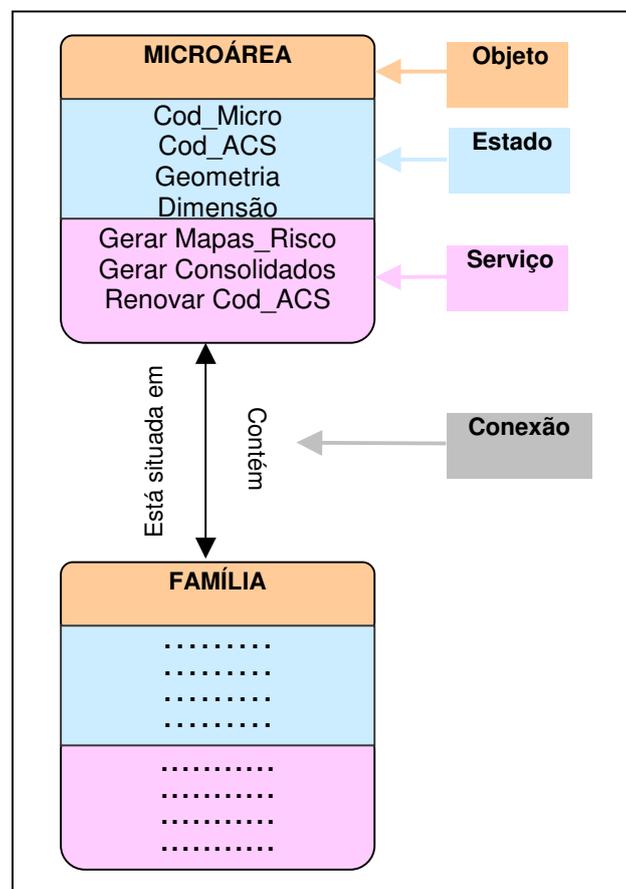


Figura 10 – Estrutura orientada a objetos

A conexão dos objetos tem por finalidade compor relacionamentos e hierarquias. A partir do momento em que estas são estabelecidas, o estado e o serviço fluem, por herança, do nível hierárquico mais alto para o nível mais baixo (Yourdon & Arguila, 1999; Borges, 1997).

A alteração e consulta a um objeto se processam por meio de mensagens, solicitando um de seus *serviços*. A resposta a essas mensagens dependerá do *estado* do objeto.

Em SIG, essa estrutura tem sido estimulada porque os dados geográficos são caracterizados por uma série de atributos e comportamentos (*serviços*) que definem sua dimensão espacial, gráfica, temporal e textual (Worboys, 1994). É importante ressaltar que parte dos atributos deve descrever a natureza geométrica do objeto (ponto, linha, polígono ou célula) e também que mais de um tipo de geometria pode ser empregado, de modo a refletir alterações de forma, quando se altera a escala de representação gráfica. As ligações topológicas entre as classes são estabelecidas explicitamente por meio de ponteiros e operadores como *direção*, *interseção*, *adjacente a*, *direita de*, *esquerda de*, *sobreposição*.

3.1 ESCOLHA DA ESTRUTURA MAIS ADEQUADA

Uma vez definidas as estruturas de dados, cabe identificar a mais adequada para ser empregada no SIG para o PSF. Apresentam-se como candidatas em potencial a estrutura georrelacional e a orientada a objetos. É pertinente, nesse momento, detalhar as características de cada uma delas, para nortear a escolha.

Estrutura georrelacional:

- Os dados espaciais que já compõem o banco de dados podem ser alterados com relativa facilidade;
- A funcionalidade do sistema e a recuperação dos dados são resolvidas pelo SGBD;
- A integração de dados com dados de outros sistemas quando processada por meio de atributos é facilitada;
- As operações de consulta ou modelagem das informações são determinadas pela posição espacial ou pelo valor dos atributos,

limitadas pela funcionalidade do SIG (caso contrário os dados devem ser exportados);

- O manuseio dos dados temporais é limitado.

Estrutura orientada a objetos

- O afastamento semântico entre os objetos do mundo real e sua representação no banco de dados é menor que o existente com o modelo georrelacional;
- As estruturas *raster* e vetorial podem ser integradas no mesmo banco;
- A permuta de dados dos objetos é permitida;
- O espaço de memória consumido em disco é minimizado;
- O emprego de funções definidas pelo usuário é permitido;
- A existência de um modelo único de dados universalmente aceito ainda não está definida. Assim, diferentes bancos de dados apresentam diferentes padrões e tendem a fixar-se numa particular linguagem orientada a objetos;
- A identificação de objetos é por vezes difícil, particularmente quando se trabalha com superfícies contínuas;
- A definição de funções e topologia dos objetos é necessária;
- Não possui padrões como SQL e consultas otimizadas podem se tornar dificultadas, em função da complexidade dos tipos de objeto no sistema.

Comparando-se as duas estruturas, observa-se uma característica extremamente limitante da orientada a objetos para o PSF. Trata-se do armazenamento conjunto do *estado* e *serviços*. Tal fato exige que esses sejam conhecidos, *a priori*. O problema que se apresenta é que os *atributos* não são rigidamente especificados e os *serviços* não estão disponíveis.

De acordo com as atividades previstas para o PSF (MS, 2000), cabe à ESF realizar o planejamento local das ações. Para tanto, é previsto o diagnóstico da realidade, onde aspectos demográficos, sócio-econômicos, sócio-culturais, de meio ambiente e de saúde são coletados. Com base nesses dados, o mapeamento da área de atuação deve contemplar grupos populacionais específicos e microáreas de risco a serem priorizadas. Entretanto, não é definido para cada aspecto coletado o universo de riscos

a ser considerado, o modelo a ser empregado para descrição do risco no espaço e os atributos (variáveis) que devem ser inseridos nesse modelo.

Tomando-se como exemplo a identificação de microáreas de risco, deve-se levar em consideração para o desenvolvimento de um SIG para o PSF o seguinte:

- O programa tem abrangência nacional, havendo grande diversidade entre macrorregiões, estados e municípios onde o mesmo está implantado. Além disso, existe heterogeneidade interna de saúde e de seus determinantes em cada município, entre áreas urbanas e rurais, etc;
- A definição de microáreas de risco embute um conceito coletivo que só pode ser definido a partir do conhecimento dos diversos aspectos que influem na saúde da população da área de atuação do PSF.

Outro fato a ser considerado é não se deseja alterar a estrutura do SIAB, que se encontra implantado nacionalmente e possui uma série histórica, que é relacional e facilmente adaptada a uma estrutura híbrida.

Com base nesses aspectos, optou-se pela estrutura georrelacional, onde os atributos descritivos serão disponibilizados pelo SIAB podendo ser conectados à geometria e topologia das componentes espaciais a serem tratadas.

3.2 Dado Geográfico

O desenrolar desse capítulo se destina a uma breve revisão dos métodos de entrada, armazenamento e análise de dados geográficos, sob a ótica do modelo georrelacional e dos programas de SIG mais utilizados.

Cabe inicialmente definir dado geográfico, também chamado de entidade geográfica (Laurini & Thompson, 1992; Burrough & Frank, 1996).

Essa pode ser entendida como sendo qualquer objeto concreto, que possui as seguintes características: posição geográfica, atributos, relações espaciais e temporais.

- Posição Geográfica

A posição geográfica tem por fim identificar, univocamente, determinada entidade (Tomlin, 1990). Isso pode ser obtido lançando-se sobre a área trabalhada um sistema de coordenadas

perfeitamente definido a partir da especificação dos seguintes parâmetros: posição da origem, orientação dos eixos coordenados e escala de medida (Maling, 1992). Desse modo, qualquer objeto localizado sobre a superfície do terreno será definido por suas coordenadas, conforme mostrado na Figura 11 e terá garantida ou controlada sua geometria. No caso brasileiro, o sistema de referência empregado para o mapeamento sistemático do território é o Universal Transverso de Mercator (UTM).

Deve-se considerar, que caso determinada região tenha sido levantada em sistemas de referência diferentes, é possível, por meio de transformações geométricas, que são operações de rotação, translação e escala (Efimov, 1972), compatibilizá-los num único, de modo a tratar a dita região como um todo. Tais informações estão implementadas na maioria dos programas de SIG e recebem o nome de georreferenciamento.

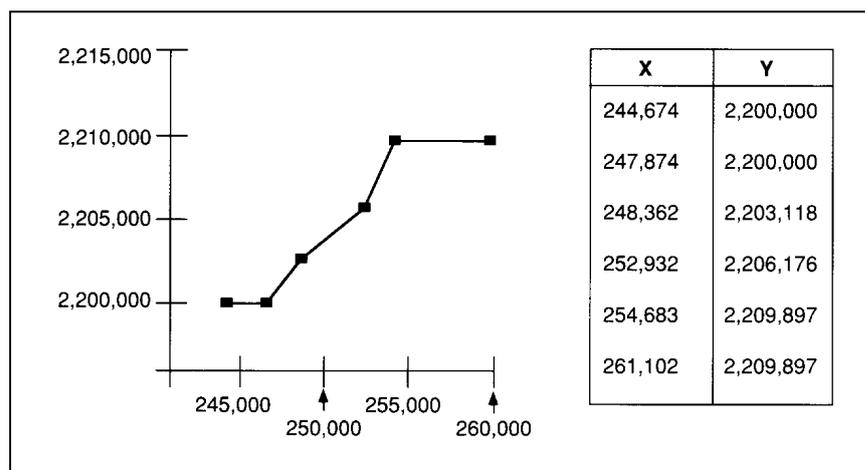


Figura 11 – Sistema de Referência Plano

Atenção especial deve ser dada ao tipo de informação gráfica chamada de *croqui*. Trata-se de um esboço, em breves traços, de um desenho, empregado para posicionar, grosseiramente, uma entidade em relação a outra. Esse documento carece de sistema de coordenadas e por isso tem sua geometria comprometida, impedindo seu georreferenciamento.

- Atributos

Os atributos representam as características descritivas da entidade. São tidos como não espaciais porque, raramente, por si só, representam informações de localização. Tomando-se como exemplo a entidade “residência”, ela pode ter por atributos: tipo de habitação, número de cômodos, energia elétrica, entre outros.

Existem inúmeros atributos para uma entidade, cabendo ao projetista e ao usuário do sistema definirem aqueles necessários à aplicação desejada.

Um modo eficiente de trabalhar com atributos é empregando-se tabelas, conforme mostrado na Figura 12.

Residência	Tipo	Nº Cômodos	Energia Elétrica
A	Taipa	4	Não
B	Madeira	6	Sim

Figura 12 – Atributos de Residência

- Relações Espaciais

As relações espaciais identificam como uma entidade se comporta em relação à outra, segundo alguma propriedade previamente definida. Quando se observa um mapa, essas relações são, muitas vezes, intuitivas e quando se emprega computador, devem ser expressas por meio de proposições.

As relações espaciais são fundamentais para operações de análise e estão baseadas na Teoria dos Conjuntos.

Sejam, por exemplo, os conjuntos $A = \{1,2,3\}$ e $B = \{1,2\}$. Chama-se *produto cartesiano de A em B*, um novo conjunto, cujos elementos são os pares ordenados com o primeiro elemento de A e o segundo de B, indicando-se por $A \times B$. Assim, o produto cartesiano $A \times B = \{(1,1), (1,2), (2,1), (2,2), (3,1), (3,2)\}$. A representação gráfica do produto cartesiano é mostrada na Figura 13.

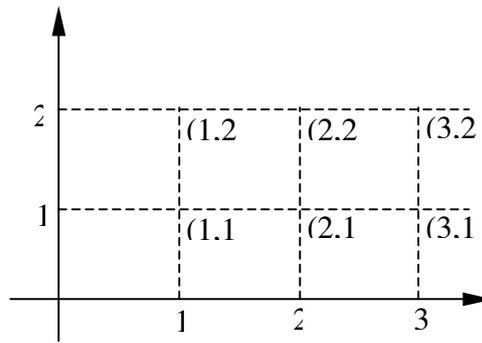


Figura 13 – Produto Cartesiano $A \times B$

Chama-se *relação R de A em B* a qualquer subconjunto do produto cartesiano $A \times B$. Uma relação exige a especificação, a priori, de uma propriedade, também chamada de sentença aberta. Exemplificando, sejam A e B os conjuntos já definidos e a propriedade $x \geq y$, onde $x \in A$ e $y \in B$, então a relação R de A em B é $\{(2,2), (3,1), (3,2)\}$.

Numa situação prática, seja o conjunto $A = \{a_1, a_2, a_3\}$, representando a posição geográfica da residência dos agentes comunitários e $P = \{p_1, p_2, p_3, p_4\}$, representando a posição geográfica da residência dos pacientes, então o produto cartesiano $A \times P = \{(a_1, p_1), (a_1, p_2), (a_1, p_3), (a_1, p_4), (a_2, p_1), (a_2, p_2), (a_2, p_3), (a_2, p_4), (a_3, p_1), (a_3, p_2), (a_3, p_3), (a_3, p_4)\}$. Definindo-se a propriedade $(a-p) \leq d$, onde $(a-p)$ é a distância entre a residência dos agentes e pacientes e d um valor pré-estabelecido, então a relação pode ser útil para selecionar as menores distâncias entre agentes e pacientes. A Figura 14 fornece a interpretação geométrica dessa situação.

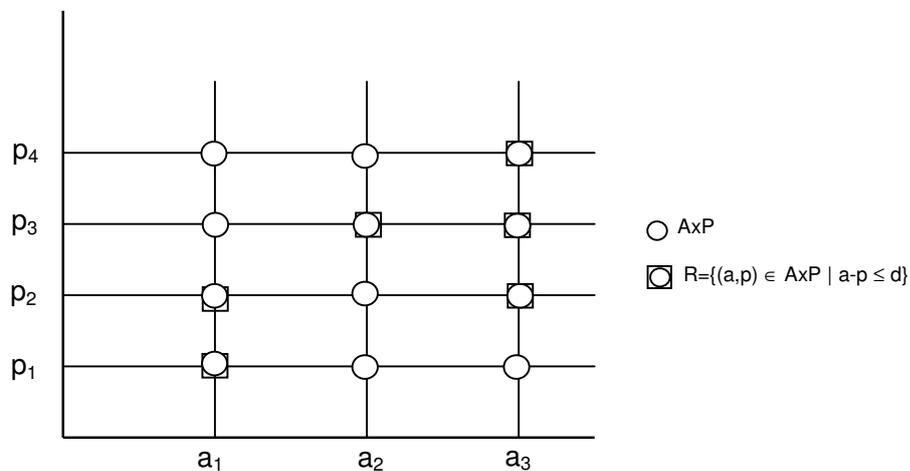


Figura 14 – Produto Cartesiano $A \times P$

- Relações temporais

As relações temporais têm por finalidade identificar quando uma condição da entidade aconteceu e também servir de parâmetro para prever sua ocorrência futura.

3.3 ESTRUTURAS DE ARMAZENAMENTO E ANÁLISE DOS DADOS GRÁFICOS

Os dados gráficos podem ser armazenados em meio magnético de forma explícita ou implícita (Burrough & Frank, 1996). A forma explícita é conhecida por estrutura de dados matriciais ou *raster* e a implícita por estrutura de dados vetorial.

3.3.1 ESTRUTURA DE DADOS *RASTER*

A estrutura de dados *raster* é empregada para representar fenômenos que se distribuem continuamente no espaço, através da discretização deste espaço em uma matriz de células (Holroyd & Bell, 1992; Peuquet, 1994). Cada célula tem sua posição e valor. A posição é definida pela linha e coluna da matriz, e o valor é um número que identifica o fenômeno incidente em toda a célula. Essa estrutura matricial é adequada à manipulação e representação de superfícies tridimensionais. A linha e coluna identificam, univocamente, a célula em 2 dimensões e o valor identifica a terceira dimensão.

O tamanho da célula define sua resolução espacial e deve ser conhecido. A resolução espacial é que vai determinar a menor dimensão da entidade capaz de ser identificada e tem efeitos diretos na estimativa de áreas e perímetros.

Vale ressaltar que nessa estrutura existe somente uma entidade espacial que é a célula. Um ponto é representado por uma célula, uma linha é representada por células de mesmo valor que espalham numa direção definida e uma área é representada por células vizinhas de mesmo valor. A Figura 15 mostra a visualização da estrutura de dados *raster*.

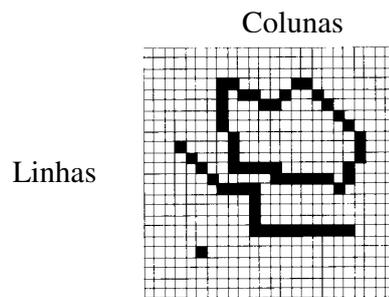


Figura 15 – Estrutura de Dados *Raster*

No que diz respeito à posição geográfica, ela é resolvida através de operações de rotação, translação e escala. Os parâmetros dessas transformações são obtidos relacionando-se a posição de linha e coluna do pixel (sistema de referência antigo) com sua posição no sistema desejado (sistema de referência novo).

As imagens de satélites, obtidas por sensoriamento remoto, são exemplos de produtos que empregam essa estrutura. Elas discretizam determinada superfície do terreno em matriz de células. Cada uma delas captura um número (valor) que representa sua energia refletida, segundo determinado comprimento de onda. A cada valor é associado um tom de cinza ou uma cor, que enviada ao monitor, torna analógica a superfície numérica. O algoritmo, a seguir, identifica esse procedimento.

Leia no cabeçalho número de linhas, número de colunas

Faça i de 1 a número de linhas

Faça j de 1 a número de colunas

Leia valor(i,j)

Associe cor para valor(i,j)

Ascenda Pixel(i,j) com cor

Fim

Fim

A Figura 16 mostra um mapa em estrutura *raster*, detalhando o procedimento de associação Cor-Valor.

MAPA EM ESTRUTURA DE DADOS RASTER

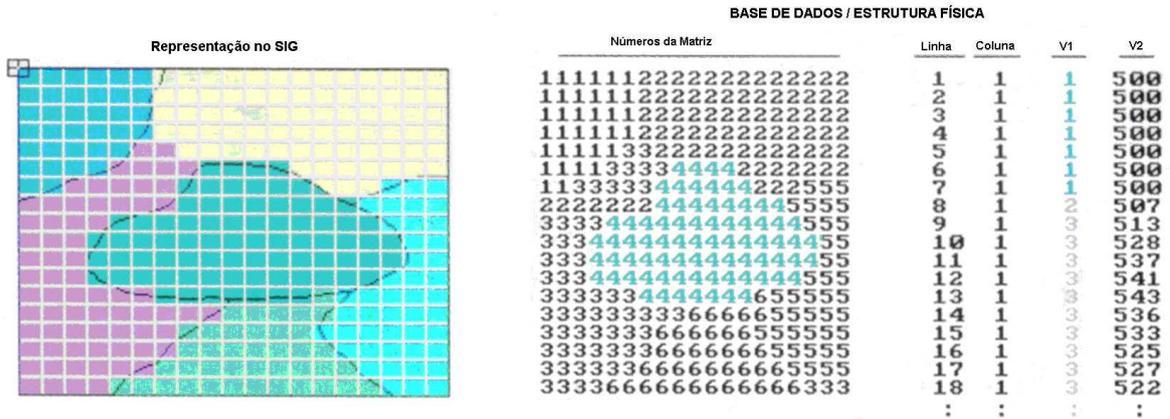


Figura 16 – Associação Cor-Valor

3.3.1.1 ANÁLISE ESPACIAL NA ESTRUTURA RASTER

A análise espacial, com dados na estrutura *raster*, processa-se através de operações matemáticas que podem ser divididas em pontuais e espaciais (Burrough, 1998).

As operações pontuais são aquelas em que células homólogas de diferentes camadas (*overlays*) podem ser logicamente selecionadas, adicionadas, subtraídas ou manipuladas por algum procedimento matemático, função do dado em questão. A Figura 17 representa, esquematicamente, uma operação desse tipo.

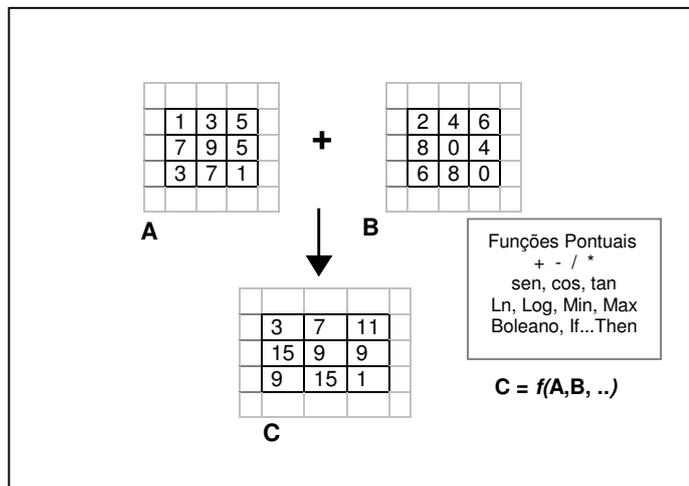


Figura 17 – Operação pontual

Uma aplicação prática de operações pontuais é a obtenção de imagens coloridas que são formadas pela combinação de três camadas (A, B e C), formando uma quarta (camada de saída). Esse tipo de operação facilita a interpretação qualitativa dos fenômenos naturais e artificiais imageados, conforme mostrado na Figura 18.

As operações espaciais são aquelas em que o valor de uma célula é obtido em função dos valores das células vizinhas. Como exemplo dessas operações, cita-se a filtragem espacial, que se caracteriza pela varredura da superfície, por uma janela retangular (filtro), onde um novo valor da célula central da janela $C_{i,j}$ é calculado, em função dos valores das células cobertas pela janela. Essa operação, que pode ser realizada por convolução, cujo modelo é dado pela Equação 1.

$$C_{i,j} = f(\sum_{i-m}^{i+m} \sum_{j-n}^{j+n} c_{i,j} \cdot \lambda_{i,j}) \quad (1)$$

onde f é uma função aplicável à janela de dimensões $2m+1$, $2n+1$ e $\lambda_{i,j}$ são os pesos.

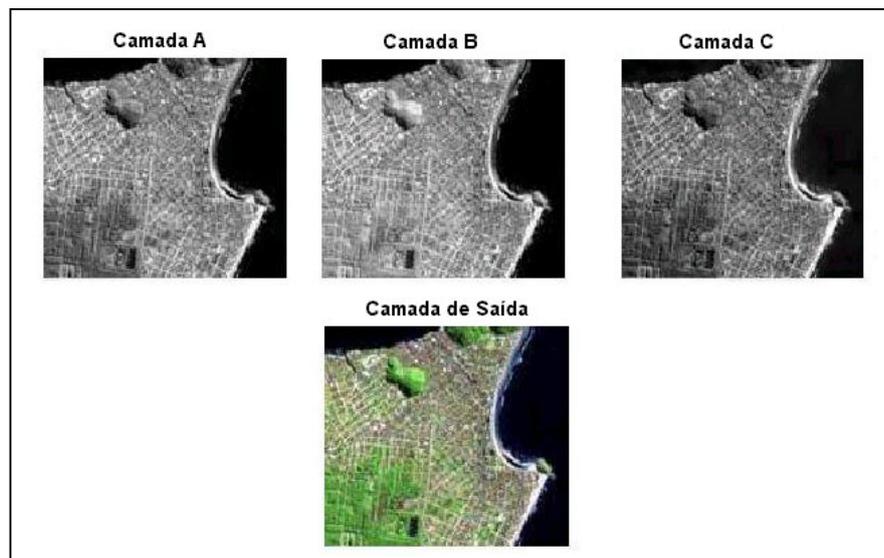


Figura 18 - Combinação de camadas

As funções f mais comumente empregadas são os filtros passa baixa e passa alta. Os filtros passa baixas têm a propriedade de suavizar as regiões de transição da superfície original e os filtros passa altas têm a propriedade de realçar tais regiões. Esses últimos são particularmente interessantes para este trabalho, principalmente por facilitar a identificação de estradas,

massas d'água, áreas de mata, alagados, etc. A Figura 19 mostra os resultados obtidos com um filtro passa altas em uma imagem de satélite.

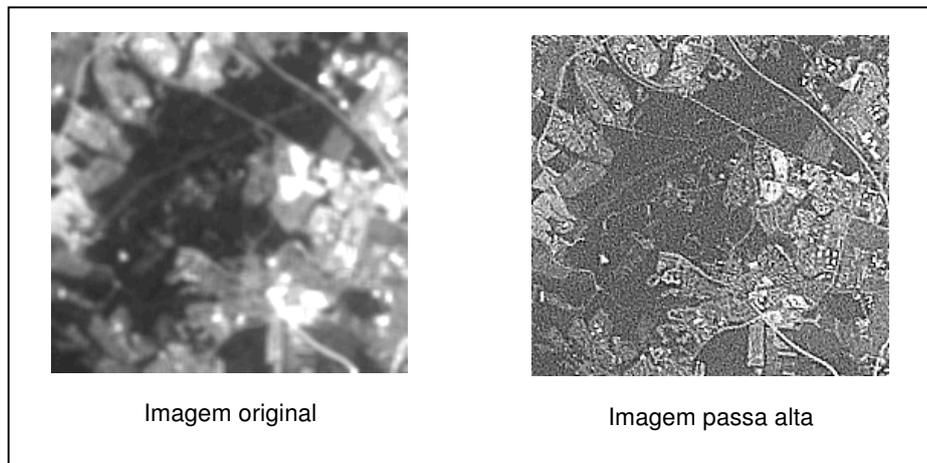


Figura 19 - Filtro Passa Altas

3.3.2 ESTRUTURA DE DADOS VETORIAL

A estrutura de dados vetorial identifica as entidades espaciais através de suas coordenadas, num espaço bi ou tri-dimensional, em um sistema de referência pré-definido (Maguire et al., 1994). Assume-se que o espaço a ser representado é contínuo e a posição dos objetos nele contidos, pode ser definida tão acuradamente quanto se desejar.

Quando se trabalha em duas dimensões, essa estrutura prevê a ocorrência de três entidades espaciais distintas: ponto, linha e polígono (Peuquet, 1994). O ponto tem sua posição definida por um único par de coordenadas (x,y) , a linha por uma lista de coordenadas de pontos que a especificam $((x_1,y_1), (x_2,y_2), \dots, (x_n,y_n))$ e o polígono por uma lista de coordenadas de pontos que definem seu perímetro, de modo que o primeiro par de coordenadas é repetido ao final da lista $((x_1,y_1), (x_2,y_2), \dots, (x_n,y_n), (x_1,y_1))$. A Figura 20 identifica essa estrutura.

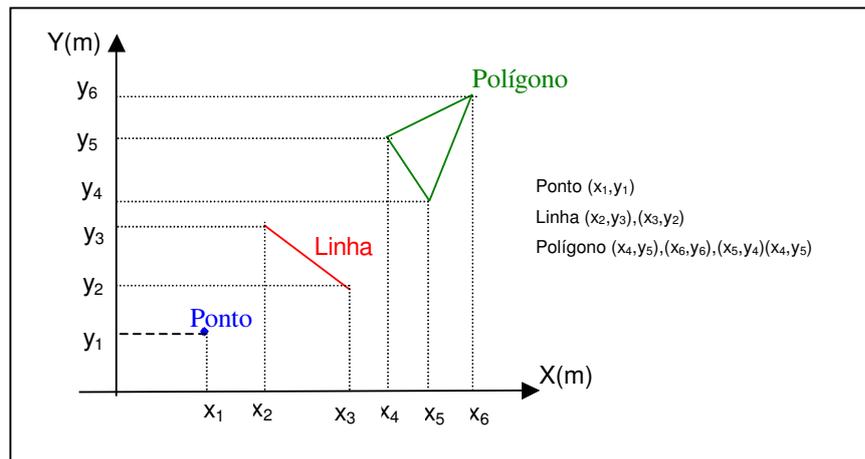


Figura 20 – Estrutura Vetorial

Para a confecção de mapas nessa estrutura, cabe ao projetista associar as feições a serem representadas à uma das 3 entidades, levando-se em consideração a finalidade e escala do mapa. Assim, uma residência pode ser um polígono para uma prefeitura estimar o valor do imposto territorial e um ponto ao estudarmos ocorrência de casos de doenças.

A estrutura vetorial permite o emprego de dois modelos de dados: modelo topológico e o modelo *spaghetti* (Aronoff, 1989).

- Modelo *spaghetti*

No modelo *spaghetti* os pontos, linhas e polígonos são materializados pelas coordenadas de seus vértices, sendo esses armazenadas em arquivos, como uma lista. Caso seja necessária mudança de sistema de referência, é possível por meio de transformações geométricas, atuando-se diretamente sobre as coordenadas dos vértices.

- Modelo Topológico

O modelo topológico emprega conceitos de topologia para descrever relacionamentos entre entidades espaciais, definidas, explicitamente, numa lista ou tabela (Montgomery & Shuch, 1993). Um destes, bastante empregado em SIG, é o chamado “Arco-Nó” (Burrough, 1998, Aronoff, 1989), onde a entidade básica é o arco e são consideradas as seguintes premissas:

- um arco é identificado por uma série de vértices que começam e terminam em um nó;
- um nó é um ponto de interseção de 2 ou mais arcos, ou então, início e fim de um arco aberto;

- nós isolados, não conectados a arcos, representam pontos; e
- um polígono é uma figura formada por uma cadeia de arcos que se fecham.

A Figura 21 esboça um exemplo deste modelo. A topologia é armazenada em 3 tabelas de dados, cada uma para um tipo de elemento espacial (ponto, linha e polígono). Numa quarta tabela são armazenadas as coordenadas dos arcos, tendo por fim garantir a veracidade de suas formas.

A **tabela topológica de polígonos** caracteriza uma definição de área e mostra os arcos que compõem o perímetro de cada um. Por exemplo, o polígono A tem como perímetro os arcos a1, a3 e a5. Polígonos podem ter ilhas em seu interior. O polígono C é uma ilha dentro do polígono B, isto é indicado na lista de arcos do polígono B por um zero, precedendo os arcos que compõem esta ilha. Neste caso, existe apenas um arco, a7, limitante do polígono C. O ponto D contido no polígono B é tratado como um polígono de área zero, composto pelo arco 6 e nó 5. De forma a completar as definições espaciais, deve haver um modo de se referir à área que se encontra do lado de fora do perímetro do mapa. Esta área é o polígono E, para o qual os arcos não estão explicitamente definidos.

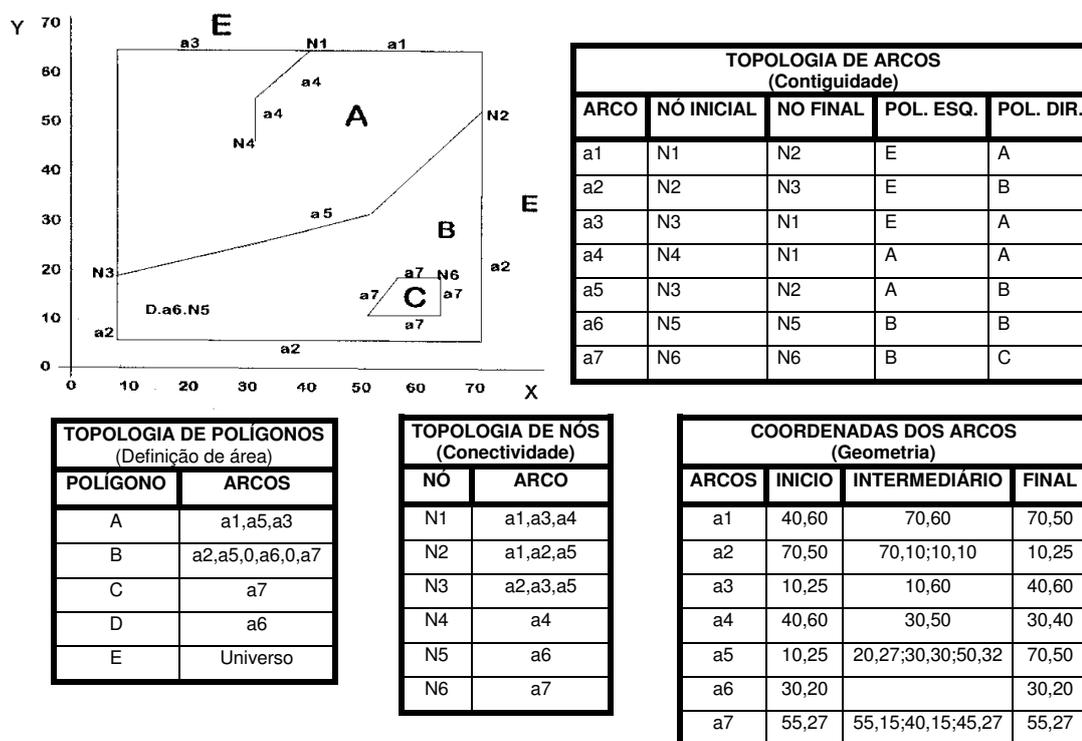


Figura 21 – Topologia Arco-Nó

A **tabela de topologia de nós** caracteriza conectividade, onde cada nó é definido pelos arcos que lhe concorrem. Por exemplo, para o nó N1 concorrem os arcos a1, a3 e a4.

A **tabela de topologia de arcos** caracteriza contigüidade e define as relações entre nós e polígonos aos arcos. Os pontos extremos são identificados pelos atributos de “nó inicial” e “nó final”. Por exemplo, o arco a5 inicia no nó N3 e termina no nó N2. Movendo-se do nó N3 para o N2, o polígono à esquerda é o A, e o polígono à direita é o B.

A partir da topologia podem ser feitas análises das posições relativas dos elementos de um mapa. Por exemplo, todos os polígonos adjacentes ao polígono B podem ser encontrados por uma pesquisa na topologia de arco. Cada polígono que faz par com B, nesta tabela lhe é adjacente, pois ambos tem um arco em comum. Os polígonos A e B são pares no que diz respeito ao arco 5, então A e B são adjacentes.

