

Ministério da Saúde

Volume 1 - Livro Texto



ASIS - Análise de Situação de Saúde

Brasília, 2014



© 2013 Ministério da Saúde. Universidade Federal de Goiás.

Todos os direitos reservados. É permitida a reprodução parcial ou total desta obra, desde que citada a fonte e que não seja para venda ou qualquer fim comercial. Venda proibida. Distribuição gratuita. A responsabilidade pelos direitos autorais de textos e imagens desta obra é da área técnica. A coleção institucional do Ministério da Saúde pode ser acessada, na íntegra, na Biblioteca Virtual em Saúde do Ministério da Saúde: <www.saude.gov.br/bvs>.

Tiragem: 1ª edição – 2013 – 500 exemplares

Elaboração, distribuição e informações:

MINISTÉRIO DE SAÚDE

Secretaria de Vigilância em Saúde

Coordenação-Geral de Informações e Análise Epidemiológica

SAF Sul, Trecho 2, lotes 05/06, bloco F,

Edifício Premium, Torre I, sala 14

CEP: 70070-600 – Brasília/DF

Telefone: (61) 3315 7708

UNIVERSIDADE FEDERAL DE GOIÁS

Instituto de Patologia Tropical e Saúde Pública

Departamento de Saúde Coletiva

Rua 235, S/N, Esq. 1ª avenida, sala 404, Setor Leste Universitário

CEP: 74605-050 – Goiânia/GO

Telefone: (62) 3209-6109 / 3209-6115

Coordenação:

Ana Lúcia Sampaio Sgambatti de Andrade – UFG, IPTSP,
Departamento de Saúde Coletiva

Elaboração de texto:

Alessandra Corrêa Tomé Teixeira de Oliveira – PUC-GO

Ana Lúcia Sampaio Sgambatti de Andrade – UFG, IPTSP

Celina Maria Turchi Martelli – UFPE, Professora visitante

Elier Broche Cristo – CGIAE, SVS, Ministério da Saúde

Elisabeth Barboza França – UFMG

Elisabeth Carmen Duarte – UnB, OPAS

José Leopoldo Ferreira Antunes – USP, FSP

Marta Rovey de Souza – UFG, IPTSP

Noêmia Teixeira de Siqueira Filha – CPqAM – Fiocruz

Otaliba Libânio de Moraes Neto – UFG, IPTSP

Ricardo Arraes de Alencar Ximenes – UFPE, UPE

Ruth Minamisava – UFG, FEN

Tatiana Sugita – UFG

Walter Massa Ramalho – UnB, Faculdade da Ceilândia

Wayner Vieira de Souza – CCPqAM – Fiocruz

Capa, projeto gráfico e diagramação:

Silvestre Linhares da Silva

Normalização:

Delano de Aquino Silva – CGDI/Editora MS

Impresso no Brasil / Printed in Brazil

Ficha Catalográfica

Brasil. Ministério da Saúde.

Análise de situação de saúde : livro-texto / Ministério da Saúde, Secretaria de Vigilância em Saúde, Universidade Federal de Goiás – Brasília : Editora do Ministério da Saúde, 2013.

xx p. : il.

ISBN

1. Diagnóstico da Situação de Saúde. 2. Análise de situação. 3. Epidemiologia. I. Título.

CDU 616-036.22

Catalogação na fonte – Coordenação-Geral de Documentação e Informação – Editora MS – OS 2013/0338

Títulos para indexação:

Em inglês: Health situation analysis: textbook

Em espanhol: Análisis de situación de salud: libro texto

Sumário

Módulo 1

10

Introdução à Análise de Situação de Saúde (ASIS)

Unidade 1 12

Conceito, Objetivo e Finalidade de ASIS

Unidade 2 14

Atributos de ASIS

Unidade 3 18

A Prática de ASIS e da Saúde Coletiva Informada pelas Evidências

Unidade 4 21

Áreas e Itens Essenciais para que as Evidências Influenciem Políticas

Unidade 5 23

Fases e Tipos de ASIS

Unidade 6 26

Vantagens do Uso dos Dados da Vigilância para as ASIS

Referências 28

Módulo 2

30

Análise de Dados dos Sistemas de Informação em Saúde

Unidade 1 33

Aspectos Conceituais do Sistema de Informação de Saúde (SIS)

Unidade 2 40

O Sistema de Informação sobre Mortalidade (SIM)

Referências 73

Módulo 3

78

Análise de Dados Demográficos

Unidade 1 80

Apresentando os Conceitos e Componentes da Dinâmica Demográfica

Unidade 2 107

Fontes de Dados Demográficos

Unidade 3 112

Tendências Populacionais e os Desafios para o Século XXI

Referências 120

Módulo 4

122

Análise das Desigualdades em Saúde

Unidade 1 125

Desigualdades em Saúde: Contexto e Conceito

Unidade 2 133

Modelos de Causalidade das Iniquidades em Saúde

Unidade 3 138

Medindo as Desigualdades

Referências 157

Módulo 5

160

Análise de Inquéritos Populacionais

Unidade 1	162
Contextualização e Exemplos de Inquéritos Epidemiológicos de Base Populacional	
Unidade 2	171
Tipos de Inquéritos	
Unidade 3	173
Conceitos Básicos de Amostragem	
Unidade 4	181
Seleção e Alocação da Amostra	
Unidade 5	188
Análise e Apresentação de Resultados	
Unidade 6	191
Vantagens e Limitações dos Inquéritos de Base Populacional	
Referências	193
Glossário	195

Módulo 6

198

Análise de Séries Temporais na Epidemiologia

Unidade 1	200
Séries Temporais: Aspectos Conceituais da Organização das Medidas de Doenças no Tempo	

Unidade 2	211
Séries Temporais: Aspectos Metodológicos da Organização das Medidas de Doenças no Tempo	
Referências	228
Glossário	229

Módulo 7

232

Sistema de Informação Geográfica Aplicada à Análise da Situação de Saúde

Unidade 1	234
Abordagem Espacial da Situação de Saúde	
Unidade 2	237
Representação Geográfica em Computadores	
Unidade 3	249
Visualização de Distribuição Espacial	
Unidade 4	264
Análise Exploratória de Dados Espaciais	
Referências	282



Prefácio

Apresento o Material Instrucional em Análise de Situação de Saúde, publicação produzida e organizada pela Secretaria de Vigilância em Saúde do Ministério da Saúde (SVS/MS) com a Universidade Federal de Goiás (UFG).

Para o enfrentamento adequado dos principais problemas e desafios postos aos profissionais e gestores do Sistema Único de Saúde (SUS) é fundamental contar com o conhecimento sobre a situação de Saúde e de seus determinantes e condicionantes.

Esse material tem como objetivo capacitar os profissionais de Saúde para o aperfeiçoamento da produção de informações, conhecimentos e evidências, que promovam uma gestão do SUS cada vez mais qualificada. Para cumprir esse objetivo, o conteúdo utilizado em Análise de Situação de Saúde valoriza as fontes de dados secundárias, produzidas pelo próprio SUS, criando a retroalimentação das fontes notificadoras e o aprimoramento das informações em Saúde por meio da utilização crítica de seus dados e promovendo sua ampla visibilidade. Dessa maneira, as informações produzidas e analisadas ganham vitalidade para o aperfeiçoamento da gestão e do exercício pleno do controle social.

Tenho certeza de que essa publicação contribuirá para ampliar a qualificação dos profissionais de Saúde nesse tema, bem como fortalecer a capacidade de formulação de políticas, programas e estratégias mais eficazes para responder às necessidades de saúde da população, fortalecendo e aprimorando o SUS.

Secretaria de Vigilância em Saúde

Apresentação

O Material Instrucional em Análise da Situação de Saúde é resultado de uma parceria entre o Ministério da Saúde e a Universidade Federal de Goiás.

Esse material destina-se a instrumentalizar profissionais e gestores de Saúde, de forma crítica e reflexiva no uso de informações e conhecimentos no campo da epidemiologia. Tem o objetivo de subsidiar o planejamento de políticas e ações em Saúde que respondam aos atuais desafios que interferem na saúde dos brasileiros, bem como contribuir para ampliar a resolutividade do Sistema Único de Saúde em todo o País.

A Secretaria de Vigilância à Saúde por intermédio da Coordenação-Geral de Informações e Análise Epidemiológica em Saúde (CGIAE) tem investido na produção de informações e divulgação de análises cada vez mais consistentes, visando possibilitar análises epidemiológicas que possibilitem a elaboração do planejamento em Saúde baseado em melhores evidências, assim como o redirecionamento das ações já em curso, que resultem na melhoria da qualidade da atenção prestada pelos serviços de Saúde. Assim, o lançamento deste material de Análise em Situação de Saúde pela SVS soma-se a uma série de outras iniciativas voltadas para esse tema.

Esta primeira edição é composta por três volumes. O primeiro contempla aspectos teóricos e conceituais, o segundo é dedicado a atividades práticas de análise da situação de Saúde utilizando dados dos Sistemas de Informação em Saúde do Ministério da Saúde e o terceiro volume apresenta a resolução das atividades, utilizando o R, que é um software livre para computação estatística e construção de gráficos.

O primeiro volume incorpora sete módulos que abordam os seguintes temas relacionados à análise de situação da saúde. O Módulo 1 – Introdução à Análise de Situação de Saúde – apresenta conceitos, objetivos, atributos e vantagens da utilização dos dados da vigilância à Saúde.

O Módulo 2 trata dos Sistemas de Informação do Ministério da Saúde, especificamente o SIM e o Sinasc, detalhando seus aspectos conceituais, formas de cálculo, Declaração de óbitos e de nascidos vivos e principais indicadores produzidos pelos sistemas.

O Módulo 3 traz como tema a Análise de Dados Demográficos, apresentando conceitos

e componentes da dinâmica demográfica brasileira, fontes de dados demográficos, tendências e desafios para o século XXI. O Módulo 4 contempla uma questão importante da análise de situação que é o estudo das desigualdades em Saúde, discutindo sua contextualização, marcos conceituais, modelos de causalidade, principais medidas das iniquidades em Saúde e o estudo da interpretação material e psicossocial das desigualdades em Saúde. Além disso, trata do papel do setor Saúde e da renda como protagonista de iniquidades.

O Módulo 5 aborda os tipos de inquéritos populacionais, como delineamento de estudo para mensurar eventos de saúde/doença visando ao planejamento e à gestão em saúde pública. São apresentados exemplos de inquéritos epidemiológicos, conceitos de amostragem, seleção e alocação da amostra, análise e apresentação de resultados e vantagens e limitações desse delineamento de estudo.

O Módulo 6 aborda a Análise das Séries Temporais em Epidemiologia, os aspectos conceituais da organização das medidas de doenças no tempo, a estimativa de tendências, a sazonalidade e o alisamento das séries temporais, com objetivo de antever futuros cenários da distribuição de doenças na população e os fatores que podem modificar esta distribuição.

Finalizando o material instrucional, no Módulo 7 são apresentados os componentes da Análise Espacial da Situação de Saúde, destacando-se os marcos teóricos, o uso de mapas em Saúde, os usos do sistema de informação geográfica para análise de situação de Saúde e a introdução à cartografia para mapeamento digital.

Dessa forma, espera-se que, os profissionais e gestores de Saúde incorporem a avaliação da situação de Saúde como ferramenta fundamental de trabalho, e que usando as informações epidemiológicas como matéria-prima para a gestão, possam ampliar seus conhecimentos sobre as principais características do sistema de Saúde dos seus municípios, estimular a elaboração de estudos na área, o monitoramento das condições de Saúde e dos principais fatores de risco e tornar obrigatória a avaliação do impacto das políticas e programas de Saúde em suas áreas de atuação.

Coordenação-Geral de Informações e Análise Epidemiológica

Módulo 1

*Introdução à Análise de Situação de Saúde**

Elisabeth Carmen Duarte

Universidade de Brasília (UnB), Faculdade de Medicina, Brasília/DF, Brasil. Organização Pan-Americana da Saúde (Opas), Brasília/DF, Brasil.

Otaliba Libânio de Moraes Neto

Universidade Federal de Goiás (UFG), Instituto de Patologia Tropical e Saúde Pública (IPTSP), Departamento de Saúde Coletiva, Goiânia/GO, Brasil.

Introdução

Este módulo propõe uma pequena revisão teórica sobre esse importante instrumento de gestão, a Análise de Situação de Saúde (ASIS), e tem como objetivos:

- Apresentar e discutir as bases conceituais da ASIS.
- Descrever as fases e os tipos da ASIS.
- Discorrer sobre as relações existentes entre a ASIS e a prática em saúde coletiva informada pelas evidências.
- Discutir as vantagens do uso dos dados da vigilância para a elaboração da ASIS no contexto da experiência brasileira.

Por que o gestor do setor Saúde precisa conhecer a situação de saúde de uma população, além da pura curiosidade acadêmica?

Talvez existam circunstâncias em que ele não necessite conhecer para intervir. Vamos imaginar que em situações muito especiais, já se saiba tudo o que é necessário saber. Nesse caso,

* Algumas reflexões contidas nesse capítulo foram preliminarmente debatidas em oficina de trabalho promovida pela Organização Panamericana da Saúde, realizada em Assunção / Paraguai, 7 a 9 de Novembro de 2011.



a prevenção ou controle de um determinado problema de saúde é mera questão da aplicação adequada de uma determinada intervenção segundo normas e princípios, sendo o seu resultado determinístico. Difícil identificar alguma circunstância em que tal pressuposto se aplique.

Uma segunda situação menos incomum é: o gestor precisa conhecer a situação de saúde de uma população por questões normativas ou burocráticas. Por exemplo, por exigência legal e/ou atendimento a requisitos de órgãos de fomento. Frequentemente fontes internacionais financiadoras de projetos governamentais exigem garantias da instituição acerca da existência de processos de monitoramento da situação de saúde em especial, de avaliação de resultados alcançados pelas intervenções, sendo essas exigências, por vezes, contratuais. Essa situação, porém, pode ser desprovida de interesse legítimo, por parte do gestor, no conhecimento da situação de saúde como instrumento útil para a melhoria da qualidade da intervenção e será cumprida, por vezes, de maneira acrítica e burocrática.

Outra situação bem mais frequente e legítima é: o gestor sabe que precisa conhecer a situação de saúde de uma população para melhor intervir. Nesse caso, o uso crítico e estratégico da Análise da Situação de Saúde tem explícito comprometimento com a melhoria da qualidade da intervenção, facilita a identificação de necessidades e prioridades em Saúde, retroalimenta os sistemas de atenção à Saúde, assim como permite o monitoramento da efetividade das intervenções, entre outras finalidades¹.

Unidade 1

Conceito, Objetivo e Finalidade de ASIS

Segundo a Opas², a Análise de Situação de Saúde (ASIS) é um processo analítico-sintético que permite caracterizar, medir e explicar o perfil de saúde-doença de uma população, incluindo os danos ou problemas de saúde, assim como seus determinantes, que facilitam a identificação de necessidades e prioridades em saúde, a identificação de intervenções e programas apropriados e a avaliação de seu impacto.

Em um curso de ASIS em Costa Rica, o módulo instrucional traz que

a ASIS é um processo pelo qual se busca explicar o estado de saúde dos habitantes, de um determinado espaço geográfico, em um dado momento. Isso é alcançado por meio da análise do entorno segundo o ponto de vista dos diferentes atores sociais³ [tradução livre dos autores].

De maneira mais formal, o mesmo autor define que

a ASIS é uma metodologia de análise, na qual o observador tem a intenção de definir a realidade de um determinado espaço populacional, partindo dos problemas identificados para posteriormente analisar seus determinantes sociais, econômicos, biológicos, ambientais ou ecológicos e de serviços de Saúde, levando em conta a perspectiva dos atores sociais para definir com eles prioridades comuns³ [tradução livre dos autores].

Em resumo, a ASIS objetiva produzir informação e conhecimento útil para orientar a ação em saúde coletiva. Essa prática é relevante para os diversos níveis de decisão (serviço de Saúde, comunidade, município, estado e federação), de modo a permitir a utilização das informações e do conhecimento produzido nas atividades de planejamento, definição de prioridades, alocação de recursos, avaliação dos programas implementados, entre outras. A ASIS assume, ainda, valor inestimável como instrumento de suporte ao controle social na medida em que amplia o acesso às informações e conhecimentos criados por essa prática e informa a comunidade e os profissionais de Saúde em todos os níveis⁴.

Além disso, assim como toda análise que valoriza as fontes de dados secundários, a ASIS cria como ganho adicional a retroalimentação das fontes notificadoras e o aprimoramento das informações em Saúde, na medida em que usa de maneira crítica seus dados e dá visibilidade para seus limites e qualidades.

Com base nessas reflexões, podemos concluir que a ASIS tem como objetivo criar informações e conhecimentos válidos sobre a situação de saúde de uma população em determinado território/contexto, mas tem como finalidade principal informar a tomada de decisão em Saúde de maneira oportuna em todas as suas instâncias. Além disso, temos como ganhos adicionais dessa prática o apoio ao controle social em Saúde e a retroalimentação das fontes notificadoras em Saúde.

Unidade 2

Atributos da ASIS

Afinal, existe alguma diferença entre o que chamamos de ASIS e os estudos acadêmicos que produzem conhecimento sobre a situação de saúde de um grupo populacional?

Para que, de fato, a ASIS seja mais do que um processo acadêmico e/ou burocrático de produção de conhecimento, ela deve apresentar certos atributos desejáveis, que serão discutidos a seguir. Não se pretende nesse momento esgotar todas as características desejáveis da ASIS, mas apenas promover uma reflexão sobre algumas delas.

1) *Processos contínuos*: As ASIS são consideradas processos contínuos, já que apresentam aproximações sucessivas e contínuas dos objetos de análise⁵. Isso é necessário principalmente porque também é dinâmico o processo de saúde-doença-atenção que a ASIS busca abordar, assim como o processo de gestão que ela busca informar.

2) *Estratégicos*: Segundo Bateman⁶, a estratégia envolve um exercício de definição de conceitos e análise da realidade, ou é “a seleção dos meios para realizar objetivos“. Assim, a ASIS é um processo estratégico na medida em que auxilia na análise da realidade e na busca dos meios para o alcance dos objetivos institucionais. Para tanto, a informação e o conhecimento criados pela ASIS devem ter o potencial para auxiliar na mobilização de recursos de toda ordem, incluindo recursos humanos, recursos financeiros e, principalmente, a decisão política para alcançar os objetivos que interessam ao setor Saúde. A ASIS deve influenciar os níveis estratégicos de decisão, incluindo o nível operacional, o tático (estratégias da alta diretoria) e, principalmente, o nível político, pois é nessa esfera em que reside a decisão em Saúde quanto às questões de longo prazo. Enfim, a ASIS deve ser vista não apenas em seu eixo metodológico, mas também em seu eixo político-estratégico no processo de decisão em Saúde.

3) *Oportunos* (em tempo-espaço-população-contexto): Para ter o papel estratégico anteriormente descrito, capaz de mobilizar a decisão política e técnica, a ASIS deve vir a tempo e a propósito de apoiar a tomada de decisão, produzindo o conhecimento necessário e útil no momento, no território e no contexto em que o gestor necessite dele. Enfim, deve ser oportuna.

4) *Análise e síntese*: Embora aparentemente contraditórias, são absolutamente complementares. A análise é entendida como o “exame de cada parte de um todo, tendo em vista conhecer sua natureza, suas proporções, suas funções, suas relações etc.”⁷ No contexto da ASIS, a análise deve promover a observação do problema/pergunta da gestão em suas partes, incluindo a busca de dados, a identificação daqueles dados que são relevantes e a transformação desses dados em informação útil. Por outro lado, a síntese é a “reunião de elementos concretos ou abstratos em um todo através de fusão ou composição”⁷. No momento que se segue ao da análise dos dados e produção da informação, é necessário descartar o que não tem relevância, agregar, priorizar e concentrar a atenção nos elementos importantes para a tomada de decisão e, com isso, produzir um todo a partir de suas partes relevantes: é necessário produzir uma síntese. Os relatórios que apresentam os achados da ASIS devem incluir a síntese e, portanto, apresentar de forma concisa o conhecimento produzido pertinente à questão sendo abordada na gestão. Vale lembrar ainda a restrição de tempo dos profissionais envolvidos na decisão em Saúde de dados. Assim, a análise dos dados, a produção de informação e de conhecimento útil e a capacidade de síntese são condições essenciais para ampliar o potencial e transformar o conhecimento produzido em ação concreta em Saúde.

5) *Analisa a situação de saúde*: O guia da ASIS de Costa Rica considera que “situação” se refere à realidade em que está imerso um “ator social” qual tem importância para as atividades que ele exerce; e que toda situação tem um contexto temporal, geográfico, social (incluindo o cultural), econômico, ecológico e biológico³. Por sua vez, “atores sociais” são

sujeitos individuais ou coletivos a partir de seus interesses, que detectam a capacidade de intervir de forma significativa em uma situação, ou que contam com recursos de poder que os tornem estratégicos no processo de construção coletiva³.

6) *Metodologicamente adequadas* (válidas e precisas): Mesmo sendo análises que visam atender à agilidade das demandas advindas da gestão em Saúde, por vezes impedindo o uso de metodologias mais sofisticadas e procedimentos de coleta de dados mais demorados, os acha-

dos produzidos pela ASIS devem ser críveis e originados de metodologias adequadas (válidas e precisas) para as perguntas elaboradas pelos gestores. O conhecimento produzido a partir de procedimentos metodológicos com baixa acurácia tem apenas dois destinos prováveis: ou será descartado pelos gestores atentos a essa questão, ou – caso sejam usados devido à baixa capacidade crítica dos tomadores de decisão – não terão os resultados esperados ou poderão produzir influências desastrosas para a gestão.

7) *Valorização da validade externa*: Enquanto a validade interna se refere à ausência de vieses nas análises feitas e à qualidade das inferências produzidas para a população de onde a amostra foi extraída, a validade externa refere-se ao potencial de aplicação dos achados nos mais diferentes contextos, e em populações distintas daquelas estudadas. Dada a complexidade das intervenções no campo da Saúde pública em contextos de grande diversidade econômica, cultural e demográfica, estudos e análises que privilegiam a validade externa são essenciais no contexto da tomada de decisão em saúde coletiva^{8,9}.

8) *Preferencialmente multidisciplinar*: Várias são as disciplinas que podem colaborar, em termos metodológicos, na produção da ASIS. Castellanos¹⁰ destaca, em especial, a necessidade de articular saberes da Epidemiologia, com a Sociologia e a Economia nos processos de ASIS. Tendo Honduras como contexto, o autor afirma que se trata de integrar, quantas vezes sejam necessárias, as facetas existentes do conhecimento da realidade social para perceber entrelaçamentos, fraturas, interdependência e independência dos fenômenos. A pesquisa qualitativa, por exemplo, pode ser útil para o entendimento da determinação social dos processos de adoecimento, para a identificação dos contextos de vulnerabilidades e a abordagem das desigualdades sociais, para aprofundar no conhecimento dos estressores sociais que causam adoecimento, entre outros. Porém, é a Epidemiologia que frequentemente oferece a base metodológica da ASIS em diferentes situações. Isso por ser a Epidemiologia uma ciência claramente orientada ao estudo dos eventos relacionados à Saúde no âmbito coletivo e por privilegiar a validade externa de seus achados, em especial, nos seus delineamentos observacionais de base populacional.

9) *Preferencialmente participativo*: Por vezes, as definições incluem explicitamente a participação, na ASIS, de diferentes atores sociais, e a valorização de processos participativos para a definição dos temas e metodologias a serem utilizadas nas ASIS³. Um processo participativo auxilia na adequada definição da pergunta a ser respondida e dos elementos de análise que devem ser abordados, promovendo a adequação dela. Além disso, o envolvimento de atores-

chave do processo de gestão, desde as primeiras etapas da ASIS, facilita a apropriação das evidências criadas por parte da gestão e a tradução do conhecimento em ação.

10) *Preferencialmente institucionalizados*: Assim como a avaliação em saúde, a ASIS com ou sem a intenção avaliativa, sofre profundas influências a depender de quem a realiza: participação ativa do grupo que participa do processo decisório ou grupos externos, academia, institutos de pesquisa, por vezes alheios a esse processo. Ambas as situações podem ser importantes e fornecer informações relevantes. No entanto, entendendo o papel fundamental de influência da ASIS no processo decisório, terá mais chances disso ocorrer quanto mais próximo estiver do ambiente institucional envolvido na decisão, assim como quanto mais mobilizados os tomadores de decisão estiverem em torno desse processo. Dessa forma é desejável promover o acolhimento institucional dessa prática, ou seja, inserir no contexto diário das instituições de Saúde a prática das ASIS como instrumento de gestão (em oposição a um instrumento de produção de conhecimento alheio aos processos de gestão).

11) *Favorável razão de custo-efetividade para a produção do conhecimento*: Finalmente, a ASIS deve produzir informação e conhecimento para a gestão implicando em esforços – gastos financeiros, de tempo, com capacitações, entre outras (= custo) aceitáveis – tendo em vista a utilidade e relevância dessa informação e conhecimento produzidos (= efetividade). Esse atributo é particularmente importante porque, por vezes, a prática da ASIS estará institucionalizada, e os esforços destinados a ela irão “competir” com os esforços exigidos pelas ações, programas e políticas de Saúde instituídas.

A revisão desses atributos nos permite revisitar o conceito de ASIS, tornando explícitas algumas de suas qualidades desejáveis para potencializar a sua influência na tomada de decisão em Saúde, como segue:

Assim,

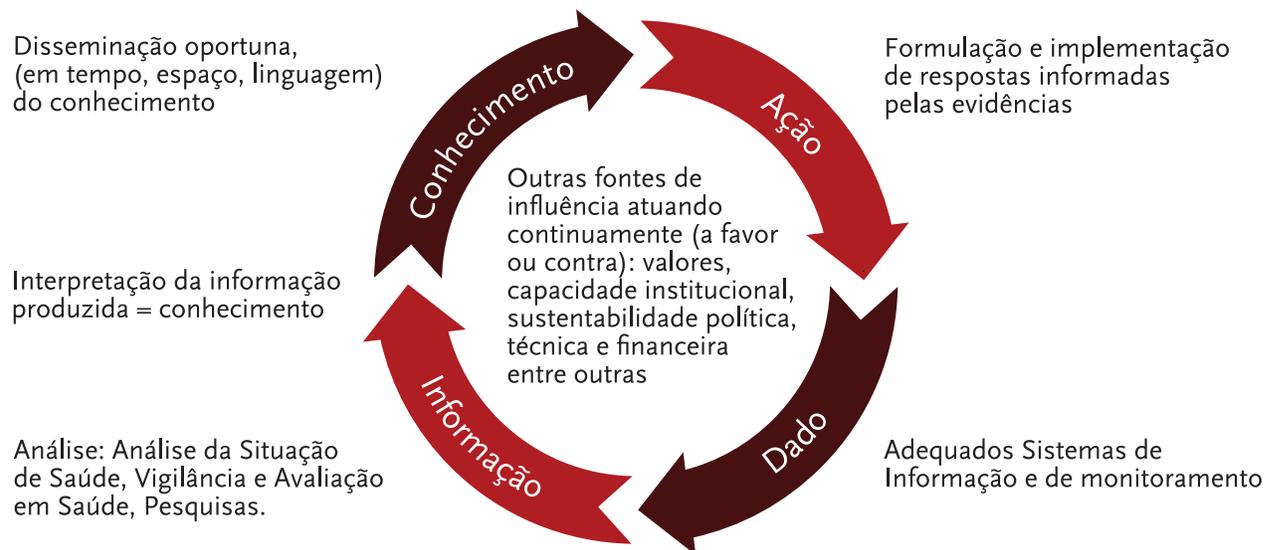
ASIS são processos contínuos e estratégicos, de análise e síntese, que permitem descrever, explicar e avaliar a tríade saúde-doença-atenção em uma população e contexto definidos, tendo em conta os seus determinantes sociais, com a finalidade principal de criar evidências válidas e oportunas para informar a decisão em Saúde pública.

Unidade 3

A Prática de ASIS e a da Saúde Coletiva Informada pelas Evidências

Como dito anteriormente, o fim último das ASIS – que talvez seja o elemento que mais a diferencie da pesquisa em Saúde em geral – é influenciar o processo decisório, auxiliando na priorização, formulação e avaliação das políticas de Saúde. Ela pode, então, ser entendida como um dos instrumentos da saúde coletiva baseada em evidências (SCBE) em um movimento espiral, de ação, produção de dados, informações, conhecimento e sua interferência para a modificação da ação (Figura 1.1).

Figura 1.1: Tomada de decisão baseada em evidências: ciclo da produção de evidências e respostas.



Fonte: Adaptado de Institute for Health Metrics and Evaluation.

A SCBE fundamenta-se no uso consciente, explícito e crítico da melhor evidência corrente disponível para a tomada de decisão sobre a atenção à comunidade e às populações no campo da proteção da saúde, prevenção de doenças e promoção da saúde. Nesse escopo, a evidência pode ser entendida como o dado, a informação ou o conhecimento acurados o suficiente para ser relevante para entender o problema ou para tomar decisão a seu respeito.

Merece ser lembrado que “criar evidências” e “criar ação em saúde” são processos fundamentalmente distintos. Por um lado, a descrição e a explicação dos processos saúde-doença-atenção – as ASIS; as pesquisas científicas e as outras formas de produzir evidências – são atos de produção de conhecimento¹¹. Por outro lado, a tomada de decisão, mediante a utilização ou não dessas evidências produzidas, constitui em atos políticos ou atos resultantes das relações sociais de poder.

Nota

Segundo Max Weber, a “política” pode ser entendida como o produto das forças exercidas entre os estados, ou no interior de um estado ou entre os grupos humanos que nele existem; e “poder” pode ser entendido como toda chance, seja ela qual for, de impor a própria vontade em uma relação social, mesmo contra a relutância dos outros.

Entendendo a complexidade do momento decisório, deve ser reconhecido que existe um conjunto de fatores que influenciam, cotidianamente, as equipes de gestão para a tomada de decisão em saúde coletiva. São exemplos desses fatores os valores e a capacidade institucional, a sua sustentabilidade política, técnica e financeira da intervenção, a aceitação social, legal e ética da intervenção, entre outros (Figura 1.1). Fatores esses que constituem em relevantes

fontes de resistência ou de suporte para a decisão em Saúde, e que interagem com projetos, valores e interesses de vários atores sociais envolvidos nesse processo.

Dessa forma, a “evidência” deve ser considerada não a resposta definitiva para determinado problema, e sim um elemento a mais de argumentação e debate, a fim de levantar pautas e definir agendas¹². De fato, segundo Lomas,¹³ diferentes áreas de influência, “legítimas” (ou não), são “acomodadas” para a tomada de decisão em saúde coletiva, incluindo aquelas ligadas à estrutura institucional (desenho institucional, atores envolvidos, regras de conduta, recursos), aos valores (crenças, ideologias e interesses) e às informações (evidências científicas, experiências, propagandas/mídia)¹³. Porém, é importante que se amplie a influência das evidências (advindas de estudos, pesquisas, ASIS entre outras) consideradas nessa “mesa de negociações” no momento do processo decisório em saúde coletiva.

Devido a esse número de fatores que, ao lado das evidências, interagem para produzir determinada ação em saúde coletiva em um dado momento, alguns autores consideram que não podemos falar em SCBE, e sim em saúde coletiva informada pelas evidências. Apesar de considerar relevante esse argumento, nesse curso – por questões de simplicidade – intercambiaremos essas duas expressões sem distinção. Entendendo a complexidade do momento decisório, deve ser reconhecido que existe um conjunto de fatores que influenciam, cotidianamente, as equipes de gestão para a tomada de decisão em saúde coletiva. São exemplos desses fatores os valores e a capacidade institucional, a sua sustentabilidade política, técnica e financeira da intervenção, a aceitação social, legal e ética da intervenção, entre outros (Figura 1.1). Fatores esses que constituem em relevantes fontes de resistência ou de suporte para a decisão em Saúde, e que interagem com projetos, valores e interesses de vários atores sociais envolvidos nesse processo.

Unidade 4

Áreas e Itens Essenciais para que as Evidências Influenciem as Políticas

Segundo Choi¹⁴, existem três áreas e 12 itens essenciais que oportunizam a influência da produção do conhecimento nas políticas, conforme descritos no Tabela 1.1.

Tabela 1.1: As três áreas e os 12 elementos essenciais para a definição de políticas baseadas em evidências.

Produção de conhecimento	Intercambio do conhecimento	Uso do Conhecimento
1 – Método adequado	1 – Conteúdo relevante	1 – Informação acessível
2 – Dados validos	2 – Tradução apropriada	2 – Mensagem compreensível
3 – Análise robusta	3 – Disseminação oportuna	3 – Usuário motivado
4 – Síntese abrangente	4 – Difusão Modular	4 – Desfecho gratificantes

Fonte: Choi, 2005.

Traduzindo esses itens essenciais para o campo das evidências criadas pela ASIS, podemos dizer que quanto mais influência terão essas análises na tomada de decisão, quanto mais permeadas por essas qualidades essenciais elas estiverem¹⁴:

a) *Na produção do conhecimento:*

- adequação do método utilizado prevenindo estudos falsos-positivos;
- validade dos dados utilizados e prevenção dos vieses;
- robustez analítica e ajuste de fatores externos;
- síntese abrangente para a apresentação resumida dos achados científicos.

b) *No intercâmbio/comunicação do conhecimento:*

- apresentar conteúdo relevante para a gestão;
- traduzir as evidências de maneira apropriada para a linguagem da gestão, simplificando a relação ciência-usuário;
- disseminar a informação de forma criativa e oportuna e vincular à necessidade para a ação;
- distribuir os achados em módulos, criando novas formas de divulgar a informação prioritária.

c) *Na absorção/uso do conhecimento:*

- tornar a informação acessível, criando meios para publicizar *market* as evidências em saúde – fazer propaganda dos achados;
- produzir a mensagem em um formato compreensível e com significado (relevante) para a audiência;
- motivar e capacitar os gestores e tomadores de decisão de maneira a torná-los ativos usuários das evidências para o processo decisório;
- desenvolver meios para efetivamente demonstrar como, o uso das evidências, podem promover desfechos recompensadores (gratificantes) para o gestor.

Unidade 5

Fases e Tipos de ASIS

Alguns autores apresentam formas de sistematizar os elementos estratégicos e metodológicos de construção das ASIS^{3,14,15}. Por questões didáticas, inspirados na proposição desses autores, discutiremos algumas fases e tipos da ASIS.

1. Fases de ASIS sob o ponto de vista estratégico

Considerando a proposição de Choi¹⁴ já comentada anteriormente, podemos identificar pelo menos três momentos relevantes para serem discutidos no contexto da ASIS: i) momento de mobilização e planejamento, ii) momento de produção de conhecimento, iii) momento de intercâmbio e uso do conhecimento (Figura 1.2).

Figura 1.2: Momentos e componentes da ASIS.



a) *Momento de mobilização e planejamento*: Um primeiro momento deve promover a mobilização, o empoderamento e a escuta dos atores sociais envolvidos no processo decisório em saúde. Quando se identifica a necessidade de se abordar determinado tema, em determinado momento e espaço, ou seja, a ASIS é identificada como necessária para atender a determinada necessidade da gestão em saúde, e essa construção deverá ser coletiva e participativa.

b) *Momento de produção de conhecimento*: Um segundo momento privilegiará a construção metodológica da ASIS propriamente dita, a depender da pergunta de interesse da gestão, das prioridades estabelecidas, dos valores e interesses dos grupos envolvidos, da disponibilidade de recursos financeiros e de tempo para a realização da ASIS e da capacidade analítica institucional existente. Nessa etapa, deve se lançar mão do arsenal metodológico disponível, elegendo os métodos mais adequados para responder a pergunta atrelada ao processo de decisão de interesse. Como mencionado, naturalmente a Epidemiologia traz um arsenal metodológico relevante nesse contexto (e essa abordagem será privilegiada nesse curso). No entanto, não se pode descartar a possibilidade de uso de outros métodos, tal como o qualitativo para lidar com o aprofundamento de questões fora do alcance da Epidemiologia. De fato, o material instrucional de Costa Rica⁵ ressalta que para se fazer ASIS se usa e analisa informações quantitativas e qualitativas, e realiza-se o processo de ASIS por meio da participação social.

c) *Momento de intercâmbio e uso do conhecimento*: E, finalmente, um terceiro momento de desenvolvimento, quando se comunica adequadamente os resultados encontrados aos diferentes públicos-alvo da ASIS, e busca vinculá-la à tomada de decisão em saúde, influenciando a gestão para a identificação de necessidades e prioridades de ações e na busca de intervenções efetivas e custo-efetivas. Segundo Morel⁶, tão importante quanto definir as prioridades nacionais na pesquisa em Saúde, é garantir que o conhecimento produzido e as intervenções sanitárias resultantes sejam efetivamente incorporados em políticas e ações de Saúde pública. E, sem dúvida, essa é uma das mais difíceis etapas dos processos da ASIS e da SCBE.

Esses três momentos são naturalmente permeados por eixos transversais, incluindo o eixo estratégico que deverá facilitar o processo participativo, o empoderamento dos atores envolvidos e a institucionalização da prática; e o eixo metodológico que diz respeito aos procedimentos científicos especificamente adotados para a produção de conhecimento (Figura 1.2).

É importante enfatizar que todos os processos de ASIS devem ser percebidos como “estratégicos”, sendo capaz de mobilizar pessoas e recursos para o alcance dos objetivos propos-

tos. Assim, o eixo estratégico é transversal em todos esses três momentos descritos, a fim de ampliar a pertinência da abordagem adotada e a capacidade da ASIS para influenciar a decisão em saúde e o acolhimento da prática no âmbito das instituições de Saúde.

2. Tipos de ASIS sob o ponto de vista metodológico

Sob o ponto de vista da abordagem metodológica – momento de produção de conhecimento – podemos dizer que existem vários tipos de ASIS, quase tantos quantos são os tipos de pesquisas e estudos aplicados, cujos delineamentos privilegiam a produção de evidências para a ação, incorporando os atributos da ASIS mencionados anteriormente. Entre os diferentes tipos de ASIS, podemos destacar²:

a) *Análise de tendências*: Analisa as mudanças da situação de saúde, a médio e longo prazos, e auxilia na construção de cenários prospectivos, podendo *eventualmente* identificar mudanças de tendências associadas às intervenções em saúde.

b) *Avaliação em saúde (efeito/impacto)*: Analisa de maneira sistemática a *influência* das ações de saúde sobre o estado de saúde para determinar – por exemplo – a efetividade dessas ações.

c) *Análise de conjuntura*: Analisa o contexto de situações definidas transversalmente ou em curtos períodos de tempo. Nesse sentido, tem especial importância o monitoramento de desigualdades em saúde (tema que será tratado em módulo específico desse curso). As análises das desigualdades em saúde abordam as diferenças nos níveis de saúde de grupos socioeconômicos distintos e fornecem elementos para o debate das iniquidades em saúde.

Exercício

Buscar na literatura artigos que exemplifiquem cada um dos tipos descritos anteriormente de ASIS. Identificar o que os caracteriza como ASIS e, em especial, ao tipo de ASIS descrito anteriormente.

Unidade 6

Vantagens do Uso dos Dados da Vigilância para as ASIS

O uso dos dados secundários e dados primários, por sua natureza, apresentam vantagens e perdas em termos de validade, oportunidade, custo e tempo despendido para a análise que devem ser analisados diante do objetivo proposto. Na Tabela 1.2 são apresentados alguns atributos da ASIS que são – usualmente – favorecidos pelo uso dos dados secundários (em contraponto com os dados primários). A Figura 1.3 resume os resultados dessa reflexão do Tabela 1.2.

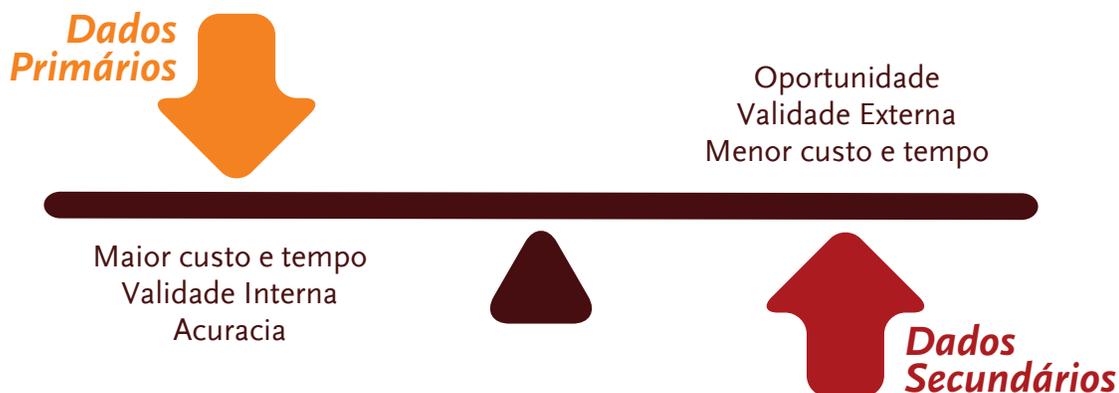
Os dados secundários, principalmente aqueles originados de sistemas contínuos de informação nacionais no Brasil (ex.: SIM, Sinasc etc.) favorecem atributos relevantes da ASIS (alguns deles discutidos anteriormente), tais como: oportunidade e continuidade das análises, precisão e validade externa dos achados, simplicidade de uso das bases de dados e menor custo, respeitabilidade dos dados, entre outros. Como ganho secundário, podemos destacar a retroalimentação dos sistemas de informação em Saúde em um processo circular de uso dos dados que produz crítica dos dados, qualidade dos dados, mais uso dos dados e assim por diante. Evidentemente, a qualidade das análises dependerá fortemente da qualidade dos dados coletados nas fontes secundárias.

Tabela 1.2: Atributos da ASIS que podem ser favorecidos pelo uso dos dados secundários.

Atributos dos dados secundários	Atributos da ASIS que são favorecidos pelo uso dos dados secundários
<p>Dado prontamente disponível</p> <p>Amplas coberturas (usualmente se refere a todo o território nacional, podendo ser universais)</p> <p>Continuidade da coleta e longos períodos de seguimento</p> <p>Dados nominais permitindo o relacionamento de bases de dados de diferentes sistemas</p> <p>Finalidade epidemiológica dos sistemas de informação</p>	<p>Oportunidade da evidência</p> <p>Favorável relação de custo-efetividade da informação/evidência</p> <p>Processo</p> <p>Precisão para as análises de tendência</p> <p>Validade externa para a construção de cenários prospectivos em saúde</p> <p>Validade interna, controle de fatores externos, ganho de acurácia da informação</p> <p>Validade interna</p>

Fonte: Elaborado pelo autor.

Figura 1.3: O que é mais importante para o gestor no momento da decisão?



Fonte: Elaborado pelo autor.

Referências

- 1 – OPAS, IBM, Universidade de Antioquia (2009) Curso Virtual de las Funciones Esenciales de Salud Publica. (Cuba, Organización Panamericana de la Salud. Instituto del Banco Mundial. Universidade de Antioquia).
- 2 – OPAS (1999) Resúmenes metodológicos en epidemiología: análisis de la situación de salud, Boletín Epidemiológico Washington. Organización Panamericana de la Salud, 1-3.
- 3 – Fernández EA., Morera PG. (2004) Análisis de situación integral de salud, Gestion local en salud para técnicos de atención primaria. Centro de Desarrollo Estratégico e Información en Salud y Seguridad Social, 62.
- 4 – Nicoletti RHA, Santos R, Rehem R, Scotti RF, Cataneli RCB, Gomes ABF et al. (2007) Vigilância em Saúde, in: J. B. d. Silva-Junior (Ed) Coleção progestores: para entender a gestão do SUS. CONASS, 278.
- 5 – Centro de Desarrollo Estratégico e Información en Salude y Seguridad Social (CENDEISSS). Curso Gestión Local en Salud para Técnicos de Atención Primária. Unidad Modular Nueve: Análisis de Situación Integral de Salud (ASIS). Universidad de Costa Rica. Costa Rica. (Autoria de conteúdos: E A Fernández & PG Morera). Primera edición, 2004.
- 6 – Bateman TS, Snell SA. (1998) Administração: construindo vantagem competitiva. São Paulo. Atlas.
- 7 – Ferreira ABH. (2010) Dicionário Aurélio da língua portuguesa. Positivo.
- 8 – Habicht JP, Victora CG, Vaughan JP. (1999) Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact, Int J Epidemiol, 28(1), 10-18.

- 9 – Victora CG, Habicht JP, Bryce J. (2004) Evidence-based public health: moving beyond randomized trials, *Am J Public Health*, 94(3), 400-405.
- 10 – Castellanos, PL. (1993) El análisis de situaciones de salud-enfermedad según condiciones de vida em Honduras. Organización Panamericana de la Salud.
- 11 – PAHO. 1999. Boletín epidemiológico ? 20 anos. Resúmenes Metodológicos en Epidemiología: Análisis de la situación de salud (ASIS). Vol. 20, No. 3.
- 12 – Black D, Townsend P, Davidson N, Whitehead M, Britain G. (1990) Inequalities in health. Penguin Books.
- 13 – Lomas J. (2000) Connecting research and policy. Disponível em: http://portals.wi.wur.nl/files/docs/ppme/lomas_e.pdf (acessado em 17 Dez 2012).
- 14 – Choi BC. (2005) Twelve essentials of science-based policy, *Prev Chronic Dis*, 2(6), A16.
- 15 – Ministerio de la Salud (Peru) (2008) Metodología para el análisis de situación de salud regional, Análisis de Situación de Salud del Perú Lima. Dirección General de Epidemiología, 229.
- 16 – Morel CM 2002. Geração de conhecimento, intervenções e ações de saúde. São Paulo em *Perspectiva* 16(4):57- 63. In Morel, CM. 2004. A pesquisa em saúde e os objetivos do milênio: desafios e oportunidades globais, soluções e políticas nacionais. *Ciência e Saúde Coletiva*. 9(2):261-270, 2004.

Módulo 2

Análise de Dados dos Sistemas de Informação em Saúde

Elisabeth Barboza França

Universidade Federal de Minas Gerais (UFMG), Faculdade de Medicina, Departamento de Medicina Preventiva e Social, Belo Horizonte/MG, Brasil.

Introdução

Este módulo aborda os Sistemas de Informação em Saúde (SIS) utilizados na análise de situação de saúde. Trataremos em particular de discutir o Sistema de Informações sobre Mortalidade (SIM) que, com o Sistema de Informação sobre Nascidos Vivos (Sinasc), representa o local onde são armazenados e processados dados sobre as estatísticas vitais. Serão abordados os principais indicadores de saúde, além de aspectos conceituais e metodológicos da avaliação dos SIS. O módulo tem um formato essencialmente instrumental, pois o objetivo é ajudar o aluno a utilizar os sistemas de informação disponíveis para as análises de situação de saúde e de avaliação dos serviços. Mas principalmente usá-los de maneira crítica, cada vez mais e melhor, de forma a explorar o enorme potencial desses sistemas. E dessa forma também se estará contribuindo para a melhoria da qualidade deles.

O módulo está dividido em duas seções complementares. A primeira seção trata de aspectos conceituais da informação em Saúde e da disponibilidade dos SIS em geral, com enfoque na relação entre os SIS, a Saúde pública e a análise de situação de saúde (ASIS). A seção sobre o SIM tem como objetivo principal facilitar a utilização crítica das estatísticas de morta-



lidade na análise da situação de saúde das populações, fundamentais também para o planejamento, o monitoramento e a avaliação em saúde. Para isso serão inicialmente apresentadas as medidas de mortalidade (taxas e mortalidade proporcional) mais utilizadas, e também os conceitos básicos sobre padronização por idade das taxas de mortalidade. A seguir será discutido o processo de coleta de dados para o SIM, abordando em especial a fonte de dados básica, a Declaração de Óbito (DO). Em seguida, será dada especial atenção à avaliação de qualidade dos dados existentes no SIM, pois apesar de ser um sistema de informação universal sobre óbitos relativamente antigo no país, a sua cobertura e a qualidade das informações sobre causas de óbito são ainda desiguais nos estados e municípios. Portanto, os dados brutos de mortalidade muitas vezes necessitam de ajustes e estimativas para possibilitarem comparações adequadas.

Há mais de 20 anos, o Ministério da Saúde reconhecia que “é o uso crítico e consequente dos dados disponíveis, o caminho adequado para melhorá-los”. E cada vez mais essa preocupação se torna pertinente devido ao amplo processo de descentralização dos serviços de Saúde

que cria necessidades específicas de análises da situação de saúde que precisam ser atendidas por ferramentas epidemiológicas adequadas. Portanto, avaliar problemas de qualidade dos dados tem importância fundamental na atual fase de consolidação do SIM no Brasil e a esse aspecto foi dado especial relevância neste módulo. Pois somente reconhecendo as limitações dos SIS, estes podem ser utilizados em toda sua potencialidade em estudos de base populacional com baixo custo. Espera-se assim contribuir para o fomento da utilização do enorme potencial dos nossos sistemas de informação em saúde, considerados experiência ímpar entre os países em desenvolvimento principalmente pela agilidade na sua produção, pela abrangência e pela disponibilidade de informações.

Unidade 1

Aspectos Conceituais de um Sistema de Informação de Saúde (SIS)

Define-se informação como o ato de informar, segundo o dicionário Aurélio. Por ser “algo” advindo de uma ação, a informação poderia ser conceituada como o significado ou sentido que se dá a determinados dados, por meio de convenções e representações. Nosso cotidiano é um processo permanente de informação. O dado, portanto, é diferente da informação, pois é uma sequencia de sinais dispersos e muitas vezes ambíguos que necessitam da mente para sua ordenação². A informação constitui-se em suporte básico para toda atividade humana e deve criar uma decisão que, por sua vez, desencadeará uma ação. Por isso, pode-se dizer que há um consenso de que não é possível exercer gerência em nenhum setor social se não houver um sistema de apoio à decisão que se sustente na informação³.

Sistema de informação é a expressão utilizada para descrever qualquer sistema automatizado ou manual, que abrange pessoas, máquinas, e/ou métodos organizados para coletar, processar, transmitir e disseminar dados que representam informação para o usuário. No caso da saúde, a Organização Mundial da Saúde (OMS) define **Sistema de Informação em Saúde (SIS)** como um mecanismo de coleta, processamento, análise e transmissão da informação necessária para planejar, organizar, operar e avaliar os serviços de Saúde. Dessa forma, os SIS devem ser delineados no seu desenho, na sua organização e estruturação de forma a responder aos objetivos estratégico-políticos, às funções, ao formato organizacional e ao grau de descentralização do Sistema de Saúde em seus níveis operacionais e de tomada de decisões⁴.

Conceitualmente, o SIS pode ser entendido como um instrumento para adquirir, organizar e analisar dados necessários à definição de problemas e riscos para a saúde, avaliar a eficácia, eficiência e influência que os serviços prestados possam ter no estado de saúde da população, além de contribuir para a produção de conhecimento acerca da saúde e dos assuntos a ela ligados⁵.

A finalidade da informação em saúde consiste em identificar problemas individuais e coletivos do quadro sanitário de uma população, propiciando elementos para análise da situação encontrada e subsidiando a busca de possíveis alternativas de encaminhamento. Assim, o âmbito das informações em saúde não se restringe ao setor Saúde em si, tendo interação com

sistemas de informação de outros setores.

Os SIS devem fornecer indicadores em três grandes áreas principais⁶:

1) determinantes de saúde, que incluem desde fatores socioeconômicos aos genéticos;

2) indicadores sobre sistema e serviços de Saúde (SS), que devem abranger desde os relativos à infraestrutura e recursos financeiros, como também a cobertura e utilização dos SS, a disponibilidade e qualidade da informação; e

3) indicadores de situação de saúde, que incluem os relacionados à mortalidade, morbidade/incapacidade e estado de saúde/qualidade de vida. Todos esses indicadores são essenciais ao processo de tomada de decisão no setor.

No Brasil, o SIS é composto por diferentes subsistemas, que produzem uma enorme quantidade de dados referentes a atividades setoriais em saúde, criando grandes bancos de dados nacionais, como o SIM, Sinasc, Sistema de Informações sobre Agravos de Notificação (Sinan), Sistema de Informações Hospitalares (SIH), Sistema de Informações Ambulatoriais do SUS (SIA-SUS) e outros. Também existem grandes bancos de dados criados em outros setores, como os relativos ao censo e pesquisas amostrais do Instituto Brasileiro de Geografia e Estatística (IBGE).

A publicação do MS/Opas/Fiocruz intitulada “A experiência brasileira em sistemas de informação em saúde” disponível em: <http://bvsmms.saude.gov.br/bvs/publicacoes/experiencia_brasileira_sistemas_saude_volume1.pdf> e em: <http://bvsmms.saude.gov.br/bvs/publicacoes/experiencia_brasileira_sistemas_saude_volume2.pdf> aborda o desenvolvimento histórico e conceitual dos SIS dentro do contexto nacional, com trabalhos de professores, pesquisadores e técnicos que tiveram papel ativo na história dos SIS. Recomendamos a leitura em especial dos capítulos nos quais são abordados os SIS de abrangência nacional sobre nascimentos (Sinasc), doenças (Sinan e SIH) e a morte (SIM).

Para refletir

Quais são as principais características dos SIS necessárias para seu uso adequado na ASIS?

A utilidade dos SIS depende da disponibilidade de informação apoiada em dados válidos (medem o que pretendem medir) e confiáveis (resultados reproduzíveis em condições semelhantes). Para a análise da situação de saúde da população tem especial relevância a cobertura dos sistemas sobre os eventos de saúde para possibilitar a estimativa do risco, ou seja, a probabilidade dos membros de uma determinada população desenvolverem alguma doença ou evento relacionado à saúde em um período de tempo. Avaliar, portanto, o desempenho dos SIS e a qualidade dos dados criados tem importância fundamental para a tomada de decisões baseadas em evidências.

1. Etapas de um sistema de informações

Segundo Moraes², no processo de produção da informação várias etapas devem ser contempladas e “[...] o sistema que produz a informação intervém em três grandes setores: onde se originam os dados, onde são processados e onde são avaliados [...]”. Para um SIS funcionar adequadamente, o mesmo autor cita as seguintes recomendações internacionais de processamento das estatísticas:

- a) Os procedimentos de coleta de dados devem estar convenientemente normatizados;
- b) Os manuais de operação devem prever todas as situações possíveis;
- c) As pessoas responsáveis pelas atividades devem conhecer a importância do que fazem em relação a todo o sistema;
- d) Deve haver supervisão e assessoria adequadas.

Dessa forma, conhecer as etapas de um Sistema de Informação é de fundamental importância para garantir a fidedignidade das bases de dados e sua utilização adequada². São as seguintes:

a) Coleta

- Origem e registro dos dados.
- Ordenamento dos documentos de coleta.
- Controle da quantidade e do conteúdo.
- Transmissão.

b) Processamento

- Recebimento e controle.
- Codificação.
- Pedido de informação adicional, se necessário.
- Transcrição.
- Classificação e tabulação.
- Controle de erros e inconsistências.
- Cálculos básicos.
- Apresentação.

c) Decisão e controle

- Análise preliminar dos dados.
- Comparação com parâmetros.
- Identificação e análise das discrepâncias.
- Opções de decisão.

De acordo com esta proposta, o objetivo final dos SIS é a produção de informação de qualidade para subsidiar as políticas de Saúde.

2. Os sistemas de informação em saúde, a Saúde pública e a análise da situação de saúde

Paim e Almeida Filho consideram a saúde coletiva

um campo de conhecimento de natureza interdisciplinar cujas disciplinas básicas são a epidemiologia, o planejamento/administração de saúde e as ciências sociais em saúde. Este

contempla o desenvolvimento de atividades de investigação sobre o estado sanitário da população, a natureza das políticas de saúde, a relação entre os processos de trabalho e doenças e agravos, bem como as intervenções de grupos e classes sociais sobre a questão sanitária⁷.

Seria, portanto, resultado da articulação de três campos disciplinares: a Epidemiologia, de produção do conhecimento científico, o planejamento e a gestão da Saúde, de aplicação tecnológica e o campo social, da promoção da Saúde.

Segundo o modelo de avaliação de desempenho dos sistemas de saúde componentes do SUS⁸, este desempenho estaria relacionado com os determinantes das condições de saúde que definem, de forma importante, as necessidades de saúde que devem orientar o financiamento e os recursos materiais e humanos indispensáveis ao bom desempenho do sistema, do qual dependerá, em parte, da melhoria das condições de saúde da população. As condições de saúde da população (segunda dimensão do modelo de avaliação) são expressas pela morbidade, mortalidade, limitação de atividade física e qualidade de vida.

Para analisar as condições de saúde da população, a Epidemiologia utiliza geralmente dados de morbidade e mortalidade. Apesar de se referirem à ocorrência de eventos indesejáveis, doença e morte, e não propriamente à qualidade de vida ou saúde no seu sentido mais amplo, são usados pela facilidade operacional e por servirem de alerta para a intervenção dos serviços de saúde.

Os indicadores de morbidade como a prevalência e, principalmente, a incidência das doenças, tornam-se úteis para análise das condições de saúde quando são relacionados à população sob risco. Mas em geral os dados de morbidade provenientes dos SIS no Brasil não têm cobertura populacional universal, exceto no caso das doenças ou agravos sob vigilância (de notificação compulsória) como o Sinan, ou com registros de base populacional, como o caso das neoplasias malignas em alguns locais. Entre os SIS existentes, o SIH representa fonte importante de dados de morbidade com abrangência nacional, mas tem cobertura estimada de 70% a 80%^{9,10} e refere-se somente a casos de doenças que demandam internação, ou seja, a morbidade hospitalar. O SIH é de suma importância, entretanto, utilizado como subsídio à avaliação de políticas de Saúde, ao planejamento e gestão, e para a mensuração de desigualdades das condições de saúde e de oferta de serviços¹¹. Nos últimos anos, os dados do SIH têm sido cada vez mais aplicados para a construção de indicadores de internações evitáveis por condições sensíveis à atenção primária na avaliação da efetividade dos serviços de saúde, como as taxas de internação por gastroenterite, pneumonia bacteriana, asma e insuficiência cardíaca¹². Ressalta-se ainda o papel do Sinasc que, apesar de ter sido implantado em 1990 com o objetivo de obter um perfil de nascimentos vivos, segundo variáveis epidemiológicas importantes, e suprir falhas da cobertura do Registro Civil¹³, tem sido cada vez mais utilizado como

fonte dados de morbidade, no caso a prevalência de anomalias congênitas¹⁴.

As medidas de mortalidade são utilizadas internacionalmente como indicadores da situação de saúde, e sua importância podem ser corroboradas pela inclusão de taxas de mortalidade precoce – como a mortalidade infantil e a materna – entre os Objetivos de Desenvolvimento do Milênio¹⁵. A morte, ao contrário da doença, representa evento único o qual é bem definido, o que facilita operacionalmente a análise no nível populacional. Além disso, a mortalidade apresenta uma regularidade bem conhecida, como por exemplo, em relação à idade, com os riscos de mortalidade sempre maiores nos extremos da vida, ou seja, ao nascimento e para os mais idosos. Para a Saúde pública, interessam mais particularmente os indicadores de mortalidade que detectam as mortes precoces e evitáveis e as causas de mortalidade, consideradas a mais básica e importante informação da Declaração de Óbito para a definição de políticas de Saúde¹⁶.

O SIM, com mais de 30 anos de existência, configura-se como um sistema de informação universal sobre óbitos já bastante consolidado no país. Entretanto, a cobertura e a qualidade das informações sobre causas de óbito são desiguais tanto entre estados quanto em relação a municípios e grupos populacionais específicos, com subenumeração de óbitos e alta proporção de óbitos registrados com causas mal definidas em algumas áreas. Sabe-se que o valor das taxas ou coeficientes de mortalidade – que estimam o risco ou a probabilidade de morte nas populações – é afetado quando há sub-registro de óbitos no SIM. E no caso das taxas de mortalidade por causas, para que reflitam melhor a realidade, é necessária também uma assistência médica de boa qualidade, com recursos propedêuticos adequados que possibilitem o diagnóstico, além do correto preenchimento da DO. Portanto, apesar da melhora importante nos últimos anos¹⁷, indicadores de saúde válidos, confiáveis e comparáveis ainda representam um desafio para o Brasil como para outros países em desenvolvimento.

O quadro abaixo apresenta um sumário dos principais SIS de abrangência nacional existentes no Brasil. O nosso intuito foi mostrar os principais SIS que coletam dados sobre eventos de saúde e possibilitam informações epidemiológicas sobre a situação de saúde.

Nos últimos anos, o Ministério da Saúde (MS) tem feito investimento específico importante nos serviços de Saúde estaduais e municipais para a melhoria das informações vitais, conforme será abordado mais adiante neste módulo. Outras iniciativas para melhoria dos SIS foram também implementadas ou reforçadas por outras instituições do país. Entre essas, merece destaque a Rede Interagencial de Informações para a Saúde (Ripsa), instituída em 1996 pelo MS em cooperação com a Organização Pan-Americana da Saúde, com o objetivo de articular instituições nacionais responsáveis pela produção, análise e disseminação de informações em saúde das três esferas do Sistema Único de Saúde, órgãos de outros setores de governo e in-

stituições de ensino e pesquisa. O produto finalístico primário da Ripsa é “uma base de indicadores configurada para fornecer panorama consistente e integrado da situação geral de saúde e suas tendências, segundo o conceito de conjunto de indicadores básicos”¹⁸. São aproximadamente 100 indicadores relacionados na matriz dos Indicadores e Dados Básicos (IDB), que são consolidados em base eletrônica na internet disponíveis no *site* do Datasus <www.datasus.gov.br informações de saúde indicadores de saúde Indicadores e Dados Básicos – IDB – 2011>.

Tabela 2.1: Principais Sistemas de Informação em Saúde Nacionais no Brasil.

SIS	Início implantação	Fonte de dados	Variável básica
Sistema de informações sobre Mortalidade-SIM	1975-1979	Declaração de óbito	Óbito
Sistema de Informações sobre Nascidos Vivos-SINASC	1991-1994	Declaração de nascido vivo	Nascido vivo
Sistema de Informação de Agravos de Notificação-SINAN	1991-1998*	Ficha de notificação e ficha de investigação	Doença de notificação compulsória
Sistema de Informações sobre Internações hospitalares-SIH-SUS	1991	Formulário Autorização de Internação Hospitalar	Internações (morbidade hospitalar)
Sistema de Informações Ambulatoriais do SUSSIA-SUS		Boletim de Produção Ambulatorial e Atendimentos de média e alta complexidade (APAC)	Procedimentos (TABNET) e atendimentos de média e alta complexidade
Sistema de Informações de Atenção Básica (Siab)		Formulários do Programa de Saúde da Família	Variáveis sobre situação de saúde e assistenciais

*O Sistema Nacional de Vigilância Epidemiológica foi implantado em 1975-1976.

Fonte: Elaborado pelo autor.

Unidade 2

O Sistema de Informações sobre Mortalidade (SIM)

As estatísticas de mortalidade são uma das principais fontes de informação de saúde e, em muitos países, constituem o tipo mais confiável de dados de saúde. Devido à sua relevância, nesta unidade, utilizar-se-á o SIM como exemplo de SIS de abrangência nacional, o qual será então abordado de forma detalhada. A unidade divide-se em três partes com subitens:

1. O SIM na análise de situação de saúde: como calcular alguns indicadores de saúde baseados em dados de mortalidade
 - 1.1 Medidas de mortalidade
 - 1.2 Padronização das taxas de mortalidade (método direto)
2. A DO como fonte de dados do SIM
3. Indicadores de qualidade dos dados criados pelo SIM e indicadores de monitoramento do funcionamento do sistema
 - 3.1 Indicadores de qualidade dos dados criados pelo sistema: a proposta dos “Dez passos para avaliação da qualidade”
 - 3.2 Indicadores de qualidade dos dados criados pelo sistema: a proposta de avaliação da qualidade das informações sobre causas de morte
 - 3.3 Indicadores de qualidade dos dados criados pelo sistema.
 - 3.4 Indicadores de monitoramento do funcionamento do SIM
 - 3.5 Propostas do Ministério da Saúde para melhoria de qualidade do SIM: a busca ativa de óbitos para a cobertura e a proposta da autópsia verbal para a investigação das causas mal definidas.

1. O SIM na análise de situação de saúde: como calcular alguns indicadores de saúde baseados em dados de mortalidade

Apesar de serem medidas negativas, as medidas de mortalidade são ainda muito utilizadas como indicadores de saúde, pois podem refletir a situação atual e as mudanças de saúde dos grupos populacionais. Como assinalado anteriormente, para a Saúde pública tem especial importância o estudo das mortes precoces ou desnecessárias ocorridas nas populações.

Os óbitos podem ser analisados quanto ao número e em relação a uma série de variáveis, tais como sexo, idade, local, profissão e outras. Para a análise da situação de saúde de uma população, entretanto, a causa da morte constitui-se em um dos aspectos mais importantes. Seu conhecimento e seu estudo são fundamentais para o setor de Saúde tanto na análise da situação de saúde das populações, quanto no planejamento e na avaliação em saúde. Para avaliar os números de óbitos e as causas dos óbitos utilizamos as medidas de mortalidade, que são de dois tipos principais: medidas relativas aos níveis de mortalidade (magnitude) e medidas relativas às causas de mortalidade (perfil).

1.1 Medidas de mortalidade

Os óbitos podem ser analisados quanto ao número e em relação a uma série de variáveis, tais como sexo, idade, local, profissão e outras. Para a análise da situação de saúde de uma população, entretanto, a causa da morte constitui um dos aspectos mais importantes. Seu conhecimento e seu estudo são fundamentais para o setor de Saúde tanto na análise da situação de saúde das populações, quanto no planejamento e na avaliação em saúde. Para avaliar os números de óbitos e as suas causas utilizamos as medidas de mortalidade, que são de dois tipos principais: medidas relativas aos níveis de mortalidade (magnitude) e medidas relativas às causas de mortalidade (perfil).

Serão aqui abordadas as seguintes medidas: taxa geral de mortalidade, taxas específicas de mortalidade por idade, duas medidas principais de mortalidade na infância (taxa de mortalidade em menores de 5 anos e a taxa de mortalidade infantil), taxa de mortalidade perinatal, razão de mortalidade materna, mortalidade proporcional por causa e taxa de mortalidade por causa específica.

1.1.1 Taxa geral (bruta) de mortalidade

Como medida do risco de morte por todas as causas, a TBM ou a Taxa Geral de Mortalidade (TGM) é um indicador pouco sensível para análise da situação de saúde das populações. Ao se examinar as taxas de mortalidade geral de duas localidades ou da mesma localidade em épocas diferentes, a maior TBM em uma delas pode ser devido ao simples fato dessa apresen-

tar maior proporção de pessoas idosas. Por exemplo, em 2005 a TGM do Brasil era 5,5/1.000 e a dos EUA era 8,3/1.000. Ou seja, é provável que os valores dessas taxas brutas tenham sido afetados pela diferente estrutura etária das populações estudadas.

$$\text{TBM} = \frac{\text{número total de óbitos na população residente em um dado ano}}{\text{população residente estimada para o meio do ano}} \times 1.000$$

Numerador: total de óbitos (todas as idades)

Denominador: população total estimada para o meio do ano calendário (estimativa da população média exposta ao risco de morte no ano)

por 1.000 habitantes usualmente, para maior facilidade de interpretação.

Para controlar a influência da estrutura etária na interpretação das taxas gerais, utilizam-se taxas padronizadas por idade. A padronização das taxas será vista mais adiante nessa seção.

1.1.2 Taxas específicas de mortalidade por idade (m_x)

$$m_x = \frac{\text{total de óbitos em cada grupo etário no ano}}{\text{população do grupo etário estimada para o meio do ano}} \times 100.000$$

As taxas podem ser calculadas para cada sexo separadamente.

$$m_x(\text{Masc}) = \frac{\text{total de óbitos em cada grupo etário no sexo masculino no ano}}{\text{pop. do grupo etário no sexo masc. estimada para o meio do ano}} \times 100.000$$

$$m_x(\text{Fem}) = \frac{\text{total de óbitos em cada grupo etário no sexo feminino no ano}}{\text{pop. do grupo etário no sexo fem. estimada para o meio do ano}} \times 100.000$$

Para refletir

Há necessidade de padronização para comparação das taxas específicas por idade em populações com diferente estrutura etária?

1.1.3 Taxa de mortalidade em menores de 5 anos (TMM5)

Para avaliar a mortalidade das crianças, são utilizadas duas medidas principais, a taxa de mortalidade em menores de 5 anos (TMM5), comumente conhecida como taxa de mortalidade na infância e a taxa de mortalidade infantil (TMI).

$$\text{TMM5} = \frac{\text{óbitos em menores de 5 anos de idade no ano}}{\text{nascidos vivos no ano}} \times 1.000$$

1.1.4 Taxa de mortalidade infantil (TMI)

$$\text{TMI} = \frac{\text{óbitos em menores de 1 ano de idade /local/ano}}{\text{nascidos vivos /local/ano}} \times 1.000$$

A TMI indica o risco de um nascido vivo vir a falecer antes de completar 1 ano de idade. Como os óbitos infantis não se distribuem de maneira uniforme, convencionou-se analisá-los segundo dois componentes principais, as taxas (coeficientes) de mortalidade neonatal ou infantil precoce e pós-neonatal ou infantil tardia:

$$\text{Taxa de mortalidade neonatal} = \frac{\text{óbitos de menores de 28 dias}}{\text{número de nascidos vivos}} \times 1.000$$

A taxa de mortalidade neonatal subdivide-se em taxa de mortalidade neonatal precoce, relativa aos óbitos de crianças menores de 7 dias de vida e taxa de mortalidade neonatal tardia, referente aos óbitos infantis de crianças de 7 dias a 27 dias de vida.

$$\text{Taxa de mortalidade pós-neonatal} = \frac{\text{óbitos de crianças de 28 dias a < 1 ano}}{\text{número de nascidos vivos}} \times 1.000$$

Detalhes em “Manual de Vigilância do óbito infantil e fetal e do Comitê de Prevenção do óbito infantil e fetal” em: <www.saúde.gov.br>:

- 1) clicar em vigilância em saúde (em “O Ministério”);
- 2) clicar em Publicações;
- 3) clicar em “Informações e análises”;
- 4) Clicar em “Óbitos”.

A partir de meados da década de 1990 e, com maior relevância, a partir de 2004, com a elaboração do Manual dos Comitês de Prevenção do Óbito Infantil e Fetal, o Ministério da Saúde tem dado prioridade na estruturação dos comitês para investigação das mortes infantis/fetais¹. Essa importante estratégia de redução da mortalidade infantil e fetal tem também contribuído na melhoria do registro dos óbitos pelos serviços de Saúde.

1.1.5 Taxa de mortalidade perinatal

Mortes perinatais são definidas na CID-10 como o conjunto dos óbitos de crianças com menos de 7 dias de vida e das perdas fetais após a 22^a semana de gestação ou com peso ao nascer >500g, ou seja, perdas fetais classificadas como intermediárias ou tardias.

$$\frac{\text{Mortes perinatais}}{\text{nascidos vivos + perdas fetais intermediárias e tardias}} \times 1.000$$

1.1.6 Razão de mortalidade materna (RMM)

Considera-se como óbito ou morte materna a morte de uma mulher enquanto está grávida ou nos 42 dias seguintes ao término da gravidez, independente da sua duração ou local, devido a qualquer causa relacionada ou agravada pela gravidez ou pela atenção a ela, mas não por causas acidentais ou incidentais. A relação entre as mortes maternas com o número de nascidos vivos é designada muitas vezes como taxa ou coeficiente de mortalidade materna. Entretanto, Razão de Mortalidade Materna (RMM) é o termo mais correto.

$$\text{RMM} = \frac{\text{óbitos por causas maternas no ano e local}}{\text{nascidos vivos no ano e local}} \times 100.000$$

Segundo o MS (ver em <www.datasus.gov.br>; informações de Saúde; estatísticas vitais; óbitos de mulheres em idade fértil e óbitos maternos; notas técnicas), devem ser selecionados como óbitos maternos os seguintes óbitos femininos, independentemente da idade, causados por:

- afecções do Capítulo XV da CID-10 “Gravidez, Parto e Puerpério”, códigos O00-O99, com exceção das mortes fora do período do puerpério de 42 dias (códigos O96 e O97);
- doença causada pelo HIV (B20-B24), mola hidatiforme maligna ou invasiva (D39.2) ou necrose hipofisiária pós-parto (E23.0), desde que a mulher esteja grávida no momento da morte ou tenha estado grávida até 42 dias antes da morte;
- osteomalácia puerperal (M83.0), tétano obstétrico (A34) ou transtornos mentais e comportamentais associados ao puerpério (F53), nos casos em que a morte ocorreu até 42 dias após o término da gravidez ou nos casos sem informação do tempo transcorrido entre o término da gravidez e a morte.

Para o cálculo da RMM, entretanto, além do sub-registro dos óbitos em algumas áreas, também as imprecisões na declaração das causas de morte dificultam a mensuração adequada. Vários estudos têm sido promovidos pelo Ministério da Saúde para identificar fatores de correção da razão de mortalidade materna. A última pesquisa ocorreu em 2002¹⁶ e identificou um fator de correção nacional de 1,4 a ser aplicado às RMMs do período.

A partir de 2008 foi regulamentada no País a vigilância de óbitos maternos (Portaria nº 1.119, de 5 de junho de 2008), inclusive os fluxos e prazos diferenciados para captação, entrada, envio de dados, conclusão da investigação e articulação com as áreas envolvidas com a morte materna. A investigação dos óbitos das mulheres em idade fértil (10 a 49 anos de idade) pelos Comitês de Morte **Materna**ⁱⁱ tem permitido de forma cada vez mais precisa a identificação de óbitos maternos.

A investigação utiliza todas as possíveis fontes de dados para identificar as mortes maternas, tais como: o próprio registro do óbito na DO, prontuários hospitalares, registros de parteiras, de cemitérios, da imprensa, entrevistas com autoridades religiosas, coveiros e escolas, por exemplo. Após a notificação, são feitas entrevistas domiciliares com a família e com médicos ou outros profissionais que cuidaram do caso.

Para refletir

A magnitude das RMM nos Estados tem relação com a proporção de óbitos de mulheres em idade fértil investigados pelos serviços de saúde?

1.1.7 Mortalidade proporcional por causa

A mortalidade proporcional por determinada causa de óbito representa a proporção de mortes ocorridas por uma causa em relação ao total de mortes. Esta proporção, comumente expressa em porcentagem, é útil para indicar a importância relativa de determinada causa em relação ao quadro total de mortalidade. Tal medida influencia o processo de seleção de áreas para estudos epidemiológicos e a determinação de prioridades para o planejamento de medidas de saúde. Entretanto, tem valor limitado quando se trata de fazer comparações entre diferentes populações e períodos de tempo, pois não mede a probabilidade ou risco de morrer por causa específica, estimada por meio do cálculo da taxa ou coeficiente de mortalidade por causa específica.

$$\text{Mortalidade Proporcional por Causa} = \frac{\text{óbitos por determinada causa no ano}}{\text{total de óbitos no ano}} \times 100$$

1.1.8 Taxa de mortalidade por causa

A taxa de mortalidade por uma causa específica estima o risco ou probabilidade de morte por uma determinada causa ou por agrupamentos de causas. Para que estas taxas reflitam melhor a realidade, é necessária uma assistência médica de boa qualidade, com recursos propedêuticos adequados que possibilitem o diagnóstico, além do correto preenchimento da DO.

$$\text{Taxa de mortalidade por X} = \frac{\text{óbitos por determinada causa X no ano}}{\text{população estimada no ano}} \times 100.000$$

1.2 Padronização das taxas de mortalidade

Quando se comparam as taxas de mortalidade geral (TMG) de duas localidades ou da mesma localidade em épocas diferentes, a maior TMG em uma delas pode ser devido ao simples fato desta apresentar maior proporção de pessoas idosas, pois pessoas de faixas etárias mais avançadas têm maior risco de morrer do que pessoas jovens. Isto também pode ocorrer quando se trabalha com coeficientes específicos por causa ou grupos de causas, pois certas causas de óbito estão concentradas em determinados grupos etários. Por exemplo, o coeficiente de mortalidade por doenças cardiovasculares será maior em população com grande contingente de pessoas idosas do que em população mais jovem.

A comparação de taxas específicas de mortalidade (TEM) por idade é um método útil para elimi-

nar o problema da composição etária, mas muitas vezes há necessidade de uma taxa única geral para facilitar comparações. Nesse caso, utiliza-se a taxa de mortalidade padronizada por idade.

A primeira etapa da padronização ou ajustamento por idade é a escolha de uma população cuja composição, em relação à variável a ser controlada é conhecida, e considerada como padrão. As taxas de mortalidade das populações a serem comparadas são então ajustadas em relação à composição da população-padrão escolhida.

Portanto, para o cálculo das taxas padronizadas, em primeiro lugar seleciona-se a população-padrão. Como população-padrão se pode utilizar a soma das populações estudadas, somente uma delas ou uma terceira populaçãoⁱⁱⁱ. Em seguida, as taxas de mortalidade específicas por idade das populações estudadas são aplicadas nos grupos etários respectivos da população-padrão, ou seja, multiplica-se cada taxa específica pelo número de pessoas no grupo etário da população-padrão. Calcula-se assim, o número de óbitos esperados em cada grupo etário. A soma dos óbitos esperados dos grupos etários representa o número total de óbitos esperados nas populações estudadas se elas tivessem a distribuição etária-padrão. Esta soma constitui o numerador do coeficiente padronizado. Este numerador dividido pelo total da população-padrão (geralmente multiplicado por 1.000) nos dá o coeficiente de mortalidade padronizado por idade.

2. A DO como fonte de dados do SIM

No Brasil, somente após 1975 com a criação do SIM, foi normatizado um único formulário para a DO no País (até então, havia mais de 40 diferentes modelos) e iniciou-se o uso mais sistemático das medidas de mortalidade para estudos epidemiológicos^{iv}.

Em estudos brasileiros, muitas vezes utiliza-se como população-padrão a população brasileira de 2000 para facilitar comparações.

Consultar os textos básicos utilizados:

1) Laurenti R, Mello Jorge MHP, Lebrão ML, Gotlieb SLD. Estatísticas de saúde. São Paulo: EPU, 1987. 2) Ministério da Saúde – publicações sobre óbitos em <www.saude.gov.br>¹⁶

1) clicar em vigilância em saúde (em "O Ministério");

2) clicar em Publicações;

3) clicar em "Informações e análises";

4) Em "Óbitos", clicar em "A Declaração de óbito: documento necessário e importante" e "Manual de procedimentos do Sistema de Informações sobre Mortalidade".

Quando ocorre um óbito, o médico preenche uma DO em três vias: a primeira via destina-se ao serviço de epidemiologia do município, a segunda via é entregue à família para registro no cartório de registro civil e a terceira via fica arquivada na unidade onde ocorreu o óbito (geralmente o hospital). Quando a família do paciente providencia o registro do óbito, o cartório arquiva a segunda via para os procedimentos legais e emite uma Certidão de Óbito para que a família providencie o enterro do corpo.

Cabe ao médico, que assina o atestado de óbito, indicar quais afecções mórbidas conduziram diretamente à morte e declarar quaisquer afecções que contribuíram com o evento. Seu papel em relação à qualidade das estatísticas de mortalidade é essencial, pois é ele quem fornece as informações existentes na DO. Entretanto, muitas vezes o médico não preenche corretamente a DO, pois não compreende sua importância em Saúde pública, considerando que o documento tem apenas finalidade legal, obrigatória para o sepultamento. Isso faz com que a apresentação da mortalidade por causas não seja plenamente satisfatória.

Em 1967, a Vigésima Assembleia Mundial de Saúde definiu as causas de morte a serem registradas no atestado médico como “todas as doenças, afecções mórbidas ou lesões que produziram a morte ou contribuíram para esta e as circunstâncias do acidente ou violência que produziram quaisquer de tais lesões”. A definição não inclui sintomas e modos de morrer, tais como parada cardíaca ou parada respiratória. O seu propósito foi assegurar o registro de toda informação relevante, procurando evitar que o atestante selecione algumas afecções para menção e rejeite outras.

Quando o médico declara as causas do óbito, a seleção de uma causa para constar nas estatísticas de mortalidade é feita de acordo com determinadas regras, que se baseiam no conceito de causa básica de morte. Vários estudos, entretanto, têm sido realizados na perspectiva das causas múltiplas de morte – apuração de todos os diagnósticos informados na DO – com o objetivo de analisar associação de causas para melhor conhecimento das doenças crônico-degenerativas que, com o envelhecimento da população, tornaram-se importantes causas de morte.

2.1 Causa básica de morte

Segundo a Sexta Conferência Internacional para Revisão Decenal da CID, a causa de morte para tabulação primária deveria ser designada como a “causa básica de morte”. Isso porque do ponto de vista de prevenção de morte é necessário interromper a cadeia de eventos ou instituir a cura em algum ponto. O objetivo em Saúde pública é prevenir a ação da causa precipitante. Para esse fim, a causa básica tem sido definida como:

“a doença ou lesão que iniciou a cadeia de acontecimentos patológicos que conduziram diretamente à morte”

ou

“as circunstâncias do acidente ou violência que produziram a lesão fatal”

No modelo de atestado médico recomendado pela Assembleia Mundial de Saúde, a Parte I do item “Causas da morte” destina-se a doenças relacionadas com a cadeia de acontecimentos patológicos que levaram diretamente à morte^v. As causas a serem registradas na Parte I do item “Causas da morte” da declaração de óbitos são as seguintes:

- a) *Causa terminal ou direta.*
- b) *Causa consequencial.*
- c) *Causa consequencial.*
- d) *Causa básica.*

Quando o médico preenche corretamente a DO, a afecção registrada na linha mais inferior da Parte I do atestado é geralmente selecionada para tabulação como a causa básica de morte; entretanto, alguns procedimentos podem determinar a seleção de outra afecção como causa básica. Isto ocorre, por exemplo, quando não há preenchimento dos eventos em sequência lógica, como no caso de DO com

- a) anemia perniciosa e gangrena do pé;
- b) aterosclerose.

Neste caso, pelas regras de seleção de causa básica, será selecionada a anemia perniciosa.

A Parte II destina-se a qualquer outra afecção significativa que contribuiu para o evento fatal, mas que não entrou na cadeia de causas listadas na Parte I que causaram diretamente a morte. Deve ser então registrada qualquer outra doença ou lesão que, a juízo do médico, tenha influído desfavoravelmente, contribuindo assim para a morte. As causas registradas na Parte II são chamadas causas contribuintes e seu registro é importante para alguns tipos de análise, principalmente do tipo de associa-

Para maiores detalhes, consultar o Manual de Instruções para o preenchimento da Declaração de Óbito, disponível na internet: <www.saude.gov.br>, conforme referenciado na nota anterior.

ções de causas de morte. Em alguns países, o tabagismo tem sido regularmente declarado pelos médicos como causa contribuinte da morte nos óbitos por doenças tabaco-associadas.

O espaço existente para registrar o “tempo aproximado entre o início da doença e a morte” de cada causa declarada deve sempre ser preenchido, quer na Parte I, quer na Parte II (ver Figura 2.1 abaixo, relativa ao campo 49 da DO, e o Anexo relativo a uma DO completa).

Figura 2.1: Declaração de óbito: “Condições e causas, do óbito”.

CAUSAS DA MORTE		ANOTE SOMENTE UM DIAGNÓSTICO POR LINHA		Tempo aproximado entre o início da doença e a morte	CID
PARTE I Doença ou estado mórbido que causou diretamente a morte	a	Devido ou como consequência de:			
	b	Devido ou como consequência de:			
	c	Devido ou como consequência de:			
	d	Devido ou como consequência de:			
CAUSAS ANTECEDENTES Estados mórbidos, se existirem, que produziram a causa acima registrada, mencionando-se em último lugar a causa básica					
PARTE II					
Outras condições significativas que contribuíram para a morte, e que não entraram, porém, na cadeia acima.					

Fonte: Ministério da Saúde, 2009.

Verificar a publicação: Ministério da Saúde, Conselho Federal de Medicina, Centro Brasileiro de Classificação de Doenças. A declaração de óbito: documento necessário e importante, 3. ed.; Brasília: Ministério da Saúde, 2009.¹⁹

2.2 Serviço de verificação de óbitos (SVO)

Destina-se ao fornecimento de DO no caso de mortes naturais sem assistência médica ou mesmo mortes de pacientes em tratamento ambulatorial, no caso de não ser designado médico pela instituição. Em locais sem SVO, as mortes naturais sem assistência médica deverão ser atestadas pelo médico do serviço de Saúde mais próximo.

2.3 Instituto Médico Legal (IML)

Caso a morte tenha sido devido a violências (homicídio, suicídio, por exemplo) ou acidentes, a DO deverá ser fornecida pelos serviços médico-legais, com preenchimento também do campo VIII “Causas externas”. Em locais sem IML, a responsabilidade cabe ao médico perito designado e, em locais com somente um médico, este deverá se responsabilizar pelo fornecimento da DO^{vi}.

2.4 Codificação das causas de óbito

Trata-se de um dos mais importantes passos na elaboração das estatísticas de mortalidade. Consiste em, partindo-se da DO, atribuir a cada causa declarada o código correspondente na Classificação Internacional de Doenças (CID). Quando só uma causa de morte é registrada, essa causa é selecionada como causa básica para tabulação. Quando mais de uma causa é registrada e a DO foi preenchida corretamente, a causa básica selecionada é a da linha mais inferior utilizada na Parte I do atestado. O princípio geral utilizado nas regras internacionais de seleção da causa básica é: “Quando mais de uma afecção for registrada no atestado, a afecção registrada sozinha na última linha utilizada da Parte I deverá ser selecionada somente se tiver dado origem a todas as afecções registradas acima dela”.

Para seleção da causa básica da morte nas DOs, os codificadores especializados devem codificar e digitar todas as causas declaradas no programa Seleção de Causa Básica (SCB). O SCB é um sistema especializado, desenvolvido em cooperação com o Centro Brasileiro de Classificação de Doenças da USP, representando a técnica mais atual na automação da seleção da causa básica da morte, pois aplica as regras constantes da Classificação Internacional de Doenças e está adaptado para os códigos da CID-10, de forma a facilitar o processo de descentralização da codificação dos atestados de óbito.

Para refletir

No caso de um natimorto (feto de 32 semanas de gestação) e mãe com história de descolamento prematuro de placenta, como deveria ser preenchida a DO?

2.5 Classificação Internacional de Doenças (CID)

Para a análise de qualquer tipo de dados, uma das primeiras tarefas é a sua organização em agrupamentos com características semelhantes. Desde 1996, o Brasil utiliza a Décima Revisão da CID. A CID-10 ou a Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde que tem 21 capítulos. As causas de morte são classificadas em categorias (2.499 possibilidades de **categorias**)^{vii}.

Utilizam-se listas mais curtas ou abreviadas para a divul-

Detalhes sobre a CID podem ser consultados em Laurenti R. Análise da informação em Saúde:1893–1993, 100 anos da Classificação Internacional de Doenças. Rev. Saúde públ. 1991, 25 (6):407-17.²⁰

gação dos dados. Na internet <www.datasus.gov.br> -> informações de Saúde -> estatísticas vitais -> mortalidade, os dados tabulados são disponibilizados conforme várias propostas de listas, como por exemplo, a lista com os capítulos da CID (21 capítulos pela CID-10), segundo categorias de três dígitos ou segundo uma tabulação mais detalhada, a CID-BR.

3. Indicadores de qualidade dos dados criados pelo SIM e indicadores de monitoramento do funcionamento do sistema

Conforme já visto, para o cálculo das taxas de mortalidade utilizam-se os dados registrados no SIM e os dados populacionais estimados (para as taxas de mortalidade de maiores de um ano e geral) ou dados de nascidos vivos (taxa de mortalidade infantil, razão de mortalidade materna, por exemplo). Sempre que os dados do SIM são de boa qualidade em determinada área (captam todos os óbitos ocorridos, ou seja, têm boa cobertura e as causas registradas e outras variáveis são confiáveis) se realiza o cálculo direto das taxas sem correções, pois se sabe que o cálculo direto é o método ideal, menos sujeito a erros de estimação²¹.

Em geral, para avaliar a qualidade, são usados indicadores de avaliação de:

- cobertura dos óbitos registrados e
- proporção de causas mal definidas de óbito.

Para estimar a cobertura dos dados de mortalidade, podem ser utilizados dois tipos de métodos:

- *Métodos diretos ou métodos de captura-recaptura*: Requerem duas fontes independentes de dados sobre óbitos, geralmente o sistema de informação de rotina e dados de um censo ou inquérito²². Apesar de mais confiáveis, são menos utilizados por exigirem pareamento individual de registros das duas fontes, com maior custo.

- *Métodos indiretos ou métodos demográficos*: Avaliam a cobertura do registro de óbitos comparando a distribuição etária de mortes com a distribuição etária da população (proveniente de censos ou *surveys*). Um censo: métodos da Equação de Balanço de Brass e de Preston-Coale. Dois censos: método de Bennett-Horiuchi e método da razão de sobrevivência intercensitária ou *Growth-Balance* generalizado. Devido ao sub-registro diferenciado em geral, opta-se por corrigir separadamente o sub-registro de óbitos dos menores de 1 ano ou 5 anos e daqueles com mais de 5 anos (mortalidade adulta).

Os métodos demográficos para mortalidade adulta estimam a cobertura dos óbitos registrados a partir de avaliação da consistência interna dos dados. A ideia básica é comparar o padrão por idade dos óbitos e população exposta ao risco de morte. Os pressupostos para sua aplicação são:

- 1) população estável, ou seja, taxas de fecundidade e mortalidade constantes;
- 2) população fechada, com migração pouco importante ou mensurável;
- 3) grau de cobertura de mortes constante para as diferentes idades **adultas**^{viii}.

No Brasil, a Ripsa divulga as estimativas de cobertura de óbitos baseadas em vários métodos demográficos, utilizando informações do Censo Demográfico e da Pesquisa Nacional de Amostra por Domicílio (PNAD) para Brasil, regiões e estados. No IDB-2011, disponível em <www.datasus.gov.br> informações de saúde, indicadores de saúde, estão as coberturas do SIM para óbitos totais (ambos os sexos) e infantis.

Para avaliar a qualidade do SIM, além da estimativa da cobertura, é comum recorrer à análise da **proporção de óbitos por causas mal definidas**. Para isso, analisa-se a distribuição dos óbitos de causas mal definidas, tradicionalmente aqueles classificados como códigos R cujas causas básicas não foram ou não puderam ser identificadas¹³. Níveis elevados de proporções de causas mal definidas podem comprometer a consistência e o uso das estatísticas de mortalidade por causas.

3.1 Indicadores de qualidade dos dados criados pelo sistema: a proposta dos “Dez passos para avaliação de qualidade”

Recentemente foi proposto um método de avaliação da qualidade dos dados, que se baseia em 10 etapas de análise dos dados de mortalidade²².

Detalhes em: Agostinho, C.S.; Queiroz, B.L. Estimativas da mortalidade adulta para o Brasil no período 1980/2000: uma abordagem metodológica comparativa. In: XVI Encontro Nacional de Estudo de População da ABEP, 2008, Caxambu, Anais do Encontro Nacional de Estudos de População, 2008. p.1-20.²³

Etapa 1: Realizar tabulações básicas de óbito por idade, sexo e causa

Essas tabulações devem incluir o número de óbitos para cada ano específico analisado, segundo as seguintes variáveis:

- Por sexo, isto é, homens e mulheres separadamente.
- Por idade do óbito, utilizando os seguintes grupos etários:
 - menos de 28 dias de vida (óbitos neonatais);
 - de 28 dias a menos de 1 ano (óbitos pós-neonatais);
 - de 1 a 4 anos de idade;
 - 5-9 anos de idade;
 - 10-14 anos etc., até 75-80 anos;
 - 80 e mais anos.
- Por causa do óbito, utilizando listas condensadas.

Essa tabulação básica deve incluir também a população estimada para o ano, segundo sexo e grupo etário.

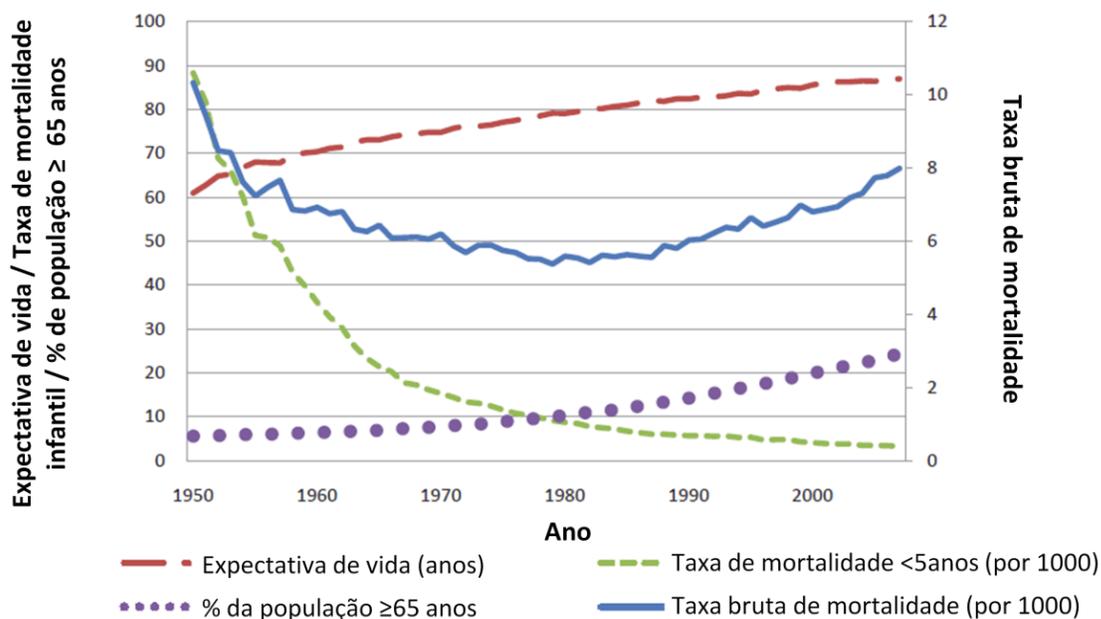
Etapa 2: Calcular a taxa geral (bruta) de mortalidade (TBM ou TGM)

As TGMs podem ser indicadoras da qualidade (no caso a completude) dos dados disponíveis sobre óbitos de uma população. Em geral, uma TGM menor que 5/1.000 sugere subnotificação de óbitos ao sistema de informação sobre óbitos²², no caso do Brasil o SIM. Exceções são pouco frequentes e podem ocorrer no caso de populações com altas taxas anuais de crescimento (TAC) devido ao excesso de nascimentos, migração ou ambos (2,5% ou maior), combinadas com baixas taxas de mortalidade por idade, inclusive a taxa de mortalidade infantil, levando a uma esperança de vida ao nascer (e^o) comparativamente alta (70 anos ou mais). Por exemplo, no caso de uma e^o de 70 anos, a TAC teria que ser de 5,0%.

Para avaliar se existe sub-registro de óbitos, é importante calcular as TGMs para homens e mulheres separadamente. Geralmente se esperam maiores taxas para homens. Também a análise da tendência temporal das TGMs pode indicar problemas de qualidade dos dados. Pode ser avaliada de duas maneiras diferentes:

- de um ano para o outro, pois não se esperam grandes variações;
- em períodos maiores (tendência secular), comparando com a tendência de outros indicadores, como a taxa de mortalidade de menores de 5 anos (TMM5) e a proporção da população com 65 anos e mais. No Japão (Figura 2.2), a evolução temporal das TGMs mostra um modelo bem definido, com padrão decrescente relacionado à diminuição das TMM5 até 1980 e crescente a partir de então, devido ao aumento da proporção de idosos na população e da esperança de vida ao nascer.

Figura 2.2: Principais tendências demográficas no Japão 1950-2007 (mulheres).



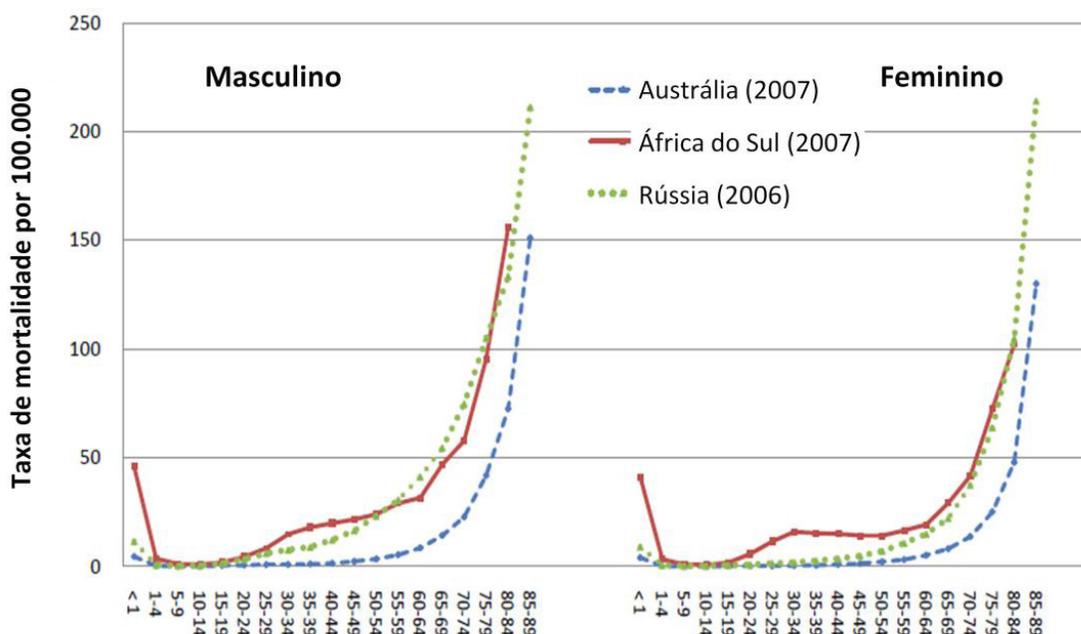
Fonte: Health Information Knowledge hub, 2010 apud Abouzahr et al., 2010.²²

Etapa 3: Avaliar se as taxas específicas de mortalidade têm padrão esperado segundo idade e sexo

- Segundo a idade: as taxas de mortalidade são maiores em crianças, atingem o menor nível nas idades de 5-14 anos, depois aumentam exponencialmente após os 30 anos de idade (possível exceção pode ocorrer devido a taxas altas de mortalidade por causas externas em adultos jovens, por causas maternas em mulheres ou por HIV/aids nas idades de 15-49 anos).
- Segundo o sexo: as taxas de mortalidade são maiores em homens em todas as idades.

Na Figura 2.3, compara-se o padrão das taxas de mortalidade por idade na Austrália, onde os registros de óbitos são completos com a África do Sul e Rússia. Na África do Sul, existe maior risco de morte em crianças e ocorre um abaulamento da curva nas idades reprodutivas em ambos os sexos devido às mortes prematuras por HIV/aids. A avaliação de gráficos como esse pode indicar possível sub-registro de óbitos em certos grupos etários.

Figura 2.3: Taxas específicas de mortalidade por idade para Austrália, Rússia e África do Sul, homens e mulheres, 2000.



Fonte: Health Information Knowledge hub, 2010 apud Abouzahr et al., 2010.²²

Como pode ser visto na Figura 2.3, após a idade de 35 anos, as taxas de mortalidade aumentam exponencialmente com a idade. Portanto, o logaritmo (ln) das taxas específicas de mortalidade por idade (m_x) deve ser uma linha reta quando a idade (x) aumenta. Portanto, uma maneira melhor de avaliar problemas de qualidade dos dados seria preparar um gráfico do log das taxas específicas de mortalidade por idade – $\ln(m_x)$.

Pois se

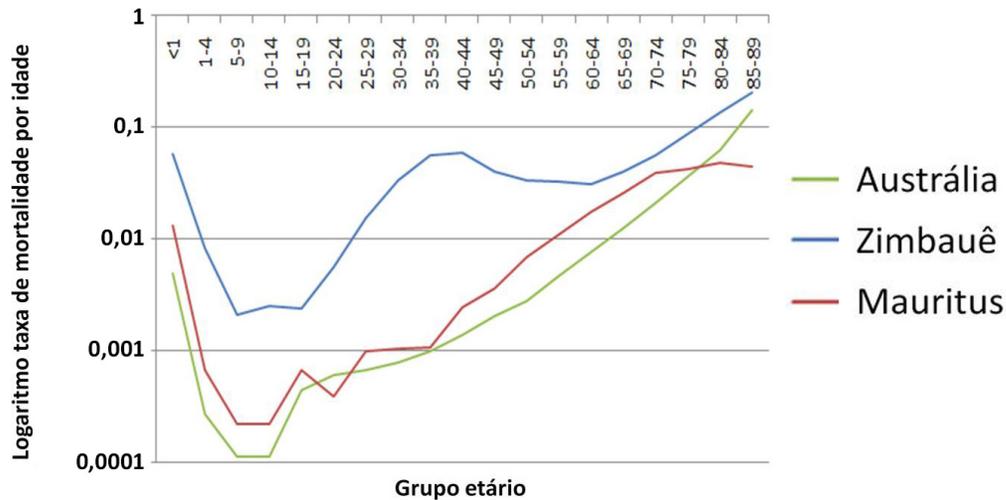
$$m_x = e^x, \text{ para idades } x > 30 \text{ ou mais}$$

Então

$$\ln(m_x) = \ln(e^x) = x$$

i.e. O log natural das taxas específicas de mortalidade por idade (m_x) deve ser uma linha reta após 30 anos de idade, conforme pode ser visto na Figura 2.4, relativo à Austrália.

Figura 2.4: Taxas de mortalidade específica por idade de Zimbauê (2006), Mauritius (2000) e Austrália (2005).



Fonte: Health Information Knowledge hub, 2010 apud Abouzahr et al., 2010.²²

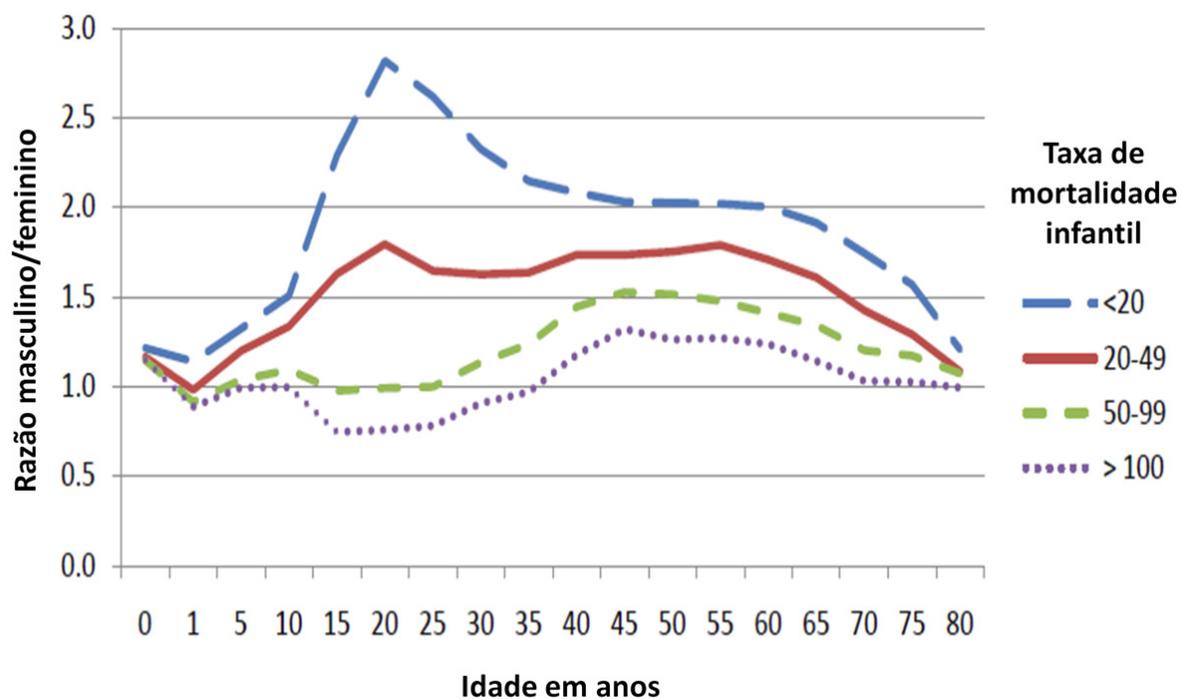
Cálculo da razão homem/mulher das taxas específicas de mortalidade por idade

$$\text{Razão homem/mulher} = \frac{m_x \text{ em cada grupo etário no ano para homens}}{m_x \text{ em cada grupo etário no ano para mulheres}}$$

Essa razão é sempre maior que 1,0 em qualquer país, exceto no caso de sociedades em que a mulher tem um status muito baixo. Entretanto, essa razão varia de acordo com o grupo etário e geralmente é maior na idade de 15 a 29 anos devido às mortes por causas externas em homens jovens. Aumenta também na idade de 55-64 anos, devido a maiores taxas de mortalidade por doenças crônicas entre os homens por maior exposição a fatores de risco como tabagismo, dieta inadequada e obesidade.

Na Figura 2.5 apresenta-se o padrão típico da razão homem/mulher das taxas de mortalidade por idade em países com diferentes níveis de mortalidade, conforme as diferentes taxas de mortalidade infantil (TMIs) existentes (Atenção: as TMIs são estimadas utilizando fonte independente dos dados de mortalidade analisados).

Figura 2.5: Razão (*ratio*) homem–mulher das taxas específicas de mortalidade por idade em locais com diferentes taxas de mortalidade infantil.

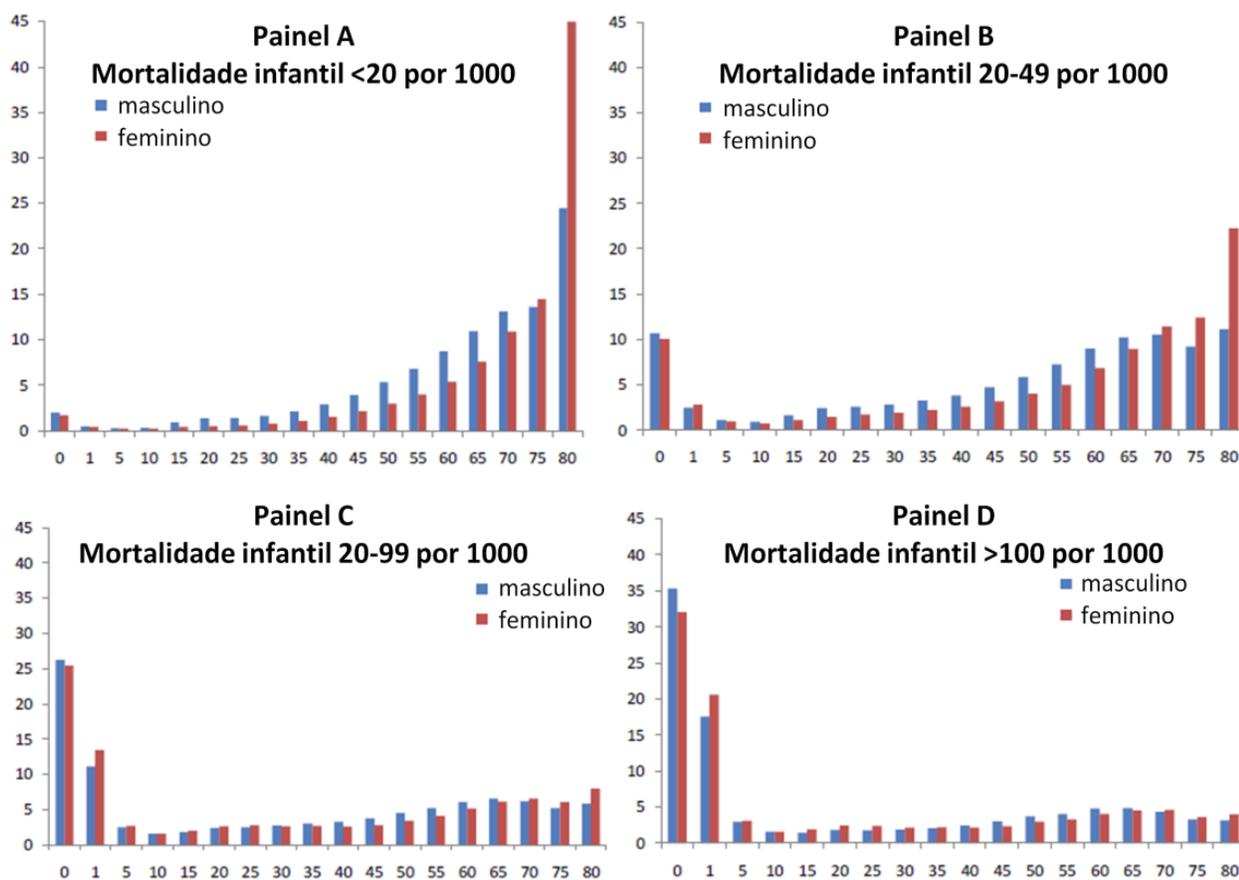


Fonte: Health Information Knowledge hub, 2010 apud Abouzahr et al., 2010.²²

Etapa 4: Revisar a distribuição dos óbitos por idade

A distribuição proporcional dos óbitos por idade na população estudada deve ser semelhante a uma das distribuições esperadas de acordo com o nível de mortalidade infantil, apresentadas na Figura 2.6. Por exemplo, se a TMI for menor que 20/1.000, espera-se que a distribuição por idade dos óbitos registrados seja semelhante ao “Painel A”.

Figura 2.6: Típica distribuição da proporção dos óbitos por idade segundo sexo (em %) em locais com diferentes níveis de mortalidade infantil.



Fonte: Health Information Knowledge hub, 2010 apud Abouzahr et al., 2010.²²

Diferenças importantes do padrão esperado indicam haver problemas na idade registrada dos óbitos. Nesse caso, pode ser devido à preferência digital, ou seja, a preferência em relatar a idade do óbito como um número terminado em zero ou 5 (ex.: 45, 50, 55). Também pode ter ocorrido relato incorreto da idade do óbito, pois geralmente se prefere relatar que o falecido era mais velho do que realmente era.

Nascido vivo: É a expulsão ou a extração completa de um produto da concepção do corpo materno, independentemente da duração da gestação, o qual, depois da separação do corpo materno, respire ou dê qualquer outro sinal de vida, tais como: batimento do coração, pulsação do cordão umbilical ou movimentos efetivos dos músculos da contração voluntária, estando ou não cortado o cordão umbilical e estando ou não desprendida a placenta (CID-10).

Etapa 5: Avaliar os indicadores de mortalidade na infância

As estimativas das taxas de mortalidade infantil (TMIs) e na infância são mais complexas que no caso da mortalidade de adultos, pois dependem de se ter informações completas tanto no numerador (óbitos infantis) quanto no denominador (nascidos vivos). Em geral, o sub-registro de óbitos infantis é muito maior que em outras faixas etárias, principalmente no caso das mortes ocorridas no primeiro mês de vida (neonatais), muitas vezes devido à aplicação incorreta do conceito de nascido vivo^{ix}. Por exemplo, em estudo realizado na Tailândia para estimar a proporção de sub-registro de óbitos nas estatísticas vitais, verificou-se ser esta bem maior para crianças menores que 5 anos de idade (Tabela 2.2).

Tabela 2.2: Proporção de sub-registro de óbitos por idade. Tailândia, 2005.

	0-4 anos	5-49 anos	50-74 anos	75 e + anos	Total
Proporção de sub-registro	42,8	14,8	7,7	5,9	8,7

Fonte: Popakkam et al., 2010.²⁴

Pelas razões acima comentadas, em grande parte dos países do mundo, a taxa de mortalidade infantil e outros indicadores de saúde na infância não são provenientes dos sistemas de informação de rotina (como o SIM e Sinasc, no caso do Brasil), sendo os dados provenientes de censos e inquéritos populacionais considerados fontes de informação mais confiáveis. As estimativas de mortalidade na infância provenientes dessas fontes são utilizadas para avaliar a cobertura dos sistemas de rotina de estatísticas vitais (óbitos e nascimentos).

No Brasil, a Ripsa publica estimativas indiretas das TMIs para 19 UFs, pois em somente oito, o SIM e o Sinasc são considerados com qualidade suficiente para que o cálculo das taxas seja realizado com o método direto.

Etapa 6: Avaliar a distribuição das causas de morte segundo grandes grupos de causas

A CID contém cerca de 3.000 possíveis causas de morte. Todas essas causas podem ser condensadas em três grandes grupos, segundo a lista de classificação de causas de morte utilizada no Estudo da Carga Global de Doenças ou Global Burden of Disease Study (GBD) proposta com o objetivo de converter os dados sobre causas de morte em informação útil para as políticas de Saúde²⁵:

- a) Grupo I: Doenças infecciosas, causas maternas, perinatais e nutricionais;
- b) Grupo II: Doenças crônico-degenerativas e
- c) Grupo III: Causas externas^x.

Para avaliar a qualidade dos dados sobre causas de óbito, compara-se a distribuição proporcional dos três grupos de causas de morte da população estudada com a distribuição esperada de acordo com a esperança de vida ao nascer das populações (ver tabela a seguir). Um desvio maior do padrão esperado que não possa ser explicado por alguma característica local (por exemplo, epidemia de aids em adultos jovens) deve ser avaliado como possível problema de qualidade dos dados sobre causas de óbito.

Grupo I (Doenças infecciosas, causas maternas, perinatais e nutricionais): A00-B99, G00-G04, N70-N73, J00-J06, J09-J18, J20-J22, H65-H66, O00-O99, P00-P96, E00-E02, E40-E46, E50, E86, D50-D53, D64, E51-E64;

Grupo II (Doenças crônico-degenerativas): C00-C97, D00-D48, D55-D63, D65-D89, E03-E07, E10-E16, E20-E34, E65-E85, E87-E90, F00-F99, G06-G98, H00-H61, H68-H93, I00-I99, J30-J99, K00-K92, N00-N64, N75-N99, L00-L98, M00-M99, Q00-Q99;

Grupo III (Causas externas): V01-Y98

Tabela 2.3: Distribuição esperada de causas de óbito por grandes grupos, segundo a expectativa de vida ao nascer.

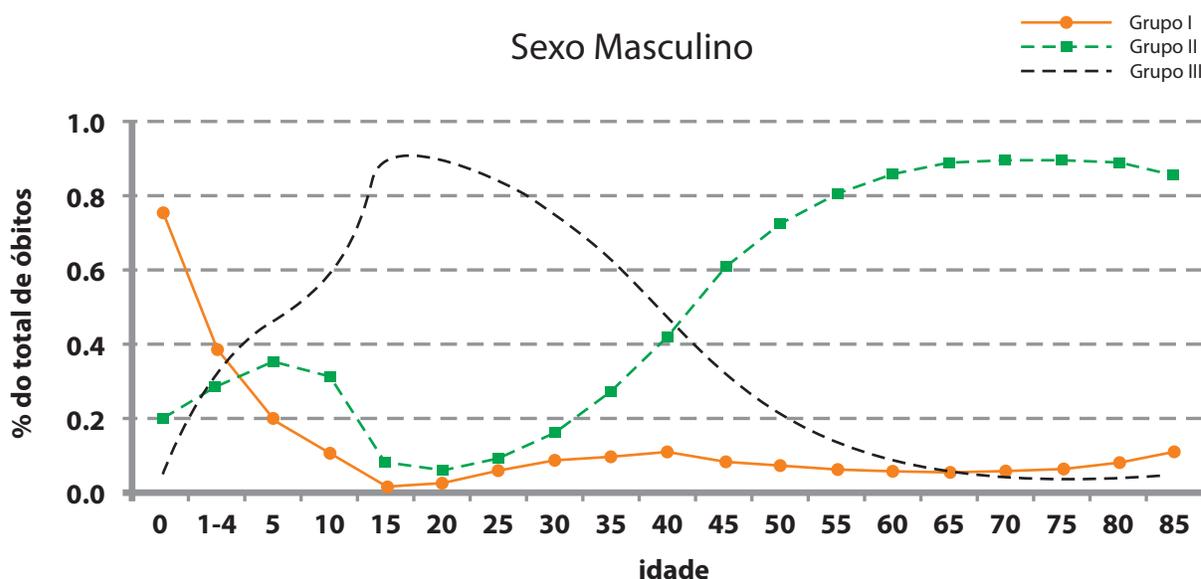
Grupo de causa de morte	Expectativa de vida			
	55 anos	60 anos	65 anos	70 anos
Grupo I	22%	16%	13%	11%
Grupo II	65%	70%	74%	78%
Grupo III	13%	14%	13%	11%

Fonte: AbouZahr et al., 2010.²²

Etapa 7: Avaliar a distribuição por idade das causas de morte segundo grandes grupos de causas

A distribuição dos três grandes grupos de causas de morte segue um determinado padrão de acordo com a idade do óbito, com proporção maior de óbitos para as causas do Grupo I em crianças, para o Grupo II em idosos e Grupo III em adultos jovens. A Figura 2.7 a seguir apresenta o padrão esperado de distribuição das causas para homens segundo idades. Qualquer desvio maior na população estudada é sugestivo de problemas na declaração das causas e da idade do óbito.

Figura 2.7: Distribuição de grandes causas de morte por grupo e idade.



Fonte: AbouZahr et al., 2010.²²

Importante: os cálculos são feitos para o total de óbitos por causas definidas, ou seja, as causas mal definidas não são consideradas no cálculo.

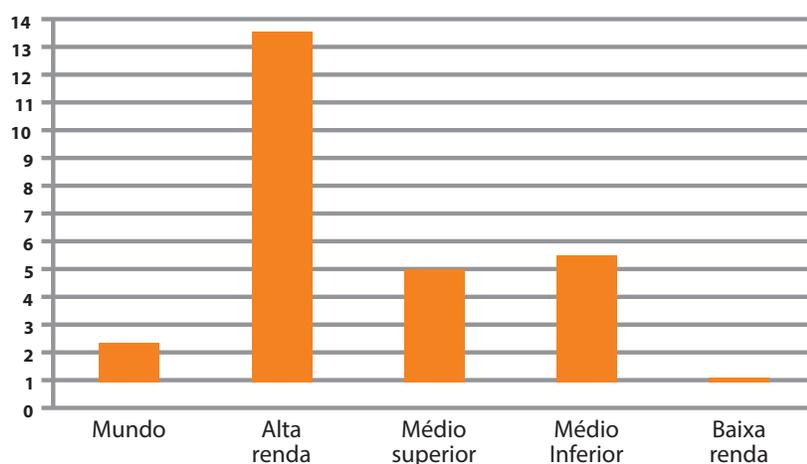
Etapa 8: Avaliar a distribuição das principais causas de morte segundo grupos de causas específicas

A distribuição das principais causas de mortalidade, segundo grupos específicos, deve ser analisada, pois em países de renda alta em geral as doenças isquêmicas do coração representam a principal causa de morte, enquanto em países de renda média são as doenças cerebrovasculares. Qualquer desvio maior na população estudada é sugestivo de problemas na declaração das causas.

Etapa 9: Razão entre óbitos por doenças do Grupo II (crônicas) e doenças do Grupo I (infecciosas, maternas, perinatais e nutricionais)

Em geral esta razão amplia com o aumento da esperança de vida ao nascer e, em países de alta renda, as doenças crônicas são responsáveis por cerca de 14 vezes mais óbitos que as do Grupo I. Por outro lado, esta razão é próxima de um em países subdesenvolvidos. A Figura 2.8 a seguir apresenta o padrão esperado e permite avaliar os dados sobre causas de morte da população estudada.

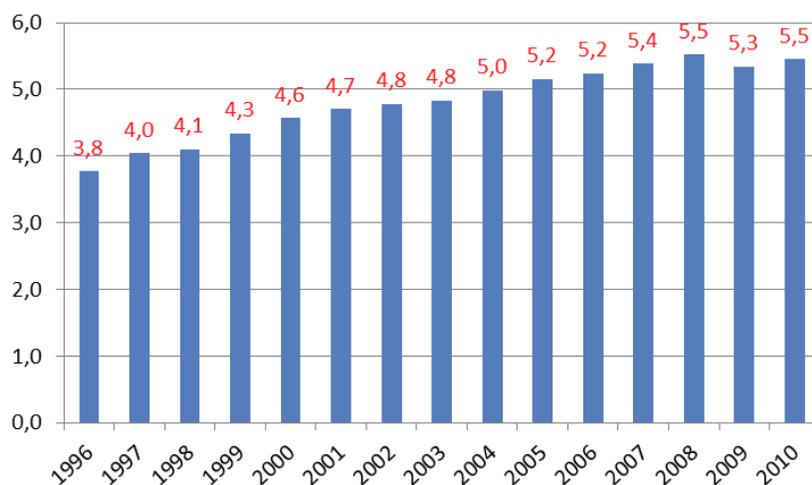
Figura 2.8: Razão de doenças não notificáveis por notificáveis por grupo de países, 2004.



Fonte: AbouZahr et al., 2010.²²

Para o Brasil, essa razão apresenta um comportamento condizente com o esperado, conforme a Figura 2.9 a seguir.

Figura 2.9: Razão entre óbitos por causas não transmissíveis e causas transmissíveis, Brasil, 1996 a 2010.



Fonte: SIM–CGIAE/DASIS/SVS/MS e IBGE *dados do SIM preliminares para 2010 (Apresentação Dácio Rabello – novembro de 2011).

Etapa 10: Causas mal definidas de óbito

As causas de óbito classificadas como causas mal definidas (CMD), ou seja, com um dos códigos R do Capítulo XVIII (sintomas, sinais e achados anormais de exames clínicos e de laboratório não classificados em outra parte), de acordo com a CID-10, prejudicam a análise dos riscos e as tendências das principais causas de morte.

É importante identificar a percentagem de CMD (capítulo da CID), total e por idade. Em geral, espera-se que essa proporção seja menor que 5% para menores de 65 anos e até 10% para 65 anos e mais. É importante também analisar essa proporção no tempo, pois mudanças bruscas sinalizam para possíveis problemas de qualidade, e.g. alteração de práticas de codificação, mais que mudanças reais no padrão de mortalidade por causas.

Não existe consenso sobre os códigos a serem considerados como códigos inespecíficos ou garbage, mas os códigos aqui considerados são os mais utilizados e aceitos internacionalmente.

Além dos códigos classificados no Capítulo XVIII da CID-10, outros códigos de outros capítulos também têm sido considerados inespecíficos ou “códigos-lixo” por não caracterizarem a causa da morte. Entre esses, estão códigos incluídos nos capítulos das doenças cardiovasculares, câncer e causas externas^{xii, 26}. São os seguintes:

- Doenças cardiovasculares: I46 (Parada cardíaca), I47.2 (Taquicardia ventricular), I49.0 (Flutter e fibrilação ventricular), I50 (Insuficiência cardíaca), I51.4 (Miocardite NE), I51.5 (Degeneração miocárdica), I51.6 (Doença cardiovascular NE), I51.9 (Doença não especificada do coração), I70.9 (Aterosclerose generalizada e a não especificada).
- Câncer: C76 (Neoplasia maligna de outras localizações e de localizações mal definidas), C80 (Neoplasia maligna, sem especificação de localização) e C97 (Neoplasias malignas de localizações múltiplas).
- Causas externas: Y10-Y34 (Eventos de intenção indeterminada).

Na tabela a seguir, são apresentados dados para o Brasil e regiões em 2002-2004. Nesse período, as regiões tinham proporções de CMD muito diferentes, com o Nordeste e o Norte com altas proporções²⁷.

Tabela 2.4: Porcentagem do total de óbitos por causas mal definidas e por outras não especificadas. Brasil e regiões, 2002-2004.

Grupo de causas não especificadas	Norte	Nordeste	Sudeste	Sul	Centro-Oeste	Brasil
Sintomas ^a	21,2	25,4	8,8	6,4	6,0	13,1
Causas externas ^b	0,5	1,3	1,5	0,6	0,4	1,2
Câncer ^c	0,6	0,6	0,8	0,9	0,8	0,8
Doenças cardiovasculares ^d	2,4	3,1	3,1	3,8	3,9	3,3
Total mal definidas	25,0	30,5	14,3	11,7	11,1	18,3

a – Códigos da CID10: R00-R99

b – Codigos da CID10: Y10-Y34 e Y87.2

c – Codigos da CID10: C76, C80, C97

d – Codigos da CID10: I01.0, I01.1, I01.2, I01.3, I01.4, I01.5, I01.6, I01.7, I01.8, I01.9, I02.0, I02.1, I02.2, I02.3, I02.4, I02.5, I02.6, I02.7, I02.8, I02.9, I03.0, I03.1, I03.2, I03.3, I03.4, I03.5, I03.6, I03.7, I03.8, I03.9, I04.0, I04.1, I04.2, I04.3, I04.4, I04.5, I04.6, I04.7, I04.8, I04.9, I05.0, I05.1, I05.2, I05.3, I05.4, I05.5, I05.6, I05.7, I05.8, I05.9, I06.0, I06.1, I06.2, I06.3, I06.4, I06.5, I06.6, I06.7, I06.8, I06.9, I07.0, I07.1, I07.2, I07.3, I07.4, I07.5, I07.6, I07.7, I07.8, I07.9, I08.0, I08.1, I08.2, I08.3, I08.4, I08.5, I08.6, I08.7, I08.8, I08.9, I09.0, I09.1, I09.2, I09.3, I09.4, I09.5, I09.6, I09.7, I09.8, I09.9, I10.0, I10.1, I10.2, I10.3, I10.4, I10.5, I10.6, I10.7, I10.8, I10.9, I11.0, I11.1, I11.2, I11.3, I11.4, I11.5, I11.6, I11.7, I11.8, I11.9, I12.0, I12.1, I12.2, I12.3, I12.4, I12.5, I12.6, I12.7, I12.8, I12.9, I13.0, I13.1, I13.2, I13.3, I13.4, I13.5, I13.6, I13.7, I13.8, I13.9, I14.0, I14.1, I14.2, I14.3, I14.4, I14.5, I14.6, I14.7, I14.8, I14.9, I15.0, I15.1, I15.2, I15.3, I15.4, I15.5, I15.6, I15.7, I15.8, I15.9, I16.0, I16.1, I16.2, I16.3, I16.4, I16.5, I16.6, I16.7, I16.8, I16.9, I17.0, I17.1, I17.2, I17.3, I17.4, I17.5, I17.6, I17.7, I17.8, I17.9, I18.0, I18.1, I18.2, I18.3, I18.4, I18.5, I18.6, I18.7, I18.8, I18.9, I19.0, I19.1, I19.2, I19.3, I19.4, I19.5, I19.6, I19.7, I19.8, I19.9, I20.0, I20.1, I20.2, I20.3, I20.4, I20.5, I20.6, I20.7, I20.8, I20.9, I21.0, I21.1, I21.2, I21.3, I21.4, I21.5, I21.6, I21.7, I21.8, I21.9, I22.0, I22.1, I22.2, I22.3, I22.4, I22.5, I22.6, I22.7, I22.8, I22.9, I23.0, I23.1, I23.2, I23.3, I23.4, I23.5, I23.6, I23.7, I23.8, I23.9, I24.0, I24.1, I24.2, I24.3, I24.4, I24.5, I24.6, I24.7, I24.8, I24.9, I25.0, I25.1, I25.2, I25.3, I25.4, I25.5, I25.6, I25.7, I25.8, I25.9, I26.0, I26.1, I26.2, I26.3, I26.4, I26.5, I26.6, I26.7, I26.8, I26.9, I27.0, I27.1, I27.2, I27.3, I27.4, I27.5, I27.6, I27.7, I27.8, I27.9, I28.0, I28.1, I28.2, I28.3, I28.4, I28.5, I28.6, I28.7, I28.8, I28.9, I29.0, I29.1, I29.2, I29.3, I29.4, I29.5, I29.6, I29.7, I29.8, I29.9, I30.0, I30.1, I30.2, I30.3, I30.4, I30.5, I30.6, I30.7, I30.8, I30.9, I31.0, I31.1, I31.2, I31.3, I31.4, I31.5, I31.6, I31.7, I31.8, I31.9, I32.0, I32.1, I32.2, I32.3, I32.4, I32.5, I32.6, I32.7, I32.8, I32.9, I33.0, I33.1, I33.2, I33.3, I33.4, I33.5, I33.6, I33.7, I33.8, I33.9, I34.0, I34.1, I34.2, I34.3, I34.4, I34.5, I34.6, I34.7, I34.8, I34.9, I35.0, I35.1, I35.2, I35.3, I35.4, I35.5, I35.6, I35.7, I35.8, I35.9, I36.0, I36.1, I36.2, I36.3, I36.4, I36.5, I36.6, I36.7, I36.8, I36.9, I37.0, I37.1, I37.2, I37.3, I37.4, I37.5, I37.6, I37.7, I37.8, I37.9, I38.0, I38.1, I38.2, I38.3, I38.4, I38.5, I38.6, I38.7, I38.8, I38.9, I39.0, I39.1, I39.2, I39.3, I39.4, I39.5, I39.6, I39.7, I39.8, I39.9, I40.0, I40.1, I40.2, I40.3, I40.4, I40.5, I40.6, I40.7, I40.8, I40.9, I41.0, I41.1, I41.2, I41.3, I41.4, I41.5, I41.6, I41.7, I41.8, I41.9, I42.0, I42.1, I42.2, I42.3, I42.4, I42.5, I42.6, I42.7, I42.8, I42.9, I43.0, I43.1, I43.2, I43.3, I43.4, I43.5, I43.6, I43.7, I43.8, I43.9, I44.0, I44.1, I44.2, I44.3, I44.4, I44.5, I44.6, I44.7, I44.8, I44.9, I45.0, I45.1, I45.2, I45.3, I45.4, I45.5, I45.6, I45.7, I45.8, I45.9, I46.0, I46.1, I46.2, I46.3, I46.4, I46.5, I46.6, I46.7, I46.8, I46.9, I47.0, I47.1, I47.2, I47.3, I47.4, I47.5, I47.6, I47.7, I47.8, I47.9, I48.0, I48.1, I48.2, I48.3, I48.4, I48.5, I48.6, I48.7, I48.8, I48.9, I49.0, I49.1, I49.2, I49.3, I49.4, I49.5, I49.6, I49.7, I49.8, I49.9, I50.0, I50.1, I50.2, I50.3, I50.4, I50.5, I50.6, I50.7, I50.8, I50.9, I51.0, I51.1, I51.2, I51.3, I51.4, I51.5, I51.6, I51.7, I51.8, I51.9, I52.0, I52.1, I52.2, I52.3, I52.4, I52.5, I52.6, I52.7, I52.8, I52.9, I53.0, I53.1, I53.2, I53.3, I53.4, I53.5, I53.6, I53.7, I53.8, I53.9, I54.0, I54.1, I54.2, I54.3, I54.4, I54.5, I54.6, I54.7, I54.8, I54.9, I55.0, I55.1, I55.2, I55.3, I55.4, I55.5, I55.6, I55.7, I55.8, I55.9, I56.0, I56.1, I56.2, I56.3, I56.4, I56.5, I56.6, I56.7, I56.8, I56.9, I57.0, I57.1, I57.2, I57.3, I57.4, I57.5, I57.6, I57.7, I57.8, I57.9, I58.0, I58.1, I58.2, I58.3, I58.4, I58.5, I58.6, I58.7, I58.8, I58.9, I59.0, I59.1, I59.2, I59.3, I59.4, I59.5, I59.6, I59.7, I59.8, I59.9, I60.0, I60.1, I60.2, I60.3, I60.4, I60.5, I60.6, I60.7, I60.8, I60.9, I61.0, I61.1, I61.2, I61.3, I61.4, I61.5, I61.6, I61.7, I61.8, I61.9, I62.0, I62.1, I62.2, I62.3, I62.4, I62.5, I62.6, I62.7, I62.8, I62.9, I63.0, I63.1, I63.2, I63.3, I63.4, I63.5, I63.6, I63.7, I63.8, I63.9, I64.0, I64.1, I64.2, I64.3, I64.4, I64.5, I64.6, I64.7, I64.8, I64.9, I65.0, I65.1, I65.2, I65.3, I65.4, I65.5, I65.6, I65.7, I65.8, I65.9, I66.0, I66.1, I66.2, I66.3, I66.4, I66.5, I66.6, I66.7, I66.8, I66.9, I67.0, I67.1, I67.2, I67.3, I67.4, I67.5, I67.6, I67.7, I67.8, I67.9, I68.0, I68.1, I68.2, I68.3, I68.4, I68.5, I68.6, I68.7, I68.8, I68.9, I69.0, I69.1, I69.2, I69.3, I69.4, I69.5, I69.6, I69.7, I69.8, I69.9, I70.0, I70.1, I70.2, I70.3, I70.4, I70.5, I70.6, I70.7, I70.8, I70.9, I71.0, I71.1, I71.2, I71.3, I71.4, I71.5, I71.6, I71.7, I71.8, I71.9, I72.0, I72.1, I72.2, I72.3, I72.4, I72.5, I72.6, I72.7, I72.8, I72.9, I73.0, I73.1, I73.2, I73.3, I73.4, I73.5, I73.6, I73.7, I73.8, I73.9, I74.0, I74.1, I74.2, I74.3, I74.4, I74.5, I74.6, I74.7, I74.8, I74.9, I75.0, I75.1, I75.2, I75.3, I75.4, I75.5, I75.6, I75.7, I75.8, I75.9, I76.0, I76.1, I76.2, I76.3, I76.4, I76.5, I76.6, I76.7, I76.8, I76.9, I77.0, I77.1, I77.2, I77.3, I77.4, I77.5, I77.6, I77.7, I77.8, I77.9, I78.0, I78.1, I78.2, I78.3, I78.4, I78.5, I78.6, I78.7, I78.8, I78.9, I79.0, I79.1, I79.2, I79.3, I79.4, I79.5, I79.6, I79.7, I79.8, I79.9, I80.0, I80.1, I80.2, I80.3, I80.4, I80.5, I80.6, I80.7, I80.8, I80.9, I81.0, I81.1, I81.2, I81.3, I81.4, I81.5, I81.6, I81.7, I81.8, I81.9, I82.0, I82.1, I82.2, I82.3, I82.4, I82.5, I82.6, I82.7, I82.8, I82.9, I83.0, I83.1, I83.2, I83.3, I83.4, I83.5, I83.6, I83.7, I83.8, I83.9, I84.0, I84.1, I84.2, I84.3, I84.4, I84.5, I84.6, I84.7, I84.8, I84.9, I85.0, I85.1, I85.2, I85.3, I85.4, I85.5, I85.6, I85.7, I85.8, I85.9, I86.0, I86.1, I86.2, I86.3, I86.4, I86.5, I86.6, I86.7, I86.8, I86.9, I87.0, I87.1, I87.2, I87.3, I87.4, I87.5, I87.6, I87.7, I87.8, I87.9, I88.0, I88.1, I88.2, I88.3, I88.4, I88.5, I88.6, I88.7, I88.8, I88.9, I89.0, I89.1, I89.2, I89.3, I89.4, I89.5, I89.6, I89.7, I89.8, I89.9, I90.0, I90.1, I90.2, I90.3, I90.4, I90.5, I90.6, I90.7, I90.8, I90.9, I91.0, I91.1, I91.2, I91.3, I91.4, I91.5, I91.6, I91.7, I91.8, I91.9, I92.0, I92.1, I92.2, I92.3, I92.4, I92.5, I92.6, I92.7, I92.8, I92.9, I93.0, I93.1, I93.2, I93.3, I93.4, I93.5, I93.6, I93.7, I93.8, I93.9, I94.0, I94.1, I94.2, I94.3, I94.4, I94.5, I94.6, I94.7, I94.8, I94.9, I95.0, I95.1, I95.2, I95.3, I95.4, I95.5, I95.6, I95.7, I95.8, I95.9, I96.0, I96.1, I96.2, I96.3, I96.4, I96.5, I96.6, I96.7, I96.8, I96.9, I97.0, I97.1, I97.2, I97.3, I97.4, I97.5, I97.6, I97.7, I97.8, I97.9, I98.0, I98.1, I98.2, I98.3, I98.4, I98.5, I98.6, I98.7, I98.8, I98.9, I99.0, I99.1, I99.2, I99.3, I99.4, I99.5, I99.6, I99.7, I99.8, I99.9

Fonte: França et al., 2008.²⁷

3.2 Indicadores de qualidade dos dados criados pelo sistema: avaliação da qualidade das informações sobre causas de morte

Metodologias de avaliação da qualidade das informações sobre causas de morte nas estatísticas vitais para uso em análises de mortalidade e que sejam capazes de refletir as inter-relações entre as deficiências de cobertura/representatividade e a qualidade da informação sobre as causas registradas ainda representam um desafio. Recentemente, um modelo de avaliação dos registros de óbitos baseado em quatro atributos principais – representatividade, confiabilidade, validade e transcendência – foi proposto por Rao e colaboradores²⁸ como arcabouço teórico para a classificação de nove critérios de julgamento da qualidade das estatísticas de mortalidade por causas, já testado para o nível de grandes regiões do País²⁷. O modelo baseia-se em quatro critérios maiores (atributos) e nove indicadores:

- Qual a representatividade (cobertura, completude dos eventos)?
- Os dados sobre causas são consistentes (segundo nível da mortalidade geral, consistência temporal)?
- As causas de morte registradas são válidas (códigos maldefinidos, validade das causas declaradas pelos médicos, consistência segundo idade e sexo)?
- Os dados são utilizados (desagregação geográfica, oportunidade)?

A representatividade dos dados de mortalidade é avaliada de duas formas: em relação à cobertura (coverage) de toda a população pelo sistema e, na população coberta, se há registro de todos os óbitos ocorridos (completeness). No caso do Brasil, como o SIM é um sistema universal de coleta de dados em todos os municípios do País, tem importância avaliar o indicador (completeness) (traduzido como cobertura entre nós), ou seja, entre os óbitos ocorridos de fato na população, qual proporção foi registrada no SIM. Como os métodos de avaliação da cobertura já foram abordados nas seções anteriores, não nos deteremos novamente neles.

Para verificar a consistência, ou seja, a plausibilidade dos resultados, um procedimento simples se refere a avaliar a consistência temporal dos dados sobre causas. Para isso, compara-se a distribuição proporcional temporal de algumas causas de óbito, que se espera variarem pouco em curto espaço de tempo. Na tabela a seguir, apresenta-se a distribuição dos óbitos por diabetes em mulheres em três anos. O aumento da proporção observado em 2004 pode ser real ou devido a alterações na declaração ou codificação das causas de óbito.

Tabela 2.3: Número e proporção de óbitos definidos e por diabetes em mulheres do Nordeste brasileiro, 2002-2004.

Ano	Número de óbitos por causa definida	Número de óbitos por diabetes	% de óbitos por diabetes	Intervalo de confiança
2002	75789	5698	7,52	7,33 – 7,72
2003	78127	5859	7,50	7,31 – 7,69
2004	82131	6323	7,70	7,51 – 7,89

Fonte: Elaborado pelo autor.

As causas de morte registradas devem também ser analisadas para detectar possíveis inconsistências segundo idade e sexo, como por exemplo, óbitos por causas maternas em homens, câncer de próstata em mulheres, óbitos maternos em mulheres com mais de 55/60 anos etc. No Brasil, apesar da crítica dos dados ser feita já há algum tempo nos níveis municipal e estadual²⁹, além do federal, é interessante ter em conta esse critério, que pode ser um marcador de qualidade de dados a ser utilizado, por exemplo, pelo gestor do SIM municipal.

A validade das causas registradas no SIM com diagnóstico definido é em geral avaliada por meio da comparação das causas básicas de óbito da DO declaradas pelos médicos com a causa definida utilizando outra fonte de dados como referência. No Brasil, estudos de validação da causa do óbito utilizam em geral prontuários hospitalares, associados ou não a entrevistas com médicos, entrevistas com familiares e/ou dados de autópsia²⁷.

Em geral, os resultados dos estudos de validação de causas de óbito realizados no Brasil indicam melhor qualidade no caso das doenças crônicas. Por exemplo, em estudo de validação do câncer de útero como causa de óbito em Belém, em que Nunes e colaboradores em 2004 compararam as causas registradas em 188 DOs com as causas definidas após revisão de prontuários e/ou resultados anátomo-patológicos, foi verificada uma alta concordância, com índice Kappa de 0,87. Vale lembrar que a proporção de óbitos por causas mal definidas e inespecíficas é também um indicador da validade das causas de morte registradas, conforme visto anteriormente.