As tabelas topológicas podem também ser usadas para se encontrar todas as feições contidas num polígono. Isso é feito, consultando-se na tabela topologia de polígonos quais são os arcos precedidos por zero. Exemplificando, o polígono B contém duas ilhas, uma definida pelo arco a6 e outra pelo arco a7. A partir das tabelas topológicas de arcos e nós, identifica-se o arco a6 como um ponto, pois possui os mesmos polígonos tanto à direita quanto à esquerda e apenas 1 nó. O arco a7 compõe um polígono ilha, pois possui polígonos diferentes à esquerda e à direita, apesar de um único nó.

Consultas com dados espaciais podem ser processadas com muito mais velocidade, usando tabelas topológicas do que se fossem realizadas através de cálculos, como no modelo spaghetti.

A geometria das entidades espaciais é obtida da tabela “Coordenadas de Arcos”. Cada arco é representado por um ou mais segmentos de reta definidos por uma série de coordenadas. Caso seja necessária mudança de sistema de referência, é possível através de transformações geométricas, aplicadas diretamente à essa tabela.

3.3.2.1 ANÁLISE ESPACIAL VETORIAL

A análise espacial, com dados na estrutura vetorial, compreende as

seguintes operações: combinação de planos de informação (*map overlay topológico*), geração de áreas de influência (*buffer*), extração de feições e, combinação de feições.

- Geração de área de influência

A área de influência tem por objetivo a determinação de proximidade ou vizinhança espacial. Numa situação prática, considere-se a camada A, na Figura 22, representativa de focos de determinado vetor. Conhecendo-se a distância máxima que esse vetor pode percorrer, identificam-se áreas de risco, sujeitas à sua influência, materializadas pela camada B, na mesma Figura. Esta operação é empregada para obter respostas do tipo: “A feição A está a uma distância D além ou aquém da feição B”.

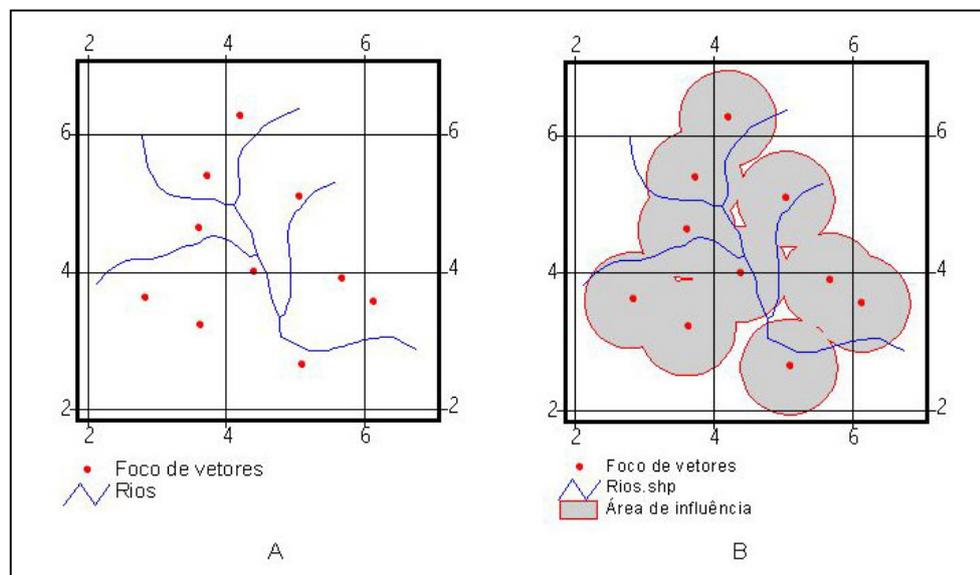


Figura 22 – Área de influência

- Combinação de planos de informação

A combinação de planos de informação produz uma combinação espacial de duas camadas homólogas (entrada), gerando uma terceira (saída), de modo que esta última contenha os atributos descritivos das duas primeiras. São exemplos dessa operação a união, a identidade e a interseção. A união retém na camada de saída todas as feições contidas nas camadas de entrada. A interseção retém na camada de saída somente as feições comuns às camadas de entrada. A identidade é

similar à união, com a diferença que a camada de saída será limitada por uma das camadas de entrada.

Numa situação prática, considere-se a camada A representando distritos sanitários e a camada B representando áreas de risco. A união de A com B gerará uma camada C, identificando regiões de risco e não risco, dentro e fora dos limites dos distritos, conforme mostrado na Figura 23.

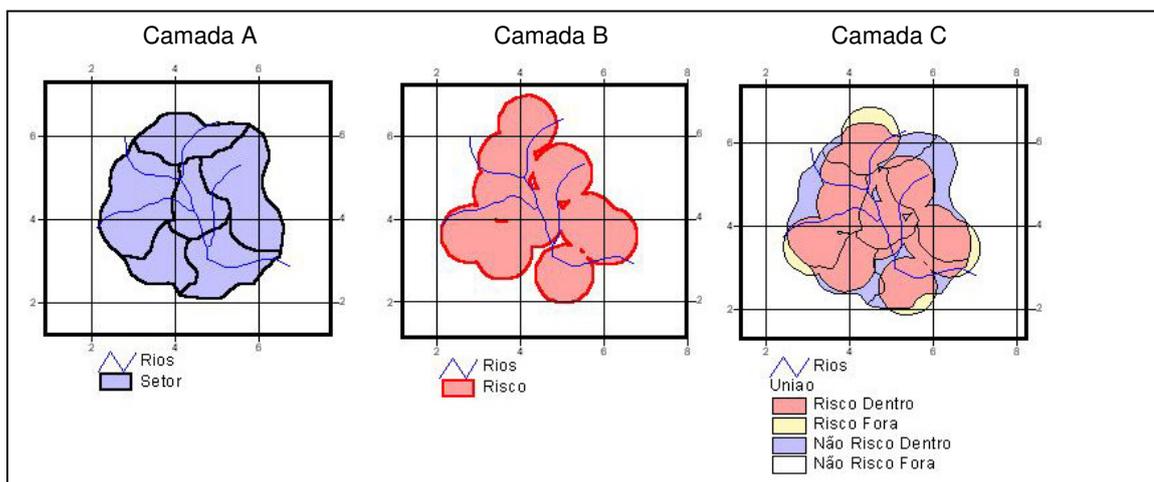


Figura 23 - União

A identidade da camada A com a camada B gerará a camada C, identificando regiões de risco e não risco dentro dos distritos, conforme mostrado na Figura 24.

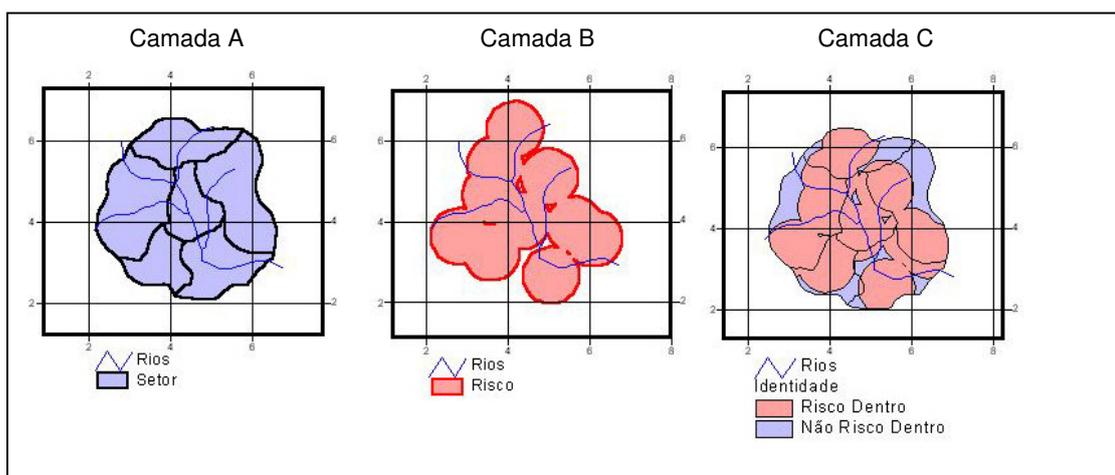


Figura 24 – Identidade

A interseção da camada A com a camada B gerará a camada C, identificando as regiões de risco dentro dos distritos, conforme mostrado

na Figura 25.

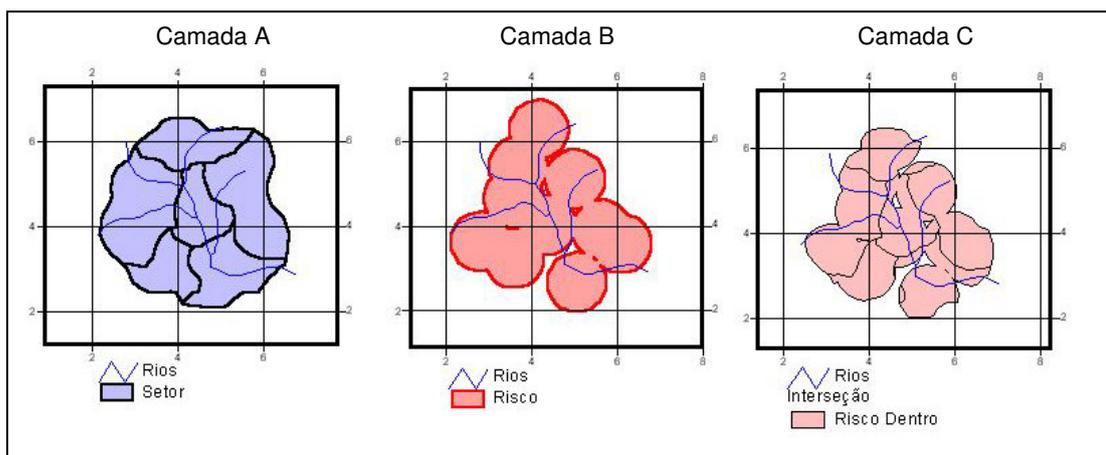


Figura 25 -Interseção

- Extração e combinação de feições

A extração e combinação de feições tem por objetivo selecionar um subconjunto da camada de entrada e enviá-lo à camada de saída. Isso pode ser efetuado por intermédio de algum critério lógico nos atributos da camada de entrada, ou então se definindo uma segunda camada que sirva de recorte para a primeira.

Numa situação prática, considere-se a camada A representativa dos distritos sanitários e hidrografia. Com a operação de combinação de feições pode-se selecionar, logicamente, os distritos banhados e não banhados por rios, criando-se uma camada B, correspondente a esse critério, conforme mostrado na Figura 26.

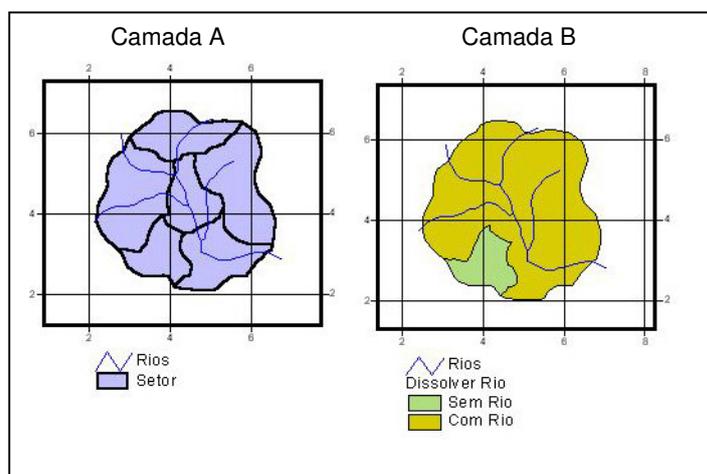


Figura 26 - Combinação de feições

Considere-se agora a camada A representativa dos distritos sanitários e a camada B regiões de risco. A operação de extração gerará a camada C, identificando regiões de não risco dentro dos distritos, conforme mostrado na figura 27.

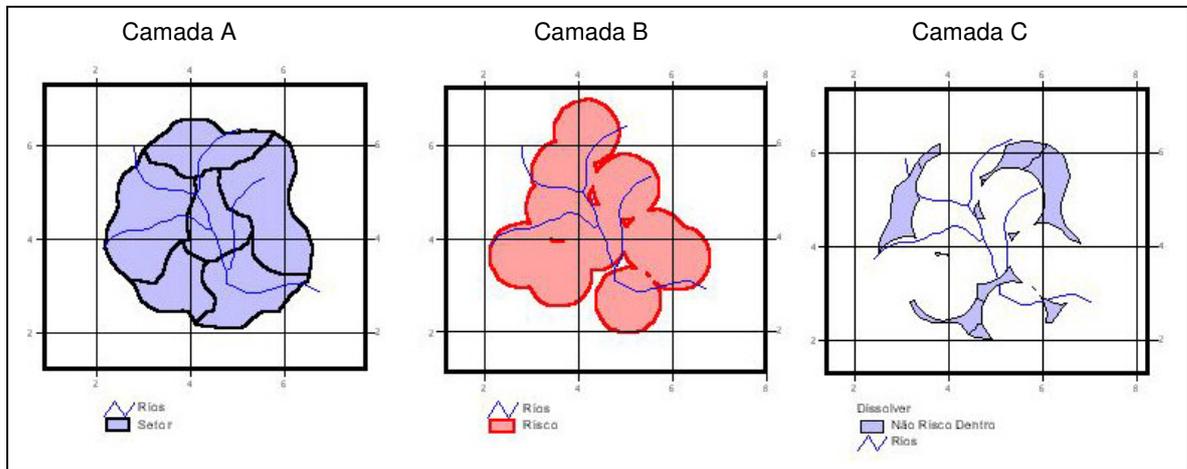


Figura 27 - Extração

3.4 ARMAZENAMENTO DOS ATRIBUTOS EM MEIO MAGNÉTICO

A entrada e o armazenamento dos atributos, também chamados de dados não gráficos, processa-se através de um sistema de gerenciamento de banco de dados (Frank, 1988; Ullman, 1989, Burrough, 1998). Este opera com duas espécies de informação: descrição das entidades e descrição dos relacionamentos. As entidades são descritas por seus atributos e os relacionamentos por uma associação, com algum significado, entre as entidades .

Para uniformização de linguagem, será empregada a terminologia listada abaixo (Furtado & Santos,1989):

- Arquivo: coleção de registros, cada um deles representando um objeto ou entidade;
- Registro: seqüência de itens, sendo cada um deles chamado campo ou atributo. Cada atributo corresponde à uma característica ou propriedade do objeto representado. Cada campo possui um nome, um tipo (cadeia de caracteres, valor numérico, data, valor booleano) e um comprimento (número de palavras, bytes ou bits);
- Chave: seqüência de um ou mais campos em um arquivo;

- Chave Primária: chave que apresenta um valor diferente para cada registro do arquivo, de tal forma que, dado um valor da chave primária, é identificado um único registro do arquivo;
- Chave secundária: chave que identifica um conjunto de registros;
- Chave de Acesso: chave utilizada para identificar o(s) registro(s) desejado(s) em uma operação de acesso a um arquivo;
- Argumento de Pesquisa: é o valor da chave de acesso em uma operação;
- Chave de um Registro: é o valor de uma chave primária em um particular registro do arquivo;
- Chave de Ordenação: é a chave primária usada para estabelecer a seqüência na qual devem ser dispostos os registros em um arquivo;
- Elo: ligação entre dois registros de um mesmo arquivo ou de arquivos diferentes. Existem dois tipos de elos, os implícitos e os explícitos.
 - O elo implícito existe quando há um item (ou seqüência de itens) em um dos registros, cujos valores são comparados aos de um item em outro registro, conforme mostrado na Figura 28. A comparação mais usual é a igualdade, ou seja, no arquivo **Moradores**, um registro tem no item correspondente à Microárea o valor “4758” e no arquivo **Agentes**, um registro tem no item correspondente à Atuação o valor “4758”.

Moradores		Agentes	
Nome morador	Microárea	Atuação	Nome agente
João	4758	4758	Ana

Figura 28 – Elo implícito

- O elo explícito pode ser implementado de várias maneiras, uma delas é esquematizada na Figura 29. Supõe-se que uma das ocorrências do valor comum é substituída por algum tipo de referência ao outro registro. Esta referência pode tomar diversas

formas, como o endereço físico em que se encontra o registro. Em qualquer caso, representar-se-á uma referência por uma seta.

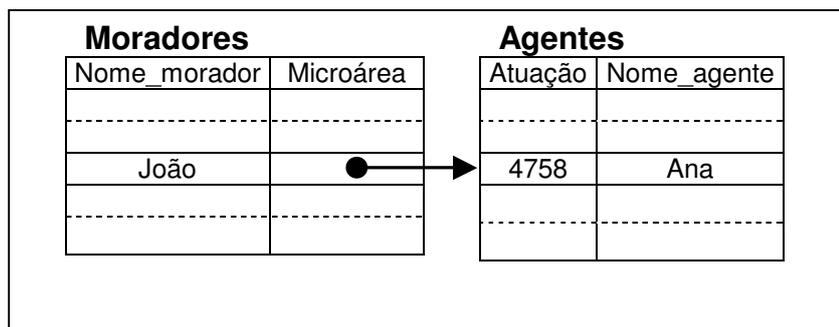


Figura 29 – Elo explícito

Exemplificando a terminologia citada, seja o arquivo “Dados.xxx” mostrado na Figura 30, nessa situação tem-se:

- Arquivo: Dados.xxx;
- Registro: Cada uma das linhas 1,2,3,4 e 5;
- Campos: PESSOA, 25 (tamanho em caracteres) ,C (tipo caracter);
ENDEREÇO, 30 (tamanho em caracteres), C (tipo caracter);
CÓDIGO, 1 (tamanho em caracteres), C (tipo caracter);
IDENTIFICADOR, 4 (tamanho em caracteres), N (tipo numérico);
- Chave Primária: IDENTIFICADOR;
- Chave Secundária: PESSOA ou ENDEREÇO ou CÓDIGO;
- Chave de Acesso: PESSOA ou ENDEREÇO ou CÓDIGO ou IDENTIFICADOR;
- Argumento de Pesquisa: Conteúdo da chave de acesso escolhida;
- Chave de um registro: Conteúdo do Campo IDENTIFICADOR;
- Chave de Ordenação: IDENTIFICADOR;
- Elo: CÓDIGO no arquivo Dados.xxx e CÓDIGO no arquivo Codificação.xxx (elo implícito).

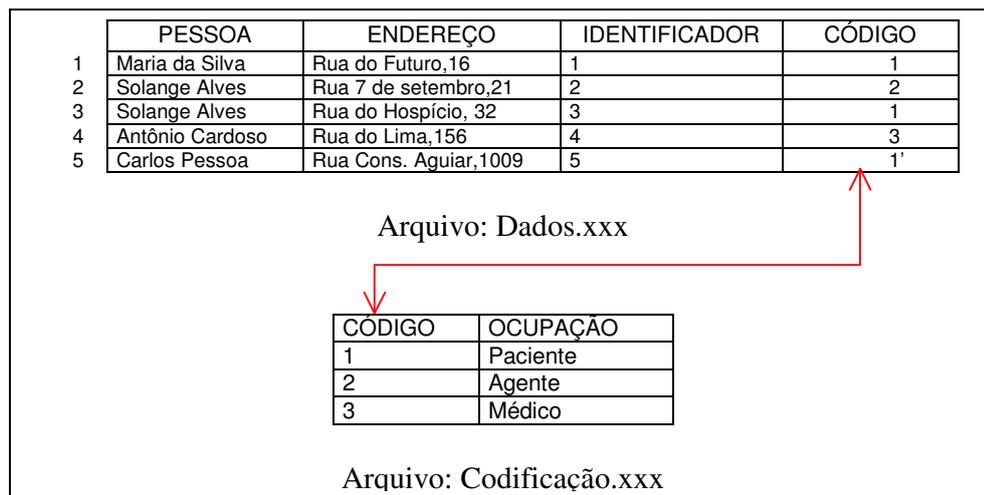


Figura 30 – Arquivos de Dados

Os SGBD empregam um modelo de banco de dados que tem por finalidade organizar logicamente as informações (Heusser, 2000). Esses modelos distinguem-se, essencialmente, pelo tipo de elo utilizado e neste trabalho, será tratado o relacional (Kroenke, 1988).

3.4.1 MODELO DE BANCO DE DADOS RELACIONAL

Antes de detalhar o modelo, convém uma breve introdução sobre sua base conceitual. Identifica-se na Figura 31, 4 conjuntos: NOMES, IDENTIFICADORES, CÓDIGO OCUPAÇÃO e OCUPAÇÃO.

O domínio do conjunto NOMES é formado por todos os nomes possíveis de pessoas e seu valor de domínio é formado por {Maria, Solange, Carlos, Paulo}.

O domínio do conjunto IDENTIFICADORES é formado por todos os valores numéricos possíveis que representam identificadores e seu valor de domínio é formado por {35, 48, 32, 28}.

O domínio do conjunto CÓDIGO OCUPAÇÃO é formado por todos os valores numéricos possíveis que representam código de ocupação e seu valor de domínio é formado por {1, 2, 3}.

O domínio do conjunto OCUPAÇÃO é formado por todas as ocupações possíveis e seu valor de domínio é formado por {Médico, Agente, Paciente}.

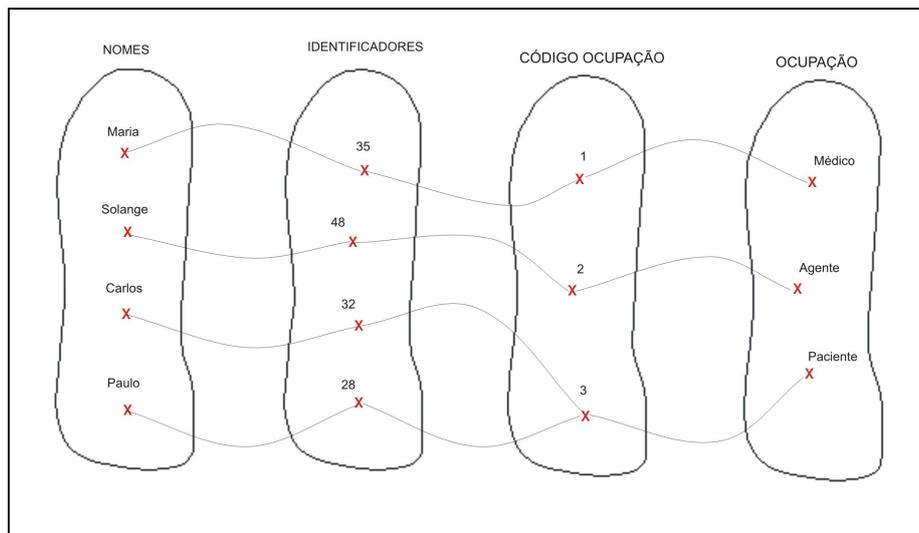


Figura 31 - Relação

As linhas identificam como os diferentes valores de domínio se associam.

Diante do exposto, verifica-se que os domínios, os valores de domínios e as ligações podem ser armazenados na forma de tabelas que aqui recebem o nome de relações. Os conjuntos e respectivas associações representados na Figura 31 são agora identificados pela *Relação A*, mostrada na Figura 32, sendo esse o princípio do modelo relacional. Cabe ainda registrar a seguinte correspondência de terminologia: arquivo com relação, registro com tupla, item com domínio, valor de item com valor de domínio.



Figura 32 – Relação A

Sob a ótica de análise e projeto de sistemas, identifica-se o modelo de dados relacional como um modelo conceitual de dados no qual estes são representados por uma coleção de relacionamentos e elos implícitos (Elmasri & Navathe, 2000).

Na terminologia desse modelo, são considerados conjuntos de valores chamados domínios. Se D_1, D_2, \dots, D_n são domínios, define-se uma “Relação R”, como um sub-conjunto do produto cartesiano dos domínios, ou seja, $R \subseteq D_1 \times D_2 \times \dots \times D_n$. Os resultados da relação são suas n -tuplas ou simplesmente tuplas, que são constituídas de valores de domínios.

A Figura 33 mostra a representação das relações Atendimentos_Antigos (Relação F) \subseteq Código do Profissional (Cod_Prof) x Nome do Profissional (Nome_Prof) x Cadastro da Família (Cad_Fam), Características da Residência (Relação P) \subseteq Cad_Fam x Endereço (End) x Número de Cômodos (Nr_Com) x Tipo de Revestimento (Tipo_Rev) e Atendimentos Novos (Relação G) \subseteq Cod_Prof x Nome_Prof x Cad_Fam. Verifica-se que cada linha numerada corresponde a uma tupla e os valores originários de diferentes domínios aparecem em colunas designadas pelos nomes dos domínios.

	Cod_Prof	Nome_Prof	Cad_Fam
1	12	Ana	109
2	17	Eva	201
3	17	Eva	202
4	1	Ada	109
5	1	Ada	202

	P#	End	Nr_Com	Tipo_Rev
1	109	a	4	Taipa
2	201	b	6	Madeira
3	202	c	3	Taipa

	Cod_Prof	Nome_Prof	Cad_Fam
1	12	Ana	109
2	1	Ada	201

Figura 33 – Relações

Os relacionamentos entre entidades aparecem de duas maneiras:

- As entidades aparecem juntas em tuplas da mesma relação, como em F o profissional “12” e a família “109”;
- Há um elo (implícito) através da ocorrência de valores comparáveis nos domínios de duas relações. Assim obtém-se o relacionamento entre o profissional “12” e o endereço “a” associando-se a primeira tupla de F com a primeira tupla de P, através do valor comum “109”.

A manipulação do Banco de Dados oferece as possibilidades da Álgebra Relacional e consta de oito operações, descritas adiante. A Figura 33 será tomada como referência para exemplificação. Ressalta-se que o resultado de uma operação da álgebra relacional é sempre uma nova relação, que, neste caso, será chamada de T.

- União ($T \leftarrow F \cup G$): o resultado T é uma relação que contém todas as tuplas de F e G, sem repetição, conforme mostrado na Figura 34. Essa operação só é possível em relações união-compatíveis (relações que tem o mesmo domínio). Pode ser empregada, por exemplo, para responder a seguinte pergunta: *Quais são os atendimentos novos e antigos?*

T	Cod_Prof	Nome_Prof	Cad_Fam
1	12	Ana	109
2	17	Eva	201
3	17	Eva	202
4	1	Ada	109
5	1	Ada	202
6	1	Ada	201

Figura 34 – União

- Interseção ($T \leftarrow F \cap G$): o resultado T é uma relação que contém apenas as tuplas que aparecem em F e G, conforme mostrado na Figura 35. Essa operação só é possível em relações união-compatíveis. Pode ser empregada, por exemplo, para responder a seguinte pergunta: *Quais dos atendimentos antigos foram feitos novamente?*

T	Cod_Prof	Nome_Prof	Cad_Fam
1	12	Ana	109

Figura 35 – Interseção

- Diferença ($T \leftarrow F - G$): o resultado T é uma relação que contém as tuplas que aparecem em F mas não em G, conforme mostrado na Figura 36. Essa operação só é possível em relações união-compatíveis. Pode ser empregada, por exemplo, para responder a seguinte pergunta: *Quais dos atendimentos antigos não foram repetidos?*

T	Cod_Prof	Nome_Prof	Cad_Fam
1	17	Eva	201
2	17	Eva	202
3	1	Ada	109
4	1	Ada	202

Figura 36 – Diferença

- Produto Cartesiano ($T \leftarrow G \times P$): o resultado T é uma relação que contém todas as tuplas de G concatenadas com as tuplas de P, conforme mostrado na Figura 37. Pode ser empregada, por exemplo, para responder a seguinte pergunta: *Qual a informação existente sobre atendimentos novos e características da residência?* Ressalta-se que o produto cartesiano não costuma ser usado isoladamente. Neste exemplo, apenas daria toda a informação em G e toda a informação em P, de forma repetitiva e inconveniente.

T	Cod_Prof	Nome_Prof	Cad_Fam	P#	End	Nr_Com	Tipo_Rev
1	12	Ana	109	109	a	4	Taipa
2	12	Ana	109	201	b	6	Madeira
3	12	Ana	109	202	c	3	Taipa
4	1	Ada	201	109	a	4	Taipa
5	1	Ada	201	201	b	6	Madeira
6	1	Ada	201	202	c	3	Taipa

Figura 37 – Produto Cartesiano

- Projeção ($T \leftarrow F [\text{Cod_Prof}, \text{Cad_Fam}]$): o resultado T é uma relação que contém apenas os domínios indicados, no caso Cod_Prof e Cad_Fam. As tuplas em duplicata que seriam criadas, por não serem considerados os outros domínios, são eliminadas. A Figura 38 identifica essa operação. Pode ser empregada, por exemplo, para responder a seguinte pergunta: *Quais profissionais trabalham para quais famílias?*

T	Cod_Prof	Cad_Fam
1	12	109
2	17	201
3	17	202
4	1	109
5	1	202

Figura 38 - Projeção

- Restrição ($T \leftarrow F [\text{Cod_Prof} = 1]$): o resultado T é uma relação que contém as tuplas de F que satisfazem a comparação indicada, conforme mostrado na Figura 39. As comparações são efetuadas empregando-se operadores de comparação (=, ≠, ≤, <, ≥, >). Pode ser empregada, por exemplo, para responder a

seguinte pergunta: *Quais os atendimentos efetuados pelo profissional "1"?*

T	Cod_Prof	Nome_Prof	Cad_Fam
1	1	Ada	109
2	1	Ada	202

Figura 39 - Restrição

- Junção ($T \leftarrow G [Cad_Fam = P\#] P$): o resultado T é uma relação que contém concatenadas as tuplas de G e P que satisfazem a comparação sobre os domínios indicados, conforme mostrado na Figura 40. A comparação pode ser efetuada empregando-se os comparadores utilizados para a restrição. Pode ser empregada, por exemplo, para responder a seguinte pergunta: *Quais os novos atendimentos executados por cada profissional?* Observa-se que a junção é sub-conjunto do produto cartesiano.

T	Cod_Prof	Nome_Prof	Cad_Fam	P#	End	Nr_Com	Tipo_Rev
1	12	Ana	109	109	a	4	Taipa
2	1	Ada	201	201	b	6	Madeira

Figura 40 – Junção

- Divisão: Essa operação é mais complexa que as anteriores. Para melhor entendimento, reporte-se à Figura 41. Seja A uma relação binária com atributos x e y, e B uma relação unária com atributo z, com y e z definidos sobre o mesmo domínio. Define-se a operação divisão T, como sendo o conjunto dos elementos x, com pares (x,y) pertencentes a A para todos os valores y pertencentes a B. Pode ser empregada, por exemplo, para responder a seguinte pergunta: *Quais os profissionais que atenderam todas as famílias?*

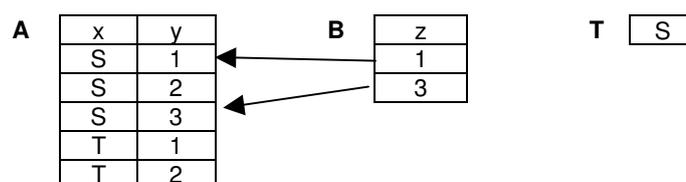


Figura 41 – Divisão ($T \leftarrow A [y \div z] B$)

3.5 LIGAÇÃO DOS DADOS GRÁFICOS AOS ATRIBUTOS

A ligação das entidades geográficas aos atributos levará em consideração as seguintes premissas:

- Emprego da estrutura de dados gráficos vetorial;
- Emprego do modelo de dados gráficos topológico;
- Emprego do modelo de dados relacional.

É conveniente lembrar que a estrutura vetorial representa as entidades geográficas por pontos, linhas e polígonos, num sistema de referência em duas dimensões; o modelo de dados topológico identifica cada entidade geográfica (ponto, linha ou polígono) por uma chave primária; o modelo relacional prevê a existência de elo implícito.

Para melhor compreensão do processo, reporte-se à Figura 42. Nela está esquematizada determinada região formada por polígonos, a correspondente tabela (relação) Topologia de Polígonos, a tabela (relação) Atributos de Polígonos e a tabela (relação) Coordenadas de Arcos.

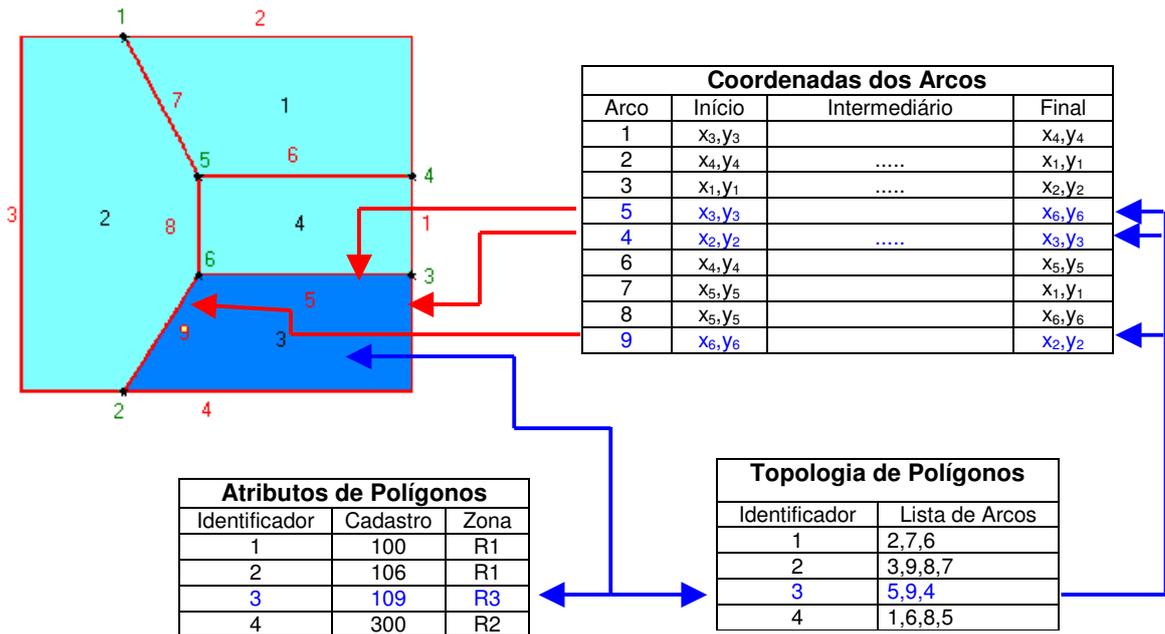


Figura 42 – Ligação Entidades Espaciais – Atributos

Cada polígono e cada arco da região recebe um identificador único, chave primária dos polígonos e chave primária dos arcos, respectivamente. A chave primária dos polígonos é utilizada como elo na relação Topologia de Polígonos e na relação Atributos de Polígonos. A relação Topologia de

Polígonos contém um domínio chamado Lista de Arcos, cujos valores de domínio são o elo para a relação Coordenadas dos Arcos.