Para refletir

Você considera que as causas de óbito atestadas pelos médicos em seu estado são confiáveis? Você conhece estudos de validação da causa básica realizados nos últimos cinco anos?

O último critério refere-se à relevância política das informações sobre mortalidade para definição de políticas públicas e propostas de programas de Saúde. Os dois indicadores desse critério são a oportunidade, em que se considera que dados oportunos seriam dados disponibilizados em até dois anos após a ocorrência do óbito, e a desagregação geográfica dos dados disponíveis, suficiente para possibilitar a identificação de diferenciais na população para subsidiar políticas públicas. Em relação a esses critérios, o SIM tem sido classificado como plenamente adequado²⁷. O critério de relevância política pode ser considerado uma medida da acessibilidade do sistema, pois avalia a disponibilidade dos dados segundo o tipo de informação disponível (tipo de desagregação dos dados) e segundo a oportunidade, ou seja, o período de

referência estabelecido para a disponibilização dos dados ao usuário.

Em relação a esse último critério, merece destaque o programa Tabwin de processamento de dados e análise desenvolvido para o setor público de Saúde pelo DATASUS. Além de permitir tabular dados disponíveis na internet, possibilita utilizar arquivos de outras bases de dados (DBF, CNV e TXT) e realizar análises desagregadas de um grande número de variáveis com aproveitamento de informações específicas locais.

3.3 Indicadores de qualidade dos dados criados pelo sistema: avaliação das informações sobre mortalidade infantil em municípios

Em avaliação nacional encomendada pelo Ministério da Saúde sobre a qualidade do SIM e Sinasc, Szwarcwald 2002 e colaboradores²¹ classificaram os municípios em 1996-1998 segundo quatro indicadores:

- Taxa geral de mortalidade padronizada (TGMP).
- Desvio médio da taxa geral de mortalidade (DTGM).
- Desvio médio da taxa de natalidade (DTN).
- Proporção de causas mal definidas entre os óbitos infantis (%MD).

A taxa geral de mortalidade padronizada (TGMP) representa a TGM (ou TBM) padronizada por idade pelo método direto para controlar o efeito das diferentes distribuições etárias dos municípios. Conforme visto anteriormente, o primeiro passo da padronização é a seleção da população-padrão. No estudo, foi escolhida a população do Rio de Janeiro por ser a mais idosa do Brasil. Em seguida, as taxas de mortalidade específicas por idade dos municípios foram multiplicadas as da população-padrão do grupo etário, obtendo-se o número de óbitos esperados. A soma do total de óbitos esperados dividida pela população total padrão representa a TGMP. No estudo, TGMPs menores do que 4,0 por 1.000 habitantes foram consideradas como não adequadas.

A análise do desvio da taxa geral de mortalidade (DTGM) baseou-se na premissa de que, em geral, as TGMs não variam muito de um ano para o outro. De forma semelhante ao indicador DTGM, espera-se que não haja grandes alterações na taxa de natalidade (calculada pela razão entre o número de nascidos vivos e a população total em determinado ano), em um período de três anos. A proporção de óbitos infantis classificados como sintomas, sinais e achados anormais de exames clínicos e de laboratório não classificados em outra parte (CID-10: Roo-R99), ou seja, óbitos infantis em que não houve definição da causa básica de óbito, indica irregularidades do registro das informações de óbitos.

Os quatro indicadores foram utilizados para a classificação de cada um dos municípios em três categorias:

- Categoria 1 ou Adequada: satisfaz todos os critérios (TGMP, DTGM, DTN, %MD);
- Categoria 2 ou Intermediária: adequada para TGMP e não adequada para pelo menos um dos outros;
- Categoria 3 ou Inadequada: TGMP < 4,0/1.000.

Posteriormente, os valores-limite dos indicadores foram revistos²⁶ com base nas suas distribuições nos municípios das oito UFs com informações vitais consideradas adequadas segundo a Ripsa, ver no IDB em <www.datasus.gov.br>. Foram também considerados pontos de corte diferenciados segundo a categoria populacional do município (menos de 50 mil habitantes; 50 mil habitantes ou mais).

Em estudo recente foi utilizada a taxa de mortalidade padronizada por idade (TGMP) para avaliar a cobertura do SIM, utilizando-se a população do Brasil em 2008 como população-padrão³⁰. A distribuição das TGMPs nos municípios das oito UFs com estatísticas vitais adequadas, segundo a Ripsa, foi utilizada para classificar a adequação das informações de óbitos nos demais municípios (pontos de corte para insuficiente e insatisfatório definidos pelos percentis 5% e 20%, respectivamente).

Para refletir

Se Szwarcwald e colaboradores²⁵ realizaram a pesquisa de busca ativa na Amazônia Legal e no Nordeste, como puderam inferir os resultados para cálculo das coberturas de óbitos nos demais estados?

3.4 Indicadores de monitoramento do funcionamento do SIM

A Portaria nº 116 do Ministério da Saúde de 11 de fevereiro de 2009 regulamenta a coleta de dados, fluxo e periodicidade de envio das informações sobre óbitos e nascidos vivos para os SIS sob gestão da Secretaria de Vigilância em Saúde. Segundo essa portaria, são unidades notificadoras dos óbitos:

- 1) os estabelecimentos e serviços de Saúde, inclusive os de atendimento ou internação domiciliar;
- 2) os institutos médicos legais (IMLs);
- 3) os serviços de verificação de óbitos (SVOs) e
- 4) os médicos cadastrados pelas secretarias municipais de Saúde. No caso dos cartórios

de Registro Civil, é permitida a distribuição de formulários de DO somente em localidades onde não exista médico³¹.

A DO é impressa como um conjunto de três vias coloridas autocopiativas em uma sequência numérica única. As primeiras vias das DOs devem ser coletadas pela secretaria municipal de saúde (SMS) no local onde foram preenchidas, ou seja, nas unidades notificadoras como hospitais, outros estabelecimentos de Saúde, cartórios e médicos cadastrados que tenham prestado o atendimento fora dos estabelecimentos de Saúde, além dos IMLs e SVOs. Essa via deve ficar arquivada no setor da SMS responsável pelo processamento dos dados. A segunda via deverá ser entregue à família para que seja realizado o registro do óbito no Cartório de Registro Civil, onde ficará retida. A terceira via deve permanecer arquivada nas unidades notificadoras. Segundo a Portaria nº 116/2009, a primeira via da DO destina-se ao processamento de dados na secretaria de Saúde do município. No caso de óbito natural ocorrido em localidade sem médico, a DO preenchida pelo Cartório do Registro Civil deverá ser coletada posteriormente pela SMS^{17,31}.

O modelo de operacionalização do SIM é formado por seis componentes²⁸:

- notificação do óbito, por meio do preenchimento oportuno e emissão da DO, de responsabilidade do médico;
- coleta das DOs nas fontes notificadoras, avaliação preliminar das variáveis e processamento dos dados pelas SMS;
- análise de indicadores de cobertura e qualidade;
- busca ativa em fontes alternativas como cemitérios, Programa Saúde da Família (PSF) e informantes-chave;
- fluxo entre as esferas de gestão do SIM (municipal, estadual e federal) e
- divulgação das informações por meio de boletins e disponibilização dos dados na internet pelo DATASUS.

O acompanhamento do processo de coleta de dados do SIM é de importância fundamental, pois a coleta de dados inconsistentes leva a análises com viés, de pouca utilidade para a avaliação da situação de saúde. Já há algum tempo, pesquisadores têm chamado a atenção sobre a importância de “mecanismos gerenciais que promovam a normatização do sistema e estabeleçam instrumentos de controle de qualidade [...] principalmente nos níveis de gestão estadual, regional e municipal”²⁹ e as secretarias de Saúde têm procurado atuar nesse sentido. Sabe-se que um procedimento importante no momento da coleta das DOs nas fontes notificadoras é a crítica cuidadosa do seu preenchimento para identificar incorreções e/ou omissões de variáveis. A identificação de falhas

no preenchimento das DOs pode apontar quais os profissionais/estabelecimentos necessitam de treinamentos, além de permitir corrigir/complementar a informação registrada³³.

Além do pareamento com outras bases de dados como o SIH proposto pelo Ministério da Saúde para detectar óbitos não notificados, alguns indicadores específicos têm sido propostos por pesquisadores para avaliação do desempenho do SIM. Romero & Cunha³⁴, em estudo sobre a qualidade das variáveis socioeconômicas e demográficas das DOs de óbitos de crianças menores de 1 ano registrados no SIM em 2001–2006, utilizam como um dos indicadores de avaliação a *clareza metodológica*, que definem como instruções de coleta adequadas, manuais de preenchimento e documentação da base de dados com conceitos e definições precisos, assim como a comparabilidade das variáveis entre os diferentes documentos metodológicos do sistema e com outras fontes de informação.

O critério *incompletude* (ou não completude), ou seja, a proporção de informação ignorada (campos em branco e códigos atribuídos à informação ignorada) foi também um critério utilizado pelas autoras acima citadas. Os seguintes graus de avaliação foram definidos para avaliar a incompletude: excelente (menor de 5%), bom (5% a 10%), regular (10% a 20%), ruim (20% a 50%) e muito ruim (50% ou mais). É importante avaliar as proporções de informações ignoradas para todas as variáveis existentes no sistema, pois o achado de persistentes proporções altas indica problemas de qualidade do dado que devem ser investigados. Conforme mencionado anteriormente, variáveis com alta proporção de valores ignorados podem levar a análises com viés, de pouca utilidade para a avaliação da situação de saúde.

3.5 Propostas de intervenção para melhoria de qualidade da informação do SIM: busca ativa de óbitos e a proposta da Autópsia Verbal

O Ministério da Saúde tem realizado investimento específico importante para a melhoria das informações vitais³⁵, destacando-se, entre as várias iniciativas propostas, o estímulo à busca ativa de óbitos e à investigação dos óbitos por causas mal definidas, além da inclusão de metas relacionadas à qualidade das informações sobre óbitos na Programação Pactuada e Integrada (PPI) e implantação dos comitês de morte materna e comitês de prevenção do óbito infantil/fetal nos últimos anos. Inúmeros projetos específicos ou locais para captação de óbitos não notificados ao SIM, por busca ativa em fontes alternativas, têm sido também desenvolvidos pelos serviços de Saúde municipais e pesquisadores da academia e apresentados em publicações, congressos nacionais ou eventos sobre experiências dos serviços de Saúde como a ExpoEpi. Vale ressaltar que essa importante iniciativa do MS, desde seu início em 2001, tem como um dos blocos temáticos os sistemas de informação e sua qualificação³⁶.

A busca ativa é considerada oficialmente como um dos componentes do modelo de operacionalização do SIM, sendo especificado pela Portaria nº 116 do Ministério da Saúde, de 11 de fevereiro de 2009, em que “as Secretarias Municipais de Saúde deverão utilizar-se dos meios disponíveis na busca ativa de casos não notificados, valendo-se de todos os meios disponíveis para esta finalidade”. Vários estudos publicados sobre o tema destacaram a importância dos agentes comunitários de Saúde do Programa de Saúde da Família (PSF), estabelecimentos de Saúde, cemitérios e outros locais como fontes potencialmente importantes para realização de busca ativa de óbitos^{32,37}. Chama a atenção o achado recente em Minas Gerais em que se verificou que cerca de 50% dos óbitos não notificados ao SIM nos municípios investigados tiveram as DOs localizadas em cartórios (arquivos de DO ou livro de registro dos cartórios), o que aponta para uma falha na busca ativa pelos profissionais de Saúde dos municípios³⁸, reflexo de possível pouca importância dada ao SIM municipal³⁹.

Em relação às causas mal definidas, o “Programa de redução do percentual de óbitos com causa mal definida”, desenvolvido pela Secretaria de Vigilância em Saúde do MS principalmente nas Regiões Norte e Nordeste do País, conseguiu resultados promissores em curto prazo. Alguns estados em que a proporção de óbitos por causas mal definidas ultrapassavam 20% em 1997³⁶, apresentaram, em 2006, percentagens menores que 10%.

No caso de óbitos domiciliares sem assistência médica, desde 2008 o MS tem proposto a utilização do método da autópsia verbal (AV) para a definição das causas de morte. Os formulários da AV da OMS foram adaptados para as causas mais prevalentes no Brasil e sua análise permite ao médico identificar a sequência de eventos que levou à morte e assim definir a causa básica do óbito⁴⁰. Os formulários específicos da autópsia verbal (AV-1, AV-2 e AV-3) estão disponíveis no *site* do Ministério da Saúde <www.saude.gov.br> em “vigilância em saúde”, clicar em “publicações”, “informações e análises”, “óbitos” e clicar em “manual de investigação do óbito por causa mal definida”.

Para refletir

Existe relação temporal entre a qualidade das estatísticas de mortalidade sobre causas e a proporção de óbitos por causas mal definidas investigadas em seu município?

Referências

- 1 – Becker RA. (1991) Análise de mortalidade: delineamentos básicos, Brasília. Ministério da Saúde, 85.
- 2 – Moraes IHS. (1994) Informações em saúde: da prática fragmentada ao exercício da cidadania. São Paulo. Hucitec.
- 3 – Carvalho AO, Eduardo MBP. (1998) Sistemas de informação em saúde para municípios, Série Saúde & Cidadania São Paulo. Faculdade de Saúde Pública da Universidade de São Paulo.
- 4 – Souza M de. (2011) Sistema de información de salud: conceptos básicos e implicaciones para las políticas de salud, International Standards for Civil Registration and Vital Statistics Systems. New York.
- 5 – White KL. (1980) Information for health care: an epidemiological perspective, *Inquiry*, 17(4), 296-312.
- 6 – Health Metrics Network (2008) Framework and standards for country health information systems. Disponível em: <http://www.who.int/healthmetrics/about/en/> (acessado).
- 7 – Paim JS, Almeida-Filho N. (1998) Saúde coletiva: uma “nova saúde pública” ou campo aberto a novos paradigmas?, *Rev Saude Publica*, 32(4), 299-316.
- 8 – Viacava F. (2004) Uma metodologia de avaliação do desempenho do sistema de saúde brasileiro, *Ciência & Saúde Coletiva*, 9(3), 711-724.
- 9 – Bittencourt SA, Camacho LA, Leal MC. (2006) O Sistema de Informacao Hospitalar e sua aplicação na saúde coletiva, *Cad Saude Publica*, 22(1), 19-30.
- 10 – Drumond, E., Machado, C., Vasconcelos, M. & França, E. (2009) Utilização de dados secundários do SIM, Sinasc e SIH na produção científica brasileira de 1990 a 2006, *R bras Est Pop*, 26(1), 7-19.

- 11 – Ministério da Saúde (Brasil) (2008) Indicadores básicos para a saúde no Brasil: conceitos e aplicações. Brasília. Rede Interagencial de Informação para a Saúde. Ministério da Saúde.
- 12 – Viacava F. (2012) Avaliação de desempenho de sistemas de saúde: um modelo de análise, *Ciência & Saúde Coletiva*, 17(4), 921-934.
- 13 – Mello-Jorge, MHP, Laurenti, R, Gotlieb, SL. (2007) Análise da qualidade das estatísticas vitais brasileiras: a experiência de implantação do SIM e do SINASC, *Ciência & Saúde Coletiva*, 12(3), 643-654.
- 14 – Guerra FA, Llerena JC Jr, Gama, SG, Cunha CB, Theme Filha MM. (2008) Defeitos congênitos no Município do Rio de Janeiro, Brasil: uma avaliação através do SINASC (2000-2004), *Cad Saude Publica*, 24(1), 140-149.
- 15 – UN (2000) United Nations Millennium Declaration. Disponível em: <http://www.un.org/millennium/declaration/ares552e.htm> (acessado em 22 Nov 2012).
- 16 – Laurenti R, Mello-Jorge MHP, Lebrão ML, Gotlieb SLD (2005) Estatísticas de saúde. São Paulo. EPU.
- 17 – Ministério da Saúde (Brasil) (2011) Manual de instruções para o preenchimento da declaração de óbito, Série A. Normas e Manuais Técnicos Brasília. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde, 44.
- 18 – Risi Júnior JB. (2006) Informação em saúde no Brasil: a contribuição da RIPSAs, *Ciência & Saúde Coletiva*, 11(4), 1049-1053
- 19 – Ministério da Saúde (Brasil) (2009) Conselho Federal de Medicina, Centro Brasileiro de Classificação de Doenças. A declaração de óbito: documento necessário e importante, 3. ed.; Brasília: Ministério da Saúde.
- 20 – Laurenti R, (1991) Análise da informação em Saúde:1893–1993, 100 anos da Classificação Internacional de Doenças. *Rev. Saúde públ.* 1991, 25 (6):407-17.

- 21 – Szwarcwald CL, Leal MC, Andrade CL, Souza PRJ. (2002) Estimação da mortalidade infantil no Brasil: o que dizem as informações sobre óbitos e nascimentos do Ministério da Saúde?, *Cad Saude Publica*, 18(6), 1725-1736.
- 22 – AbouZahr C, Mikkelsen L, Rampatige R, Lopez A. Mortality statistics: a tool to enhance understanding and improve quality. Disponível em: [www. uq.edu.au/hishub/wp13](http://www.uq.edu.au/hishub/wp13).
- 23 – Agostinho, C.S.; Queiroz, B.L. Estimativas da mortalidade adulta para o Brasil no período 1980/2000: uma abordagem metodológica comparativa. In: XVI Encontro Nacional de Estudo de População da ABEP, 2008, Caxambu, Anais do Encontro Nacional de Estudos de População, 2008. p.1-20.
- 24 – Porapakkam Y, Rao C, Pattaraarchachai J et al (2010). Estimated causes-of-death in Thailand, 2005: implications for health policy. *Population Health Metrics* 8:14.
- 25 – Murray CJL, Lopez AD. (1996) Estimating causes of death: new methods and global and regional applications for 1990, in: C. J. L. Christopher J L Murray & A. D. Lopez (Eds) *The global burden of disease: comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Boston. Harvard School of Public Health, 118-200.
- 26 – Mathers CD, Fat DM, Inoue M, Rao C, Lopez, AD. (2005) Counting the dead and what they died from: an assessment of the global status of cause of death data, *Bull World Health Organ*, 83(3), 171-177.
- 27 – França E, Abreu DX, Rao C, Lopez AD. (2008) Evaluation of cause-of-death statistics for Brazil, 2002-2004, *Int J Epidemiol*, 37(4), 891-901.
- 28 – Rao C, Lopez AD, Yang G, Begg S, Ma J. (2005) Evaluating national cause-of-death statistics: principles and application to the case of China, *Bull World Health Organ*, 83(8), 618-625.
- 29 – Almeida M, Alencar, G. (2000) Informações em saúde: necessidade de introdução de mecanismos de gerenciamento dos sistemas, *Inf. Epidemiol. SUS*, 9(4), 241-249.

30 – Szwarcwald CL, Morais Neto OL, Frias PG, Souza Junior PRB, Escalante JJC, Lima RB, Viola RC. Busca ativa de óbitos e nascimentos no Nordeste e Amazônia Legal: Estimação das coberturas do SIM e do Sinasc nos municípios brasileiros. In: Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise da Situação de Saúde. Saúde Brasil 2010. Brasília: Ministério da Saúde, 2011. p. 79-98.

31 – Ministério da Saúde (Brasil) (2009) Portaria nº 116. Regulamenta a coleta de dados, fluxo e periodicidade de envio das informações sobre óbitos e nascidos vivos para os Sistemas de Informações em Saúde sob gestão da Secretaria de Vigilância em Saúde, DOU, 11 de fevereiro de 2009.

32 – Frias, P. G., Pereira, P. M., Andrade, C. L. & Szwarcwald, C. L. (2008) Sistema de informações sobre mortalidade: estudo de caso em municípios com precariedade dos dados, Cad Saude Publica, 24(10), 2257-2266.

33 – SES-SC (2004) Rotinas de gerenciamento do sistema de informações sobre mortalidade, Florianópolis. Santa Catarina. Secretaria de Estado da Saúde. Divisão de Informações. Sistema de Informações sobre Mortalidade.

34 – Romero DE, Cunha CB. (2007) Avaliação da qualidade das variáveis epidemiológicas e demográficas do sistema de informações sobre nascidos vivos, 2002, Cad Saude Publica, 23(3), 701-714.

35 – Ministério da Saúde (2003) Monitoramento da acurácia dos sistemas de informações sobre mortalidade e nascidos vivos, 3ª Expoepi Mostra Nacional de Experiências Bem-sucedidas em Epidemiologia, Prevenção e Controle de Doenças. Salvador, BA. Ministério da Saúde, 173-180.

36 – Ministério da Saúde (Brasil) (2001) 1ª Expoepi Mostra Nacional de Experiências Bem-sucedidas em Epidemiologia, Prevenção e Controle de Doenças. Brasília.

37 – Façanha, M., Pinheiro, A. & Fauth, S. (2003) Busca ativa de óbitos em cemitérios da Região Metropolitana de Fortaleza, 1999 a 2000, Epidemiol Serv Saude, 112(3), 131-136.

38 – Cunha CC, Campos D, França E. (2011) Uso da busca ativa de óbitos na avaliação do Sistema de Informações sobre mortalidade em Minas Gerais, Brasil Epidemiol Serv Saúde, 20(3), 275-286.

39 – Campos D, Hadad SC, Abreu C, Franca EB . Sistema de Informações sobre Mortalidade em municípios de pequeno porte de Minas Gerais: concepções dos profissionais de saúde. Ciênc. saúde coletiva, Maio 2013, vol.18, no.5, p.1473-1482. ISSN 1413-8123.

40 – Ministério da Saúde (Brasil) (2009) Manual para investigação do óbito com causa mal definida, Série A. Normas e Manuais Técnicos. Brasília. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde, 48.

Módulo 3

Análise de Dados Demográficos

Marta Roverly de Souza

Universidade Federal de Goiás (UFG), Instituto de Patologia Tropical e Saúde Pública, Goiânia/GO, Brasil.

Introdução

As informações sócio demográficas são de suma importância no processo de planejamento e tomada de decisão. O efetivo conhecimento da interação entre as tendências de crescimento, migração e estrutura etária da população é determinante para o planejamento da distribuição de bens e serviços e atendimento das demandas sociais. A compreensão das tendências de distribuição da população por idade e sexo, por exemplo, é de fundamental importância no processo de delineamento e elaboração de políticas voltadas para o atendimento das demandas sociais. Tal importância decorre do fato de que cada política social tem um público-alvo, diferenciado em termos de volume, ritmo de crescimento, composição e distribuição espacial pelo território. Essa configuração demográfica dos públicos-alvos está estreitamente relacionada ao comportamento das necessidades e requerimento de serviços dos indivíduos ao longo do ciclo de vida de indivíduos e famílias.

Diante deste contexto, as informações e análises demográficas são imprescindíveis no processo de tomada de decisão. Por outro lado, observa-se que é ainda muito incipiente o conhecimento/uso de tal potencialidade por parte do setor privado e até mesmo do setor público, o Instituto Brasileiro de Geografia e Estatística (IBGE), responsável pelas estatísticas oficiais de população têm enfrentado grandes desafios para atender as mais diversas demandas por infor-



mações populacionais detalhadas e prospectivas. Vale destacar que estas estatísticas têm revelado mudanças nos padrões de sexo e idade oriundos da queda da fecundidade e da mortalidade e chamado a atenção dos estudiosos da área da saúde para uma análise mais robusta sobre um novo perfil epidemiológico que vem se apresentando influenciado por este novo padrão populacional brasileiro. Conhecer este novo padrão populacional, as estatísticas populacionais, as respectivas projeções e cenários populacionais constituem-se em uma ferramenta fundamental na elaboração e adequação de políticas públicas eficazes e resolutivas.

Neste sentido o Módulo Análise de Dados Demográficos têm como objetivos: Apresentar e estabelecer as relações entre os componentes da dinâmica demográfica brasileira e sensibilizar profissionais e gestores de saúde para a percepção de que as informações demográficas devem ser consideradas no contexto da Saúde Pública.

Para facilitar a apreensão dos conteúdos fundamentais optou-se por dividi-los em três unidades sendo o primeiro intitulado **Apresentando os Conceitos e Componentes da Dinâmica Demográfica**, o segundo **Fontes de Dados Demográficos** e o terceiro **Tendências Populacionais e os Desafios para o Século XXI**.

Unidade 1

Apresentando os Conceitos e Componentes da Dinâmica Demográfica

Demografia (dêmos=população, *graphein*=estudo) refere-se ao estudo das populações humanas e sua evolução temporal. Em um estudo demográfico as principais variáveis que devem ser consideradas são: **o tamanho da população; sua distribuição por sexo, idade, estado conjugal; natalidade, fecundidade, mortalidade e migrações (distribuição da população segundo região geográfica de residência atual, anterior e de nascimento).**

Ao caracterizar uma população humana, pensamos inicialmente no seu tamanho:

Quantas pessoas existem numa localidade, num determinado momento?

Pensamos também sobre a sua composição: Quantas pessoas maiores de 50 anos existem? Quantas são do sexo feminino? Quantas são economicamente ativas?

Outro aspecto que vem a baila seriam os elementos que afetam o tamanho da população:

Quantas pessoas nascem num determinado período, quantas morrem, quantas migram?

Quando refletimos sobre cada um destes elementos, outras questões surgem imediatamente. Por exemplo, no que se refere a nascimentos: Quantas são as mulheres em idade reprodutiva? Quantas estão casadas? Que proporção usa métodos anticoncepcionais efetivos? etc.

Uma questão importante que surge seria: Como é que as mudanças em um ou mais destes componentes poderiam afetar os demais?¹

1. Tamanho da população

Para entender o que seja o tamanho de uma população, imaginemos os moradores de uma determinada área geográfica, em um determinado momento. Suponhamos também que, a partir de uma população inicial, não tenha havido entrada e saída de pessoas da área. Trata-se de uma população fechada, isto é, sem movimentos migratórios. A trajetória entre aquela população inicial e a população atual é totalmente explicada pelas mortes e nascimentos ocorridos no período, assim o tamanho da população em qualquer momento desse período pode ser representado pela seguinte equação:

$$Q_n = Q_0 + N_t - O_t$$

onde: Q_n = população num instante n ;
 Q_0 = população inicial, instante 0;
 N_t = nascimentos no período t ($t = n - 0$);
 O_t = mortes no período t ($t = n - 0$).

Conhecer o número total de pessoas residentes e a sua estrutura relativa pode nos ajudar a dimensionar a população-alvo de ações e serviços; a contribuir para o planejamento, gestão e avaliação de políticas públicas relacionadas à saúde, educação, trabalho, previdência e assistência Social para os diversos segmentos de idade e a subsidiar as discussões sobre alocação de recursos.

A população brasileira, segundo o Censo Demográfico 2010, atingiu o total de 190.755.799 habitantes na data de referência. A população urbana com 160.925.792 habitantes foi predominante, representando 84,4% da população total, enquanto 29.830.007 habitantes residiam em áreas rurais (Tabela 3.1).

Tabela 3.1: População residente e participação relativa, por situação do domicílio – Brasil, 1950–2010.

Data	População residente			População relativa %		
	Total	Urbana	Rural	Total	Urbana	Rural
01.07.1950¹	51.944.397	18.782.891	33.161.506	100,0	36,2	63,8
01.09.1960	70.070.457	31.303.034	38.767.423	100,0	44,7	55,3
01.09.1970	93.139.037	52.084.984	41.054.053	100,0	55,9	44,1
01.09.1980	119.002.706	80.436.409	38.566.297	100,0	67,6	32,4
01.09.1991	146.825.475	110.990.990	35.834.485	100,0	75,6	24,4
01.08.2000	169.799.170	137.953.959	31.845.211	100,0	81,2	18,8
01.08.2010	190.755.799	160.925.792	29.830.007	100,0	84,4	15,6

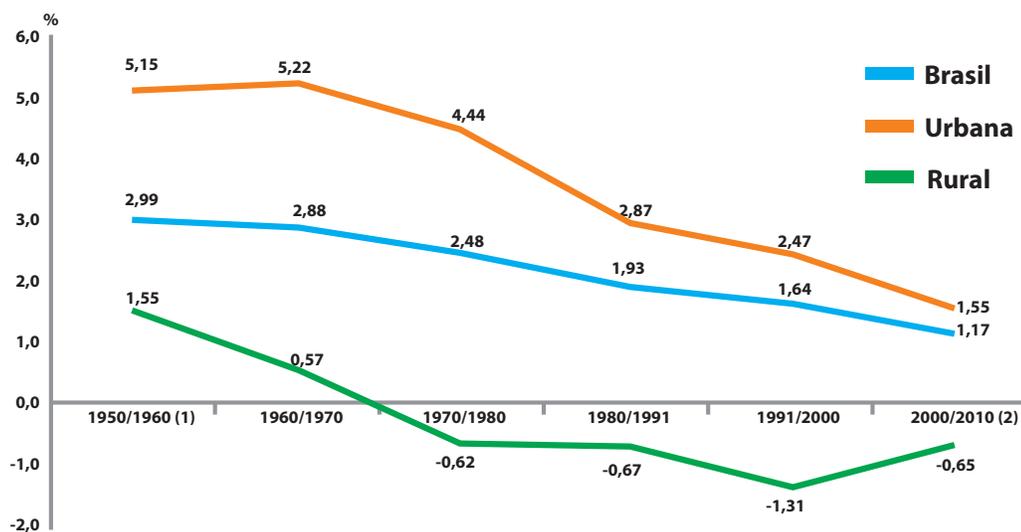
1 – Utilizada a população presente em 1950, enquanto para os anos seguintes foi utilizada a população residente.

Fonte: IBGE, Censo Demográfico 1950/2010.

Desde os anos 1960 que a taxa de crescimento da população brasileira – *percentual de incremento médio anual da população residente em determinado espaço geográfico, no período considerado, o valor da taxa refere-se à média anual obtida para um período de anos compreendido entre dois momentos, em geral correspondentes aos censos demográficos* – vem experimentando contínuos declínios. O último período 2000-2010 é uma continuidade da tendência observada a partir da década de 1960, com a população brasileira passando a crescer a um ritmo menos acentuado, 1,17% ao ano, aumentando o tempo estimado em que o volume populacional se duplicaria, 59,6 anos.

As taxas de crescimento das áreas urbana e rural declinam em relação ao período 1991–2000, 37,1% e 50,4%, respectivamente. A área urbana passa a crescer a uma taxa de 1,55% e a rural diminui o ritmo de perda para 0,65% ao ano, acarretando diminuição do número de habitantes residindo em áreas rurais (Figura 3.1).

Figura 3.1: Taxa média geométrica de crescimento anual da população residente, por situação do domicílio – Brasil, 1950–2010.



1 – Para o cálculo da taxa foi utilizada a população presente em 1950, enquanto para os anos seguintes foi utilizada a população residente.

2 – Para a obtenção da taxa do período 2000/2010 foram utilizadas as populações residentes em 2000 e 2010, sendo que, para este último ano, foi incluída a população estimada de 2,8 milhões de habitantes para os domicílios fechados.

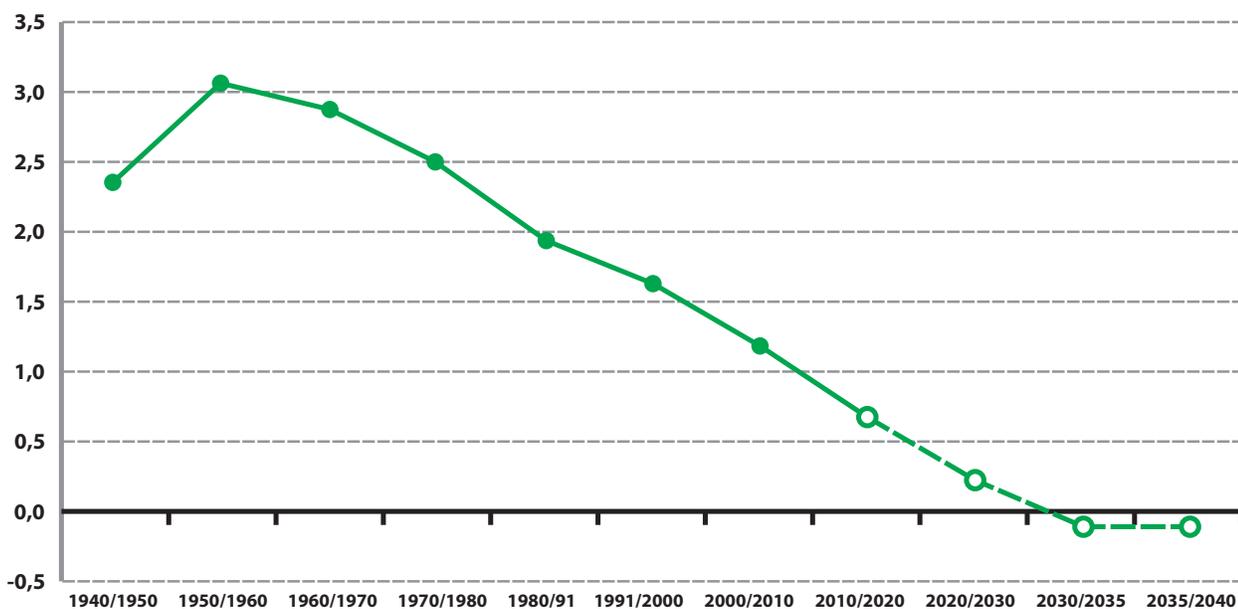
Fonte: IBGE, Censo Demográfico 1950/2010.

As estimativas de crescimento da população são realizadas pelo método geométrico. Em termos técnicos, para se obter a taxa de crescimento (r), subtrai-se 1 da raiz enésima do quociente entre a população final (Pt) e a população no começo do período considerado (P0), multiplicando-se o resultado por 100, sendo “n” igual ao número de anos no período.

$$r = \left[\left(\sqrt[n]{\frac{P_t}{P_0}} \right) - 1 \right] \times 100$$

As projeções indicam que em 2040, a taxa de crescimento cairá para - 0,291%, o que representa uma população de 215,3 milhões de habitantes. O País apresentará um potencial de crescimento populacional até 2030, quando se espera que a população atinja o chamado “crescimento zero”. A partir desse ano serão registradas taxas de crescimento negativas, que correspondem à queda no número da população (Figura 3.2).

Figura 3.2: Taxa de crescimento da população brasileira, 1940 a 2040.



Fonte: IBGE / Censo Demográficos. Elaborado pela autora.

2. Distribuição por sexo e idade

A distribuição por sexo e idade de uma população é, na realidade, reflexo da história da sua dinâmica populacional, desde um passado relativamente longínquo. O número de pessoas de uma população, em uma determinada idade, é o resultante do número de nascimentos que ocorreram anos atrás e dos níveis de mortalidade aos quais estes indivíduos estiveram sujeitos desde que nasceram.

Vale destacar que para expressar a relação quantitativa entre os sexos, bem como para analisar variações geográficas e temporais na distribuição da população por sexo, utiliza-se um indicador demográfico chamado Razão de sexos – se ela for igual a 100, o número de homens e de mulheres equivalem-se; acima de 100 há predominância de homens e, abaixo, predominância de mulheres. Observa-se também que este indicador sofre influência das taxas de migração e de mortalidade diferenciadas por sexo e idade.

$$\frac{\text{Número de residentes sexo masculino}}{\text{Número de residentes sexo feminino}} \times 100$$

Apesar de nascerem mais crianças do sexo masculino do que do feminino, na população como um todo, tem-se mais mulheres que homens, em virtude dos diferenciais de mortalidade existentes entre os sexos. A mortalidade masculina é superior à feminina ao longo de toda a vida.

Nesse sentido, os resultados do cálculo deste indicador revelam que a maioria da população brasileira é composta de pessoas do sexo feminino, reflexo da sobremortalidade masculina, especialmente nas faixas etárias jovens e adultas, decorrentes da alta incidência de óbitos por causas violentas (Tabela 3.2).

Tabela 3.2: Razão de sexos – Brasil, 1991, 1996, 2000 e 2005.

Ano	1991	1996	2000	2005
Razão de sexos	97,5	97,3	96,9	96,6

Fonte: IBGE, Censo Demográfico 1991 e 2000, Contagem de população 1996 e Estimativas Demográficas 2005.

Uma forma bastante ilustrativa de representar a estrutura da população por idade e sexo é por meio da pirâmide etária (ela pode ser confeccionada segundo região, município, áreas rural e urbana). O **eixo horizontal** de uma pirâmide etária representa o número absoluto ou a proporção da população, enquanto o **eixo vertical** representa os grupos de idade. O lado direito do eixo horizontal é destinado à representação do contingente ou à proporção de mulheres e o esquerdo, à dos homens.

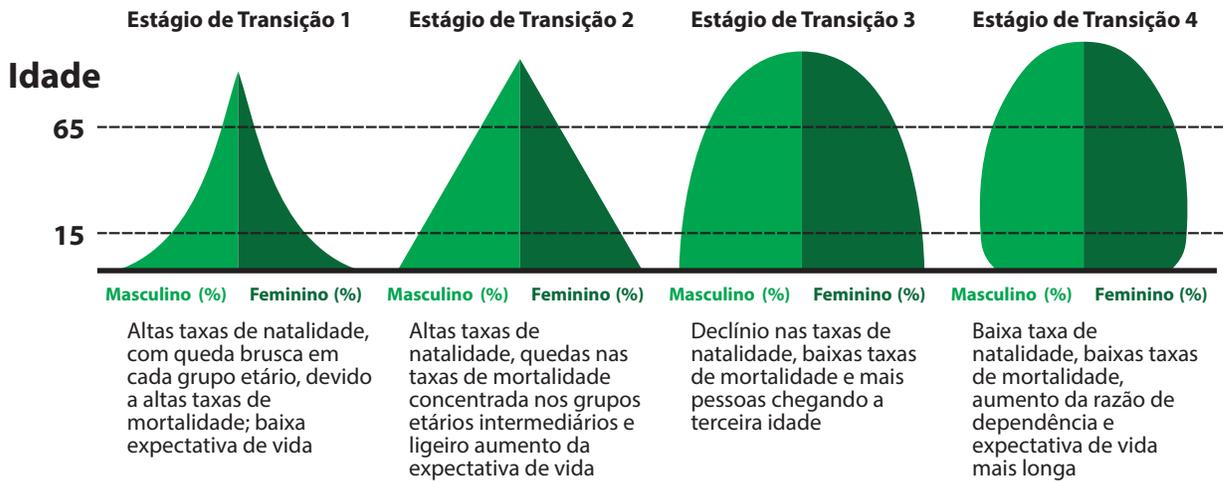
Nota

As pirâmides etárias são usadas, não só para monitorar a estrutura de sexo e idade, mas como um complemento aos estudos da qualidade de vida, já que podemos visualizar a média do tempo de vida, a taxa de mortalidade e a regularidade, ou não, da população ao longo do tempo. Quanto mais alta a pirâmide, maior a expectativa de vida e, conseqüentemente, melhor as condições de vida daquela população. É possível perceber que quanto mais desenvolvido econômica e socialmente é o país, mais sua pirâmide terá uma forma retangular.

Quando a pirâmide populacional tem base larga e ápice estreito, a pirâmide retrata uma população bastante jovem. Na medida em que a fecundidade declina – menos crianças nascem – a base da pirâmide vai se estreitando, com tendência a forma retangular, passando a caracterizar uma população envelhecida.

As análises de pirâmides populacionais também são fundamentais para entender em que estágio de transição demográfica se encontra determinado país (Figura 3.3). Alguns fatores interferem significativamente na pirâmide etária: uma guerra, por exemplo, provoca distúrbios visíveis sendo o mais comum a queda no número de jovens e adultos do sexo masculino. Normalmente, após uma crise como essa, é notável uma reposição populacional estimulada pelo governo, chamada de *baby boom*. Outros fatores podem afetar o formato da pirâmide: um controle populacional explícito, um êxodo, um movimento populacional importante, uma epidemia, uma catástrofe ambiental, entre outros.

Figura 3.3: Estágios da transição demográfica.

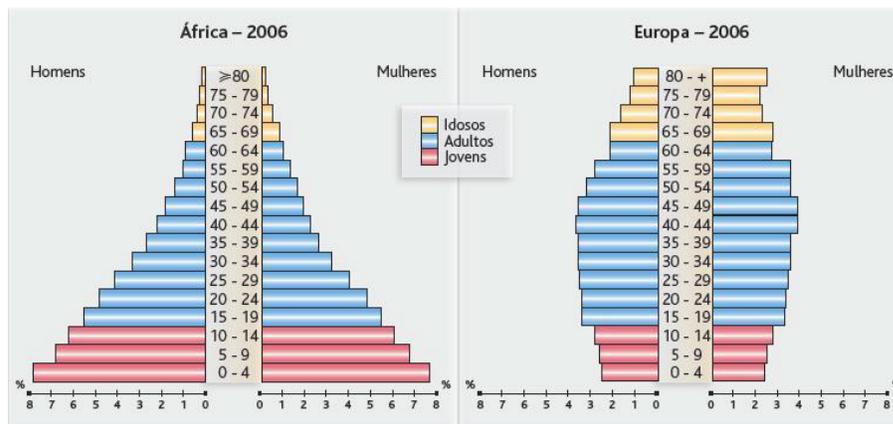


Nesse tipo de gráfico, cada uma das metades representa um sexo; a base representa o grupo jovem (até 19 anos); a área intermediária ou corpo representa o grupo adulto (entre 20 e 59 anos); e o topo ou ápice representa a população idosa (acima de 60 anos).

Fonte: Wikipédia – Pirâmides populacionais.

Para entender melhor o que estamos dizendo com relação aos estágios de transição demográfica seguem dois exemplos: o primeiro da África que pode ser identificado como primeiro estágio da transição demográfica e a pirâmide da Europa no quarto estágio da transição (Figura 3.4).

Figura 3.4: Pirâmides populacionais da África e da Europa, 2006.

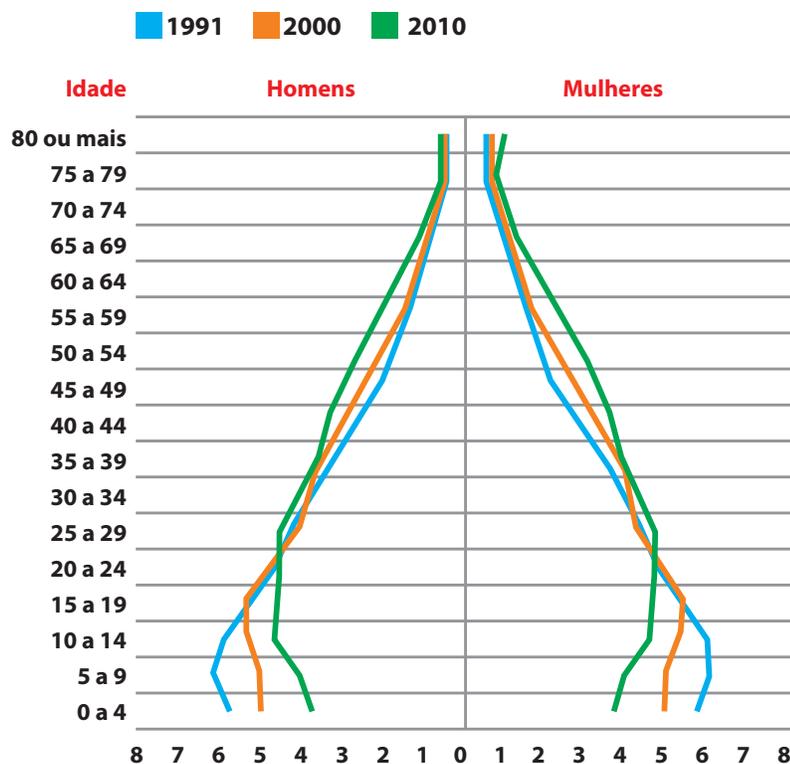


Fonte: <<http://cocgeografia.blogspot.com.br/2011/05/piramides-etarias.html>>.

Nesse sentido, destaca-se que desde o século XIX até meados da década de 1940, o Brasil caracterizou-se por altas taxas de natalidade e de mortalidade. A partir desse período, com a incorporação às políticas de Saúde pública dos avanços da medicina, particularmente os antibióticos descobertos na época e importados no pós-guerra, o país experimentou a primeira fase de sua transição demográfica, caracterizada pelo início da queda das taxas de mortalidade.

No caso do Brasil, a mudança no formato da pirâmide foi relacionada diretamente com a queda da fecundidade e da mortalidade. Veja a Figura 3.5 a seguir.

Figura 3.5: Composição da população residente total, por sexo e grupos de idade, Brasil, 1991, 2000 e 2010.

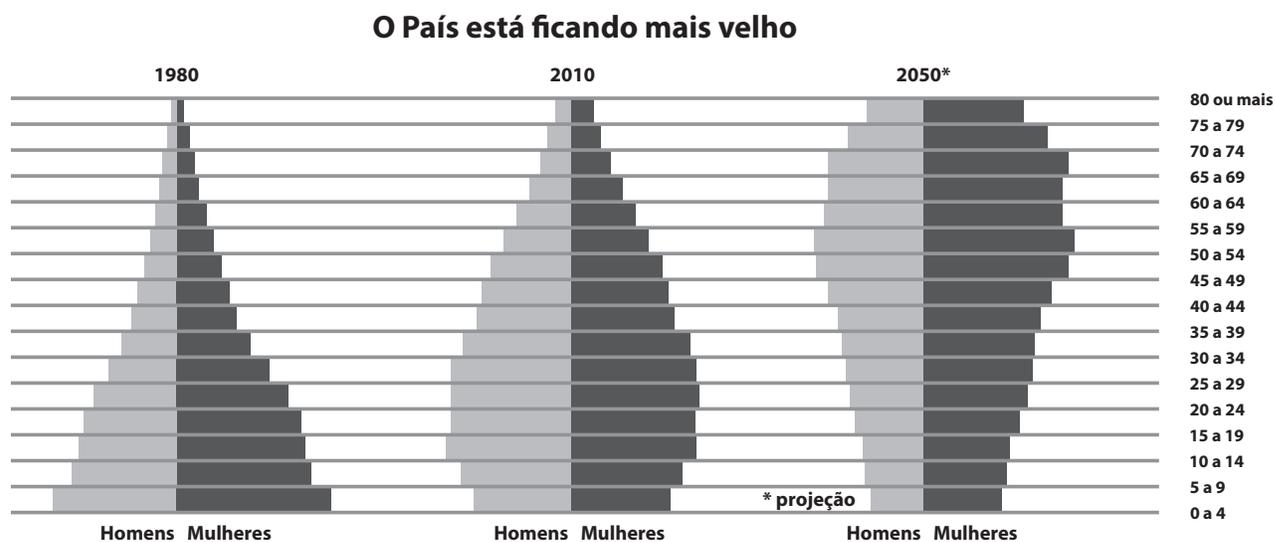


Fonte: Censo Demográfico, IBGE, 2012.

Nas últimas décadas, o Brasil tem registrado redução significativa na participação da população com idades até 25 anos e aumento no número de idosos. E a diferença é mais evidente se comparada às populações de até 4 anos de idade e acima dos 65 anos. Em 2010, conforme o Censo Demográfico, o País tinha 13,8 milhões de crianças de até 4 anos e 14 milhões de pessoas com mais de 65 anos².

Com relação ao estágio da transição demográfica, pode-se observar que em 2010 o formato da pirâmide mostra que o Brasil se encontra no terceiro estágio, e que também segundo as projeções, alcançaremos o quarto estágio em menos de 48 anos se forem mantidas as quedas nas taxas de fecundidade e de mortalidade apresentadas atualmente. (Figura 3.6)

Figura 3.6: Pirâmides etárias do Brasil, 1980, 2010 e 2050.



Fonte: Revista Atualidades. São Paulo: Abril, n. 14, p. 141, 2011.

Inserida e intensamente articulada a esse contexto de desenvolvimento desequilibrado, a transição demográfica não é autônoma. Ela é um processo social que não se resume aos efeitos combinados das variáveis estritamente demográficas. Pelo contrário, imersa nas profundas mudanças sociais e econômicas pelas quais tem passado o Brasil, é, simultaneamente, uma de suas causas e um de seus efeitos. Como tal, está longe de ser considerada neutra: tanto pode criar possibilidades demográficas que potencializem o crescimento da economia, aumentando o bem-estar social, quanto pode potencializar as adversidades econômicas e sociais, ampliando as graves desigualdades sociais que marcam a sociedade brasileira.

Isso é fundamental, pois, se a transição não é neutra e pode favorecer consequências sociais diversas, ela depende de políticas que podem colocá-la nos trilhos que a levará a um destino ou outro, e se as possibilidades abertas pela transição demográfica devem significar o desafio de ultrapassar esses limites, ampliando as dimensões dos caminhos que podem levar à redução das desigualdades sociais.

3. Mortalidade, natalidade, fecundidade e migrações

3.1 Mortalidade

Ao estudar a mortalidade não se pode perder de vista que, se por um lado a morte é encarada como fenômeno individual dependente de fatores biológicos, quando vista sob o ângulo de um fenômeno coletivo está afetada pelo contexto social em que os indivíduos realizam suas trajetórias de vida. Assim a interação do social com o biológico determina modificações que acabam por alterar os riscos de morrer dos indivíduos.

O cálculo da **taxa bruta de mortalidade (TBM)** é uma importante medida e corresponde ao total de óbitos ocorridos durante o ano sobre a população total, indicando o risco que tem uma pessoa dessa população morrer no decorrer desse ano (por mil). É importante salientar que o nível da TBM dependerá de dois componentes básicos: **a intensidade com que se morre a cada idade e a distribuição etária proporcional da população**. O primeiro componente explica-se porque em diferentes idades as pessoas estão sujeitas a diferentes riscos de morte, por exemplo: os recém-nascidos e os idosos têm maior chance de morrer do que os adolescentes. O segundo decorre do primeiro porque se os riscos são diferenciados por idades, há de se levar em conta o maior ou menor peso dos diversos grupos etários.

Vale ainda destacar que tão importante quanto a TBM é a taxa de mortalidade infantil (TMI), indicador este muitas vezes usado como medida, assim como outros, para avaliar o desenvolvimento de um país, que é o caso do Índice de Desenvolvimento Humano (IDH). Neste sentido, a TMI corresponde ao risco que um nascido vivo tem de vir a falecer antes de completar 1 ano de idade e é considerado um indicador mais robusto, pois tem uma forte correlação com as condições de vida em geral.

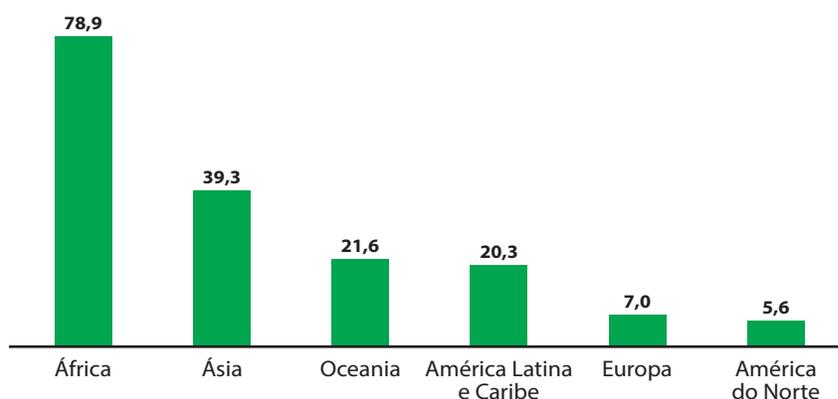
Analisando o comportamento da mortalidade infantil, observa-se que o aumento da escolaridade feminina, a elevação do percentual de domicílios com saneamento básico adequado (esgotamento sanitário, água potável e coleta de lixo) e o acesso aos serviços de Saúde contribuíram para a queda da taxa de mortalidade infantil em todo o País. A taxa de mortalidade infantil no Brasil continua em declínio, passando de 31,7%, em 1999, para 22,5%, em 2009, o que corresponde a uma queda de 29,0% no período.

A melhoria das condições de habitação, particularmente o aumento relativo do número de domicílios com saneamento básico adequado, vem contribuindo para reduzir as mortes infantis. O Rio Grande do Sul foi o estado que registrou a menor taxa de mortalidade infantil em 2009 (12,7%), e Alagoas, com 46,4%, apresentou a mais elevada. Contudo, ainda há um longo

percurso pela frente, uma vez que a mortalidade infantil no Brasil, estimada em 23,30 óbitos de menores de 1 ano para cada mil nascidos vivos, em 2008, é alta quando comparada com os indicadores correspondentes aos países vizinhos do cone sul para o período 2005-2010. No mesmo período, os países como, por exemplo, Argentina (13,40 por 1.000), Chile (7,20 por 1.000) e Uruguai (13,10 por 1.000) registraram taxas bem menores. Vale lembrar que, em 1970, a taxa de mortalidade infantil no Brasil estava próxima de 100 óbitos de crianças menores de 1 ano por 1.000 nascidos vivos.

Ainda segundo as Nações Unidas, em relação à mortalidade infantil, indicador também muito utilizado para aferir os níveis de desenvolvimento econômico e social dos países, no continente americano, a América do Norte se sobressai com a menor taxa, 5,6 óbitos por cada 1.000 nascidos vivos. O Brasil (21,8‰) ainda apresenta uma taxa ligeiramente superior à da América Latina, 20,3‰³ (Figura 3.7).

Figura 3.7: Taxa de mortalidade infantil, segundo a América Latina e Caribe, América do Norte, Ásia, África, Europa e Oceania – 2010.



Fonte: World population prospects: the 2008 revision. In: ONU, Population Division. Population Database. New York, 2010. Disponível em: <<http://esa.un.org/unpp>>. Acesso em: set. 2010.

O Brasil por algum tempo experimentou declínios nas taxas de mortalidade em todas as idades, mas a partir de meados dos anos 1980, as mortes associadas às causas externas (acidentes de qualquer natureza e violência) passaram a desempenhar papel de destaque, e infelizmente de forma desfavorável, sobre a estrutura por idade das taxas de mortalidade, particularmente dos adultos jovens do sexo masculino.

As causas externas de morbidade e mortalidade compreendem as lesões decorrentes de acidentes (relacionados ao trânsito, afogamento, envenenamento, quedas ou queimaduras) e de violências (agressões/homicídios, suicídios, tentativas de suicídio, abusos físicos, sexuais e psicológicos), as quais se impõem como importante desafio às autoridades de saúde pública. Anualmente, essas causas são responsáveis por mais de 5 milhões de mortes em todo o mundo, representando cerca de 9% da mortalidade mundial. No mundo, as causas externas apresentam comportamento de constante crescimento, ocupando usualmente as primeiras posições entre as mais frequentes causas de morte. Porém, esses agravos não afetam a população de maneira uniforme.

Evidências já demonstraram que há grupos populacionais mais vulneráveis, o que pode ser percebido pela distribuição desigual das mortes por causas externas, as quais atingem, sobretudo, pessoas de 5 a 44 anos, do sexo masculino e residente em países de baixa e média renda, com diferentes gradações entre áreas pobres e ricas de um mesmo país ou cidade⁴.

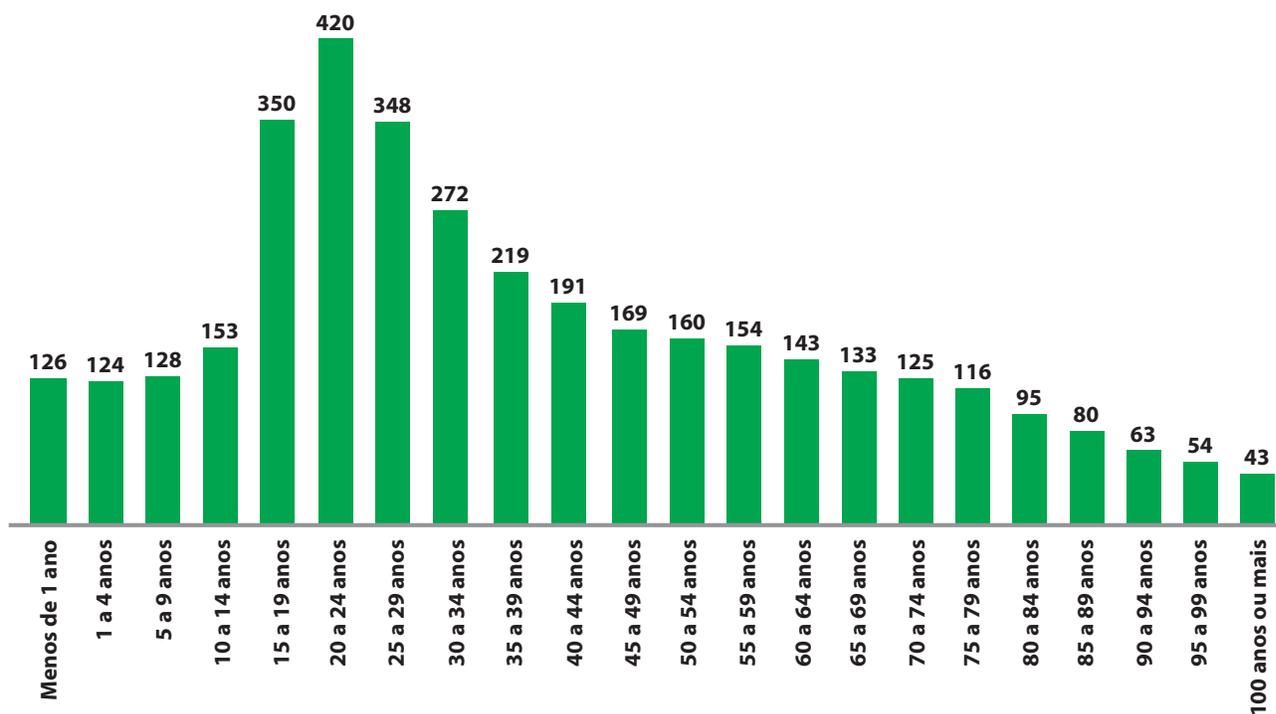
Assim como as causas acidentais de morbimortalidade, a violência apresenta-se sob diversas manifestações e com diferente distribuição nas populações. A partir de levantamentos consolidados pela OMS, as principais manifestações da violência podem ser ilustradas pelos homicídios e suicídios. As taxas de mortalidade por homicídio entre os homens são três vezes mais altas do que entre as mulheres.

Nota

No Brasil, atualmente, as causas externas representam a terceira causa mais frequente de morte, passando a ocupar a primeira posição quando se restringe a análise ao grupo de pessoas de 1 a 39 anos, configurando-se como inquestionável desafio aos gestores públicos, especialmente os dirigentes e profissionais do setor Saúde, e conduzindo a inclusão do tema na agenda de prioridades em saúde.

A esperança de vida no Brasil continuou se elevando, mas poderia, na atualidade, ser superior em 2 ou 3 anos à estimada, se não fosse o efeito das mortes prematuras de jovens por violência. Basta constatar que, em 2000, a mortalidade masculina no grupo etário 20 a 24 anos era quase quatro vezes superior à da feminina e, este indicador estaria se elevando com o passar dos anos (Figura 3.8).

Figura 3.8: Razão entre o número de óbitos masculinos e femininos, segundo os grupos de idade – Brasil, 2010.



Nota: Excluímos os óbitos sem declaração da idade das pessoas ao falecerem.

Fonte: IBGE, Censo Demográfico 2010.

Do total de 1.103.088 mortes notificadas em 2009, 138.697 (12,5%) foram decorrentes de causas externas, representando a terceira causa mais frequente de morte no Brasil.

Na demografia, a noção de risco tradicionalmente é vista como a probabilidade de ocorrer um evento da dinâmica demográfica (fecundidade, migração e mortalidade). No campo específico da saúde, muitos estudos se voltam para os riscos de morte ou de contrair uma doença. Estudos recentes sobre a aids têm procurado ampliar as discussões para ultrapassar a dimensão comportamental do risco, incorporando o contexto social – vulnerabilidades. Nesses

estudos, são consideradas as diferentes chances que cada indivíduo ou grupo populacional particular tem de se contaminar, dado o conjunto formado por certas características individuais e sociais de cotidiano, “julgadas relevantes para a maior exposição ou a menor chance de proteção diante do problema”.

Ressalta-se, ainda, a importância de verificar a capacidade do indivíduo de se proteger diante da materialização do risco (*empowerment*). Assim, *no campo da epidemiologia aliada à demografia, um avanço na noção de risco*, enquanto medida objetiva, quantitativa e comportamental se dá pela incorporação da biface vulnerabilidade – *empowerment* como duas faces do mesmo processo, que interagem na equação do *risco e da saúde*. No entanto, a conceituação de vulnerabilidade ainda continua em construção e deve ser uma dimensão importante de ser incorporada nas discussões de mortalidade.

Outro ponto fundamental a ser considerado nos estudos de mortalidade é a ênfase nos processos coletivos, sociais e demográficos, e na face política da doença e do risco, influenciando a capacidade das pessoas e grupos de se protegerem e/ou se tratarem.

3.2 Natalidade

A natalidade refere-se à relação entre os nascimentos vivos e a população total. Observação: adota-se, no numerador, o número de nascidos vivos informados no Sinasc, desde que igual ou superior a 90% do número de nascidos vivos estimado por métodos demográficos. Sendo inferior, recomenda-se adotar o número estimado. Os totais para as regiões e o Brasil combinam os dados diretos e indiretos. A taxa bruta de natalidade (TBN) depende da maior ou menor intensidade com que as mulheres têm filhos a cada idade, do número das mulheres em idade fértil, como proporção da população total, e da distribuição etária relativa das mulheres dentro do período reprodutivo. Portanto, não é um bom indicador para se analisar diferenciais de níveis de fecundidade entre populações.

$$\frac{\text{Número total de nascidos vivos residentes}}{\text{População total de residentes}} \times 1000$$

No período, observam-se valores decrescentes para todas as regiões brasileiras. Os dados padronizados evidenciam as diferenças regionais, sendo que as regiões Norte e Nordeste apresentam os mais elevados níveis de natalidade e as regiões Sul e Sudeste os mais reduzidos (Tabela 3.4).

Tabela 3.4: Taxas bruta e padronizada de natalidade. Brasil e grandes regiões, 1991, 1995, 2000 e 2004.

	1991	1995	2000		2004	
Regiões	Bruta	Bruta	Bruta	Padronizada	Bruta	Padronizada
Brasil	23,39	21,97	21,06	20,94	18,17	18,11
Norte	31,93	30,14	28,63	28,15	23,62	22,75
Nordeste	26,81	25,13	24,29	24,23	21,66	21,26
Sudeste	20,23	19,20	18,71	18,52	15,87	15,84
Sul	21,49	19,77	17,96	18,41	14,83	15,61
Centro-Oeste	24,38	22,34	20,70	19,22	19,16	17,95

Fonte: Estimativa: IBGE/Projeções demográficas. Dados Diretos: MS/SVS/Sistema de Informações sobre nascidos vivos.

3.3 Fecundidade

É importante não confundir fecundidade com fertilidade. A fertilidade diz respeito ao potencial reprodutivo das mulheres, enquanto fecundidade é o resultado concreto da capacidade reprodutiva delas (filhos tidos). Vale destacar que quanto maior o controle exercido pelas mulheres sobre o tamanho de sua prole (utilização de métodos contraceptivos) maior será a distância entre a fertilidade e a fecundidade. O estado conjugal (duração das uniões), a idade do início da vida sexual, a frequência das relações sexuais e perdas fetais, bem como o uso de métodos contraceptivos são alguns elementos que são levados em consideração nas análises e cálculos da fecundidade¹.

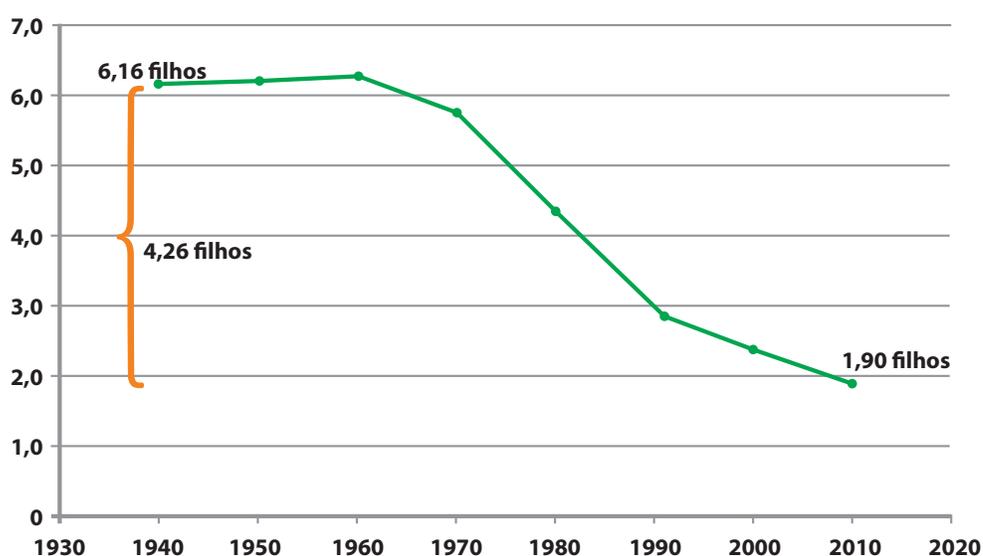
A fecundidade refere-se à relação entre nascimentos vivos e mulheres em idade reprodutiva (15 a 49 anos), e a medida mais usada para o cálculo desta relação é chamada de taxa de fecundidade total (TFT), que é uma estimativa do número médio de filhos que uma mulher teria até o final de seu período reprodutivo, mantidas constantes as taxas observadas na referida data. Também pode ser definida como o número médio de filhos por mulher em idade de procriar, ou seja, de 15 a 49 anos.

Um aspecto a se chamar a atenção quanto à TFT é que ela não é influenciada pela distribuição etária das mulheres da população a qual se refere, pois a TFT é construída a partir das taxas específicas de fecundidade (TEFs), que correspondem às médias de nascimentos vivos por

mulher nos diversos grupos etários. As TFTs de diferentes populações podem ser usadas para comparação de níveis de fecundidade, pois dependem apenas das TEFs e não das distribuições etárias concretas.

Segundo os resultados da amostra do Censo Demográfico 2010, o número médio de filhos tidos nascidos vivos por mulher ao final de seu período fértil, no Brasil, foi de 1,90 filho. Diminuição de 69,2% em relação ao valor de 1940 (Figura 3.10).

Figura 3.10: Taxa de Fecundidade Total, Brasil 1940-2010.



Fonte: IBGE, Censo Demográfico 2000/2010.

A queda da fecundidade estende-se a todas as regiões do País e a todos os grupos sociais, mas com ritmo diferenciado. O primeiro diferencial considerado foi por regiões, como mostra o gráfico. Este compara a taxa de fecundidade total das cinco regiões brasileiras em 2000 e 2010. Tanto em 2000 como 2010, a mais alta taxa foi verificada na Região Norte. Nos dois anos considerados, a taxa mais baixa foi observada na Região Sudeste, sendo que no último ano o valor alcançado foi de 1,7 filho, valor este semelhante ao observado na Região Sul (1,78). Os diferenciais regionais também diminuíram no período. Com exceção da Região Norte, todas as demais estão abaixo do nível de reposição (TFT = 2,10 filhos). O maior decréscimo na taxa, entre 2000 e 2010, foi observado na Região Nordeste (23,4%). Sumarizando, pode-se dizer que a fecundidade de todas as regiões brasileiras ou já atingiram ou já estão bem próximas de atingir

o nível de reposição, bem como está tendendo a uma convergência.

Filha de Ana Maria de Freitas Tavares, de 64 anos, que teve 11 filhos, a pernambucana Ieda de Freitas, 39, é o retrato da diferença de tamanho das famílias brasileiras. Mãe de duas meninas, ela engravidou cedo, aos 16 e 17 anos, mas aos 19 decidiu que não teria mais filhos. Moradora da Favela do Pilar, em Recife, Ieda criou as meninas sozinha. Das duas, uma já casou, também só tem dois filhos e não quer mais. A outra não pretende ter nenhum⁵.

Segundo o IBGE, no Brasil, a taxa de fecundidade registrada em 2010 é menor do que a registrada em 2000 em todas as regiões, todos os graus de escolaridade e grupos raciais. Mas as maiores quedas aconteceram entre as mulheres pretas no Nordeste (29,1%), Norte (27,8%) e Sul (25,3%) (Tabela 3.4).

Tabela 3.4: Taxas bruta e padronizada de natalidade. Brasil e grandes regiões, 1991, 1995, 2000 e 2004.

Grande Regiões	Taxa de Fecundidade Total		Diferença Relativa
	2000	2010	2000/2010 (%)
Brasil	2,38	1,90	-20,1
Norte	3,16	2,47	-21,8
Nordeste	2,69	2,06	-23,4
Sudeste	2,10	1,70	-19,0
Sul	2,24	1,78	-20,6
Centro-Oeste	2,25	1,92	-14,5

Fonte: IBGE, Censo Demográfico 2000/2010.

Há uma tendência à convergência das taxas no País. As taxas no Sudeste e no Sul, por exemplo, começaram a cair no fim da década de 1960, enquanto no Norte e no Nordeste elas só começaram a cair nos anos 1980. A divergência foi maior nos anos 70, mas a partir dos anos 1980 veio a convergência. O Sudeste continuará a cair, mas em um ritmo menor.

O provável é que a taxa não suba mais, assim como é tendência das regiões não terem mais tanta disparidade. Conforme as mulheres forem ficando mais instruídas e ocupando mais cargos no mercado de trabalho, elas vão adiando a maternidade e tendo menos filhos. A tendência observada é que as famílias serão pequenas, independente da classe econômica e do quanto estudaram. Embora a taxa tenha caído, o censo mostra que ainda há distância entre o número de filhos quando são consideradas as mulheres nos extremos das faixas de renda (3,9 filhos para as que ganham até um quarto do salário mínimo e 0,97 filho para aquelas com mais de cinco mínimos).

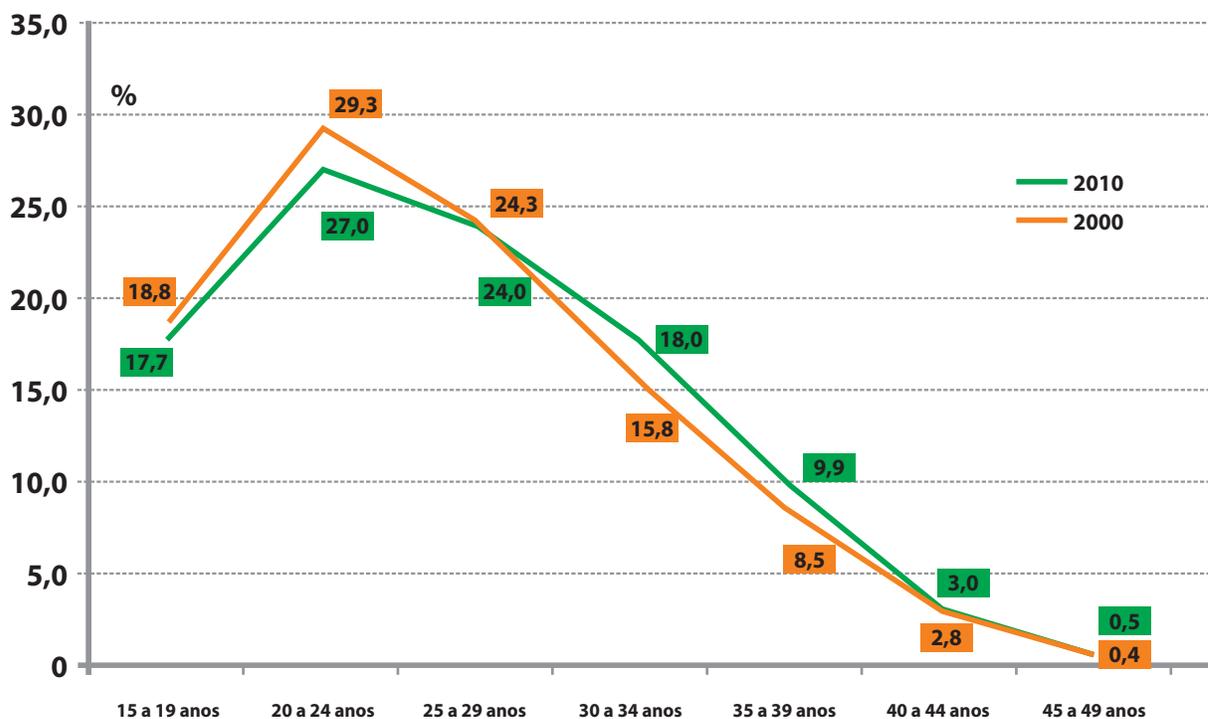
Apesar de estar havendo queda também entre as mais pobres, o que ocorreria em muitos casos é uma “reprodução do ciclo da pobreza”, o que impediria uma queda maior.

Falta às mulheres mais pobres mais acesso a serviços de saúde e políticas de planejamento. Uma menina que engravida na adolescência compromete sua possibilidade de sair da pobreza, estimulando, por sua vez, o não acesso a esses serviços. Nesse sentido, a política de transferência de renda do Bolsa Família, apesar de ter um desenho pró-natalista, impulsiona a queda de fecundidade ao estimular o acesso à renda e à educação. O interessante é que a queda não é imposta como na China. Mas, tendo menos filhos, a criação dessas crianças poderia ter mais qualidade.

A caminho de um padrão de fecundidade mais tardio, o gráfico abaixo mostra que a fecundidade das mulheres menores de 30 anos representava 72,4% da fecundidade total em 2000. Os grupos etários em que os aumentos mais significativos ocorreram foram de 30-34 anos primeiramente e 35-39 anos em seguida. Em 2010, esta participação foi de 68,6%. Considerando o período analisado, a queda mencionada da fecundidade ocorreu em todos os grupos de idade, inclusive entre as mulheres de 15 a 19 anos. Esse decréscimo não foi linear no período, pois se observou um aumento nesse indicador na segunda metade dos anos 1990 e uma reversão nessa tendência a partir de 2001 (Figura 3.11).

Não se trata de introduzir a discussão sobre níveis ou metas a serem alcançadas com respeito ao número ideal de filhos por família no Brasil. Porém, a convergência generalizada para uma fecundidade muito abaixo do nível de reposição das gerações certamente implicará em adequações e alterações substanciais nos atuais Sistemas de Saúde e, em especial, no de Previdência e Assistência Social, como consequência direta da intensificação da velocidade com a qual a população envelhece.

Figura 3.11: Distribuição relativa das taxas específicas de fecundidade (padrão de fecundidade) – Brasil, 2000-2010



Fonte: IBGE, Censo Demográfico 2000/2010.

Por outro lado, não se deve desconsiderar que a permanência de níveis marcadamente diferenciados de fecundidade, no Brasil, envolve questões associadas, não somente aos típicos condicionantes desta variável, mas também às maiores ou menores oportunidades de acesso que as mulheres, em idade fértil, possuem para fazerem uso dos mecanismos disponíveis para a regulação do número de filhos. Isto porque dificilmente deixará de existir por completo aquela parcela de mulheres, cujas necessidades que proporcionem as tomadas de decisão sobre o número desejado de filhos não serão satisfeitas.

3.4 Migrações

3.4.1 Migrações internas

Os deslocamentos de população de um determinado espaço a outro – ou seja, os chamados movimentos migratórios – distinguem-se dos demais componentes da dinâmica popula-

cional vistos até agora (natalidade, fecundidade e mortalidade), pelo menos sob dois aspectos:

a) não apresentam dimensão biológica como os anteriores e

b) por meio da redistribuição espacial da população, influenciam os efeitos decorrentes da magnitude das taxas de fecundidade e mortalidade encontradas em determinada população.

O fenômeno migratório é social, assume a dimensão de classe social, que estaria respondendo aos processos social, econômico e político ao migrar. Para o autor, “as migrações internas são sempre historicamente condicionadas, sendo o resultado de um processo global de mudança, do qual elas não devem ser separadas”⁶.

Se considerarmos o espaço de deslocamento temos as migrações internas ou nacionais (são aqueles deslocamentos que se realizam dentro do mesmo país) e a externa ou internacional (que se realiza de um país para o outro). Entre as migrações internas estão a inter-regional (de uma região para outra) a intrarregional (dentro da mesma região). Ainda devemos considerar o êxodo rural (transferência de população rural para o espaço urbano), o urbano-urbano, a migração sazonal (ligada às estações do ano), a migração de retorno (retorno para a região de origem ou cidade) e atualmente uma das mais importantes que é a migração diária ou pendular (ocorre fundamentalmente nos grandes centros urbanos, onde milhares de trabalhadores saem todas as manhãs em direção ao trabalho e retornam no final do dia ou final da semana).

A mobilidade espacial da população no território nacional insere-se em um contexto mais amplo de transformações da sociedade em seu conjunto. Cabe destacar que os distintos contextos históricos, econômicos, sociais, demográficos e políticos tiveram implicações nos processos de redistribuição da população e de urbanização ao longo do século XX.

A crise econômica mundial em 1929 e a conseqüente crise do café contribuíram para o incipiente processo de industrialização nacional. Neste período a população das áreas de cafeicultoras rurais se transferiu para as áreas urbanas. Os planos de desenvolvimento industrial, pós-1930, exigiram a unificação do mercado. Assim as migrações internas entre 1930 e 1950, seguiram, basicamente, rumo ao meio urbano dos municípios, às fronteiras agrícolas e aos centros industriais do Sudeste; a migração rural-urbana chegou a 3 milhões de pessoas na década de 1940 passando a 7 milhões na década de 1950 dado o processo de urbanização e industrialização pesada.

Em 1960 esgotam-se as antigas áreas de fronteira, totalizando 12,8 milhões de pessoas que saíram do campo entre 1960 e 1970. A década de 1970 intensificou essas tendências apontadas na década anterior, ressaltando-se que, mesmo com a fronteira amazônica, as migrações passaram a ser predominantemente em direção ao meio urbano; esse período marcou a con-

solidação dos grandes centros urbanos com o crescente processo de metropolização⁷.

Entende-se assim, que os deslocamentos de população no Brasil tiveram um período intenso, marcado pelos anos 1960–1980, quando grandes volumes de migrantes se deslocaram do campo para a cidade, delineando um processo de intensificação da urbanização e caracterizando áreas de expulsão ou emigração: Região Nordeste e os estados de Minas Gerais, Espírito Santo, Santa Catarina e Rio Grande do Sul e áreas de atração ou forte imigração populacional – núcleo industrial – formadas pelos estados de São Paulo e do Rio de Janeiro⁸.

Esses deslocamentos, típicos da primeira fase da transição demográfica⁹, quando as taxas de fecundidade eram altas e a mortalidade começava a declinar, criando excedentes populacionais que favoreciam a migração do campo para a cidade, começaram a perder importância no Brasil a partir dos anos de 1980.

A partir da década de 1980, o comportamento da mobilidade espacial da população sofreu importantes transformações nos países desenvolvidos e em desenvolvimento. Aqueles movimentos que tinham, de modo geral, como características básicas a migração para os grandes centros, passaram a ter como destino as cidades médias e serem cada vez mais de curta duração. Por outro lado, os deslocamentos pendulares ganham importância ainda maior, deixando de ser um fenômeno meramente metropolitano¹⁰. De modo que surgem novos eixos de deslocamentos envolvendo expressivos contingentes populacionais, destacando:

- i) a inversão nas correntes principais nos estados de Minas Gerais e do Rio de Janeiro;
- ii) a redução da atratividade migratória exercida pelo Estado de São Paulo;
- iii) o aumento da retenção de população na Região Nordeste;
- iv) os novos eixos de deslocamentos populacionais em direção às cidades médias no interior do País;
- v) o aumento da importância dos deslocamentos pendulares;
- vi) o esgotamento da expansão da fronteira agrícola e
- vii) a migração de retorno para o Paraná.

Nos últimos anos da década de 1980 e nos anos de 1990, observou-se uma diminuição no volume desses migrantes e a formação de novos fluxos migratórios, incluindo-se, nesse contexto, as migrações a curta distância e aquelas direcionadas às cidades médias. O Censo Demográfico de 1991, muito rico em informações a respeito das migrações internas, já havia apontado algumas transformações no comportamento dos fluxos que antes predominavam no Brasil, como o arrefecimento das migrações do Nordeste para o Sudeste e algumas reversões nos saldos migratórios das unidades da Federação.

O Censo de 2000 confirmou algumas tendências nos fluxos migratórios apontados e mostrou novos espaços de redistribuição populacional. Apresentou que os deslocamentos entre as regiões brasileiras envolvem cerca de 3,3 milhões de pessoas, as quais entre entradas e saídas, destacou-se a Região Nordeste que apresentou a maior perda absoluta (760 mil pessoas), tendo as trocas com o Sudeste contribuído com cerca de dois terços dessa perda. Nos últimos anos da década passada, o Nordeste continuou sendo uma região de expulsão populacional, visto que as trocas com as outras regiões brasileiras foram negativas, sendo a Região Sul a que apresentou o menor saldo nas trocas com o Nordeste brasileiro.

Além do Nordeste, a Região Sul também apresentou pequeno saldo negativo, tendo os maiores volumes de trocas com o Sudeste. A Região Sudeste foi a que apresentou o maior saldo líquido absoluto, fruto da imigração nordestina, pois as trocas com as outras regiões não foram expressivas em termos quantitativos.

O Centro-Oeste destacou-se por ter apresentado saldo positivo na troca com todas as regiões, ou seja, ele pode estar se tornando pólo de atração de população das demais regiões brasileiras. A Região Norte apresentou saldo positivo nas trocas com as outras regiões, sendo o maior volume de imigrantes nordestinos. A migração de retorno representava 19% do total de imigrantes no quinquênio.

As trocas entre as unidades da Federação mostram que as principais correntes migratórias, observadas no passado, mantiveram-se, destacando-se a migração de retorno no contrafluxo, tendo a unidade da Federação do Nordeste brasileiro apresentado os maiores percentuais de retornados entre os imigrantes, como o Ceará e a Paraíba com quase a metade dos imigrantes, nesses cinco anos, composta por retornados. Esses dois estados têm como principais unidades da Federação de origem, no quinquênio 1995-2000, São Paulo e Rio de Janeiro. Com exceção do Rio Grande do Norte e de Sergipe, os demais estados da Região Nordeste apresentaram percentuais de retornados acima de 40% do total de imigrantes em seus estados. Paraná, Rio Grande do Sul e Minas Gerais também tiveram mais de um terço de retornados no total de imigrantes.

Veja na Figura 3.12 as informações do Censo de 2000 que ilustram o apresentado até o momento.

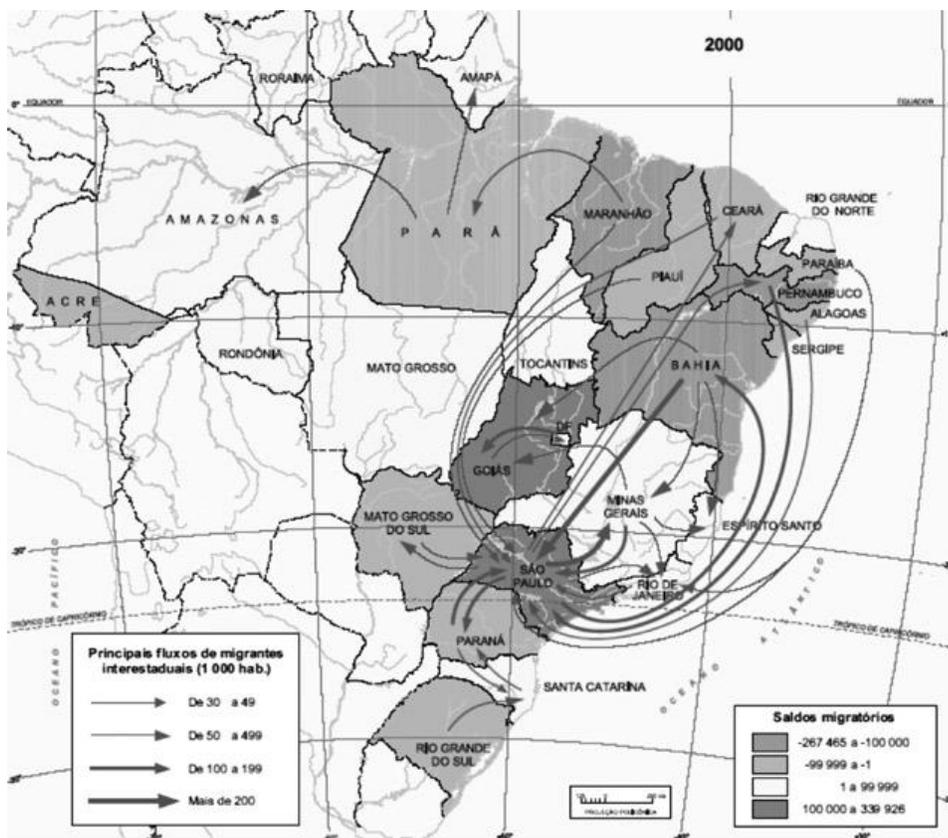
O volume de migrantes entre as unidades da Federação praticamente manteve-se inalterado entre os quinquênios 1995-2000 e 2005-2010: 5.196.093 e 5.018.898 migrantes, respectivamente. Observa-se uma redução na mobilidade espacial da população. Entre 1995-2000, movimentaram-se 30,6 migrantes para cada 1.000 habitantes, já no período 2005-2010, observaram-se 26,3 migrantes para cada 1.000 habitantes.

Os primeiros resultados divulgados do Censo Demográfico de 2010 apresentam apenas

os volumes populacionais desagregados por município. Com esses dados é possível se estabelecer, a partir do cálculo da taxa média geométrica entre os anos de 2000 e 2010, os eixos de crescimento populacional no País e, por meio de uma proxy, especular sobre as áreas que ganham ou perdem população, de modo a inferir se houve alteração no comportamento dos deslocamentos de população na década passada.

As cidades com menos de 500 mil habitantes são as que mais crescem no País, o que demonstra a influência da migração, muito embora as grandes cidades continuem concentrando parcela expressiva da população (aproximadamente 30%). Esse fenômeno vem ocorrendo nas últimas três décadas, o que reforça o caráter de “desconcentração concentrada” dois na distribuição populacional no Brasil. Os municípios com 500 mil habitantes ou mais aumentaram em quantidade quando comparados com o ano de 2000, passando de 31 para 38. Outro aspecto a ser destacado é que o ritmo de fragmentação do território, nos anos 2000, foi menos intenso que nas décadas passadas, tendo sido instalados 58 municípios, contra 501 nos anos de 1980 e 1.016 nos anos de 1990.

Figura 3.12: Fluxos migratórios no Brasil, 2000.



Fonte: IBGE, Censo Demográfico, 2000.

As evidências empíricas sinalizam que os movimentos de população no País seguem se processando como nos últimos dois decênios. Desta forma, quando levamos em consideração distâncias maiores, como na escala inter-regional, observamos que os fluxos migratórios apresentaram uma tendência de redução nos seus volumes, muito embora a direção dos principais fluxos seja mantida, com as maiores correntes ocorrendo no eixo Nordeste-Sudeste.

Na mesma linha de observação, notamos, de modo geral, que os principais movimentos ocorrem dentro das próprias regiões, sugerindo deslocamentos a menores distâncias.

Outro aspecto da migração que permanece é o retorno, que ganha emergência no auge das transformações que impactaram os movimentos internos no País que ao ter a saturação dos espaços do início da industrialização no Centro-Sul do País, reduzindo a capacidade de criação de emprego e de novas oportunidades ocupacionais, o que coloca o movimento de retorno na pauta das estratégias de reprodução e circulação dos migrantes.

3.4.2 Migrações internacionais

As migrações internacionais, no século XXI, adquirem, cada vez mais, papel importante no cotidiano social, nos mercados de trabalho, nas sociedades de chegada e de partida, nos fluxos financeiros e na mobilidade da força de trabalho¹¹, tornando-se a expressão social dos processos recentes da divisão internacional do trabalho e de seus impactos territoriais.

Alguém poderia especular que, se não houve fugas de cérebros *brain drain* significativas nas duas décadas passadas, um período de estagnação econômica, então, não pareceria razoável desenhar um quadro sombrio para o futuro. Este ponto pode ser questionado em dois aspectos. O primeiro aspecto refere-se ao número crescente de jovens da classe média que contam com alguma experiência internacional, via intercâmbio ou turismo, uma vez que a barreira da língua é cada vez menor. E o segundo aspecto refere-se à demanda por mão de obra nos países europeus que deve aumentar substancialmente nas próximas décadas, como resultado da fecundidade abaixo do nível de reposição e do envelhecimento populacional.

Neste contexto, do ponto de vista dos países europeus, o perfil do imigrante latino-americano é bastante atrativo. Se isto for verdadeiro, então a possibilidade de se perder um segmento dos jovens qualificados brasileiros (ensino médio ou mais) para o mercado de trabalho internacional é algo bastante concreto, o que coloca o *brain drain* na pauta de estudos futuros. Vale destacar também a entrada recente de estrangeiros no País, em especial bolivianos, peruanos e coreanos, os primeiros em sua maior parte clandestinos que trabalham para estes últimos na indústria da confecção.

A migração de brasileiros para outros países tornou-se tema importante nos últimos anos,

tanto nos meios acadêmicos e governamentais, como a repercussão que ganhou na mídia e nos meios de comunicação. De país tipicamente receptor de imigrantes, a partir da década de 1980 o Brasil passou a ser uma importante região de origem de emigrantes internacionais. As preocupações com os fluxos de migrantes estrangeiros para o Brasil também têm ganhado destaque, principalmente de sul-americanos e africanos.

Apesar da dedicação ao estudo das migrações internacionais no Brasil, há um desafio em estimar o número de brasileiros que vivem no exterior, e também o volume de entradas e saídas de população no País. Isso ocorre porque parte significativa destes fluxos migratórios, tanto de imigrantes quanto de emigrantes, é constituída do que se convencionou chamar de “ilegais” ou “clandestinos”, impossibilitando que se conheçam seus verdadeiros números por meio dos registros administrativos e pesquisas domiciliares. Algumas estimativas dão conta de que mais de 60% dos emigrantes brasileiros para os Estados Unidos entraram de forma irregular naquele país.

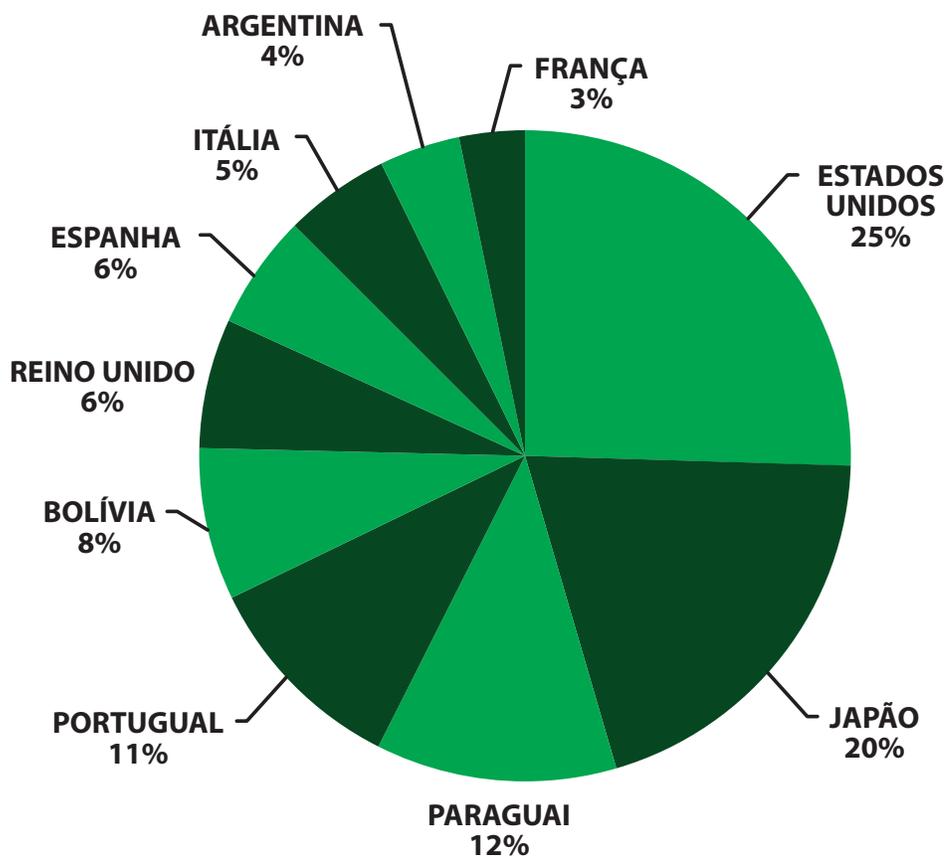
As principais regiões de destino dos migrantes são os Estados Unidos, Japão e alguns países da América do Sul e da Europa¹². Esses números são estimados com base em dados das embaixadas do Brasil no exterior.

Podemos perceber que, embora a migração internacional seja um tema de relevância crescente no país, ainda existem desafios para sabermos qual o número real de migrantes que deixam e que ingressam no Brasil a cada ano.

Os registros administrativos não dão conta desses números e, por isso, faz-se necessário lançar mão de dados censitários para a realização das estimativas. Nesse sentido, os censos vêm sendo aprimorados para poder aumentar a precisão das estimativas de migração internacional, culminando com a inclusão de novos quesitos sobre o tema em 2010.

Os dados divulgados pelo IBGE mostram, segundo Censo de 2010, que o número de imigrantes internacionais do Brasil passou de 143 mil entre 1995 e 2000 para 268 mil entre 2005 e 2010. Dos imigrantes internacionais que chegaram ao Brasil entre 1995 e 2000, 61% eram brasileiros, ou seja, imigrantes internacionais de retorno, enquanto entre 2005 e 2010 o percentual de brasileiros alcançou 65,5% dos imigrantes. Dos 51.933 imigrantes provenientes dos Estados Unidos, 84,2% eram brasileiros. Dos 41.417 imigrantes provenientes do Japão, 89,1% eram brasileiros. Já entre os 15.753 imigrantes provenientes da Bolívia, apenas 25% eram brasileiros (Figura 3.13).

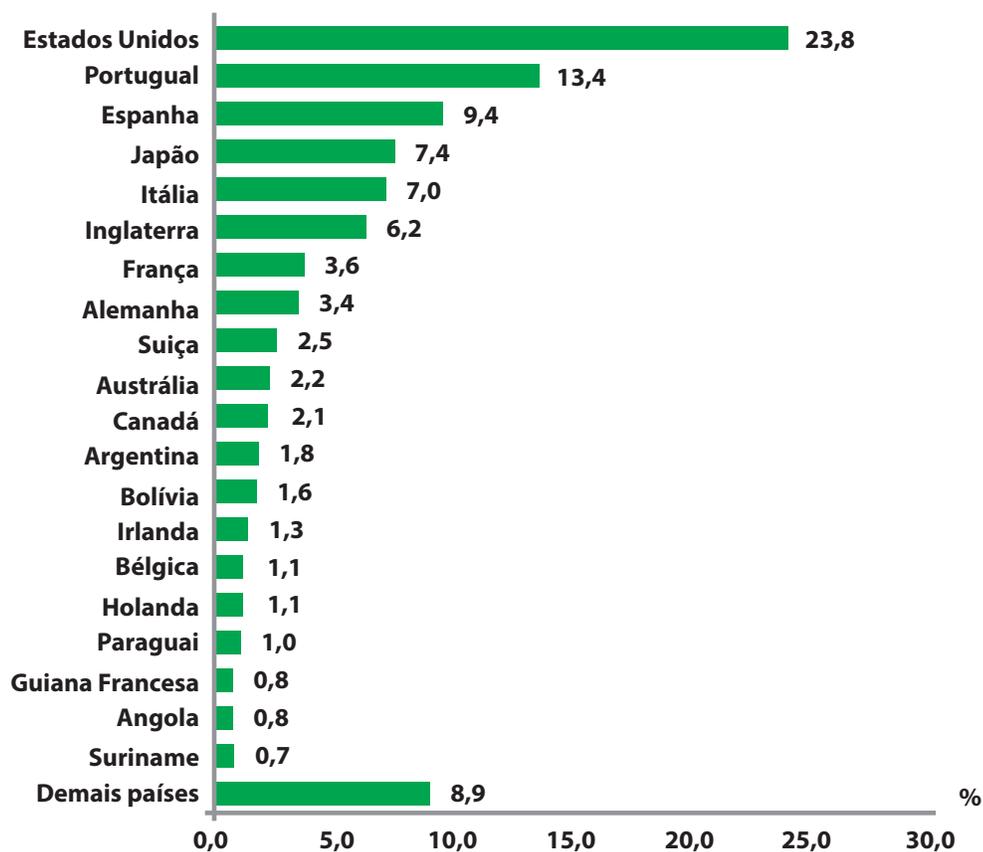
Figura 3.13: Principais países de origem dos imigrantes internacionais do Brasil entre 2005-2010.



Fonte: IBGE– Censo Demográfico 2010; Elaboração MF-SPE.

O Censo Demográfico de 2010 identificou a presença de brasileiros residindo em 193 países ao redor do mundo. Muito embora os destinos sejam bem diversificados, majoritariamente esses fluxos se dirigem aos Estados Unidos (23,8%); Portugal (13,4%); Espanha (9,4%); Japão (7,4%); Itália (7,0%) e Inglaterra (6,2%), ou seja, esses seis países receberam aproximadamente 70,0% dos emigrantes brasileiros. (Figura 3.14).

Figura 3.14: Percentual de emigrantes internacionais, segundo o país de residência Brasil – 2010.



Fonte: IBGE- Censo Demográfico 2010.

Muito embora os deslocamentos entre países vizinhos sejam pouco expressivos, como já assinalados, é importante ressaltar que os movimentos transfronteiriços foram proporcionalmente mais significativos nas unidades da Federação localizadas no Norte do País. Assim, a Guiana Francesa é o principal destino da emigração que parte do Amapá, a Venezuela recebe a maior parte dos fluxos que partem de Roraima e a Bolívia atrai o maior volume de emigrantes oriundos do Acre. Nas fronteiras do Centro-Sul do País, os países vizinhos, Argentina, Uruguai, Paraguai e Bolívia aparecem como quinta opção preferencial².

Unidade 2

Fontes de Dados Demográficos

Não é por coincidência que os textos introdutórios de demografia sempre começam com um capítulo sobre fontes de dados. Historicamente, a demografia tem se caracterizado mais pelo desenvolvimento de técnicas de análise para descrever quantitativamente como as populações se transformam, do que pelo esforço teórico de explicar tais transformações. Grande parte da chamada *demografia formal* dedica-se exclusivamente à crítica e correção de dados empíricos. Por outro lado, o dado demográfico tem aplicações práticas importantes para fins de planejamento (projeção de números de crianças em idade escolar ou trabalhadores em idade de aposentar-se), de diagnóstico (níveis da mortalidade infantil, materna, por causa etc.), de avaliação de programas e estudos socioeconômicos em geral (desemprego, pobreza, moradia etc.).

Seria inútil, neste momento, tentar descrever todos os aspectos relevantes de todos os tipos de informação demográfica levantada nas diversas fontes de dados. No que segue, pretende-se apenas apresentar uma visão global das principais fontes de informação demográficas existentes no Brasil.

Censo Demográfico – Apesar das desvantagens de alto custo, divulgação demorada e frequência reduzida, e a despeito de existirem hoje várias alternativas de coleta de informação, o Censo Demográfico ainda é o principal instrumento para obter dados sobre a população, principalmente nos países em desenvolvimento, onde existem relativamente poucas alternativas. Os censos brasileiros têm variado consideravelmente na sua qualidade e escopo. Embora já tenham sido realizados levantamentos parciais anteriores, como o Censo Colonial de 1808 (o primeiro recenseamento completo do Brasil independente, em 1872, depois da criação da Diretoria Geral de Estatística, em 1870).

Formalmente, este órgão foi encarregado da organização decenal de um recenseamento geral em todo o Território Nacional, mas de fato esta periodicidade foi apenas alcançada a partir de 1940, o mesmo ano em que o Brasil adotou os padrões internacionais explicitados pela Liga das Nações. Antes já tinham sido realizados quatro censos nacionais (1872, 1890, 1900 e 1920), um do Distrito Federal, em 1906, e um censo estadual, de São Paulo, em 1934. O censo do Distrito Federal foi necessário em função da má qualidade do levantamento de 1900. O de 1890 também foi muito deficiente, demorou muito para ser publicado devido à complexidade do questionário e, mesmo assim, foi divulgado incompleto.

O padrão de qualidade dos censos melhorou muito a partir de 1940, devido, em grande medida, ao desempenho pessoal de Giorgio Mortara, que chefiou a execução dos censos pelo IBGE, criado em 1936, órgão sucessor da extinta Diretoria Geral de Estatística. A partir de 1940, a exceção do Censo de 1991, a periodicidade tem sido decenal.

Registro Civil – Enquanto o censo consiste no levantamento de dados sobre todos os indivíduos de uma população em momentos preestabelecidos, o registro civil visa acompanhar as ocorrências de eventos que modificam o tamanho ou a composição da população ao longo do tempo. A unidade de enumeração do registro civil, portanto, é o evento demográfico, enquanto a unidade de enumeração do censo é o indivíduo. Além das suas finalidades estatísticas, o registro civil cumpre uma função legal, já que os eventos registrados modificam a situação das pessoas perante leis que variam de país para país e têm uma maior especificidade e durabilidade do que a legislação referente aos censos.

Segundo as definições das Nações Unidas, um sistema de estatísticas vitais, ou registro civil, deve compreender: o registro oficial dos eventos vitais: óbitos, nascimentos, casamentos, divórcios e, eventualmente, adoções, legitimações e mudanças de ocupação e residência. No caso concreto do Brasil, registram-se nascidos vivos, nascidos mortos, óbitos, óbitos fetais, casamentos, divórcios, adoções e legitimações. A publicação das primeiras estatísticas vitais no Brasil data de 1894, só cobria uma pequena parte do Território Nacional, limitando-se praticamente às capitais estaduais. A omissão naquela época era muito elevada, de modo que os dados publicados representavam apenas 20% dos nascimentos e 25% dos óbitos ocorridos no País.

A publicação foi interrompida entre 1900 e 1931, quando a Diretoria Geral de Estatística foi extinta e a responsabilidade pela divulgação das estatísticas passou para o Serviço de Estatística Demográfica, Moral e Política (SEDMP), do Ministério de Justiça. Este órgão só retomou a publicação dos dados em 1963, quando se iniciou a série *Registro Civil do Brasil*, com as informações relativas a 1959. A série foi novamente interrompida em 1966, com a publicação dos dados de 1964. Depois de alguns anos de indefinição das responsabilidades, a atribuição de apurar as estatísticas vitais passou para o IBGE em 1972. Em 1975, o IBGE reiniciou a publicação da série *Registro Civil do Brasil*, com a divulgação dos dados preliminares do período de 1974 a 1978, concluída em 1979. Neste ano iniciou-se a publicação dos dados definitivos referentes aos mesmos anos, com a série “*Estatísticas do Registro Civil*.”

Pesquisa Nacional por Amostra de Domicílios (PNAD) da Fundação IBGE. A PNAD foi implantada no segundo trimestre de 1967, inicialmente como pesquisa trimestral que visava levantar dados socioeconômicos básicos para todo o País, menos o Distrito Federal, o Norte e o

Centro-Oeste, sendo que os dados referentes ao Distrito Federal eram levantados anualmente. Depois de uma interrupção por causa do Censo de 1970, o sistema foi reiniciado no quarto trimestre de 1971, passando a ser anual. No final de 1973, foram acrescentadas as regiões Norte e Centro-Oeste. Em 1974 e 1975, a PNAD foi temporariamente suspensa por causa da realização do *Estudo Nacional de Despesas Familiares* (ENDEF), um levantamento semelhante ao da PNAD, mas com enfoque e metodologia específicos. A partir de 1976, o questionário básico da PNAD, que antes tinha variado de um ano para outro, foi uniformizado, contendo quesitos sobre os membros do domicílio, ocupação, renda e instrução (Tabela 3.6).

Tabela 3.6: Instituições realizadoras de contagem da população, ano e periodicidade.

Pesquisa	Ano/Periodicidade	Instituição Realizadora
Censo Demográfico	Desde 1872 – Decenal	IBGE
Contagem de população	Realizada em períodos intercensitários 1996 foi a primeira	IBGE
Estatística do Registro Civil	São divulgados pelo IBGE desde 1974	Cartórios de Registro Civil
Pesquisa por amostra de domicílios (PNAD)	Desde 1967 realizada anualmente	IBGE

Fonte: Pesquisas Nacionais Citadas (sistematização para este capítulo feita pela autora).

A **Contagem de População** aconteceu em 1996 e 2007. A Contagem da População, assim denominada por ser uma operação muito mais simples do que um censo demográfico, é planejada para ser realizada no meio da década, com o objetivo de atualizar as estimativas de população, incorporando as mudanças demográficas ocorridas no Território Nacional, desde o último Censo Demográfico. Sua importância está naquilo que as estimativas de população representam para a sociedade, pois se trata de um exemplo expressivo da exigência de informações estatísticas que orientam diretamente a ação pública local. A limitação de recursos orçamentários para os Censos 2007 implicou na necessidade de se fazer um corte na abrangência da Contagem da População. Após alguns estudos, decidiu-se fazer a Contagem nos municípios com até 170 mil habitantes, faixa onde os efetivos de população causam impacto direto nos valores repassados pelo Fundo de Participação dos Municípios – FPM.

Entretanto, como em algumas Unidades da Federação apenas um ou dois municípios ficariam fora da cobertura da Contagem por terem mais de 170 mil habitantes, o IBGE decidiu incluí-los também na pesquisa. Dessa forma, além dos municípios com até 170 mil habitantes, a Contagem da População 2007 abrangeu também um conjunto de 21 municípios com população acima dessa faixa.

O universo pesquisado pela Contagem da População está distribuído em cerca de 30 milhões de domicílios, alcançando 57% do total de domicílios existentes no Brasil.

Os órgãos de planejamento produzem, analisam e divulgam informações para subsidiar o planejamento do governo municipal, estadual e nacional. Desde 1940, os censos demográficos são realizados no início de cada década e monitorados pelo IBGE, órgão que substituiu a Diretoria Geral de Estatística.

A seguir, é apresentada uma lista com alguns centros de estudos demográficos e órgãos de planejamento nacionais e internacionais.

Navegando

IUSSP (International Union for the Scientific Study of Population), promove estudos científicos de demografia e de assuntos relacionados a população. <http://www.iussp.org/>

Population Division – United Nations / Department of Economic and Social Affairs, é responsável para a monitoração e a avaliação de áreas em larga escala referentes a população. <http://www.un.org/esa/population/unpop.htm>

U.S. Census Bureau, serve como principal fonte de dados sobre o povo e a economia da nação americana. <http://www.census.gov/>

CELADE (Centro Latinoamericano y Caribeño de Demografía) – División de Población de la CEPAL (Comisión Económica para América Latina), proporciona assistência técnica, capacitação e informação em população nos países da região. <http://www.cepal.org/celade/>

IBGE (Instituto Brasileiro de Geografia e Estatística), com sede no Rio de Janeiro, constitui-se no principal provedor de dados e informações do país, que atendem às necessidades dos mais diversos segmentos da sociedade civil, bem como dos órgãos das esferas governamentais federal, estadual e municipal. <http://www.ibge.gov.br/>

ENCE (Escola Nacional de Ciências Estatísticas), criada inicialmente para suprir a necessidade da formação de Estatísticos para o IBGE. Tem formado bacharéis em Estatística e mestres em Demografia. <http://www.ence.ibge.gov.br/>

ABEP (Associação Brasileira de Estudos Populacionais) <http://www.abep.org.br/>

IPEA (Instituto de Pesquisa Econômica Aplicada), com a atribuição de elaborar estudos, análises e pesquisas nas áreas econômica e social brasileiras, com sedes em Brasília e Rio de Janeiro. <http://www.ipea.gov.br/>

CEDEPLAR (Centro de Desenvolvimento e Planejamento) é um órgão suplementar da UFMG que abriga um programa de pesquisa e ensino de pós-graduação em Economia regional e urbana e em Demografia. É também onde se encontra a sede da Associação Brasileira de Estudos Populacionais. <http://www.cedeplar.ufmg.br/>

NEPO (Núcleo de Estudos Populacionais) é uma unidade de pesquisa interdisciplinar e multidisciplinar na área de Demografia e Estudos de População da UNICAMP. <http://www.unicamp.br/nepo/>

FUNDAJ (Fundação Joaquim Nabuco), em Recife, tem como missão produzir, acumular e difundir conhecimentos; resgatar e preservar a memória; e promover atividades científicas e culturais, visando à compreensão e ao desenvolvimento da sociedade brasileira, prioritariamente a do Norte e do Nordeste do país. <http://www.fundaj.gov.br/>

IPARDES (Instituto Paranaense de Desenvolvimento Econômico e Social), vinculado a Secretaria de Planejamento do Paraná. <http://www.ipardes.gov.br/>

IPPUC (Instituto de Pesquisa e Planejamento Urbano de Curitiba) tem como atribuição “participar na reformulação da política urbana municipal”. <http://www.ippuc.org.br/>

Com relação aos dados existentes sobre a migração internacional, vale destacar que teoricamente, ela poderia ser estudada com a ajuda da documentação mantida pela Polícia Federal, enquanto os registros hospitalares podem ser aproveitados para diversos estudos sobre a mortalidade e a fecundidade. Da mesma forma, existem fontes alternativas de dados sobre migração, através do Ministério do Interior, e sobre emprego, através do Ministério do Trabalho. Diversos órgãos da administração pública e até empresas privadas coletam informações que às vezes podem ser aproveitadas para o estudo de fenômenos demográficos específicos.

Nos Estados Unidos, os registros das Receitas Federal e Estadual do imposto de renda, bem como o cadastro das autorizações para a construção civil, já se tornaram fontes valiosas para a estimação intercensitária das características populacionais de pequenas áreas. Da mesma forma, o registro das ligações e desligamentos domésticos das utilidades públicas pode ser aproveitado para o estudo da migração. O fluxo de passageiros transportados pelas companhias de ônibus ou as remessas postais de dinheiro também podem servir de base para estimativas de determinados tipos de mobilidade espacial. Embora o sigilo comercial possa constituir um obstáculo ao acesso a diversos bancos de dados particulares, as pesquisas de mercado e os registros das companhias de seguro de vida devem ser mencionados como fontes potenciais de dados demograficamente relevantes. Os registros de pessoal das grandes companhias também podem providenciar informação importante. Para o caso do Brasil, os registros do Banco do Brasil, se mostram como uma fonte comprovadamente confiável sobre a fecundidade, mortalidade e nupcialidade dos empregados. Como todas estas fontes são parciais e sujeitas a vieses, os melhores resultados podem ser obtidos quando se usam diversas fontes simultâneas para estudar um mesmo fenômeno.

Nos últimos anos, as alternativas para a coleta de informação têm sido objeto de uma atenção renovada, na medida em que as fontes tradicionais (principalmente o censo e o registro civil) vêm sofrendo problemas e críticas crescentes. No entanto outros Censos e Pesquisas estão sendo realizados para tentar minimizar algumas lacunas. Neste sentido vale destacar uma pesquisa feita em parceria do Ministério da Saúde com outras instituições a saber como é a PNDS.

- Pesquisa Nacional sobre Demografia e Saúde (PNDS)

A PNDS foi promovida pela Sociedade Bem-Estar Familiar no Brasil (Bemfam) em 1996, e contou com o apoio de várias instituições nacionais e internacionais, fazendo parte do programa mundial de Pesquisas de Demografia e Saúde (DHS). Trata-se de um estudo domiciliar, utilizando subamostra da PNAD. Foram coletadas informações sobre os níveis de fecundidade, mortalidade infantil e materna, anticoncepção, saúde da mulher e da criança, conhecimentos e atitudes relacionadas com doenças sexualmente transmissíveis e AIDS. A PNDS foi realizada também em 1986 e 1991 (neste ano, apenas na Região Nordeste). Nova pesquisa foi realizada em 2006 coordenada pelo Ministério da Saúde e pelo Centro Brasileiro de Análise e Planejamento – CEBRAP.

Unidade 3

Tendências Populacionais e os Desafios para o Século XXI

1. Envelhecimento Populacional

É importante destacar que convivemos neste início de século com quatro principais mudanças nos contextos de vida e comportamento mundial: a primeira grande questão refere-se ao papel desempenhado pela globalização na difusão dos hábitos e padrões de comportamento; a segunda ocorrida na área nutricional e que configurou mudanças na alimentação e aumento do sedentarismo; o perfil epidemiológico mostra predomínio da mortalidade por doenças não transmissíveis, e no que tange ao comportamento demográfico, observa-se o envelhecimento populacional acelerado. E será sobre este aspecto que nos deteremos aqui.

A população brasileira experimentou uma queda na mortalidade antes da queda na fecundidade, conforme indica o aumento na esperança de vida ao nascer (corresponde à quantidade de anos que vive em média a população), que passou de 43,6 anos na década de 1940 a 53,7 anos na década de 1960 – uma variação de cerca de dez anos durante o período. A mortalidade continuou sua tendência de declínio na década de 1970, com a esperança de vida passando a 59,9 anos nos anos de 1970 – um ganho de 6,2 anos em apenas dez anos. O cálculo da esperança de vida se dá a partir de tábuas de vida elaboradas para cada área geográfica, toma-se o número correspondente a uma geração inicial de nascimentos (1_0) e determina o tempo cumulativo vivido por essa mesma geração (T_0) até a idade limite. A esperança de vida ao nascer é o quociente da divisão de T_0 por 1_0 .

As projeções para o Brasil em 2030 mostram que a **esperança de vida** alcançará 78,33 anos para o total da população, com um diferencial para as mulheres de quase 7 anos. Já a expectativa de vida aos 60 anos de idade, que representa o número médio de anos de vida esperados para uma pessoa ao completar 60 anos de idade, mostra que no Brasil, esta vem aumentando progressivamente em todas as regiões brasileiras em ambos os sexos e em média, em 2005, seria de 20,9 anos. Vale ainda ressaltar que em todas as regiões se observa uma sobre mortalidade masculina, o que pode ser percebido pela expectativa de vida aos 60 anos das mulheres ser superior a dos homens.

De acordo com estudos do IBGE, o **diferencial de esperança de vida por sexo** vem aumentando, em parte devido ao peso das mortes por causas externas.

O Brasil experimenta um processo chamado de transição epidemiológica que está diretamente influenciado pelo novo perfil etário da população brasileira mostrado aqui. Representado pela queda nos níveis da mortalidade devido a doenças infectocontagiosas, embora estes ainda estejam em patamares superiores aos encontrados em países desenvolvidos, e por uma elevação na participação dos óbitos por doenças de natureza crônico-degenerativa.

A mortalidade no Brasil passou por importantes mudanças ao longo dos anos, com diminuição de algumas doenças infecciosas e o ressurgimento de outras, e com mudanças na frequência de algumas doenças e eventos não infecciosos, como as enfermidades cardíacas, a diabetes, as neoplasias e as mortes violentas.

Estas descobertas refletem a crescente concentração de mortes no grupo de idades mais avançadas e expressam a melhoria das condições de acesso aos serviços de Saúde e imunização, na saúde da infância, com a conseqüente redução de mortes de menores de idade. O envelhecimento da população brasileira resulta da combinação das altas taxas de fecundidade das décadas passadas e da diminuição da mortalidade em idades mais jovens observada também nas últimas décadas. Ou seja, as pessoas estão morrendo mais tarde, ou melhor, vivendo mais.

Segundo Moreira¹³ a ampliação da idade cronológica é uma compreensão bastante simplista do envelhecimento populacional:

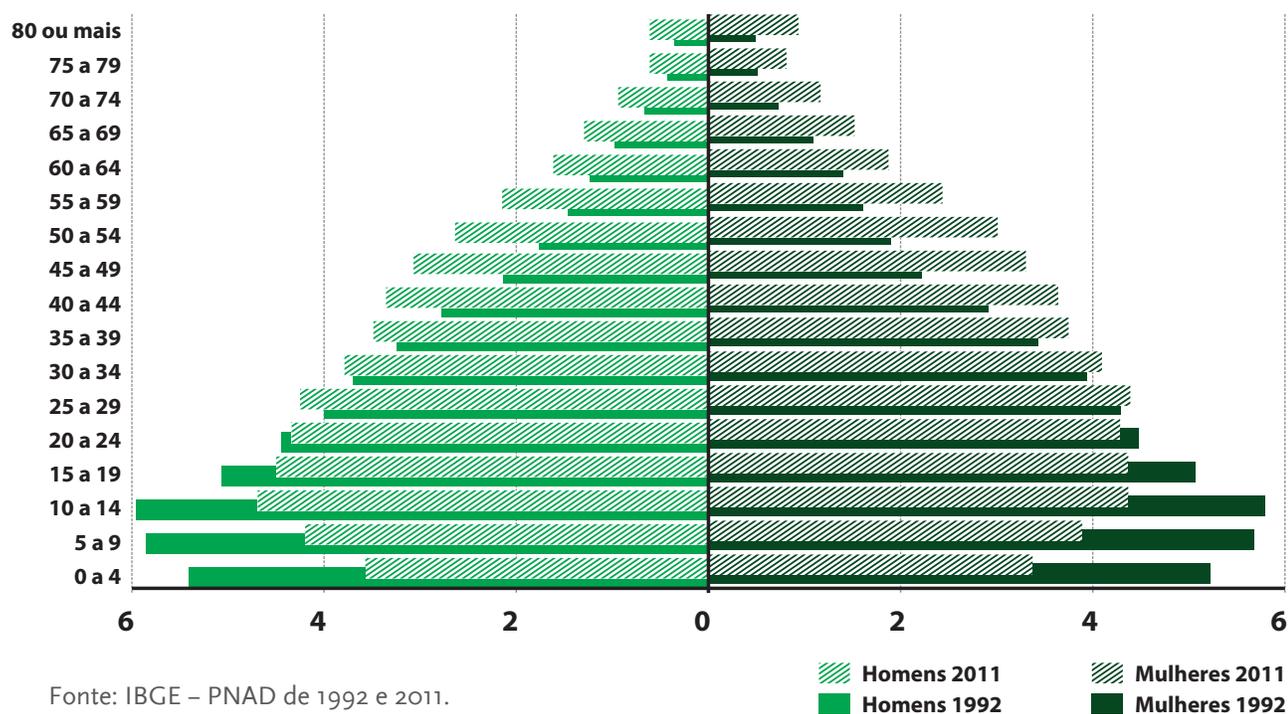
“Individualmente a idade de uma pessoa é mensurada pelo intervalo de tempo transcrito entre a data atual e a de seu nascimento. A sobrevivência média individual varia não só de pessoa para pessoa como também ao longo do tempo. Sendo amplas as diferenças individuais no processo de envelhecimento individual, enquanto processo biológico de declínio e deterioração, centrar o processo de envelhecimento na ampliação da idade cronológica oferece apenas uma compreensão superficial. Assim como o sexo, a cor e a classe social, a idade é um dos mais importantes fatores de diferenciação social e uma característica essencial da vida sociocultural de qualquer sociedade, variando entre as populações a percepção sobre o envelhecimento, o valor e o papel do idoso.

É usual, em demografia, definir os 60 ou 65 anos como o limiar que define a população idosa, sendo, portanto, considerada como população idosa a população de 60 ou de 65 anos e mais.

Em demografia, entende-se por envelhecimento populacional o crescimento da população considerada idosa em dimensão tal que, de forma sustentada, amplia a sua participação relativa no total da população. A ampliação do peso relativo da população idosa pode se dever a uma redução do grupo etário jovem, em consequência da queda da fecundidade, configurando o que se denomina envelhecimento pela base. Alternativamente, o aumento do percentual da população idosa pode ocorrer por um crescimento do grupo idoso. Independente de eventual redução da população jovem, em razão da queda da mortalidade, definindo o envelhecimento pelo topo. Há uma estreita relação entre os estágios da transição demográfica e o processo de envelhecimento populacional. Desta forma, o período de declínio da fecundidade marca o início do processo de envelhecimento populacional pela base, sendo típico daqueles países onde os níveis de fecundidade são relativamente elevados, desempenhando esta o papel principal na evolução do envelhecimento. A queda de mortalidade, como determinante do envelhecimento pelo topo, define a consolidação do processo de envelhecimento e é próprio das populações que já atingiram a maturidade demográfica, nas quais os níveis de fecundidade já são baixos e os níveis de mortalidade continuam a declinar, concentrando-se, entretanto, entre as faixas etárias mais elevadas, pois os níveis de mortalidade infantil e infanto-juvenil já atingiram os limites passíveis de redução e os dos adultos estão fortemente associados a comportamentos difíceis de serem mudados.”

Os resultados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2011 confirmam a tendência demográfica em curso no País desde os anos de 1970: desaceleração no ritmo de crescimento de sua população e mudanças expressivas em sua estrutura etária, no sentido de seu envelhecimento. O envelhecimento populacional significa uma alteração na proporção da população dos diversos grupos etários no total da população. Por exemplo, em 1940 a população idosa representava 4,1% da população total brasileira, e passou a representar 12,1% em 2011. O contingente, em valores absolutos, aumentou de 1,7 milhões para cerca de 23,5 milhões no mesmo período. Por outro lado, diminuiu a proporção da população jovem. A população menor de 20 anos passou a apresentar uma diminuição no seu contingente em termos absolutos e relativos. Esta tendência acentuar-se-á nas próximas décadas segundo as análises realizadas (Figura 3.15).

Figura 3.15: Distribuição por sexo e idade da população – Brasil, 1992 e 2011.



As altas velocidades da queda da fecundidade e da mortalidade resultam, também, em mudanças rápidas na distribuição etária da população brasileira no sentido do seu envelhecimento. Isso significa uma mudança nos pesos dos diversos grupos etários no total da população.

Do ponto de vista demográfico, o envelhecimento populacional é o resultado da manutenção, por um período de tempo razoavelmente longo, de taxas de crescimento da população idosa superiores às da população mais jovem. Isto implica uma mudança nos pesos dos diversos grupos etários no total da população. Além do envelhecimento da população total, a população idosa também envelheceu. A proporção da população “mais idosa”, de 80 anos ou mais, está aumentando também, alterando a composição etária no próprio grupo. Sua participação na população brasileira passou de 0,9% para 1,7%, entre 1992 e 2011. Embora o per-

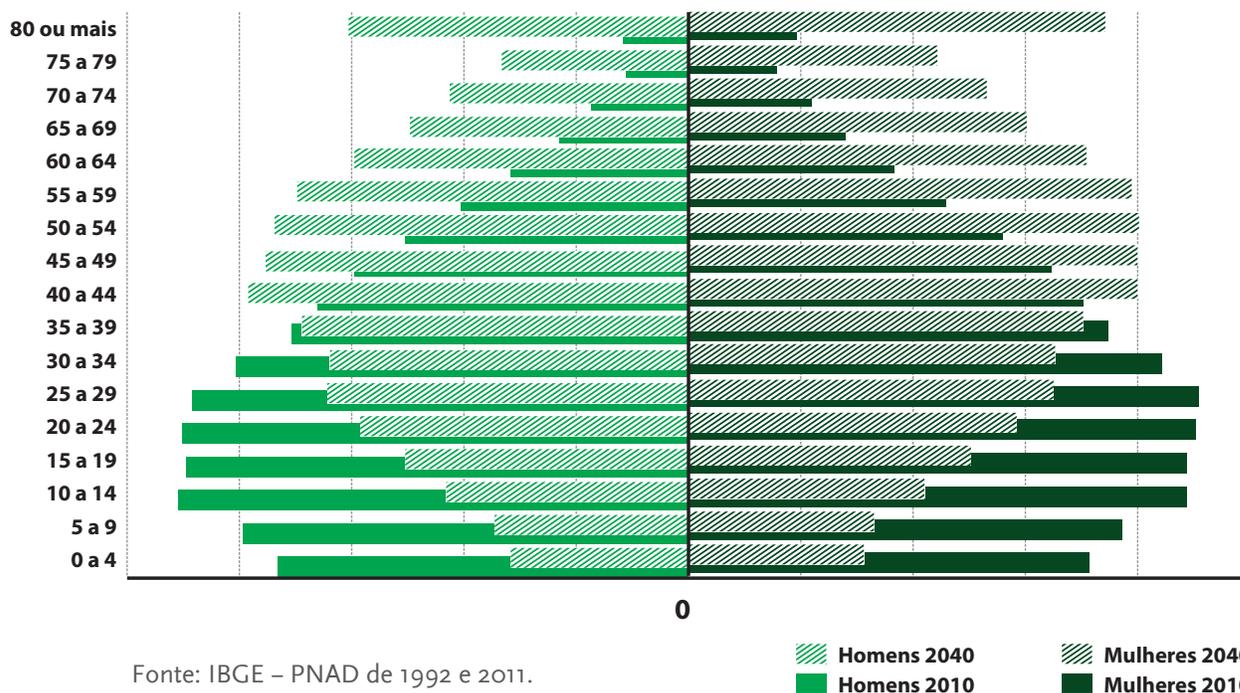
centual seja baixo, fala-se de 3,2 milhões de pessoas com 80 anos ou mais. Isso leva a uma heterogeneidade do segmento idoso, pois este passa a incluir pessoas de 60 anos a pessoas com mais de 100 anos de idade.

Nota

Em termos de políticas públicas, pode-se esperar um aumento na demanda por cuidados de longa duração e por serviços de Saúde, além de requerer pagamentos de benefícios previdenciários e assistenciais por um período de tempo maior.

As perspectivas vislumbradas nas projeções mencionadas são de um acelerado envelhecimento populacional, como mostra a Figura 3.16.

Figura 3.16: Distribuição por sexo e idade da população – Brasil, 2010 e 2040.



Uma nova e fundamental dimensão que vem sendo incorporada à análise sobre a esperança de vida é a análise acerca da **incapacidade**. O aumento da esperança de vida faz com que uma porcentagem maior de pessoas viva parte de sua vida em estado de incapacidade física ou mental. Interessa saber quantos anos de vida ativa uma pessoa terá, quer dizer, sem incapacidade.

Assim importa analisar não somente o aumento da esperança de vida, mas também, como está se dando o processo de envelhecimento.

Concluído o processo do envelhecimento é muito mais amplo do que uma modificação de pesos de uma determinada população, dado que altera a vida dos indivíduos, as estruturas familiares, a sociedade etc. Altera, também, a demanda por políticas públicas e a pressão pela distribuição de recursos na sociedade. Por isso, suas consequências têm sido, em geral, vistas com preocupações por acarretarem pressões para transferência de recursos na sociedade, colocando desafios para o Estado, o mercado e as famílias. Assim o enfrentamento das questões trazidas por este tema se constitui em um dos desafios do século XXI.

2. Estruturas familiares e o novo papel social da mulher

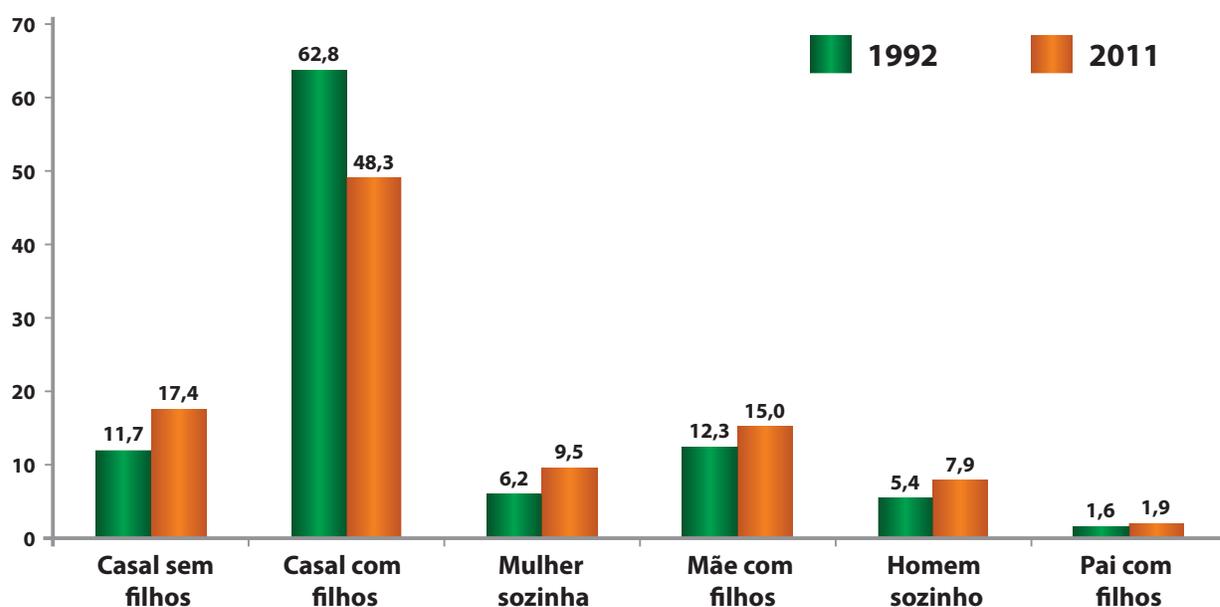
Estão em curso três revoluções que afetam, profundamente, as estruturas familiares e, conseqüentemente, a social:

- revolução contraceptiva: dissociação da sexualidade da reprodução;
- revolução sexual, principalmente, para as mulheres: separação entre sexualidade e casamento;
- revolução no papel social da mulher e nas relações de gêneros tradicionais: homem provedor *versus* mulher cuidadora.

Essas revoluções estão em curso em quase todo o mundo desenvolvido e, também, no Brasil. Discute-se a seguir, brevemente, o impacto das segunda e terceira revoluções já mencionadas nos arranjos familiares à luz dos resultados da PNAD de 2011.

O arranjo familiar predominante no Brasil é o do tipo casal com filhos, mas essa predominância vem decrescendo ao longo do tempo. Constituíam 62,8% do total de 17 arranjos em 1992 e passaram a constituir 48,3% em 2011 (Figura 3.17). Neste ano, 7,4 milhões de famílias brasileiras encontravam-se nessa categoria.

Figura 3.17: Distribuição percentual dos arranjos familiares brasileiros pelo tipo de arranjo.



Fonte: IBGE/PNAD de 1992 e 2011. Elaboração IPEA.

A redução na proporção de casais com filhos tem sido compensada pelo aumento das famílias constituídas por casais sem filhos e das monoparentais, principalmente as chefiadas por mulheres e de homens morando sozinhos. No segundo tipo de arranjo, o mais frequente era o do tipo mãe com filhos, que também cresceu no período. Passou de 12,3% para 15,0%. No total de arranjos brasileiros, a proporção dos formados por homens morando sozinhos cresceu de 3,7% para 6,2%. Ou seja, 3,7 milhões de homens brasileiros viviam sozinhos. A proporção para os domicílios formados por mulheres sozinhas passou de 3,9% em 1992 para 6,7% em 2011 e significa 4,0 milhões de mulheres vivendo sozinhas.

O aumento da proporção de domicílios chefiados por mulher guarda estreita relação com o aumento da participação feminina no mercado de trabalho. Esses fatores provocaram algumas mudanças nas características dos domicílios brasileiros, alterando as relações tradicionais de gênero: mulher cuidadora e homem provedor.

Sintetizando, a família brasileira está mudando e a mulher é uma das grandes responsáveis por isso. Ela, hoje, está assumindo novos papéis sociais, como o de provedora, dada a sua participação no mercado de trabalho, mas ainda mantém os papéis tradicionais, como o de responsável pelas tarefas domésticas e cuidados com os membros dependentes.

Por outro lado, a atual legislação previdenciária é ainda baseada nos papéis tradicionais, o

que leva as mulheres se aposentarem mais cedo que os homens, aproximadamente cinco anos, embora vivam sete anos a mais. Além disso, ela permite o acúmulo dos benefícios previdenciários, como o de pensão por morte e o da aposentadoria.

Aproximadamente 11% das mulheres brasileiras de 60 anos e mais recebiam, em 2011, o benefício da aposentadoria e o da pensão por morte. Embora a legislação estabeleça que isso possa ocorrer tanto para homens quanto para mulheres, elas são as maiores beneficiárias, pois vivem mais e se recasam menos quando viúvas.

A inserção crescente das mulheres nas atividades econômicas levará, a um futuro próximo, que mais mulheres passem a receber o benefício devido ao seu trabalho/contribuição. Isso pode resultar no crescimento da proporção de mulheres recebendo duplo benefício. No entanto, as mudanças nos arranjos familiares, especialmente na nupcialidade (separações), podem implicar uma redução da demanda por pensões por morte. Além disso, a queda da fecundidade, ou melhor, da maternidade, também deve ser considerada uma reestruturação dos sistemas de previdência social. O sistema vigente que estabelece prazos diferenciados de trabalho/contribuição para o recebimento do benefício entre homens e mulheres tem como um dos objetivos compensar as últimas pelo tempo passado na maternidade.

Todas essas mudanças requerem repensar as formas (tempo, alíquota) de contribuição por parte das mulheres, os tradicionais benefícios (duplo ou não), o valor das pensões por morte (igual ao benefício do cônjuge ou fração deste) e fazer adaptações, diante da nova realidade das famílias com mais de um provedor, das mulheres que mesmo casadas não têm filhos etc.

Neste sentido e diante do quadro apresentado de um envelhecimento populacional com grande diferencial de gênero é que o tema dos arranjos familiares se coloca também no centro das atenções e preocupações que precisam ser enfrentadas neste século.

Referências

- 1 – Carvalho JAM, Sawyer DO, Rodrigues RN (1998) Introdução a alguns conceitos básicos e medidas em demografia. São Paulo. 2 Ed. Associação Brasileira de Estudos Populacionais.
- 2 – IBGE (2011) Censo demográfico 2010: características da população e dos domicílios. Resultados do universo, Rio de Janeiro. IBGE.
- 3 – IBGE (2010) Síntese de indicadores sociais: uma análise das condições de vida da população, Estudos e pesquisas informação demográfica e socioeconômica. IBGE.
- 4 – Ministério da Saúde (Brasil) (2011) Saúde Brasil 2010: uma análise da situação de saúde e de evidências selecionadas de impacto de ações de vigilância em saúde, Brasília. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação em Saúde.
- 5 – von Manteuffel H. Jornal O Globo, 18 out. 2012.
- 6 – Singer P (1980) Migrações internas: considerações teóricas sobre seu estudo, in: H. A. d. Moura (Ed) Migração interna: textos selecionados. Fortaleza. Banco do Nordeste do Brasil – BNB, Escritório Técnico de Estudos Econômicos do Nordeste, 211-244.
- 7 – Berquó E (2001) Evolução demográfica, in: IS et al. (Ed.) Brasil um século de transformações. São Paulo. Companhia das Letras, 14-38.
- 8 – Ervatti LR. (2003) Dinâmica migratória no Estado do Rio de Janeiro na década de 90: uma análise mesorregional, Rio de Janeiro, Escola Nacional de Ciências Estatísticas, 155.
- 9 – Brito F (2009) As migrações internas no Brasil: um ensaio sobre os desafios teóricos recentes. Disponível em: <http://www.abep.nepo.unicamp.br/docs/anais/outros/> (acessado em 30 Out 2012).
- 10 – Oliveira LAP, Oliveira ATR (Eds) (2011) Reflexões sobre os deslocamentos populacionais no Brasil Rio de Janeiro. IBGE.

11 – Sassen S (1998) As cidades na economia mundial. São Paulo. Studio Nobel.

12 – OIM (2010) Perfil Migratório do Brasil 2009. Disponível em: http://www.mte.gov.br/trab_estrang/perfil_migratorio_2009.pdf (acessado em 25 Set 2012).

13 – Moreira M (2000) Envelhecimento da população brasileira: aspectos gerais, mimeo Brasília. Rede Interagencial de Informação para a Saúde. Organização Panamericana da Saúde, 349.

Módulo 4

Análise das Desigualdades em Saúde

Walter Massa Ramalho

Universidade de Brasília (UnB), Faculdade da Cielândia, Brasília/DF, Brasil.

Elisabeth Carmen Duarte

Universidade de Brasília (UnB), Faculdade de Medicina, Brasília/DF, Brasil. Organização Pan-Americana da Saúde (Opas), Brasília/DF, Brasil.

Introdução

O presente módulo tratará de aspectos teóricos e conceituais das desigualdades em Saúde e a sua aplicação ao caso brasileiro, valorizando – sempre que possível – o uso das bases de dados dos Sistemas Nacionais de Informação em Saúde e a discussão de exemplos práticos. Um aplicativo computacional será ainda apresentado, com o propósito de automatizar os cálculos utilizados normalmente para os estudos das desigualdades em Saúde, facilitando ainda mais a possibilidade de atender oportunamente às demandas dos gestores nos diversos cenários do SUS. Os objetivos deste módulo são:

- Conhecer o referencial histórico, os conceitos básicos, as teorias mais relevantes e a literatura pertinente para o entendimento de desigualdades e iniquidades em Saúde.
- Identificar, estimar e interpretar os indicadores de Saúde e de utilização dos serviços de Saúde úteis para a análise de desigualdades em Saúde.
- Identificar, estimar e interpretar os indicadores socioeconômicos úteis para a análise de desigualdades em Saúde.
- Identificar, estimar e interpretar as principais medidas de desigualdades em Saúde.



ASIS são processos contínuos e estratégicos, de análise e síntese, que permitem descrever, explicar e avaliar a tríade saúde-doença-atenção em uma população e contexto definidos, tendo em conta os seus determinantes sociais. Seu objetivo é criar informações e conhecimentos válidos sobre a situação de saúde de uma população em determinado território/contexto a fim de informar a decisão em Saúde pública. Dessa forma, observamos que no próprio conceito de ASIS é identificada a necessidade de levar em conta os determinantes sociais na compreensão dos fenômenos ligados à saúde em populações humanas, identificando a relevância dessa temática para o gestor em saúde.

Não é recente a notícia de que pobres adoecem mais, adquirem mais incapacidades, morrem mais precocemente e tem menos acesso aos serviços de Saúde de qualidade, do que as pessoas com melhores níveis socioeconômicos. Porém, talvez mais importante, é o fato de que essas desigualdades refletem mais frequentemente diferenças na distribuição das carências e restrições entre pobres e ricos, do que apenas diferenças em preferências individuais^{1,2,3,4,5,6}. Essas diferenças tendem, então, a ser vistas como mais do que simples desigualdades, e são

interpretadas como o resultado de relações sociais injustas – ou seja, refletem o que denominamos de “iniquidades” em Saúde⁷. Consequentemente, na medida em que uma sociedade reconhece como injustas essas desigualdades, e sente-se incomodada com elas, emerge a necessidade de abordar esse tema no âmbito das ASIS e da gestão em Saúde.

Unidade 1

Desigualdades em Saúde: Contexto e Conceito

Nesta unidade visitaremos brevemente marcos históricos selecionados sobre o pensamento referente às desigualdades em Saúde e a contextualização da abordagem das desigualdades sociais e em Saúde no Brasil, bem como o conceito e as diferenças adotados nos estudos da temática.

Leitura recomendada

Barbosa-Silva, J e Barros, M B A. Epidemiologia e desigualdade: notas sobre a teoria e a história. Rev Panam Salud Publica/Pan Am J Public Health 12(6), 2002.

Kawachi I, Subramanian SV & Almeida-Filho. A glossary for health inequalities. J Epidemiol Community Health. 2002. 56:647-52

Braveman, P & Gruskin, S. Defining equity in health J Epidemiol Community Health 2003;57:254–258

1. Contextualizando as desigualdades

As reflexões sobre as desigualdades entre os seres humanos devem ter surgido com a filosofia na **Grécia dos séculos VI e VII a.C.**⁸. Aristóteles, um dos primeiros a refletir sobre as desigualdades, apresentou princípios de igualitarismo que justifica os desníveis existentes entre os seres humanos como naturalmente determinados, já que alguns estariam destinados a comandar e outros, a obedecer. Esse pensamento legitima as práticas sociais vigentes. Mais especificamente no campo da Saúde, Hipócrates opôs-se ao pensamento mágico e supersticioso de sua época, e defendeu que as doenças estavam relacionadas ao meio ambiente, ao clima, à determinada raça e à alimentação.