Verifica-se que se tem três relações e um desenho ligados por elos. Consultando-se uma das relações é possível recuperar as outras e o desenho ou consultando-se o desenho é possível recuperar as relações.

Esse tipo de ligação e organização dos dados gráficos e não gráficos constitui o princípio do modelo georrelacional.

4. MÉTODOS ESTATÍSTICOS NA ANÁLISE ESPACIAL EM SAÚDE

Os métodos de estatística espacial são, particularmente, úteis para analisar a distribuição de doenças em determinada região, com o objetivo de identificar áreas de risco, analisar acessibilidade a serviços de saúde, descrever a população adscrita, entre outros.

Tais métodos podem ter a seguinte classificação, em função do tipo de dado disponível e da finalidade desejada (Bailey & Gatrell, 1995):

- **Eventos Pontuais** – esses métodos têm por objetivo estudar a distribuição espacial de eventos associados às representações pontuais, verificando se sua disposição apresenta características de aleatoriedade, de regularidade ou de aglomeração (Diggle, 2003). Como aplicações, cita-se a identificação de padrões de distribuição de doenças, de tipos de crimes, de espécies vegetais, etc. A Figura 43 exemplifica a distribuição espacial de residências em determinada região.

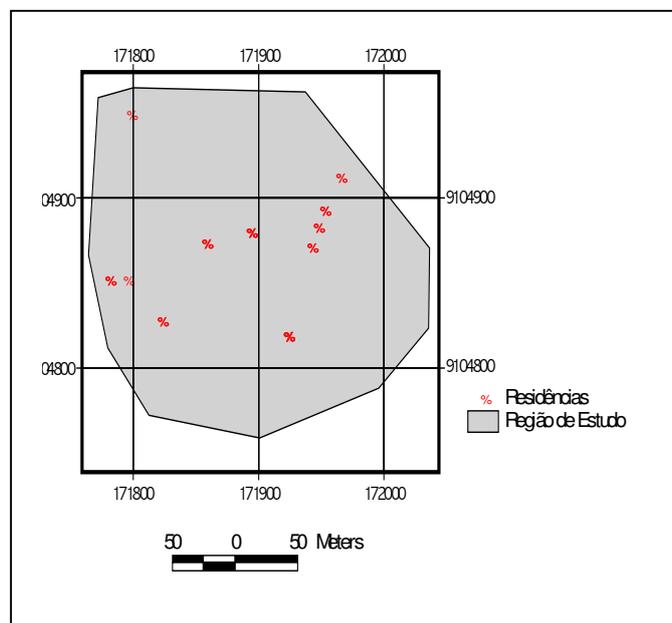


Figura 43 – Distribuição de Residências

- **Superfícies contínuas** – esses métodos têm por objetivo estudar a distribuição espacial dos valores de um atributo sobre toda a região de estudo, a partir de uma amostra coletada pontualmente. É útil na

obtenção de estimativas de valores do atributo em pontos não amostrados. Como aplicações citam-se a modelagem do relevo topográfico, o dimensionamento volumétrico de barragens e represas, a obtenção de superfícies de distribuição de poluentes aéreos, entre outras. A Figura 44 exemplifica a superfície contínua dos nascidos vivos, na cidade do Recife-PE, no período de 1995 a 1999, gerada a partir do número de nascimentos em cada bairro, identificados, pontualmente, no centróide desses.

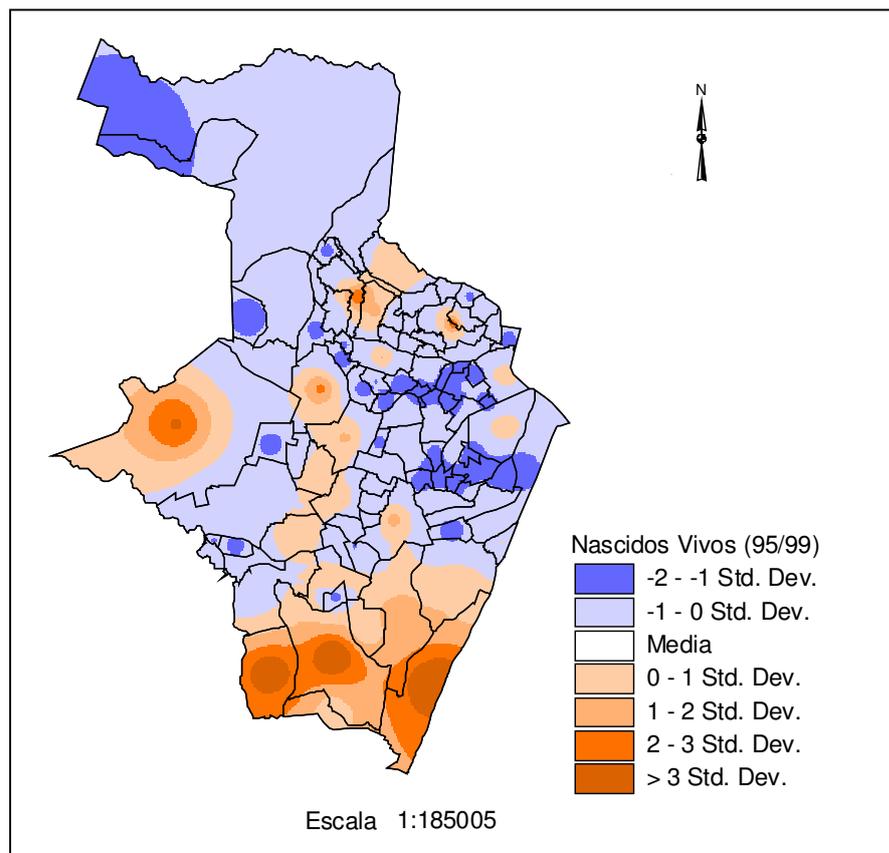


Figura 44 – Superfície dos nascidos vivos, Recife-PE, 1995 -1999

- Áreas com contagens ou taxas agregadas – tem por objetivo estudar a distribuição espacial de eventos, agregados em áreas discretas, dentro de uma região previamente delimitada. Essas regiões são municípios, bairros, setores censitários, cujos atributos são, por exemplo, estatísticas de mortalidade, morbidade, indicadores socioeconômicos, etc. A Figura 45 exemplifica a distribuição espacial de nascidos vivos, por bairro, na cidade do Recife-PE, no período de

1995 a 1999 (Secretaria de Saúde do Recife), usando como pontos de corte unidades de desvio padrão.

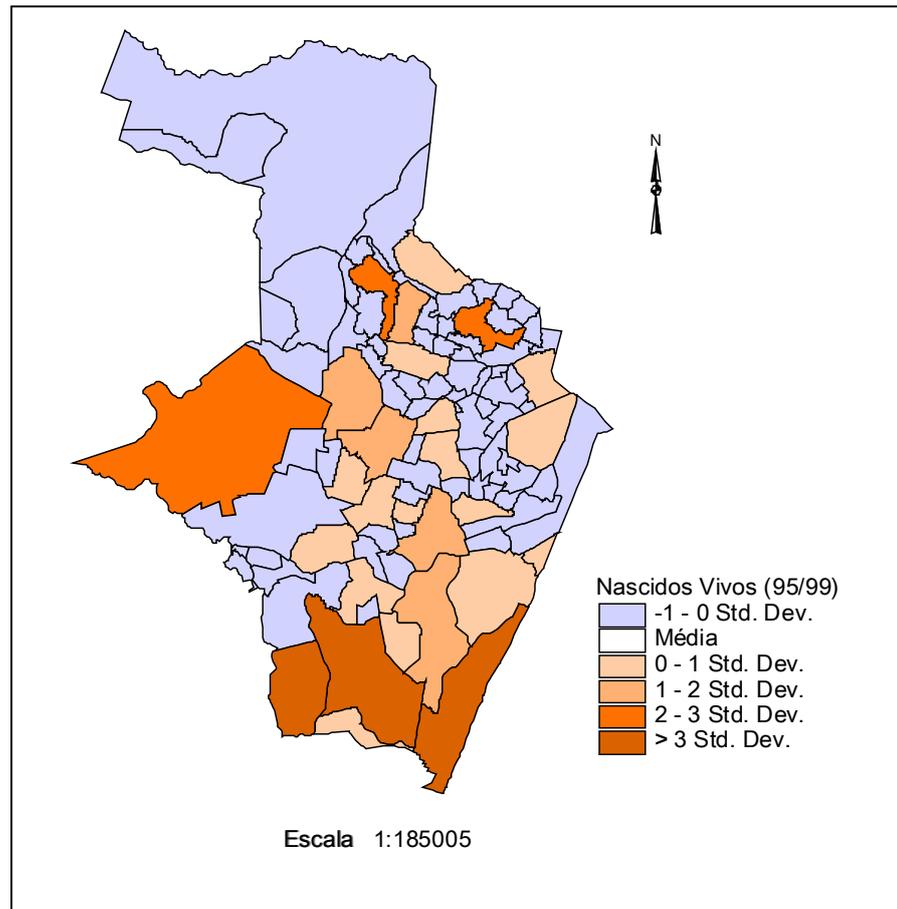


Figura 45 – Nascidos Vivos por bairro, Recife-PE, 1995-1999

- Interação espacial – esse método tem por objetivo modelar um padrão de fluxo, em termos de acessibilidade geográfica, relacionando origens e destinos. Como aplicações citam-se o planejamento de transportes urbanos, a análise de demanda aos serviços de saúde (Lapa, 2003), estudos sobre migrações, etc. A Figura 46 exemplifica o fluxo entre regiões administrativas de residência e de óbitos pós-neonatal, sobreposto ao padrão sócio-econômico dos bairros do município do Rio de Janeiro - 1995.

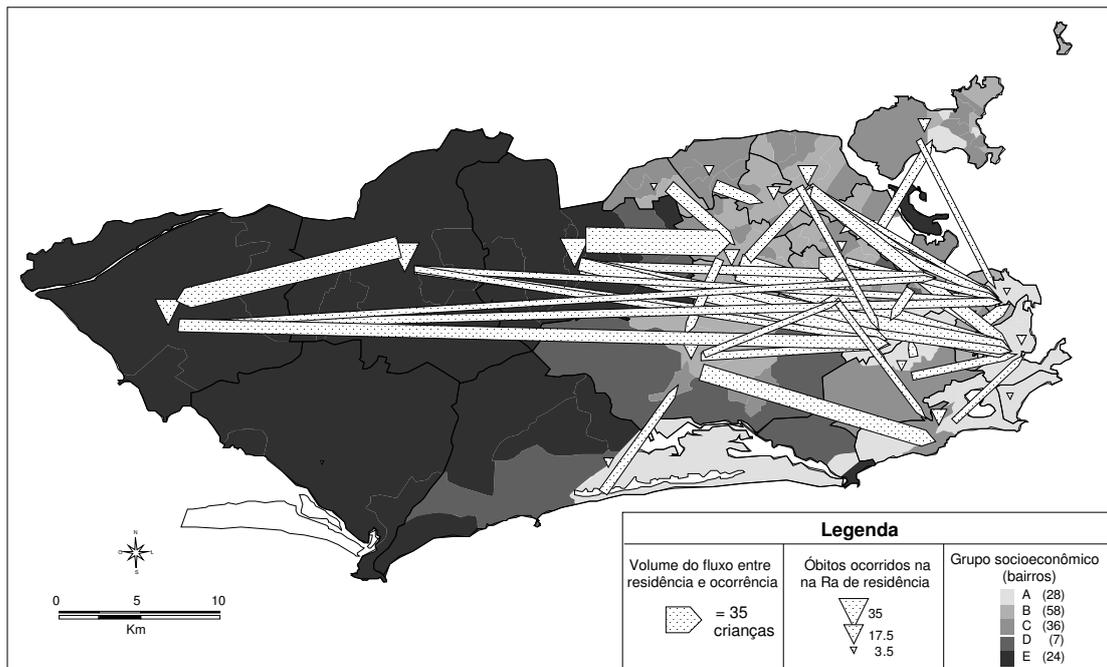


Figura 46 - Fluxo entre regiões de residência e de óbitos pós-neonatal, Rio de Janeiro-RJ, 1995

Diversas outras aplicações são possíveis e um tipo de dado pode ser convertido para outro. Entretanto, considerando que dados pontuais são pouco freqüentes nos estudos em saúde e que o evento tratado nessa tese é a ocorrência de doenças registradas pelo PACS/PSF, localizadas espacialmente pelas coordenadas pontuais das residências de seus moradores, serão detalhados somente métodos apropriados para análise de padrões de pontos.

4.1 PADRÕES DE PONTOS

O fenômeno que gera uma distribuição pontual, por exemplo de casos de doença em uma localidade, é um processo estocástico que pode ser descrito em termos de **efeitos de primeira ordem** e de **efeitos de segunda ordem** (Bailey & Gatrell, 1995).

Os **efeitos de primeira ordem** são descritos pela variação do valor médio do processo no espaço e por isso são considerados de pequena escala⁵ ou globais. Tais efeitos são obtidos por medidas de intensidade $\lambda_{(s)}$, que

⁵ Escala: relação entre as dimensões dos elementos representados em um mapa, carta, fotografia ou imagem e as correspondentes dimensões no terreno (IBGE, 2003). Isso indica que uma escala pequena é empregada para representar grandes áreas e escala grande para representar pequenas áreas do terreno.

correspondem à média, ou valor esperado, de eventos por unidade de área, quando esta área tende a zero, ou seja, tendo a um ponto \mathbf{s} . Essa intensidade é modelada segundo a Equação 2.

$$\lambda_{(s)} = \lim_{ds \rightarrow 0} \left\{ \frac{E(Y_{(ds)})}{ds} \right\} \quad (2)$$

Onde:

$E(Y_{(ds)})$ é o valor esperado do número de eventos na região ds ;

ds é a área de uma pequena região em torno do ponto \mathbf{s} .

Merece registro o fato de que $\lambda_{(s)}$ é função exclusivamente da posição absoluta da ocorrência dos eventos e essa é obtida por meio de um sistema de coordenadas perfeitamente definido sobre a região.

Os **efeitos de segunda ordem** permitem analisar a estrutura de dependência espacial do processo. Envolvem a relação entre pares de eventos segundo distâncias que guardam entre si e por isso são considerados efeitos locais ou de grande escala. Tais efeitos podem ser descritos através da esperança conjunta de ocorrência do evento em duas pequenas regiões, ds_1 e ds_2 , quando suas áreas tendem a zero, materializando assim um ponto. Matematicamente, os efeitos de segunda ordem são modelados segundo a Equação 3.

$$\gamma_{(s_i, s_j)} = \lim_{ds_i, ds_j \rightarrow 0} \left\{ \frac{E(Y_{(ds_i)} Y_{(ds_j)})}{ds_{(i)} ds_{(j)}} \right\} \quad (3)$$

Nessa situação, caso o processo seja estacionário, a intensidade independe da posição absoluta da ocorrência dos eventos, mas da distância que os separa. Ressalta-se que um processo espacial $Y_{(s)}, s \in R$ é estacionário quando sua média $E(Y_{(s)})$ e sua variância $Var(Y_{(s)})$ são constantes em R e assim independem da localização S .

A análise das propriedades de segunda ordem, ou seja, do padrão da distribuição, permite detectar a presença de aglomerados de casos ou de regularidade excessiva, com base no conceito de completa aleatoriedade espacial (CAE). Na área da saúde, o emprego mais usual dessa análise, destina-se à detecção de *cluster* espaço-temporal. Define-se aglomerados (Knox, 1988) como um número excessivo de casos em dada sub-área de uma região, quando comparada com as outras sub-áreas da mesma região. Existem dezenas de testes para detecção deste “número excessivo

da casos”, que podem ser divididos em testes focados, quando se testa a presença de aglomerados em torno de um ponto suspeito, e testes genéricos, que permitem verificar afastamentos da hipótese de CAE. Esse tema não será mais desenvolvido nesta tese, uma vez que esses métodos não foram aplicados, ainda que possam ser usados no modelo de SIG proposto, particularmente na detecção de epidemias, de doenças causadas por fontes de contaminação ambiental, entre outras.

4.2 EXPLORAÇÃO DO PADRÃO DE PONTOS

Uma das possibilidades da exploração do padrão de pontos é sua visualização em mapas. Para os efeitos de primeira ordem citam-se como pertinentes para confecção de tais mapas o **método do reticulado** e o **estimador kernel** (Bailey & Gatrell, 1995).

O **método do reticulado** consiste em lançar sobre a região estudada uma grade regular, com dimensões definidas a priori, de modo que todas as células que a compõem tenham a mesma área. A intensidade do processo é obtida dividindo-se a quantidade de eventos que ocorrem em cada célula pela área correspondente. O resultado desse procedimento é uma superfície contínua, representada num mapa, que mostra a variação da dita intensidade. A Figura 47 exemplifica a distribuição da densidade demográfica em determinado local, dado que essa é uma medida de intensidade.

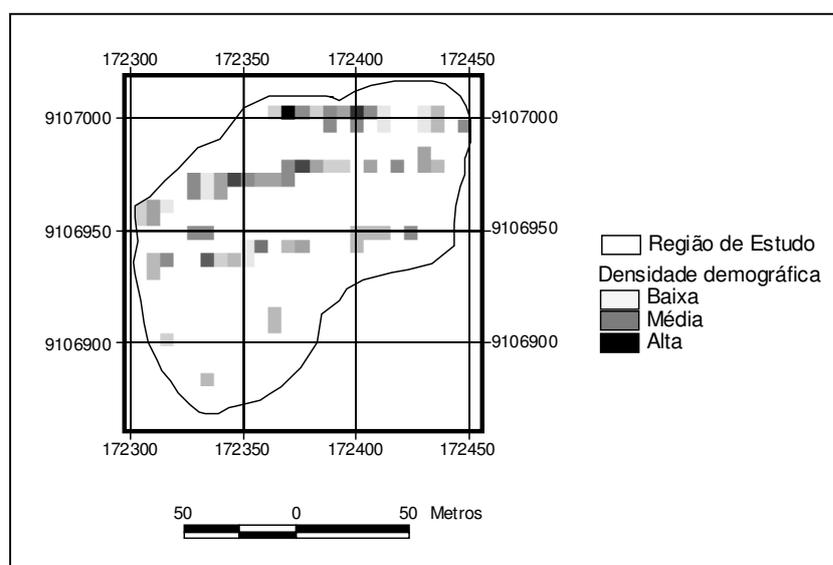


Figura 47 – Variação da densidade demográfica, segundo grade de 5 metros

Apesar do método fornecer uma idéia global de sub-regiões com alta ou baixa intensidade, tem por limitação a perda de detalhes, a medida que se aumenta o tamanho das células, tornando a superfície significativamente descontínua.

A primeira vista, tal limitação seria resolvida diminuindo-se a dimensão das células, entretanto, caso isso aconteça, haverá grande variabilidade de contagens entre elas, contribuindo para a ocorrência de muitas células vazias e, no caso extremo, retorna-se à localização dos pontos. Nessa situação, a interpretação da superfície torna-se comprometida, conforme pode ser observado na Figura 48.

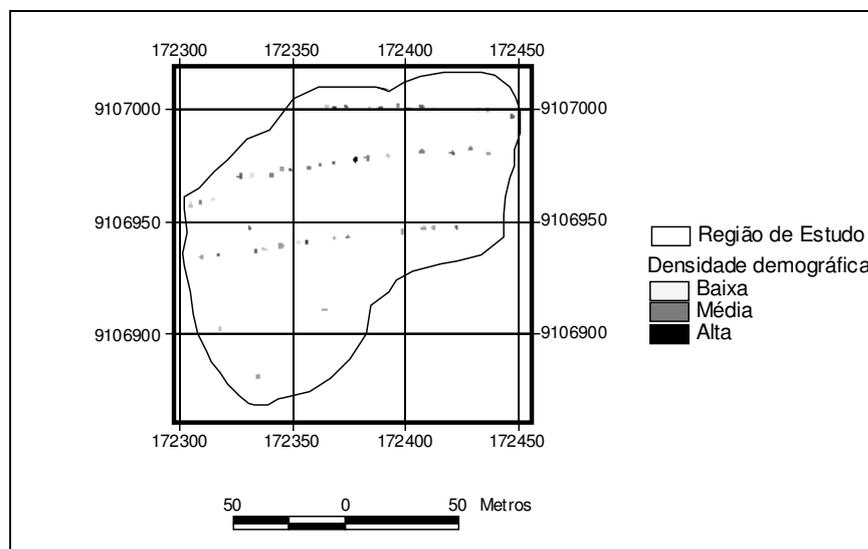


Figura 48 – Variação da densidade demográfica, segundo grade de 0,5 metros

O **estimador de densidade Kernel** consiste em ajustar uma função densidade de probabilidade bidimensional sobre os eventos considerados, compondo uma superfície suavizada, cujo valor será proporcional à intensidade de amostras por unidade de área (Santos et al., 2001). Tal intensidade é calculada segundo o modelo constante da Equação 4 e sua interpretação geométrica é mostrada na Figura 49, extraída de (Bailey, 1995).

$$\hat{\lambda}_{\tau(s)} = \frac{1}{\tau^2} \sum_{i=1}^n k\left(\frac{s-s_i}{\tau}\right) \quad (4)$$

onde:

s é a localização do ponto onde se deseja estimar a intensidade $\hat{\lambda}$;
 τ é a largura de banda (raio de busca ou área de influência), definida *a priori*;
 k é a função densidade de probabilidade bivariada (Kernel);
 $s_{(i=1,n)}$ é a posição dos n eventos que ocorrem dentro da largura de banda.

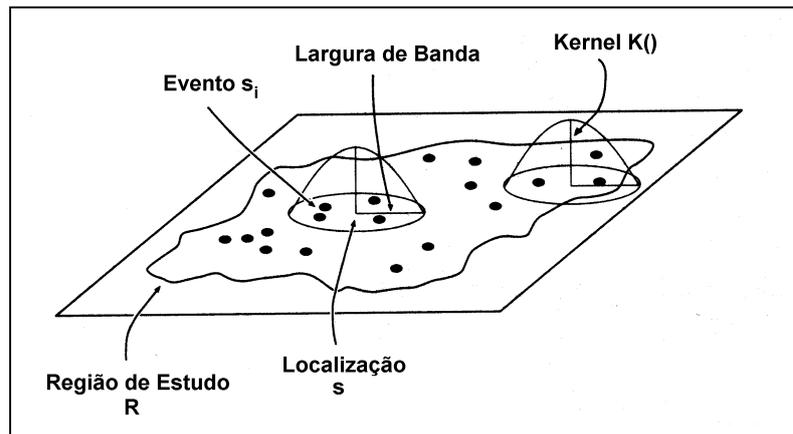


Figura 49 – Interpretação geométrica do estimador Kernel

No que diz respeito à escolha da função de suavização $k()$, pode ser empregada uma de quarta ordem (Bailey, 1995) e assim o estimador da intensidade passa a ser modelado pela Equação 5.

$$\hat{\lambda}_{\tau(s)} = \sum_{i=1}^n \frac{3}{\pi\tau^2} \left(1 - \frac{h_i^2}{\tau^2}\right)^2 \quad (5)$$

onde:

$\pi\tau^2$ é o ponderador do estimador, que corresponde à área de um círculo cujo raio é a largura de banda;

$h_{(i=1,n)}$ são as distâncias entre os n pontos localizados dentro da largura de banda e o ponto s onde se deseja estimar a intensidade.

Deve-se observar que o ponderador do estimador é função de τ e que esse é inversamente proporcional a $\hat{\lambda}_{\tau(s)}$. Isso indica que quanto maior for a largura de banda, menor será o valor da estimativa e, conseqüentemente, maior será o grau de suavização da superfície gerada. No limite, este estimador permite verificar a presença de tendência global.

A vantagem do estimador kernel sobre o método do reticulado é que, dada sua propriedade de suavização, este reduz, por meio de interpolação,

os efeitos de descontinuidade entre as células. Isso facilita a interpretação da variação da intensidade do processo sobre a região, conforme mostrado na Figura 50.

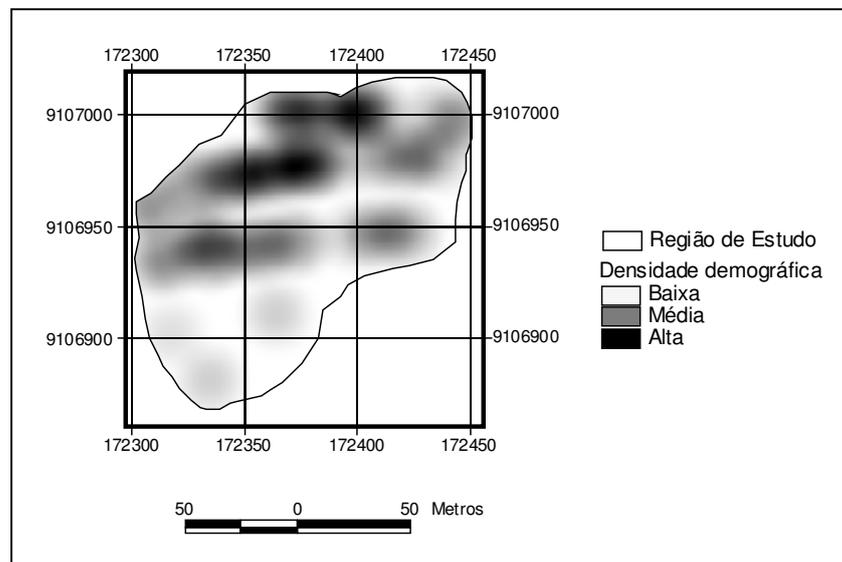


Figura 50 – Variação da densidade demográfica, segundo raio de busca de 20 metros

As técnicas apresentadas até este ponto somente permitem analisar as propriedades dos padrões de pontos, onde não há qualquer atributo associado ao ponto. Quando a cada localização pode-se associar alguma característica, por exemplo, ter ou não uma doença, denomina-se Processo Pontual Marcado. O atributo ou “marca” é o elemento de estudo.

Tal processo será detalhado a seguir, face sua pertinência para o PACS/PSF, uma vez que a cada indivíduo – pontos no espaço – estão registradas suas características – marcas.

4.3 PROCESSO PONTUAL MARCADO

O processo pontual marcado pode ser descrito como a co-existência de dois processos espaciais caracterizados pelo atributo. Sua importância decorre do fato de permitir a obtenção de medidas de risco, através da comparação, por exemplo, entre crianças que morrem no primeiro ano de vida (casos) e as que não morrem (controles), identificadas espacialmente (Câmara & Carvalho, 2001). A idéia é semelhante aos estudos

epidemiológicos tipo caso-controle, mas onde a variabilidade de localização de casos e controles é o objeto de estudo.

Considere-se $(s_1, s_2, \dots, s_{n_1})$ as localizações na região de estudo R dos n_1 casos e $(s_{n_1+1}, s_{n_1+2}, \dots, s_n)$ as localizações dos $(n-n_1)$ controles como observações de dois processos de Poisson I e II, com intensidades $\lambda_1(s)$ e $\lambda_2(s)$, respectivamente. Assim, o logaritmo do risco na localidade s , é dado pela Equação 6, sendo o objetivo da análise investigar a variação desse risco na região R .

$$\rho(s) = \log \left\{ \frac{\lambda_1(s)}{\lambda_2(s)} \right\} \quad (6)$$

A estimação espacial do risco, nesse caso, pode ser feita dentro da lógica do Modelo Aditivo Generalizado. Este tipo de modelo permite a incorporação de covariáveis que se relacionam com a variável resposta de forma não linear. Assim, é possível incorporar a variação espacial na estimação do risco controlando por fatores individuais de risco (Kelsall & Diggle, 1998, Shimakura et al., 2001).

Considerando-se então as localizações dos n_1 casos e dos $(n-n_1)$ controles da região R , assume-se y_i um indicador associado ao ponto s_i tal que $y_i = 1$ se o indivíduo i for caso ou $y_i = 0$ caso contrário. Supondo que y_i ($i=1, \dots, n$) são realizações de variáveis aleatórias independentes $Y_i \sim \text{Bernoulli}$ chega-se a Equação 7.

$$p(s) = \frac{q_1 \lambda_1(s)}{q_1 \lambda_1(s) + q_2 \lambda_2(s)} \quad (7)$$

onde q_1 e q_2 representam a proporção de casos e controles registrados, respectivamente, em relação ao total de fato existente.

Operando-se as Equações 6 e 7 obtém-se a Equação 8.

$$\log \left(\frac{p(s)}{1-p(s)} \right) = \rho(s) + c \quad (8)$$

onde $c = \log(q_1/q_2)$, significando que c é uma constante aditiva que não modifica as características da distribuição espacial do risco na região.

Dada a Equação 8 pode-se obter estimativas espaciais de risco como definido em (6), a parte de uma constante aditiva, considerando um modelo simples atribuído a variáveis binárias.

A inclusão de efeitos de covariáveis no modelo sugerido por Kelsall & Diggle é dada pela Equação 9.

$$\log\left(\frac{p(s, x)}{1 - p(s, x)}\right) = y = \beta x + g(s) \quad (9)$$

onde:

- y é a variável resposta, do tipo binária, numa posição s ;
- x é o vetor da covariáveis individuais;
- β é o vetor dos parâmetros estimados pelo modelo, representando a razão de chances (*odds ratio*) relacionada a cada covariável;
- $g(s)$ é uma função suave, porém desconhecida, das coordenadas espaciais s . Uma função que preenche essa característica é a razão do estimador *kernel* de casos e controles.

Caso o risco seja constante na região, a constante aditiva $g(s)$ será zero e o modelo se reduzirá a uma regressão logística usual (Shimakura et al, 2001, Hosmer & Lemeshow, 1989).

Os procedimentos de estimação dos parâmetros desse modelo baseiam-se em métodos iterativos usuais na estimação de modelos aditivos generalizados. Estimam-se, inicialmente, os parâmetros da regressão e sobre os resíduos estima-se a função *Kernel* e assim sucessivamente, até que as estimativas não mais se alterem. O método iterativo aqui adotado (Hastie et Tibishirani, 1997) pode ser resumido no seguinte algoritmo (Shimakura et al, 2001):

1. Faz-se $g(s) = 0$ e obtém-se as estimativas iniciais dos coeficientes β ajustando um modelo de regressão logística da forma

$$\log\left(\frac{\rho(x)}{1 - \rho(x)}\right) = \beta x;$$

2. Calcula-se $\hat{\eta}_i = \hat{\beta}x_i + \hat{g}(s)$, $\hat{p}_i = \frac{(\exp \hat{\eta}_i)}{1 + (\exp \hat{\eta}_i)}$, $z_i = \hat{\eta}_i + \frac{y_i - \hat{p}_i}{\hat{p}_i(1 - \hat{p}_i)}$;

3. Ajusta-se o modelo da forma $z = \beta x + g(s) + \varepsilon$;

- a. Estima-se $g(s)$ usando regressão *kernel* ponderada (Wand & Jones, 1995) de $u_i = z_i - \hat{\beta}x_i$ em s_i , com pesos $w_i = \hat{p}_i(1 - \hat{p}_i)$:

$$\hat{g}(s) = \frac{\sum_{i=1}^n w_i k\left(\frac{s - s_i}{\tau}\right) u_i}{\sum_{i=1}^n w_i k\left(\frac{s - s_i}{\tau}\right)},$$

onde $k()$ é a função *kernel*, definida pela Equação 5.

- b. Estima-se β por métodos de mínimos quadrados ponderados de $z_i - \tilde{g}(s_i)$ em u_i com pesos w_i ;
 - c. Repete-se os passos (a) e (b) até a convergência das estimativas;
4. Repetem-se os passos (2) e (3) até a convergência das estimativas.

A largura de banda (τ) ótima pode ser definida por *métodos de validação cruzada*, cujo princípio é selecionar tal largura de modo a minimizar determinada função dos dados e dos valores preditos pelo modelo (Shimakura et al, 2001). Entretanto, freqüentemente, como a distribuição dos pontos na região é heterogênea, a largura estimada dessa forma não é a que melhor expressa o fenômeno em estudo, cabendo ao pesquisador selecioná-la utilizando diversos outros critérios relacionados à região.

É necessário verificar se realmente existem evidências de que o risco não é constante na região, ou seja, testar a hipótese $H_0: g(s)=0$ na Equação 5. Além desse teste a construção de contornos de tolerância (isolinhas) que identifiquem áreas onde o risco é significativamente maior ou menor que a média global, é interessante para orientar a intervenção.

O teste global de risco e a identificação de áreas de baixo e alto risco podem ser feitos utilizando o método de simulação Monte Carlo, conforme o algoritmo proposto por Kelsall & Diggle (1998):

1. Ajusta-se um modelo de regressão logística convencional (aceitando-se H_0), obtendo-se as probabilidades ajustadas \hat{p}_i .
2. Fixam-se as localizações de cada ponto e amostram-se n_1 dos n indivíduos sem reposição, com probabilidade proporcional a \hat{p}_i e estes são rotulados como casos e os $n-n_1$ restantes como controles.
3. Calcula-se nova estimativa de $g(s)$, $\hat{g}_1(s)$, a estimativa centralizada em torno da média $\tilde{g}_1(s) = \hat{g}_1(s) - \bar{g}_1$, onde

$$\bar{g}_1(s) = n^{-1} \sum_{i=1}^n \hat{g}_1(s_i) \text{ e a estatística } t_1 = n^{-1} \sum_{i=1}^n (\tilde{g}_1(s_i))^2 .$$

4. Repetem-se os passos 2 a 3 m vezes.
5. Constrói-se uma superfície de p-valores que para cada s fornece a proporção dos valores de $\tilde{g}_j(s), j=1, \dots, m$, menores que a estimativa original, digamos $\tilde{g}_0(s)$.
6. Adicionam-se os contornos de 0,05 e 0,95 da superfície de p-valores ao mapa de $\tilde{g}_0(s)$ como contornos de 90% de confiança para indicar áreas de alto/baixo risco.
7. Para o teste de hipótese, define-se k o número $t_j > t_0$ (obtido a partir de $\tilde{g}_0(s)$ e o nível de significância correspondente por $p = (k+1)/(m+1)$).

Os procedimentos de obtenção da estimativa de risco e do teste global do risco serão implementados no estudo de caso, apresentado no Capítulo 7.

5. PROPOSTA DE UM MODELO CONCEITUAL DE SIG PARA O PSF

Procedeu-se a um estudo detalhado do PSF e do SIAB para que se pudesse cumprir a fase da *abstração do mundo real*⁶. Estudaram-se as potencialidades do SIG e dos métodos de estatística espacial para analisar a distribuição de doenças, observando-se que são pertinentes para atender às necessidades do Programa, sob a ótica da informação.

Esse capítulo destina-se a cumprir a fase da *definição do modelo conceitual* (Buzai & Durán, 1977) de um SIG para o PSF, com base num banco de dados georrelacional, com informações espaciais nas estruturas *raster* e vetorial e não espaciais na estrutura de banco de dados.

A estrutura *raster* é empregada no armazenamento de imagens de satélites e no armazenamento de superfícies contínuas, geradas a partir da modelagem da localização de eventos de saúde.

As imagens de satélite se destinam, em princípio, à análise qualitativa das entidades geográficas ambientais, como a identificação de massas d'água, de estradas, de áreas de plantio, entre outras, e também para pano de fundo das demais. Cabe ressaltar que essa destinação pode ser também a de identificar residências, desde que sejam empregadas imagens de alta resolução geométrica.

As superfícies contínuas de eventos de saúde são obtidas como resultados dos métodos estatísticos empregados. A amostra geradora corresponde ao morador, portador ou não de alguma doença, tendo sua posição espacial definida pelas coordenadas UTM do local onde reside. Cabe a observação de que o sistema permite mas não especifica o método estatístico a ser empregado, porque esse é realizado em programa específico.

A estrutura vetorial é empregada no armazenamento de entidades geográficas ambientais, incluindo-se as residências da região e no armazenamento de superfícies vetoriais de eventos de saúde.

As entidades geográficas ambientais são armazenadas segundo um modelo topológico, obedecendo-se o sistema de projeção UTM.

As superfícies vetoriais de eventos de saúde são materializadas por polígonos limitados por isolinhas. Podem ser provenientes das superfícies

⁶ Ver Capítulo 1

contínuas de eventos de saúde ou de modelagem de dados agregados por áreas.

A estrutura de banco de dados tradicional para as informações não espaciais tem por objetivo permitir rapidez, integridade e segurança em operações do tipo consulta, atualização e eliminação de registros, bem como emissão de relatórios consolidados, previstos pelo PSF.

Ressalta-se, ainda, que o sistema como foi projetado, permite a implementação de modelos de análise de eventos de saúde, espaciais ou não, cabendo ao usuário a escolha do mais conveniente.

5.1 PROJETO LÓGICO

O projeto lógico do SIG será efetuado na forma de um Diagrama Entidade-Relacionamento (DER), por permitir uma fácil visualização da estrutura e do relacionamento entre as entidades que compõem o sistema (Chen, 1990; Date, 1995), conforme mostrado na Figura 51.

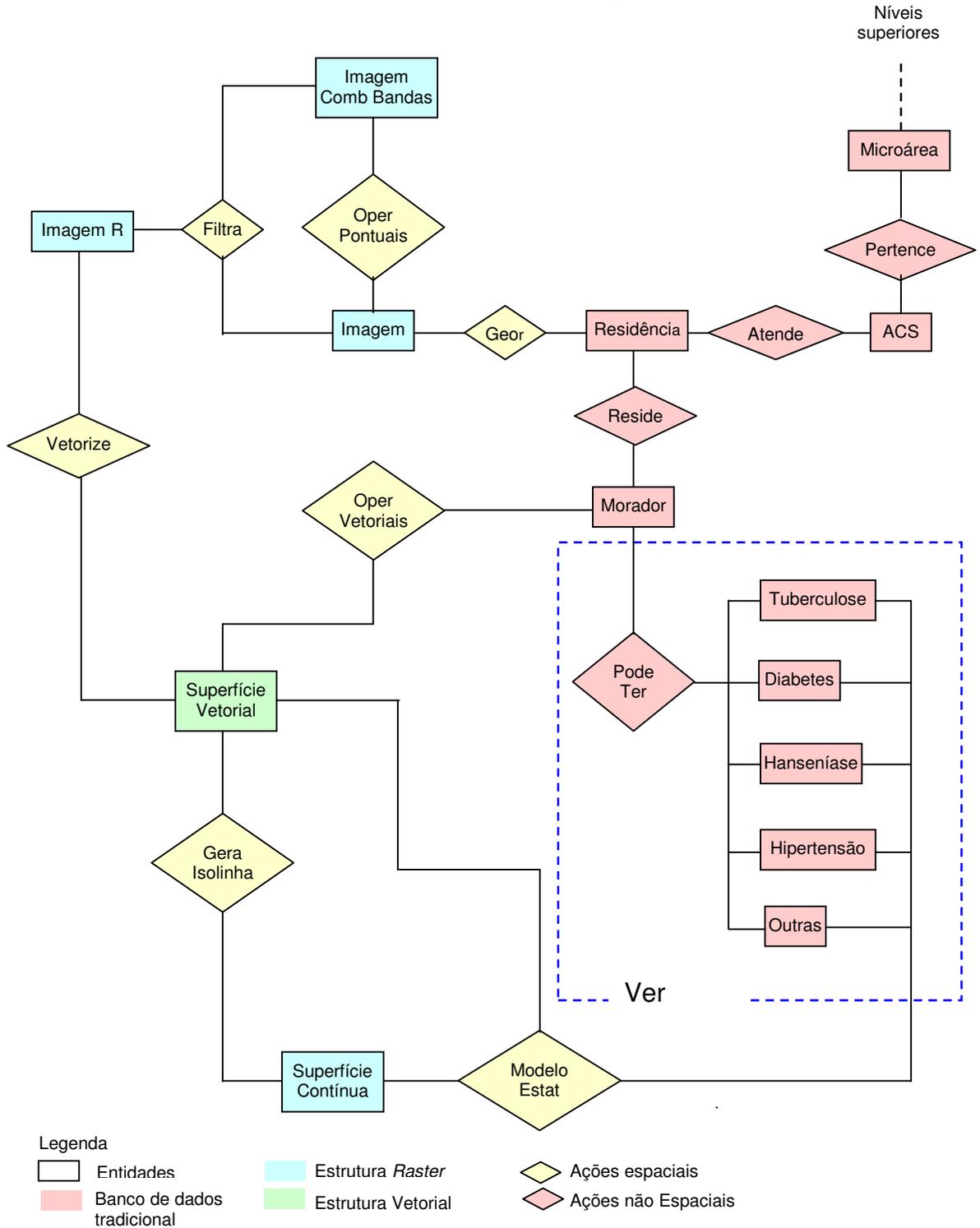


Figura 51 – Diagrama Entidade-Relacionamento

Convém observar que tratar os eventos de saúde diabetes, hanseníase, hipertensão, tuberculose e outras como entidades separadas tem a vantagem de permitir substancial economia de memória e eficiência em operações de consulta, eliminação e atualização de registros.

Tomando-se, por exemplo, o caso da hipertensão, verifica-se que seus atributos são definidos pela relação a seguir:

Hipten			
Item	Tipo	Descrição	Memória (bytes)
Data_Vis_As	D	Data da visita do agente de saúde	8
Faz_Dieta	1 C	Faz dieta	1
Faz_Exerc	1 C	Faz exercícios	1
Toma_Med	1 C	Toma medicação	1
Pres_Art	5 C	Pressão arterial	5
Data_UI_C	D	Data da última consulta	8
			Total 24

Caso não fosse tratada como entidade distinta, esses atributos estariam incluídos na relação **Morador** e seriam consumidos 24 bytes de memória por morador, que fosse ou não hipertenso. Se existirem 2000 moradores, serão consumidos 48.000 bytes de memória, só para esse evento.

Tratando-se agora a hipertensão como entidade distinta, faz-se necessário incluir na sua relação mais um atributo (Cod_Pront_Mor), que será o elo implícito para ligação com a relação **Morador**, conforme mostrado a seguir:

Morador			
Item	Tipo	Descrição	
Cod_Pront	4 N 0	Código do prontuário da família	
Cod_Pront_Mor	6 N 2	Código do prontuário/morador	
Nome	45 C	Nome do morador	
Sexo	1 C	Sexo do morador	
Idade	3 N 0	Idade do morador	
C_Ocup	1 N 0	Ocupação do morador (código)	
C_Renda	1 N 0	Renda do morador (código)	
Escolar	1 N 0	Escolaridade do morador	
Fuma	1 N 0	Fumante (S), não fumante (N)	

Hipten			
Item	Tipo	Descrição	Memória (bytes)
Data_Vis_As	D	Data da visita do agente de saúde	8
Faz_Dieta	1 C	Faz dieta	1
Faz_Exerc	1 C	Faz exercícios	1
Toma_Med	1 C	Toma medicação	1
Pres_Art	5 C	Pressão arterial	5
Data_UI_C	D	Data da última consulta	8
Cod_Pront_Mor	6 N 2	Código do prontuário/morador	4
			Total 28

Nessa situação, serão empregados 28 bytes de memória por morador portador de hipertensão. Se dos 2.000 moradores 10% forem hipertensos,

serão consumidos 5.600 bytes, uma redução de 78% em relação à situação anterior.

Vale salientar que essa economia será tanto maior quanto maior for a quantidade de atributos por doença e mais rara seja sua ocorrência.

5.2 DICIONÁRIO DE DADOS

O dicionário de dados tem por objetivo registrar a descrição dos dados e informar seus formatos físicos (Jones, 1988). No presente caso, está dividido em duas partes. A primeira trata exclusivamente das informações espaciais e a segunda das informações não espaciais.

5.2.1 DICIONÁRIO DAS INFORMAÇÕES ESPACIAIS

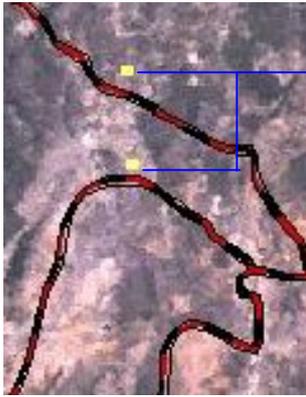
- Imagem: imagem de satélite da microárea coberta pela USF, na estrutura *raster*.
- Imagem Comb Bandas: imagem de satélite da microárea coberta pela USF resultante de combinação de bandas, de modo a facilitar a análise qualitativa das entidades geográficas ambientais, na estrutura *raster*.
- Imagem R: imagem de satélite da microárea coberta pela USF com realce de feições com desenvolvimento linear, como estradas e limites de uso do solo e massas d'água, na estrutura *raster*.
- Geor: operação que visa o georreferenciamento da imagem de satélite da microárea coberta pelo USF.
- Oper Pontuais: operação que visa combinar bandas de imagens de satélite da microárea coberta pela USF.
- Filtra: operação que visa suavizar ou realçar as imagens de satélite da microárea coberta pela USF.
- Superfície Vetorial: superfície vetorial da microárea coberta pela USF, composta por pontos linhas e polígonos.
- Vetorize: operação que visa vetorizar as feições realçadas da Imagem R.
- Oper Vetoriais: Operações que visam análise espacial vetorial, envolvendo as operações de combinação de planos de informação,

geração de áreas de influência, extração de feições, combinação de feições.

- Gera Isolinha: Operação que visa a transformação da superfície na estrutura *raster* para superfície na estrutura vetorial, através da construção de linhas de iso-valor.
- Superfície Contínua: superfície contínua de eventos de saúde, dentro da microárea coberta pela USF, na estrutura *raster*.
- Modelo Estat: operações que se destinam à análise espacial, segundo modelos estatísticos definidos pelo usuário.

5.2.2 DICIONÁRIO DAS INFORMAÇÕES NÃO ESPACIAIS

Tema: RESIDÊNCIA



Cobertura: Residência
 Tipo: Ponto
 Tabela de Atributos: Residência
 Elo: **Cod_Pront**
 Escala de Aquisição: 1:1 (GPS)

Arquivo Gráfico: Residência

Pano de Fundo: Caru.img
 Escala de Aquisição: 1:25000

Residência

Item	Tipo	Descrição
E	9,N,2	Coordenada este (identificador espacial)
N	10,N,2	Coordenada norte (identificador espacial)
Cod_Pront	4,N,0	Código do prontuário da família
Nome_ACS	45,C	Nome do agente comunitário de saúde
End	45,C	Endereço
Num	4,N,0	Número
Bairro	45,C	Bairro
Localidade	45,C	Localidade
C_Mat_Pred	1,N,0	Material predominante (código)
C_Piso	1,N,0	Piso (código)
C_Coberta	1,N,0	Cobertura da residência (código)
N_Com	2,N,0	Número de cômodos
N_Com_D	2,N,0	Número de cômodos usados para dormir
C_Água_Enc	1,N,0	Tipo de encanamento do domicílio (código)
C_Or_Agu_Be	1,N,0	Origem da água para beber (código)
C_Dist_Agua	1,N,0	Distância da água à residência (código)
C_Saneam	1,N,0	Tipo de saneamento (código)
C_Dest_Lixo	1,N,0	Destino do lixo (código)
C_Elet	1,N,0	Existência eletrificação (código)
C_Rádio	1,N,0	Existência de rádio (código)
C_TV	1,N,0	Existência de televisão (código)
C_Gelad	1,N,0	Existência de geladeira (código)
N_cão	2,N,0	Número de cães
N_gato	2,N,0	Número de gatos

Conteúdo de Dist_Ag

C_Dist_Agua	DISTANCIA DA AGUA
1	ATE 50 m
2	DE 50 A 199 m
3	DE 200 A 999 m
4	MAIS DE 1 km

Conteúdo de Or_A_Be

C_Or_Agu_Be	OR A BE
1	REDE PUBLICA
2	POÇO
3	RIO/RIACHO
4	CISTERNA
5	BARREIRO
6	OUTROS

Conteúdo de Agua_En

C_Água_Enc	ÁGUA ENCANADA
1	SIM DENTRO DE CASA
2	SIM USO COMUM
3	SIM NO VIZINHO
4	NÃO

Conteúdo de Coberta

C_Coberta	COBERTA
1	TELHA
2	LAJE
3	OUTRO

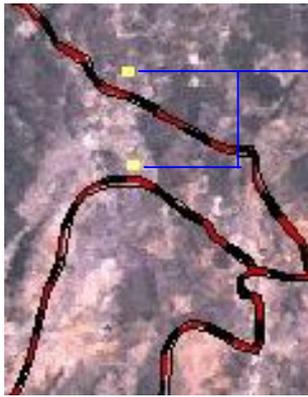
Conteúdo de Piso

C_Piso	PISO
1	CIMENTADO
2	BARRO
3	CERÂMICA
4	OUTRO

Conteúdo de Mat_Pred

C_Mat_Pred	MATERIAL PREDOMINANTE
1	TIJOLO
2	TAIPA
3	TIJOLO E TAIPA
4	LATA/PAPELÃO
5	OUTRO

Tema: RESIDÊNCIA



Cobertura: Residência
 Tipo: Ponto
 Tabela de Atributos: Residência
 Elo: **Cod_Pront**
 Escala de Aquisição: 1:1 (GPS)

Arquivo Gráfico: Residência

Pano de Fundo: Caru.img
 Escala de Aquisição: 1:25000

Residência

Item	Tipo	Descrição
E	9,N,2	Coordenada este (identificador espacial)
N	10,N,2	Coordenada norte (identificador espacial)
Cod_Pront	4,N,0	Código do prontuário da família
Cod_ACS	45,C	Código do agente comunitário de saúde
End	45,C	Endereço
Num	4,N,0	Número
Bairro	45,C	Bairro
Localidade	45,C	Localidade
C_Mat_Pred	1,N,0	Material predominante (código)
C_Piso	1,N,0	Piso (código)
C_Coberta	1,N,0	Coberta da residência (código)
N_Com	2,N,0	Número de cômodos
N_Com_D	2,N,0	Número de cômodos usados para dormir
C_Água_Enc	1,N,0	Tipo de encanamento do domicílio (código)
C_Or_Águ_Be	1,N,0	Origem da água para beber (código)
C_Dist_Água	1,N,0	Distância da água à residência (código)
C_Saneam	1,N,0	Tipo de saneamento (código)
C_Dest_Lixo	1,N,0	Destino do lixo (código)
C_Elet	1,N,0	Existência eletrificação (código)
C_Rádio	1,N,0	Existência de rádio (código)
C_TV	1,N,0	Existência de televisão (código)
C_Gelad	1,N,0	Existência de geladeira (código)
N_cão	2,N,0	Número de cães
N_gato	2,N,0	Número de gatos

Conteúdo de Gel

C_Gelad	GELADEIRA
1	SIM
2	NÃO

Conteúdo de Telev

C_TV	TELEVISÃO
1	SIM
2	NÃO

Conteúdo de Radio

C_Radio	RADIO
1	SIM
2	NÃO

Conteúdo de Elet

C_Elet	ELETRICIDADE
1	SIM
2	NÃO

Conteúdo de Lixo

C_Dest_Lixo	DEST LIXO
1	JOGADO A CÉU ABERTO
2	RECOLHIDO PELO GARI
3	DEPÓSITO DA PREFEITURA
4	QUEIMADO
5	ENTERRADO

Conteúdo de Saneam

C_Saneam	SANEAMENTO
1	PRIVADA DENTRO DE CASA COM DESCARGA
2	PRIVADA DENTRO DE CASA SEM DESCARGA
3	PRIVADA NA CASINHA
4	NÃO TEM SANEAMENTO

ACS(21)

Item	Tipo	Descrição
Cod_ACS	45 C	Código do ACS
Município	7 C	Código do município
Segmento	2 C	Código do segmento
Unidade	7 C	Código da unidade
Área	3 C	Código de área
Microárea	2 C	Código da microárea

Tema: MORADOR

.....
.....

Cobertura: Indivíduo

Tipo: Registro

Elos: **Cód_Pront****Cód_Pront_Mor**

Arquivo Gráfico: Residência

Pano de Fundo: Caru.img

Escala de Aquisição: 1:25000

Morador

Item	Tipo	Descrição
Cód_Pront	4 N 0	Código do prontuário da família
Cód_Pront_Mor	6 N 2	Código do prontuário/morador
Nome	45 C	Nome do morador
Sexo	1 C	Sexo do morador
Idade	3 N 0	Idade do morador
C_Ocup	1 N 0	Ocupação do morador (código)
C_Renda	1 N 0	Renda do morador (código)
Escolar	1 N 0	Escolaridade do morador
Fuma	1 N 0	Fumante (S), não fumante (N)

Renda

C_Renda	RENDA
1	ATE 1/2 SALÁRIO MÍNIMO
2	MAIS DE 1/2 ATE 1 SALÁRIO MÍNIMO
3	MAIS DE UM ATE 2 SALÁRIOS MÍNIMOS
4	MAIS DE 2 ATE 3 SALÁRIOS MÍNIMOS
5	MAIS DE 3 SALÁRIOS MÍNIMOS

Ocup

C_Ocup	OCUPAÇÃO
1	DESEMPREGADO
2	AUTÔNOMO
3	EMPREGADO SEM CARTEIRA ASSINADA
4	EMPREGADO COM CARTEIRA ASSINADA
5	TRABALHADOR EVENTUAL
6	APOSENTADO
7	AGRICULTOR
8	FUNCIONÁRIO PÚBLICO

Tema: TUBERCULOSE

.....
.....

Cobertura: Tub

Tipo: Registro

Elos: **Cód_Pront_Mor**

Arquivo Gráfico: Residência

Pano de Fundo: Caru.img

Escala de Aquisição: 1:25000

Tub

Item	Tipo	Descrição
Data_Vis_As	D	Data da visita do agente de saúde
Toma_Med_D	1 C	Toma medicação diária
Reações_I	1 C	Reações indesejáveis
Dat_Ult_C	D	Data última consulta
Ex_Esc	D	Data de exame de escarro
Com_Ex	2 N 0	Número de comunicantes examinados
<5_BCG	2 N 0	Número de < 5 anos que receberam BCG
N_Com	2 N 0	Número de comunicantes
N_Com_<5	2 N 0	Número de comunicantes <5 anos
Cód_Pront_Mor	6 N 2	Código do prontuário da família/morador

Tema: HIPERTENSÃO

.....
.....

Cobertura: Hipten
 Tipo: Registro
 Elos: **Cód. Pront. Mor**

Arquivo Gráfico:
 Residência

Pano de Fundo: Caru.img
 Escala de Aquisição: 1:25000

Hipten

Item	Tipo	Descrição
Data_Vis_As	D	Data da visita do agente de saúde
Faz_Dieta	1 C	Faz dieta
Faz_Exerc	1 C	Faz exercícios
Toma_Med	1 C	Toma medicação
Pres_Art	5 C	Pressão arterial
Data_UI_C	D	Data da última consulta
Cód. Pront. Mor	6 N 2	Código do prontuário da família/morador

Tema: DIABETES

.....
.....

Cobertura: Diab
 Tipo: Registro
 Elos: **Cód. Pront. Mor**

Arquivo Gráfico: Residência

Pano de Fundo: Caru.img
 Escala de Aquisição: 1:25000

Diab

Item	Tipo	Descrição
Data_Vis_As	D	Data da visita do agente de saúde
Faz_Dieta	1 C	Faz dieta
Faz_Exerc	1 C	Faz exercícios
Usa_Insul	1 C	Usa insulina
Toma_HGM	1 C	Toma hipoglicemiantes oral
Data_UI_C	D	Data da última consulta
Cód. Pront. Mor	6 N 2	Código do prontuário da família/morador

Tema: HANSENIASE

.....
.....

Cobertura: Hans
 Tipo: Registro
 Elos: **Cód. Pront. Mor**

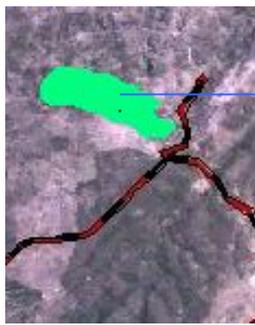
Arquivo Gráfico: Residência

Pano de Fundo: Caru.img
 Escala de Aquisição:
 1:25000

Hans

Item	Tipo	Descrição
Data_Vis_As	D	Data da visita do agente de saúde
Toma_Med_D	1 C	Toma medicação diária ?
Dat_Ult_Dos_S	D	Data última dose supervisionada
Faz_A_Cui	1 C	Faz auto cuidados?
Dat_Ult_C	D	Data última consulta
Com_Ex	2 N	Número de comunicantes examinados
Com_BCG	2 N	Número de comunicantes que receberam BCG
N_Com	2 N	Número de comunicantes
Cód. Pront. Mor	6 N 2	Código do prontuário da família/morador

Tema: USO DO SOLO



Cobertura: Uso_Solo
 Tipo: Polígono
 Tabela de Atributos: Uso_Solo
 Escala de Aquisição: 1:25000

Arquivo Gráfico: Uso_Solo

Pano de Fundo:
 Caru.img

Uso_Solo

Item	Tipo	Descrição
Área	13,N,6	Área do polígono
Perímetro	13,N,6	Perímetro do polígono
Uso_Solo	11,N,0	Seqüência Interna
Uso_Solo_Id	11,N,0	Identificador
Tipo	1,N,0	Tipo de uso (código)

Conteúdo de Tipo

C_Tipo	Tipo
1	Solo Natural
2	Cultura
3	Mata
4	Caatinga

Tema: MASSA D'ÁGUA



Cobertura: Massa_Ag
 Tipo: Polígono
 Tabela de Atributos: Massa_Ag
 Escala de Aquisição: 1:25000

Arquivo Gráfico: Massa_Ag

Pano de Fundo: Caru.img
 Escala de Aquisição: 1:25000

Massa_Ag

Item	Tipo	Descrição
Área	13,N,6	Área do polígono
Perímetro	13,N,6	Perímetro do polígono
Massa_Ag	11,N,0	Seqüência Interna
Massa_Ag_Id	11,N,0	Identificador
Tipo	1,N,0	Tipo de uso (código)

Conteúdo de Tipo

C_Tipo	Tipo
1	Rio
2	Lagoa
3	Barreiro
4	Tanque

6. IMPLEMENTAÇÃO DO PROJETO LÓGICO

Este capítulo trata da implementação de parte do projeto lógico inicialmente previsto, em uma área piloto. As razões de não implementá-lo na sua totalidade são devidas, por um lado, à indisponibilidade de imagens de satélite em escala compatível com as dimensões geográficas da área estudada, impossibilitando a localização precisa das residências e de outras variáveis ambientais de interesse e, por outro lado, à grande diversidade de análises e aplicações possíveis num ambiente integrado SIG-PACS/PSF.

6.1 CARACTERIZAÇÃO DA ÁREA E POPULAÇÃO DE ESTUDO

A área piloto pertence ao município de Caruaru-PE e corresponde à região do Juá, conforme mostrada na Figura 52.

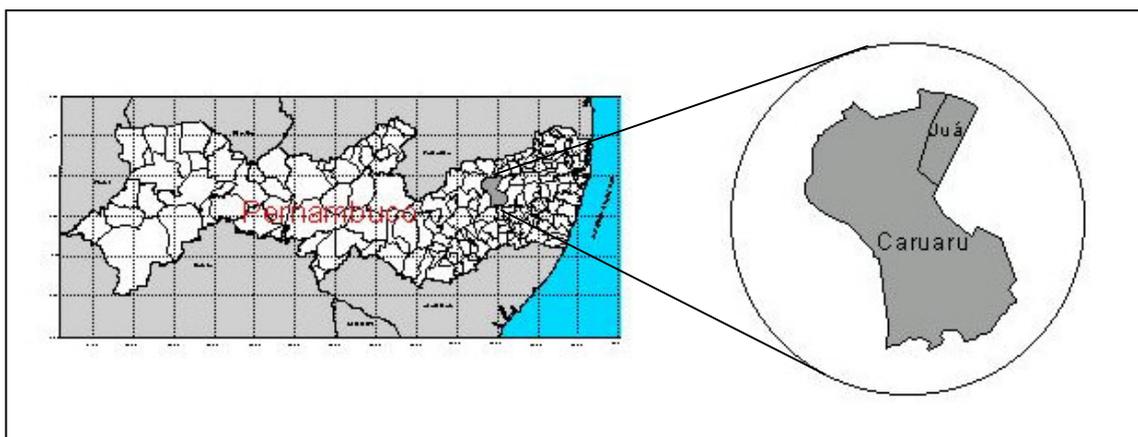


Figura 52 - Área Piloto

A cidade de Caruaru foi fundada em 1849, está localizada na região Agreste, no Vale do Ipojuca. Conta com uma área de 932 km², ocupando parte da Vertente da Cordilheira da Borborema, à margem esquerda do rio Ipojuca, onde se encontra distribuída uma população de 253.634 habitantes (IBGE, 2000), entre a zona urbana e rural. A temperatura média é de 26 graus centígrados, com uma altitude de 550 metros acima do nível do mar. Apesar de suas chuvas serem regulares, é parte integrante de área muito seca, com vegetação constituída pelo tipo de caatinga baixa, caracterizada por plantas arbustivas, como cactáceas e leguminosas. Seus limites são: ao

Norte, Toritama, Vertentes e Frei Miguelino; ao Sul Agrestina e Altinho; a Leste, Riacho das Almas e Bezerros; a Oeste São Caetano e Brejo da Madre de Deus.

A região do Juá é carente em termos econômicos e sua USF dista, aproximadamente 35 Km da sede do município. O acesso se dá por 25 Km de estrada asfaltada e os restantes 10 km, por estrada não pavimentada. Trata-se de área rural, com pequenos aglomerados residenciais. A população atendida pela USF local é de 2.058 pessoas (dado de 2001). A principal fonte de renda corresponde à confecções domiciliares, sobretudo de calças, que são vendidas na feira de Caruaru. Os costureiros trabalham em suas residências, utilizando máquinas emprestadas e recebendo em torno de R\$ 0,20 por calça pronta (dado de 2001). A agricultura é inconstante devido à seca e muitos moradores se candidatam a trabalhos nas frentes de emergência. O abastecimento de água é precário e, apesar de cerca de 60% das residências estarem ligadas à rede geral, é comum a falta d'água nas torneiras, passando o abastecimento a ser proveniente de barreiros (pequenos açudes que coletam água da chuva) e caminhões pipa, mantidos pelo município. A USF tem sua lotação de pessoal completa (1 médico, 1 enfermeira, 1 auxiliar de enfermagem, 5 agentes comunitários de saúde além de pessoal de apoio). Foi reestruturada em 1998 e levou 6 meses para completar o cadastro das famílias e adaptação ao Programa de Saúde da Família (informações fornecidas pela equipe). Vale ressaltar o criterioso trabalho de todos os componentes da equipe de saúde na confecção do cadastro e demais atividades. Essa boa organização da unidade facilitou os trabalhos de campo dessa pesquisa.

6.2 PROJETO LÓGICO IMPLEMENTADO

Detalha-se na Figura 53 o Diagrama Entidade-Relacionamento do projeto lógico implementado, que é parte do previsto no Capítulo 5 e sua respectiva implementação. O evento de saúde a ser trabalhado nesta aplicação será a hipertensão arterial, por ser a doença crônica mais prevalente na área, entre as que o PSF atua, e registrada no SIAB.

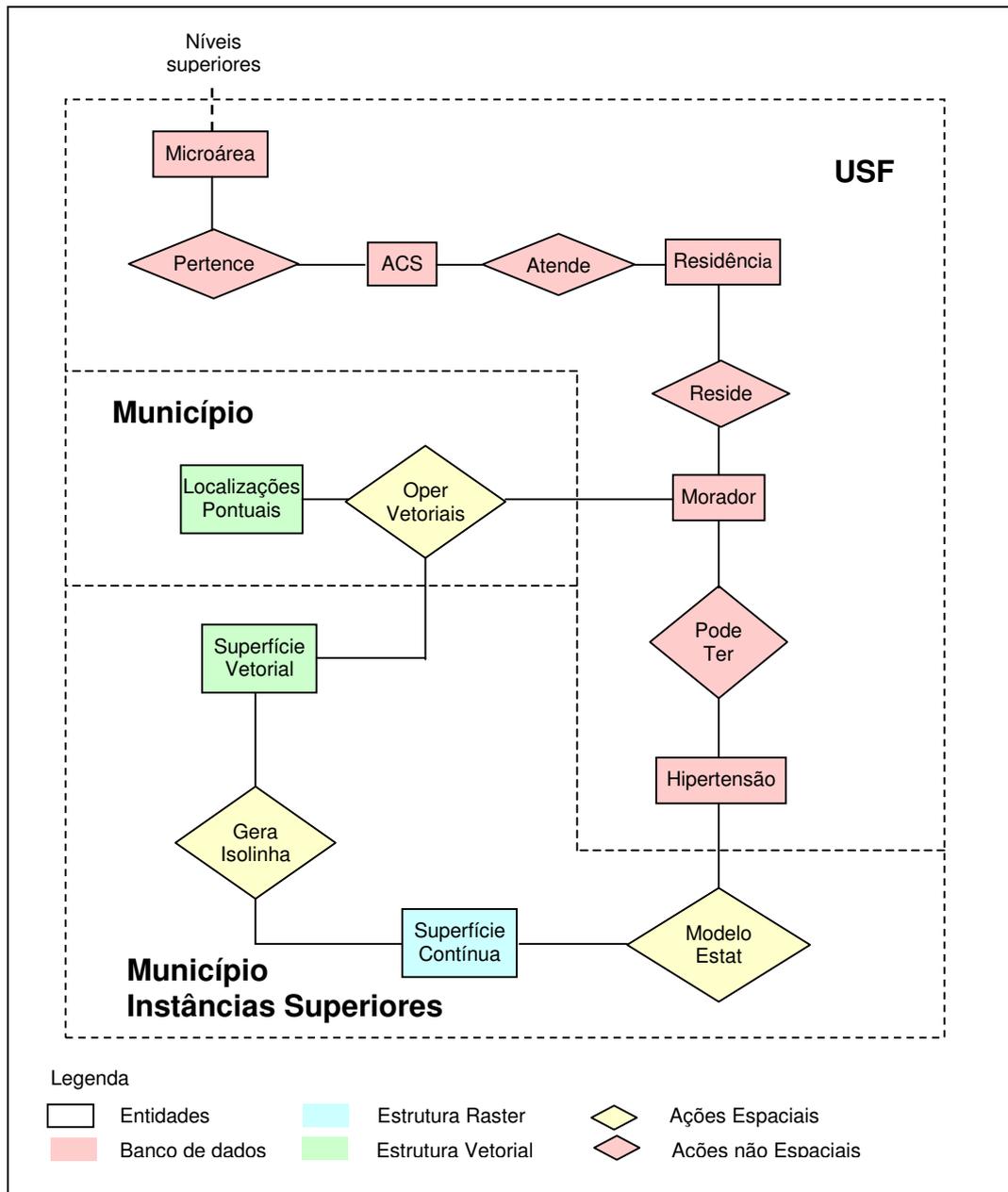
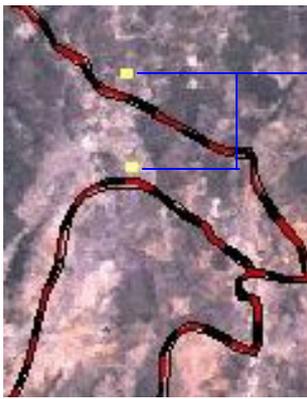


Figura 53 – Diagrama Entidade-Relacionamento Implementado

6.2.1 DICIONÁRIO DE DADOS

O dicionário de dados contempla, exclusivamente, as informações não espaciais, visto que as espaciais são um sub-conjunto do que foi especificado no Capítulo 5.