Posteriormente, Luis Villermé e William Farr são considerados os precursores das aplicações da Epidemiologia na discussão de grupos de maior risco de adoecer e morrer. Luis Villermé, entre **1826 e 1840**, analisa a mortalidade, segundo os setores da Cidade de Paris, observa as diferenças intraurbanas de saúde, estuda operários das indústrias de algodão e vincula a pobreza às enfermidades. Em **1839**, William Farr organiza o primeiro registro de mortalidade da Inglaterra,

e a partir de **1860** desenvolve estudos que representaram a primeira tentativa de estratificação social e que apontaram excessos de mortalidade entre homens mineiros dos distritos da Inglaterra. Também Snow, em **1854**, estudando a epidemia de cólera em Londres, apontou diferenciais sociais na incidência da doença, com excessos afetando a classe trabalhadora, e relacionou essa distribuição não homogênea às condições de moradia dessa população proletária.

Além dos estudos de Snow, as descobertas de Pasteur e Koch reacenderam as teorias contagionistas como explicação hegemônica para a causalidade das doenças, aderindo a um modelo biologicista e monocausal: surge a era bacteriológica. Esse momento, por um lado derrota os conceitos fantasmagóricos dos miasmas, mas por outro ofusca o entendimento da determinação social das doenças.

Conseqüentemente, o período de **1910 a 1960** foi marcado por uma fase “silenciosa” no que se refere à abordagem social da doença. No final do século XIX e nas primeiras décadas do século XX, o processo clássico de transição epidemiológica instala-se no Hemisfério Norte. Observa-se expansão econômica, melhoria das condições de vida, aumento da expectativa de vida, progressiva substituição das doenças infecciosas e parasitárias pelas doenças degenerativas e doenças provocadas pelo ser humano. Surge, então, a identificação da necessidade de dar conta de uma maior complexidade dos processos de causalidade das doenças, motivando algumas exceções de estudos que abordaram a determinação social das doenças nessa fase silenciosa. São exemplos de estudos dessa época os de Goldberger, que discute a relação entre a pelagra e a pobreza e a alimentação dos pobres; os estudos de Stevenson, que cria o primeiro sistema de classificação social baseado em grupos ocupacionais na Inglaterra; e os estudos de Mayor Greenwood, que desenvolve a estatística para estudos epidemiológicos e buscavam associação entre câncer e estratificação social.

Entre **1970 e 1980**, principalmente na América Latina, e em especial no México, Brasil e Argentina, o movimento da Saúde pública tenta resgatar a discussão da determinação social do processo saúde/doença. Os pensadores desse período procuram identificar meios para transformar as realidades sociais, econômica e política e reenquadram o processo saúde-doença como fenômeno socialmente determinado. Nesse contexto, importantes críticas dirigidas à Epidemiologia Clássica dá lugar ao desenvolvimento de uma Epidemiologia Social.

Na década de **1980**, surge ainda o polêmico Relatório Black (Douglas Black), um estudo sobre as desigualdades sociais em Saúde na Europa, que foi desautorizado pelo novo governo britânico.

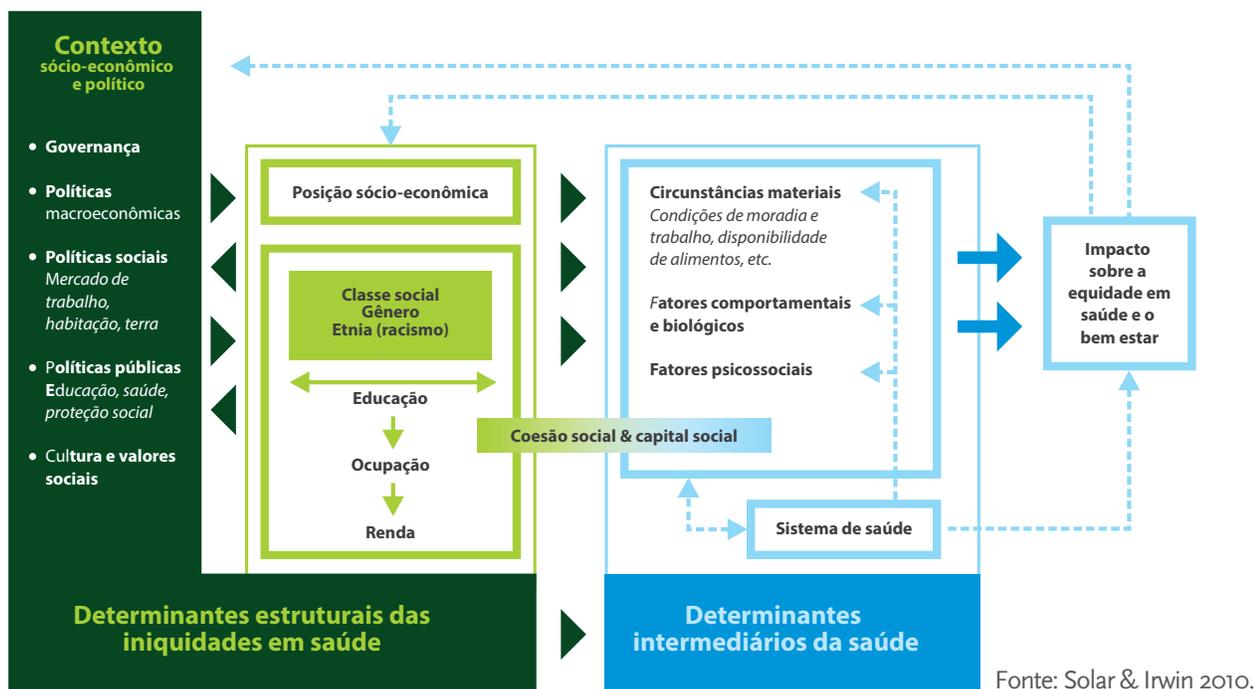
Em geral, nos anos seguintes, a abordagem da equidade em Saúde torna-se cada vez mais popular, sendo alvo de grande desenvolvimento de instrumental metodológico, elaboração de indicadores compostos, lançamento de programas (governamentais e não governamentais) para monitoramento e enfrentamento das iniquidades em Saúde, entre outras iniciativas no mundo. Por exemplo, o

desenvolvimento de abordagens, tais como, as análises multiníveis para dar conta da complexidade da determinação social das doenças e das desigualdades têm ampliado sobremaneira as possibilidades analíticas, que respeitam os diferentes níveis de determinação como o marco (Figura 4.1) adotado no relatório final da Conferência Mundial sobre Determinantes Sociais da Saúde⁹.

Essa mudança importante, dada a abordagem da determinação social da doença, pode ter sido influenciada também pelo aumento da sensibilidade acerca da temática e da demanda por estudos a partir dos tomadores de decisão, instituições doadoras de recursos financeiros, organizações governamentais e não governamentais; maior disponibilidade de dados (maiores coberturas e maior escopo de variáveis) por meio do aumento de interesse pela realização de inquéritos demográficos e de Saúde nos países; maior capacidade computacional dos novos computadores e *softwares* capazes de analisar maior volume de dados, com maior eficiência (mais rapidamente, com maior complexidade computacional e menor custo), entre outros⁷.

No Brasil, o monitoramento e o enfrentamento das **desigualdades sociais** têm sido pautados nos discursos das mais variadas áreas, tais como a Economia, a Educação, a Saúde, a Habitação, entre outras. Essa pauta vem sendo fortemente influenciada por uma agenda política estratégica direcionada, há pouco mais de uma década, para crescimento econômico vinculado à distribuição de renda, à redução das desigualdades sociais e da pobreza extrema no Brasil.

Figura 4.1: Marco conceitual dos determinantes sociais da Saúde.



Contemplamos no Brasil redução importante da miséria e da pobreza extrema nos últimos dez anos. Segundo o Instituto de Pesquisa Econômica Aplicada¹⁰, a proporção de pessoas extremamente pobres caiu de 25,4% em 2001 para 13,5% em 2009. Além disso, e mais importante, deve ser ressaltada a inédita redução das desigualdades sociais nesse mesmo período: a renda dos 50% mais pobres no Brasil cresceu quase seis vezes (580%) mais que a renda dos 10% mais ricos¹¹. O coeficiente de Gini, um indicador de distribuição de renda, havia permanecido estável no Brasil desde os princípios da década dos anos de 1980, e oscilava em torno de 0,60. Desde 2001 esse indicador tem reduzido sistematicamente chegando a 0,543 em 2009; quase um quarto dessa redução tem sido atribuído ao Programa Bolsa Família.

A temática do enfrentamento das desigualdades sociais ainda tem se destacado, no Brasil, na agenda ambiental, como elemento relevante para a melhoria da qualidade de vida, articulado ao conceito da “sustentabilidade” com um significado mais amplo de desenvolvimento. Debates presentes na Conferência das Nações Unidas sobre Desenvolvimento Sustentável (UNCSD), a Rio+20, identificam a promoção da saúde e a resolução dos problemas ambientais, como elementos inseridos em movimentos socialmente justos, economicamente viáveis, ecologicamente prudentes e politicamente emancipadoras¹², por “uma economia verde no contexto do desenvolvimento sustentável e da erradicação da pobreza”¹³.

No entanto, apesar dos avanços, o Brasil ainda figura entre os 12 países mais desiguais do mundo.

As desigualdades sociais refletem também nas condições de vida e de saúde do povo brasileiro. Isso porque os riscos e os recursos em Saúde podem também ser distribuídos de forma desigual a depender do acesso aos bens e serviços, e das posições ocupadas pelos indivíduos na sociedade. As ASIS, no que se refere às análises da distribuição dos eventos e dos recursos de Saúde em relação aos grupos sociais, têm sido úteis para mensurar tais desigualdades em Saúde.

A ASIS com foco nas desigualdades em Saúde tem, no Brasil, o apoio de diferentes Sistemas de Informação em Saúde e de inquéritos e pesquisas em Saúde. Fontes essas que têm grande potencial por usualmente apresentarem ampla cobertura, acesso público e gratuito, oportunidade e universalidade dos dados.

Finalmente a superação das desigualdades em Saúde é exercício em três direções: uma questão teórica/conceitual, uma problemática metodológica e um desafio político¹⁴.

2. Marcos conceituais

Apesar de muitas vezes os termos equidade e igualdade estarem se referindo aos mesmos conceitos, e ambos partirem do princípio da pluralidade social, eles traduzem regras distintas para as suas aplicações. Desigualdade e igualdade são conceitos que se referem a conteúdos mensuráveis e valorados, a iniquidade ou a equidade são conceitos político-normativos, que expressam um compromisso moral de justiça social.

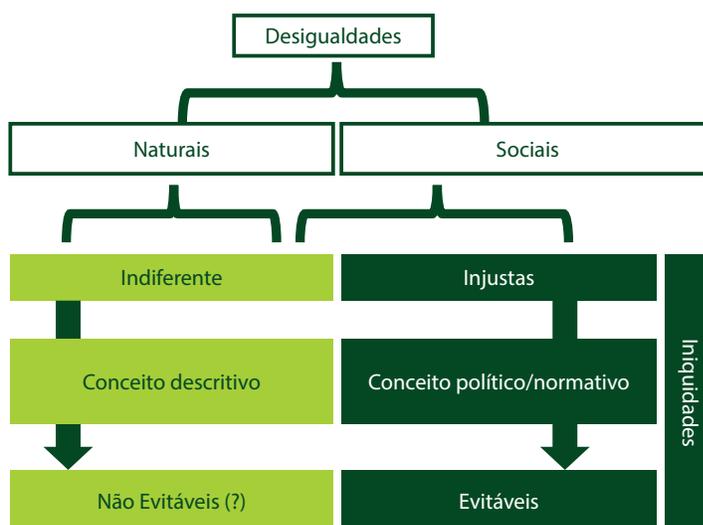
2.1 Igualdade e desigualdade

“Igualdade” significa a qualidade daquele ou daquilo que é igual e “desigualdade” significa qualidade daquele ou daquilo que apresenta diferenças. Assim, por princípio, a qualidade de ser igual ou desigual possui caráter descritivo apenas⁸. Desigualdade trata-se, portanto, de termo genérico usado para designar diferenças, variações ou disparidades nos desfechos de saúde de indivíduos e grupos de toda ordem¹⁵.

Segundo o Instituto de Saúde dos Estados Unidos, em seu plano estratégico para reduzir e finalmente eliminar as desigualdades em Saúde (2001), “desigualdades em saúde são diferenças na incidência, prevalência, mortalidade ou carga de doenças e de outros danos à saúde que ocorrem entre grupos populacionais”.

Entre as formas de classificar as desigualdades, pode ser adotada a categorização em desigualdades naturais e sociais em Saúde (Figura 4.2).

Figura 4.2: Classificação das desigualdades em Saúde – síntese.



Fonte: Silva & Barros 2002.

Desigualdades naturais em Saúde ocorrem quando há diferenças ou dessemelhanças nos desfechos em saúde entre os indivíduos devido a atributos tais como sexo, etnia, idade, força física, inteligência, habilidade inatas, entre outros⁸. As pessoas são distintas no sentido biológico, e fazem escolhas distintas, apresentando comportamentos distintos – ainda que façam parte do mesmo grupo social –, e todos esses fatos podem levar a desfechos em Saúde distintos. Por exemplo, no Brasil, é conhecido o fato de mulheres viverem, em média, seis anos a mais do que os homens. Também é conhecido o fato de elas usarem mais os serviços de Saúde, terem comportamentos mais cuidadosos em relação à própria saúde, e estarem menos expostas às mortes por causas violentas do que os homens no Brasil. Esses são alguns exemplos de desigualdades em Saúde entre homens e mulheres.

As **desigualdades sociais em Saúde** referem-se a diferenças, variações e disparidades nos padrões de saúde de grupos sociais distintos. Estão presentes quando são determinadas pela estrutura da sociedade na qual os indivíduos estão inseridos. As desigualdades sociais podem ser definidas como a distribuição desigual de bens materiais e não materiais e serviços entre os grupos sociais¹⁶. Como decorrência desse conceito, na área da Saúde, temos que a distribuição da saúde e da doença entre os grupos sociais também é uma forma de desigualdade social¹⁷.

As desigualdades naturais podem ou não definir as desigualdades sociais em Saúde. As desigualdades naturais propiciam o surgimento de desigualdades sociais quando uma sociedade escolhe os atributos de estratificação (sexo, etnia, idade, força física, inteligência, habilidade inatas, entre outros) como critério para atribuição de papéis sociais⁸. Esse é o caso das desigualdades no risco de morte precoce por causas violentas no Brasil. Isso porque se origina em papéis socialmente proibitivos para as mulheres e valorizados entre os homens, tais com os comportamentos mais agressivos e de exposição a riscos rotineiramente encorajados entre os homens e inibidos entre as mulheres. Nessa situação, uma desigualdade natural passa a ser uma desigualdade social.

Uma desigualdade social não implica, necessariamente, um julgamento moral, podendo ou não ser considerada injusta. Pois esta tem um sentido descritivo (não normativo) sem necessário julgamento de valor, na medida em que não necessariamente enuncia um “deve-se”⁸. Dessa forma, surge mais uma maneira de classificar as desigualdades sociais: elas podem ser consideradas injustas ou não (ou indiferentes). Quando as desigualdades sociais são consideradas injustas por uma sociedade, surge o conceito das iniquidades.

2.2 Equidade e iniquidade

Segundo Kawachi et al.¹⁵, iniquidades são diferenças nos níveis de saúde de grupos po-

pulacionais distintos socialmente (desigualdades sociais), consideradas injustas ou emanadas de alguma forma de injustiça.

Três características relevantes, quando combinadas, transformam uma mera variação ou desigualdade em Saúde em iniquidade social em Saúde. Elas são sistemáticas, produzidas socialmente (e, portanto modificáveis) e injustas^{2,18,19}. As iniquidades são sistemáticas porque as diferenças em Saúde não refletem uma distribuição aleatória, mas sim apontam para padrões sistemáticos de distribuição: por exemplo, a mortalidade e morbidade aumentam sistematicamente na medida em que se observa decréscimo do nível socioeconômico de uma população. As iniquidades são produzidas socialmente – em oposição a serem produzidas por diferenças biológicas entre os sujeitos. Essa característica implica que as diferenças socialmente determinadas não são fixas e, portanto, são modificáveis e evitáveis. Finalmente, e amplamente reconhecidas, as iniquidades são injustas, uma vez que são criadas por “arranjos sociais injustos” que ofendem a noção vigente de justiça¹⁸.

No entanto, existe debate na literatura sobre o uso da ideia de evitabilidade nesse contexto. Braveman & Gruskin²⁰ considera que por um lado a ideia de “ser evitável” apoia a noção amplamente compartilhada de justiça social do conceito das iniquidades e pode mobilizar a opinião pública, dos gestores e políticos para o seu enfrentamento. Porém, os autores mencionam que a evitabilidade é redundante, já que a noção de justiça implica noção de evitabilidade. Além disso, e mais relevante, os autores mencionam que algumas iniquidades podem ser extremamente desafiadoras em termos de seu enfrentamento, exigindo para tanto mudanças fundamentais nas estruturas econômicas e sociais. Dessa forma, os autores não recomendam que a evitabilidade constitua em uma medida de grau de iniquidade. Adicionalmente, tendo a evitabilidade como um atributo da iniquidade, os autores mencionam que caberia, então, perguntar: evitável por ações de quem? Pela população já sujeita às condições de desvantagem social? Pela comunidade como um todo? Pelo governo? Por organismos internacionais?

Evidentemente, paralelo ao conceito de iniquidade, o conceito de equidade também aparece com sua conotação de justiça social. Assim, um conceito mensurável e operacional de equidade em Saúde afirma que esta é a ausência de disparidades sistemáticas em Saúde (ou nos determinantes sociais em Saúde mais relevantes) entre grupos com diferentes níveis (vantagem /desvantagem) de condições sociais – ou seja, riqueza, poder e prestígio²⁰.

Os arranjos sociais injustos, que promovem as iniquidades em Saúde, têm seu conceito vinculado à justiça social e, portanto, vincula-se também aos direitos humanos. Segundo a Declaração Universal dos Direitos Humanos (Artigo 25) *“Everyone has the right to a standard of living adequate for the health and well being of himself and of his family, including food, clothing,*

*housing and medical care*²¹. Além disso, saúde é um recurso indispensável para o alcance de outros objetivos – educação, emprego, moradia, entre outros – e é indispensável, portanto, para promover cidadania¹⁸.

Diferente das desigualdades, que tem caráter descritivo apenas, as iniquidades têm caráter político-normativo. É um conceito normativo na medida em que se refere a valores e normas aceitas socialmente; e é um conceito político porque é o resultado das relações de poder existentes em uma sociedade.

Segundo a International Society for Equity in Health, as iniquidades dizem respeito às diferenças sistemáticas e potencialmente remediáveis em um ou mais aspectos da saúde entre grupos populacionais definidos socialmente, economicamente, demograficamente ou geograficamente²².

Por expressar compromisso moral com justiça de uma sociedade, o conceito de iniquidades vincula-se a um momento histórico dessa sociedade e, portanto, depende: da teoria de justiça; da teoria de sociedade e das explicações para a etiologia das desigualdades observadas por uma sociedade¹⁵.

Fique por dentro

Desigualdades em Saúde:

- *Naturais*, ocorre quando há diferenças ou dessemelhanças nos desfechos em saúde entre os indivíduos devido a atributos, tais como: sexo, etnia, idade, força física, inteligência, habilidade inatas, entre outros.

- *Sociais*, ocorre quando há diferenças, variações e disparidades nos padrões de saúde de grupos sociais distintos. Essas não remetem necessariamente à ideia de injustiça (ou iniquidade), uma vez que tem um sentido descritivo (não normativo) sem necessário julgamento de valor, na medida em que não necessariamente enuncia um “deve ser”.

Iniquidades em saúde:

São diferenças nos níveis de saúde de grupos populacionais distintos socialmente (desigualdades sociais), consideradas injustas ou emanadas de alguma forma de injustiça, e que são desnecessárias e evitáveis.

Unidade 2

Modelos de Causalidade das Iniquidades em Saúde

O que explica as desigualdades (e as iniquidades) em Saúde? E como o setor Saúde pode ser um elemento promotor ou inibidor das desigualdades sociais em Saúde no Brasil? Agora que os conceitos estão sedimentados, trataremos a respeito das teorias relativas à etiologia das desigualdades em Saúde em uma sociedade. Essas teorias, apesar de vistas eventualmente como argumentos contraditórios, elas são na verdade complementares e, muito possivelmente incompletas, sobre a complexa rede de causalidade das desigualdades em Saúde em populações humanas. Nessa seção serão discutidas as principais teorias que tentam explicar a etiologia das desigualdades em Saúde, seus conceitos e complementaridades.

Leitura recomendada

Donaldson. *Health Services and Public Health*. *J Epidemiol Community Health* 2002;56:835-840.

Woodward & Kawachi. *Why reduce health inequalities*. *J Epidemiol Community Health* 2000;54:923-929.

Whitehead. M & Dahlgren, Göran. Concepts and principles for tackling social inequities in health: *Levelling up Part 1*. WHO. 2006.

1. A Interpretação das desigualdades em Saúde

Ou neomaterial explica a existência denexo causal de desigualdades em Saúde por meio do gradiente de estratificação social e a divisão social do trabalho. Os seus resultados são oriundos da combinação da exposição negativa, da falta de recursos financeiros e da sistemática ausência de investimentos por políticas públicas.

Segundo essa teoria, as diferenças em acesso aos bens e serviços – desde comida, abrigo e serviços de Saúde, até lazer, carro, telefone, internet, entre outros – com as decisões políticas e econômicas produzem doenças e desigualdades em Saúde¹⁵.

Figura 4.3: Esquema representativo da teoria material.



Fonte: Elaborado pelos autores.

1.2 A teoria psicossocial

Baseada em pesquisas sobre percepções sociais, segundo a teoria psicossocial, desigualdades são produzidas por estressores sociais que podem causar doenças direta ou indiretamente, e desigualdades em Saúde baseadas na forma relativa de como o indivíduo está inserido na sociedade, e não na forma absoluta de renda. Nesta abordagem, sentimentos de desvalorização e desvantagem pela percepção ao ambiente social ao qual está inserido criam conflitos e desfechos desfavoráveis em saúde, em consequência à insegurança (falta de controle) por estar em hierarquia socialmente inferior e/ou por viver em condições de relativa desvantagem socioeconômica.

Os efeitos diretos dos estressores sociais são aqueles que agem no sistema fisiológico humano, incluindo o desgaste físico causado pela exposição diária a circunstâncias adversas da vida.

Os efeitos indiretos são aqueles mediando por comportamentos e atitudes adversas (de risco) como consequência dos estressores sociais, tais como: fumo, bebida, violência, entre outros. E, por sua vez, esses comportamentos e atitudes ampliam a probabilidade de adoecer e morrer mais precocemente.

Figura 4.4: Esquema representativo da teoria psicossocial.



Fonte: Elaborado pelos autores.

1.3 A complementaridade da interpretação material e interpretação psicossocial das desigualdades em Saúde

Segundo Kawachi et al.¹⁵, toda privação material relevante para a vida diária parece ter algum significado psicossocial. Dessa forma, não existe nada de contraditório nessas teorias, e sim são mais bem vistas como explicações complementares (e não excludentes), portanto a interpretação recai sobre a interação entre aspectos biológicos e sociais, possibilitando diferentes abordagens na perspectiva de saúde populacional.

Figura 4.5: Esquema representativo da complementaridade das teorias material e psicossocial.



Fonte: Elaborado pelos autores.

2. A relação entre os serviços de Saúde e a saúde em populações

Obviamente, a atenção à saúde pode contribuir para melhorar o estado de saúde do indivíduo e da população. Os serviços de Saúde podem contribuir para a instituição de ações que visam a vigilância em saúde, a oferta de serviços preventivos, a melhoria da saúde ambiental, a avaliação das ações em Saúde e a redução de iniquidades em Saúde²³. Em especial, para ter um papel na redução das desigualdades e iniquidades em Saúde, é necessário um claro compromisso com a promoção da equidade na atenção à saúde.

A depender da forma com que a atenção à saúde é ofertada aos indivíduos e às comunidades, essa pode se tornar em um dos muitos fatores promotores ou inibidores das desigualdades sociais em Saúde. Isso porque acesso, oferta e uso desigual dos serviços de Saúde são características comuns aos sistemas de Saúde²³. A equidade no acesso à atenção à saúde de qualidade e a oferta adequada às necessidades é um modelo teorizado e perseguido, porém raramente alcançado em sua plenitude em nossos países. Pelo contrário, o que se observa, não raras vezes, é a oferta de serviços de Saúde em menor cobertura e de pior qualidade para populações com maiores necessidades, e vice-versa.

A percepção desse fenômeno – o fato de que a disponibilidade da atenção à saúde adequada tende a variar inversamente com a necessidade da população atendida – foi descrita por Tudor Hart em 1971, ficando conhecida como a lei da atenção inversa. Na verdade, é de se supor que os mesmos determinantes que contribuem com os arranjos sociais injustos também têm papel sobre a oferta e a qualidade dos bens e serviços e, entre eles, dos serviços de Saúde ofertados a essas populações em desvantagem social.

Sobre a lei da atenção inversa (the inverse care law)²⁴

A disponibilidade da atenção à saúde adequada tende a variar inversamente com a necessidade da população atendida.

As razões para essas desigualdades na atenção à saúde são complexas e precisam ser enfrentadas por meio de rigorosa análise do problema e adoção de políticas de Saúde e planejamento orientadas para sua solução, com clara intencionalidade de discriminação positiva em relação às necessidades de saúde²³. Somar evidências (medir) pode ajudar a atrair atenção para o problema, e o uso dessas evidências para orientar intervenções pode propiciar maior eficiência no enfrentamento das iniquidades em Saúde decorrentes da não equidade na atenção à saúde.

Nesse sentido, Woodward & Kawachi²⁵ discutem pelo menos quatro razões que justificam por que os serviços de Saúde devem se preocupar com as desigualdades em Saúde²⁵: primeiro por ser uma questão de justiça, segundo porque as desigualdades afetam toda a sociedade negativamente, terceiro porque as desigualdades são evitáveis, e finalmente porque as intervenções para reduzir as desigualdades em Saúde podem ser eficientes.

As quatro razões para que os serviços de Saúde se preocupem com as desigualdades em Saúde²⁵

a) Muitas desigualdades em Saúde são injustas e indesejáveis (= iniquidade)

As desigualdades injustas são consequência de uma distribuição desigual de determinantes sociais essenciais, que não dependem (vão além) do controle do indivíduo. Esse é o caso da oportunidade desigual no acesso aos serviços de Saúde de qualidade. Esta razão, discutida anteriormente, deveria ser suficiente para justificar por que os serviços de Saúde devem se preocupar com as desigualdades em Saúde. Isso porque justiça é um valor moral/ético das nossas sociedades e, portanto, desejável, independente de suas consequências.

Fazendo um paralelo, podemos lembrar que Rothman et al.²⁶ observa que críticos da Epidemiologia usam a melhoria na saúde de populações como argumento para fins sociais, e defendem a erradicação da pobreza com o argumento da saúde. No entanto, o autor destaca que o desejo de buscar a erradicação da pobreza tem justificativa em si mesma, já que é um valor que deve ser perseguido independente de suas consequências para a Saúde pública.

b) As desigualdades afetam toda a sociedade

Os grupos afetados negativamente pelas desigualdades em Saúde não se restringem àqueles em desvantagem social. Na verdade as desigualdades em Saúde afetam toda a sociedade. Por exemplo, a disseminação de doenças infecciosas, as consequências do abuso de drogas e álcool, a ocorrência de violências e crimes, entre outros, são eventos agravados pelas desigualdades em Saúde cujas consequências vão além do grupo especialmente afetados por essas desigualdades.

c) Algumas desigualdades são evitáveis

São evitáveis na medida em que se originam nas opções políticas identificáveis, exercidas pelos governantes. Por exemplo: a política tributária e trabalhista, financiamento previdenciário e da saúde, entre outras. Então as desigualdades em Saúde são, em princípio, sensíveis às intervenções.

d) Intervenções para reduzir desigualdades podem ser eficientes (custo-efetivas)

Apesar de pouco avaliados, é intuitivo pensar que políticas e programas de saúde coletiva para reduzir desigualdades em saúde podem ser custo-efetivos. Por exemplo, para a sociedade, como um todo, ampliar o acesso ao rastreamento de câncer cervicouterino entre mulheres de baixa renda pode repercutir positivamente no financiamento do setor Saúde no que se refere à redução de gastos com o tratamento do câncer em estágios avançados, nos recursos destinados a atender aos impactos sociais dessas mortes evitáveis, entre outros benefícios para a sociedade.

Unidade 3

Medindo as Desigualdades

Partindo do princípio de que as desigualdades sociais em saúde podem refletir injustiças não toleradas por uma sociedade, cabe perguntar a essa sociedade: O que gostaríamos de ver equitativamente distribuído? A resposta seria a saúde! Mas como medir saúde?

E se estamos falando de desigualdades em saúde entre diferentes grupos sociais, como podemos classificar os indivíduos segundo grupos sociais?

Nesse capítulo iremos discutir brevemente a medição dos indicadores de saúde e de serviços de saúde (detalhamento desse tema foi discutido anteriormente nesse curso).

Leitura recomendada

Kawachi, Subramanian, Almeida-Filho. A glossary for health inequalities. *J Epidemiol Community Health* 2002;56:647-652

O'Donnell O, Doorslaer E, Wagstaff A & Lindelow M. Analyzing Health Equity Using Household Survey Data. The World Bank, Washington DC, 2008, www.worldbank.org/analyzinghealthequity

1. Desigualdades de que?

Como sabemos, saúde é mais facilmente medida com base na sua ausência, Assim, são desfechos desfavoráveis em saúde: número de anos de vida perdidos, número de anos de vida perdidos por incapacidade, risco de morte segundo causas e segundo idade (morte precoce), risco de adoecimento segundo causas (evitáveis), entre outros. Além disso, por saber que os serviços de Saúde de qualidade promovem saúde no nível individual e coletivo, interessa também ter os indicadores associados à qualidade dos serviços de Saúde equitativamente distribuídos em uma população. Segundo Donabedian²⁷, esses indicadores de Saúde podem ser classificados em: indicadores de estrutura, processo e resultado, como descritos a seguir.

Indicadores de estrutura

Referem-se às condições físicas, humanas e organizacionais em que o cuidado se dá. Parte do pressuposto que bons indicadores de estrutura ampliam as chances de se ter bons indicadores de processo e de se ter um cuidado com qualidade. Exemplos: presença e adequação de equipamentos, de área física, de instalações, de insumos e de recursos humanos.

Indicadores de processo

Dizem respeito à inter-relação entre prestador e receptor dos cuidados. Referem-se ao “como” fazer. Parte do pressuposto que os indicadores de processo são influenciados pelos indicadores de estrutura, e que bons indicadores de processo tendem a influenciar positivamente os indicadores de resultados. Exemplos: prática e diretrizes para solicitação de exames, escolhas terapêuticas baseadas em evidência, acolhimento e humanização do paciente, entre outros.

Indicadores de resultado

Produto final da atenção prestada, considerando saúde, satisfação de padrões e de expectativas. Reflete a mudança no estado de saúde do paciente e na população que pode ser atribuída ao cuidado. Parte do pressuposto que bons indicadores de estrutura e de processo contribuíram com melhores desfechos na saúde individual e coletiva. Exemplos: taxas de mortalidade infantil, taxas de letalidade por doenças imunopreveníveis, taxas de incidência de sífilis congênita, entre outros.

2. Desigualdades segundo o quê?

Voltando ao princípio de que as desigualdades sociais em Saúde podem refletir injustiças não toleradas por uma sociedade, a estratificação dos grupos exige um referencial teórico. Nesse sentido é importante discutir os critérios de elegibilidade e medição dos indicadores socioeconômicos, tais como: gênero, raça e etnia, local de residência, deficiências físicas, escolaridade, renda, posição na sociedade entre outros.

As hipóteses da renda absoluta *versus* relativa

Hipótese da renda absoluta

Segundo Kawashi et al.¹⁵, pobreza absoluta refere-se à incapacidade de acessar necessi-

dades humanas básicas, tais como alimentação, moradia, e saúde. O conceito operacional de pobreza absoluta geralmente é o da linha da pobreza, que define o indispensável para atender minimamente às necessidades humanas.

Esses autores também descrevem a hipótese da renda absoluta, que afirma que a saúde do indivíduo depende de sua (e apenas de sua) renda. Por exemplo, a renda média familiar per capita, se o indivíduo está ou não abaixo da linha de pobreza ou da linha de indigência. Mas, e se todos passam a ganhar mais, menos você? É impossível imaginar que a saúde das pessoas não seria afetada, especialmente se pensarmos que o padrão de consumo considerado necessário é bem provável que também mude. Por exemplo, se bens e serviços de luxo passar a serem incorporados entre as necessidades do dia a dia (como já aconteceu anteriormente com a internet)?

Hipótese da renda relativa

Por contraste, segundo Kawashi et al.¹⁵ (2002), pobreza relativa refere-se a padrões que existem na sociedade em algum de seus grupos. Por exemplo, a linha de pobreza nesse caso é definida pela distribuição da pobreza na sociedade. Sobre a hipótese da renda relativa, os autores afirmam que a saúde do indivíduo depende não apenas de sua renda, mas também de sua posição (de seu *rank*) quanto à distribuição da renda da sociedade em que vive. Refere-se à magnitude da distância entre os mais ricos e os mais pobres. Então, se esta hipótese for verdadeira, é esperado que a saúde de uma pessoa com baixa renda seja pior em sociedades mais desiguais do que em sociedades mais equitativas. Vários estudos mostram essa relação. Apenas para descrever que um estudo ecológico, analisando os efeitos das desigualdades sociais (Gini municipal) e da pobreza absoluta (renda média per capita) nas taxas municipais de homicídios em homens jovens no Brasil, apontou que as primeiras são melhores preditoras desses eventos (2012). Exemplo de medidas relativas de renda são: Razão de *decis* (renda dos 10% superior/renda dos 10% inferior ou outro ponto de corte) e coeficiente Gini, Curva de Lorenz.

Que indicador escolher?

Depende:

1. Plausibilidade: renda, escolaridade etc.
2. Hipótese da renda absoluta *versus* hipótese da renda relativa.
3. Disponibilidade do indicador no nível de agregação da análise.
4. Adequada variabilidade do indicador entre as unidades de análise.

A escolha dos indicadores tem peso considerável nos resultados dos estudos de desigualdades sociais em Saúde. Como os indicadores socioeconômicos são utilizados na estratificação, devem possuir o poder discriminante da população, agrupando indivíduos semelhantes e separando grupos distintos. Devem ainda estar respaldados por contexto teórico e metodológico robusto para defesa da hipótese das desigualdades.

Como as desigualdades sociais em Saúde são multifacetadas, um conjunto grande de indicadores serve para os mais diversos propósitos, mas devem ser priorizadas as seguintes características²⁸:

1. refletir a dimensão socioeconômica das desigualdades em Saúde – ou seja, as variáveis utilizadas devem estar amparadas *a priori* por uma reflexão de pertinência causal, e não buscar nos resultados suposições sobre os nexos encontrados *a posteriori*;

2. incorporar informações sobre todos os grupos populacionais definidos pelo indicador – busca-se que o indicador possa aferir as variações entre todos os grupos estudados. Nem todos os indicadores possuem esta qualidade por permitirem comparações entre os grupos extremos;

3. ser sensível às mudanças na distribuição e tamanho da população ao longo da escala socioeconômica – ou o poder discriminante do indicador, ou a qualidade com que pode separar a heterogeneidade entre os grupos, e juntar a homogeneidade intra-grupos. Observamos por exemplo, que alguns indicadores tradicionais, em que facilmente as desigualdades eram visíveis, têm perdido a sua capacidade discriminatória entre os grupos sociais, como abastecimento de água tratada ou coleta de lixo, serviços hoje disponíveis para uma grande parcela da população nas diversas regiões do Brasil.

3. Principais medidas de desigualdades

A escolha dos indicadores tem peso considerável nos resultados dos estudos das desigualdades sociais em Saúde, constituindo em uma das principais preocupações no desenho. Como os indicadores socioeconômicos são utilizados na estratificação, devem possuir o poder discriminante da população, agrupando indivíduos semelhantes e separando grupos distintos. Devem ainda estar respaldados por contexto teórico e metodológico robusto para defesa da hipótese das desigualdades.

O diagrama a seguir descreve essas etapas de construção de uma medida de desigualdade.

Figura 4.6: Diagrama das etapas de construção de medidas de desigualdades.



Fonte: Elaborado pelos autores.

Quais são as hipóteses por trás dessa teoria?

Hipótese nula = H_0 : os indicadores de saúde são homogeneamente distribuídos por todos os grupos sociais, ou seja: a ocorrência de eventos de Saúde é consequência apenas da distribuição do denominador (população de referência).

Hipótese alternativa = H_a : Fatores socioeconômicos interferem (estão associados) na distribuição do indicador de saúde.

Assim, as medidas de efeito ou de associação são utilizadas para representar e quantificar a ocorrência de desfechos em saúde (incapacidades, enfermidades ou mortes) em relação a uma exposição desigual a um fator socioeconômico.

Descreveremos as principais medidas de desigualdades em Saúde: conceitos, cálculos, interpretações e limitações. Medidas relativas do tipo “razão” (riscos relativos ou razão de taxas), medidas absolutas do tipo “diferença” (diferença de riscos ou riscos atribuíveis), medidas

dependentes de ranqueamento (curva de Lorenz e índice de Gini, curva e índice de concentração), medidas derivadas de análises multivariadas.

Medidas relativas do tipo “razão”: riscos relativos e razão de taxas

É a razão entre os coeficientes de incidência entre grupos de indivíduos expostos e não expostos, cujos resultados estimam a probabilidade ou a susceptibilidade de ocorrência de uma incapacidade, enfermidade ou morte em uma população dada as condições socioeconômicas estudadas.

Para o cálculo da Razão de Taxas, faz-se necessária a estratificação dos grupos segundo o indicador socioeconômico, cujo grupo melhor classificado será atribuída a não exposição. Os resultados dos cálculos devem ser interpretados como um estudo de risco relativo para coortes em epidemiologia, como indicado no gráfico a seguir:



$$RT = \frac{I_e}{I_o}$$

Onde:

I_e , grupo de melhor situação socioeconômica

I_o , grupo de pior situação socioeconômica

Exemplo

Dadas as diferentes condições socioeconômicas entre as regiões brasileiras, o risco de óbitos infantis tende a serem distintos também. É possível aferir a probabilidade de mortes na faixa etária segundo diferentes condições socioeconômicas utilizando a estatística de Razão de Taxas.

A Tabela 4.1 apresenta os valores dos coeficientes de mortalidade infantil, estratificada pela variável socioeconômica. Observamos que a região de maior rendimento médio é o Centro-Oeste e a de menor o Nordeste.

Aplicamos a fórmula da Razão de Taxas, utilizando as regiões de melhor e pior renda:

Tabela 4.1: Rendimento nominal médio e coeficiente de mortalidade infantil (por 1.000), segundo regiões brasileiras, 2010.

Região	Indicador socioeconômico*	Indicador de saúde		
		Coefficiente de Mortalidade Infantil	Numerador Óbitos Infantis	Denominador Nascidos Vivos
Nordeste (NE)	804,89	15,69	13.197	841.160
Norte (N)	954,4	17,26	5.289	306.422
Sul (S)	1.280,51	11,41	4.220	369.905
Sudeste (SE)	1.369,88	12,57	14.120	1.123.593
Centro-Oeste (CO)	1.419,80	13,79	3.044	220.788

*Valor do rendimento nominal médio mensal das pessoas de 10 anos ou mais de idade, com rendimento (Reais).

Fonte: Ministério da saúde, Sistema de Informação sobre Mortalidade, 2010, Instituto Brasileiro de Geografia e Estatística, censo 2010.

$$RT = CO / NE = 15,69 / 13,79 = 1,13$$

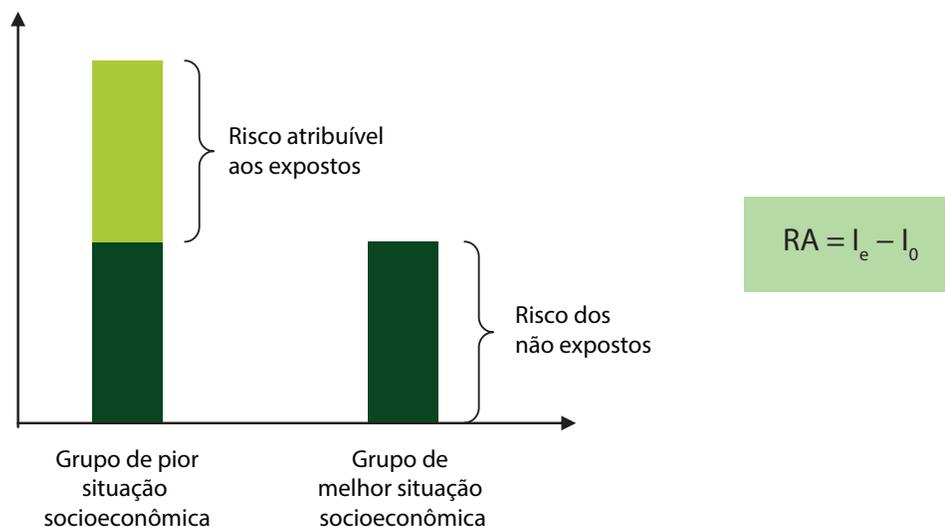
Portanto, a região Nordeste tem uma probabilidade de mortes infantis 13% ($RT^{-1} = 13\%$) maior que a região Centro-Oeste.

Medidas absolutas do tipo “diferença”: diferença de riscos ou riscos atribuíveis

Diferença de riscos

Trata-se da diferença entre os coeficientes de incidência dos grupos de indivíduos expostos e os não expostos. Indica o quanto é a diferença entre os grupos estudados e quantas ocorrências poderiam ser evitadas no caso de ausência de exposição das desigualdades socioeconômicas, ou quantos casos são atribuídos às desigualdades, finalmente, qual o risco “excedente” que é possível evitar em uma condição de igualdade.

Para o cálculo, como na Razão de Taxas, é realizada uma estratificação dos grupos segundo o indicador socioeconômico, e as taxas subtraídas entre eles. Observando o gráfico a seguir, o risco de desenvolver uma doença nunca é nula, mesmo para os não expostos; o que temos na área laranja é o “excesso” de risco provocado pelo fator socioeconômico de estudo.



Onde:

I_e , grupo de melhor situação socioeconômica

I_0 , grupo de pior situação socioeconômica

Exemplo

No exemplo anterior observamos que a probabilidade de óbitos infantis entre as regiões de pior e melhor renda média foi de 13%, em termos relativos. Mas em termos absolutos, qual seria esta diferença?

Para o cálculo do “excesso” de risco entre as regiões Nordeste e Centro-Oeste, descrito na Tabela 4.1, utilizamos o método proposto para o risco atribuído.

$$RA = CO - NE = 15,69 - 13,79 = 1,9$$

Portanto, se a Região Nordeste experimentasse as mesmas condições que a Centro-Oeste, poderia haver uma redução de 1,9 mortes infantis para cada grupo de 1.000 nascidos vivos.

Risco Atribuível Percentual (RA%)

Trata-se da diferença percentual dos coeficientes de incidência dos grupos de indivíduos expostos e os não expostos. Diferente da diferença de taxas ou do risco atribuível absoluto, o RA% indica a proporção percentual de casos do evento teoricamente prevenível pela elimina-

ção/modificação do fator de exposição na população entre os grupos estudados, ou a proporção do risco das pessoas expostas que é devida à exposição de um fator socioeconômico desfavorável.

O RA% pode expressar o excedente percentual de risco, ou como uma proporção do coeficiente de incidência total no grupo exposto que é atribuível à exposição, matematicamente expressando a razão da fórmula do RA pelo coeficiente de incidência no grupo exposto: (RA% ou fração etiológica).

$$RT = \frac{I_e - I_0}{I_e}$$

Onde:

I_e , grupo de melhor situação socioeconômica

I_0 , grupo de pior situação socioeconômica

Exemplo

No exemplo da Diferença de Taxas ou Risco Atribuível Absoluto, esperaríamos uma redução de 1,9 mortes infantis para cada grupo de 1.000 nascidos vivos, se houvesse as mesmas condições socioeconômicas entre as regiões Centro-Oeste e Nordeste, que representavam as maiores e menores rendas médias mensais respectivamente. Mas em termos relativos, qual seria esta diferença?

Para o cálculo do “excesso proporcional” de risco entre as regiões Nordeste e Centro-Oeste, descrito na Tabela 4.1, utilizamos o método proposto para o Risco Atribuído Percentual.

$$RA = CO - NE / CO = 15,69 - 13,79 / 15,69 = 0,12 \text{ ou } 12\%$$

Portanto, se a Região Nordeste experimentasse as mesmas condições que a Centro-Oeste, poderia haver redução de 12% de mortes infantis para cada grupo de 1.000 nascidos vivos.

Eventos Evitáveis Totais

Número absoluto de eventos (uma incapacidade, enfermidade ou morte, isto é, o numerador do Indicador de Saúde) que poderiam ser evitados caso todos os grupos tivessem a experiência do melhor grupo socioeconômico.

Exemplo

Será possível quantificar a possível evitabilidade, no caso de igualdade na situação socioeconômica. Pegaremos o exemplo da Diferença de Taxas, cujo valor é de 1,9, e aplicaremos a fórmula, que é uma recomposição do coeficiente pela população, aqui representado pelo número de nascidos vivos do grupo de pior situação socioeconômica (NE):

$$841.160 \times 1,9 / 1000 = 1.598$$

Ou seja, em números absolutos, esperaria que o Nordeste apresentasse 1.598 mortes a menos entre crianças menores de um ano do o Centro-Oeste se a situação socioeconômica entre as regiões fossem do mesmo nível.

Medidas dependentes de ranqueamento: curva de Lorenz e índice de Gini, curva e índice de concentração

Curva de Lorenz e Coeficiente de Gini

A curva de Lorenz e o índice de Gini são medidas utilizadas pelos economistas para avaliar as propriedades de distribuição de renda familiar e da riqueza dos grupos populacionais e por demógrafos para quantificar o grau de concentração populacional.

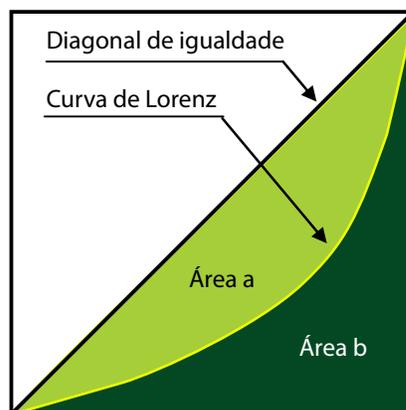
Na saúde encontramos ainda algumas aplicações direcionadas a análises de concentrações temporais na ocorrência de doenças e na caracterização de causa-efeito²⁹. Devido à propriedade de descrição na variação de riscos da doença na população, os índices têm sido utilizados também nos estudos de desigualdades em saúde, por estratificação de coeficientes (por incapacidades, enfermidades ou mortes) e os dados absolutos que deram origem a estes coeficientes.

A curva de Lorenz é um gráfico de frequência acumulada que compara a distribuição de uma dada variável com uma distribuição uniforme esperada no caso de completa. Essa distribuição uniforme de igualdade é representada por uma linha diagonal no gráfico. Quanto mais a curva de Lorenz se distanciar da diagonal, maior será a desigualdade na distribuição da variável analisada. Quanto mais a curva se aproximar da diagonal, mais homogênea será a distribuição da variável analisada.

O coeficiente de Gini é uma medida-resumo do afastamento da curva de Lorenz, que quantifica a área entre a curva de Lorenz e a diagonal da igualdade, que pode ser extraído a partir da razão entre as áreas hachuradas a e b, representadas na Figura 4.2, sendo igual a “a/

(a+b)”. O seu resultado se expressa como percentagem ou equivalência, que resulta em um valor entre 0 e 1, portanto quanto maior a “área a” maior será a desigualdade, e o seu valor mais próximo a 1. O índice de Gini é o coeficiente expresso em pontos percentuais, ou seja, o valor do coeficiente multiplicado por 100³⁰.

Figura 4.7: Representação gráfica da curva de Lorenz e área para o cálculo do coeficiente de Gini.



Fonte: Adaptado de Wikipédia.

O coeficiente de Gini pode ser calculado com a Fórmula de Brown:

$$G = 1 - \sum_{i=0}^{k-1} (Y_{i+1} - Y_i) (X_{i+1} + X_i)$$

Onde:

Y_i , a proporção acumulada da variável de saúde até o grupo i

X_i , a proporção acumulada da variável de população até o grupo i

Exemplo

A mortalidade de crianças menores de 5 anos ainda é um problema importante e se distribui de forma diferentes nas regiões brasileiras. Poderemos aferir a forma de distribuição dos casos em relação à população de crianças menores de 5 anos por região.

Dados necessários para o cálculo estão descritos na Tabela 4.2, composta de grupos, que no caso são representados pelas regiões, o coeficiente de mortalidade segundo as doenças diarreicas agudas em crianças menores de 5 anos, o número absoluto de óbitos e a população. Observe que na tabela, os dados foram estratificados segundo o risco, de forma negativa, isto é, os piores indicadores estão nas linhas superiores.

Tabela 4.2: Coeficiente de mortalidade por doenças diarreicas agudas na infância (por 100.000), segundo regiões brasileiras, 2010.

Região	Indicador de saúde Coeficiente de Mortalidade	Numerador Óbitos(Y)	Numerador Óbitos(Y)
Norte (N)	14,55	226	1.553.548
Nordeste (NE)	10,73	454	4.231.910
Centro-Oeste (CO)	8,19	87	1.062.663
Sudeste (SE)	3,57	185	5.184.294
Sul (S)	3,00	53	1.763.744

Fonte: Ministério da saúde, Sistema de Informação sobre Mortalidade, 2010.

Para a criação do gráfico, a estratificação dos grupos segundo a variável de saúde precede de transformação das variáveis em unidades normalizadas (0 a 1), cálculo das frequências acumuladas, e finalmente o coeficiente de Gini é obtido pela somatória do produto dos indicadores normalizados.

A representação gráfica das populações dos grupos estudados devem preencher as abscissas (eixo x) e as frequências acumuladas no número de eventos da variável de saúde no eixo das ordenadas (eixo y), demonstradas na Tabela 4.3.

Tabela 4.3: Estatísticas para a representação da curva de Lorenz e o coeficiente de Gini dos óbitos por doenças diarreicas agudas em crianças menores de 5 anos, segundo regiões brasileiras, 2010.

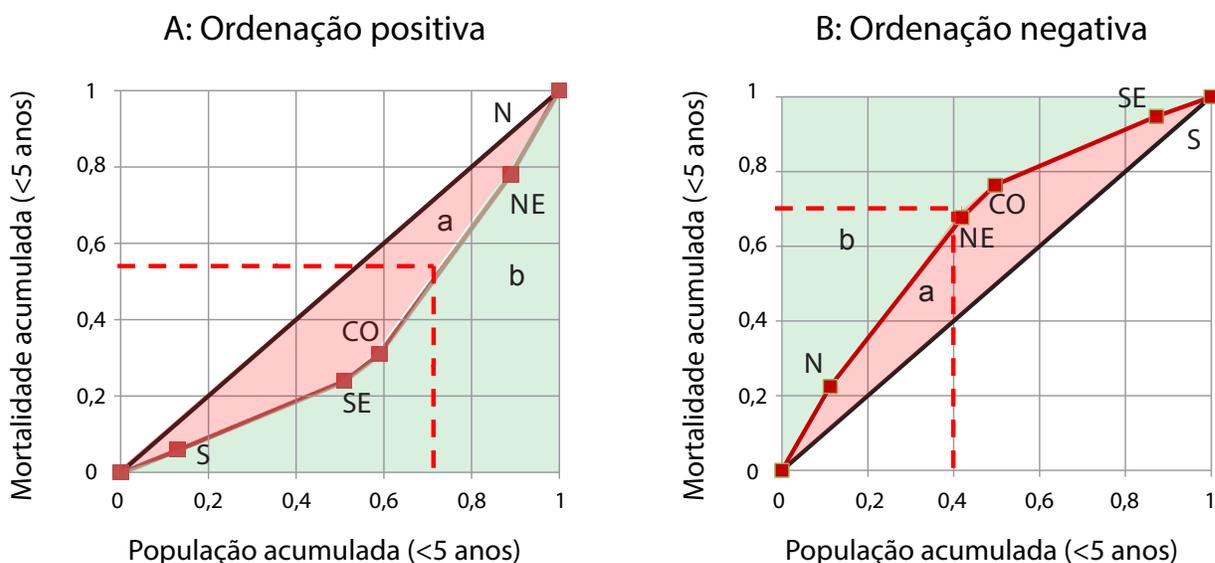
Região	FR (X)	FRA(X')	$X'_{i+1} - X_i (X'')$	FRA(Y)(X'')	FRA(Y)(X'')	$Y'_{i+1} + Y_i (Y'')$	$Y'' \times X''$
Norte (N)	0,11	0,11	0,11	0,22	0,22	0,22	0,03
Nordeste (NE)	0,31	0,42	0,31	0,45	0,68	0,90	0,28
Centro-Oeste (CO)	0,08	0,50	0,08	0,09	0,76	1,44	0,11
Sudeste (SE)	0,38	0,87	0,38	0,18	0,95	1,71	0,64
Sul (S)	0,13	1,00	0,13	0,05	1,00	1,95	0,25
Total (Σ)	1,00		1,00	1,00			1,31*

Nota: FR: Frequência relativa dos dados. No exemplo, X compreende a população menor de 5 anos e Y ao total de óbitos por doenças diarreicas agudas em menores de 5 anos. FRA: Acumulado da frequência relativa das variáveis estudadas.

* Representa o cálculo do coeficiente de Gini, porém subtraído em módulo da unidade 1, portanto o coeficiente será 0,31.

Fonte: Elaborado pelos autores.

Figura 4.8: Distribuição de mortes na infância provocadas por doenças diarreicas agudas, por regiões brasileiras, 2010.



Fonte: Elaborado pelos autores.

A interpretação da curva de Lorenz é bastante intuitiva. No exemplo, mostrado na Tabela 4.3, o coeficiente de mortalidade das doenças diarreicas agudas foi ordenado de forma decrescente, pois o indicador é negativo (quanto maior o seu valor, pior o estado de saúde). É importante que o analista tenha em mente o sentido do indicador de saúde devido à interpretação dos resultados.

Na Figura 4.8A, a ordenação escolhida foi a positiva, cuja interpretação deverá ser levado em conta que os grupos com melhor indicador estará no começo. Observamos que 30% dos óbitos por doenças diarreicas agudas em crianças menores de 5 anos estão distribuídos em 60% da população. Portanto os grupos estão em uma situação favorável, e representam as regiões Sul e Sudeste basicamente.

Na Figura 4.8B, a ordenação escolhida foi a negativa, que representa a ordem correta do indicador. Na análise, observamos que 70% dos óbitos brasileiros pela causa estudada estão concentrados nos grupos cuja população representa 40% do total, portanto passando uma imagem bastante direcionada do problema.

Aplicando a fórmula de Brown com os dados da Tabela 4.3, obteremos o coeficiente de Gini.

$$G = 1 - \sum_{i=0}^{k-1} (Y_{i+1} - Y_i) (X_{i+1} + X_i)$$

$$G = 1 - 1,31 = 0,31$$

O Índice de Concentração encontrado é de 0,31, 31%.

Índice de Concentração e Curva de Concentração

O Índice de Concentração e a Curva de Concentração são muito semelhantes ao Coeficiente de Gini e Curva de Lorenz, com a diferença de que agora a estratificação é realizada pelo indicador socioeconômico, e não pela variável de saúde. O índice e a curva de Concentração incorpora a dimensão socioeconômica de forma explícita, seus resultados e interpretações portanto devem ser levados em consideração.

O cálculo do Índice de Concentração segue a mesma lógica do cálculo do Coeficiente de Gini, variando, porém, entre -1 e +1, porém como o Gini, quanto mais distante do valor 0, maior será a desigualdade observada. Valores negativos ocorrem dependendo da estratificação adotada, em que a curva se apresenta acima da diagonal de igualdade. Nos casos de coincidências nas áreas de desigualdades apresentadas pelas curvas de Lorenz e de Concentração, os resultados entre os índices de Gini e de Concentração serão semelhantes.

Exemplo

Necessitamos agora saber a distribuição da mortalidade de crianças menores de 5 anos em relação à proporção de domicílios com abastecimento de água proveniente de rede geral de abastecimento em pelo menos uma cômoda na residência, dados oriundos do censo de 2010.

A Tabela 4.4 apresenta os dados agora com a inclusão da variável socioeconômica utilizada, além dos demais dados necessários para o cálculo do Índice e Curva de Concentração, apresentados anteriormente. Observe que na tabela, os dados foram estratificados segundo o indicador socioeconômico, de forma negativa, isto é, os piores indicadores estão nas linhas superiores.

Tabela 4.4: Proporção de domicílios abastecidos com água de rede geral de abastecimento e coeficiente de mortalidade por doenças diarreicas agudas na infância (por 100.000), segundo regiões brasileiras, 2010.

Região	% de abastecimento	Coefficiente de Mortalidade	Óbitos (Y)	População (X)
Norte	47,99	14,55	226	1.553.548
Nordeste	70,85	10,73	454	4.231.910
Centro-Oeste	79,17	8,19	87	1.062.663
Sul	84,46	3,00	53	1.763.744
Sudeste	88,48	3,57	185	5.184.294

Fonte: Ministério da saúde, Sistema de Informação sobre Mortalidade, 2010 Instituto Brasileiro de Geografia e Estatística, Censo 2010.

Para a criação do gráfico, a estratificação dos grupos segundo a variável socioeconômica precede de transformação das variáveis em unidades normalizadas (0 a 1) tal qual a curva de Lorenz, cálculo das frequências acumuladas, e finalmente o Índice de Concentração é obtido pela somatória do produto dos indicadores normalizados.

A representação gráfica das populações dos grupos estudados devem preencher as abscissas (eixo x) e as frequências acumuladas no número de eventos da variável de saúde no eixo das ordenadas (eixo y), demonstradas na Tabela 4.5.

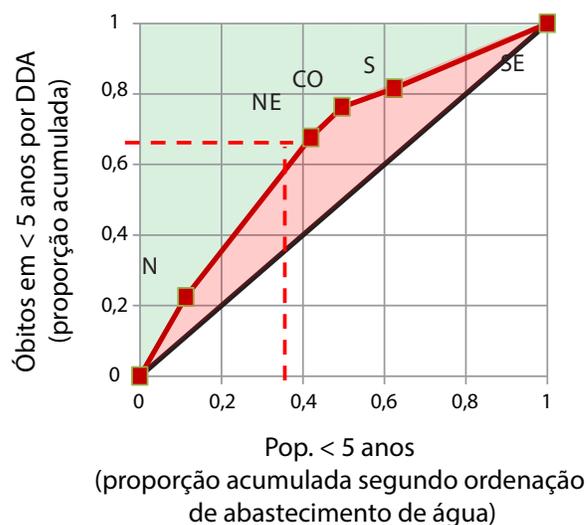
Tabela 4.5: Estatísticas para a representação da Curva de Concentração e o Índice de Concentração dos óbitos por doenças diarreicas agudas em crianças menores de 5 anos, segundo a proporção de domicílios abastecidos com água de rede geral, por regiões brasileiras, 2010

Região	FR (X)	FRA(X')	$X'_{i-1} - X_i (X'')$	FRA(Y)(X'')	FRA(Y)(X''')	$Y'_{i-1} + Y_i (Y''')$	$Y'' \times X''$
Norte	0,11	0,11	0,11	0,22	0,22	0,22	0,03
Nordeste	0,31	0,42	0,31	0,45	0,68	0,90	0,28
Centro-Oeste	0,08	0,50	0,08	0,09	0,76	1,44	0,11
Sul	0,13	0,62	0,13	0,05	0,82	1,58	0,20
Sudeste	0,38	1,00	0,38	0,18	1,00	1,82	0,68
Total (Σ)	1,00		1,00	1,00			1,29*

Nota: FR: Frequência relativa dos dados. No exemplo, X compreende a população menor de 5 anos e Y ao total de óbitos por doenças diarreicas agudas em menores de 5 anos. FRA: Acumulado da frequência relativa das variáveis estudadas.
 * Representa o cálculo do Índice de Concentração, porém subtraído da unidade 1 e multiplicado por 100, portanto o Índice de Concentração será 27%.

Fonte: Elaborado pelos autores.

Figura 4.9: Concentração de mortes na infância provocadas por doenças diarreicas agudas, segundo ordenação de proporção de domicílios abastecidos por água através de rede geral, segundo regiões brasileiras, 2010.



Fonte: Elaborado pelos autores.

A interpretação da curva de Concentração também é intuitiva e no exemplo, mostrado na Tabela 4.5 em que o indicador socioeconômico representado pela proporção de domicílios abastecidos com água de rede geral, foi utilizada para estratificação de forma descendente, pois o indicador é negativo (quanto maior o seu valor, pior o estado socioeconômico).

Na Figura 4.9, observa-se que pouco menos que 70% dos óbitos brasileiros pela causa estudada estão concentrados nos grupos cuja população representa 40% do total, utilizando a variável socioeconômica como estratificadora, portanto passando uma imagem bastante direcionada do problema.

Aplicando a fórmula de Brown com os dados da Tabela 4.5, obteremos o Índice de Concentração.

$$G = 1 - \sum_{i=0}^{k-1} (Y_{i+1} - Y_i) (X_{i+1} + X_i)$$

$$G = 1 - 1,29 = 0,29$$

O Índice de Concentração encontrado é de 0,29, 29%.

Risco Atribuível Populacional Percentual (RAP %)

O Risco Atribuível Populacional Percentual (RAP) é dado pela subtração da taxa do evento na população total da taxa do grupo de indivíduos com a melhor situação para o indicador em questão (grupo não exposto). O RAP% indica a proporção percentual de casos do evento teoricamente prevenível pela eliminação/modificação do fator de exposição na população estudada.

Eventos Evitáveis (TOTAL)

Eventos Evitáveis (TOTAL) é o número absoluto de eventos (agravos, isto é, o numerador do indicador de Saúde) que poderiam ser evitados, caso toda a população tivesse a experiência do grupo de indivíduos com a melhor situação para o indicador em questão.

Medidas derivadas de análises multivariadas

Índice de Efeito baseado em regressão

As medidas de Razão de Taxas e Diferença de Taxas baseiam-se na comparação de dois grupos por vez, portanto não leva em consideração todos os grupos que poderiam haver nos

estratos socioeconômicos estabelecidos. Uma medida baseada em regressão simples do indicador de Saúde sobre o indicador socioeconômico absoluto é uma metodologia aplicável a todos os grupos dos estratos, permitindo descrever a relação de desigualdade como um todo.

O Índice de Efeito resulta da estatística da inclinação da reta, descrita na equação de regressão, e poderá responder a questões sobre o que esperar de mudanças na saúde dadas as alterações no indicador socioeconômico e qual a proporção de variação do indicador de Saúde correspondente ao indicador socioeconômico.

Ainda que as relações entre as variáveis estudadas sejam complexas devido a muitos fatores intervenientes, em se tratando de um método mais sofisticado, os resultados passam a ser satisfatórios a partir da adequação das variáveis à hipótese, a linearidade entre as variáveis e o tamanho similar entre os grupos montados, portanto algumas premissas básicas devem ser respeitadas como requisito na utilização de modelos de regressão linear, caso contrário, outros modelos podem ser utilizados, como modelo logístico ou de Poisson³¹.

$$Y_i = \alpha + \beta X_i + \epsilon_i \text{ ou}$$
$$Y_i = \beta X_i = \alpha + \epsilon_i$$

No qual:

Y_i é a variável de efeito ou dependente (saúde);

α , é a constante, que representa a interceptação da reta com o eixo vertical;

β é a constante, que representa a inclinação da reta, no caso, a Medida de Efeito;

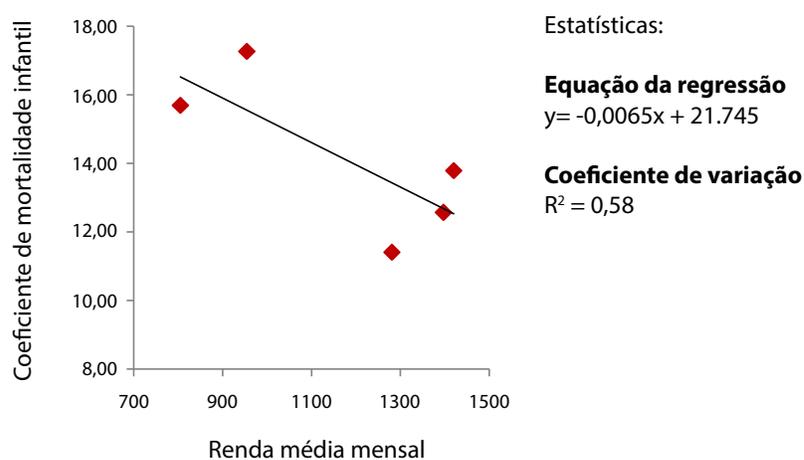
X_i variável explicativa ou independente (socioeconômica);

ϵ_i variável aleatória de efeito residual e outros erros.

Exemplo

Como varia a mortalidade infantil em relação à renda média nas regiões brasileiras? Para satisfazer a pergunta, será necessário utilizar um método que utilize todas as regiões, e resulte um indicador capaz de estabelecer a relação entre as variáveis.

Figura 4.10: Gráfico de efeito de regressão baseado na Tabela 4.1.



Fonte: Elaborado pelos autores.

A quantidade de estatísticas de saída no cálculo da regressão depende do software utilizado, caso seja um programa de estatística, deverá ser incluído na rotina os erros-padrão das estimativas ou os intervalos de confiança.

A inclinação da reta, equivalente ao Índice de Efeito, é descendente, corroborando com o valor negativo de β , mostrando que de forma geral, quanto maior a renda média mensal, menor a mortalidade infantil. A mesma estatística mostra que, em média, o coeficiente diminui 0,0065 por mil nascidos vivos para cada real de aumento na renda. O Coeficiente de Variação por outro lado, fornece a indicação de que 58% da variação das mortes infantis pode ser atribuída às variações da renda.

Referências

- 1 – Le Grand J (1987) Equity, health, and health care, *Social Justice Research*, 1(3), 257-274.
- 2 – Whitehead M (1992) The concepts and principles of equity and health, *Int J Health Serv*, 22(3), 429-445.
- 3 – Alleyne GA, Casas JA, Castillo-Salgado C (2000) Equality, equity: why bother?, *Bull World Health Organ*, 78(1), 76-77.
- 4 – Braveman P, Starfield B, Geiger, HJ (2001) World Health Report 2000: how it removes equity from the agenda for public health monitoring and policy, *BMJ*, 323(7314), 678-681.
- 5 – Evans T, Whitehead M, Diderichsen F, Bhuiya A, Wirth M (2001) *Challenging inequities in health: from ethics to action*. Oxford University Press, USA.
- 6 – Wagstaff, A. (2001) Economics, health and development: some ethical dilemmas facing the World Bank and the international community, *J Med Ethics*, 27(4), 262-267.
- 7 – O'Donnell O, van Doorslaer E, Wagstaff A, Lindelow M (2008) *Analyzing health equity using household survey data: a guide to techniques and their implementation*, WBI Learning Resources Series Washington, D.C. The International Bank for Reconstruction and Development. The World Bank, 220.
- 8 – Silva JB, Barros MBA (2002) *Epidemiologia e desigualdade: notas sobre a teoria e a história*, *Rev Panam Salud Publica*, 12(6), 375-383.
- 9 – WHO (2011) *World conference on social determinants of health. Meeting report at <http://www.who.int/sdhconference/en/>*. (Rio de Janeiro, World Health Organization).
- 10 – IPEA (2011) *O Brasil em 4 décadas*. Instituto de Pesquisa Econômica Aplicada.
- 11 – Neri MC. (2012) *De volta ao país do futuro: projeções, crise européia e a nova classe média*. Rio de Janeiro. FGV/CPS.

- 12 – Freitas CMD. (2003) Problemas ambientais, saúde coletiva e ciências sociais, *Ciência & Saúde Coletiva*, 8(1), 137-150.
- 13 – Schutz GE, Tambellini AT, Asmus CI, Meyer A, Camara VM. (2012) A agenda da sustentabilidade global e sua pauta oficial: uma análise crítica na perspectiva da saúde coletiva, *Ciência & Saúde Coletiva*, 17(6), 1407-1418.
- 14 – Almeida-Filho N. (2009) A problemática teórica da determinação social da saúde (nota breve sobre desigualdades em saúde como objeto de conhecimento), *Saúde em Debate Rio de Janeiro. Cebes*, 349-370.
- 15 – Kawachi I, Subramanian SV, Almeida-Filho N. (2002) A glossary for health inequalities, *J Epidemiol Community Health*, 56(9), 647-652.
- 16 – Wood CH, Carvalho JAM (1988) *The demography of inequality in Brazil*. Cambridge. Cambridge University Press.
- 17 – Townsend P, Davidson N (1998) Inequalities in health, in: R. C. Ian Marsh, Mike Keating (Ed) *Classic and contemporary readings in sociology*. Harlow UK, New York. Longman, 202.
- 18 – Whitehead M, Dahlgren G (Ed.) (2007) *Concepts and principles for tackling social inequities in health: Levelling up Part 1* Copenhagen. WHO Collaborating Centre for Policy Research on Social Determinants of Health. University of Liverpool.
- 19 – Macinko JA, Starfield B (2002) Annotated bibliography on equity in health, 1980-2001, *Int J Equity Health*, 1(1), 1.
- 20 – Braveman P, Gruskin S (2003) Defining equity in health, *J Epidemiol Community Health*, 57(4), 254-258.
- 21 – UN (1948) Universal declaration of human rights. Resolution adopted by the General Assembly United Nation. Disponível em: <http://www.un.org/en/documents/udhr/> (acessado em 12 Jan 2013).
- 22 – Toronto declaration on equity in health (2002). *Pan Am J Public Health*, 12(6), 465-467.