Tema: RESIDÊNCIA



Cobertura: Residência
 Tipo: Ponto
 Tabela de Atributos: Residência
 Elo: **Cod_Pront**
 Escala de Aquisição: 1:1 (GPS)

Arquivo Gráfico: Residência

Pano de Fundo: Caru.img
 Escala de Aquisição: 1:25000

Residência

Item	Tipo	Descrição
E	9,N,2	Coordenada este (identificador espacial)
N	10,N,2	Coordenada norte (identificador espacial)
Cod_Pront	4,N,0	Código do prontuário da família
Cod_ACS	45,C	Código do agente comunitário de saúde
Microárea	45,C	Microárea
Mor_c	1,N,0	Tipo de moradia (código)
Teto_c	1,N,0	Tipo de teto (código)
Ban_c	1,N,0	Existência de banheiro (código)
Piso_c	1,N,0	Tipo de piso (código)
Casa_c	1,N,0	Situação da casa (código)
Abas_c	1,N,0	Tipo de abastecimento d'água (código)
Trat_c	1,N,0	Tipo de tratamento d'água (código)
Dej_c	1,N,0	Destino dos dejetos (código)
Lix_c	1,N,0	Destino do lixo (código)
Ilum_c	1,N,0	Tipo de iluminação (código)
Anim_c	1,N,0	Existência de animais (código)
Aces_c	1,N,0	Acesso a terra (código)
Uso_c	1,N,0	Uso da terra (código)

Conteúdo de Uso_c

Cod	Uso_c
1	Usa
0	Não usa

Conteúdo de Aces_c

Cod	Aces_c
1	Proprietário
0	Outros

Conteúdo de Anim_c

Cod	Anim_c
1	Sim
0	Não

Conteúdo de Ilum_c

Cod	Ilum_c
1	Rede elétrica
0	Outros

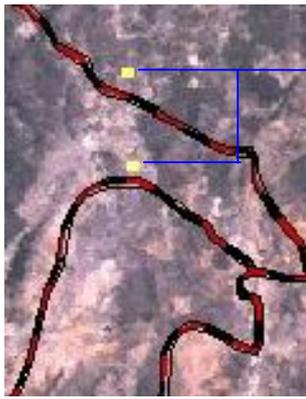
Conteúdo de Lix_c

Lix_c	Destino
1	Col.Pub/Enterrado/Queimado
0	Céu Aberto

Conteúdo de Dej_c

Cod	Dej_c
1	Rede esgoto/fossa
0	Outros

Tema: RESIDÊNCIA (cont)



Cobertura: Residência
 Tipo: Ponto
 Tabela de Atributos: Residência
 Elo: **Cod_Pront**
 Escala de Aquisição: 1:1 (GPS)

Arquivo Gráfico: Residência

Pano de Fundo: Caru.img
 Escala de Aquisição: 1:25000

Residência

Item	Tipo	Descrição
E	9,N,2	Coordenada este (identificador espacial)
N	10,N,2	Coordenada norte (identificador espacial)
Cod_Pront	4,N,0	Código do prontuário da família
Cod_ACS	45,C	Código do agente comunitário de saúde
Microárea	45,C	Microárea
Mor_c	1,N,0	Tipo de moradia (código)
Teto_c	1,N,0	Tipo de teto (código)
Ban_c	1,N,0	Existência de banheiro (código)
Piso_c	1,N,0	Tipo de piso (código)
Casa_c	1,N,0	Situação da casa (código)
Abas_c	1,N,0	Tipo de abastecimento d'água (código)
Trat_c	1,N,0	Tipo de tratamento d'água (código)
Dej_c	1,N,0	Destino dos dejetos (código)
Lix_c	1,N,0	Destino do lixo (código)
Ilum_c	1,N,0	Tipo de iluminação (código)
Anim_c	1,N,0	Existência de animais (código)
Ter_c	1,N,0	Possui terra (código)
Aces_c	1,N,0	Acesso a terra (código)
Uso_c	1,N,0	Uso da terra (código)

Conteúdo de Trat_c

Cod	Trat_c
1	Tratada
0	Não tratada

Conteúdo de Abas_c

Cod	Abas_c
1	Rede Geral
0	Outros

Conteúdo de Casa_c

Cod	Casa_c
1	Própria
0	Outras

Conteúdo de Piso_c

Cod	Piso_c
1	Tijolo/cimento/cerâmica
0	Areia/barro

Conteúdo de Ban_c

Cod	Ban_c
1	Sim
0	Não

Conteúdo de Teto_c

Cod	Teto_c
1	Telha
0	Outros

Conteúdo de Mor_c

Cod	Mor_c
1	Alvenaria
0	Outros

Tema: MORADOR

.....
.....

Cobertura: Indivíduo
 Tipo: Registro
 Elos: **Cód_Pront**

Arquivo Gráfico: Residência

Morador

Item	Tipo	Descrição
Cod_Pront	4,N,0	Código do prontuário da família
Cod_Pront_Mor	6,N,2	Código do prontuário/morador
Nome	45,C	Nome do morador
Sex_c	1,N,0	Sexo do morador (código)
Nas	Data	Data de nascimento
Renda_c	1,N,0	Renda do morador (código)
Esc_al	1,N,0	Escolaridade do morador
Ativ_c	1,N,0	Atividade Econômica
Par_c	1,C	Parentesco(código)
Hipertensão	1,N,0	Tem=1; Não tem=0
Diabetes	1,N,0	Tem=1; Não tem=0
Nat	45,N,0	Naturalidade

Conteúdo de Par_c

Cod	Parentesco
P	Pai
M	Mãe
A	Avô/avó
O	Nora
N	Neto
G	Genro
I	Irmão
S	Sobrinha
R	Prima

Conteúdo de Ativ_c

Cod	Ocupação
1	Desempregado
2	Autônomo
3	Empregado sem carteira assinada
4	Empregado com carteira assinada
5	Trabalhador eventual
6	Aposentado
7	Agricultor
8	Funcionário publico

Conteúdo de Esc_Al

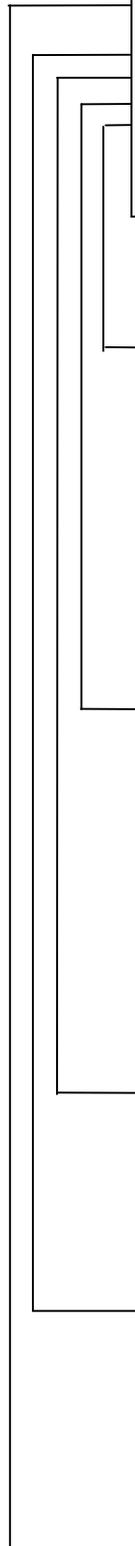
Cod	Esc_c
0	Analfabeto
1	Alfabetizado

Conteúdo de Renda_c

Cod	Renda
0	< 1 SM
1	≥ 1 SM

Conteúdo de Sex_c

Cod	Sexo
1	Feminino
0	Masculino



6.3 CARGA DO BANCO DE DADOS

A carga do banco de dados englobou trabalho de campo e sua validação em gabinete. Teve por base o projeto lógico implementado, prévia e exaustivamente testado por meio de simulações, antes de ser colocado em prática.

6.3.1 TRABALHO DE CAMPO

O trabalho de campo envolveu as fases de reconhecimento, de levantamentos espaciais e de levantamentos não espaciais.

- Reconhecimento

Essa fase se caracterizou por visita prévia à região do Juá onde aconteceu o primeiro contato com a USF e seus membros. Foi verificado o funcionamento e organização da unidade, explicado aos membros da equipe a finalidade da pesquisa e definido, em comum acordo, como se procederiam aos levantamentos.

- Levantamentos Espaciais

Caracterizou-se pela obtenção das coordenadas UTM das residências, definidas como pontos, com precisão de determinação da ordem de 0,5 metros. Como não se dispunha de mapeamento prévio compatível com esse valor, a solução encontrada foi o emprego de rastreadores de satélites GPS (Kennedy, 1996). O equipamento utilizado foi o modelo 4000-SST da Trimble, de dupla frequência e o método de determinação foi pseudo-cinemático com a técnica *Stop and Go* (Santos, 2001; Leick, 1990).

A escolha do equipamento deu-se pela sua elevada precisão e por permitir que os pontos levantados possam ser nomeados a critério do usuário, no momento de seu rastreamento. Esse último aspecto foi relevante do ponto de vista prático, uma vez que cada residência, por ocasião do levantamento, pode receber como nome seu código do prontuário da família.

A escolha da técnica *Stop and Go* deve-se à rapidez permitida para o trabalho e está esquematizada na Figura 54. Nela são empregados 2 rastreadores, sendo um deles instalado no ponto fixo de coordenadas (E,N) conhecidas e permanece ligado durante toda a jornada. O outro é chamado itinerante e é o que percorre as residências, cujas coordenadas se quer determinar, devendo permanecer em torno de 1 minuto em cada uma, desde que não se perca o sinal do satélite por ocasião do deslocamento entre elas.

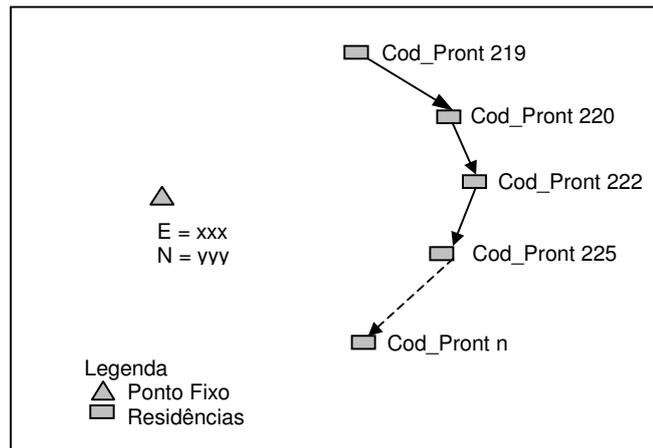


Figura 54 - Método Stop and Go

Terminada a jornada do dia, os dados dos dois equipamentos são descarregados, pós-processados e o resultado final é uma tabela com três atributos: Coordenada E, Coordenada N e Código do Prontuário da Família, conforme mostrado na Figura 55.

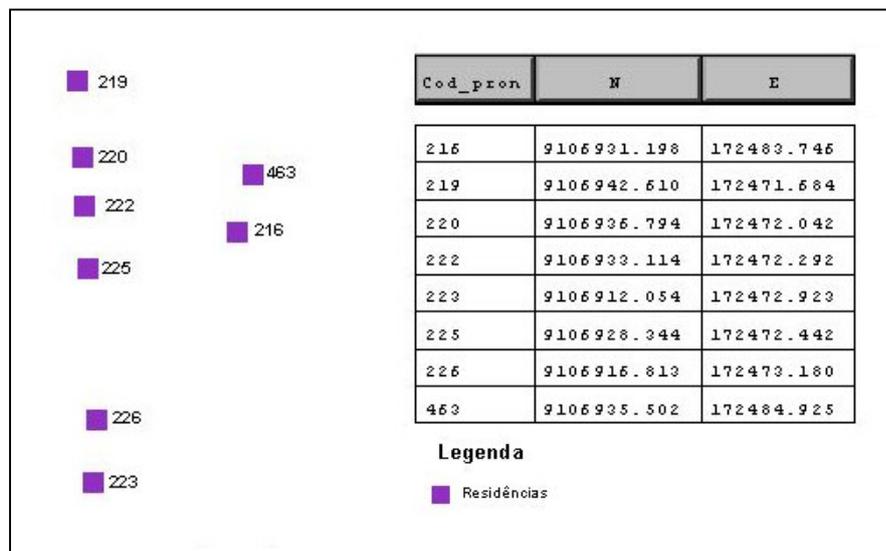


Figura 55 - Resultados dos levantamentos espaciais

Convém observar a importância dessa tabela que é formada essencialmente por elos: as coordenadas E,N são responsáveis pela ligação das residências com o mapa e o Código do Prontuário responsável pela ligação das residências com as informações não espaciais.

O levantamento completo da região foi executado em 25 dias úteis.

- Levantamentos não espaciais

Corresponde à digitação dos dados não espaciais em um programa de banco de dados. Tais informações estavam organizadas na USF em envelopes e fichas. Cada envelope continha informações sobre determinada família e era identificado pelo seu código de prontuário. Em cada envelope existiam tantas fichas quantos fossem os componentes da família, que eram identificados pelo código de prontuário do morador e continham informações individuais, registradas entre 1998 e 2002 pela ESF. Com base nessa organização, estruturou-se um banco de dados conforme mostrado na Figura 56.

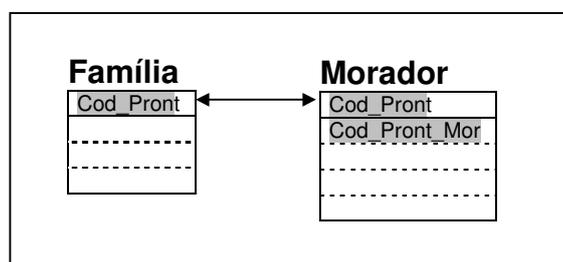


Figura 56 - Estrutura do Banco de Dados

Em termos práticos, essa estrutura permite identificar, univocamente, cada morador através da chave **Cod_Pront_Mor** e também efetuar a ligação deste com a família a que pertence através do elo **Cod_Pront**.

A digitação deu-se na própria USF porque os envelopes e as fichas são manuseados diariamente, por ocasião do atendimento ao público. Para facilitar, criou-se um formulário conforme mostrado na Figura 57. O tempo necessário para essa atividade foi de 10 dias úteis, com 2 pessoas trabalhando.

The screenshot shows a software window titled "Residencia" with two main sections: "Residencia" and "Morador".

Residencia Section:

- Cod_Pront: 102
- Ano: 98
- Tipo_Mor: 1
- Teto: 1
- Ban: 2
- Piso: 1
- Casa: 1
- Ab_Agua:
- Trat_Ag: 3
- Dest_Dej: 1
- Dest_Lix: 1
- Ilum: 3
- P_Anim: 1
- Pos_Ter: 1
- Ac_Ter: 1
- Uso_Ter: 3
- Gest: 0
- Tub: 0
- Hans: 0
- S_Ment: 1
- P_Cancer: 0
- P_Fam: 0
- Hiper: 1
- Diab: 0
- Risco: 0
- T_Anim: 1
- T_Plant: MILHO FEIJOAO PALM

Morador Section:

- Cod_ACS: 3
- Cod_Pront: 102
- Cod_Pront_Mor: 102.1
- Nome: JOAO RENOVARO DO NASCIM
- Parent: P
- Nasc: 18/12/22
- Sexo: M
- Est_Civil: C
- Natural: CARUARU
- Escolar: 3
- Ativ_Econ: APOSENTADO
- Renda: 3
- Hiper_T: 1
- Diab:

At the bottom right, there is a "Registro:" field showing "1 de 7".

Figura 57 - Formulário

6.3.2 VALIDAÇÃO DOS TRABALHOS DE CAMPO

A validação dos trabalhos de campo envolveu as fases de revisão dos levantamentos e implementação do modelo georrelacional propriamente dito.

A revisão dos levantamentos teve por objetivo identificar possíveis erros de digitação e perdas de informação. Sobre erros de digitação, a quase totalidade pode ser resolvida em gabinete, o que não aconteceu no que diz respeito à perda de informações. Nesse caso, foram detectados 28 prontuários da família sem a informação espacial (coordenadas E,N), exigindo retorno ao campo e novo levantamento para esses casos.

A implementação do modelo georrelacional propriamente dito, deu-se pela ligação das informações espaciais às não espaciais, conforme esquematizado na Figura 58 e já detalhado no Dicionário de Dados.

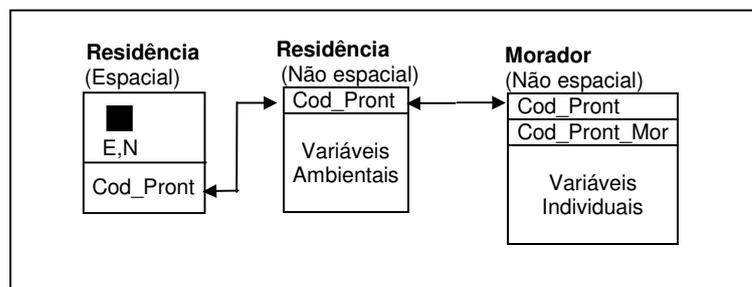


Figura 58 - Esquema do modelo georrelacional

6.4. EXPLORAÇÃO DO SISTEMA

Uma vez implementado o modelo georrelacional torna-se possível realizar as 8 operações previstas num sistema de banco de dados⁷ e também é permitido recuperar informações não espaciais, tendo como argumento de pesquisa informações espaciais, bem como recuperar informações espaciais, tendo como argumento de pesquisa informações não espaciais.

Sugere-se que na USF, a exploração do sistema restrinja-se a operações com informações não espaciais, cujo é permitir que a equipe de saúde identifique, na população da área adscrita, situações de risco. Tal exploração pode se processar pelo exame de tabelas e gráficos dos dados coletados de modo à formar um quadro da realidade mais confiável que a usual impressão qualitativa. Vale aqui ressaltar que os programas de computador exigidos para essas tarefas são inúmeros e muitos deles são disponibilizados gratuitamente.

A título de ilustração passa-se a seguir à uma análise exploratória dos dados de Juá.

Inicialmente mostra-se na Tabela 1 as distribuições de freqüência das variáveis disponíveis, como forma de fazer a descrição inicial da área, segundo algumas características demográficas, sócio-econômicas, sócio-culturais e sanitárias da população.

⁷ Ver Capítulo 3

Tabela 1 - Distribuições de freqüência das variáveis disponíveis, Juá, 2001

Variável	n	%	Variável	n	%
Hipertensão arterial			Destino do lixo		
Tem	110	5,3	Coleta publica, enter- rado, queimado	875	42,5
Não tem	1948	94,7	Céu aberto	1183	57,5
Diabetes			Destino dos dejetos		
Tem	23	1,1	Rede de esgoto/fossa	887	43,1
Não tem	2035	98,9	Outros	1171	56,9
Sexo			Tratamento d'água		
Feminino	1041	50,6	Tratada	1890	91,8
Masculino	1017	49,4	Não tratada	168	8,2
Escolaridade			Abastecimento d'água		
Alfabetizado	1092	53,1	Rede geral	1237	60,1
Analfabeto	966	46,9	Outros	821	39,9
Renda			Casa		
Maior ou igual a 1 SM	430	20,9	Própria	1840	89,4
Menor que 1 SM	1628	79,1	Outros	218	10,6
Uso da terra			Banheiro dentro de casa		
Usa	1577	76,6	Sim	943	45,8
Não usa	481	23,4	Não	1115	54,2
Acesso a terra			Piso da casa		
Proprietário	456	22,2	Tijolo, cimento, cerâmica	1831	89
Não proprietário	1602	77,8	Areia, barro	227	11
Animais			Teto da casa		
Possui	1132	55	Telha	2032	98,7
Não possui	926	45	Outros	26	1,3
Iluminação			Revestimento da casa		
Rede elétrica	1932	93,9	Alvenaria	1931	93,8
Outros	126	6,1	Outros	127	6,2

Os dados mostrados na Tabela 1, são de fundamental importância para o conhecimento do perfil da área de atuação da equipe de saúde. Como pode ser observado, a USF do Juá possui uma população adscrita de 2058 pessoas, com ligeira maioria de mulheres, elevado percentual de analfabetismo (46,9%) e de indivíduos com renda mensal inferior a 1 salário mínimo (79,1%). Cerca de 94% de sua população mora em casa de alvenaria ligada à rede elétrica de energia. Quanto a prevalência de doenças crônicas que são objeto do PSF, não foram detectados portadores de hanseníase e tuberculose, entretanto, foram constatados 110 indivíduos com hipertensão arterial (5,3% da população) e 23 com diabetes (1,1%).

As informações constantes da Tabela 1 podem ser combinadas e visualizadas na forma gráfica. Por exemplo, verifica-se na Figura 59 a distribuição da população do Juá por faixa etária e sexo.

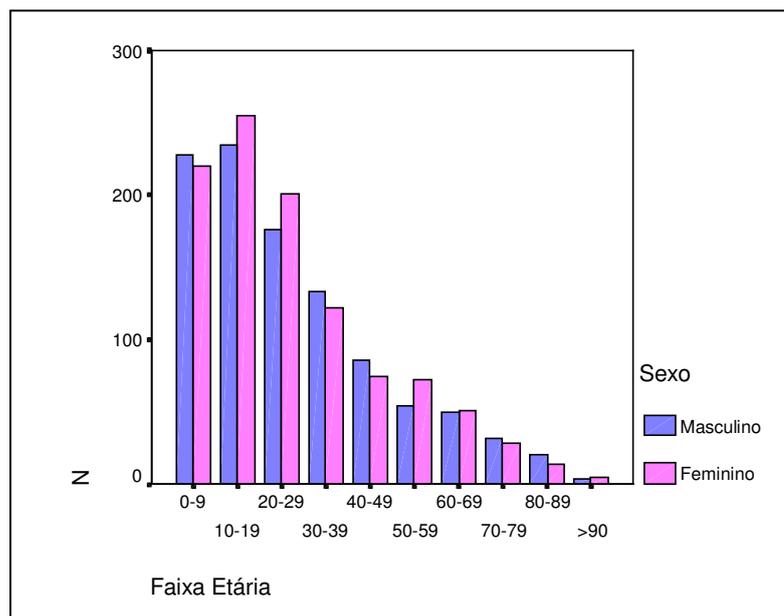


Figura 59 – Distribuição da população por faixa etária e sexo – Juá, 2001

A Figura 60, identifica a distribuição da população do Juá por faixa etária, segundo a renda mensal em termos monetários. Pode ser observado que até a faixa etária de 50-59 anos, predomina a renda individual inferior a 1 salário mínimo e, a partir de então, a situação se inverte.

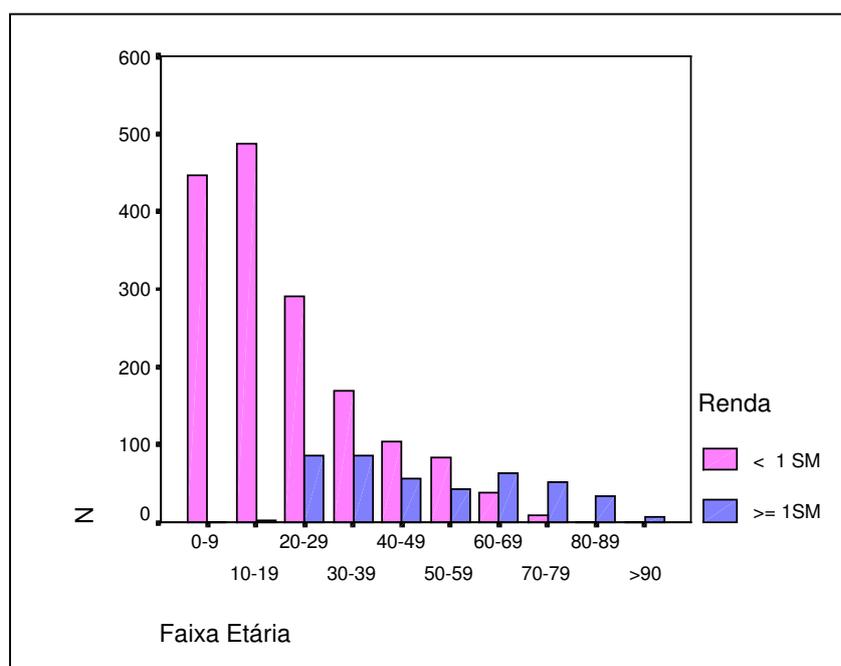


Figura 60 – Distribuição da população por faixa etária e renda, Juá, 2001

A Figura 61, identifica a distribuição da população do Juá, portadora de hipertensão arterial, por faixa etária e sexo. Nela observa-se que o número de hipertensos é maior nas mulheres, até a faixa etária de 80-89 anos. A partir dos 60 anos, a diferença entre os sexos diminui consideravelmente em termos de número absoluto de casos.

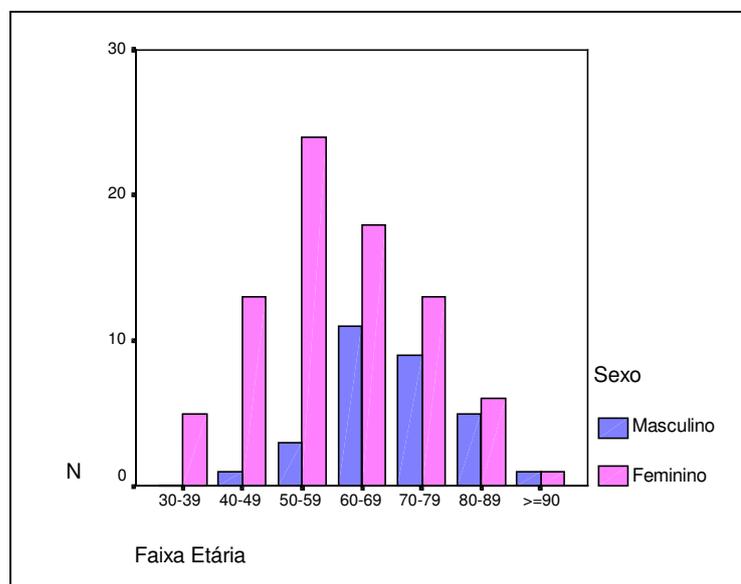


Figura 61 – Distribuição dos hipertensos por faixa etária e sexo, Juá, 2001

A Figura 62, identifica a distribuição da população do Juá por faixa etária, segundo o nível de escolaridade. Nela observa-se que a população até os 39 anos de idade é alfabetizada em sua maior parte, invertendo-se a situação a partir daí.

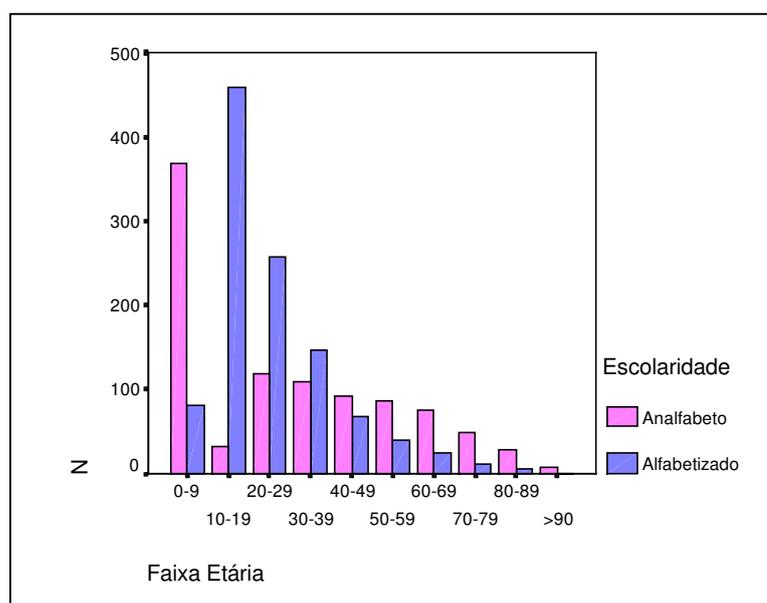


Figura 62 - Distribuição da população por faixa etária e escolaridade – Juá, 2001

Observando-se a Figura acima, verifica-se que existem 80 crianças com idade inferior a 9 anos alfabetizadas. A partir daí, é permitido conhecer a totalização por agente e respectiva média etária, na forma de tabela, conforme mostrado na Figura 63.

Cod. Agente	Número de crianças	Idade Média
1	17	8.1176
2	28	7.7857
3	20	8.2000
4	15	8.2667

Figura 63 – Totalização de crianças, por agente, com idade inferior a 9 e anos alfabetizadas, Juá, 2001

Os números acima indicam que a microárea do agente 2 possui o maior número de crianças alfabetizadas (28) e a menor média de idade (7,7857). Ressalta-se que o único estabelecimento escolar do Juá está localizado nessa microárea.

6.4.1. ANÁLISE ESPACIAL EM INSTÂNCIAS SUPERIORES

Os dados levantados pelo PSF merecem, necessariamente, análise em áreas maiores, abrangendo várias unidades ou mesmo municípios. O tipo de análise possibilitado nestes níveis é diferenciado, em parte pelo grau de sofisticação nos modelos, em parte por ter um volume maior de dados. Do ponto de vista da estrutura dos serviços de saúde, as instâncias superiores seriam as secretarias de saúde do município, as regionais de saúde, as secretarias de saúde do estado e o Ministério de Saúde. Propõe-se que os dados coletados nas áreas sejam repassados, na íntegra, para o município e daí para os níveis hierárquicos superiores, seguem somente os consolidados, conforme previsto pelo SIAB.

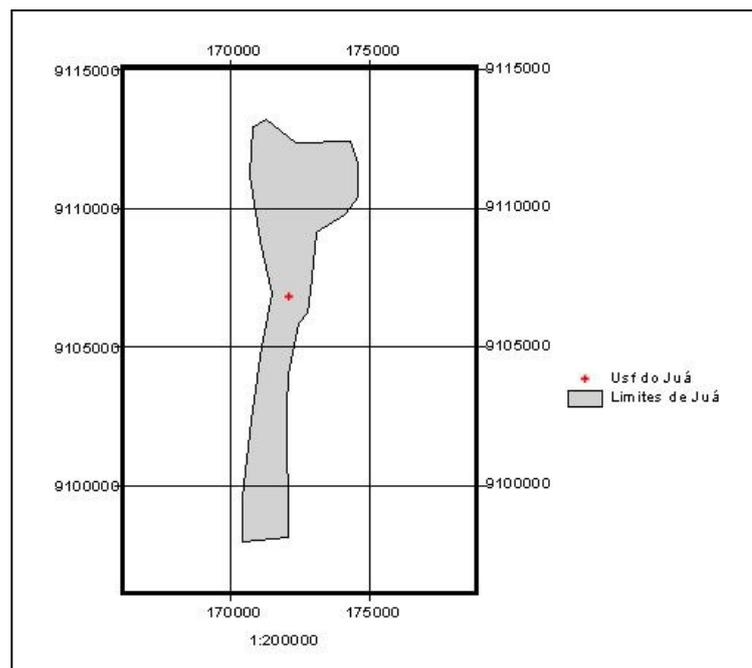
Com base nessa premissa, observa-se que a análise exploratória básica, executada na USF, também pode e deve se processar a nível municipal, propondo-se ainda para essa instância, o emprego de técnicas de análise espacial, utilizando funções básicas de SIG, sobre as coordenadas das residências.

É evidente que tais análises, sejam estas realizadas na própria unidade ou em nível superior, devem ser discutidas com a equipe local.

As técnicas de análise espacial sugeridas referem-se às operações vetoriais previstas no DER da Figura 53, caracterizam-se por disponibilizar informações de localização, realizadas com relativa simplicidade num ambiente SIG, como as exemplificadas a seguir:

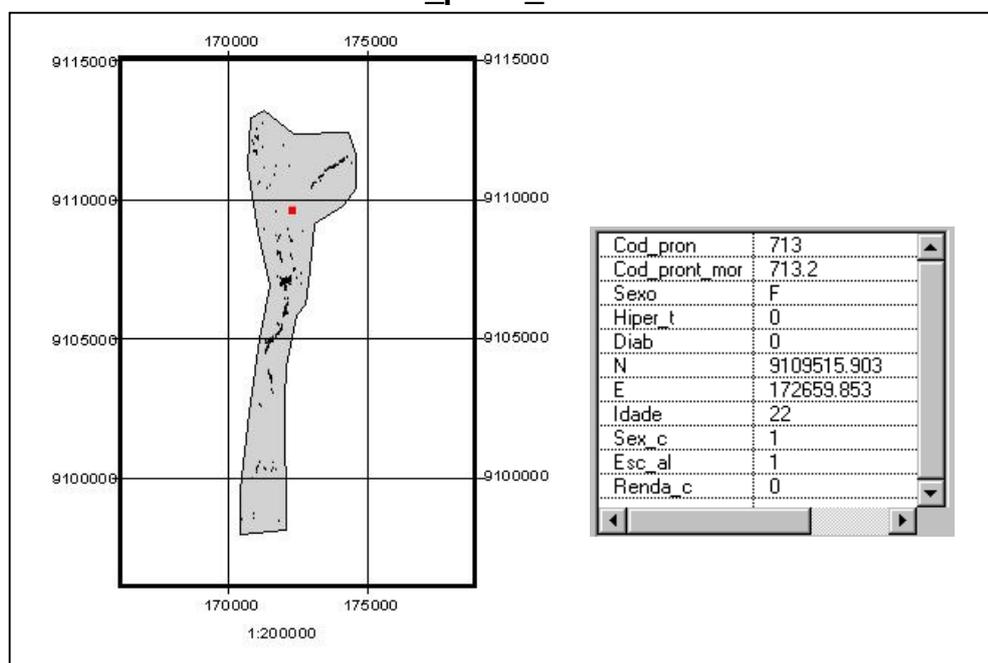
1. Onde se localiza a USF do Juá?

Selecione Tema: “USF do Juá”



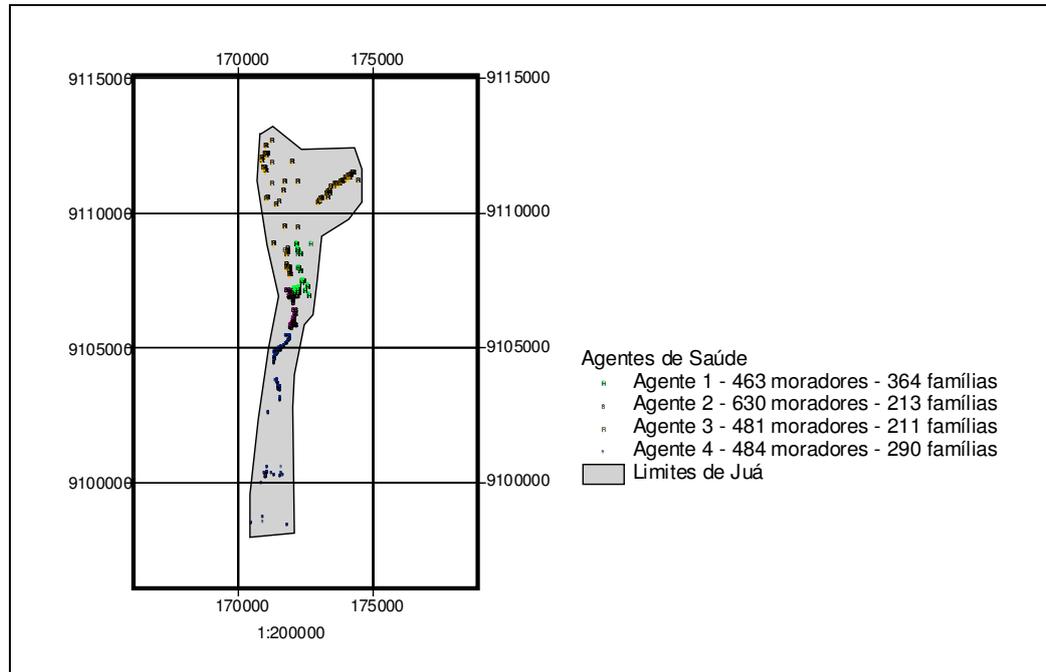
2. Onde se localiza a pessoa cujo código de prontuário é “y” e quais suas informações disponíveis?

Selecione: “Morador.Cod_pront_mor = 713.2”



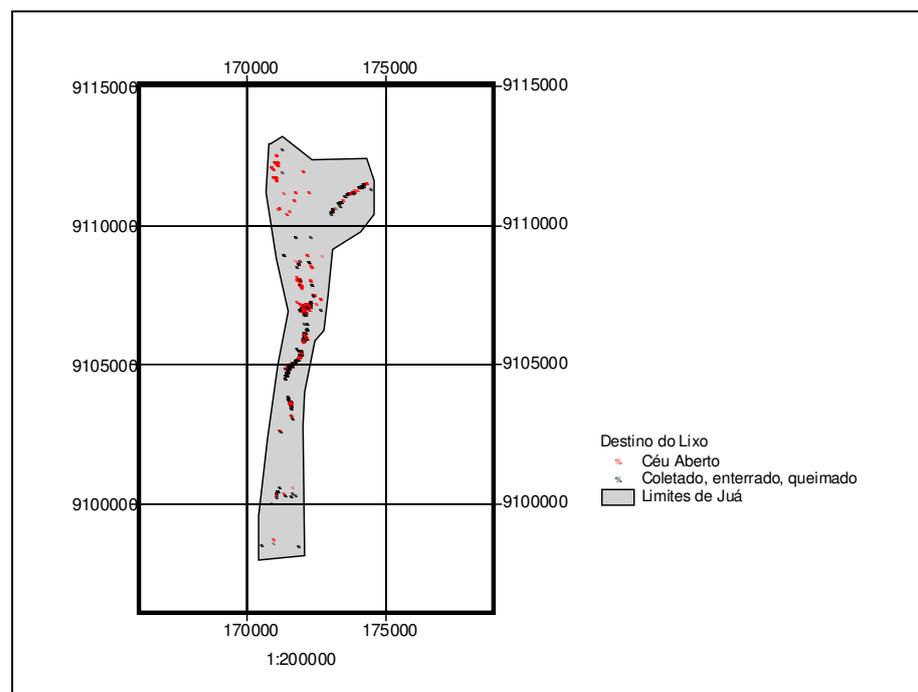
3. Onde se localizam as residências atendidas pelos agentes comunitários, quantas famílias e quantos moradores são atendidos por eles?

Classifique: **"Morador.Cod_acs"**



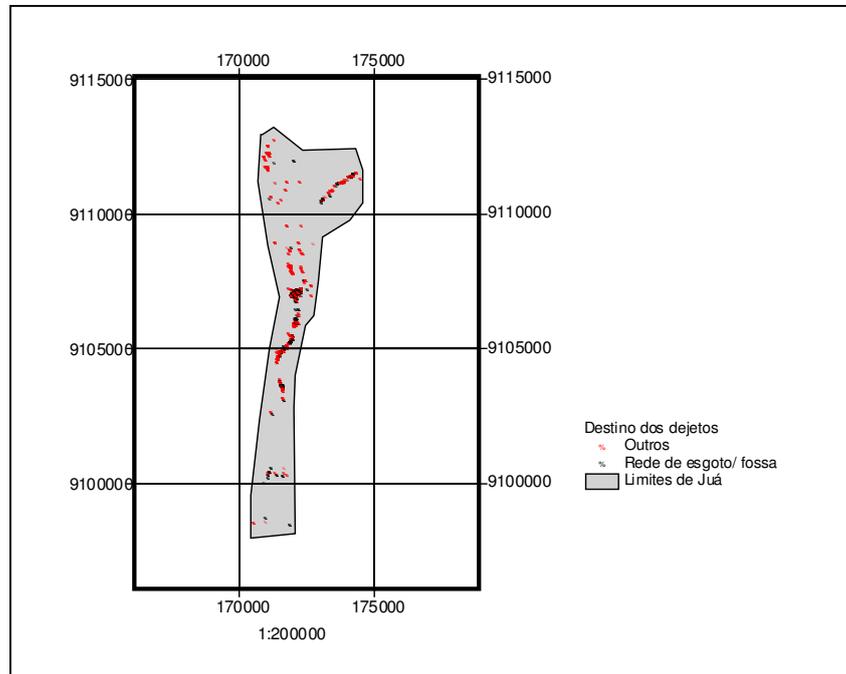
4. Onde se localizam as residências de acordo com o destino do lixo?

Selecione: **"Morador.Lix_c"**



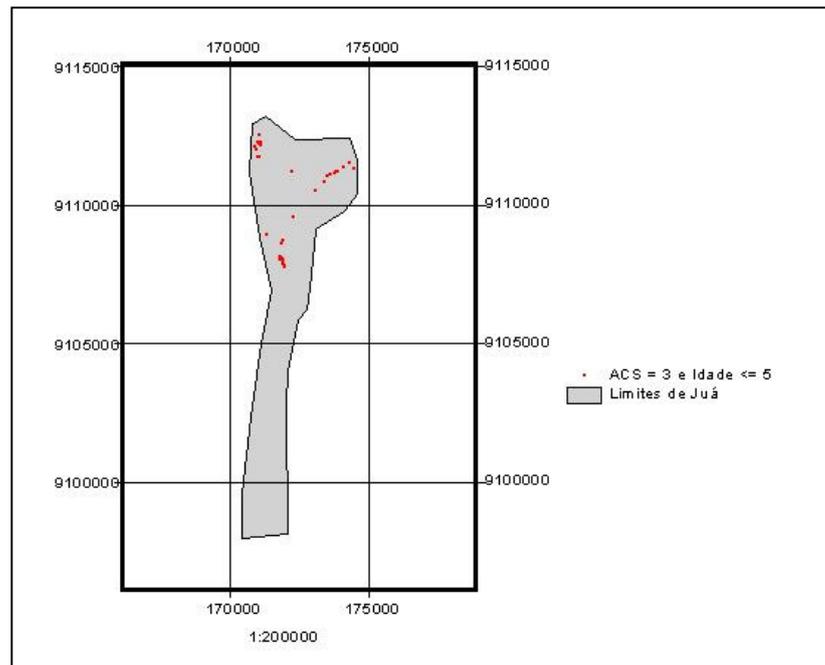
5. Onde se localizam as residências de acordo com o destino dos dejetos?

Selecione: **“Morador.Dej_c”**

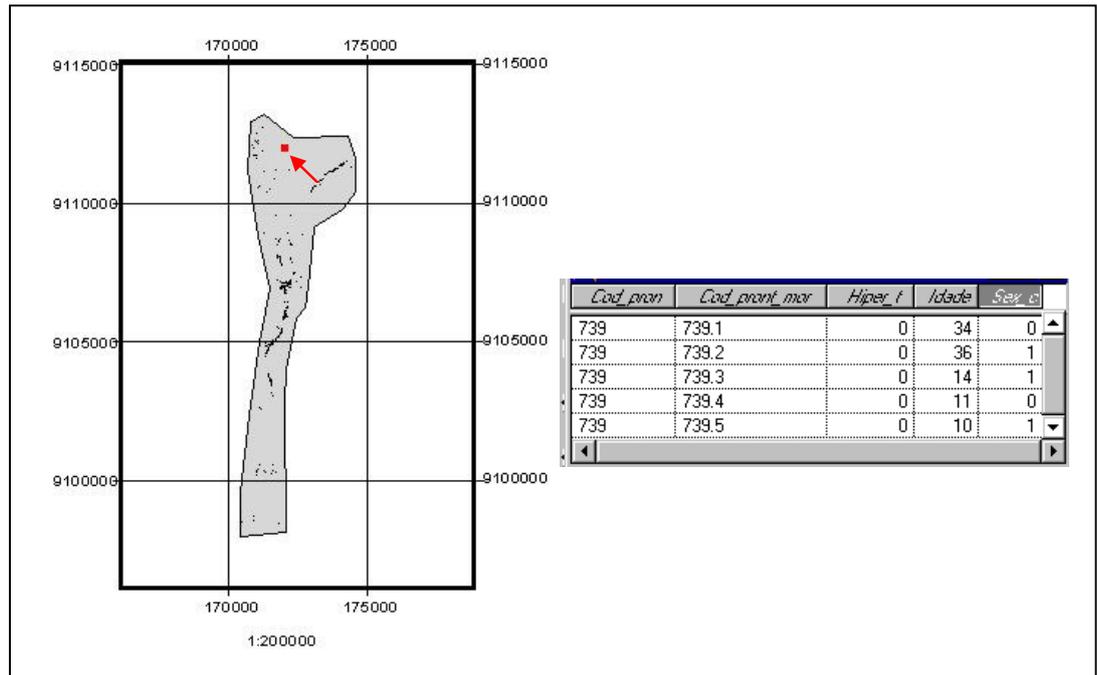


6. Onde se localizam as crianças menores de 5 anos, atendidas pelo agente comunitário de saúde cujo código é “z”?

Selecione: **“Morador.Cod_acs = 3” e “Morador.Idade ≤ 5”**



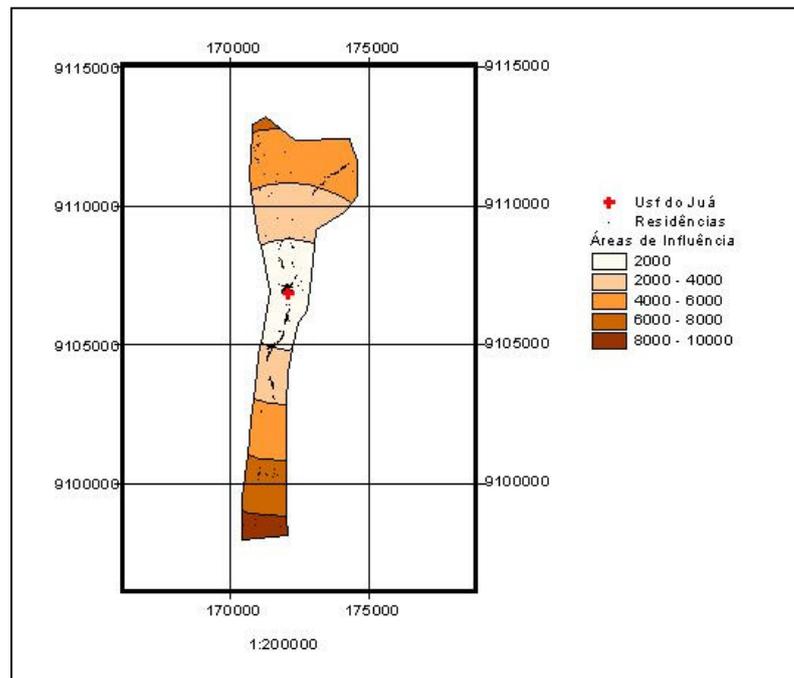
7. Quais são as pessoas que moram numa residência identificada no mapa?



8. Identifique áreas de influência que distam 2, 4, 6 e 8 Km da USF.

Selecione Tema: "USF do Juá"

Crie: "Buffer de 2 a 8, intervalo 2"



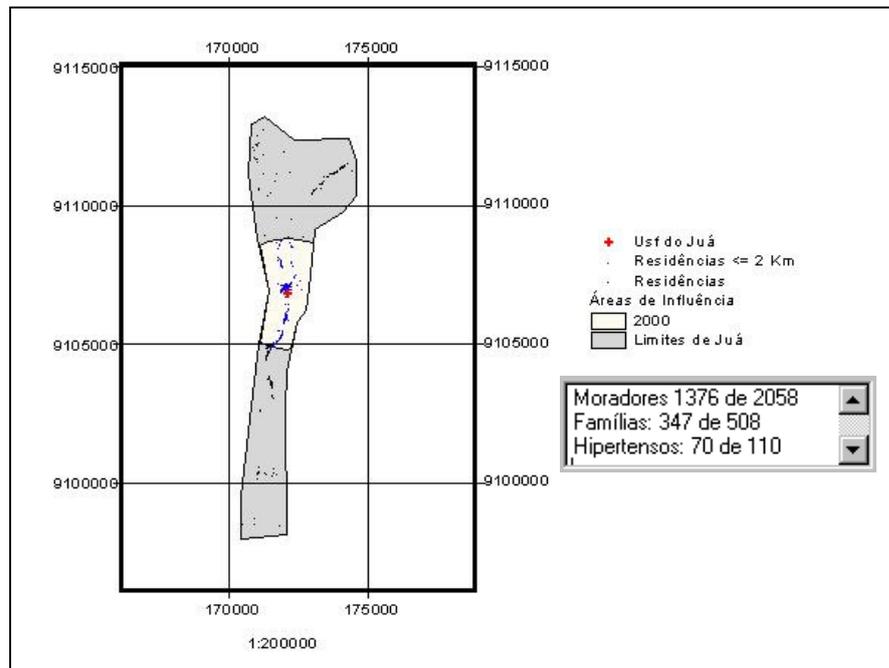
9. Quantos moradores, quantas famílias e quantos hipertensos residem a uma distância menor ou igual a 2 Km da USF?

Selecione Tema: "USF do Juá"

Crie: "Buffer de 0 a 8 intervalo 2"

Selecione: "Morador.Cod_pront_mor" contido em "Buffer de 0 a 2"

Estatística: "Morador.Cod_pront_mor"



Cabe ainda a observação que, da forma como está estruturado o banco de dados, é possível identificar o cidadão com todos os seus atributos, inclusive nome e local de residência. Por motivos éticos, esse tipo de consulta não foi exemplificado.

7. ESTUDO DE CASO: A HIPERTENSÃO ARTERIAL NO JUÁ

A hipertensão arterial é um importante problema de saúde pública, com elevada prevalência na população economicamente ativa e com alta participação nos óbitos acima de 50 anos. Estimativas dessa prevalência indicam que, em algumas regiões brasileiras, seus valores variavam entre 11% e 25%, na década de 90 (Faerstein et al, 2001), podendo ser decorrente do aumento da frequência de alguns fatores de risco, como aspectos da dieta (Mondini & Monteiro, 1994), excesso de peso (Coitinho et al, 1991) e estresse de natureza psicossocial (INCLLEN, 1994; Marmot et al, 1978). Além disso, esta patologia consome grande volume de recursos assistenciais: 13% das internações do Sistema Único de Saúde e 25% da verba para assistência à saúde (MS,2000; Mion Jr et al, 2001).

Apesar da alta prevalência, estima-se que no país, somente cerca de 30% dos casos estão controlados, o que se relaciona a: idade, sexo, escolaridade, nível sócio-econômico, ocupação, distância ao serviço de saúde, ausência de sintomas, relacionamento com a equipe de saúde (Mion Jr & Pierin, 2001).

Com base nestas considerações, escolheu-se este agravo para o presente estudo de caso, onde se procurou desenvolver um conjunto de análise de dados, que sirvam de exemplo das potencialidades do modelo proposto. Parte dessas análises pode ser realizada pela própria ESF, com base no banco de dados disponível e outra parte, em instâncias superiores das Secretarias de Saúde, por exigir conhecimento mais sofisticado, utilizando funções básicas dos programas de SIG⁸.

Saindo deste nível mais simples de uso do SIG, diversas aplicações podem ser desenhadas. A aplicação apresentada a partir desse momento se restringe ao estudo da hipertensão, analisada com métodos estatísticos de processo pontual marcado⁹.

Vale ressaltar que não foram questionados os procedimentos empregados para aferição da pressão arterial nem os critérios de diagnóstico da doença. Partiu-se do princípio que estes estão corretos e as

⁸ Análises exemplificadas no Capítulo 6

⁹ Ver Capítulo 4

peças identificadas como hipertensas nos prontuários foram assim consideradas.

O primeiro passo, para implementar o processo pontual marcado, é identificar entre as variáveis disponíveis¹⁰, quais delas podem ser consideradas de risco para a hipertensão arterial. Nesta etapa seguem-se as rotinas usuais de modelagem multivariada, até que seja selecionado um conjunto de variáveis explicativas, desconsiderando-se a componente espacial. Isso pode ser feito em qualquer software que opere com modelos logísticos.

A Tabela 2 mostra os resultados da regressão, tendo por variável dependente a hipertensão arterial (ter ou não ter) e por independentes todas as demais.

Tabela 2 – Resultados da Regressão com todas as variáveis

Variável	β	Erro estimado de β	Nível de Significância
Idade	0.0819	0.0082	0.0000
Sexo	1.4371	0.2759	0.0000
Renda	-0.2433	0.3056	0.4261
Escolaridade	0.0131	0.2934	0.9643
Diabetes	2.7069	0.5984	0.0000
Uso da terra	-0.2831	0.3171	0.3720
Acesso a terra	0.3624	0.2798	0.1952
Animais	0.2648	0.2843	0.3515
Iluminação	0.7080	0.6576	0.2816
Destino do lixo	0.2662	0.2595	0.3048
Destino dos dejetos	0.2241	0.5045	0.6570
Tratamento d'água	-0.1915	0.4158	0.6451
Abastecimento d'água	-0.4466	0.2683	0.0960
Piso da casa	0.4535	0.5230	0.3859
Casa	-0.1696	0.6315	0.7883
Teto da casa	5.0977	9.5827	0.5947
Banheiro dentro da casa	-0.1032	0.4982	0.8359
Revestimento da casa	0.6047	0.5694	0.2882
Constante	-13.7778	9.6683	0.1541

Os valores acima, indicam que as variáveis **Idade, Sexo, e Diabetes** são as únicas significativas ao nível de 5%, entretanto a literatura identifica que além dessas, **Renda e Escolaridade** também são influentes (Mion Jr et al, 2001; Lotufo, 2001).

Por ocasião da digitação do banco de dados, verificou-se que a maior parte dos jovens era alfabetizada e na sua maioria não possuía renda.

¹⁰ Ver Tabela 1

Verificou-se também a existência de muitos idosos aposentados com renda de 1 salário mínimo ou pouco mais.

Tais observações fornecem indícios que as variáveis **Idade**, **Escolaridade** e **Renda** podem estar interagindo. Essa hipótese é testada e os resultados são apresentados na Tabela 3.

Tabela 3 - Resultados da Regressão com todas as variáveis e interações

Variável	β	Erro estimado de β	Nível de Significância
Constante	-13.6214	9.5068	0.1519
Idade	0.0791	0.0121	0.0000
Sexo	1.4589	0.2716	0.0000
Renda	1.7246	0.9273	0.0629
Escolaridade	-2.5205	0.8909	0.0047
Diabetes	2.7083	0.5845	0.0000
Uso da terra	-0.3319	0.3163	0.2940
Acesso à terra	0.4597	0.2820	0.1031
Animais	0.0827	0.2851	0.7718
Iluminação	0.6753	0.6611	0.3070
Destino do lixo	0.2410	0.2612	0.3561
Destino dos dejetos	0.2361	0.5109	0.6439
Tratamento d'água	-0.1237	0.4202	0.7684
Abastecimento d'água	-0.4337	0.2703	0.1086
Piso da casa	0.4280	0.5085	0.3999
Casa	-0.0769	0.6176	0.9009
Teto da casa	5.2819	9.4104	0.5746
Banheiro dentro de casa	-0.2114	0.5023	0.6739
Revestimento da casa	0.6595	0.5770	0.2531
Escolaridade x Idade	0.0473	0.0155	0.0023
Renda x Idade	-0.0329	0.0155	0.0332

Analisando-se os resultados acima, pode-se concluir o seguinte:

- As interações sugeridas se confirmaram;
- A variável **Escolaridade**, quando controlada por **Idade** assume caráter protetor para a hipertensão, materializada pelo sinal negativo de seu coeficiente;
- A variável **Renda**, quando controlada por **Idade** é estatisticamente significativa ao nível de 6.29% e identifica que os moradores de maior renda são os mais propensos à hipertensão.

Uma vez selecionadas as variáveis passa-se a considerar a componente espacial de acordo com a Equação 10. Sobre a largura de banda, escolheu-se o valor de 1000m, que corresponde a $\frac{1}{4}$ da menor dimensão do retângulo que circunscreve a região do Juá.

$$y = \beta_0 + \beta_1 idade + \beta_2 sexo + \beta_3 diabetes + \beta_4 escolaridade + \beta_5 renda + \beta_6 idade * renda + \beta_7 idade * escolaridade + g(s) \quad (10)$$

O resultado desse modelo é apresentado no Tabela 4.

Tabela 4- Resultados da regressão incluindo-se a variável espaço

Variáveis	β	Erro estimado de β	Nível de significância
(Intercept)	-7.15494963	0.380006391	0.000000e+00
Idade	0.07967457	0.006769248	0.000000e+00
Sexo	1.44354021	0.155565946	0.000000e+00
Diabetes	2.77909772	0.335098865	2.220446e-16
Escolaridade	-2.37720190	0.507582337	3.006763e-06
Renda	1.59765093	0.531239855	2.666963e-03
Idade:Escolaridade	0.04598330	0.008771268	1.747104e-07
Idade:Renda	-0.03053218	0.008801584	5.333056e-04

Deve ser observado que a partir da inclusão da variável espaço, todas as demais apresentaram nível de significância inferior a 5%, melhorando-se assim os resultados estatísticos em relação à situação anterior.

Antes de concluir pela aceitação desse modelo, convém verificar como se comportam os coeficientes e seus níveis de significância, variando-se a largura da banda. Para isso, ao invés de empregar-se validação cruzada¹¹, testam-se os valores de 500m, 1000m, 1500m e 2000m e os resultados são mostrados na Tabela 5.

Tabela 5 – resultados obtidos para diversos valores de largura de banda

Variável	h = 500 m		h = 1000 m		h = 1500 m		h = 2000 m	
	B	Pr(> t)	B	Pr(> t)	B	Pr(> t)	B	Pr(> t)
(Intercept)	-7.32	0.0	-7.15	0.0	-7.11	0.0	-7.08	0.0
Idade	0.08	0.0	0.08	0.0	0.08	0.0	0.08	0.0
Sexo	1.46	0.0	1.44	0.	1.44	0.0	1.44	0.0
Diabetes	2.77	0.0	2.78	2.22e-16	2.77	4.44e-16	2.75	4.44e-16
Escolaridade	-2.29	5.70e-06	-2.38	3.00e-06	-2.40	2.35e-06	-2.41	2.06e-06
Renda	1.63	2.02e-03	1.60	2.66e-03	1.59	2.68e-03	1.58	2.72e-03
Idade x Escolaridade	0.04	3.61e-07	0.05	1.74e-07	0.05	1.11e-07	0.05	8.41e-08
Idade x Renda	-0.03	2.62e-04	-0.03	5.33e-04	-0.03	6.34e-04	-0.03	6.70e-04

Diante desses resultados, verifica-se que as alterações ocorridas nos valores dos coeficientes e dos respectivos níveis de confiança são, praticamente, irrelevantes, podendo-se aceitar como válido o modelo identificado pela Equação 10, com largura de banda de 1000m.

A partir desse modelo, passa-se a obtenção da superfície contínua, representando a estimativa da componente não paramétrica $g(s)$,

¹¹ Ver Capítulo 4

interpretada como superfície de risco estimada¹². Tal superfície é gerada, tomando-se por referência uma grade regular com dimensões de 26,4 m x 26,4 m, onde foram criadas 6 classes, materializadas por isolinhas, com equidistância de 0,23 unidades de chance e então vetorizadas. O resultado final é o mapa vetorial mostrado na Figura 64-A, identificando áreas de proteção com 3 classes (valores inferiores à unidade) e de não proteção com 3 classes (valores superiores à unidade).

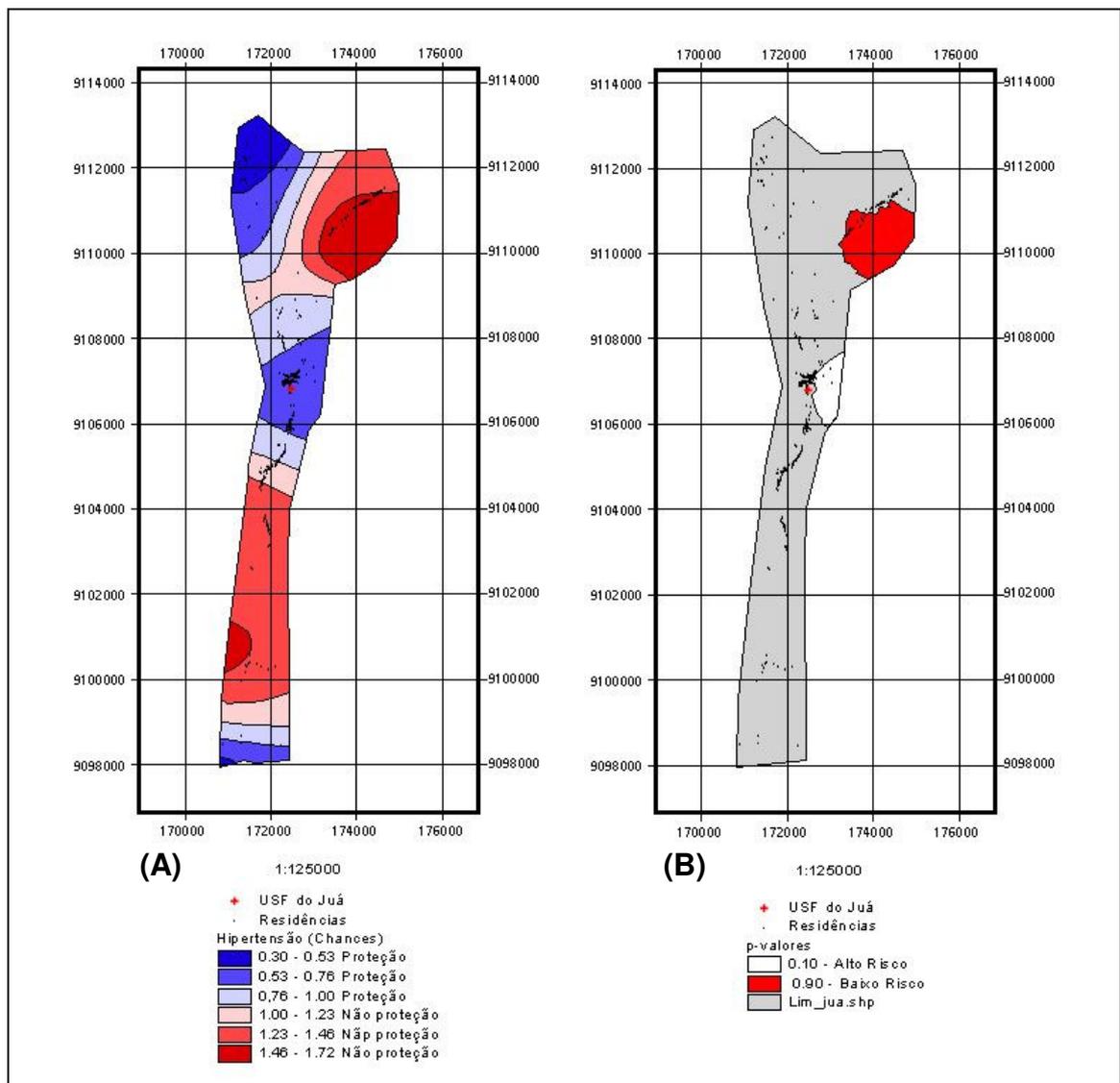


Figura 64 – Mapas da componente não paramétrica $g(s)$ e da superfície de p -valores

Além do mapa do risco, avaliou-se se a superfície estimada varia significativamente na região. Isso foi executado empregando-se o modelo

¹² Ver Capítulo 4

de simulação Monte Carlo¹³, testando-se a hipótese de que o risco não é constante na região, ou seja, $H_0: g(s) = 0$. O resultado do teste forneceu um p_{valor} de 0.43, identificando que há evidências de risco constante no Juá.

Tal resultado mostra que mesmo existindo aglomerados residenciais na região, a distribuição espacial dos portadores de hipertensão arterial tem a mesma média em todos eles, caracterizando, como esperado, a homogeneidade na população.

Convém ainda destacar, por observação *“in loco”*, que os ditos aglomerados eram função do acesso por estradas, da proximidade da terra onde as pessoas trabalham e de laços familiares.

Quanto à pertinência do modelo espacial, verifica-se que não é adequado no nível de USF, porque a população atendida por uma USF, normalmente apresenta características sócio-econômicas, culturais, sanitárias e demográficas semelhantes, além de espalhar-se por uma área geográfica pequena. Julga-se, entretanto, que o mesmo torna-se viável para ser implementado a partir do nível municipal, comparando-se as USF subordinadas, dado que nessa situação é possível ocorrerem diferenças que orientariam a intervenção.