- 23 – Donaldson LJ (2002) Health services and the public health, *J Epidemiol Community Health*, 56(11), 835-840.
- 24 – Hart JT (1971) The inverse care law, *Lancet*, 1(7696), 405-412.
- 25 – Woodward A, Kawachi I. (2000) Why reduce health inequalities?, *J Epidemiol Community Health*, 54(12), 923-929.
- 26 – Rothman KJ, Adami HO, Trichopoulos D (1998) Should the mission of epidemiology include the eradication of poverty?, *Lancet*, 352(9130), 810-813.
- 27 – Donabedian A (1988) The quality of care. How can it be assessed?, *JAMA*, 260(12), 1743-1748.
- 28 – Schneider MC, Castillo SC, Bacallao J, Loyola E, Mujica OJ, Vidaurre M, Roca A (2002) Métodos de medición de las desigualdades de salud, *Rev Panam Salud Publica*, 12(6), 398-414.
- 29 – Lee WC. (1998) Characterizing exposure–disease association in human populations using the Lorenz curve and Gini index *Stat Med*, 16(7), 729-739.
- 30 – Coulter PB (1989) *Measuring inequality: a methodological handbook*. Boulder. Westview Press.
- 31 – Schneider MC, Castillo-Salgado C, Bacallao J, Loyola E, Mujica OJ, Vidaurre M, et al. Métodos de mensuração das desigualdades em saúde. *Rev Panam Salud Publica*. 2002;12(6). Disponível em: http://www.scielosp.org/scielo.php?script=sci_arttext&pid=S1020-49892002001200006&lng=en&nrm=iso&tlng=pt

Módulo 5

Inquéritos Populacionais

Celina Maria Turchi Martelli

Professora visitante da Universidade Federal de Pernambuco (UFPE), PE, Brasil.

Ricardo Arraes de Alencar Ximenes

Universidade Federal de Pernambuco (UFPE) e Universidade de Pernambuco, Recife/PE, Brasil.

Wayner Vieira de Souza

Centro de Pesquisa Ageu Magalhães (CPqAM), FIOCRUZ, Recife/PE, Brasil.

Noêmia Teixeira de Siqueira Filha

Centro de Pesquisa Ageu Magalhães (CPqAM), FIOCRUZ, Recife/PE, Brasil.

Introdução

O objetivo desse módulo é contextualizar os inquéritos populacionais como delineamento de estudo para mensurar eventos de saúde/doença visando o planejamento e gestão em Saúde pública. Na primeira e segunda unidades serão discutidos os tipos de inquéritos populacionais de abrangência nacional e regional com exemplos, bem como suas potencialidades. Os conceitos básicos de amostragem e representatividade serão apresentados na terceira unidade. As questões referentes à seleção e à alocação da amostra são discutidas na quarta unidade, utilizando-se como exemplo o Inquérito Nacional das Hepatites Virais. Na quinta unidade, a análise e apresentação dos resultados de inquéritos populacionais serão discutidas. Finalmente, as vantagens e as limitações e considerações éticas dos inquéritos populacionais serão abordadas na sexta unidade.



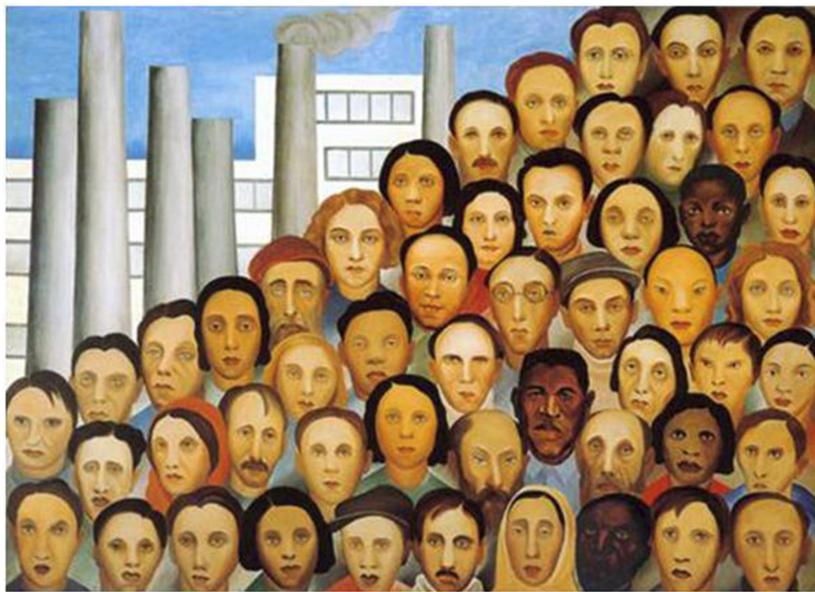
Unidade 1

Contextualização e Exemplos de Inquéritos Epidemiológicos de Base Populacional

Vamos iniciar este módulo com duas perguntas. Observem as figuras 5.1 e 5.2.
Por que realizar inquéritos populacionais?

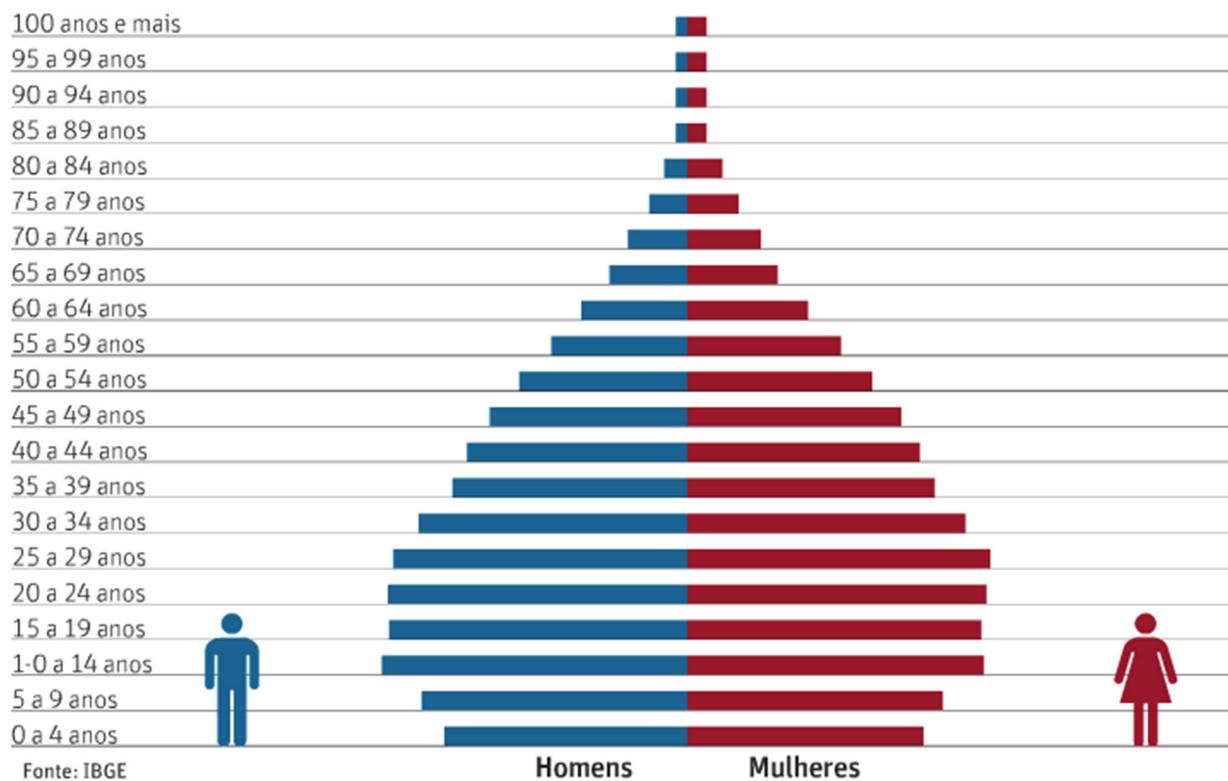
Quais os inquéritos nacionais de base populacional que conhecemos e suas aplicações na avaliação de eventos de saúde e do desempenho do setor saúde?

Figura 5.1: Quadro “Operários” de Tarsila do Amaral.



Fonte: Google images <<http://dalinhares.blogspot.com.br/2011/05/dia-do-trabalho-com-tarsila-do-amaral-e.html>>.

Figura 5.2: Pirâmide etária Brasil 2010.



Fonte: IBGE 2010. Pirâmide populacional

Uma das justificativas para a realização dos inquéritos populacionais é que por meio desse delineamento é possível obter informações sobre eventos relacionados à saúde-doença e aos fatores de risco da população, em determinada região geográfica. Trata-se de um delineamento tradicional que vem sendo amplamente usado nos últimos dois séculos, com incorporação dos aprimoramentos metodológicos e tecnológicos referentes ao período de estudo. Apesar do seu uso, ainda apresenta desafios metodológicos e operacionais como veremos neste módulo. Os inquéritos são classificados como estudos de prevalência, transversais ou seccionais^{1,2,3}.

Nas últimas décadas, inúmeros inquéritos de base populacional vêm sendo também realizados com o objetivo de determinar as características de assistência da prestação de serviços de Saúde como acesso, qualidade e satisfação dos usuários. Assim informações produzidas por inquéritos de base populacional representam uma demanda crescente dos gestores da área de Saúde para avaliar a morbidade e o desempenho do setor. Esse tipo de informação vem sendo considerada essencial para o planejamento de Saúde. Fica fácil entender a aplicação desse retrato do estado de saúde-doença e da avaliação do setor por meio dos inquéritos populacionais no planejamento e gestão em Saúde⁴.

O texto intitulado “Inquéritos populacionais: aspectos metodológicos, operacionais e éticos” reflete as questões metodológicas inerentes desse tipo de estudo dentro do contexto nacional. Essa discussão foi realizada por especialistas da área de Epidemiologia durante o seminário promovido pela Abrasco e Universidade de São Paulo². Recomendamos também a leitura do Capítulo 16 do livro “Epidemiologia & Saúde”³ que traz, de forma abrangente, os aspectos conceituais dos estudos de prevalência, com exemplos de estudos realizados em regiões brasileiras.

Para refletir

Quais são as medidas utilizadas para estimar a frequência dos eventos de Saúde?

Em Epidemiologia há duas categorias abrangentes para avaliar a frequência dos eventos: prevalência e incidência. A prevalência é a medida obtida nos inquéritos e é calculada pela razão entre o número existente de casos e o total da população de estudo. Veja no glossário a terminologia para prevalência mensurada no ponto e no período. Já a incidência mede as novas ocorrências do evento (doença, infecção etc.) que ocorrem em uma população. Dessa forma, a razão entre casos novos em uma população de não infectados e a população sob risco corresponde à taxa de incidência. Nesse aspecto, a incidência mensura o risco de se tornar infectado por unidade de tempo¹.

Exemplo: Na infecção pelo HIV-1, que apresenta longo período de incubação, a prevalência mede a carga geral da doença/infecção e a incidência mede as novas infecções e, portanto, avalia a dinâmica da infecção em populações. Para o gestor interessa a estimativa dos casos de pessoas vivendo com HIV/aids (prevalência) e também quantos casos novos anuais ocorrem na população refletindo a transmissão da infecção. Com os dados de incidência, avalia-se a tendência da epidemia e das medidas de controle adotadas no período⁵.

1. Inter-relações entre Incidência e Prevalência

Um dos conceitos básicos em Epidemiologia é que a prevalência depende de dois fatores: incidência e duração do evento de Saúde, que é expresso matematicamente por:

$$\text{Prevalência} = \text{incidência} \times \text{duração da doença}$$

A mudança na prevalência de um período para outro pode resultar de mudanças na incidência ou na duração da doença ou ambas. A introdução de novos tratamentos que previnam a morte, mas não a recuperação, pode resultar em aumento da prevalência da doença¹.

A Tabela 5.1 apresenta outras denominações para prevalência e as perguntas formuladas para obtenção dos dados e fontes de dados.

Tabela 5.1: Exemplos de tipos de prevalência com perguntas e potenciais fontes de dados.

Medidas	Pergunta	Fontes
Prevalência pontual	Você já teve...?	Internações/dia
Prevalência na vida	Você já teve...?	Qualquer admissão desde o nascimento
Prevalência no período	Você já teve...?	Internações + admissões no intervalo
Duração do evento	Há quanto tempo...?	Duração da internação

Fonte: Estudos de prevalência. In: Andrade ALSS & Zicker F. (Org.). Métodos de investigação epidemiológica em doenças transmissíveis - Volumes I. Ministério da Saúde e Organização Panamericana da Saúde 1997. Brasília.

Para refletir

Pense em exemplos de eventos aplicáveis aos diferentes tipos de prevalência. Sugerimos:

Hanseníase – duração do tratamento e prevalência no ponto

Infecção pelo HIV-1 – incidência e prevalência da infecção

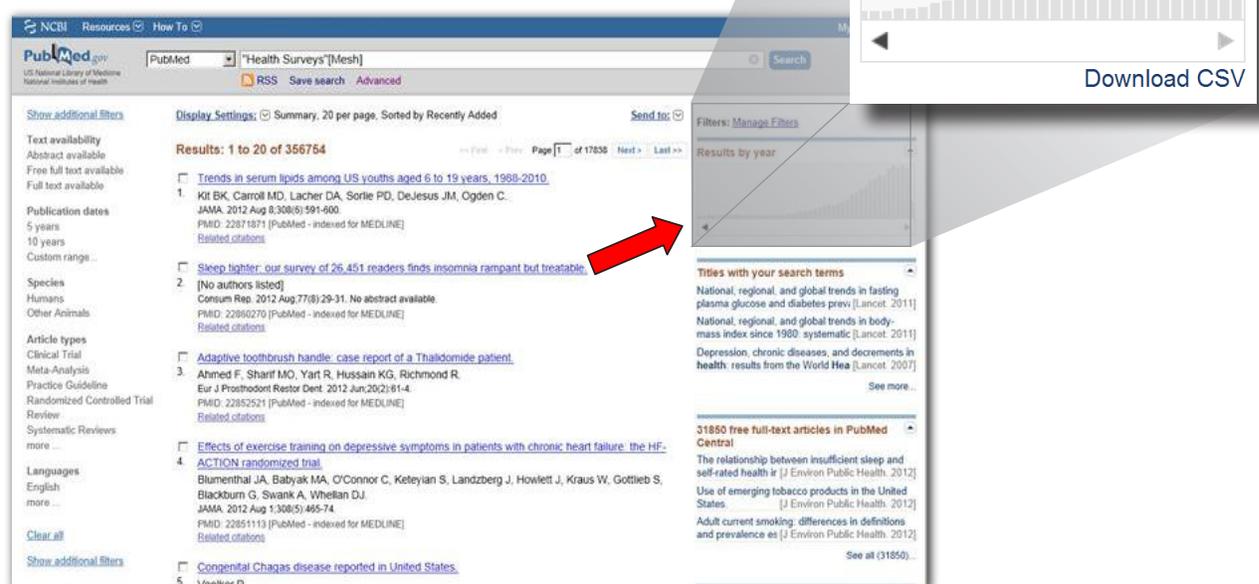
Tuberculose – duração e aderência ao tratamento e sobrevida

Diabetes do adulto – incidência e letalidade

Qual o efeito na estimativa da prevalência (superestimava ou subestimava) nos exemplos da Tabela 5.1?

Em busca realizada na base de dados bibliográficas PubMed <<http://www.ncbi.nlm.nih.gov/pubmed/>>, em agosto de 2012, foram encontradas 356.754 referências quando utilizado o termo de busca “Health Survey” (MeShTerm). Verifique o gráfico com a estatística anual dos estudos realizados no link <[http://www.ncbi.nlm.nih.gov/pubmed?term=“Health Surveys”\[Mesh\]](http://www.ncbi.nlm.nih.gov/pubmed?term=“Health Surveys”[Mesh])>.

Figura 5.3: Tela de consulta PubMed.



Fonte: <[http://www.ncbi.nlm.nih.gov/pubmed?term=“Health Surveys”\[Mesh\]](http://www.ncbi.nlm.nih.gov/pubmed?term=“Health Surveys”[Mesh])>.

Veja no gráfico de barras anterior (indicado na Figura 5.3), que o número de inquéritos publicados aumenta conforme o ano, mostrando a extensa produção de informações por esse tipo de delineamento.

No Brasil, por meio do suplemento de saúde da PNAD inserido no censo, é possível obter informações de acesso e utilização de serviços de Saúde para os anos de 1998, 2003 e 2008. O objetivo é produzir dados de base populacional sobre o acesso a serviços de Saúde no País; conhecer a cobertura dos grupos populacionais por diferentes modalidades de planos de seguro-saúde, dimensionando a população segurada; delinear o perfil de necessidades de saúde da população brasileira avaliado subjetivamente por meio de restrição de atividades habituais por motivo de saúde; autoavaliação da situação de saúde da limitação de atividade física e doenças crônicas referidas; produzir dados de base populacional sobre a utilização de serviços de Saúde; estimar o gasto privado em saúde das famílias brasileiras com planos de saúde, o consumo de bens, serviços e medicamentos. No primeiro bloco do suplemento, existem quatro medidas de morbidade referida: autoavaliação do estado de saúde, restrição de atividade por motivo de saúde, incapacidade funcional física e presença de doença crônica^{6,7,8}.

A Tabela 5.1 apresenta um sumário de inquéritos populacionais realizados no Brasil e em outros países. O nosso intuito foi mostrar que inquéritos de abrangência nacional são, geralmente, iniciativas institucionais/governamentais, podendo ser realizados uma única vez – Inquérito das Hepatites Virais⁹ ou mesmo instituído como programa de avaliação contínua como parte das políticas de Saúde para tomada de decisão – Saúde Bucal/Brasil; NHANES/EUA^{10,11}.

Tabela 5.2: Exemplos de inquéritos populacionais realizados no Brasil e em outros países.

Inquérito	País /Instituição	Ano/ periodicidade	
Pesquisa Nacional por Amostra de Domicílios (PNAD)	Brasil MS	1998, 2003, 2008	
Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (VIGITEL)	Brasil MS	2003 /anual	
Saúde Bucal (SBBrazil)	Brasil MS	2000 a 2003 / 2010	
Estudo de Prevalência de Base Populacional das Infecções pelos vírus das Hepatites A (VHA), B (VHB) e C (VHC) nas Capitais do Brasil	Brasil MS	2005 a 2009	
NHANES (National Health and Nutrition Examination Survey)	EUA CDC	Desde 1999 como programa institucional/anual	
National Population Health Survey – Household Component – Longitudinal (NPHS)	Canadá National Health Information Council	1992/bianual	
World Health Survey	Países membros WHO	2003 no Brasil	

MS = Ministério da Saúde; CDC = Center for Disease Control and Prevention; WHO = World Health Organization; DCNT = Doenças Crônicas Não Transmissíveis

Fontes: <<http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/PNAD2009/>>
 <http://portal.saude.gov.br/portal/saude/profissional/area.cfm?id_area=1521>
 <<http://dab.saude.gov.br/cnsb/sbbrasil/index.html>>
 <http://www.aids.gov.br/publicacao/2010/estudo_de_prevalencia_de_base_populacional_das_infecoes_pelos_virus_das_hepatites_b>
 <<http://www.cdc.gov/nchs/nhanes.htm/>>
 <<http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3225&lang=en&db=imdb&adm=8&dis=2>>
 <<http://www.who.int/healthinfo/survey/en/>>

Para refletir

Pense sobre o potencial dos inquéritos populacionais para estudos epidemiológicos e para o desenvolvimento de políticas de saúde pública.

Objetivo	População estudada	Site
Produzir dados de base populacional sobre o acesso a serviços de saúde, cobertura e morbidade referida	150.000 domicílios	http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/PNAD2009/
Monitorar a frequência e a distribuição de fatores de risco e proteção para DCNT	2.000 entrevistas nas capitais Faixa etária: ≥18 anos	http://portal.saude.gov.br/portal/saude/profissional/area.cfm?id_area=1521
Avaliar a situação do país em relação às Metas em Saúde Bucal	38.000 Faixas etárias: 5 e 12; 15-19; 35-44 e 65-74 anos	http://dab.saude.gov.br/cnsb/sbbrasil/index.html
Estimar a prevalência das infecções virais A, B e C por macrorregião	19.634 VHB e VHC Faixa etária: 10-69 anos 6.468 VHA Faixa etária: 5-19 anos	http://www.aids.gov.br/publicacao/2010/estudo_de_prevalencia_de_base_populacional_das_infeccoes_pelos_virus_das_hepatites_b
Avaliar a saúde e estado nutricional de adultos e crianças, com a responsabilidade de produzir estatísticas vitais e de saúde	5.000 exames	http://www.cdc.gov/nchs/nhanes.htm/
Coletar informações de saúde e sócio-demográfica da população	17.276 entrevistas	http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3225&lang=en&db=imdb&adm=8&dis=2
Fornecer informações sobre a saúde das populações e avaliação de desempenho dos sistemas de saúde	5.000 domicílios por país	http://www.who.int/healthinfo/survey/en/

O reconhecimento do valor dos dados obtidos por inquéritos populacionais no conhecimento do estado de saúde e das necessidades da população para os gestores de Saúde pública pode ser exemplificado pelo National Health and Nutrition Examination Survey, EUA (NHANES).

Historicamente, o Congresso dos EUA promulgou em 1956 um ato estabelecendo inquéritos periódicos para obter dados da prevalência de doenças agudas e crônicas, incapacidades, utilização de recursos e características demográficas e pessoais com o objetivo de planejamento e administração. Nesse tipo de survey de base populacional se combinam entrevistas de saúde, exames físicos e testes laboratoriais. Atualmente o NHANES é considerado o maior programa do National Center for Health Statistics (NCHS) que é parte do Center for Disease Control and Prevention (CDC), órgão responsável pela estatística de Saúde para os EUA. Alguns indicadores

de Saúde relacionados às condições médicas registrados nesses inquéritos incluem: anemia, doenças cardiovasculares, diabetes, exposições ambientais, doenças sexualmente transmissíveis, nutrição, obesidade, saúde oral, osteoporose, doenças respiratórias entre outras. A amostra para o NHANES é selecionada para representar a população dos EUA de todas as idades¹¹. Veja na Unidade 3 as questões relativas a amostras populacionais complexas em inquéritos populacionais.



Um exemplo de inquérito realizado no Brasil é a “Vigilância de Fatores de Risco e Proteção para Doenças Crônicas” (Vigitel), por entrevistas telefônicas. Conforme descrito no *site*, a Vigitel faz parte do Sistema de Vigilância de Fatores de Risco de Doenças Crônicas Não Transmissíveis (DCNTs) do Ministério da Saúde e tem por objetivo: “monitorar as principais determinantes das DCNTs no Brasil, contribuindo para a formulação de políticas públicas que promovam a melhoria da qualidade de vida da população brasileira”¹². O Ministério da Saúde do Brasil vem desenvolvendo o projeto de planejamento da Pesquisa Nacional de Saúde¹³.

Unidade 2

Tipos de Inquéritos

Vamos discutir nesta seção os tipos mais frequentes de inquéritos populacionais. Utilizaremos os conceitos do livro-texto de epidemiologia¹⁴.

Inquérito domiciliar de morbidade: Neste tipo de inquérito existe uma base populacional *sampling frame* para o estudo que geralmente é realizado por meio de amostragem. Também pode ser realizado um diagnóstico censitário (recenseamento) em situações especiais. Dessa forma, estabelece-se pela amostra o denominador populacional que servirá o cálculo das frequências e coeficientes de morbidade e mortalidade.

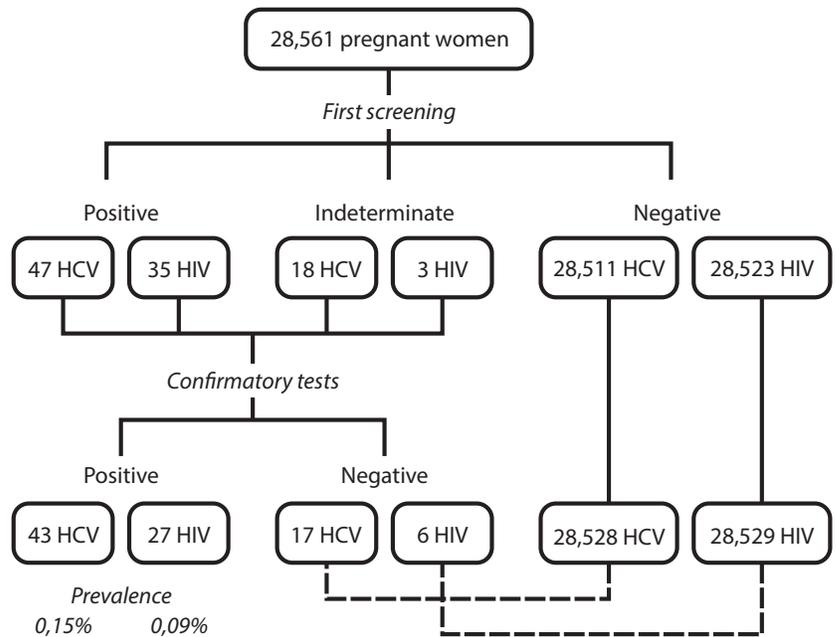
Estudo seccional multifásico: Trata-se da aplicação de instrumento de coleta de dados simplificado a toda a população para selecionar um subgrupo de maior suspeição para investigação mais detalhada.

Pense em um programa de triagem populacional como o Programa da Mamãe que oferece gratuitamente rastreamento sorológico para as principais doenças de transmissão vertical durante a gestação. Inicialmente a triagem sorológica em papel de filtro e um pequeno questionário são aplicados a toda população-alvo. Na etapa seguinte, as gestantes com exames positivos são convidadas para retestagem sorológica em sangue venoso, investigação clínica e aconselhamento, se necessário. Informações adicionais do cálculo de prevalência e de formas de análise nesse tipo de delineamento foram apresentadas no artigo “Prevalência e fatores de risco para infecções para HCV e HIV-1 entre gestantes na Região Centro-Oeste”¹⁵.

Inquéritos populacionais seriados

Quando os inquéritos são realizados mais de uma vez ou, periodicamente, recebem a denominação de seriados. Note que a população estudada em diferentes momentos não é a mesma, mas devem refletir uma amostra representativa da população. Como exemplo, podemos citar dois inquéritos populacionais conduzidos na Cidade de Goiânia nos anos de 2001 e 2002 que tiveram como principal objetivo a identificação de potenciais áreas de risco de infecção para dengue. No primeiro inquérito foram selecionados, por meio de amostra probabilística, 1.586 indivíduos. Foram realizadas entrevistas domiciliares e coleta de amostra de sangue para detecção de anticorpos anti-dengue. No ano seguinte, uma nova amostra probabilística de 2.906 indivíduos foi selecionada para entrevista e coleta de sangue. A prevalência de anticorpos antidengue foi de 37,3% no ano de 2001, com incremento de 7,8% no ano seguinte^{16,17}.

Figura 5.5: Triagem para HCV e infecção pelo HCV e HIV entre grávidas no setor público de saúde na região Centro-Oeste do Brasil.



Fonte: Costa et al., 2009.

Para refletir



Veja na Tabela 5.2 anterior quais dos inquéritos podem ser classificados como inquéritos seriados. Liste as vantagens do ponto de vista de análise de dados da repetição dos inquéritos em determinada população.

Um exemplo de análise de inquéritos seriados encontra-se na publicação da série Saúde no Brasil – Lancet (2011). Informações oriundas de inquéritos seriados foram utilizadas na avaliação da morbidade, mortalidade e de cobertura de intervenções da saúde materno-infantil, possibilitando avaliar a tendência dos indicadores em diferentes períodos¹⁸.

Unidade 3

Conceitos Básicos de Amostragem

A amostragem é um processo por meio do qual se pretende conhecer o que se passa no todo (população ou universo), investigando um subconjunto deste (amostra). Devemos ter claras algumas definições:

População – grupo inteiro de objetos (unidades) dos quais se pretende obter informações. A população deve ser definida claramente e em termos do que se pretende conhecer.

Amostra – parte ou subconjunto da população usada para obter informação acerca do todo.

Unidade amostral – qualquer elemento da população.

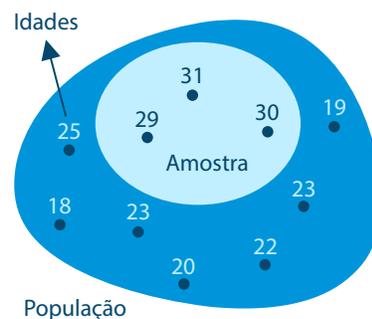
Variável – característica de uma unidade que será medida a partir daquela unidade da amostra.

Figura 5.6: População e amostra.

População: Pode ser finita ou infinita.

Média da população = 24 anos

Média da amostra = 30 anos



Fonte: Elaborado pelos autores

Os processos de amostragem aplicam-se em várias áreas do conhecimento e constituem, muitas vezes, a única forma de obter informações sobre determinada realidade que desejamos conhecer.

A teoria da amostragem é um dos instrumentos que possibilita esse conhecimento científico da realidade, em que outros processos (censo por exemplo), por razões diversas, não se mostram adequados ou até mesmo possíveis. Estuda ainda as relações existentes entre uma população e as amostras extraídas dessa população. É útil para avaliação de grandezas desconhecidas ou para determinar se as diferenças observadas entre duas amostras são devidas ao acaso ou se são verdadeiramente significativas.

A amostragem é o processo de determinação de uma amostra a ser pesquisada, enquanto que a amostra é uma parte dos elementos selecionada de uma população.

A amostragem consiste, portanto, em selecionar parte de uma população e observá-la com vistas a estimar uma ou mais características para a totalidade da população. Enquanto o censo envolve exame de todos os elementos de um dado grupo, a amostragem envolve estudo de apenas uma parte dos elementos. “Para se saber se o bolo de chocolate está bom, basta comer uma fatia.”

1. Amostragens em Estudos Transversais

Em estudos transversais, na maioria das vezes, estamos interessados em conhecer a proporção com que um fenômeno ocorre ou a média de uma determinada medida em uma população. Por exemplo, podemos estar interessados em conhecer a prevalência de uma determinada doença ou os níveis médios de uma medida bioquímica (colesterol, glicose etc.) nesta população. Adicionalmente podemos também estar interessados em além de estimar parâmetros globais, comparar essas médias ou proporções em subgrupos da população estudada.

Para tal, decidindo-se por uma investigação por amostragem, deve-se ter em conta as fases deste processo. Tal processo consiste em identificar os dados que deverão ser coletados e o instrumento a ser utilizado para essa coleta, definindo um processo de amostragem adequado ao tipo de dados e ao instrumento de análise.

No processo de coleta de dados, é necessário desenvolver uma sistemática que assegure a confiabilidade e a comparabilidade desses dados. Mais especificamente, é necessário que se estabeleça um plano de amostragem de acordo com a população-alvo, com a definição da população a pesquisar e com um método adequado de administração do inquérito.

O plano de amostragem deverá começar por determinar qual o nível de extensão geográfica em que o processo de amostragem deverá ser conduzido (mundial, nacional, regional, urbano, rural, grupo de indivíduos etc.).

A construção da amostra propriamente dita envolve várias etapas igualmente importantes e que são:

- A identificação da população-alvo.
- O método de seleção da amostra.
- A dimensão da amostra.

Inferência

Quando realizamos inquéritos por amostragem, dispomos de parâmetros calculados para a amostra praticada, sendo que nosso objetivo é conhecer o valor, ou intervalo mais provável de valores que esse parâmetro assume na população (universo).

Sendo assim, podemos dizer que o objetivo de qualquer pesquisa por amostragem é fazer *inferências* a respeito da população de interesse, baseado nas informações obtidas de uma amostra, extraída dessa população. Essas inferências visam à estimação de certas características numéricas da população, assim como proporções médias, totais, etc. Essas medidas numéricas descritivas da população são chamadas *parâmetros*.

Um *estimador* é uma função de variáveis aleatórias observáveis, usada para estimar um parâmetro. Por exemplo, a média de uma amostra \bar{x} pode ser usada como estimador da média da população, μ . \bar{x} é um estimador dado ser uma função de observações de uma amostra, mas note que \bar{x} é uma variável aleatória que tem uma distribuição de probabilidade, que depende do mecanismo de amostragem adotado.

Alguns dos valores possíveis que \bar{x} pode assumir serão pouco desviados de μ e outros poderão ser mais desviados, tanto para menos como para mais.

Se tomarmos uma amostra adequada e calcularmos um valor específico de como nossa melhor estimativa de μ , é importante saber que, em média, a variável aleatória cria valores que estarão centrados em torno de μ e, em geral, muito próximos de μ .

Imagine se extraíssemos várias amostras independentes e de mesmo tamanho de uma população e calculássemos a média de cada uma delas. A distribuição de frequência dessas médias teria (para n suficientemente grande) média igual à média da população e desvio-padrão igual a σ/\sqrt{n} , que chamamos de erro-padrão da média amostral e estimamos como **se = σ/\sqrt{n}** .

No caso de estimação de prevalências, considere-se p esta prevalência na amostra, que pode ser descrita como $p=r/n$ a proporção de vezes que determinado evento ocorre em uma amostra e seja a proporção na população.

Estritamente falando $\frac{p - \pi}{\text{se}(p)}$ tem distribuição binomial com **se(p) = $\sqrt{\pi(1-\pi)/n}$**

2. Elementos do Problema Amostral

Devemos novamente enfatizar que nosso objetivo é fazer inferência sobre a população, tendo por base informações de uma amostra.

Consideremos o problema particular de extrair uma amostra de uma população de medidas (população), com o objetivo de fazer inferências sobre o parâmetro populacional como total média ou proporção, porém impondo um limite de erro **estimativas**. Lembremos que diferentes amostras produzirão diferentes estimativas de determinado parâmetro populacional ao que chamaremos de **variável**. O limite para os erros de estimativa é garantir que as diversas amostras produzirão **não se afastarão além do limite estabelecido do parâmetro**.

Cada observação, ou item, extraído de uma população custa algum dinheiro ou esforço (custo) sobre determinado parâmetro populacional de interesse. Portanto, o custo e o pesquisador deverá decidir sobre a quantidade de informações a obter com os recursos que dispõe. Pouca informação impede a inferência, enquanto que o excesso de informação leva ao desperdício de recursos. A quantidade de informação obtida de uma amostra é consequência do número de observações. Além disso, a vez também depende da quantidade de variação nos dados (ou seja, da **dispersão na população**). Portanto a quantidade de informação que vamos obter depende dos elementos que a amostra terá (**tamanho da amostra**), que por sua vez depende da **dispersão** do fenômeno na população e dos custos envolvidos.

Quanto à precisão, vamos explicitar melhor o que devemos garantir.

Seja θ o parâmetro populacional e seja $\hat{\theta}$ sua estimativa amostral. Queremos que $|\theta - \hat{\theta}| < B$. Como $\hat{\theta}$ é criado por um estimador que é uma variável aleatória, devemos assegurar que $p(|\theta - \hat{\theta}| < B) = 1 - \alpha$, ou seja $p(|\theta - \hat{\theta}| < B) = 1 - \alpha$.

A **dispersão** do fenômeno a ser estudado influencia o **tamanho da amostra** e também a forma de **selecionar a amostra** da população, que por sua vez influencia o custo.

Em síntese, devemos estabelecer um **desenho amostral** que seja adequado para os parâmetros populacionais, considerando a dispersão do fenômeno estudado, o custo estabelecido e com um método de seleção que permita atingir tais objetivos.

Outra questão importante na amostragem refere-se ao **método de amostragem**. Existem diferentes formas de se coletar os dados amostrais cujo **desenho amostral** e as técnicas amostrais e procedimentos de estimação estão baseados em diferentes princípios.

que os *dados da amostra foram extraídos da população de interesse*. Sendo assim, “o como” extrair a amostra é de importância capital e existem alguns problemas clássicos que devemos examinar:

- A não resposta.
- Erros de aferição e/ou informação.
- Substituição de elementos da amostra.
- A ocorrência de qualquer um desses problemas pode tornar nossos *resultados enviesados*.

Quanto aos métodos de coleta de dados, podemos descrever alguns usuais e discutir os cuidados que devemos ter com cada um deles:

- Entrevista pessoal.
- Entrevista por telefone.
- Questionário autoadministrado.
- Observação direta.
- Todos esses métodos pressupõem a existência de um instrumento (questionário ou de um elenco de perguntas e observações a serem realizadas), o que também deve ser planejado com bastante critério.

Desenhos Amostrais

Amostra Probabilística – Todos devem ter probabilidade (chance) conhecida e diferente de zero de pertencer à amostra.

Se uma amostra de tamanho n é extraída de uma população de tamanho N de forma que toda e qualquer amostra de tamanho n , possível de ser extraída, **tenha a mesma chance de ser selecionada**, chamaremos o procedimento amostral de **Amostra Aleatória Simples (AAS)**.

Seleção de uma AAS:

- Listas/sorteio.
- Tabela de números aleatórios.
- Fração de amostragem (tamanho da amostra $[n]$ /tamanho da população $[N]$) – seleção sistemática. Deve-se registrar que a seleção sistemática atende estritamente aos princípios da AAS quando as unidades são listadas aleatoriamente.
- Limitação da AAS: pode ter alto custo

Amostra Aleatória Simples – desenho básico que consiste na seleção de n unidades amostrais de forma que cada amostra de tamanho n tenha **a mesma chance de ser selecionada**.

Amostra Estratificada – se pudermos dividir os elementos de uma população em **grupos relativamente homogêneos**, ou estratos, e selecionar uma amostra aleatória simples de cada grupo, extrairemos uma amostra chamada amostra aleatória **estratificada**. Uma **amostra aleatória estratificada** é aquela obtida pela separação dos elementos de uma população em subconjuntos mutuamente excludentes, chamados estratos, para então selecionar uma amostra aleatória simples de cada estrato.

Ex.: média de permanência em internações por clínica.

Se o número de elementos de cada amostra estiver de acordo com a proporção do estrato na população, as observações podem ser misturadas para se obter os resultados globais. Se, no entanto, todas as amostras tiverem o mesmo número de elementos, os resultados de cada estrato têm de ser reponderados considerando a proporção desse estrato na população.

A estratificação de uma população faz sentido quando é possível identificar subpopulações que variam muito entre si no que diz respeito à variável em estudo, mas que variam pouco dentro de si. Nestas condições, uma amostra estratificada pode fornecer resultados mais precisos do que uma amostra aleatória simples. Em princípio, na amostragem estratificada todos os estratos devem ser representados na amostra e a etapa seguinte consistirá em selecionar uma AAS dentro de cada estrato, utilizando-se as técnicas já mencionadas.

Amostra por Conglomerados – tipo de amostragem usado por seu caráter de economia, principalmente em áreas urbanas, onde a localidade é dividida em blocos que serão selecionados para investigação. Uma amostra por conglomerados é uma amostra aleatória simples em que cada unidade amostral é um conjunto, ou conglomerado, de elementos da população. Ela será chamada de amostra por conglomerados com um estágio (AC_1), se incluirmos todos os elementos integrantes dos conglomerados selecionados.

Na amostragem por *clusters*, a população é dividida em grupos ou *clusters*. Este tipo de amostragem se torna particularmente útil quando a população se encontra dividida em número finito de grupos, caracterizados por terem uma dispersão idêntica à população total, isto é, os grupos deverão, tanto quanto possível, ser "microcosmos" da população a estudar. Vantagens: amostras por conglomerados são menos dispendiosas do que amostras aleatórias simples ou estratificadas (racionalização dos trajetos). Desvantagens: passamos a ter que controlar dois tipos de dispersão – dentro e entre conglomerados. Seleção: AC_1 é uma AAS de conglomerados.

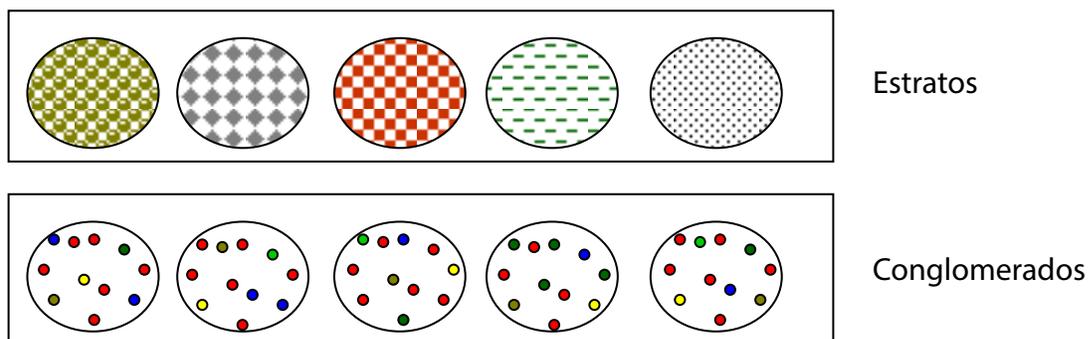
Amostras em vários estágios – por exemplo, é usual, tanto nas amostras estratificadas como nas amostras por conglomerados, praticá-las em vários estágios. Isto se dá pela seleção, em primeiro estágio, de certo número de conglomerados ou estratos para compor a amostra e, no segundo estágio, pela seleção de determinado número de domicílios para inclusão. Pode-se também trabalhar com o terceiro estágio com amostragem aleatória dos residentes no domicílio, por exemplo.

3. Considerações sobre Estratos e Conglomerados

Cabe aqui salientar a diferença primordial a ser considerada na construção de estratos e conglomerados. Os estratos devem ser tão homogêneos quanto possível internamente, devendo haver bastante heterogeneidade de um estrato para outro. Contrariamente, os conglomerados deverão apresentar bastante heterogeneidade interna, sendo um conglomerado bastante parecido com outro.

Note-se ainda que podemos desenhar uma amostra estratificada, em primeiro estágio e posteriormente por conglomerados em segundo.

Figura 5.7: Diagrama representando conglomerados e estratos.



Fonte: Elaborado pelos autores.

4. Dimensionamentos de Amostras

O objetivo de qualquer pesquisa por amostragem é fazer **inferências** a respeito de um parâmetro da população de interesse, baseado nas informações obtidas de uma amostra extraída dessa população.

Logo toda a discussão levada até aqui sobre dispersão e precisão, como condicionantes do tamanho de uma amostra, está atrelada à estimação de um único parâmetro.

Sendo assim, na pesquisa em geral e particularmente na por amostragem, devemos ter claro nosso objetivo que indicará o parâmetro a ser estimado, que por conseguinte, delineará o dimensionamento da amostra.

Uma amostra, assim dimensionada, fornecer-nos-á boa estimativa para o parâmetro a partir do qual foi dimensionada, sendo capaz de fornecer estimativas de parâmetros adjacentes com precisão variável dependendo da dispersão de outros parâmetros na população.

Parte essencial no planejamento de uma investigação é decidir quantos elementos serão incluídos na amostra a ser estudada, no sentido de atender aos objetivos do estudo.

Tirar esse número da “cartola” ou simplesmente dimensionar a amostra com base na capacidade operacional é uma péssima prática.

Por outro lado, incluir muito mais elementos na amostra do que o necessário é perda de tempo e de recursos.

a) Dimensionamento de uma amostra para estimação de uma proporção

$$n = \frac{z^2 \cdot p(1 - p)}{d^2}$$

Onde d é a semi-amplitude do IC e z o valor da normal para o grau de confiança desejado. Isto é válido apenas se N é suficientemente grande, dado que esta relação é uma simplificação de:

$$n = \frac{N \cdot z^2 \cdot p(1 - p)}{z^2 \cdot p(1-p) + (N - 1) \cdot d^2}$$

b) Dimensionamento de uma amostra para estimação de uma média

$$n = \frac{N \cdot \sigma^2}{(N - 1) \cdot d^2 + \sigma^2}$$

Onde σ^2 representa estimativa do desvio-padrão do universo.

Para o dimensionamento de uma amostra para comparação de duas proporções ou duas médias, ver o livro-texto indicado¹⁹.

Unidade 4

Seleção e Alocação da Amostra

1. Fontes de Dados para Seleção de Amostras (listas, bancos de dados etc.)

A base de amostragem *sampling frame* corresponde a uma lista que contém todas as unidades da população que se deseja estudar. A construção ou a obtenção de uma base de amostragem adequada é indispensável para que se evite viés de seleção (consultar glossário). Se as unidades amostrais na base de amostragem diferem, em suas características, daquelas da população-alvo, os parâmetros a serem obtidos na amostra também irão diferir daqueles da população-alvo, comprometendo assim a validade do estudo. Moser and Kalton referem-se a alguns problemas relativamente frequentes relacionados à base de amostragem²⁰. A base de amostragem pode não conter todas as unidades que compõem a população-alvo, seja porque a cobertura dela não é completa ou porque algumas das unidades que deveriam compor essa base não estão aí presentes. Um exemplo da primeira situação seria se tomássemos como base de amostragem uma lista das residências ligadas à rede de eletricidade, o segmento da população sem acesso a esse serviço ficaria excluído da nossa lista e da amostra. A segunda situação pode ser exemplificada pela ausência dos imigrantes recentes em uma lista de moradores de determinada localidade. O segundo problema da base de amostragem ocorre quando essa contém agrupamentos de elementos *clusters*, mas o objetivo é realizar uma análise a nível individual. Nessa situação a maior semelhança entre os elementos de mesmo agrupamento (intragrupo) e as diferenças entre os agrupamentos (intergrupo) terão que ser consideradas na análise e interpretação dos resultados. O terceiro problema refere-se à incorporação na lista de elementos que não deveriam fazer parte dela (por exemplo: moradores temporários na lista de residentes permanentes ou a permanência na lista de indivíduos que tenham emigrado para outra localidade). O quarto problema seria a existência de duplicidade de elementos na mesma lista. Todos esses problemas podem vir a comprometer a representatividade da amostra.

Para refletir

Leia nas próximas seções algumas estratégias para não comprometer a representatividade da amostra populacional.

2. Fração de Amostragem, Pesos e Ponderação

Ao se estratificar a análise por determinado fator, objetiva-se que todas as categorias daquele fator estejam representadas na amostra. Uma vez definido(s) o(s) fator(es) de estratificação, a etapa seguinte é a divisão da amostra nos diferentes estratos considerados. Duas alternativas são possíveis. Uma seria fazer essa divisão sem considerar o que a fração correspondente a cada estrato representa do todo. Nessa situação, para se obter a estimativa de um parâmetro para o todo, terá que ser usado um fator de correção que corresponde a um peso. Por exemplo, se uma amostra é estratificada por sexo e a proporção de homens e mulheres na amostra difere da população de origem, para se estimar um parâmetro para a amostra como um todo (homens e mulheres) será necessário utilizar um peso para reestabelecer a proporcionalidade original ou a estimativa estará enviesada. O uso de pesos é frequente em inquéritos populacionais, e exemplos podem ser facilmente acessados na literatura em trabalhos que usam esse desenho de estudo. A segunda alternativa seria fazer uma partilha proporcional utilizando-se uma fração constante de amostragem. Se, no exemplo anterior, tivessem sido selecionados $1/10$ dos homens e $1/10$ das mulheres, a proporcionalidade original teria sido mantida, dispensando o uso de pesos no momento da análise.

3. Seleção Proporcional ao Tamanho

Em amostras complexas, em múltiplos níveis, pode ocorrer que em alguns níveis sejam selecionados agrupamentos e, em outros, indivíduos. Para manter controle sobre o tamanho da amostra é necessário definir o número de indivíduos a serem selecionados em cada agrupamento. Se os agrupamentos diferem em tamanho e é estabelecido um tamanho de amostra fixo a ser selecionado em cada agrupamento, a amostra não será equiprobabilística; a probabilidade de cada agrupamento ser sorteado será semelhante, mas a probabilidade de um indivíduo ser sorteado será diferente nos diferentes agrupamentos. A estratégia que permite manter controle sobre o tamanho da amostra e, ao mesmo tempo, manter a amostra equiprobabilística é fazer o sorteio com probabilidade proporcional ao tamanho do agrupamento. Dessa forma, os agregados com maior número de unidades terão maior chance de serem sorteados, o que será balanceado pela menor probabilidade das unidades amostrais desses agrupamentos vierem a ser sorteadas.

4. Seleção Sistemática

Na seleção sistemática após definir-se a fração da população que comporá a amostra (por exemplo 1 em cada 100), e, com isso, o intervalo amostral, realiza-se o sorteio aleatório de um número contido no intervalo amostral (no exemplo um número entre 1 e 100) e adiciona-se consecutivamente o intervalo amostral a esse número sorteado, definindo-se assim todos os elementos da amostra. Em uma amostra sistemática, ao se definir o primeiro elemento da amostra, são definidos todos os demais, diferentemente de uma amostra aleatória simples na qual a seleção de cada unidade amostral é independente das demais. Por outro lado, a amostra obtida pode ser mais dispersa do que aquela proveniente de sorteio aleatório simples.

5. Amostragem por Quotas

A amostra por quotas é um tipo de amostra estratificada no qual a escolha final dos componentes dela, em cada estrato, é feita pelo entrevistador. Se for um processo de amostragem em múltiplos níveis, poderá haver sorteio de unidades amostrais em um ou mais níveis, porém não no último nível, ou seja, aquele no qual são efetivamente selecionados os elementos da amostra. Na seleção das unidades amostrais, se são considerados vários fatores de estratificação, a seleção final dos indivíduos de cada estrato poderá ser independente ou inter-relacionada. Na independente define-se, por exemplo, que na amostra determinado percentual de indivíduos será do sexo masculino e determinado percentual pertencerá a um grupo etário específico, respeitando a proporcionalidade existente na população de origem. Na forma inter-relacionada seria especificado o percentual de indivíduos do sexo masculino dentro de cada grupo etário. Os pros e os contras da amostragem por quotas, notadamente a praticidade *versus* os problemas de precisão e representatividade, têm sido amplamente discutidos na literatura específica.

O estudo de prevalência de base populacional das infecções pelos vírus das hepatites A, B e C, nas capitais do Brasil, exemplifica a utilização de várias estratégias de amostragem⁹.

Tomando-se como exemplo a hepatite B, um dos objetivos do estudo foi o de estimar a soroprevalência da hepatite B para cada uma das regiões do País, para as faixas etárias de 10 a 19 e de 20 a 69 anos. O tamanho da amostra foi definido tomando como base os dados dos bancos de sangue, por constituírem informações padronizadas para todo o País. Foi utilizada uma amostra probabilística, em múltiplos estágios.

Primeiro estágio: uma vez estimado o tamanho da amostra total para cada região e para cada grupo etário, ela foi dividida entre as capitais de cada região utilizando-se uma fração constante de amostragem, garantido-se, portanto, uma partilha proporcional. As figuras a seguir apresentam a distribuição da população e da amostra em relação às capitais da Região Nordeste.

Tabela 5.3: População total, por faixa etária e total de domicílios das capitais do nordeste do Brasil.

Capitais	Faixa etária			Total de Domicílios	População Total
	5 a 9	10 a 19	20 e mais		
Maceió	78.352	166.752	470.572	202.226	795.812
Fortaleza	206.064	454.903	1.275.874	527.905	2.137.464
João Pessoa	54.136	124.545	366.428	152.535	595.272
Teresina	67.207	167.619	411.408	170.104	714.232
Aracajú	40.636	99.099	280.542	117.011	460.522
Salvador	206.299	505.639	1.522.619	653.522	2.435.157
São Luiz	82.712	208.309	494.404	202.891	868.137
Recife	121.420	278.308	905.136	377.070	1.416.318
Natal	65.114	151.013	432.157	178.312	410.422
Total	921.940	2.156.187	6.159.140	2.581.576	9.833.336

Fonte: Universidade de Pernambuco, 2010⁹.

- Fração de amostragem da população de 5 a 9 anos para o NE = $428/921.940 = 0,00046423$.
- Fração de amostragem da população de 10 a 19 anos para o NE = $428/2.156.187 = 0,00019849$.

Tabela 5.4: Tamanho da amostra de indivíduos para inquérito de hepatite A, segundo faixa etária e capital – Nordeste, 2000.

Capitais	Tamanho da amostra por faixa etária	
	5 a 9	10 a 19
Maceió	37	34
Fortaleza	96	91
João Pessoa	26	25
Teresina	32	34
Aracajú	19	20
Salvador	96	101
São Luiz	39	42
Recife	57	56
Natal	31	30
Total	433	433

- Fração de amostragem da população de 5 a 9 anos para o NE = $428/921.940 = 0,00046423$.
- Fração de amostragem da população de 10 a 19 anos para o NE = $428/2.156.187 = 0,00019849$.

Fonte: Universidade de Pernambuco, 2010⁹.

Subsequentemente, tomando-se como base o número médio de indivíduos por domicílio em cada faixa etária considerada, a amostra de indivíduos foi transformada em amostra de domicílios.

Tabela 5.5: Tamanho da amostra de indivíduos para inquérito de hepatite B e C, segundo faixa etária e capital – Nordeste, 2000.

Capitais	Tamanho da amostra por faixa etária	
	10 a 19	20 e mais anos
Maceió	201	188
Fortaleza	548	538
João Pessoa	150	155
Teresina	202	174
Aracajú	120	119
Salvador	609	642
São Luiz	251	209
Recife	336	382
Natal	182	182
Total	2.599	2.600

- Fração de amostragem da população de 10 a 19 anos para o NE = $2.598/2.156.187 = 0,00120490$.
- Fração de amostragem da população de 20 e mais anos para o NE = $2.600/6.159.140 = 0,00019849$.

Fonte: Universidade de Pernambuco, 2010⁹.

Na segunda etapa, foi retirada uma amostra de setores censitários proporcional ao número de domicílios. A completude dessa etapa implicou na execução de algumas fases de preparação. Ao se comparar o número de domicílios e de indivíduos por setor censitário, verificou-se que esses não eram homogêneos e, conseqüentemente, o sorteio de setores e o posterior sorteio de domicílios não produziram amostra equiprobabilística. Para contornar esse problema, duas estratégias foram seguidas. A primeira foi tornar os setores censitários mais semelhantes em suas dimensões. Definiu-se então um limite do coeficiente de variação a ser aceito e os setores censitários foram divididos ou agrupados para se enquadrarem dentro dos limites propostos. O agrupamento de setores censitários menores foi feito não por proximidade geográfica, mas por similitude das condições socioeconômicas, expressa pela média de anos de estudo do chefe da família. Foram constituídas assim as unidades primárias de amostragem.

Tabela 5.6: Formação das UPA's – Ordenação dos Setores Censitários, em ordem crescente, pela média de anos de estudos dos responsáveis pelos domicílios particulares permanentes.

Anos de estudo do Chefe da família	Número de domicílios	UPA	Setor Censitário
2,5	215	1	261160605230117
2,6	23	2	261160605180060
2,8	120	2	261160605190154
2,9	462	3	261160605180077
3,0	153	2	261160605230074
3,0	220	4	261160605220016
3,0	172	5	261160605210164
3,3	153	5	261160605200208
3,1	262	6	261160605220017
3,2	452	7	261160605180064
3,3	191	8	261160605220101
3,3	77	8	261160605180064
3,3	244	9	261160605220101
3,3	373	10	261160605200210

Fonte: Universidade de Pernambuco 2010⁹.

Para sorteio das UPAs, com probabilidade proporcional ao número de domicílios, os setores censitários foram ordenados em ordem crescente em relação à média de anos de estudo do chefe da família e feita contagem cumulativa dos domicílios por UPA. Foi realizado então um sorteio sistemático das UPAs e definido um intervalo amostral. Considerando-se o sorteio de 30 UPAs por capital, procedeu-se o sorteio de um número menor do que o valor do intervalo amostral e adicionando-o, consecutivamente, obteve-se as UPAs, que compuseram a amostra em cada capital.

Tabela 5.7: Contagem cumulativa dos domicílios particulares permanentes para realização do sorteio das UPAs com probabilidade proporcional ao número de domicílios.

Anos de estudo do Chefe da família	Número de domicílios	Número de domicílios por UPA	Número da UPA	Setor Censitário
2,5	215	215	1	261160605230117
2,6	23	296	2	261160605180060
2,8	120		2	261160605190154
3,0	153		2	261160605180078
2,9	462	462	3	261160605180077
3,0	220	220	4	261160605230074
3,0	172	325	5	261160605220016
3,3	153		5	261160605210106
3,1	262	262	6	261160605210164
3,2	452	452	7	261160605200208
3,3	191	268	8	261160605220017
3,3	77		8	261160605180064
3,3	244	244	9	261160605220101
3,3	373	373	10	261160605200210

Fonte: Universidade de Pernambuco, 2010⁹.

No quarto estágio foi feita uma amostra dos quarteirões também com probabilidade proporcional ao número de domicílios. No quinto estágio, dentro de cada quarteirão foi realizado um sorteio sistemático dos domicílios. A seleção dos indivíduos dentro de cada faixa etária e para cada tipo de hepatite em cada domicílio seguiu uma estratégia previamente definida.

Unidade 5

Análise e Apresentação de Resultados

1. Inferência e Intervalos de Confiança

Na primeira etapa da análise de um corte transversal, dependendo do objetivo do estudo, são calculadas as médias com respectivos intervalos de confiança.

A média e o erro-padrão para amostra randômica simples obtida de uma população infinita são calculados, respectivamente, por $\bar{x} = \Sigma x/n$ e $EP = s/\sqrt{n}$, em que s corresponde ao desvio-padrão e n ao tamanho da amostra. Se a amostra é obtida de uma população finita, torna-se necessário um ajuste do erro-padrão, utilizando-se a seguinte fórmula:

$$EP = s/\sqrt{n} \times \sqrt{(1-f)}, \text{ sendo } f = n/N$$

A proporção e o erro-padrão de uma proporção para amostra randômica simples obtida de uma população finita é estimada, respectivamente, por $p = r/n$ e $EP = \sqrt{pq/n}$, sendo $q = 1-p$. Para amostra obtida de uma população finita, o ajuste é alcançado utilizando-se a seguinte fórmula:

$$SE = \sqrt{pq(1-f)/n}$$

Para cálculo do intervalo de confiança multiplica-se o valor de Z ou t (quando indicado), correspondendo a determinado nível de probabilidade, pelo erro-padrão: $\pm Z \times EP$ (intervalo de confiança para a média) ou $p \pm Z \times ER$ (intervalo de confiança para a proporção).

Para outros desenhos amostrais como amostra estratificada com ou sem partilha proporcional, amostra em múltiplos níveis etc., as fórmulas serão adaptadas para levar em consideração o efeito do desenho sobre a precisão das estimativas obtidas.

2. Razão de Prevalência e *Odds Ratio*

As medidas de associação utilizadas nos estudos transversais são a razão de prevalência e a *Odds Ratio*. Kirkwood e Sterne¹⁹ apontam que, na literatura médica moderna, a análise estatística de desfechos binários é comumente feita pela *Odds Ratio*, independentemente do desfecho ser frequente ou raro.

A razão de prevalência e o respectivo erro-padrão são estimados por:

$$RP = a/a + b / c/c + d$$

e para o cálculo do erro-padrão é necessário inicialmente calcular o logaritmo da razão de prevalência e do respectivo erro-padrão:

$$EP(\log RP) = \sqrt{[1/a - 1/a+c + 1/c - 1/c+d]}$$

para subsequentemente estimar o intervalo de confiança do logaritmo da razão de prevalência e o antilogaritmo do intervalo obtido.

A *Odds Ratio* é estimada por axd/bxc e o erro-padrão do logaritmo do OR por

$$EP(\log OR) = \sqrt{1/a + 1/b + 1/C + 1/d}$$

3. Confusão e Interação

A distorção da estimativa da medida de associação, por meio de fatores de confusão ou a presença de interação, pode ser testada utilizando-se a técnica de estratificação ou os modelos de regressão. Resumidamente, para avaliar a existência de confundimento, compara-se a medida de associação bruta com aquela ajustada, essa última correspondendo a uma média ponderada (pelo tamanho do estrato) das medidas de cada um dos estratos. A diferença nos valores entre a medida bruta e a ajustada indica a presença de fator de confundimento. Para avaliar a presença de interação, comparam-se as estimativas da medida de associação nos diferentes estratos utilizando-se um teste de homogeneidade. A diferença nessas medidas expressa a presença de interação.

4. Modelos de Regressão (Poisson, Logístico)

As regressões de Poisson ou Logística podem ser utilizadas para comparar dois ou mais grupos de exposição, e ajustar pelo efeito de um ou mais fatores de confusão. Ambos os modelos utilizam uma escala logarítmica, e o cálculo do antilogaritmo fornece as medidas de associação e os respectivos intervalos de confiança. A estrutura dos dois modelos é semelhante:

Poisson

$$\log(\text{taxa}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

Logístico

$$\log(\text{odds do desfecho}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

O uso de um desses modelos tem sido alvo de debate na literatura e, para maior aprofundamento sobre a seleção de um ou de outro modelo sugeriríamos os artigos de Barros e Hirakata²¹; Coutinho et al.²² e Reichenheim e Coutinho²³. Vários programas estatísticos disponibilizam os dois modelos de regressão. Como a *Odds Ratio* vem sendo utilizada com mais frequência na análise dos estudos transversais do que a razão de prevalência, o modelo logístico é também o mais utilizado.

5. Modelo Multinível

Existem várias técnicas estatísticas que permitem incorporar diferentes níveis de análise em um mesmo modelo possibilitando, por exemplo, incluir em uma mesma análise fatores ligados ao indivíduo e fatores ligados ao ambiente. Uma alternativa é o modelo contextual que corresponde a uma extensão dos modelos individuais, como o modelo logístico, ao qual se adiciona um preditor nível de grupo. O modelo contextual tem como limitação não levar em consideração o agrupamento *clustering* dos indivíduos, ignorando que o desfecho de um indivíduo possa ser dependente do desfecho dos demais do mesmo grupo.

Outra alternativa, que ultrapassa essa limitação, é a modelagem multinível que introduz os efeitos aleatórios como forma de lidar com a dependência intragrupo. Um exemplo dessa abordagem é a análise do inquérito nacional de hepatites virais para a hepatite A no qual foram considerados três níveis: o primeiro nível correspondendo ao indivíduo, o segundo nível correspondendo à família e o terceiro nível correspondendo ao setor censitário. As variáveis socioeconômicas de cada um dos níveis estão ajustadas pelas correlatas dos demais²⁴.

Unidade 6

Vantagens e Limitações dos Inquéritos de Base Populacional

Vantagens

- Estudos de Prevalência podem ser mais rápidos, baratos e operacionalmente mais simples que estudos caso-controle ou de coorte, na dependência da pergunta e da abrangência geográfica.
- Estudos de Prevalência devem ser baseados, sempre que possível, em amostras representativas da população em geral.

Limitações

- Não são adequados para doenças raras e de curta duração devido a baixa frequência dos eventos a serem estudados.
- Não podem, de forma geral, testar hipóteses etiológicas (causa e efeito) já que exposição e o status da doença são medidos simultaneamente.
- Não medem incidência, mas em estudos de prevalência com determinação de faixa etária pode-se estimar indiretamente a incidência por meio de técnicas estatísticas.

Veja na tabela ao lado os potenciais erros relacionados ao tipo de inquérito realizado e seleção da população de estudo.

Tabela 5.8: Potenciais erros e problemas segundo a população dos estudos de prevalência para infecção pelo HIV-1.

Inquérito de base populacional	Inquérito em populações de alto risco	Inquéritos em clínicas de pré-natal
Viés de não resposta	Não representatividade	Não representatividade <ul style="list-style-type: none">• Só mulheres• Idade reprodutiva• Período sexualmente ativo
Grande tamanho de amostra	Viés de não resposta	Limitado a cobertura dos serviços e da área de abrangência
Alto custo	Incerteza no tamanho das populações	Incerteza nos fatores de ajuste

Fonte: Brookmeyer 2010²⁵.

Para refletir

Quais as vantagens e limitações de um inquérito, como a Vigitel, que é realizado por via telefônica? Acesse o link: <http://portal.saude.gov.br/portal/saude/profissional/area.cfm?id_area=1521>.

Considerações Éticas

Na Plataforma Brasil, podemos encontrar a Resolução nº 196/96 do Conselho Nacional de Saúde que rege os princípios éticos das pesquisas envolvendo seres humanos no Brasil. Assim como qualquer projeto de pesquisa realizado em seres humanos, o protocolo deve ser submetido ao Comitê de Ética e Pesquisa por meio da Plataforma Brasil <<http://aplicacao.saude.gov.br/plataformabrasil/login.jsf>>.

Após o cadastro e a aprovação do projeto de pesquisa, a assinatura do Termo de Consentimento Livre e Esclarecido é solicitada durante a realização da pesquisa de campo. Projetos envolvendo populações consideradas vulneráveis (ver glossário) devem ser analisados também pela Comissão Nacional de Ética em Pesquisa (Conep).

Referências

- 1 – Hennekens, C., Buring, J. & Mayrent, S. (1987) *Epidemiology in medicine*. Boston. Library of Congress Catalog Card.
- 2 – Waldman, E., Novaes, H., Albuquerque, M., Latorre, M., Ribeiro, M., Vasconcelos, M. & al., e. (2008) Inquéritos populacionais: aspectos metodológicos, operacionais e éticos, *Rev Bras Epidemiol*, 11(11), 168-179.
- 3 – Cunha, V. S. S. S. (2011) Estudos transversais, in: N. d. A. Filho & M. L. Barreto (Eds) *Epidemiologia & Saúde: fundamentos, métodos e aplicações*. Rio de Janeiro. Guanabara Koogan.
- 4 – Szwarcwald, C. L. & Damacena, G. N. (2008) Amostras complexas em inquéritos populacionais: planejamento e implicações na análise estatística dos dados, *Rev Bras Epidemiol*, 11(1), 38-45.
- 5 – Brookmeyer, R. (2010) Measuring the HIV/AIDS epidemic: approaches and challenges, *Epidemiol Rev*, 32(1), 26-37.
- 6 – Travassos, C., Viacava, F. & Laguardia, J. (2008) Os suplementos saúde na pesquisa nacional por amostra de domicílios (PNAD) no Brasil, *Rev Bras Epidemiol*, 11(1), 98-112.
- 7 – Viacava, F. (2009) Suplemento saúde da PNAD. Disponível em: <http://www.pns.icict.fiocruz.br/arquivos/Atualiza%C3%A7%C3%A3o%20da%20PNAD-Saude%20Viacava.pdf> (acessado em 20 Ago 2012).
- 8 – IBGE (2010) *Pesquisa Nacional por Amostra de Domicílios: um panorama da saúde no Brasil. Acesso e utilização dos serviços, condições de saúde e fatores de risco e proteção à saúde*, 2008, Rio de Janeiro, 245.
- 9 – Universidade de Pernambuco (2010) Estudo de prevalência de base populacional das infecções pelos vírus das hepatites A, B e C nas capitais do Brasil, Brasília, 295.

10 – Ministério da Saúde (2011) Projeto SB Brasil 2010 – Pesquisa Nacional de Saúde Bucal. Disponível em: http://189.28.128.100/dab/docs/geral/projeto_sb2010_relatorio_final.pdf (acessado em 20 Dez 2012).

11 – Center for Disease Control and Prevention (2012) National Health and Nutrition Examination Survey. Disponível em: <http://www.cdc.gov/nchs/nhanes.htm/> (acessado em 03 Jul 2012).

12 – Ministério da Saúde (2010) Vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico. Disponível em: <http://www2.datasus.gov.br/DATASUS/index.php?area=0207> (acessado em 19 Ago 2012).

13 – Ministério da Saúde (2010) Projeto de planejamento da pesquisa nacional de saúde. Disponível em: <http://www.pns.icict.fiocruz.br/> (acessado em 11 Dez 2012).

14 – Almeida-Filho, N. & Barreto, M. L. (2011) *Epidemiologia & Saúde: Fundamentos, Métodos, Aplicações*. Rio de Janeiro

15 – Costa, Z. B., Machado, G. C., Avelino, M. M., Gomes Filho, C., Macedo Filho, J. V., Minuzzi, A. L., Turchi, M. D., Stefani, M. M., de Souza, W. V. & Martelli, C. M. (2009) Prevalence and risk factors for Hepatitis C and HIV-1 infections among pregnant women in Central Brazil, *BMC Infect Dis*, 9, 116.

16 – Siqueira, J. B., Martelli, C. M., Maciel, I. J., Oliveira, R. M., Ribeiro, M. G., Amorim, F. P., Moreira, B. C., Cardoso, D. D., Souza, W. V. & Andrade, A. L. (2004) Household survey of dengue infection in central Brazil: spatial point pattern analysis and risk factors assessment, *Am J Trop Med Hyg*, 71(5), 646-651.

17 – Siqueira-Júnior, J. B., Maciel, I. J., Barcellos, C., Souza, W. V., Carvalho, M. S., Nascimento, N. E., Oliveira, R. M., Morais-Neto, O. & Martelli, C. M. (2008) Spatial point analysis based on dengue surveys at household level in Central Brazil, *BMC Public Health*, 8, 361.

- 18 – Victora, C. G., Barreto, M. L., do Carmo Leal, M., Monteiro, C. A., Schmidt, M. I., Paim, J., Bastos, F. I., Almeida, C., Bahia, L., Travassos, C., Reichenheim, M. & Barros, F. C. (2011) Health conditions and health-policy innovations in Brazil: the way forward, *Lancet*, 377(9782), 2042-2053.
- 19 – Kirkwood, B. R. & Sterne, J. A. C. (2003) *Essential Medical Statistics*. Victoria. Blackwell Science.
- 20 – Moser, C. A. & Kalton, G. (1975) *Survey methods in social investigation*. London. Heinemann Educational Books.
- 21 – Barros, A. J. & Hirakata, V. N. (2003) Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio, *BMC Med Res Methodol*, 3, 21.
- 22 – Coutinho, L. M., Scazufca, M. & Menezes, P. R. (2008) Methods for estimating prevalence ratios in cross-sectional studies, *Rev Saude Publica*, 42(6), 992-998.
- 23 – Reichenheim, M. E. & Coutinho, E. S. (2010) Measures and models for causal inference in cross-sectional studies: arguments for the appropriateness of the prevalence odds ratio and related logistic regression, *BMC Med Res Methodol*, 10, 66.
- 24 – Ximenes, R. A. A., Martelli, C. M., Merchan-Hamann, E., Montarroyos, U. R., Braga, M. C., de Lima, M. L., Cardoso, M. R., Turchi, M. D., Costa, M. A., de Alencar, L. C., Moreira, R. C., Figueiredo, G. M. & Pereira, L. M. (2008) Multilevel analysis of hepatitis A infection in children and adolescents: a household survey in the Northeast and Central-west regions of Brazil, *Int J Epidemiol*, 37(4), 852-861.
- 25 – Brookmeyer R. (2010) Measuring the HIV/AIDS epidemic: approaches and challenges. *Epidemiol Rev*. Apr;32(1):26-37.