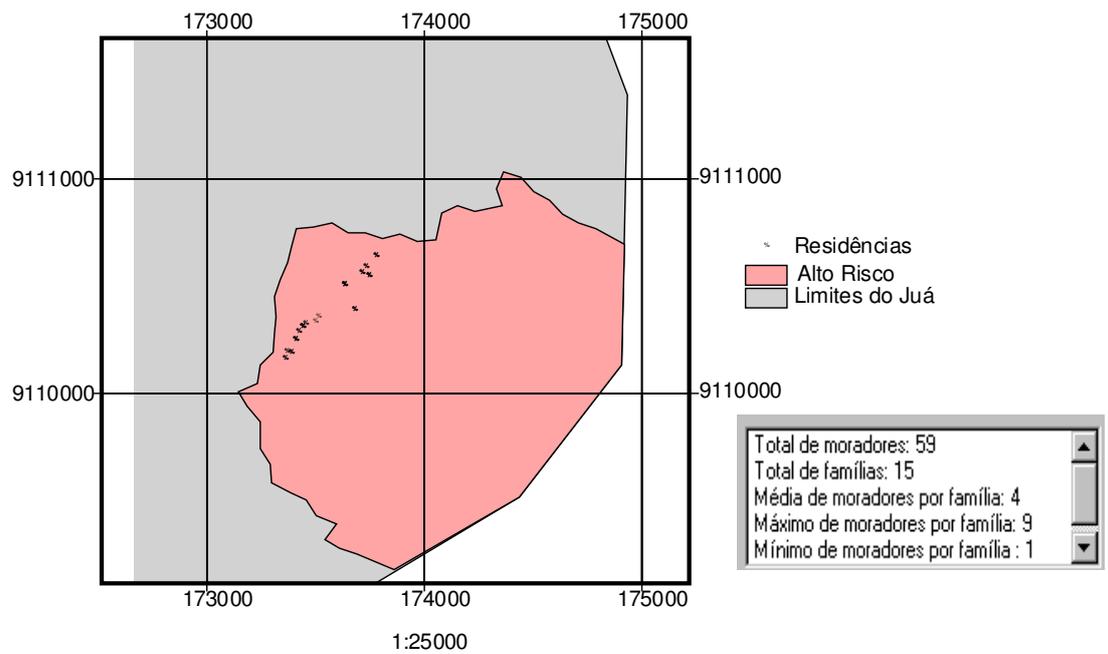
Mesmo com o resultado não significativo da simulação, gerou-se um mapa de $p_{valores}$ (valores de significância), com contornos de 0,10 e 0,90 para indicar superfícies de baixo e alto risco, respectivamente, conforme mostrado na Figura 64-B. Isso teve por objetivo mostrar como tal mapa pode ser explorado no ambiente de SIG, conforme exemplificado a seguir:

1. Onde se localizam, quantos moradores e quantas famílias residem na região de alto risco?

Selezione: “**Morador.Cod_pront_mor**” *contido em* “ **p-valores = 0.80 – 0.90**”

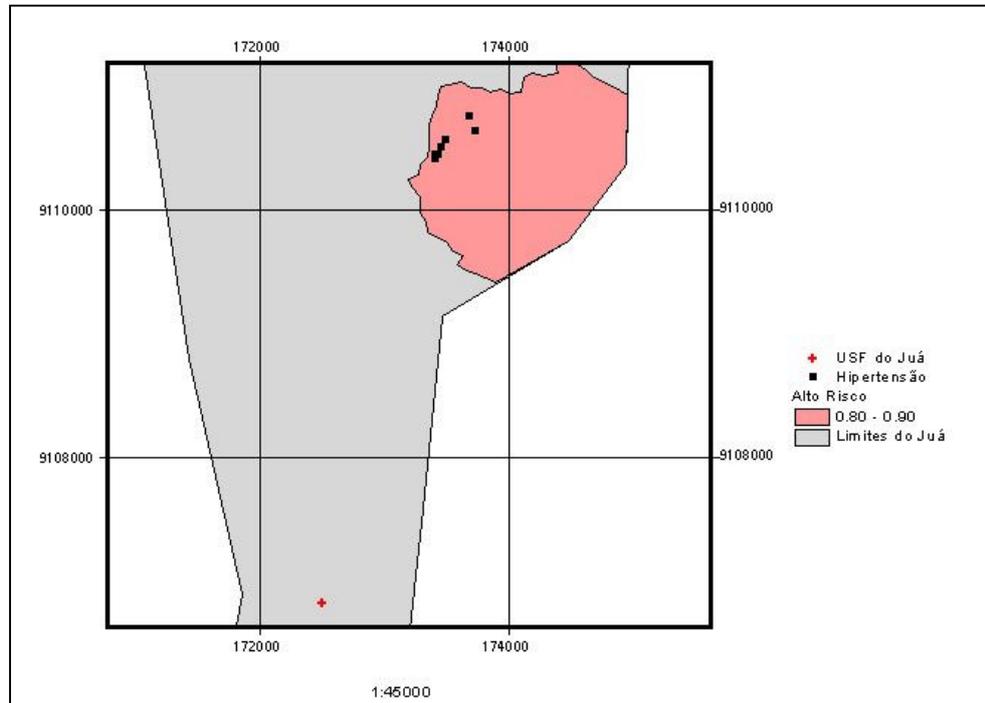
Estatística: “**Morador.Cod_pront_mor**”

¹³ Ver Capítulo 4



2. Onde se localizam as residências de hipertensos dentro da região de alto risco?

Selecione: [“Moradores.Hipertensão =1”] interseção [“p-valores = 0.80 – 0.90”]



3. Quem são os hipertensos residentes na região de alto risco?

Selecione: ["**Morador.Cod_pront_mor**" interseção "**p-valores = 0.89 – 0.90**"] e ["**Hipertensão = 1**"]

<i>Cod_pront_mor</i>	<i>Sexo</i>	<i>Hiper</i>	<i>Idade</i>
132.1	M	1	59
133.1	F	1	69
133.4	F	1	100
136.1	F	1	50
137.1	M	1	84
139.1	F	1	52
147.2	F	1	47
168.1	M	1	50

4. Quantas pessoas com idade igual ou superior a 40 anos residem na região de alto risco?

Selecione: *Estatística* ["**Morador.Cod_pront_mor**" interseção "**p-valores = 0.80 – 0.90**"] e ["**Morador.Idade ≥ 40**"]

Total de moradores: 22
Total de famílias: 13
Média de moradores por família: 2
Máximo de moradores por família: 2
Mínimo de moradores por família: 1

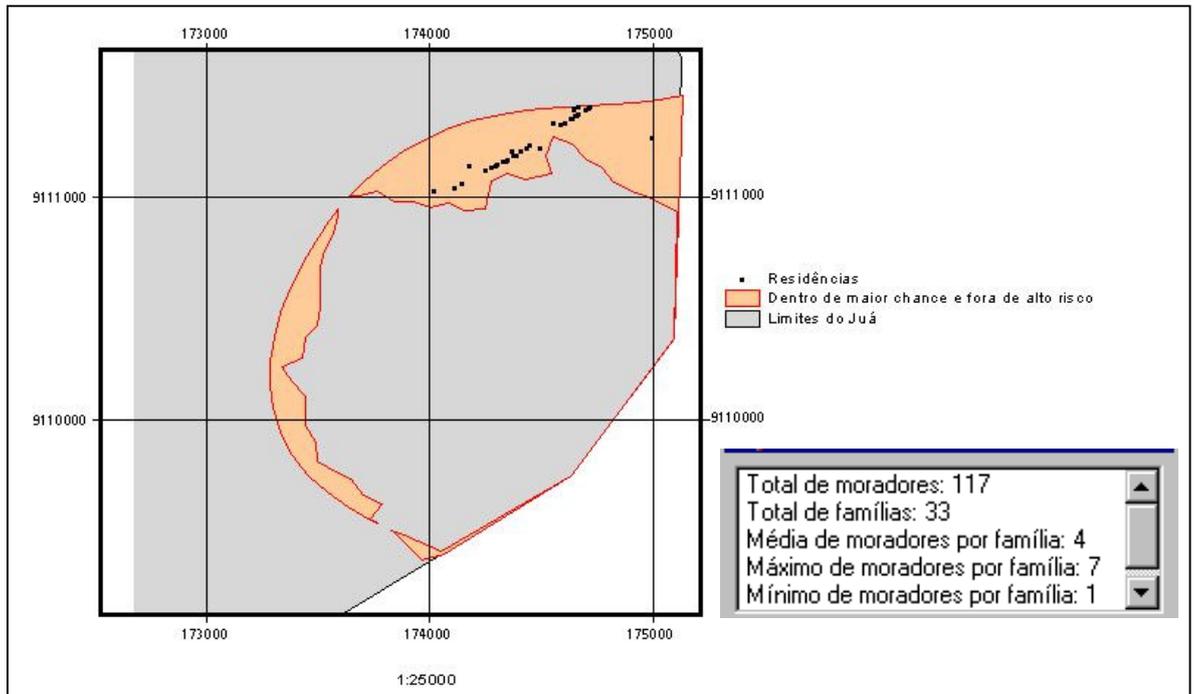
5. Quantas mulheres com idade igual ou superior a 40 anos residem na região de alto risco?

Selecione: *Estatística* ["**Morador.Cod_pront_mor**" interseção "**p-valores = 0.80 – 0.90**"] e ["**Morador.Idade ≥ 40**"] e ["**Morador.sexo = 1**"]

Total de mulheres: 13
Total de famílias: 11
Media de mulheres por família: 1
Máximo de mulheres por família: 2
Mínimo de mulheres por família: 1

6. Quantas pessoas e onde se localizam as que residem na região de maior chance e fora da região de alto risco?

Selecione: Estatística ["Morador.Cod_pront_mor"] interseção [{"Hipertensão(chances) = 1.46 -1.72"}] extração [{"p-valores = 0.80 - 0.90"}]



Convém ainda ressaltar que qualquer que seja o método estatístico empregado para analisar dados de saúde, o SIG apresenta-se como instrumento facilitador. No presente caso, por exemplo, as superfícies de risco foram geradas em um programa específico de estatística espacial e posteriormente exportadas para um programa de SIG onde foram vetorizadas. A partir daí foi possível o emprego das funções de análise espacial vetorial e de banco de dados já disponibilizadas sistema¹⁴.

¹⁴ Ver Capítulo 3

8. CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES

O desenvolvimento de qualquer sistema de informação exige uma fase de testes que, na maioria das vezes, está interligada com a implementação, de modo a identificar problemas não detectados nas fases anteriores. Este trabalho não fugiu a regra, com a aplicação do sistema desenvolvido a uma área piloto (Juá – Caruaru – PE).

Pode-se afirmar que o projeto lógico mostrou-se eficiente, do ponto de vista de organização dos dados, por apresentar um desenho compatível com a realidade do PSF, por manter inalterada a estrutura do SIAB e por permitir o uso de diversas técnicas analíticas.

Dunn e colaboradores (2001), empregando um mesmo banco de dados de saúde georreferenciado - tal qual o utilizado na implementação do sistema desenvolvido - compararam diferentes abordagens de análise (da epidemiologia tradicional, das disponíveis em programas de SIG e das de padrões de pontos), utilizadas de acordo com o interesse em questão, encontrando resultados complementares. Nessa situação, a estrutura de dados híbrida apresenta-se como um facilitador, por permitir o compartilhamento dos dados com diferentes programas de computador.

A escolha das coordenadas como chave para identificação espacial pontual das residências, do código de prontuário da família como chave para identificação unívoca das famílias/residências e do código de prontuário do morador como chave para identificação unívoca dos indivíduos foi fundamental para o funcionamento do sistema, que pode ser implementado tanto em zona rural quanto em urbana.

No que diz respeito ao emprego de SIG em nível urbano para o PSF, existem algumas experiências em implantação no País, como o “Projeto Mapeando a Saúde” (Pereira & Paranhos, 2002), da Secretaria Municipal de Saúde de Montes Claros - MG. O objetivo principal desse projeto é desenvolver tecnologias para identificar e mapear áreas e populações de risco e, também, de integrar as bases de dados do Cadastro Técnico Municipal com as do Programa de Saúde da Família. O chaveamento empregado por esse projeto para localização das famílias assistidas pelo PSF é similar ao utilizado pelo sistema ora desenvolvido, diferindo quanto à identificação das famílias e dos indivíduos. No projeto de Montes Claros

essa identificação é baseada no endereço e no nome dos componentes da família, respectivamente, permitindo também o registro do número do Cartão Nacional de Saúde. Deve ser observado que tal chaveamento corresponde a uma situação particular, sendo viável apenas quando o arruamento é bem definido, o que não acontece, por exemplo, no interior de favelas. Por essa razão, a proposta implementada nesta tese – com o uso do código do prontuário da família e do morador - possui maior potencial de uso em regiões rurais e em cidades com áreas não urbanizadas. Por ocasião da adoção do Cartão SUS, seu código pode substituir o do prontuário do morador.

Sobre a carga dos dados, é importante lembrar que ela envolve dois tipos de levantamentos: não espaciais e espaciais. Os não espaciais são caracterizados pela digitação das informações, contidas nas fichas das famílias atendidas pela USF, num programa de banco de dados. Os espaciais são caracterizados pela obtenção das coordenadas das residências, com o emprego de rastreadores GPS.

No que diz respeito à digitação das informações no banco de dados, elas compreendem a coleta e o armazenamento dos dados individuais e da família/domicílio, conforme identificados na estrutura organizacional da ESF, já mostrada na Figura 3. Sugere-se que tais atividades continuem a ser executadas pela equipe de saúde, criando-se formulários específicos no programa, que podem ser o espelho das fichas que já se encontram em uso. Desse modo, não será verificada alteração na rotina de trabalho da equipe e sim uma simples substituição das tarefas manuscritas pela digitação, desde que atendidas as seguintes necessidades:

- Disponibilização na USF de computador tipo PC com o programa de banco de dados e impressora. Deve-se considerar que o computador deve conter gravador de CD para permitir cópia de segurança do banco, sempre que o mesmo for atualizado, dado que se sugere o envio das informações com periodicidade mensal, conforme já acontece com o SIAB. A razão do emprego desse tipo de mídia é devido à sua durabilidade e significativa capacidade de memória;
- Treinamento de pessoal no uso do sistema computacional previsto e na operação do programa de banco de dados;

- Conexão com a *Internet*, de forma a permitir comunicação com os diversos níveis hierárquicos. Ressalta-se que esse problema tem por restrição a disponibilização de linha telefônica, identificando-se que no caso do Juá, por ocasião dos levantamentos, existia apenas um posto telefônico para atender toda a comunidade.

Como vantagens dessa nova situação, listam-se abaixo as mais pertinentes:

- Facilidade de atualização e manutenção dos dados, pelo uso de programa de banco de dados e também pela proximidade física da equipe com a população adscrita, que é a fonte da informação;
- Motivação da equipe de saúde pelo uso de recursos computacionais, sobretudo por parte das agentes, que passarão a se sentir inseridas no contexto atual da informação;
- Possibilidade de expansão do sistema proposto para atividades administrativas da Unidade, como controle de estoque de remédios, de distribuição de leite e até mesmo de implantação de prontuário médico informatizado.

As necessidades e vantagens acima descritas também foram observadas pelo “Projeto Mapeando a Saúde” (Pereira & Paranhos, 2002).

Quanto à obtenção das coordenadas das residências, a situação ideal é a pré-existência de mapeamento da região em escala compatível com a localização pontual das residências, dentro do sistema geodésico adotado pelo País, que é o SAD-69, na projeção UTM. Infelizmente, tal produto é escasso, normalmente encontrado somente para as grandes cidades brasileiras e muitas vezes desatualizado. Especificamente para o caso do Juá, não existia mapeamento nem em escala de 1:100.000.

A solução de contorno para essa problemática foi o emprego de rastreadores GPS. Como primeira idéia, ventilou-se a hipótese de usar os equipamentos ditos de navegação. Estes têm a vantagem de apresentar baixo custo de aquisição (em torno de U\$ 300,00) e de disponibilizar as coordenadas em tempo real, com precisão da ordem de 15 metros (Santos, 2001).

Tal opção foi descartada logo na fase de reconhecimento do Juá. Observou-se que, mesmo tratando-se de área rural, as famílias distribuíam-se em aglomerados populacionais, onde as residências apresentavam

significativa proximidade física e por vezes eram até germinadas. Esse fato é ilustrado na Figura 65, que mostra parte de um desses aglomerados e uma residência como origem de um círculo de raio 15 metros, que corresponde à precisão do rastreador. Verifica-se que dentro dessa região de incerteza estão contidas outras residências, que podem ser confundidas com a primeira, inviabilizando a premissa da identificação unívoca da entidade.

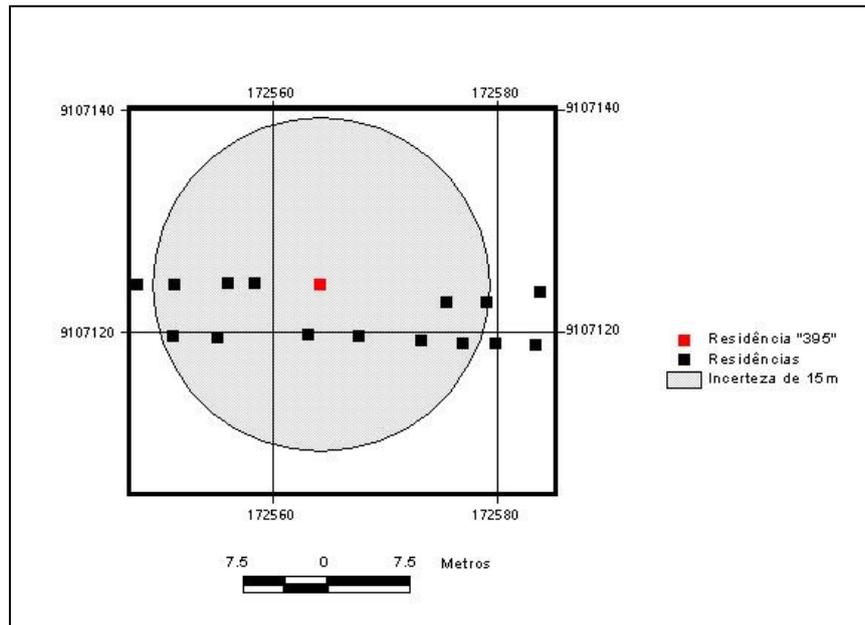


Figura 65 – Aglomerado populacional de parte da região do Juá

Com essas limitações, resolveu-se empregar rastreadores topográficos, por apresentarem precisão da ordem de 1 metro, desde que se empregando 2 estações. Esse tipo de equipamento tem custo de aquisição em torno de U\$ 8.500,00 o par, devendo ser operado por pessoal especializado, por exigir técnicas específicas de rastreamento e pós-processamento dos dados para obtenção das coordenadas (Santos, 2001). Sua produtividade é função de uma série de fatores, dentre os quais citam-se como bastante limitantes as obstruções, atenuações ou reflexões do sinal, ocasionadas por obstáculos nas proximidades da antena, conforme mostrado na Figura 66, que deterioram as informações transmitidas pelo satélite, comprometendo a precisão das determinações.

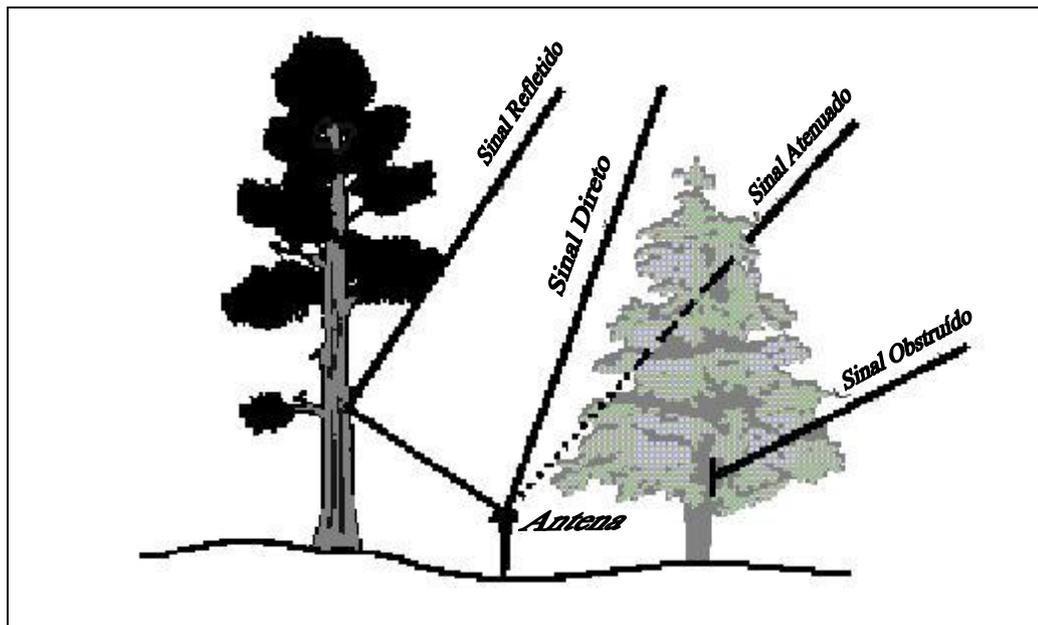


Figura 66 – Obstáculos para os sinais de satélites

No caso específico do Juá, por conta da existência de muitas árvores, os problemas representados na Figura acima foram comuns, onerando o tempo gasto nas operações de rastreamento. Apesar desses problemas essa tecnologia foi empregada por ser a mais viável no ano de 2000, época em que aconteceram os levantamentos.

Em resumo, pode-se dizer que para a localização das residências, o GPS só é plausível em situações particulares, ou seja, em terrenos livres de obstáculos e em locais onde residências distribuam-se com afastamento superior à precisão dos rastreadores de navegação, de modo que esses possam ser utilizados.

Ainda no ano 2000, entrou em operação o satélite imageador IKONOS, que disponibiliza imagens coloridas de alta resolução, devidamente georreferenciadas. Tais imagens são compatíveis com mapeamento planimétrico na escala de 1:2.500, apresentando-se como solução ótima para identificação pontual de residências e outros acidentes naturais e artificiais do terreno (Engesat, 2002). É importante observar que dada a imagem georreferenciada, as coordenadas das moradias são obtidas pela simples inserção de um ponto sobre as mesmas, conforme mostrado na Figura 67.

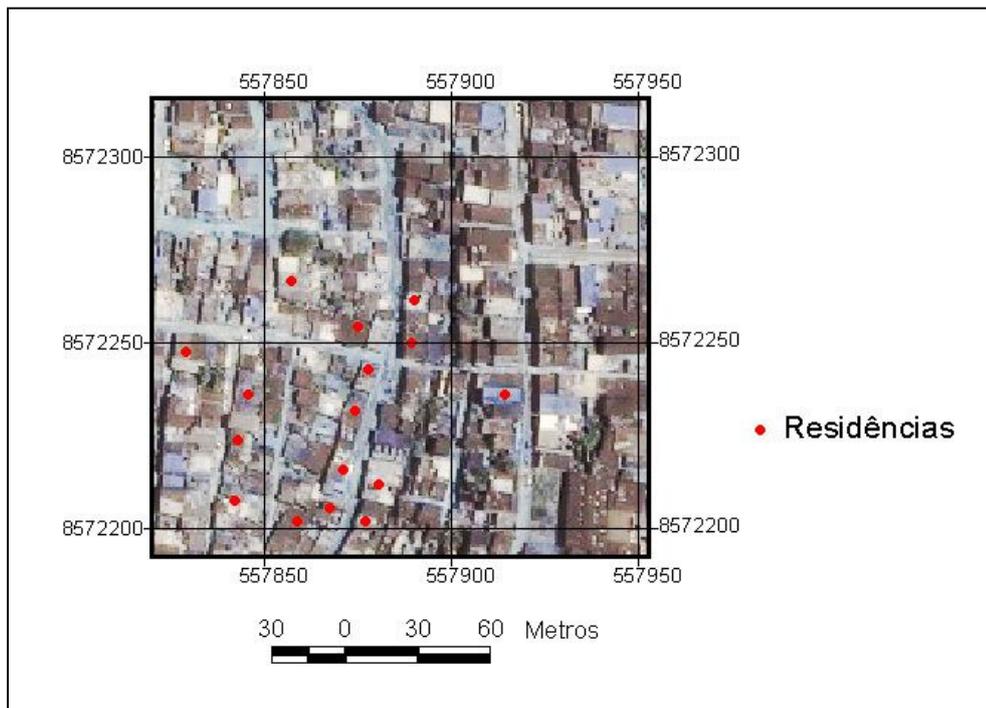


Figura 67 – Imagem IKONOS, Salvador – BA, 2002

Deve-se considerar, todavia, que mesmo com alta resolução, a identificação correta de alvos sobre uma imagem requer conhecimentos de foto-interpretação, particularmente em locais com elevada densidade residencial, sugerindo-se que tal atividade seja realizada por pessoal especializado. Essa afirmação foi comprovada em pesquisa sobre Leptospirose na região de Pau da Lima – Salvador - BA, em desenvolvimento no Centro de Pesquisas Gonçalo Muniz. Naquela ocasião, esse pesquisador observou que os encarregados da identificação de residências na imagem, tiveram relativa dificuldade em executar a tarefa.

Sobre o preço de aquisição das imagens, sabe-se que os valores variam de R\$ 50,00 a R\$ 130,00 o km², segundo informações de janeiro de 2003, constantes no *site* da empresa ENGESAT (www.engesat.com.br), representante de vendas desse produto. A primeira vista, seu custo é elevado, porém significativamente inferior ao de um mapeamento executado por métodos tradicionais, que pode chegar a R\$ 142.000,00 o Km², segundo a revista A MIRA, edição de janeiro/fevereiro de 2003. Recomenda-se o uso das ditas imagens para todo o município, ou mesmo para todo o estado, desde que não existam levantamentos cartográficos

disponíveis. Isso é justificado pelo preço das imagens que tende a diminuir a medida que se aumenta a área contratada.

No que diz respeito à execução da análise exploratória básica, julga-se ser viável na USF, uma vez que só as informações descritivas são trabalhadas. Sugere-se, para fins de uniformização da apresentação dos resultados, que seja definido em nível municipal ou em instâncias superiores, os *lay-outs* e tipos de gráficos, relatórios e tabelas de interesse. Isso se justifica porque na maioria dos programas de banco de dados existem opções que permitem adaptações ao usuário, como a criação de menus e ícones específicos, facilitando seu manuseio por parte da equipe.

Sobre as operações de análise espacial, no que se refere à informações de localização, verifica-se que se procederam com relativa simplicidade no ambiente de SIG. Recomenda-se que esse tipo de análise seja implementado no nível central municipal (secretaria de saúde), na medida que também permite o tratamento de informações descritivas, a integração horizontal e vertical com outros sistemas de informação em saúde e por ser o local onde acontece o acompanhamento e avaliação dos serviços prestados pelas USF.

Dado o SIG implantado em nível municipal, faz-se imprescindível discutir o fluxo das informações, que está esquematizado na Figura 68, considerando aspectos éticos no acesso à informação dos pacientes.

Observa-se que cada USF possui um banco de dados próprio implementado em seu computador. Esse banco pode ser lido, para que se processem as operações de análise exploratória básica, e atualizado, sempre que um novo evento ocorrer.

O município tem acesso irrestrito, para leitura, aos bancos de todas as Unidades subordinadas, criando em sua sede o banco de dados descritivo municipal, que corresponde à união de todos os anteriores. Esse banco descritivo é então ligado ao banco de dados espacial, materializado por imagens de alta resolução georreferenciadas e residências identificadas pontualmente. Essa ligação caracteriza o modelo georrelacional para o ambiente de SIG, permitindo a execução das operações de análise espacial de localização, que são enviadas na forma de mapas às USF.

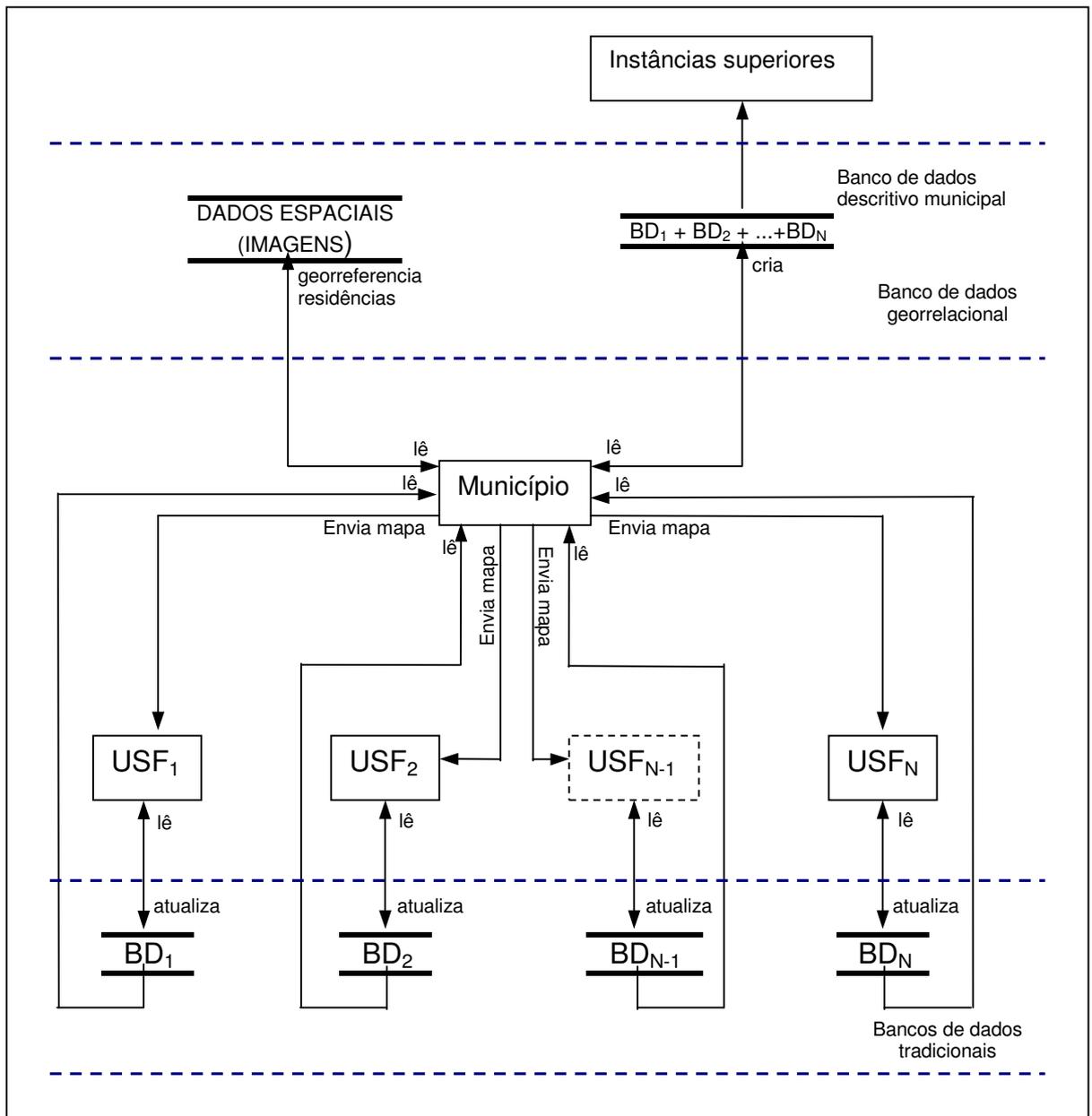


Figura 68 – Fluxo das informações

Deve-se observar que o banco municipal contempla informações individuais de toda população atendida pelo PSF e por isso seu acesso deve ser restrito, por motivos éticos, ficando a cargo do município a segurança desses dados. Verifica-se, também, que a integração vertical é facilitada, a partir da consolidação dos dados desse banco e da sua leitura pelas instâncias superiores.

No estudo de caso, analisou-se a hipertensão arterial, utilizando técnicas de processo pontual marcado. Salienta-se que essas são adequadas a, praticamente, todas as doenças que o Programa trata, como a hanseníase,

a diabetes e a tuberculose, por permitir comparar eventos espacialmente localizados, do tipo binário, materializados pela ocorrência ou não ocorrência de determinada doença.

A modelagem deste tipo de processo utilizando regressão logística, acrescida da componente espacial é interessante por estimar a variação espacial do risco, controlando-a por covariáveis influentes no evento, podendo também ser empregada para as demais doenças. No que diz respeito à componente espacial, a razão do estimador de intensidade Kernel de casos e controles permite a criação dos mapas de risco, fundamentais para melhor intervir sobre a doença.

Quanto à escolha das covariáveis, identifica-se que foram selecionadas em função das que se encontravam disponíveis no banco de dados e eram pertinentes de acordo com a literatura especializada. Esse fato apresentou-se como limitante por existirem outras, sabidamente importantes, não contempladas pelo PSF. Face a essa questão, recomenda-se que profissionais de saúde envolvidos com o Programa definam, *a priori*, as variáveis concernentes aos eventos tratados, de modo a também comporem o banco de dados das USF. A proposta de inserção de novas covariáveis pode obedecer a interesses municipais, e a forma como o sistema está pensado é flexível o suficiente para que esta inclusão seja feita sem grandes custos operacionais.

Sobre a confecção do mapa de risco e do mapa da superfície de valores de significância, observa-se sua inviabilidade de execução pelo pessoal de serviço, em nível municipal. Essa afirmação se justifica, em primeiro lugar, pela complexidade dos modelos estatísticos empregados e, em segundo lugar, por ainda não se dispor de um programa de computador único, com interface amigável, que atenda todo o desenrolar da modelagem. O estudo de caso da hipertensão arterial exemplifica esses problemas na medida que envolveu os seguintes passos:

1. Execução de análise prévia não espacial das variáveis influentes na doença, com o objetivo de verificar seus níveis de confiança e possíveis interações, num modelo de regressão logística tradicional. Esta fase, embora tenha sido feita no SPSS, pode ser realizada no Epi-Info sem qualquer prejuízo;

2. Geração das superfícies de risco e valores de significância , empregando-se o programa R, materializadas por grades regulares;
3. Importação da grade, vetorização e geração dos mapas, nesse formato, empregando-se o programa ArcView, que corresponde ao ambiente de SIG.

Face ao exposto, sugere-se que a confecção dos referidos mapas, seja efetuada em instâncias superiores, através de convênios com universidades, centros ou institutos de pesquisa e outros órgãos correlatos, onde se encontram profissionais qualificados para tal.

Um aspecto importante ainda é o custo dos programas necessários. O R é de domínio público com código aberto já o Arc-View tem custo elevado.

Uma opção que se apresenta em relação ao programa de SIG diz respeito a biblioteca TerraLib. Trata-se de uma biblioteca de classes para a construção de Sistemas de Informações Geográficas, sem custo para o usuário, disponível na Internet como código aberto, permitindo a construção de um ambiente colaborativo e de uso para o desenvolvimento de variados aplicativos.(INPE, 2002). Dadas essas características, verifica-se que pode ser empregada para a criação de um ambiente de SIG específico para o PACS/PSF, inclusive com possibilidade de acessar bancos secundários de saúde, facilitando o uso pelas secretarias de saúde.

9. CONCLUSÃO

Este estudo tem como principal produto o desenvolvimento de uma tecnologia a ser aplicada na área de atenção básica de saúde, representada por um SIG destinado ao PSF.

O conhecimento dos protocolos e das rotinas de trabalho de todos os comprometidos com o Programa, permite a identificação de três níveis funcionais (o operacional, o gerencial e o político) e a necessidade de suporte do mesmo pela informação. As restrições do atual sistema de informação do Programa - o SIAB - podem ser minimizadas ou até mesmo eliminadas, com a implantação do sistema desenvolvido em ambiente de SIG.