Glossário

Estudo de prevalência ou seccional ou transversal – Estudos nos quais a presença ou ausência de doença ou outra variável relacionada à saúde são determinadas para cada membro da população de estudo ou em amostra representativa em um período de tempo. Estudo epidemiológico nos quais fator e efeito são observados no mesmo momento.

Fator de risco – Aspecto do comportamento ou estilo de vida, exposição ambiental ou característica ao nascimento ou hereditária, a qual com base em evidências epidemiológicas, é associada às condições relacionadas à saúde consideradas importantes para prevenção.

Inquérito de saúde – Coleção sistemática de dados relativos à saúde e à doença em população humana dentro de determinada área geográfica.

Odds ratio – Em estudos de prevalência, a medida de associação pode ser denominada razão de prevalência ou prevalence odds ratio.

População vulnerável – Referente ao estado de pessoas ou grupos que por quaisquer razões ou motivos tenham a sua capacidade de autodeterminação reduzida, sobretudo no que se refere ao consentimento livre e esclarecido.

Prevalência no período ou lápsica – Abrange todos os casos presentes em período de tempo especificado (exemplo: durante o ano de 2012).

Prevalência no ponto ou instantânea ou momentânea – Quantifica a proporção de indivíduos em uma população que apresenta um evento de saúde em um momento específico. Estima probabilidade (risco) de um indivíduo apresentar o evento em dado momento de tempo (exemplo: 31 de dezembro).

Viés de amostragem – Erro relativo à falha de assegurar que todos os membros da população de referência tenham a mesma chance de serem selecionados na amostra.

Viés de causalidade reversa – O erro ocasionado devido à mensuração simultânea da exposição e desfecho nos estudos de prevalência.

Viés de observador – Pesquisadores de campo devem ser treinados para fazer perguntas de maneira uniforme, para evitar tendenciosidade.

Viés de prevalência – Qualquer desvio de resultados ou inferência do resultado obtido.

Viés de publicação – A influência dos resultados de estudo sobre a chance de sua publicação e a tendência dos investigadores, revisores e editores em submeter ou aceitar estudos que apresentam resultados em que a associação é encontrada.

Viés de seleção de participante – Amostras não randômicas podem distorcer a frequência da distribuição das doenças em uma população.

Viés de seleção – Introdução de erro devido à diferença sistemática entre as características dos selecionados e dos não selecionados para determinado estudo.

Viés de sobrevivente – Casos prevalentes observados tendem a excluir casos com rápida evolução e óbitos precoces (já que eles estão ausentes no momento da seleção) e superestimam casos de longa sobrevivência.

Módulo 6

Análise de Séries Temporais na Epidemiologia

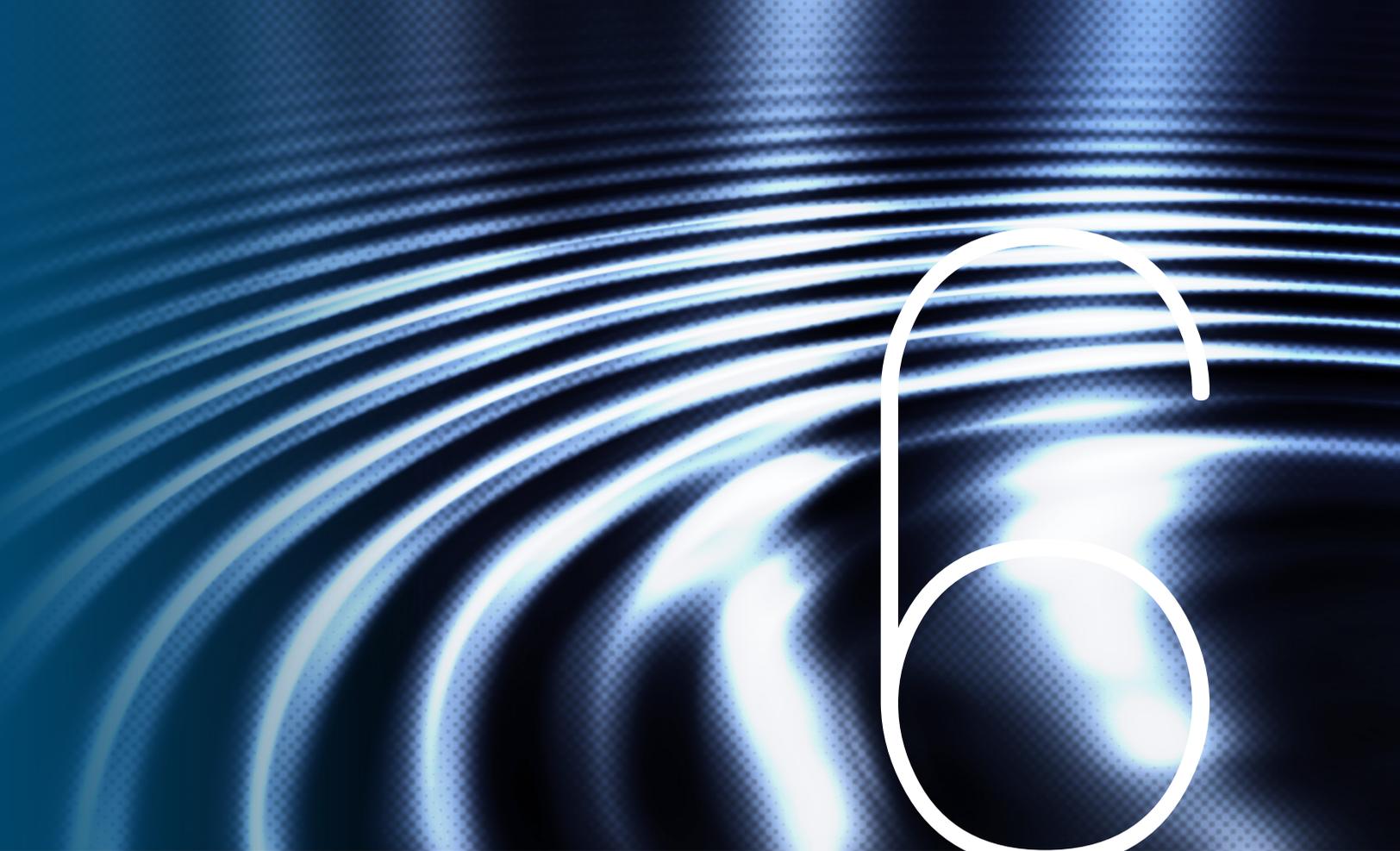
José Leopoldo Ferreira Antunes

Universidade de São Paulo (USP), Faculdade de Saúde Pública, Departamento de Epidemiologia, São Paulo/SP, Brasil.

Introdução

A perspectiva de antever o futuro sempre encantou a humanidade. Saber o que acontecerá antes mesmo que os primeiros sinais se manifestem, pode propiciar um melhor aproveitamento de efeitos benéficos dos eventos futuros e uma preparação antecipada para eventuais efeitos adversos. Talvez até mais importante que antecipar os resultados seja o reconhecimento de tudo quanto possa interferir favorável ou desfavoravelmente nos processos em curso, para permitir o planejamento. De fato, lançar os olhos para o futuro é importante para qualquer tomada de decisão no presente!

A Epidemiologia é a disciplina que se aplica ao estudo da distribuição dos problemas de saúde, dos determinantes que influenciam esses problemas e dos esforços para seu controle¹. Para essa área de estudos, a necessidade de antever o futuro e, com base nessa antevisão, intervir nos processos do presente é muito mais que mera curiosidade ou interesse mesquinho. De fato, é assunto de vida ou morte, pois a redução da carga de doença na população depende da efetividade desses esforços.



Nesse sentido, o presente livro focaliza séries temporais, o ramo da Epidemiologia que permite uma antevisão de futuros cenários da distribuição de doenças na população e dos fatores que podem modificar esta distribuição para melhor ou para pior. O livro foi organizado em duas seções. Na primeira delas, são introduzidos os aspectos conceituais sobre séries temporais, enquanto o recurso lógico para organizar as medidas epidemiológicas no tempo. Na segunda seção, discutiremos os aspectos práticos desta metodologia, exercitando as técnicas de análise das séries temporais.

Unidade 1

Séries temporais: Aspectos Conceituais da Organização das Medidas de Doença no Tempo

Imagine um gestor de Saúde que precise antever o coeficiente de mortalidade infantil em sua região, com o intuito de programar iniciativas de promoção de saúde. Sabe-se que este coeficiente é influenciado por algumas características da população, como condições socioeconômicas, saneamento básico, provisão e acesso aos serviços de Saúde. Mas estas informações podem não ser atuais, confiáveis ou facilmente recuperadas. Além disso, as relações entre a mortalidade infantil e essas características são complexas e podem não ser rapidamente equacionadas. Com isso, quais seriam as primeiras perguntas que esse gestor se faria, para prever o valor deste indicador?

A primeira pergunta parece ser instintiva: “Qual foi o valor deste coeficiente no ano anterior?” Continuando sua reflexão sobre o tema, o gestor decerto se perguntará: “É justo pensar que o que ocorreu no passado recente deva se repetir no futuro imediato?” Quer dizer, essa segunda pergunta amplia o escopo de seu esforço de antevisão do futuro. Agora, o gestor quer apurar sua previsão levando em consideração não apenas o que ocorreu no último ano, mas também perguntando sobre a movimentação recente que esta medida sofreu no tempo. Daí a terceira hipotética questão: “Como a mortalidade infantil tem se comportado ao longo dos últimos anos?” O estudo epidemiológico das séries temporais contempla justamente essa preocupação em derivar conhecimentos sobre a movimentação recente das medidas de interesse em saúde para a previsão de resultados e para o conhecimento de fatores que interferem sobre esses resultados.

Séries temporais foram definidas de modo breve e sucinto: “São sequências de dados quantitativos relativos a momentos específicos e estudados segundo sua distribuição no tempo”². Essa definição indica a aplicabilidade desse recurso de análise a diferentes finalidades e campos de conhecimento. Como aperfeiçoar o fluxo de estoque de um almoxarifado? Como programar a compra de matéria prima para uma atividade industrial? Como dimensionar o fluxo de vendas em um empreendimento comercial? Estes são apenas alguns exemplos de aplicação da análise de séries temporais para finalidades não imediatamente relacionadas à saúde. Naturalmente, nosso enfoque focalizará as medidas de saúde e a aplicação da análise de séries temporais para os estudos epidemiológicos.

Esta primeira unidade aborda aspectos conceituais relacionados a essa estratégia de análise dos estudos epidemiológicos, a qual consiste em organizar no tempo as informações sobre as doenças e derivar conhecimento desta análise.

1. Séries temporais e música

Vimos que as séries temporais são uma forma de organizar no tempo as informações quantitativas sobre aspectos relacionados à saúde. A música também é uma forma de organizar no tempo um tipo específico de informação: os registros sonoros. A teoria musical reconhece três elementos da música, os quais podem ser pensados em correspondência com a análise epidemiológica das séries temporais: melodia, harmonia e ritmo.

1.1 Melodia e tendência

Embora seja difícil definir com precisão, a melodia pode ser pensada como uma sequência de sons organizados de modo a fazer sentido musical. A Figura 6.1 apresenta um trecho melódico bastante conhecido. De fato, todos se recordam da linha melódica que acompanha o trecho “Ouviram do Ipiranga às margens plácidas”. Sua representação gráfica na forma de partitura deixa bem claro que, excetuando as duas últimas notas, a melodia tem uma progressão das notas graves, representadas pelas linhas inferiores do pentagrama, para as notas agudas, as que aparecem na parte mais elevada da partitura.

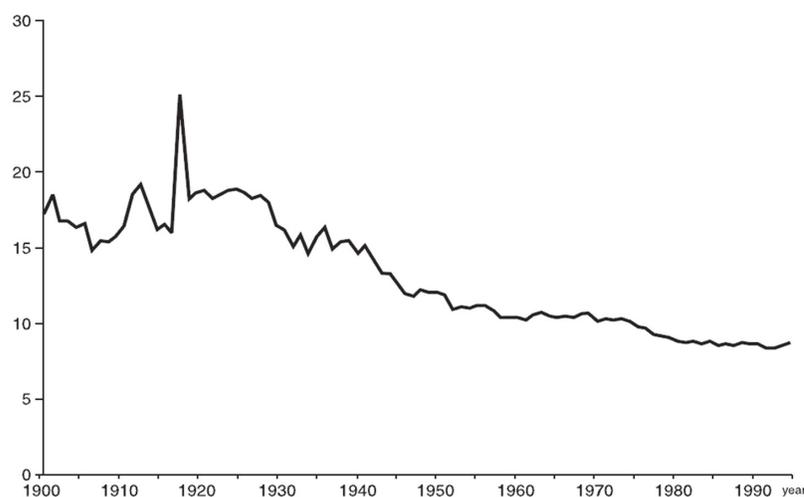
Os movimentos melódicos em direção às notas mais graves ou às mais agudas podem ser projetados aos movimentos das tendências nas séries temporais. Observe, na Figura 6.2, a série temporal do coeficiente de mortalidade geral na Cidade de São Paulo, de 1900 a 1994. A imagem da série deixa claro que, de modo geral, essa medida evoluiu ao longo do século reduzindo a magnitude de valores; isto é, passando de valores mais elevados na escala vertical para valores menos elevados. Isto configura uma tendência decrescente.

Figura 6.1: Compassos iniciais do Hino Nacional Brasileiro.



Quando estudamos séries temporais em estudos epidemiológicos, o primeiro elemento da análise deve focalizar qual é a tendência da medida. O conceito de tendência também pode ser definido de forma breve e sucinta: “um movimento prolongado em uma série ordenada”¹. Obviamente, a tendência pode ser decrescente, como na Figura 6.2, mas também pode ser crescente ou estacionária.

Figura 6.2: Coeficiente de mortalidade geral, Cidade de São Paulo, 1900–1994.



Fonte: Antunes, 1998³.

Complicando um pouco mais, uma mesma série temporal pode apresentar trechos com diferentes tendências. Observe como isso de fato ocorre na Figura 6.3, que mostra a série temporal da mortalidade infantil na Cidade de São Paulo, abrangendo o mesmo período. A tendência secular de declínio dessa medida foi claramente interrompida no período de 1961 a 1973, quando a mortalidade infantil cresceu de modo consistente. O que teria acontecido nesse período para justificar esta observação? O que mais de interessante pode ser apreendido da apreciação visual desta série temporal? Em cada um dos primeiros anos do monitoramento, ocorreram em torno de 220 a 230 óbitos de crianças com menos de 1 ano de idade, para cada 1.000 nascidas vivas. O que pensar da magnitude dessa medida?

Figura 6.3: Coeficiente de mortalidade infantil, cidade de São Paulo, 1900-1994.



Fonte: Antunes, 1998³.

1.2 Harmonia e associação

Mas a música não se limita à linha melódica. Imagine a diferença entre um violinista ressoando seu instrumento sozinho no palco e outro que toca acompanhado de toda uma orquestra sinfônica! A mesma linha melódica executada pelo violinista decerto terá outro brilho quando reproduzida em harmonia com a orquestra. Há atividades que definitivamente ficam melhor quando ocorrem de modo concomitante. O vinho, que deve ser tomado durante as refeições, para harmonizar com o sabor dos alimentos. Assistir ao final do campeonato sozinho em casa na televisão, ou no estádio com o calor de toda a torcida!

No piano, enquanto a mão direita dedilha a melodia, a mão esquerda toca as outras notas dos acordes que compõem a harmonia. Como elementos musicais complementares, a melodia é o desenvolvimento horizontal da música, enquanto a harmonia é seu desenvolvimento vertical, representado pelos acordes formados por sons concomitantes. De modo análogo, quando estudamos a tendência e sintonizamos nossa atenção no desenvolvimento horizontal da série no tempo, precisamos atentar também para a complexidade vertical da série temporal, identificando como suas medidas se harmonizam, isto é, associam-se com informações adicionais sobre os fenômenos relacionados.

Com essa ideia em mente, reveja a série temporal da Figura 6.2 e identifique o pico de mortalidade geral que ocorreu na Cidade de São Paulo no ano de 1918. Por meio da análise de tendência desta série temporal, pode-se estimar que o valor observado em 1918 foi quase 50% mais elevado que a média entre os anos imediatamente anterior e posterior. Esta informação é

compreensível por si só; ela indica que, de fato, algo muito ruim ocorreu naquele ano.

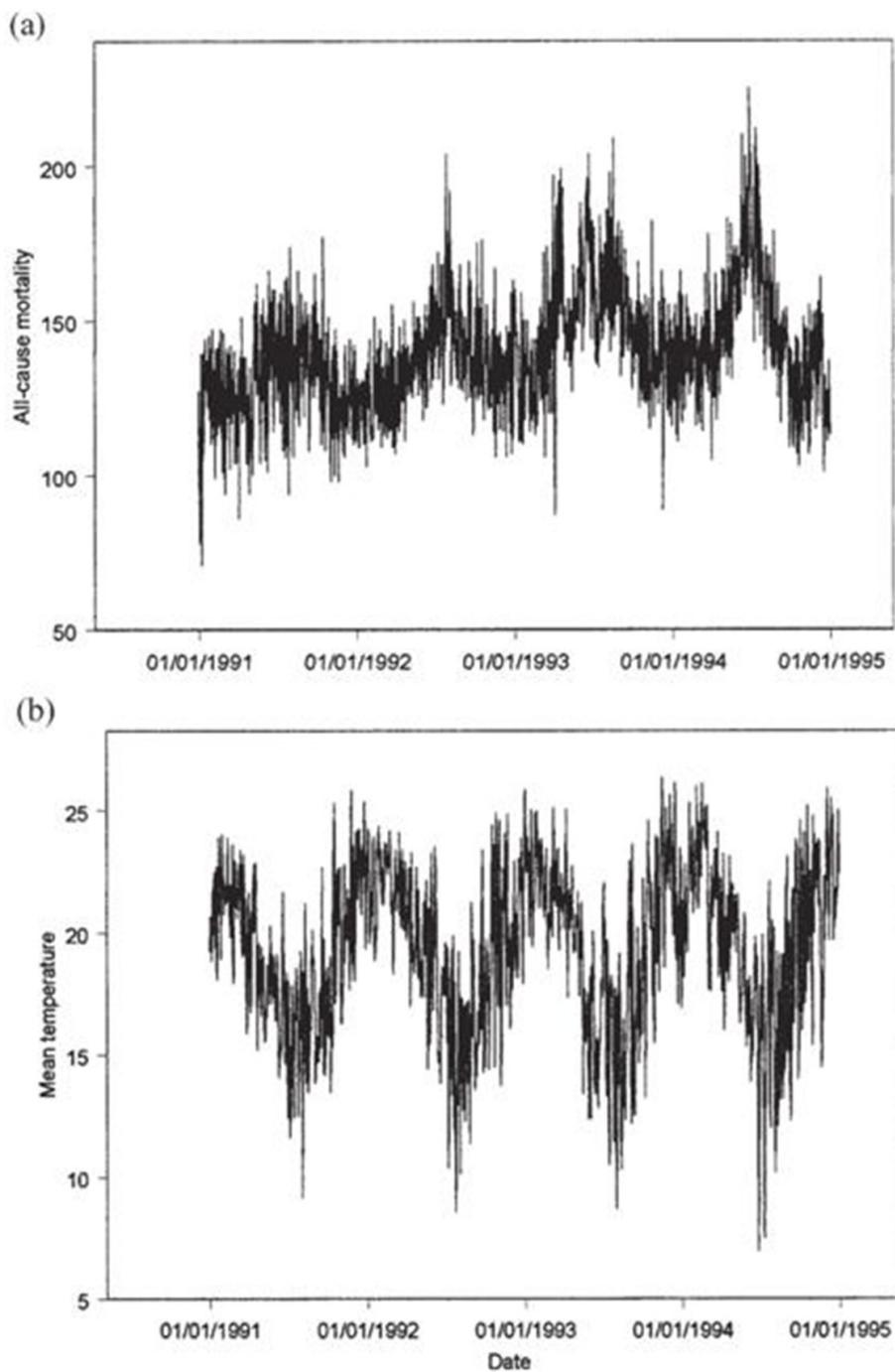
Mas o significado desta informação decerto se amplifica quando procuramos associá-la com outros conhecimentos sobre o que se passou na Cidade de São Paulo. Como se sabe, 1918 foi o ano da gripe espanhola, de triste lembrança. Naquele ano, ocorreu um grave surto de gripe, com enorme impacto na mortalidade geral. No mesmo estudo do qual foi reproduzida a Figura 6.2, foi delineada a série temporal da esperança de vida ao nascer, a qual mostrou que esta medida se reduziu fortemente. Isto é, foi cerca de 50 anos, no ano anterior, para menos de 32 em 1918, justamente em função de muitas crianças, adolescentes e adultos jovens terem sido vitimados pela gripe.

A perspectiva de harmonizar a interpretação das tendências observadas nas séries temporais com outras informações que estejam disponíveis sobre o fenômeno em questão diz respeito ao estudo de associação, que é um recurso bastante usual da análise epidemiológica. Estas informações adicionais podem ser qualitativas, auxiliando a interpretação dos motivos que justificam o aumento, a diminuição ou a persistência dos valores de uma medida de interesse para a saúde. Estas informações adicionais também podem ser quantitativas, dando ensejo à aplicação de técnicas de análise estatística para estimar sua associação com a série temporal que se tenta explicar.

Exemplificando o uso de informações qualitativas para justificar movimentos de tendência das séries temporais, observa-se que o estudo original, no qual foi delineada a série temporal da Figura 6.3, aventou duas hipóteses para explicar a inversão de tendência da mortalidade infantil e o crescimento observado nos anos 60 e início dos 70. A primeira delas foi a diminuição do valor real do salário mínimo, parâmetro que regulava a remuneração de uma parcela ponderável das famílias residentes na cidade. A outra hipótese aventava o desmame precoce, a comercialização de leite em pó e a má qualidade da água de abastecimento público, com a qual o leite em pó era preparado para o consumo infantil³.

Vejam agora outro exemplo de associação dos dados obtidos em uma série temporal para criar conhecimento em Saúde. Neste caso, o estudo de associação utilizou medidas quantitativas organizadas em outra série temporal para avaliar a concomitância entre dois fenômenos de interesse para a saúde. Examine a Figura 6.4, na qual se mostra a série temporal da mortalidade geral diária em São Paulo, do início de 1991 até o final de 1994. A segunda série temporal delineada organiza os dados de temperatura média diária na mesma cidade e no mesmo período. Considerando a associação entre as duas medidas, o que se depreende da inspeção visual das duas séries temporais? Que conhecimento pode ser inferido?

Figura 6.4: Séries temporais: (a) mortalidade geral por dia e (b) temperatura média diária na Cidade de São Paulo, 1991–1994.



Fonte: Gouveia et al., 2003⁴.

As temperaturas médias diárias são, com certeza, mais baixas no inverno (meados de cada ano) e mais elevadas no verão (período do fim de ano e início do ano seguinte). Mas a mortalidade geral é mais elevada no inverno que no verão. No estudo que delineou as séries temporais reproduzidas na Figura 6.4, os autores concluíram haver uma associação inversa entre os dois fenômenos. Isto é, nos períodos em que os dias são mais quentes, a mortalidade geral diária tende a ser menor, e vice-versa, a mortalidade é tendencialmente mais elevada nos períodos de frio.

Diz-se que há uma associação direta ou positiva entre duas séries temporais quando os valores de ambas aumentam ou diminuem concomitantemente. De modo complementar, a associação é inversa ou negativa se o aumento de uma medida corresponde à diminuição da outra, como no caso discutido no parágrafo anterior. E se diz que não há associação quando não há correspondência entre as mudanças de valor de ambas as séries. A associação entre séries temporais não necessariamente tem origem causal. O aumento de uma variável pode não ser causa do aumento ou declínio da outra, ambas podem ter causas comuns, e sua associação pode ser reflexo de processos mais complexos.

1.3 Ritmo e sazonalidade

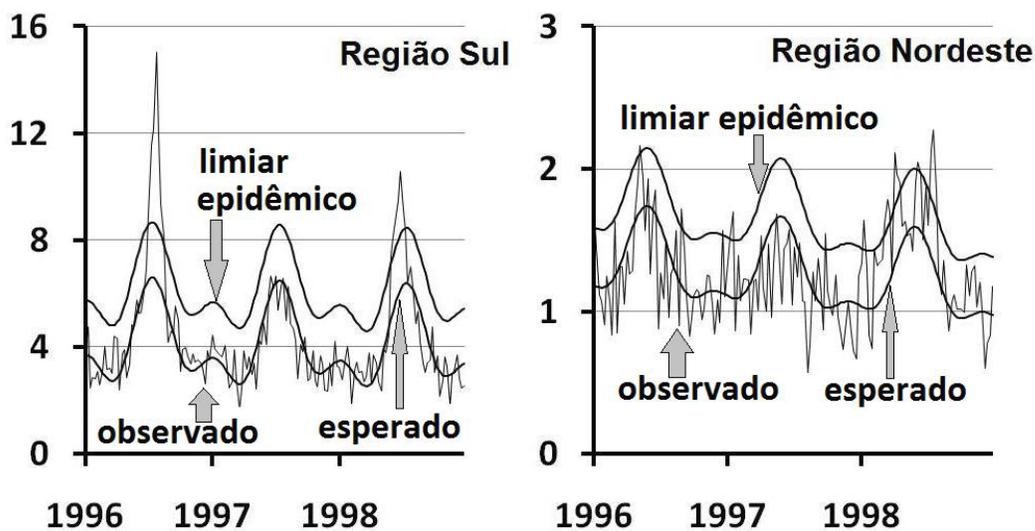
O ritmo é possivelmente o elemento mais instintivo da teoria musical. O coração bate de forma ritmada; o sol parece girar em torno da terra de forma ritmada; as estações do ano se sucedem de forma ritmada. Na música, o ritmo envolve estrofes, refrãos e repetições. Também se dão de forma ritmada os movimentos da dança, ou mesmo de quem acompanha distraidamente uma canção. Na notação musical, o ritmo é registrado por meio da marcação do tempo em que cada nota deve soar e pela divisão da partitura em compassos.

A percepção de que os fenômenos de interesse para a saúde também podem apresentar repetições de certa forma organizadas no tempo, ou seja, de que há algum ritmo a ser reconhecido na análise de séries temporais, é muito importante para a Epidemiologia. Estamos falando das variações sazonais e cíclicas que afetam a medida de muitas doenças.

O movimento sazonal da temperatura média diária que foi representado na Figura 6.4 é bem conhecido de todos: faz mais frio no inverno que no verão... A originalidade do estudo do qual foram reproduzidas as imagens da Figura 6.4 diz respeito justamente à existência de variação sazonal também no número diário de mortes. Quem recuperar o manuscrito original na página que a revista mantém na internet (o acesso é livre) poderá refletir mais detalhadamente sobre os motivos que justificam esta variação sazonal.

A Figura 6.5 fornece outro exemplo de construção de conhecimento em saúde com base na análise da variação sazonal das séries temporais. Esta figura reproduz, para as regiões Sul e Nordeste do Brasil, a mortalidade semanal de pessoas com 65 anos ou mais por pneumonia e *influenza*. O período de monitoramento foi de 1996 a 1998, os três anos que antecedem a introdução do programa nacional de vacinação de idosos contra a gripe.

Figura 6.5: Séries temporais da mortalidade semanal de idosos (65 anos ou mais) por *influenza* e pneumonia nas regiões Sul e Nordeste, 1996 a 1998. Mortalidade observada, mortalidade esperada e limiar epidêmico previsto



Fonte: Oliveira, 2012⁵.

Além da série temporal da mortalidade observada, foi delineada, para cada região, a sequência da mortalidade esperada para cada semana, se não houvesse variação aleatória na medida, e tampouco houvesse surtos de gripe causando aumento brusco da mortalidade, como este, que pode ser facilmente identificado na Região Sul, em meados de 1996. Também foi delineada a curva que projeta uma previsão para o limiar epidêmico, o qual, quando ultrapassado, configuraria os surtos de gripe. Os cálculos da mortalidade esperada e do limiar epidêmico previsto demandam o uso de técnicas avançadas de análise estatística, as quais não serão objeto deste módulo.

Exercitando o olhar sobre as duas séries temporais da Figura 6.5, o que pode ser depreendido de relevante para o conhecimento em Saúde?

Ambas as séries apresentam tendência estacionária com variação sazonal. Notando a diferença de escala no eixo vertical, percebe-se que a magnitude da mortalidade é mais elevada na Região Sul que no Nordeste. Também a amplitude de variação sazonal é mais elevada na Região Sul. Estas duas diferenças podem estar relacionadas à maior amplitude de variação climática na Região Sul que no Nordeste.

Uma análise mais detalhada indica que o período de máxima mortalidade anual estimada para a Região Sul variou entre a 28^a e a 31^a semanas epidemiológicas, respectivamente a 2^a semana de julho e a 4^a semana de julho, o que corresponde ao inverno. A hipótese subjacente é que as pessoas passam mais tempo em ambientes fechados nos dias mais frios, propiciando a transmissão da gripe. Na Região Nordeste, o período de máxima mortalidade esperada ocorreu entre a 17^a e a 22^a semanas epidemiológicas, respectivamente a 4^a semana de abril e a 4^a semana de maio. Como há menos variação de temperatura entre as estações no Nordeste, o período de maior permanência em ambientes fechados (e, portanto, mais propício à transmissão da gripe) ocorre no outono, quando chove mais. Estes dados são indicativos de diferenças no perfil epidemiológico da doença em cada região.

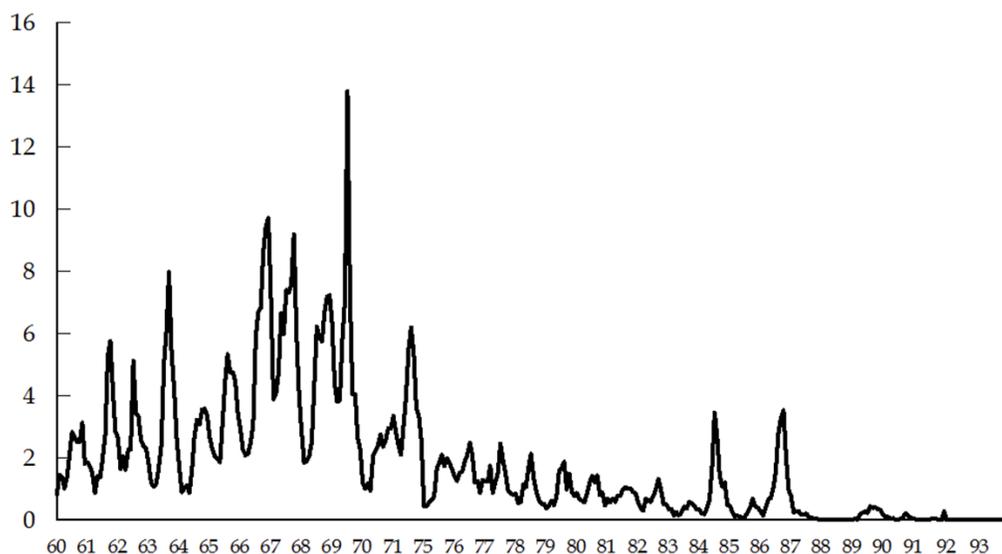
Na Figura 6.4, a observação de variação sazonal foi feita por meio de dados diários de temperatura e de mortalidade. Nas séries da Figura 6.5, a variação sazonal foi apreendida em dados discriminados por semana. Utilizar a escala mensal é outra estratégia para a percepção visual de sazonalidade nas séries temporais de interesse para a saúde. Do ponto de vista etimológico, a palavra sazonal é relativa às estações do ano. Nesse sentido, quando o ritmo ou o ciclo de repetição da série temporal se prolonga por mais de um ano, não se fala em variação sazonal, mas sim em variação cíclica.

A Figura 6.6 exemplifica uma série temporal com indicação de variação cíclica. Medidas epidemiológicas do sarampo são emblemáticas para a percepção de variação cíclica, pois a ocorrência de um surto tende a reduzir o risco de novos surtos em curto prazo. Os sobreviventes da doença adquirem imunidade e, com menos crianças susceptíveis, a transmissão torna-se menos provável. Isto é, há imunidade individual e imunidade de rebanho interagindo para a redução do risco da doença. Contudo, nos anos seguintes, as crianças imunes vão crescendo e saindo da faixa etária de maior risco para a doença, enquanto novas crianças vão nascendo e aumentando o número de suscetíveis. Esse processo, cíclico, pode facilitar a eclosão de um novo surto. A Figura 6.6 indica variação cíclica na incidência de sarampo na Cidade de São Paulo. Mesmo nos períodos em que não houve surtos epidêmicos com grande aumento do número de casos da doença, como entre 1975 e 1983, houve aumento cíclico da incidência mensal. A importância epidemiológica desta variação cíclica pode ser mais bem apreciada,

quando se recorda que no início dos anos 1980, o sarampo foi a segunda principal causa de morte (seguida da pneumonia) no grupo etário de 1 a 5 anos⁶.

Além da variação cíclica, a Figura 6.6 também apresenta acentuada tendência crescente na década de 1960, com posterior tendência decrescente. O declínio do sarampo na cidade foi associado às melhorias gerais das condições de vida e à introdução e ao aumento de cobertura da vacinação nesse período⁷.

Figura 6.6: Incidência mensal de sarampo na Cidade de São Paulo, 1960–1993.



Fonte: Waldman & Rosa, 1998⁷.

1.4 Ruído e variação aleatória

Em janeiro de 2012, o regente da Orquestra Filarmônica de Nova York interrompeu a execução da Nona Sinfonia de Gustav Mahler quando um telefone celular começou a tocar na platéia. Os músicos ficaram aturdidos: a desatenção momentânea interrompeu sua percepção de ritmo, distraiu-lhes do curso da melodia e impossibilitou a manutenção do sincronismo necessário para a harmonia.

Seguindo o paralelo entre música e análise de séries temporais, qual seria o elemento de variação das medidas epidemiológicas organizadas no tempo que propiciaria tamanha perturbação na percepção de tendências, associações e variações sazonal e cíclica? O ruído na análise de séries temporais é causado pela variação aleatória da medida, a qual se manifesta visualmente na forma de rugosidade nas linhas dos gráficos de séries temporais. Ao rever as figuras

anteriores, pode-se notar que a variação aleatória afeta todas as séries temporais delineadas, e coexiste com os demais movimentos que foram descritos neste texto.

A variação aleatória na análise de séries temporais é definida como flutuações irregulares e erráticas que não são importantes em si mesmas, e são causadas por fatores que ocorrem ao acaso e não podem ser antecipados, detectados, identificados ou eliminados⁸. Curiosamente, o termo “ruído” também é empregado para referenciar a variação aleatória na análise de séries temporais. A comparação visual das figuras anteriores propicia o reconhecimento de que há mais variação aleatória na Figura 6.4 que na Figura 6.3; isto é, a Figura 6.4 tem mais rugosidade que a Figura 6.3. Isto é fácil de compreender, uma medida tomada a cada dia (Figura 6.4) é de certo mais susceptível a variações aleatórias que uma medida tomada a cada ano (Figura 6.3).

Não é raro que os surtos epidêmicos ocorram de forma brusca e inesperada. Na análise de uma série temporal, o que diferencia um surto epidêmico da variação aleatória? Reveja a definição de variação aleatória no início do parágrafo anterior; surtos epidêmicos não se enquadram nessa definição. Mesmo quando escapam ao controle, eles não ocorrem ao acaso e o esforço de analisar as séries temporais vem justamente no sentido de antecipá-los e preveni-los. Por sua magnitude, os surtos epidêmicos não podem ser reduzidos a variações irregulares, erráticas e desimportantes.

Tendência; sazonalidade e variação cíclica; estudos de associação e variação aleatória. Você conhece agora os principais movimentos da análise de séries temporais. A definição de todos esses movimentos foi fortemente apoiada na disposição gráfica das séries temporais. O recurso gráfico é o primeiro passo para a compreensão dos processos subjacentes às medidas sequenciais ordenadas temporalmente. Esse primeiro passo é importantíssimo e não deve ser subestimado. Faça sempre o gráfico da série temporal a ser estudada, reflita sobre ele, tendo em mente os conceitos desenvolvidos nesta primeira seção. Sabendo o que deve ser procurado nas séries temporais, podemos partir para a Unidade 2, na qual são apresentados e exercitados os métodos práticos para esta análise.

Unidade 2

Séries temporais: Aspectos Metodológicos da Organização das Medidas de Doença no Tempo

Nesta unidade, serão apresentados recursos metodológicos para a análise epidemiológica das séries temporais. O primeiro item aborda a estimação de tendências, e busca operacionalizar procedimentos para avaliar se a série é crescente, decrescente ou estacionária, além de quantificar o crescimento (ou declínio) médio anual da série temporal. O segundo item focaliza a variação sazonal das séries temporais, e apresenta uma metodologia para avaliar se esta variação é estatisticamente significativa ou pode ser atribuída ao acaso. Em ambos os itens, emprega-se o estudo de associação entre as séries temporais, por meio da análise de regressão linear. Por fim, o terceiro item desta seção sintetiza um recurso técnico para explorar analiticamente a disposição visual das séries temporais. Complementando a apresentação dos métodos de análise das séries temporais, esta seção contém quatro atividades práticas com exercícios de aplicação.

1. Estimar tendências

Nunca é demais insistir que, para qualquer finalidade, na análise de séries temporais, o primeiro passo deve ser sempre a disposição gráfica da sequência de valores. Isso ajuda a organizar o conhecimento prévio que se tem sobre o assunto que está sendo estudado, e contribui para delinear os procedimentos analíticos que deverão ser realizados. Quando se trata de estimar a tendência, o simples esforço de apreciar visualmente a série temporal de interesse contribui muito para orientar a análise.

Vimos que as séries temporais podem ter tendência crescente, decrescente ou estacionária. Podem também apresentar tendências diferentes em trechos sequenciais. Para dar conta desta complexidade, uma alternativa consiste em utilizar funções matemáticas que melhor se ajustem aos pontos observados, seja para a série temporal como um todo, seja para o segmento que estiver em foco.

Para explorar a disposição gráfica e a inspeção visual das séries temporais e para introduzir os recursos de análise que permitem determinar quantitativamente suas tendências, consideremos um exemplo de aplicação, relativo à evolução temporal da aids em alguns países europeus.

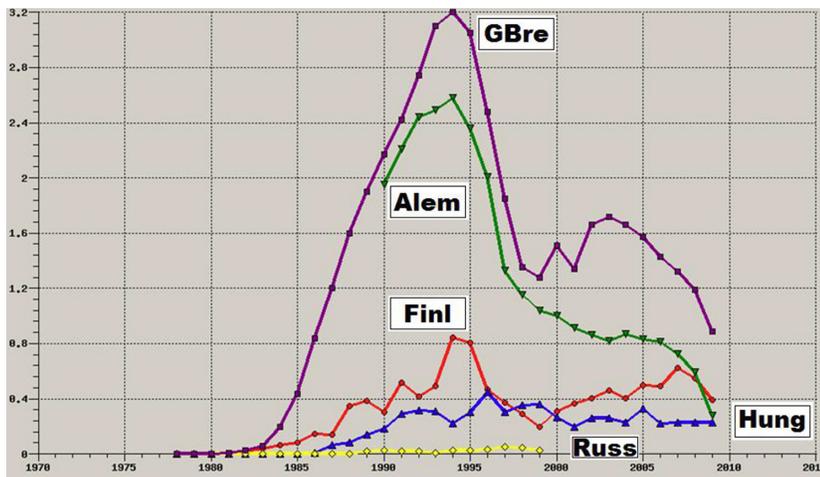
É conhecida a associação entre aids e condição socioeconômica⁹. De modo geral, estratos sociais mais submetidos à privação material tendem a ser mais reticentes às medidas de prevenção e proteção individual, o que os submete a risco mais elevado para contrair a doença. Pacientes menos escolarizados e/ou mais pauperizados tendem a ter mais dificuldade para manter a adesão ao protocolo de medicação antirretroviral, o que os submete a risco mais elevado de óbito. Com base nestas indicações, poderíamos nos perguntar sobre a evolução temporal dos indicadores epidemiológicos da aids em alguns países europeus com diferentes perfis socioeconômicos. Como a doença progrediu em países ricos e com serviços de Saúde relativamente bem organizados (como Inglaterra, Alemanha e Finlândia)? Como a doença progrediu em país mais pobre, como a Hungria? E em país emergente, como a Rússia?

A Figura 6.7 mostra as séries temporais da incidência de aids nesses países, no período para o qual havia dados disponíveis na European Health for All Database (HFA-DB), que é a base de dados sobre saúde mantida pela Organização Mundial da Saúde, em seu escritório regional para a Europa.

Nota-se marcante expansão da epidemia nos países mais ricos até meados da década de 1990. Essa expansão foi maior no Reino Unido (Inglaterra, Escócia, Irlanda do Norte e País de Gales), foi média na Alemanha e menos intensa na Finlândia. Nos países mais pobres, como Rússia e Hungria, não se chegou a notar expansão da epidemia. O que justifica o sensível crescimento da incidência no Reino Unido, na Alemanha e na Finlândia até meados dos anos 90? Por que a epidemia retrocedeu subsequentemente? Estes países teriam adotado programas efetivos de prevenção? A introdução da terapêutica antirretroviral em meados da década de 1990 teria exercido algum efeito sobre a redução de incidência? Seriam confiáveis os dados de incidência da aids na Rússia e Hungria? Uma busca no PubMed por artigos científicos que tenham se aplicado a essa temática poderia auxiliar a desdobrar essas questões.

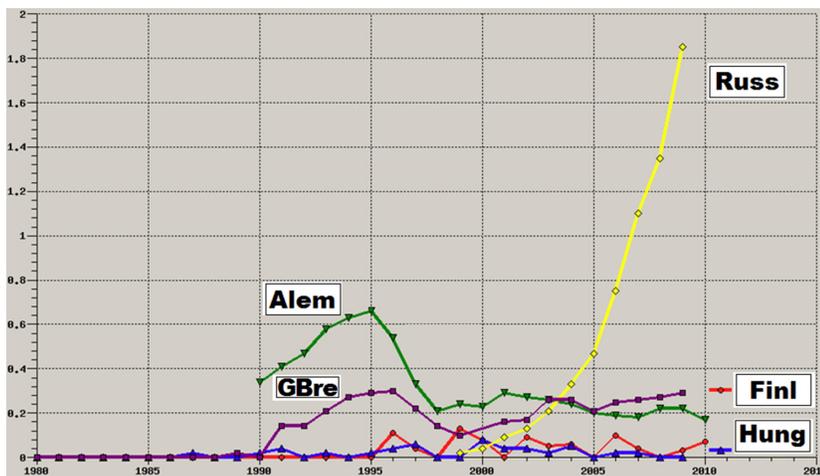
Os indicadores de mortalidade são consistentemente mais elevados para homens (Figura 6.9) que para mulheres (Figura 6.8) nos países monitorados. A Alemanha, que havia experimentado uma evolução de incidência menos elevada que o Reino Unido, manteve uma mortalidade mais elevada durante todo o período de monitoramento. Não obstante sua conhecida diferença socioeconômica, Finlândia e Hungria mantiveram um padrão análogo para as séries temporais de mortalidade por aids, tanto para homens como para mulheres. E a Rússia, surpreendentemente, apresentou enorme expansão da mortalidade por aids, tanto para homens como para mulheres, na década de 2000, período em que já era disponível a terapêutica antirretroviral, cujo efeito de redução dos indicadores populacionais de mortalidade por aids é conhecido¹⁰. Que hipóteses poderiam ser aventadas para explicar esses resultados?

Figura 6.7: Incidência de aids em alguns países europeus.



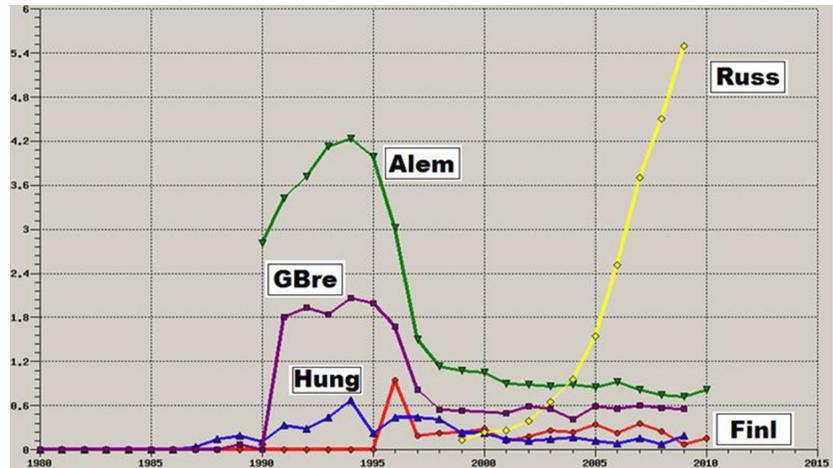
Fonte: European health for all database. <<http://data.euro.who.int/hfad/>>.

Figura 6.8: Mortalidade por aids em mulheres, em alguns países europeus.



Fonte: European mortality database. <<http://data.euro.who.int/hfamdb/>>.

Figura 6.9: Mortalidade por aids em homens, em alguns países europeus.



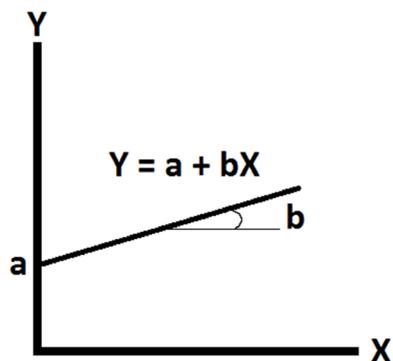
Fonte: European mortality database. <<http://data.euro.who.int/hfamdb/>>.

Na Figura 6.9, observa-se uma importante redução da mortalidade de homens por aids na Alemanha e no Reino Unido, a partir de meados dos anos 90. Independente da magnitude da mortalidade ter sido mais elevada na Alemanha, a inspeção visual das séries parece indicar que o declínio foi equivalente nos dois países. Será que isso de fato ocorreu?

A ideia de quantificar a tendência tem justamente o intuito de permitir a comparação entre diferentes séries temporais. O procedimento de estimação quantitativa da tendência demanda aplicação estatística, mas o esforço é recompensado pelo campo de possibilidades que se abre para a interpretação dos dados epidemiológicos. A metodologia descrita em seguida foi originalmente publicada por Antunes e Waldman⁶.

Considerando o traçado da reta de melhor ajuste entre os pontos da série temporal, ou de um de seus trechos, para o qual se pretende estimar a tendência, poderíamos escrever a equação da reta indicada na Figura 6.10, da seguinte forma:

Figura 6.10: Representação gráfica da reta.



$$Y = a + bX$$

Considere Y como sendo a medida dos valores da série temporal, e X a medida dos anos, ou de outra unidade temporal na qual os valores de Y tenham sido monitorados. O valor de “a” corresponde à interseção entre a reta e o eixo vertical; de fato, quando $X = 0$, temos $Y = a$. O valor de “b” corresponde à inclinação da reta; para cada mudança de uma unidade na escala de X, o valor de Y é acrescido de “b” unidades. Naturalmente, se “b” for positivo, a reta é crescente, e valores mais elevados de “b” indicam retas mais inclinadas no sentido vertical. De modo complementar, se “b” for negativo, a reta é decrescente, e se “b” for igual a zero, a reta é paralela ao eixo X e qualquer variação de X não modifica os valores de Y, configurando a condição para a tendência estacionária.

Para mensurar a taxa de variação da reta que ajusta os pontos da série temporal, é preciso focalizar o valor de “b”, cuja estimação é feita por meio da análise de regressão linear. Para esse fim, vamos trabalhar com a transformação logarítmica dos valores de Y, o que propicia vantagens de ordem estatística para a aplicação da análise de regressão linear, como a redução da heterogeneidade de variância dos resíduos da análise de regressão; isto é, dos valores da diferença entre os pontos da reta média e os pontos da série temporal. Além disso, essa transformação favorece o cômputo da razão de incremento anual da série temporal, por meio do procedimento descrito em seguida.

Consideremos $X_1, X_2, X_3, \dots, X_i, \dots, X_n$ como sendo os valores dos anos para os quais foram tomadas as medidas $Y_1, Y_2, Y_3, \dots, Y_i, \dots, Y_n$. Então, para qualquer ano “X(i)” e seu ano subsequente “X(i+1)”, podemos escrever que:

$$\begin{aligned} \log Y(i) &= a + bX(i); \text{ e} \\ \log Y(i+1) &= a + bX(i+1) \end{aligned}$$

Fazendo a diferença de ambos os termos das duas equações, teremos:

$$\log Y(i+1) - \log Y(i) = a + bX(i+1) - a - bX(i) = b[X(i+1) - X(i)]$$

Como $X(i+1)$ e $X(i)$ são anos subsequentes, sua diferença é sempre igual a um. E, usando propriedades da álgebra de logaritmos, teremos:

$$\log Y(i+1) - \log Y(i) = \log[Y(i+1)/Y(i)] = b; \text{ ou } Y(i+1)/Y(i) = 10^b$$

Nesta notação, o acento circunflexo corresponde à operação exponencial e deve ser lido como “elevado a”. Então, subtraindo 1 de ambos os lados da equação, teremos:

$$Y(i+1)/Y(i) - 1 = -1 + 10^b; \text{ ou:}$$
$$[Y(i+1) - Y(i)]/Y(i) = -1 + 10^b$$

Mas $[Y(i+1) - Y(i)]/Y(i)$ é justamente a taxa de crescimento médio anual, pois foi dimensionada para qualquer ano genérico “i”. Basta, então, estimar o valor de “b” para obtermos a taxa de crescimento médio anual. Observe que esta taxa pode ser apresentada como proporção ou como porcentagem. Se a taxa for positiva, a série temporal é crescente, se for negativa é decrescente, e será estacionária se não houver diferença significativa entre seu valor e o zero. Como o valor de “b” é calculado por meio de regressão linear, é preciso empregar o erro-padrão (EP) da estimativa de “b” para construir o intervalo de confiança (95%) desta medida. Com isso, teremos uma expressão sintética para a estimação quantitativa da tendência:

Fórmula 1: Taxa de incremento anual e intervalo de confiança (95%)
Taxa de incremento anual = $-1 + 10^b$; e IC (95%) = $-1 + 10^{(b \pm t \cdot EP)}$

Nesta fórmula, os valores de b e EP são fornecidos por análise de regressão linear, e o valor de t é fornecido pela tabela da distribuição *t de Student*, a qual pode ser encontrada em http://en.wikipedia.org/wiki/Student's_t-distribution. Procure o valor na coluna correspondente ao teste bicaudal com 95% de nível de significância e na linha correspondente ao número de graus de liberdade da análise de regressão linear (número de anos na série menos 1).

Uma última, porém importante indicação técnica estatística para este cálculo. A análise de regressão linear simples não se presta à análise de séries temporais, em função da autocorrelação serial usualmente encontrada em medidas de dados populacionais. Nesse sentido, é preciso recorrer a procedimentos específicos de análise de regressão linear, que foram especialmente delineados para essa finalidade. Os mais utilizados são as técnicas de Prais-Winsten e Cochrane-Orcutt, ambas facilmente executáveis em programas de informática para análise estatística que são empregados na área de Saúde, como o Stata e o SPSS. Sendo dois pro-

cedimentos diferentes para a mesma finalidade, seus resultados não serão exatamente coincidentes. Apesar de alguma diferença em casas decimais dos resultados; suas interpretações (tendência crescente, decrescente ou estacionária) dificilmente diferirão.

Nas séries temporais da mortalidade por aids (figuras 6.8 e 6.9), foi verificado um impressionante crescimento na Rússia, tanto para homens como para mulheres. Independente de a mortalidade ser maior em homens que em mulheres, para qual desses grupos o crescimento terá sido mais intenso? Usando a metodologia descrita anteriormente, encontra-se taxa de crescimento anual de 3,96% (Intervalo de Confiança 95%: 3,59% a 4,33%) para homens e 4,59% (IC95%: 3,77% a 5,43%) para mulheres. Esses valores são indicativos de tendência crescente para ambas as séries. Como há grande superposição entre os intervalos de confiança, reforça-se a hipótese de não ser significativa a diferença entre as duas taxas de crescimento anual.

O que poderia justificar tamanho incremento da mortalidade por aids na Rússia, em um período em que houve declínio em vários países? Para testar a hipótese de que esse aumento reflete o aumento da capacidade de reconhecer a aids como causa básica de morte, poderíamos traçar as séries temporais da proporção de óbitos por causa básica não determinada ou mal especificada. Se a magnitude desta proporção for elevada, reforça-se a hipótese de que é ruim a qualidade do registro de informações de mortalidade naquele país. Se a tendência da série temporal da proporção de óbitos sem identificação de causa for decrescente nos últimos anos, reforça-se a hipótese de que o aumento recente da mortalidade por aids reflete, ao menos em parte, o aumento de capacidade do sistema de informações em reconhecer os óbitos causados por ela.

Quando a mesma análise é aplicada para a mortalidade de homens por aids na Alemanha, a partir de 1994, encontra-se taxa de crescimento anual de -9,70% (IC95%: -15,80% a -3,16%). O sinal negativo indica que a tendência é decrescente, e a magnitude desses números indica a intensidade do declínio médio anual desta medida. Como o intervalo de confiança não inclui o valor de zero, a hipótese de tendência estacionária é excluída. Por fim, quando se analisa a mortalidade por aids em homens na Hungria, encontra-se taxa de crescimento anual de -0,15% (IC95%: -6,48% a +6,62%). Esse resultado é muito pequeno, próximo de zero, e o intervalo de confiança vai de valores negativos a valores positivos. Essas indicações são compatíveis com a hipótese já aventada na Atividade 1, durante a inspeção visual da série temporal, de que a tendência era estacionária e que a epidemia não havia progredido de modo agudo na Hungria. Naturalmente, também esse resultado é sujeito a restrições de validade relativas à qualidade do registro das causas de óbito naquele país.

2. Estimar sazonalidade

Vimos que as séries temporais podem ter variação sazonal. O termo refere-se à sazão, palavra pouco usual atualmente, que é sinônimo de estação. Com isso, quando se identifica repetições em períodos mais extensos (dois anos ou mais), isto é, que não são relativas às estações, fala-se que o fenômeno tem variação cíclica, em vez de sazonal.

As variações sazonais podem ser aferidas por meio de medidas tomadas mês a mês, semana a semana ou dia a dia. Vamos estudar uma metodologia prática para determinar se há ou não há variação sazonal estatisticamente significativa. O procedimento descrito a seguir consiste na forma mais simples de decomposição das séries temporais^{8,11,12}. Em outras palavras, vamos tentar isolar o componente sazonal e verificar se ele atende à hipótese de significância estatística.

Uma primeira consideração diz respeito à forma de registro do tempo nas avaliações de sazonalidade. É preciso reconhecer que há alguma irregularidade nesse registro. Se as medidas foram tomadas mês a mês, são utilizados os meses do calendário, enumerando-os sequencialmente (1, 2, 3,..., mês genérico “i”,..., mês final “n”). Mas sabemos que os meses não são exatamente iguais entre si. Alguns têm 30 dias, outros têm 31 e fevereiro tem 28 ou 29 dias, dependendo do ano. Apesar disso, é fácil representar os dados segundo o mês de sua ocorrência no calendário, mesmo reconhecendo que os meses não são regulares no número de dias.

Uma dificuldade análoga é encontrada quando se trabalha com dados discriminados por semana. O ano não pode ser dividido de modo exato por semanas de sete dias. Para lidar com esta dificuldade, considera-se que cada ano é dividido em 52 semanas epidemiológicas, o que obriga uma ou duas semanas a ter oito dias, dependendo se o ano é bissexto ou não. Quando os dados forem discriminados dia a dia, a solução encontrada para manter a regularidade de 365 dias ao ano consiste em considerar que o registro relativo ao dia 29 de fevereiro se soma ou é considerado em média com o dia antecedente.

Estas soluções não são ideais, mas são aproximações plausíveis, factíveis e necessárias. Trabalharemos com meses que diferem entre si quanto ao número de dias. Teremos semanas cuja extensão de sete dias terá pelo menos uma exceção a cada ano. Se os dados forem diários, nos anos bissextos, o dia 29 de fevereiro deverá ser aritmeticamente integrado ao dia anterior para não romper a frequência de 365 dias por ano.

Para avaliar se a variação existente na série temporal tem ou não tem alguma regularidade sazonal, será preciso decompor a série temporal em dois componentes, utilizando a equação de regressão linear aos moldes do que foi apresentado na Figura 6.10. Para atender ao requisito de modelar a variação sazonal, a equação de regressão é escrita como indicado na fórmula 2:

Fórmula 2: Equação de regressão linear com componente sazonal

$$Y(i) = b_0 + b_1 \cdot X(i) + b_2 \cdot \text{sen}[2\pi X(i)/L] + b_3 \cdot \text{cos}[2\pi X(i)/L]$$

Nesta equação, $Y(i)$ é a medida da série temporal para cada medida diária, semanal ou mensal correspondente ao momento genérico “ i ” e $X(i)$ é a numeração sequencial dos momentos de tomada da medida (dia, semana, mês). π é a conhecida constante 3,141592654... e L é uma constante relativa à forma da medida: 12 se a medida for mensal, 52 para medidas semanais e 365 para medidas diárias. Quanto aos coeficientes de regressão, b_0 é a constante ou intercepto, b_1 é o estimador de tendência da série temporal e b_2 e b_3 são os coeficientes que modelam a variação sazonal. Esta equação decompõe a série temporal em dois componentes, um para a tendência e outro para a sazonalidade.

Uma vez calculada a série de valores do termo sazonal, pode-se executar a análise de regressão linear para estimar os coeficientes de regressão b_0 , b_1 , b_2 e b_3 e seus respectivos erros-padrão. De modo análogo ao que fora informado para a análise de estimação da tendência, reitera-se a indicação para não utilizar regressão linear simples para análise de séries temporais, pois um dos requisitos para esta modalidade de análise é que os resíduos da equação de regressão sejam aleatórios, o que não ocorre com medidas populacionais organizadas em séries temporais, as quais frequentemente apresentam autocorrelação serial elevada. A autocorrelação serial tem como efeito superestimar os indicadores de qualidade de ajuste da equação de regressão, o que, em muitos casos, implica inversão das conclusões da análise estatística, fazendo com que variações pouco expressivas tendam a ser indevidamente indicadas como significantes. Nesse sentido, devem-se utilizar os procedimentos de Prais-Winsten ou de Cochrane-Orcutt para a análise de regressão linear em séries temporais.

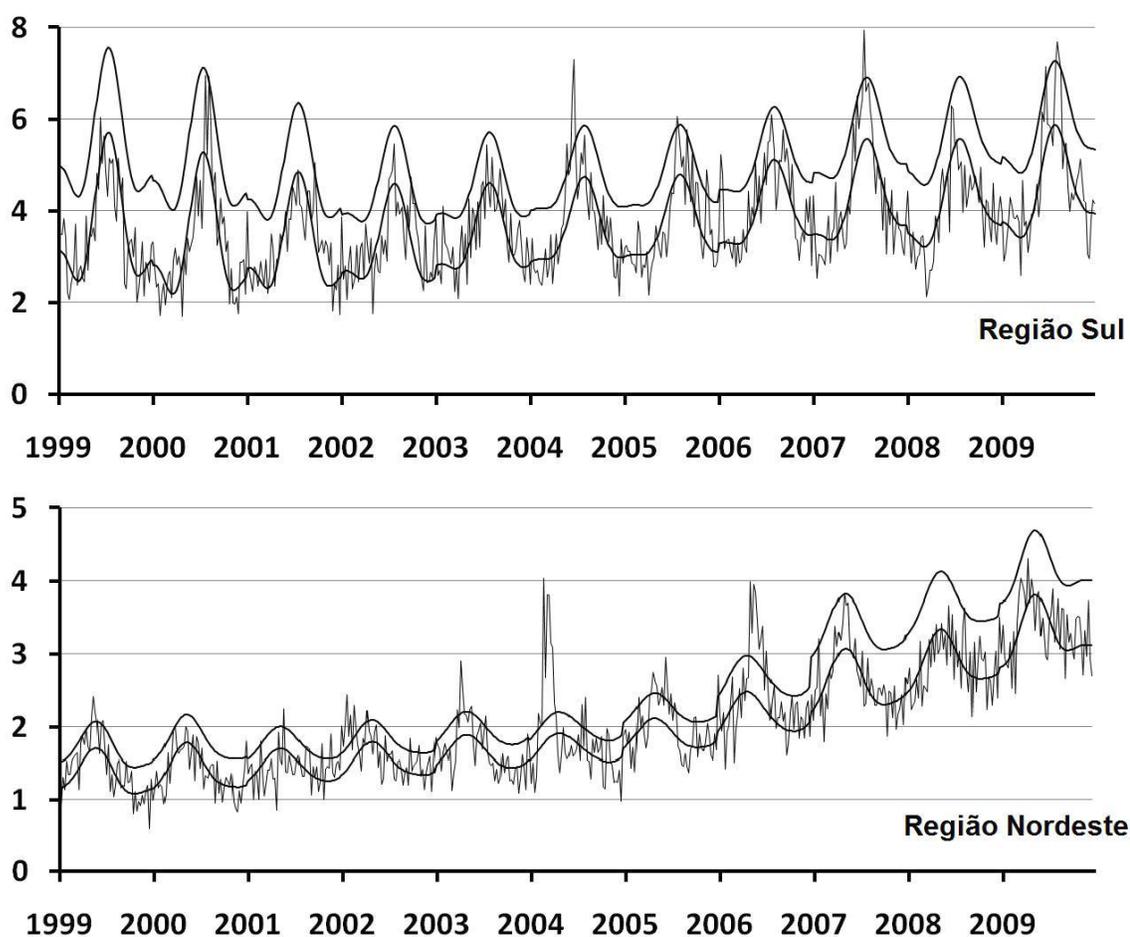
Na equação de regressão, se os coeficientes b_2 e/ou b_3 , relativos ao seno e cosseno do termo sazonal, forem estatisticamente diferente de zero ($p < 0,05$), conclui-se que há variação sazonal significativa. De modo complementar, dedui-se que a variação sazonal observada em uma série temporal pode ser atribuída ao acaso, se estes coeficientes não forem significantes, isto é, não diferirem estatisticamente de zero no teste de hipótese ($p > 0,05$).

Para exercitar a metodologia de detecção de variação sazonal nas séries temporais, vamos considerar um problema prático de pesquisa. A Figura 6.11 mostra a série temporal da

mortalidade de idosos (65 anos ou mais) por gripe e pneumonia nas regiões Nordeste e Sul do Brasil, do início de 1999 ao fim de 2009. Para favorecer a comparação visual, as séries temporais da mortalidade observada foram acompanhadas pelas séries da mortalidade esperada e pelo limiar epidêmico em cada região, de modo análogo à Figura 6.5.

Sabe-se que a mortalidade por gripe e pneumonia é maior nos períodos mais frios do ano¹³. Mas há pouca variação climática entre o inverno e o verão no Nordeste, o que justifica avaliar a hipótese de que não seja estatisticamente significativa a variação sazonal observada na mortalidade de idosos por gripe e pneumonia nesta região.

Figura 6.11: Séries temporais da mortalidade semanal de idosos (65 anos ou mais) por *influenza* e pneumonia nas regiões Sul e Nordeste, 1999–2009. Mortalidade observada, mortalidade esperada e limiar epidêmico previsto.



Fonte: Oliveira, 2012⁵.

Uma primeira observação diz respeito ao fato de que a série tem tendência estacionária na Região Sul e crescente na Região Nordeste. No estudo do qual esta imagem foi reproduzida, esta diferença foi interpretada como sendo devida ao progressivo aumento da capacidade do Sistema de Informações sobre Mortalidade em identificar as causas básicas de óbito em idosos residentes no Nordeste; enquanto, na Região Sul, esta capacidade já era elevada desde o início do monitoramento.

A magnitude dos coeficientes de mortalidade é menor na Região Nordeste que na Sul. Observe, inclusive, que a escala vertical utilizou valores menos elevados para a Região Nordeste. Também a amplitude de variação entre as estações, a cada ano, é menor na Região Nordeste que na Região Sul. Estas observações podem induzir a hipótese de que a variação sazonal da mortalidade de idosos por gripe e pneumonia na Região Nordeste não seja estatisticamente significativa. Vamos empregar a metodologia sintetizada na fórmula 2 para testar esta hipótese.

Com base na aplicação dos valores observados da série temporal da mortalidade de idosos por gripe e pneumonia na Região Nordeste, a fórmula 2 permite encontrar os seguintes coeficientes de regressão: $b_2 = 0,229$ e $b_3 = -0,221$. Os erros-padrão encontrados foram, respectivamente 0,125 e 0,123. Estas medidas correspondem a valor de $p = 0,069$ para o coeficiente relativo ao seno (b_2) e $p = 0,074$ para o coeficiente relativo ao cosseno (b_3). Com base nesses dados, concluímos que não houve variação sazonal significativa do ponto de vista estatístico ($p > 0,05$) na mortalidade de idosos por gripe e pneumonia na Região Nordeste.

Quando o mesmo procedimento é aplicado à Região Sul, os valores encontrados apontam a conclusão inversa. Os coeficientes de regressão são: $b_2 = -0,553$ (erro-padrão 0,139) e $b_3 = -0,846$ (erro-padrão 0,111). Para ambos os coeficientes, obtém-se $p < 0,001$. Em outras palavras, a variação sazonal na medida de interesse, para a Região Sul, não pode ser considerada como tendo ocorrido ao acaso, e a variação sazonal na mortalidade de idosos por gripe e pneumonia na Região Sul é considerada estatisticamente significativa.

3. Alisamento das séries temporais

Em todo este módulo, procurou-se mostrar a complexidade da análise de variação das séries temporais com as várias dimensões que são relevantes e devem ser consideradas na apreciação visual dos gráficos: tendências, variações cíclicas, sazonalidade e variações aleatórias. Salientou-se, também, que a variação aleatória nas séries temporais, como o ruído na música, pode prejudicar a percepção dos componentes mais importantes da variação nas séries temporais.

Na conclusão da primeira seção deste módulo e no início do item sobre estimação de tendências, foi sublinhada a importância da disposição gráfica das séries temporais. Neste item, será apresentado um recurso metodológico que tem como objetivos instrumentar a análise exploratória das séries temporais, favorecer sua visualização gráfica e evidenciar seus componentes de variação.

Na primeira seção, foi enfatizado que a variação aleatória das séries temporais consiste em flutuações irregulares e erráticas que não são importantes, e são causadas por fatores que ocorrem ao acaso e não podem ser antecipados, detectados, identificados ou eliminados. No entanto, assim como a manifestação de ruído prejudica a percepção da música e de seus componentes (melodia, harmonia e ritmo), a variação aleatória pode dificultar a análise da série temporal e a detecção de sua tendência, seus fatores associados e variação sazonal. Nesse sentido, pode ser importante suprimir, ao menos em parte, a variação aleatória nos gráficos.

O recurso que permite reduzir a variação aleatória é o alisamento das séries temporais. Este termo deriva, por contraste, da percepção de que a variação aleatória representa certa rugosidade no delineamento gráfico da série temporal. Isto é, quando a “rugosidade” causada pela variação aleatória é reduzida, tem-se como resultado uma série temporal “alisada”.

Fórmula 3: **Alisamento de séries temporais por médias móveis simples**

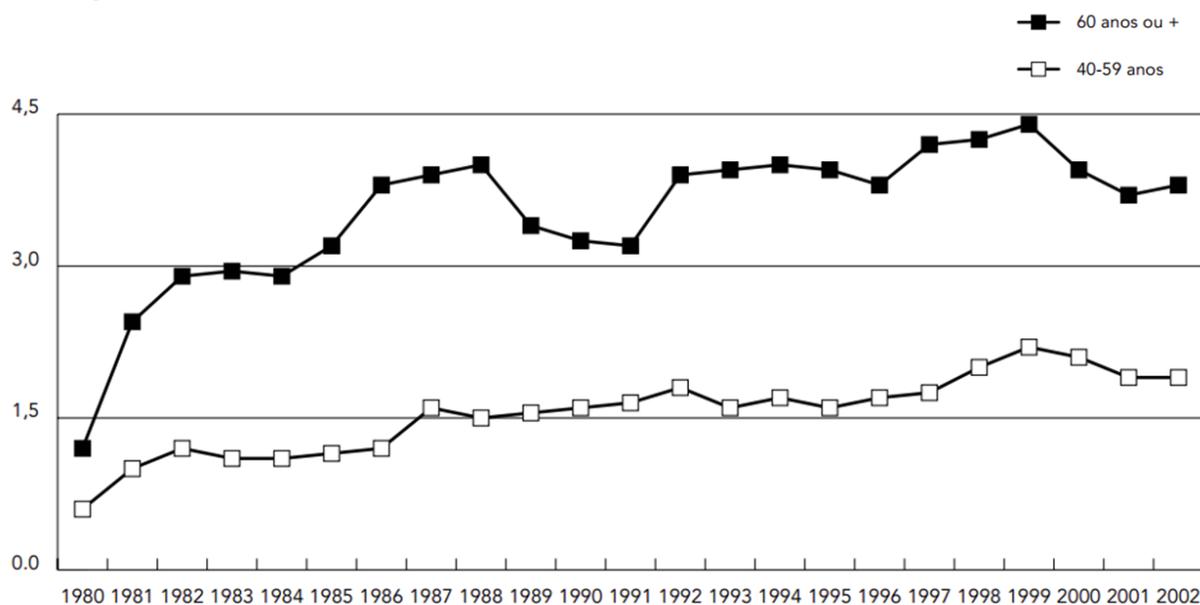
Sendo $Y(1), Y(2), Y(3), \dots, Y(i), \dots, Y(n)$ os valores da série original,
 $Y'(1), Y'(2), Y'(3), \dots, Y'(i), \dots, Y'(n)$ os valores da série modificada,
e $1, 2, 3, \dots, i, \dots, n$, os períodos de referência das medidas.