O SIG desenvolvido:

- Respeita o grau de descentralização, cobertura e especificidades do PSF;
- Representa um instrumento auxiliar no processo de implantação, gerenciamento e tomada de decisão do PSF;
- Permite o estudo criterioso das informações referentes aos protocolos e rotinas de trabalho de todos os comprometidos com o PSF;
- Permite a localização geográfica de eventos de saúde;
- Trata, de maneira integrada, tanto dados espaciais como não espaciais;
- Tem por unidade espacial básica o indivíduo, registrado pontualmente em seu local de residência;
- Possibilita, por ter o indivíduo como unidade espacial básica, a agregação de informações, sem perdas, quando se opera do nível mais baixo para o mais alto, ou seja, agregando-se indivíduos chega-se à família, agregando-se famílias chega-se à microárea e assim sucessivamente;
- Possibilita a visualização, em mapas, de padrões de pontos que mostram a distribuição espacial de doenças e permite o emprego de técnicas estatísticas de análise espacial, de modo a identificar áreas de risco, analisar a acessibilidade a serviços de saúde e obter estatísticas da população adscrita;

- Apresenta projeto lógico viável para implementação, com possibilidade de expansão gradual, sem comprometer sua estrutura. Isso indica que o Sistema tem condições de suportar funções para atender necessidades administrativas das USF e para implantação de prontuários médicos informatizados;
- Apresenta projeto físico implementado em ambiente de microcomputador com a carga das informações não espaciais processada diretamente em uma USF, com auxílio de agentes comunitários. As ACS, desde que devidamente treinadas, têm condições de manter o banco de dados atualizado, a partir de equipamentos computacionais com custo significativamente baixo. A confecção de mapas, dada sua complexidade, é recomendada a partir do nível distrital e/ou central municipal;
- Apresenta fluxo de informações que se origina nas USF, onde ocorre a carga e manutenção do banco de dados, sendo, posteriormente enviado ao nível distrital e/ou central municipal, que agrega as informações recebidas, constrói os mapas pertinentes, devolvendo-os às USF e divulgando-os aos níveis hierárquicos superiores;
- Pode, dependendo da disponibilidade de equipamentos de informática e conexões *on line*, ser alimentado, diariamente, na USF e acessado por outros níveis funcionais do sistema de saúde, permitindo o conhecimento de eventos e agravos de forma ágil por todos os envolvidos.

Em suma, o SIG desenvolvido representa um avanço importante em relação ao atual sistema de informação utilizado pelo PSF, constituindo-se num instrumento facilitador da consolidação da atenção básica no país ao permitir:

- O conhecimento ágil dos eventos de saúde de cada indivíduo e de toda a população adscrita à USF por todos os níveis funcionais, facilitando a reflexão, planejamento e adoção de ações;
- O acompanhamento e avaliação das ações desenvolvidas, possibilitando a sua reorientação ou reforço.

Para o SIG ser desenvolvido deve-se observar que:

- O município seja o local de implantação do sistema como um todo;
- As USF sejam responsáveis pela sua carga, diretamente em ambiente computacional de banco de dados;
- O fluxo dos dados das USF para o município seja periódico, preferencialmente por *Internet*;
- A obtenção das coordenadas das residências se processe a partir de mapas pré-existentes ou na sua ausência, a partir de imagens de satélite com resolução espacial da ordem de 1 metro;
- O mapeamento de áreas de risco dos agravos cobertos pelo PSF seja realizado a partir do nível distrital/central municipal;
- Que se desenvolva capacidade analítica nos níveis centrais municipais e estaduais, permitindo que todo este conjunto de dados coletados sejam usados para gerar informação que oriente a intervenção do PACS/PSF na saúde da comunidade.

10. REFERÊNCIAS BIBLIOGRÁFICAS

ARONOFF, S., 1989. *Geographic Information Systems: A Management Perspective*. Ottawa: WLD Publications.

BAILEY, T. C. & GATRELL, A. C., 1995. *Interactive Spatial Data Analysis*. London: Longman.

BARCELLOS, C. & BASTOS, F. I., 1996. Geoprocessamento, ambiente e saúde: Uma união possível? . *Cadernos de Saúde Pública*, 12:389-397.

BORGES, K. A. V., 1997. *Modelagem de dados geográficos: Uma experiência do modelo OMT para aplicações geográficas*. Dissertação de Mestrado, Belo Horizonte: Escola de Governo – Fundação João Pinheiro.

BURROUGH, P. A.; FRANK, A. U., 1996. *Geographical Objects with Indeterminate Boundaries*. London: Taylor & Francis.

BURROUGH, P. A.; McDONNELL, R. A., 1998. *Principles of Geographical Information Systems. Spatial Information Systems and Geostatistics*. New York: Oxford University Press.

BUZAI, G. & DURÁN, D., 1977. *Enseñar e Investigar com Sistemas de Información Geográfica*. Buenos Aires: Editora Troquel.

CÂMARA, G.; CASANOVA, M. A.; HEMERLY, A. S.; MAGALHÃES, G. C. & MEDEIROS, M. B., 1996. *Anatomia de Sistema de Informações Geográficas*. Campinas: Catgraf LTDA.

CÂMARA, G; CARVALHO, M. S., 2001. Análise espacial de eventos. <http://www.dpi.inpe.br/gilberto/livro/analise/cap2-eventos.pdf>.

COITINHO, D. C.; LEÃO, M. M.; RECINE, E.; SICHIERI, R., 1994. Condições nutricionais da população brasileira: adultos e idosos. *Pesquisa Nacional de Saúde e Nutrição, MS-Brasília*.

CHEN, P., 1990. *Modelagem de Dados: A abordagem entidade-relacionamento para o projeto lógico*. São Paulo: McGraw-Hill.

DANGEMOND, J., 1990. A review of digital data commonly available and some of the practical problems of entering them into a GIS. In: *Introductory readings in Geographic Information Systems*. (PEUQUET, D, J. & MARBLE, D. F, ed.), pp. 209 – 221, London: Taylor & Francis.

DATE, C. J., 1995. *An Introduction to Database Systems*. New York: Addison Wesley.

DIGGLE, P. H., 2003. *Statistical Analysis of Spatial Point Pattern*. London: Academic Press.

DUNN, C. E., KINGHAM, S. P., ROWLINGSON, B., BHOPAL, R. S., COCKINGS, S., FOY, C. J. W., ACQUILLA, S. D., HALPIN, J., DIGGLE, P., WALKER, D., 2001. Analyzing spatially referenced public health data: a comparison of three methodological approaches. *Health & Place* 7:1-12.

EFIMOV, N., 1972. *Elementos de Geometria Analítica*. Rio de Janeiro: Livraria Cultura Brasileira Editora.

ELMASRI, R., & NAVATHE, S. B., 2000. *Fundamentals of database systems*. Massachusetts: Addison Wesley.

ENGESAT, 2002. <http://www.engesat.com.br>

FAERSTEIN, E., CHOR, D., LOPES, C. S., 2001. Confiabilidade da história referida de diagnóstico e tratamento da hipertensão arterial. Diferenciais segundo gênero, idade e escolaridade. O estudo pró-saúde. *Arq Bras Cardiol*, 76(4): 297-300

FIOCRUZ/CENTRO DE PESQUISAS AGGEU MAGALHÃES, 1999. *Proposta de Implantação de um Sistema de Informações Geográficas para Apoio ao Programa de Agentes Comunitários de Saúde/Programa de Saúde da Família*. Recife, 31/05 – 02/06/99: Relatório Final. <http://www.cpqam.fiocruz.br/nicc/ofrecife.html>

FRANK, A. U., 1988. Requirements for a database management system for a GIS. *Photogrammetric Engineering and Remote Sensing*. 54: 1557-1564.

FURTADO, A. L. & SANTOS, C. S., 1989. *Organização de bancos de dados*. Rio de Janeiro: Editora Campus Ltda.

HASTIE, T. J. & TIBISHIRANI, R. J., 1997. *Generalized Additive Models*. London: Chapman & Hall.

HEUSSER, C. A., 2000. *Projeto de banco de dados*. Porto Alegre: Sagra Luzzato.

HOLROYD, F. & BELL, S. B. M., 1992. RasterGIS: Models of raster encoding. *Computers and Geosciences*, 18: 419-426.

HOSMER, D. W. & LEMESHOW, S., 1989. *Applied Logistic Regression*. New York: John Wiley & Sons. http://ids-saude.uol.com.br/psf/medicina/tema5/texto65_definição.asp

HUXHOLD, A. & LEVINSOHN, A. G., 1995. *Managing Geographic Information Systems Projects*. New York: Oxford University Press.

IBGE, 2003. <http://ibge.gov.br/home/geografia/decar/glossario>

INCLIN Multicentre Collaborative Group, 1994. Socio-economic status and risk factors for cardiovascular disease: a multicentre collaborative

study in the international clinical epidemiology network. *J Clin Epidemiol*, 47:1401-9.

INPE, 2000. <http://www.terralib.dpi.inpe.br>

JONES, M. P., 1988. *Projeto Estruturado de Sistemas*. São Paulo: McGraw-Hill.

KELSALL, J, E. & DIGGLE, P. J., 1998. Non-parametric estimation of spatial variation in relative risk. *Statistics in Medicine*, 14:2335-2342.

KENNEDY, M., 1996. *The Global Positioning System and GIS*. Ann Arbor: AnnArbor Press, Inc.

KNOX, E. G., 1964. Epidemiology of childhood leukaemia in Northumberland and Durham. *British Journal of Preventive and Social Medicine*, 18:17-24.

KROENKE, D. M., 1988. *Bancos de Dados: Fundamentos, Projetos e Implementação*. Rio de Janeiro: LTC.

LAPA, T., 2003. *Análise da demanda de casos de hanseníase aos Serviços de Saúde através do uso de Técnicas de Análise Espacial*. Tese de Doutorado, Recife: CpqAM/ENSP/FIOCRUZ.

LISBOA, A. F., 1997. *Modelos conceituais de dados para sistemas de informações geográficas*. Porto Alegre: CPGCC da UFRGS.

LAURINI, R. & THOMPSON, D., 1992. *Fundamentals of Spatial Information Systems*. London: Academic Press Limited.

LEICK, A., 1990. *GPS Satellite Surveying*. New York: John Wiley & Sons.

LORIE, R. A. & MEIER, A., 1984. Using a relational DBMS for geographical databases. *GeoProcessing*, 2:243-257.

LOTUFO, P. A., 2001. Diabetes. Instituto para o Desenvolvimento da Saúde. <http://ids-saude.uol.com.br/psf/medicina/tema5/texto66>

MAGUIRE, D. J.; GOODCHILD, M. F. & RHIND, D.W., 1994. *Geographical Information Systems – Principles and Applications*. New York: Logman Scientific & Technical.

MALING, D. H., 1992. *Coordinate Systems and Map Projections*. Oxford: Pergamon.

MION JR, D.; PIERIN, A. M. G., 2001. Hipertensão Arterial. Instituto para o Desenvolvimento da Saúde, 08 Outubro 2001. <http://ids-saude.uol.com.br/psf/medicina/tema5/texto65>

MONDRINI, L., MONTEIRO, C. A., 1994. Mudanças no padrão de alimentação da população urbana brasileira (1962-1988). *Ver Saúde Publ SP*, 28:433-9.

MONTGOMERY, G. E. & SHUCH, H., 1993. *GIS Data Conversion Handbook*. Colorado: Gisworld Book.

MS (MINISTÉRIO DA SAÚDE), 1988. *SIAB- Sistema de Informação de Atenção Básica - Manual*. Brasília-DF: Coordenação de Saúde da Comunidade, Secretaria de Assistência à Saúde, Ministério da Saúde.

MS (Ministério da Saúde), 2002. *Avaliação da implementação do Programa de Saúde da Família em dez grandes centro urbanos: Síntese dos principais resultados*. Brasília-DF: Ministério da Saúde.

MS (MINISTÉRIO DA SAÚDE), 2000. *Cadernos de Atenção Básica. Programa de Saúde da Família*. Brasília-DF: Departamento de Atenção Básica, Secretaria de Políticas de Saúde, Ministério da Saúde.

PAINHO, M., SENA, R. & CABRAL, P., 1999. Metodologias de desenvolvimento para aplicações de Sistemas de Informações Geográficas. In: *Anais do V Encontro sobre Sistemas de Informações Geográficas – ESIG'99*, CD ROM. Oeiras-Portugal.

PEREIRA, R., PARANHOS, L. S., 2002. Utilização do cadastro técnico multifinalitário no mapeamento de condições de saúde – A experiência do município de Montes Claros/MG. *Revista IP*, www.ip.pbh.gov/ip0401.

PEUQUET, D.J., 1994. Data in a Personal Computer Environment. . In: *Geographic Information Systems. The Microcomputer and modern Cartography*. (TAYLOR, D. R. F, ed.), pp. 67 - 95, Ottawa: Pergamon.

PINA, M. F. R. P., 1998. Potencialidades dos Sistemas de Informações Geográficas na Área de Saúde. In: *Saúde e Espaço. Estudos Metodológicos e Técnicas de Análise*. (A. L. Najjar & E. C. Marques, org), pp. 125-133, Rio de Janeiro: Editora FIOCRUZ.

RIPSA. Compatibilização de sistemas e base de dados (CDB) da Rede Integrada de Informações para a Saúde – Informe final. *Informe Epidemiológico do SUS*. Brasília-DF, pp35-41.

RUNBAUGH, J.; BLAHA, M., PRMERLANI, W.; EDDY, F. & LORENSEN, W., 1994. *Modelagem e Projetos Baseados em Objetos*. Rio de Janeiro: Editora Campus.

SÁ, L. A. C. M., 2001. *Modelagem de dados espaciais para sistemas de informações geográficas. Pesquisa na emergência médica*. Dissertação de Doutorado, São Carlos: Escola de Engenharia de São Carlos, Universidade de São Paulo.

SANTOS, A. A., 2001. *Geodésia Elementar. Princípios de Posicionamento Global – GPS*. Recife: Editora Universitária da UFPE.

SANTOS, S. M.; BARCELLOS, C; CARVALHO, M. S. & FLÔRES, R., 2001. Detecção de aglomerados espaciais de óbitos por causas violentas em Porto Alegre, Rio Grande do Sul, Brasil, 1996. *Cadernos de Saúde Pública*, 17(5):1141-1151.

SANTOS, V. R. B., 1977. *Curso de Cálculo numérico*. Rio de Janeiro: Livros Técnicos e Científicos.

SETZER, V. W., 1989. *Banco de Dados: conceitos, modelos, gerenciamento, projeto lógico, projeto físico*. São Paulo: Edgar Blucher.

SHIMAKURA, S. E.; CARVALHO, M. S; AERTS, R. G. C. & FLÔRES, R., 2001. Distribuição espacial do risco: modelagem da mortalidade infantil em Porto Alegre, Rio Grande do Sul, Brasil. *Cadernos de Saúde Pública*, 17(5):1251-1261.

SHOLTEN, H. J. & LEPPER, M. J. C., 1991. The benefits of the application of geographical information systems in public and environmental health. *World Health Stat Q.* 44(3):160-70. <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=Display&DB=PubMed>

SILVA, A., 2002. *Dominando a tecnologia de objetos*. Rio de Janeiro: Book Express.

SILVA, S. M. L., 1991. *Procedimento Fotogramétrico para construção de uma base de dados em sistemas de informação*. Dissertação de Mestrado, Rio de Janeiro: IME.

SOUZA, M. F.; FERNANDES, A. S.; ARAÚJO, C. L. & FERNANDES, M. C., 2000. Gestão da Atenção Básica: Redefinindo Contexto e Possibilidades. *Divulgação em Saúde para debate*, Rio de Janeiro, 21:7-14.

STARFIELD, B., 2002. *Atenção Primária. Equilíbrio entre necessidades de saúde, serviços e tecnologia*. Brasília-DF: UNESCO, MS.

STRAUCH, J. C. M. & SOUSA, J. M., 1998. Uma Metodologia para Implantação de Sistemas de Informações Geográficas. In: *Saúde e Espaço. Estudos Metodológicos e Técnicas de Análise*. (A. L. Najjar & E. C. Marques, org), pp. 109 - 123, Rio de Janeiro: Editora FIOCRUZ.

TASCA, R.; GRECO, C. & VILLAROSA, F. N., 1993. Sistemas de Informação em Saúde para Distritos Sanitários. In: *Distrito Sanitário, o Processo Social de Mudança das Práticas Sanitárias do Sistema Único de Saúde* (E. V. Mendes, org), pp. 267-310, São Paulo: Hucitec Ltda.

TOBLER, W. R., 1989. Frame independent spatial analysis. In: *Accuracy of spatial databases*. (GOODCHILD, M & GOPAL, S., ed), pp. 115 - 122, London: Taylor & Francis.

- TOMLIN, C. D., 1990. *Geographic Information Systems and Cartographic Modeling*. New Jersey: Prentice-Hall. John Wiley & Sons.
- TRAD,L.A.B. & BASTOS,A.C.S., 1998. O Impacto Sócio-cultural do Programa de Saúde da Família (PSF): Uma Proposta de Avaliação. *Cadernos de Saúde Pública*, 14:429-435.
- ULMAN, J. D., 1989. *Principles of Database Systems Computer*. New York: Computer Science Press.
- VIANA, A.L.D. & DAL POZ, M.R.,1998. A reforma do sistema de saúde no Brasil e o Programa de Saúde da Família. *Physis: Revista de Saúde Coletiva*, 8(2):11-48.
- WAND, M. P. & JONES, M. C., 1995. *Kernel Smoothing*. London: Chapman and Hall.
- WORBOYS, M. F., 1995. *GIS – A computing perspective*. London: Taylor & Francis.
- WORBOYS, M. F., 1994. *Object oriented approaches to geo-referenced information*. *International Journal of Geographical Information Systems*, 4:369:83.
- YOURDON, E., COAD, P., 1991. *Object oriented design*. New Jersey: Prentice Hall.
- YOURDON, E., ARGUILA, C., 1999. *Análise e projeto orientados a objetos – estudos de casos*. São Paulo: MAKRON Books.

ANEXO 1 – FICHA A (FRENTE)

FICHA A		SECRETARIA MUNICIPAL DE SAÚDE SISTEMA DE INFORMAÇÃO DE ATENÇÃO BÁSICA			UF __
ENDEREÇO		NÚMERO _ _ _	BAIRRO		CEP _ _ _ _ - _ _
MUNICÍPIO _ _ _ _ _ _	SEGMENTO _	ÁREA _ _	MICROÁREA _	FAMÍLIA _ _	DATA _ - _ - _

CADASTRO DA FAMÍLIA

PESSOAS COM 15 ANOS E MAIS NOME	DATA NASC.	IDADE	SEXO	ALFABETIZADO		OCUPAÇÃO	DOENÇA OU CONDIÇÃO REFERIDA (sigla)
				sim	não		

PESSOAS DE 0 A 14 ANOS NOME	DATA NASC.	IDADE	SEXO	FREQUENTE A ESCOLA		OCUPAÇÃO	DOENÇA OU CONDIÇÃO REFERIDA (sigla)
				sim	não		

Siglas para a indicação das doenças e/ou condições referidas

ALC - Alcoolismo	EPI - Epilepsia	HAN - Hanseníase
CHA - Chagas	GES - Gestação	MAL - Malária
DEF - Deficiência	HA - Hipertensão Arterial	
DIA - Diabetes	TB - Tuberculose	

ANEXO1 (CONTINUAÇÃO) – FICHA A (VERSO)

SITUAÇÃO DA MORADIA E SANEAMENTO

TIPO DE CASA		TRATAMENTO DA ÁGUA NO DOMICÍLIO	
Tijolo/Adobe		Filtração	
Taipa revestida		Fervura	
Taipa não revestida		Cloração	
Madeira		Sem tratamento	
Material aproveitado		ABASTECIMENTO DE ÁGUA	
Outro - Especificar:		Rede pública	
Número de cômodos / peças		Poço ou nascente	
Energia elétrica		Outros	
DESTINO DO LIXO		DESTINO DE FEZES E URINA	
Coletado		Sistema de esgoto (rede geral)	
Queimado / Enterrado		Fossa	
Céu aberto		Céu aberto	

OUTRAS INFORMAÇÕES

Alguém da família possui Plano de Saúde?		Número de pessoas cobertas por Plano de Saúde	
Nome do Plano de Saúde _____			

EM CASO DE DOENÇA PROCURA		PARTICIPA DE GRUPOS COMUNITÁRIOS	
Hospital		Cooperativa	
Unidade de Saúde		Grupo religioso	
Benzedeira		Associações	
Farmácia		Outros - Especificar:	
Outros - Especificar:		MEIOS DE TRANSPORTE QUE MAIS UTILIZA	
MEIOS DE COMUNICAÇÃO QUE MAIS UTILIZA		Ônibus	
Rádio		Caminhão	
Televisão		Carro	
Outros - Especificar:		Carroça	
		Outros - Especificar:	

ANEXO 1 (CONTINUAÇÃO) – FICHA B - GES

FICHA B - GES		SECRETARIA MUNICIPAL DE SAÚDE SISTEMA DE INFORMAÇÃO DE ATENÇÃO BÁSICA				ANO
MUNICÍPIO _____ ____	SEGMENTO ____	UNIDADE _____ _____	ÁREA ____	MICROÁREA ____	NOME DO ACS:	
ACOMPANHAMENTO DE GESTANTES						
Identificação da gestante	Data da última regra	Data provável do parto	Data da Vacina	Estado Nutricional : D - Desnutrida N - Nutrida		Data da consulta de pré-natal
				Mês de gestação		Mês de gestação
			1 2 3 R	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9
Nome:						
Endereço:	OBS:					Data da visita do ACS
			1 2 3 R	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9
Nome:						
Endereço:	OBS:					Data da visita do ACS
			1 2 3 R	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9
Nome:						
Endereço:	OBS:					Data da visita do ACS

Identificação da gestante	Fatores de risco	Resultado da gestação atual			Data da consulta de puerpério	
		NV	NM	AB	1	2
Nome: Endereço:	6 ou mais gestações Natimorto/Aborto 36 anos e mais Menos de 20 anos Sangramento Edema Diabetes Pressão Alta					
Nome: Endereço:	6 ou mais gestações Natimorto/Aborto 36 anos e mais Menos de 20 anos Sangramento Edema Diabetes Pressão Alta					
Nome: Endereço:	6 ou mais gestações Natimorto/Aborto 36 anos e mais Menos de 20 anos Sangramento Edema Diabetes Pressão Alta					

DIAS ⇨												Total	
P	Atendimento específico para AT												
R	Visita de Inspeção Sanitária												
O	Atend. individual prof. nível superior												
C	Curativos												
E	Inalações												
D	Injeções												
I	Retirada de pontos												
M	Terapia da Reidratação Oral												
E	Sutura												
N	Atend. Grupo - Educação em saúde												
T	Procedimentos Coletivos I (PC I)												
O	Reuniões												
S	Visita domiciliar												
N O T I F I C A Ç Õ E S	A	< 2 anos que tiveram diarreia											
	C	< 2 anos que tiveram diarreia e usaram TRO											
	S	< 2anos que tiveram Infecção Respiratória Aguda											
	I M C E D		Pneumonia em < 5 anos										
			Valvulopatias reumáticas em pessoas de 5 a 14 anos										
			Acidente Vascular Cerebral										
			Infarto Agudo do Miocárdio										
			DHEG (forma grave)										
	Õ E S	E	Doença Hemolítica Perinatal										
			Fratura de colo de fêmur em >50 anos										
		N	Meningite tuberculosa em < 5 anos										
		F	Hanseníase com incapacidade II e III										
			Citologia Oncótica NIC III (carcinoma in situ)										

ÓBITOS					
DATA	NOME	ENDEREÇO	SEXO	IDADE	CAUSA

ANEXO 2 – RELATÓRIO A1

Doença ou Condição Referida		QUANTIDADE DE PESSOAS
		TOTAL
15 ANOS OU MAIS	ALC	
	CHA	
	DEF	
	DIA	
	DME	
	EPI	
	GES	
	HA	
	HAN	
	MAL	
DE 0 A 14 ANOS	TBC	
	ALC	
	CHA	
	DEF	
	DIA	
	DME	
	EPI	
	GES	
	HA	
	HAN	
MAL		
TBC		

MORADIA E SANEAMENTO –
QUANTIDADE DE FAMÍLIAS

<i>Tipo de Casa</i>	TOTAL
Tijolo/Adobe	
Taipa revestida	
Taipa não revestida	
Madeira	
Material aproveitado	
Outros	
Energia Elétrica	
<i>Destino do Lixo</i>	
Coletado	
Queimado/Enterrado	
Céu Aberto	
<i>Tratamento d' água</i>	
Filtração	
Fervura	
Cloração	
Sem Tratamento	

Abastecimento de Água	
Rede Pública	
Poço ou Nascente	
Outros	
Destino de Fezes e Urina	
Sistema de Esgoto	
Fossa	
Céu Aberto	

ANEXO 2 (CONTINUAÇÃO) – RELATÓRIO A2

Relatório A2	SECRETARIA MUNICIPAL DE SAÚDE SISTEMA DE INFORMAÇÃO DE ATENÇÃO			UF _ _
MUNICÍPIO (nome)	MUNICÍPIO (código) _ _ _ _ _ _ _ _ _	SEGMENTO _ _	ÁREA _ _ _ _	ANO _ _ _ _

CONSOLIDADO ANUAL DAS FAMÍLIAS CADASTRADAS POR ÁREA

SEXO	FAIXA ETÁRIA	MICROÁREA										TOTAL
		01	02	03	04	05	06	07	08	09	10	
<i>M</i>	< 1											
	1 - 4											
	5 - 6											
	7 - 9											
	10 - 14											
	15 - 19											
	20 - 39											
	40 - 49											
	50 - 59											
	60 e mais											
	SUB-TOTAL											
<i>F</i>	< 1											
	1 - 4											
	5 - 6											
	7 - 9											
	10 - 14											
	15 - 19											
	20 - 39											
	40 - 49											
	50 - 59											
	60 e mais											
	SUB-TOTAL											
TOTAL												

	Nº	%
Famílias cadastradas		
Pessoas cadastradas		
Pessoas com cobertura de Plano de Saúde		

	Nº	%
Crianças de 7 - 14 anos		
Crianças de 7 a 14 anos na escola		
Pessoas de 15 anos e mais		
Pessoas de 15 anos e mais alfabetizadas		

ANEXO 2 (CONTINUAÇÃO) – RELATÓRIO A3

Relatório A3	SECRETARIA MUNICIPAL DE SAÚDE SISTEMA DE INFORMAÇÃO DE ATENÇÃO		UF [][]
MUNICÍPIO (nome)	MUNICÍPIO (código) [][][][][][][][][][]	SEGMENTO [][]	ANO [][][][]

CONSOLIDADO ANUAL DAS FAMÍLIAS CADASTRADAS POR SEGMENTO

SEXO	FAIXA ETÁRIA	ÁREA										TOTAL	
<i>M A S C U L I N O</i>	< 1												
	1 - 4												
	5 - 6												
	7 - 9												
	10 - 14												
	15 - 19												
	20 - 39												
	40 - 49												
	50 - 59												
	60 e mais												
	SUB-TOTAL												
<i>F E M I N I N O</i>	< 1												
	1 - 4												
	5 - 6												
	7 - 9												
	10 - 14												
	15 - 19												
	20 - 39												
	40 - 49												
	50 - 59												
	60 e mais												
	SUB-TOTAL												
TOTAL													

	Nº	%
Famílias cadastradas		
Pessoas cadastradas		
Pessoas com cobertura de Plano de Saúde		

	Nº	%
Crianças de 7 - 14 anos		
Crianças de 7 a 14 anos na escola		
Pessoas de 15 anos e mais		
Pessoas de 15 anos e mais alfabetizadas		

ANEXO 2 (CONTINUAÇÃO) – RELATÓRIO A4

Relatório A4	SECRETARIA MUNICIPAL DE SAÚDE SISTEMA DE INFORMAÇÃO DE ATENÇÃO		UF __ __ __
MUNICÍPIO (nome)	MUNICÍPIO (código) _ _ _ _ _ _ _ _ _ _	ZONA <input type="checkbox"/> URBANA <input type="checkbox"/> RURAL	ANO _ _ _ _

CONSOLIDADO ANUAL DAS FAMÍLIAS CADASTRADAS POR MUNICÍPIO

SEXO	FAIXA ETÁRIA	SEGMENTO TERRITORIAL										TOTAL	
<i>M A S C U L I N O</i>	< 1												
	1 - 4												
	5 - 6												
	7 - 9												
	10 - 14												
	15 - 19												
	20 - 39												
	40 - 49												
	50 - 59												
	60 e mais												
	SUB-TOTAL												
<i>F E M I N I N O</i>	< 1												
	1 - 4												
	5 - 6												
	7 - 9												
	10 - 14												
	15 - 19												
	20 - 39												
	40 - 49												
	50 - 59												
	60 e mais												
	SUB-TOTAL												
TOTAL													

	Nº	%
Famílias estimadas		
Famílias cadastradas		
População		
Pessoas cadastradas		
Pessoas com cobertura de Plano de Saúde		

	Nº	%
Crianças de 7 - 14 anos		
Crianças de 7 a 14 anos na escola		
Pessoas de 15 anos e mais		
Pessoas de 15 anos e mais alfabetizadas		

TIPO DE CASA	Nº	%
Tijolo/Adobe		
Taipa revestida		
Taipa não revestida		
Madeira		
Material aproveitado		

TRATAMENTO DA ÁGUA	Nº	%
Filtração		
Fervura		
Cloração		
Sem tratamento		
ÁGUA UTILIZADA	Nº	%

ANEXO 2 (CONTINUAÇÃO) – RELATÓRIO SSA2

Relatório SSA2	SECRETARIA MUNICIPAL DE SAÚDE SISTEMA DE INFORMAÇÃO DE ATENÇÃO BÁSICA		MÊS: __ __	ANO: __ __ __ __
MUNICÍPIO (nome):	MUNICÍPIO (código): __ __ __ __ __ __	SEGMENTO __ __	UNIDADE __ __ __ __ __ __	ÁREA __ __ __ __
RELATÓRIO DA SITUAÇÃO DE SAÚDE E ACOMPANHAMENTO DAS FAMÍLIAS NA ÁREA/EQUIPE				

MICROÁREA ⇨		1	2	3	4	TOTAL
<i>C</i>	Nascidos vivos no mês					
	RN pesados ao nascer					
	RN pesados ao nascer, com peso < 2500g					
	De 0 a 3 meses e 29 dias					
	Aleitamento exclusivo					
	Aleitamento misto					
	De 0 a 11 meses e 29 dias					
	Com as vacinas em dia					
	Pesadas					
	Desnutridas					
	De 12 a 23 meses e 29 dias					
	Com as vacinas em dia					
	Pesadas					
	Desnutridas					
	Menores de 2 anos					
	Que tiveram diarreia					
	Que tiveram diarreia e usaram TRO					
	Que tiveram infecção respiratória aguda					
<i>G</i>	Cadastradas					
	Acompanhadas					
	<i>E</i> Com vacina em dia					
	<i>S</i> Fez consulta de pré-natal no mês					
	<i>T.</i> Com pré-natal iniciado no 1º TRI					
< 20 anos cadastradas						

MICROÁREA ⇨		1	2	3	4	TOTAL
	Diabéticos	cadastrados				
		acompanhados				
	Hipertensos	cadastrados				
		acompanhados				
	Pessoas com Tuberculose	cadastradas				
		acompanhadas				
Pessoas com Hanseníase	cadastradas					
	acompanhadas					
H	Menores de 5 anos por pneumonia					
O	Menores de 5 anos por desidratação					
S	Por abuso de álcool					
P	Por complicações do Diabetes					
I	Por outras causas					
T.	Total					
	Internações em hospital psiquiátrico					
	De menores de 28 dias					
	Por diarreia					
	Por infecção respiratória aguda					
	Por outras causas					
	De 28 dias a 11 meses e 29 dias					
	Por diarreia					
Ó	Por infecção respiratória					
B	Por outras causas					
I	De menores de 1 ano					
T	Por diarreia					
O	Por infecção respiratória					
S	Por outras causas					
	De mulheres de 10 a 49 anos					
	De 10 a 14 anos					
	De 15 a 49 anos					
	Outros óbitos					
	Total de óbitos					
	De adolescentes (10-19 anos) por violência					
Total de famílias cadastradas						
Visita domiciliar - ACS						

MODELO DE ATENÇÃO ⇒	GERAL					
	URBANA		RURAL		TOTAL	
ZONA ⇒	N°	%	N°	%	N°	%
Nascidos vivos no mês						
RN pesados ao nascer						
RN pesados ao nascer, com peso < 2500g						
De 0 a 3 meses e 29 dias						
Aleitamento exclusivo						
<i>C</i> Aleitamento misto						
De 0 a 11 meses e 29 dias						
<i>I</i> Com as vacinas em dia						
<i>A</i> Pesadas						
<i>N</i> Desnutridas						
De 12 a 23 meses e 29 dias						
<i>A</i> Com as vacinas em dia						
<i>S</i> Pesadas						
Desnutridas						
Menores de 2 anos						
Que tiveram diarreia						
Que tiveram diarreia e usaram TRO						
Que tiveram infecção respiratória aguda						
Cadastradas						
<i>G</i> Acompanhadas						
<i>E</i> Com vacina em dia						
<i>S</i> Fez consulta de pré-natal no mês						
<i>T.</i> Com pré-natal iniciado no 1ºTRI						
< 20 anos cadastradas						

MODELO DE ATENÇÃO		GERAL						
ZONA ⇨		URBAN		RURAL		TOTAL		
		Nº	%	Nº	%	Nº	%	
Diabéticos	cadastrados							
	acompanhados							
	Hipertensos	cadastrados						
		acompanhados						
	Pessoas com Tuberculose	cadastradas						
		acompanhadas						
Pessoas com Hanseníase	cadastradas							
	acompanhadas							
<i>H</i>	Menores de 5 anos por pneumonia							
<i>O</i>	Menores de 5 anos por desidratação							
<i>S</i>	Por abuso de álcool							
<i>P</i>	Por complicações do Diabetes							
<i>I</i>	Por outras causas							
T.	Total							
	Internações em hospital psiquiátrico							
	De menores de 28 dias							
	Por diarreia							
	Por infecção respiratória aguda							
	Por outras causas							
	De 28 dias a 11 meses e 29 dias							
	Por diarreia							
<i>Ó</i>	Por infecção respiratória							
<i>B</i>	Por outras causas							
<i>I</i>	De menores de 1 ano							
<i>T</i>	Por diarreia							
<i>O</i>	Por infecção respiratória							
<i>S</i>	Por outras causas							
	De mulheres de 10 a 49 anos							
	De 10 a 14 anos							
	De 15 a 49 anos							
	Outros óbitos							
	Total de óbitos							
	De adolescentes (10-19 anos) por violência							
	Total de famílias cadastradas							
	Visita domiciliar - ACS							