Médias móveis de ordem 2: $Y'(i) = [Y(i) + Y(i-1)]/2$

Médias móveis de ordem 3: $Y'(i) = [Y(i) + Y(i-1) + Y(i-2)]/3$

Médias móveis de ordem k: $Y'(i) = \{Y[i] + Y[i-1] + \dots + Y'[i-(k-1)]\}/k$

Figura 6.12: Médias móveis (ordem 3) da mortalidade (por 100 mil habitantes) por câncer de orofaringe na cidade de São Paulo, 1980-2002.

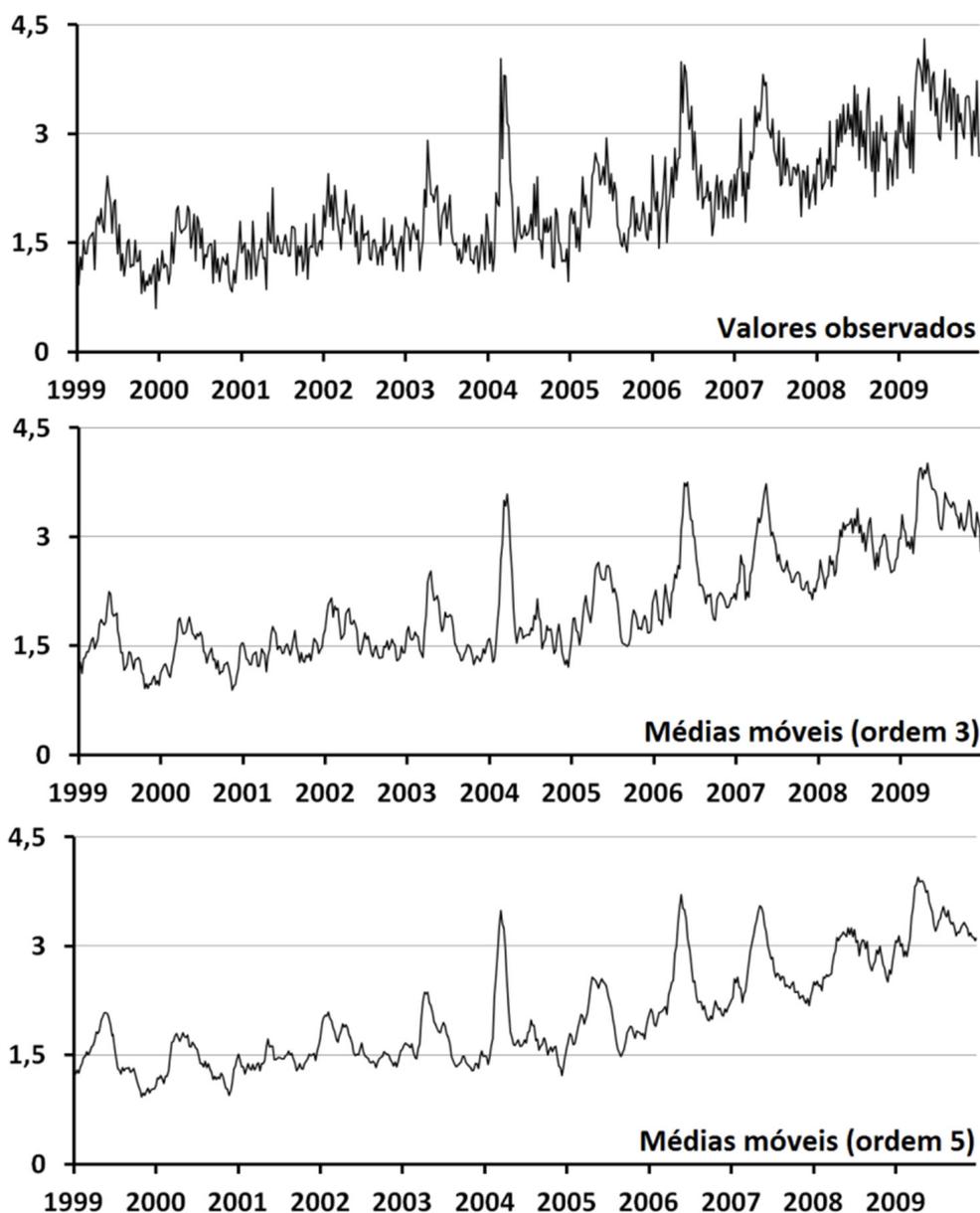


Fonte: Biazevic et al., 2006¹⁴.

Outro exemplo de aplicação pode ser fornecido para o gráfico da mortalidade de idosos por gripe e pneumonia na Região Nordeste. A apresentação de dados semanais resultava uma série com rugosidade, isto é, com acentuada variação aleatória. A Figura 6.13 reproduz a série original dos valores observados e as séries modificadas com os valores obtidos para o alisamento de médias móveis de ordem 3 e 5. Observe como, de fato, houve alisamento progressivo, sem prejuízo da percepção dos demais componentes da série temporal: tendência estacionária na primeira metade e crescente na segunda, variação sazonal em toda a série.

Recapitulando o início deste módulo, as primeiras frases da introdução propunham ser a análise de séries temporais uma perspectiva de antevisão do futuro. Retomemos agora essa ideia para refletir como esta capacidade se realiza.

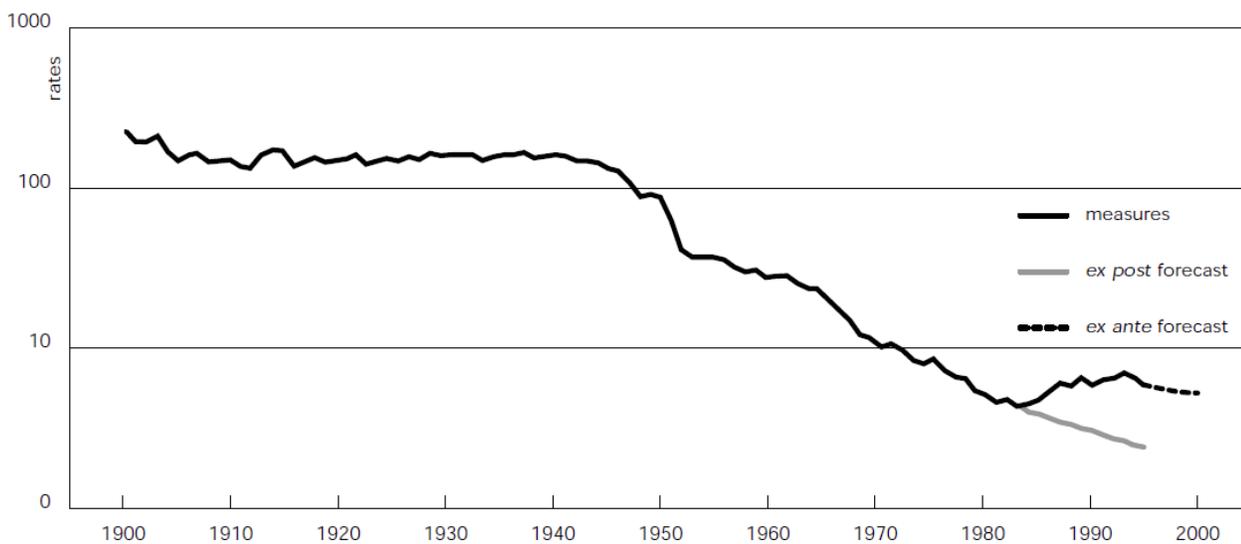
Figura 6.13: Série temporal da mortalidade semanal de idosos (65 anos ou mais) por gripe e pneumonia na Região Nordeste, 1999–2009. Alisamento por médias móveis (de ordem 3 e 5) para redução da variação aleatória.



Fonte: Oliveira, 2012⁵.

A primeira e mais óbvia forma diz respeito à previsão dos valores da série temporal em momentos futuros. Isto pode ser feito por meio da equação de regressão linear, levando em consideração a curva de melhor ajuste de seus pontos. Quando se conhece a tendência, é fácil projetar a série alguns pontos para adiante, baseados na hipótese de que a tendência se mantenha, pelo menos no futuro próximo. A previsão dos valores futuros também pode ser feita por procedimentos estatísticos mais complexos, como a metodologia Arima, cuja aplicação demanda treinamento especializado adicional. Este tipo de previsão é muito utilizado em Econometria, e tem decerto relevância também para a programação dos serviços de Saúde. Ressalta-se, contudo, que a imprecisão das estimativas sobre o futuro é tanto maior quanto mais se distancia do presente. Além disso, não é infrequente em atividades humanas que ocorra alguma intercorrência inesperada, o que pode representar fonte de erro adicional não prevista.

Figura 6.14: Série temporal da mortalidade (por 100 mil habitantes) por tuberculose na Cidade de São Paulo, 1900–1997.



Fonte: Antunes & Waldman, 1999¹⁵.

Para variáveis quantitativas, é sempre possível conhecer efetivamente os valores que foram medidos no passado, mas não os que serão medidos no futuro. Afora essa diferença óbvia, nada diferencia, tecnicamente, a previsão do futuro e a previsão do passado. A segunda forma de aplicar o instrumental de análise das séries temporais refere-se à previsão do passado. Esse tipo de previsão pode parecer estranho e desnecessário. No entanto, esta análise foi vantajosamente empregada nas séries temporais das figuras 6.5 e 6.12, as quais modelaram, com base na tendência e sazonalidade, a previsão de como seria a mortalidade esperada por gripe e pneumonia em idosos, se não houvesse a intercorrência de variação aleatória e de surtos epidêmicos. E, com base nesta previsão do passado, foi possível evidenciar os surtos de gripe manifestados nos momentos em que as medidas observadas ultrapassaram o limiar epidêmico calculado com base na mortalidade estimada, isto é, na previsão sobre o passado.

São empregadas as expressões latinas *ex ante* e *ex post* para referir, respectivamente, as modalidades de previsão do futuro e do passado. Do ponto de vista técnico, estas modalidades de previsão não são diferentes. Os mesmos instrumentos utilizados para prever valores de variáveis quantitativas no futuro servem para prever os valores do passado. Para exemplificar com um estudo epidemiológico que aplicou as duas modalidades de previsão no mesmo gráfico, a Figura 6.14 mostra a série temporal da mortalidade por tuberculose na Cidade de São Paulo, durante praticamente todo o século XX.

Até meados do século, houve tendência estacionária com magnitude elevada. Nesse período, a tuberculose foi a causa básica de cerca de 10% de todos os óbitos ocorridos a cada ano¹⁵. A partir de meados do século, a mortalidade por tuberculose declinou na cidade, até o início dos anos 80, quando subitamente inflectiu e assumiu tendência crescente, em função da associação entre a doença e aids. Além de mostrar a mudança da tendência nesses diferentes períodos, a Figura 6.14 mostra a previsão para o passado, de como a doença teria continuado a decrescer até o início dos anos 90, caso não tivesse havido a emergência da aids. A Figura 6.14 mostra também a previsão para o futuro de que a mortalidade por tuberculose voltaria a declinar lentamente a partir de meados dos anos 90, o que de fato ocorreu; de modo possivelmente associado à introdução da terapêutica antirretroviral, que modificou o perfil de mortalidade da aids na Cidade de São Paulo.

Existe, ainda, a terceira forma de se pensar na análise de séries temporais como uma modalidade de antever o futuro. Em vez de focalizar os valores que as variáveis quantitativas ordenadas temporalmente assumirão no futuro, esta terceira forma de previsão se refere ao reconhecimento dos padrões de variação da medida. Caso conheçamos esses padrões, compreenderemos o que pode interferir favorável ou desfavoravelmente para o incremento ou decréscimo da medida. Podemos dizer que antevimos o futuro, que conhecemos como determi-

nado processo se comportará, quando conseguimos caracterizar sua tendência, reconhecer sua variação sazonal e cíclica, quando dimensionamos sua variação aleatória e identificamos os fatores associados que causam impacto significativo sobre suas medidas.

Este conhecimento sobre a estrutura de variação das séries temporais não deixa de ser uma forma de antever o futuro. Se fosse possível separar estas duas alternativas e escolher entre ter uma boa estimativa do valor exato de uma variável de interesse para a saúde ou de conhecer com precisão os processos que determinam sua variação, possivelmente os profissionais de Saúde escolheriam a segunda opção.

Referências

- 1 – Porta, M. (2008) A dictionary of epidemiology. New York. Oxford University Press.
- 2 – Wiener, N. (1966) Extrapolation, interpolation and smoothing of stationary time series. Cambridge, MA. The MIT Press.
- 3 – Antunes, J. L. F. (1998) “Grow and multiply”: social development, birth rates and demographic transition in the Municipality of São Paulo, Brazil, time series for 1901-94, *Rev Bras Epidemiol*, 1(1), 61-78.
- 4 – Gouveia, N., Hajat, S. & Armstrong, B. (2003) Socioeconomic differentials in the temperature-mortality relationship in Sao Paulo, Brazil, *Int J Epidemiol*, 32(3), 390-397.
- 5 – Oliveira, J. F. M. (2012) Efetividade da vacinação contra gripe no contexto brasileiro: análise comparativa do programa nas regiões Nordeste e Sul, Faculdade de Saúde Pública São Paulo, Universidade de São Paulo, 82.
- 6 – Antunes, J. L. & Waldman, E. A. (2002) Trends and spatial distribution of deaths of children aged 12-60 months in Sao Paulo, Brazil, 1980-98, *Bull World Health Organ*, 80(5), 391-398.
- 7 – Waldman, E. A. & Rosa, T. E. C. (1998) *Vigilância em saúde pública, Série Saúde & Cidadania*. São Paulo. Instituto para o Desenvolvimento da Saúde. Faculdade de Saúde Pública da Universidade de São Paulo. Núcleo de Assistência Médico-Hospitalar. Fundação Itaú Social, 253.
- 8 – Gaynor, P. & Kirkpatrick, R. (1994) *Introduction to time series modeling and forecasting in business and economics*. New York. McGraw-Hill.
- 9 – Farias, N. & Cardoso, M. R. (2005) Mortalidade por Aids e indicadores sociais no Município de Sao Paulo, 1994 a 2002, *Rev Saude Publica*, 39(2), 198-205.
- 10 – Antunes, J. L., Waldman, E. A. & Borrell, C. (2005) Is it possible to reduce AIDS deaths without reinforcing socioeconomic inequalities in health?, *Int J Epidemiol*, 34(3), 586-592.

11 – Chow, G. C. (1983) *Econometrics*. New York. McGraw-Hill.

12 – Van Asten, L., Van den Wijngaard, C., Van Pelt, W., Van de Kasstele, J., Meijer, A., Van der Hoek, W., Kretzschmar, M. & Koopmans, M. (2012) Mortality attributable to 9 common infections: significant effect of influenza A, respiratory syncytial virus, influenza B, norovirus, and parainfluenza in elderly persons, *J Infect Dis*, 206(5), 628-639.

13 – Antunes, J. L., Waldman, E. A., Borrell, C. & Paiva, T. M. (2007) Effectiveness of influenza vaccination and its impact on health inequalities, *Int J Epidemiol*, 36(6), 1319-1326.

14 – Biazevic, M. G., Castellanos, R. A., Antunes, J. L. & Michel-Crosato, E. (2006) Tendencias de mortalidade por cancer de boca e orofaringe no Municipio de Sao Paulo, Brasil, 1980/2002, *Cad Saude Publica*, 22(10), 2105-2114.

15 – Antunes, J. L. & Waldman, E. A. (1999) Tuberculosis in the twentieth century: time-series mortality in Sao Paulo, Brazil, 1900-97, *Cad Saude Publica*, 15(3), 463-476.

Glossário

Alisamento – Recurso gráfico das séries temporais que consiste em suprimir a variação aleatória (e, portanto, a rugosidade de sua representação gráfica) para evidenciar os demais componentes de sua variação (tendência, sazonalidade e variação cíclica).

Associação – Parte do estudo das séries temporais que busca a correspondência entre seus valores e fatos ocorridos em momentos específicos. Em particular, é útil aferir a associação da série temporal com outras medidas ordenadas no tempo. A associação pode ser aferida e analisada estatisticamente, com referência concomitante ou com deslocamento no tempo para favorecer o teste de hipóteses conceituais. A verificação estatística de associações significantes, por si só, não permite sustentar inferências causais.

Prais-Winsten – Procedimento estatístico de análise de regressão linear especialmente desenvolvido para estimação de tendência e associação em séries temporais, o qual leva em consideração a relação de dependência entre valores consecutivos da série.

Previsão – Com base nos valores conhecidos da série temporal, seus valores futuros podem ser estimados levando em consideração seus diferentes componentes de variação (tendência, associação, sazonalidade, variação cíclica e variação aleatória) por meio de diferentes recursos estatísticos, como equações de regressão linear (modelo de Prais-Winsten) ou técnicas mais complexas como o procedimento Arima. Também pode ser útil calcular seus valores passados, comparando os valores estimados com aqueles que foram de fato observados. Quando a previsão se aplica a valores futuros, é denominada previsão *ex ante* (antes de acontecer). Quando se aplica a valores passados, é denominada previsão *ex post* (após acontecer).

Rugosidade – Forma usual de se referir ao componente de variação aleatória das séries temporais. A metáfora visual diz respeito ao formato das representações gráficas das séries temporais com forte componente de variação aleatória.

Ruído – Forma usual de se referir ao componente de variação aleatória das séries temporais.

Sazonalidade – Componente de variação das séries temporais, que reflete a influência de efeitos de algum modo relacionados às estações do ano. Apesar de ser esta a origem do termo, a

sazonalidade não reflete apenas aspectos geográficos e climáticos; há fatores socioeconômicos e culturais que também são sazonais e influenciam as medidas de interesse para a Saúde (veja neste glossário a diferença de definição para o conceito de variação cíclica).

Séries temporais – são sequências de dados quantitativos tomados em momentos específicos e ordenados, que podem ser estudados segundo sua distribuição no tempo. Este módulo descreve a utilidade deste método para os estudos epidemiológicos e sistematiza os procedimentos estatísticos que tornam vantajosa sua aplicação para a saúde coletiva.

Tendência – É a variação de longa duração das séries temporais. Diz-se crescente, quando seus valores aumentam de modo estatisticamente significativo ao longo do tempo, a despeito de outros componentes de variação (sazonal, cíclica, e aleatória). De modo análogo, diz-se decrescente quando ocorre o inverso. E, de modo complementar, diz-se estacionária quando a magnitude das medidas não muda substancialmente ao longo do tempo. A taxa de crescimento (ou de declínio) médio anual pode ser estimada por meio de análise de regressão linear (modelos de Prais-Winsten).

Variação aleatória – São flutuações irregulares e erráticas, sem importância em termos de magnitude (quando comparada aos valores aferidos), e que são causadas por fatores que ocorrem ao acaso e não podem ser antecipados, detectados, identificados ou eliminados.

Variação cíclica – São variações em ciclos periódicos e regulares. Conquanto as variações sazonais correspondam a repetições ao longo do ano; as variações cíclicas são diferenciadas pela referência a períodos mais prolongados, isto é, dois ou mais anos (vide a diferença de definição para o conceito de variação sazonal).

Módulo 7

Sistemas de Informação Geográfica Aplicada a Análise da Situação de Saúde

Walter Massa Ramalho

Universidade de Brasília (UnB), Faculdade da Cielândia, Brasília/DF, Brasil.

Christovam Barcellos

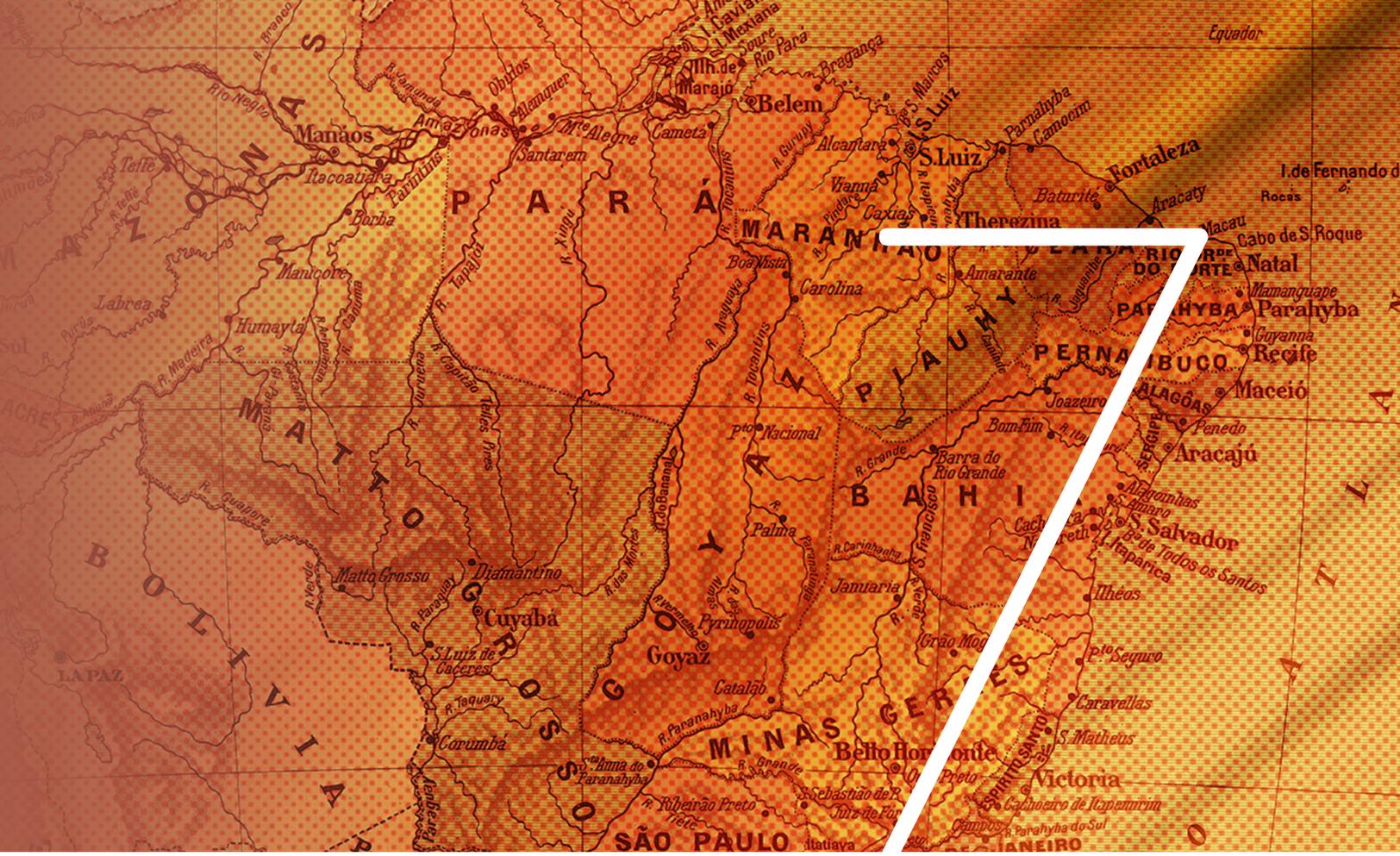
Fundação Oswaldo Cruz (Fiocruz), Centro de Informação Científica e Tecnológica, Departamento de Informação em Saúde, Rio de Janeiro/RJ, Brasil.

Introdução

A grande oferta de indicadores, bases cartográficas, o desenvolvimento de programas de computadores com interface amigável, bem como de equipamentos de baixo custo e alta capacidade de memória, possibilitaram a difusão do geoprocessamento no Brasil e possibilidades de sua utilização no cotidiano das instituições. Mas para compreender o território, e toda a complexidade que nele se desenvolve, é preciso aplicar algumas técnicas e conceitos sobre espaço geográfico e geoprocessamento. A incorporação de sistemas de geoprocessamento pelos serviços de Saúde envolve necessariamente a capacitação de profissionais para análises geográficas e entendimento dos processos que envolvem a dimensão espacial.

Neste sentido, este módulo tem os seguintes objetivos:

- Apresentar conceitos de contexto geográfico na perspectiva da saúde.
- Descrever os princípios de funcionamento de um Sistema de Informação Geográfica.
- Fornecer bases conceituais para o uso da cartografia no mapeamento digital.
- Mostrar técnicas de análise espacial que podem ser usadas para as análises de saúde.
- Capacitar o aluno a aplicar técnicas de análise espacial.



Unidade 1

Abordagem Espacial da Situação de Saúde

1. Espaço e Saúde em Breve Relato Histórico

Ao longo da história, a evolução do conhecimento científico permite compreender o processo de saúde-doença de acordo com o entendimento do processo biológico e social dominante no instante histórico. Em muitos casos foi considerada a influência do espaço geográfico na vida das coletividades humanas para o entendimento de nexos causais de diversos agravos.

Na antiguidade, Hipócrates, 480 a.C., foi um dos primeiros a questionar o paradigma da origem divina do adoecimento e da morte, considerando essencial a geografia e o meio ambiente como determinantes das condições de saúde das populações. Sua publicação “*dos ares, dos mares e dos lugares*”, é tida como um clássico na área da geografia médica¹. Bem mais tarde, o desenvolvimento de conhecimentos cartográficos e a descoberta do Novo Mundo motivaram a publicação *An Essay on Diseases Incidental to Europeans in Hot Climates: with the method of Preventing their Fatal Consequences*, na qual o médico escocês James Lind, no final do século XVIII, descreveu sobre a variação geográfica da distribuição doenças e áreas de risco². Posteriormente, trabalhos sistemáticos de geografia médica foram realizados com o propósito de subsidiar medidas preventivas nas ocupações militares em territórios considerados insalubres do mundo tropical³.

Porém, apenas com John Snow, em 1854, as bases do método epidemiológico começaram a se constituir. Em torno de uma forte epidemia de cólera em Londres, Snow publica o livro “Sobre a Maneira de Transmissão da Cólera”, baseado no estudo de dados de mortalidade, no registro de sua cronologia e na localização geográfica das residências dos óbitos, conseguindo estabelecer nexos causais entre a enfermidade e a água, mostrando que grupos populacionais situados no entorno de uma fonte de água contaminada, localizada na *Broad Street*, tiveram maiores riscos de morrer pela doença⁴.

De posse de um mapa contendo as ruas de Londres e a localização das fontes de água, Snow representou as mortes por cólera a partir de visitas domiciliares. Na sua análise, duas características o ajudaram a elucidar a forma da transmissão da cólera, até então totalmente desconhecida. A primeira característica foi o padrão espacial da ocorrência dos óbitos, bastante concentradas em determinada área da cidade, sugerindo algum tipo de relacionamento

entre eles. A segunda característica foi a proximidade geográfica com determinada fonte de abastecimento na Broad Street, além da confirmação, mediante entrevistas, da rotina do abastecimento de água da fonte em questão pela população das residências com óbitos por cólera.

Posteriormente, Snow constatou que a fonte da *Broad Street* tinha uma característica distinta das demais fontes de distribuição de água espalhadas pela cidade. A captação estava situada a jusante do rio Tâmisa, ou seja, toda a água servida pela cidade, além de dejetos, era despejada no rio, que voltava a ser ofertada novamente por meio daquela fonte específica. Apesar de não conhecer propriamente o agente etiológico, a fonte foi considerada perigosa e interdita com subsequente diminuição dos casos em Londres.

Como no exemplo de Snow e tantos outros posteriores, a utilização do espaço foi imprescindível para o estabelecimento de nexos causais do fenômeno estudado por presença de moduladores espacialmente condicionados, que muitas vezes não estão propriamente ao alcance do investigador. Os mapas diferenciam-se por oferecer ao investigador possibilidades de integração de diversas informações, no formato numérico, como os indicadores sociais, demográficos, econômicos etc., e em formato gráfico, como elementos geográficos ou ambientais, permitindo a formulação de hipóteses etiológicas por inspeção visual e acuidade investigativa diante de um mapa.

2. A Percepção Geográfica

Eventos em saúde, em sua maioria, estabelecem padrões de ocorrência com estruturas subjacentes, implícitas ou explícitas, correlacionados, denominadas de efeitos que podem ser descritos por procedimentos analíticos, permitindo evidenciar o seu arranjo espacial. Bailey e Gatrell⁵ assinalam três abordagens:

a) *Visualização*: compreende a elaboração de mapas temáticos, que permite evidenciar regularidades ou irregularidades espaciais dos fenômenos estudados, por meio de percepção de sua intensidade e proximidade, tornando as unidades observadas mais parecidas quanto mais próximas entre si. Inclui procedimentos estatísticos convencionais para explorar a variável mapeada como medidas de tendência central e de dispersão. Procura-se perceber a dependên-

cia espacial e outros efeitos nos mapas temáticos devidos à formação de gradientes de risco em mosaicos, em que as taxas mais elevadas e as menores taxas tendem a se concentrar, porém em locais diferentes, refletindo um padrão de ocorrência do fenômeno no espaço.

b) *Análise exploratória espacial*: descreve os padrões do fenômeno estudado e testa a dependência espacial, ou seja, se há interdependência entre o fenômeno de saúde em estudo e os arranjos espaciais existentes. Utiliza técnicas de identificação de *cluster* ou padrões atípicos. Possibilita também o desenvolvimento de hipóteses para serem testados em análise mais complexas.

c) *Modelagem*: envolve um conjunto diferenciado de ferramentas estatísticas para estimar e testar a extensão e a forma de dependência do fenômeno estocástico, que é representado por incertezas matemáticas, seguindo uma distribuição de probabilidade.

3. A Estética da Tecnologia

Parece certo que as tecnologias exercem grande influência na vida social, especialmente as que, de alguma forma, manejam informações. A era da informação está instalada com maior conteúdo, consumo, acesso e oportunidade, modificando a forma de como as pessoas se relacionam, aprendem e até compram. A palavra de ordem é conectividade e interatividade, que não se restringe apenas aos computadores ligados em rede física com a internet, mas também atrelado aos mais diversos aparelhos que se reconhecem e interagem entre si.

Aproveitando as interfaces atuais, a geografia está sendo impulsionada para uma nova onda, não só pela divulgação da informação propriamente dita, mas também na sua forma, proporcionando ao usuário comum o desenvolvimento de raciocínio espacial. Com a popularização da utilização de sistemas de posicionamento global em veículos, em aplicativos de *smartphones* e *tablets*, a tecnologia de busca de mapas proporcionada por empresas como o *Google*, as previsões do tempo mostrado em formato de globo, a nova linha de desenvolvimento está baseada na “realidade aumentada” pela integração entre a geografia e informações em tempo real.

Enquanto a oferta de dados e os sistemas de informações com algum tipo de atributo espacial proporcionam às empresas e instituições de Saúde a acurácia nos processos de monitoramento, planejamento estratégico e explicação de eventos relacionados à saúde com atributos espaciais, a divulgação de informação e análise para um público mais exigente e ávido por consumo será efetiva na medida da sua aproximação com a linguagem estética atual.

Unidade 2

Representação Geográfica em Computadores

Esta unidade aborda os problemas básicos de representação computacional dos dados geográficos adotado pelos *softwares* e algumas estratégias de estruturação. Apesar de observarmos diferentes maneiras de representação do universo geográfico em computadores, vamos aqui apresentar conceitos mais tradicionais e genéricos destinados a usuários.

Por outro lado, mesmo se tratando de um conteúdo destinado a profissionais de Saúde, faz parte do conteúdo alguns pontos referentes à cartografia básica, destinada à melhor compreensão da forma da terra e para a configuração dos sistemas de informações geográficas.

1. Geoprocessamento e Sistema de Informação Geográfica

O geoprocessamento conceitua-se de forma ampla e refere-se a todo processamento informatizado de dados georreferenciados. Para tanto, são incluídos os sistema de informação geográfica, o GPS para trabalhos de campo, do sensoriamento remoto para a análise de imagens do terreno obtidas por sensores acoplados a satélites e aviões, a cartografia digital para o desenho e o armazenamento de mapas em formato digital e os sistemas de gerenciamento de bancos de dados espaciais, que permitem conectar dados numéricos e de texto aos objetos dos mapas.

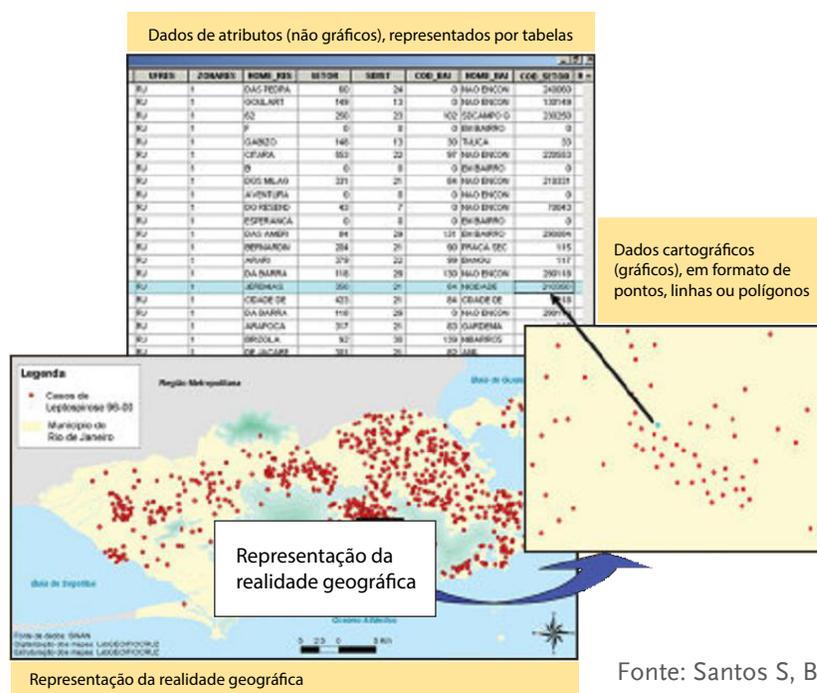
Figura 7.1: Conjunto de ferramentas do Geoprocessamento.



Fonte: Ramalho et al., 2007, p. 16⁶.

O Sistema de Informação Geográfica (SIG) é um sistema constituído por *hardware* (parte física de equipamentos), *software* (símbolos e linguagem aplicados a um modelo computacional), *peopleware* (pessoas que gerenciam, executam, programam e mantêm o sistema) e procedimentos construídos para suportar a captura, gestão, manipulação, análise, modelação e visualização de informação referenciada no espaço, com o objetivo de resolver problemas complexos de planejamento e gestão que são inerentes à realização de operações espaciais⁷. Esta definição é melhor compreendida quando pensamos na palavra Sistema como um conjunto de atividades articuladas com finalidades, objetivos e protocolos conhecidos.

Figura 7.2: Informações dos dados espaciais.



O grande diferencial do SIG em relação a outros sistemas, como por exemplo o CAD, é a propriedade de manejar dados não gráficos (dados de atributos ou tabelas) com dados gráficos (dados cartográficos). Destacam-se ainda outras duas características importantes: habilidade de sobreposição e integração de diferentes objetos geográficos e capacidade analítica, traduzida nas operações geográficas, na análise estatística e na gerência de grande volume de informações. Esta arquitetura permite estabelecer quatro funções básicas:

a) *Aquisição de dados:* Captura, importação, validação e edição são procedimentos que envolvem as etapas necessárias à alimentação do sistema. A partir de uma série de condições imposta pelo sistema relacionada à configuração de estruturas, os dados podem ser digitalizados, reconstituídos ou mesmo transformados. Operações comuns de importação e transformação de formatos também são bastante utilizados.

b) *Gerenciamento de banco de dados:* Envolve o armazenamento dos dados de forma estruturada, de modo a possibilitar e facilitar a realização de análises. A forma como os dados são estruturados é crucial para o sistema, pois dela dependem os tipos de análises que poderão ser realizados. A um mapa armazenado no SIG, sempre podem ser associadas novas informações, provenientes de diversas fontes, permitindo que se incorpore o trabalho de diferentes órgãos e instituições. Ele permite combinar vários tipos de informações, como, por exemplo, dados obtidos em campo por GPS ou por topografia convencional, tabelas, mapas, imagens, entre outros.

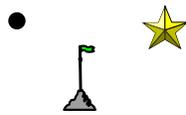
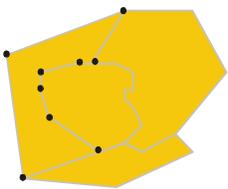
c) *Visualização e apresentação cartográfica:* Um SIG necessita ter agilidade para utilizar as diversas camadas de dados e exibir este resultado por meio de mapas de síntese com boa qualidade gráfica. Os mapas anteriormente feitos à mão, transformam-se agora em um produto de todas as operações desenvolvidas dentro do SIG, com inúmeras possibilidades de atualização mais constante.

d) *Consulta e análise:* Uma função que pode ser considerada como a principal de um SIG é a de análise, pois possibilita operações de extração e geração de novas informações sobre o espaço geográfico, a partir de critérios especificados pelo próprio usuário. As operações mais comuns são a pesquisa de dados, a busca de informações de acordo com algum critério de seleção (por exemplo, pela localização, proximidade, tamanho, valor) e a análise espacial que envolve modelagem e análise de padrões espaciais.

2. Estrutura de Dados para Edição de Mapas

a) *Feições geográficas:* As primitivas gráficas são definidas em função da escala do mapa e das dimensões espaciais do fenômeno (ou feição geográficas) com o objetivo de simplificação da forma do mundo real que melhor informa o fenômeno estudado. De acordo com a necessidade de informação e disponibilidade dos dados, o SIG oferece duas maneiras de representar graficamente o mundo: os modelos vetoriais e os matriciais.

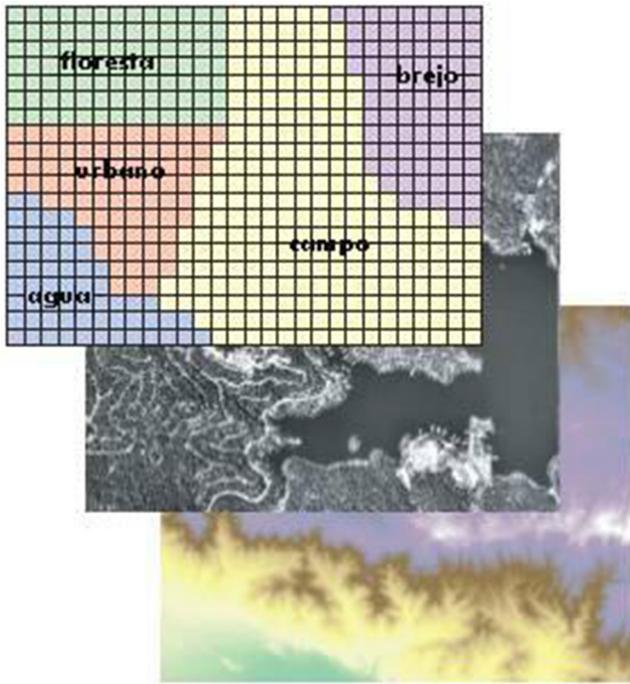
- **Modelo vetorial (ou discreto):** Os modelos vetoriais discretos assumem um valor real, finito e enumerável associado a cada unidade geográfica e desconexo de uma escala entre si, independente do seu tamanho, portanto distinto dos modelos contínuos, que podem assumir valores infinitos, correspondentes a alguma escala contínua sem interrupções ou saltos, de forma a apresentar conexões entre si. Neste modelo, os objetos geográficos são armazenados em uma estrutura de localização que são expressos pelas seguintes representações topológicas:

<p>Pontos, utilizados para a representação de cidades, em mapas de pequena escala, ou em grande escala para residências, eventos de saúde, etc. Podem ser representados por diversas formas, mas guardam apenas um par de coordenadas.</p>	
<p>Linhas, para representar rios, estradas, curvas de nível;. Sua forma expressa apenas o comprimento.</p>	
<p>Polígonos, que representam áreas (regiões, municípios, lagos) Necessita a conexão de no mínimo quatro pontos, com o primeiro e o último representado por um par de coordenadas idênticas, formando um polígono fechado.</p>	

- **Modelo matricial (ou raster):** Os dados são armazenados como uma imagem ou matriz de pixel (*picture element*), capturados por imageadores como sensores de satélites ou por equipamentos de varreduras de mesa (*scanners*). Sua estrutura é contínua, em que a superfície geográfica é representada por atributos com valores de uma escala de bits e apresentada na tela por uma cor correspondente a esta escala.

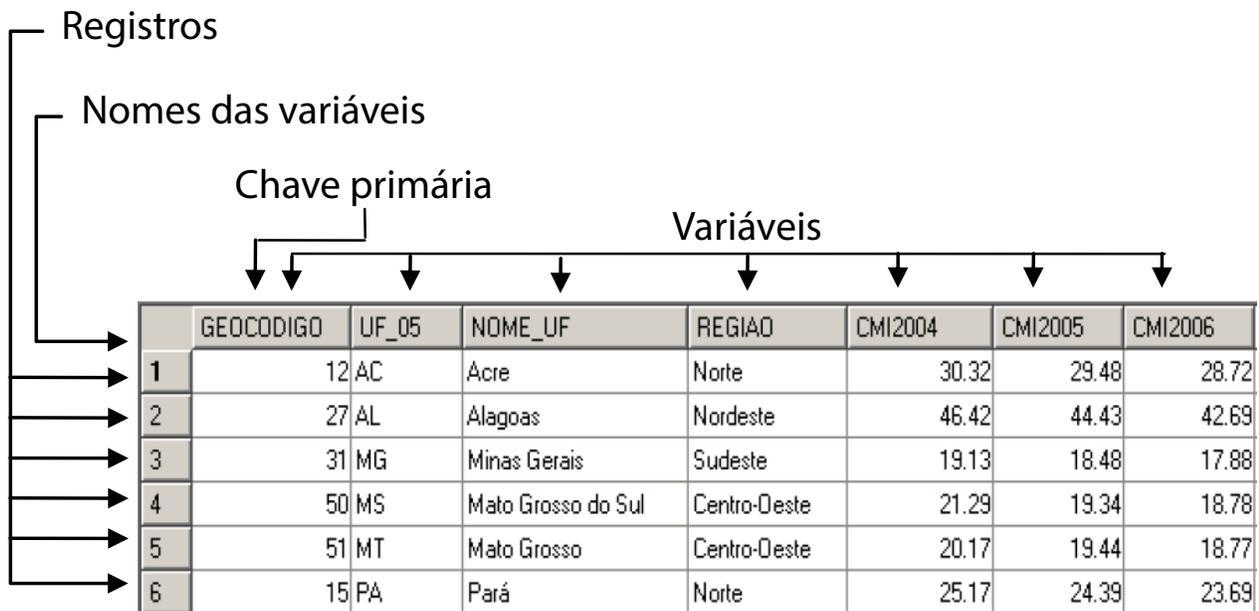
b) Atributos: Os atributos são conjuntos de dados que caracterizam os objetos em um SIG. Tem as propriedades de um banco de dados, isto é, são formados por linhas (ou registros) e colunas (ou variáveis). Cada objeto geográfico possui um elo com um registro de forma unívoca entre si (um para um) e uma chave primária que identifica a sua unicidade, isto é, não poderá ser repetida.

Figura 7.3: Modelo de representação matricial.



Fonte: Francisco et al., 2007⁹.

Figura 7.4: Modelo esquemático de uma tabela de atributos.

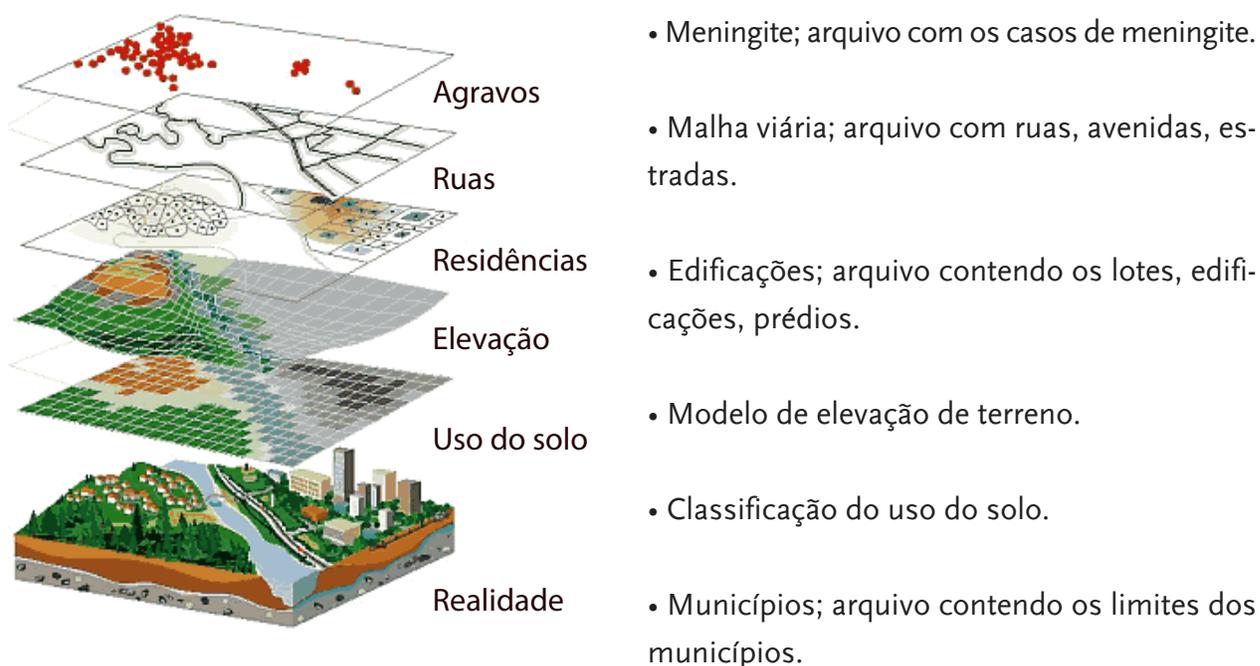


Fonte: Elaborado pelos autores.

Na tabela, existem variáveis que são chaves identificadoras (Geocodico), que funcionam como chave para ligação com os mapas e outros atributos das UFs (NOME_UF; REGIAO; CMI2004; CMI2005 e CMI2006). Algumas variáveis são numéricas (CMI2004; CMI2005 e CMI2006) e outras alfanuméricas ou textuais (UF_5, NOME_UF e REGIAO).

c) *Tema, camada ou plano de informação*: Uma camada de informação é composta por uma coleção de objetos geográficos (com geometria, topologia e atributos) com características comuns. Nela estão armazenados também informações sobre a projeção cartográfica e os seus atributos, isto é, já incorpora o banco de dados associado. O resultado das sobreposições dos planos de informações de interesse configura-se na materialização do mapa para a análise na tela de visualização.

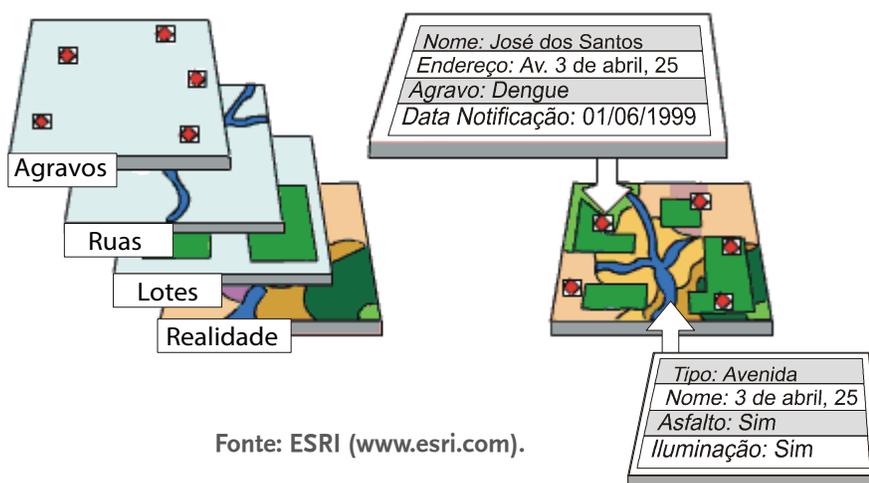
Figura 7.5: Modelo de camadas em um SIG.



Fonte: Adaptado de NOAA, 2013¹⁰.

A arquitetura de dados em forma de camadas permite que várias informações sejam visualizadas simultaneamente e que as informações estejam acessíveis. Com um *click* em um botão, o *software* pode mostrar os atributos de cada objeto visível.

Figura 7.6: Modelo operacional em um ambiente SIG.



3. Introdução à Cartografia para Mapeamento Digital

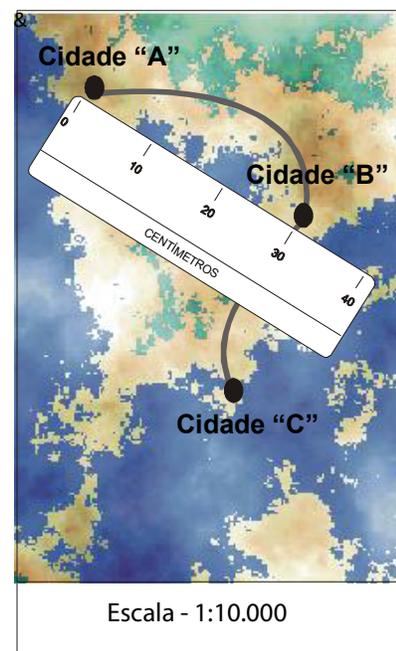
Em um SIG, os dados gráficos necessitam que possuam projeção cartográfica, sistema de coordenadas e sistema geodésico (*datum*) conhecidas e precisam ser corretamente referenciadas ao sistema de coordenadas terrestres para serem sobrepostas e representar a geografia real. Os parâmetros cartográficos de importância ao manuseio correto estão detalhados a seguir.

a) *Escala*: Como o mapa sempre é uma redução da geografia do terreno em uma representação (em papel ou digital), a escala mostra ao usuário a relação matemática das formas e distâncias dos objetos geográficos. Assim, ela estabelece uma relação entre o mundo real e a representação cartográfica da área. Em termos lineares pode ser estabelecida como a relação entre os comprimentos observados na representação e no terreno.

$$\text{Escala} = d/D = d \text{ mapa}/D \text{ terreno}$$

onde “d” e “D” representam as distâncias.

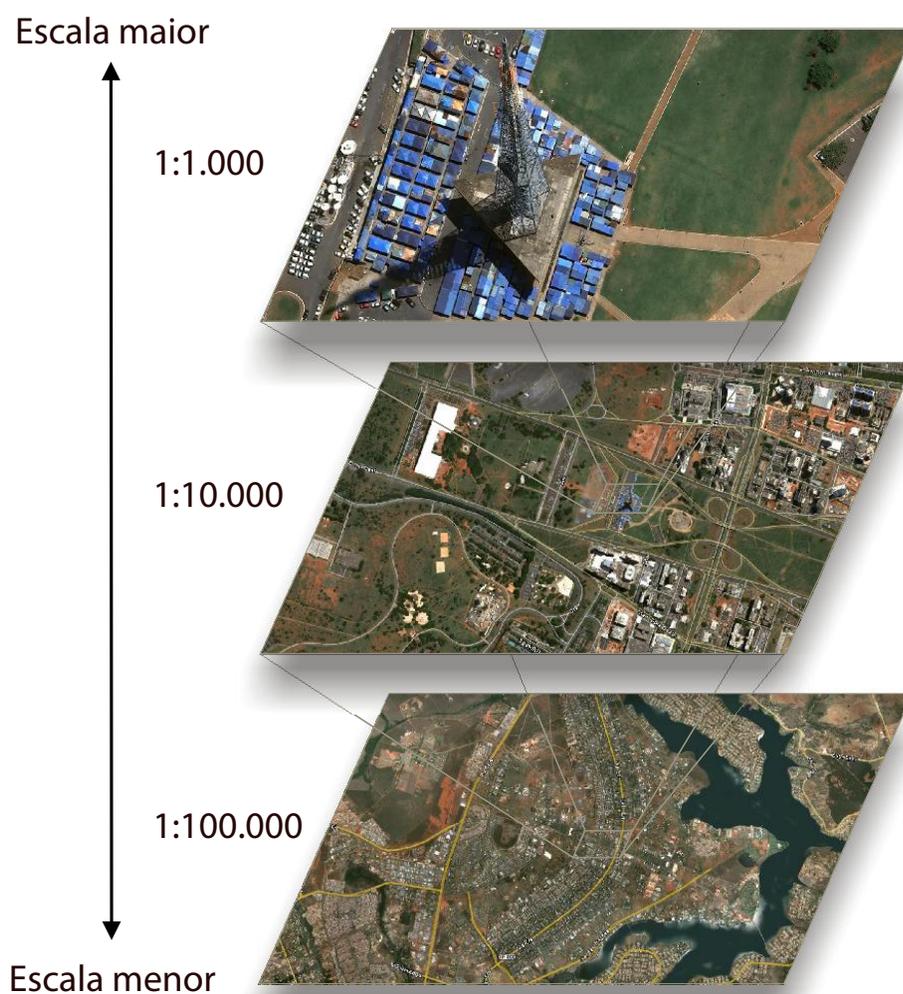
Figura 7.7: Representação de escala em um mapa.



Fonte: Elaborados pelos autores.

Normalmente encontramos a representação “d” : “D”, onde para cada “d”, unidades encontrada no mapa, equivale a “D” real. Se a escala for 1:1.000, ou 1/1.000, então para cada 1 unidade no mapa, há 1.000 unidades no terreno. Com uma régua, verificamos que no mapa da Figura 7.7, as cidades A e B se separam por 30 cm. A distância real “D”, então é : $30\text{cm} \times 10.000 = 300.000\text{cm}$ ou 3.000m ou 3km . Assim, quanto maior o denominador, menor serão os detalhes do mapa, e é representado por uma pequena escala, e inversamente, em grande escala, que é representada por um denominador pequeno, o nível de detalhamento representado no mapa é bem maior, como mostra a Figura 7.8.

Figura 7.8: Relação de representação entre as escalas.

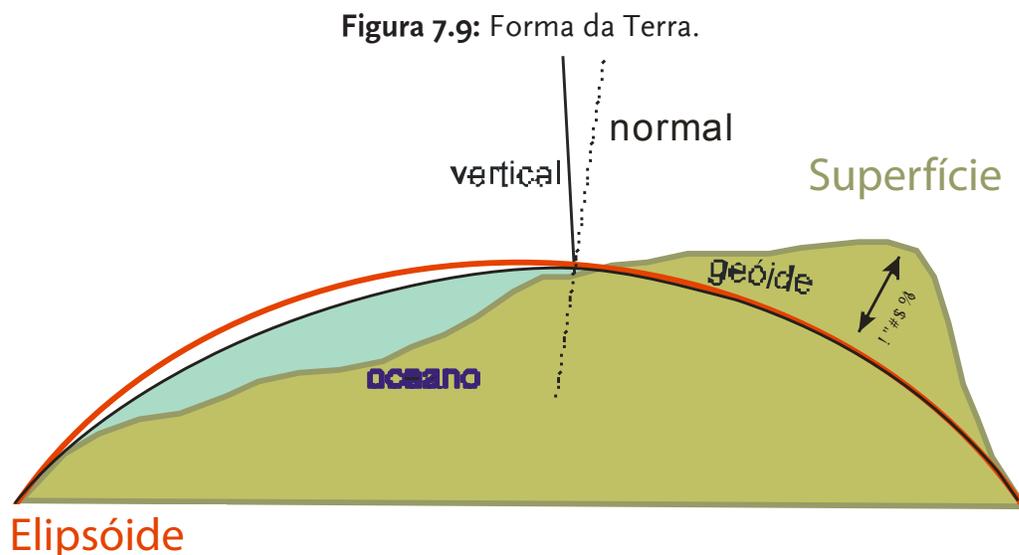


Fonte: Google Earth.

O erro cartográfico de um mapa é dado pela capacidade do olho humano em discernir linhas no processo de digitalização de mapas em papel, utilizando equipamentos de mesa digitalizadora. Aceita-se um valor médio em torno de 0,2 mm de diâmetro. Portanto, o erro cartográfico em um mapa cuja escala de 1:50.000 é de $0,2 \times 50 = 10$ m, isto quer dizer que todo detalhe do terreno, com menos de 10 m, reduz-se a um ponto nesta escala. Se o detalhe merecer representação, essa se dará pelos sinais convencionais ou cartográficos.

Este parâmetro é muito importante para o autor do mapa decidir como representará um objeto geográfico. Se, por exemplo, um domicílio tem em média menos de 10 m de largura, ele deve ser representado como um ponto em um mapa de escala 1:50.000. Para mapas de escala maior (1:1000, por exemplo), um domicílio pode ser representado como um polígono.

b) *Sistema Geodésio*: A Terra tem forma quase esférica, conhecida como globo terrestre. Apesar de assumir essa forma, em estudos em que se exige precisão de posicionamento, como é o caso da maioria das representações cartográficas em mapas e cartas, devem-se considerar mais cuidadosamente as pequenas diferenças da sua forma⁷. O modelo que mais se aproxima de sua forma real é o Geóide. Como o geóide é irregular, seu tratamento matemático é bastante difícil. A superfície matemática adotada para efeito de cálculos é o elipsóide de revolução, criado a partir de uma elipse rotacional em torno do eixo menor (eixo polar tem um raio aproximadamente 23 km menor que o equatorial).



Fonte: Adaptado de Carvalho et al., 2000⁷.

O *datum* geodésico constitui medidas usadas como subsídio para a definição do elipsóide de revolução e o sistema de referência. Centenas de *datums* têm sido concebidas, ao longo dos tempos, evoluindo desde os modelos esféricos da Terra até aos modelos elipsoidais atuais. Referir coordenadas ou projeções a um *datum* errado pode resultar em erros de posição, daí a necessidade de uma cuidadosa identificação dos *datums* em questão no momento em que se trabalha com mapas – mesmo os digitais.

Cartas brasileiras antigas guardam as referências do antigo *datum* planimétrico de Córrego Alegre. A atual caracterização do sistema geodésico brasileiro integra o *Datum* Sul-Americano de 1969 (SAD69), com referencial horizontal em Chuá/MG e vertical em Imbituba/SC, e adotou-se a figura geométrica com o Elipsóide Internacional, de referência de 1967. Atualmente, a superfície de referência está mudando para o Sistema de Referência Geocêntrico para as Américas (Sirgas), que deverá estar presente nos próximos produtos cartográficos nacional.

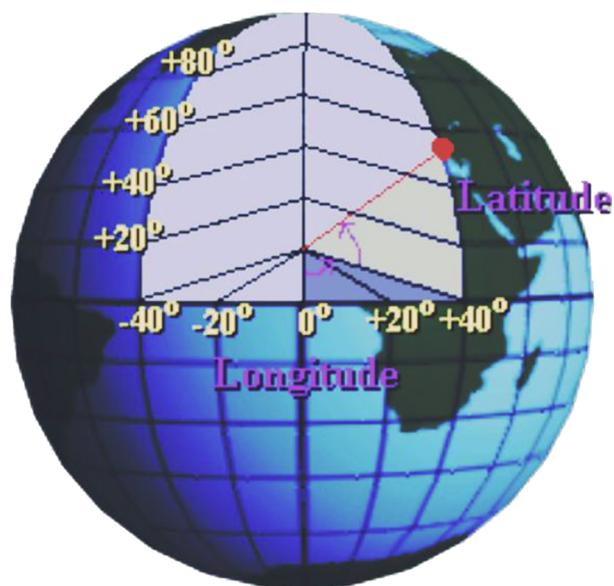
Sistemas globais, como o Sistema de Posicionamento Global (GPS), utilizam o datum planimétrico global de origem geocêntrica de referência denominado de WGS84 (*World Geodetic System* de 1984) para o cálculo de coordenadas.

c) *Sistemas de coordenadas*: O sistema de coordenadas geográficas emprega uma superfície elipsoidal ou ao geóide, e o globo terrestre é cortado por linhas imaginárias denominadas de paralelos e meridianos. Considera que pontos distribuídos na superfície da terra podem ser representados pelos ângulos em relação ao centro da Terra.

Estas medidas são chamadas coordenadas denominadas de latitude e longitude, sendo expressas em medidas angulares – graus, minutos e segundos ($nn^{\circ} nn' nn''$) ou em graus e décimos de graus ($nn,nnnn^{\circ}$).

Tendo-se os valores da latitude e da longitude de um local desejado, estarão determinadas as coordenadas geográficas dele. A grande desvantagem deste sistema é a dificuldade para confecção de mapas e o cálculo de distâncias lineares.

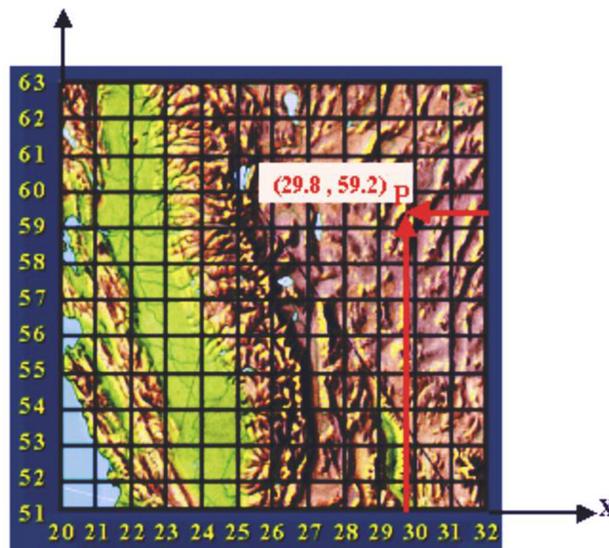
Figura 7.10: Diagrama das latitudes e longitudes.



Fonte: Carvalho et al., 2000, p. 95⁷.

Já o sistema de coordenadas planas, ou cartesianas, baseia-se no formato plano da Terra, que é projetada sobre um plano cartesiano, com eixos perpendiculares x e y. Por isso, estas coordenadas são medidas em metros em relação a um ponto de origem.

Figura 7.11: Sistemas de coordenadas planas.



Fonte: Carvalho et al., 2000, p. 94⁷.

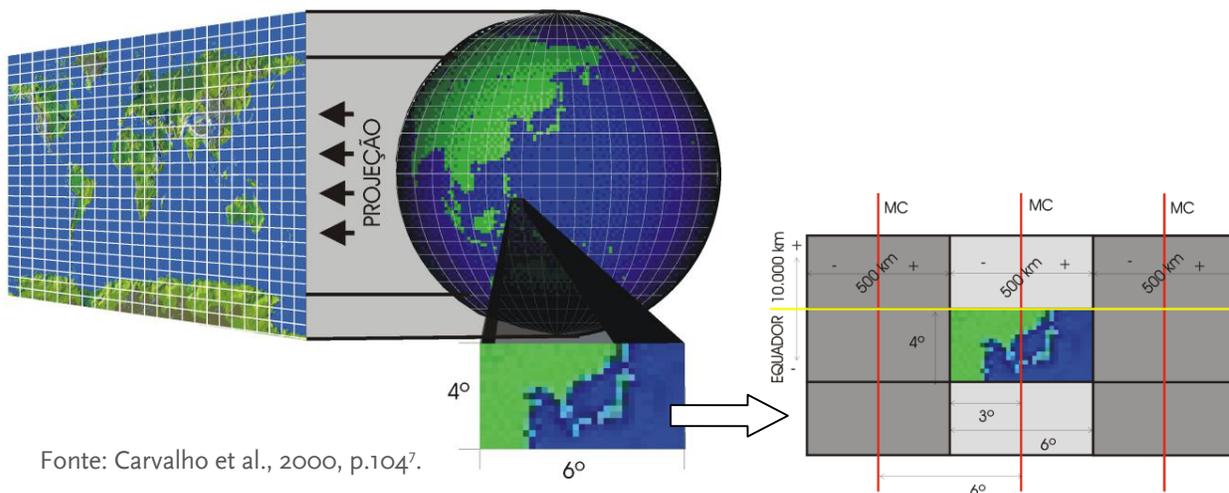
Existem diferentes projeções cartográficas: as planas cilíndricas e as cônicas, mas todas apresentam deformações, sejam em relação às distâncias, às áreas ou aos ângulos.

A projeção cilíndrica de Mercator, conhecida como Universal Transverse Mercator (UTM) foi adotada em 1955 pelos órgãos responsáveis pelo mapeamento sistemático do Brasil, e ainda hoje é bastante utilizada em vários tipos de levantamento.

Para estabelecer estas coordenadas, a Terra é recortada em pequenas quadrículas, com 6° de amplitude longitudinal e 4° na latitude. Estas quadrículas são projetadas em um plano. O mapa mundial é, então, formado pela união destas cartas planas formando uma “cocha de retalhos” contendo 60 colunas por 40 faixas, minimizando as deformações.

Diferentemente do sistema de coordenadas geográficas que possuem única origem (os valores de uma localização são únicos no mundo), a UTM possui origens parciais de coordenadas para cada amplitude longitudinal – denominado de fuso UTM. A diferenciação entre as amplitudes deve-se ao Meridiano Central, que é uma linha imaginária numerada de 1 a 60 a partir do antimeridiano de Greenwich (longitude 180°) no sentido oeste-leste, e que intersecta a amplitude a 3°. Esta interseção de um fuso com uma faixa é designada por zona UTM representada pelo número do fuso com a letra da faixa (zona 25M).

Figura 7.12: Sistema UTM.



O mapa na Figura 7.13 mostra os fusos que cobrem o território brasileiro. Dependendo do local que se deseja mapear, deve-se optar pelo fuso que contém este local. Por exemplo, a Cidade de Campo Grande está situada no fuso²¹.

Nos programas de geoprocessamento, as conversões de coordenadas dão-se automaticamente devido a parâmetros matemáticos contidos no próprio *software*, mas é preciso informar a ele qual o sistema que estamos usando, o datum e o fuso para que os mapas sejam importados corretamente.

Figura 7.13: Fusos no Brasil.



Unidade 3

Visualização de Distribuição Espacial

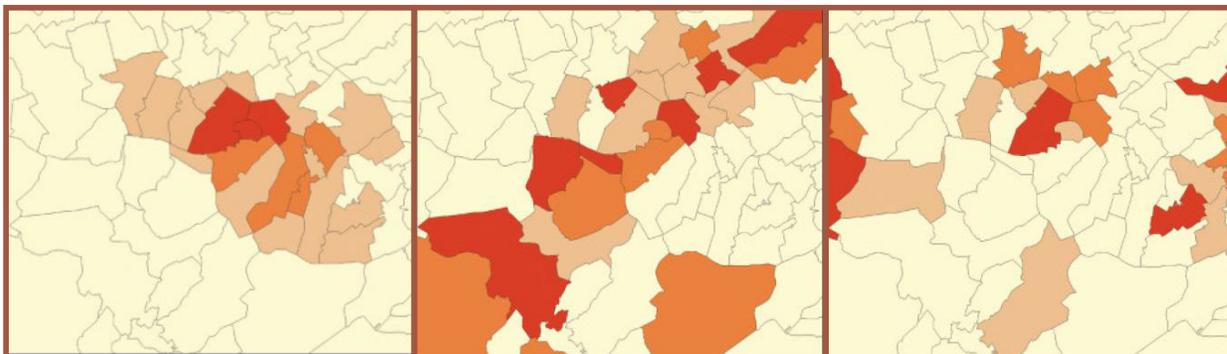
Esta unidade tratará de métodos para a elaboração de mapas temáticos para fim de exploração visual de eventos de saúde, ou seja, quais os procedimentos metodológicos para os indicadores espacializados resultem um produto de fácil compreensão e interpretação, por meio de conceitos de estatística descritiva e semiótica.

1. Forma de Distribuição de Eventos de Saúde

Dados geográficos podem criar informações quando são organizados, analisados e interpretados em formato de mapas temáticos. O propósito de visualização de uma distribuição geográfica particular é o primeiro passo para o levantamento de possíveis hipóteses etiológicas, evidenciando arranjos de proximidades ou dispersão e estabelecendo relações com diversas informações por meio de sua interface gráfica que sugerem intervenções diferenciadas.

A Figura 7.14 mostra três situações distintas. Na primeira, a concentração de taxas elevadas em torno de uma região sugere que existe uma fonte única de exposição, em casos de emissões de poluentes de uma indústria. Na segunda, taxas, especialmente elevadas de forma contígua e em formato de uma faixa, podem indicar exposição em formato linear como uma estrada ou um rio. Na terceira, observa-se ainda que padrões espaciais em formato de mosaico, ou seja, diversos aglomerados de taxas elevadas podem estar relacionados às diversas fontes de exposição distribuídas na região, muitas vezes evidenciados pela análise da situação de saúde que reflete a forma de organização social das cidades.

Figura 7.14: Altas taxas em torno de um ponto, eventos de saúde concentrados em torno de uma faixa e padrão espacial mosaico.



Fonte: Carvalho et al., 2000, p. 102, 105⁷.

Os mapas temáticos devem ser intuitivos, com informações claras e objetivas, seguindo padrões técnicos de representação da organização espacial. Segundo Pina¹¹, cabe ao autor ou analista o conhecimento técnico sobre os princípios da representação gráfica, conhecer os fenômenos que se quer retratar e dispor de dados, além de ter criatividade, bom senso e obviamente experiência. Ao construir o mapa, é necessário ainda ter em mente os objetivos da representação e do público-alvo¹².

Os elementos básicos para a criação de mapas com o intuito de visualização é a simbologia e a sua forma de representação.

2. Simbologia

Como um mapa é um meio de comunicação visual, a simbologia adotada deve representar o tema baseado na classificação dos atributos de suas feições geográficas por meio de contraste e forma, utilizando conceitos referentes à escala de medida e variáveis visuais.

a) *Escala de medidas*: Trata-se da forma de como ressaltar ou simplificar os eventos ou característica associada aos atributos das feições geográficas por alguma técnica de agregação do conjunto de símbolos ou números regida por uma regra estatística. São classificadas em:

- Escala nominal: As características da distribuição espacial são relativas a um atributo de qualidade, não numérico, e com a mesma importância. Ex.: mapa de tipo de vegetação.

- Escala ordinal: A distribuição espacial da variável, apesar de qualitativa, apresenta uma ordenação de valor, devendo ser classificada em hierarquias. Ex.: mapa de risco para malária (alto, médio e alto risco).

- Escala de intervalos: Representada por um conjunto de valores numéricos associados às feições geográficas. Diferente das variáveis qualitativas, necessita agrupar os atributos em classes ou intervalos, utilizando técnicas estatísticas. Na escala de intervalo são adotados os seguintes critérios para divisão classificação:

- Intervalos iguais ou passos iguais: classificam os valores em classes de intervalos constantes.

- Quantis: classifica em classes com o mesmo número de feições geográficas.

- Escala de Razão: Difere da escala de intervalo, pois, os dados são divididos segundo a sua dispersão em torno de uma medida de tendência central.

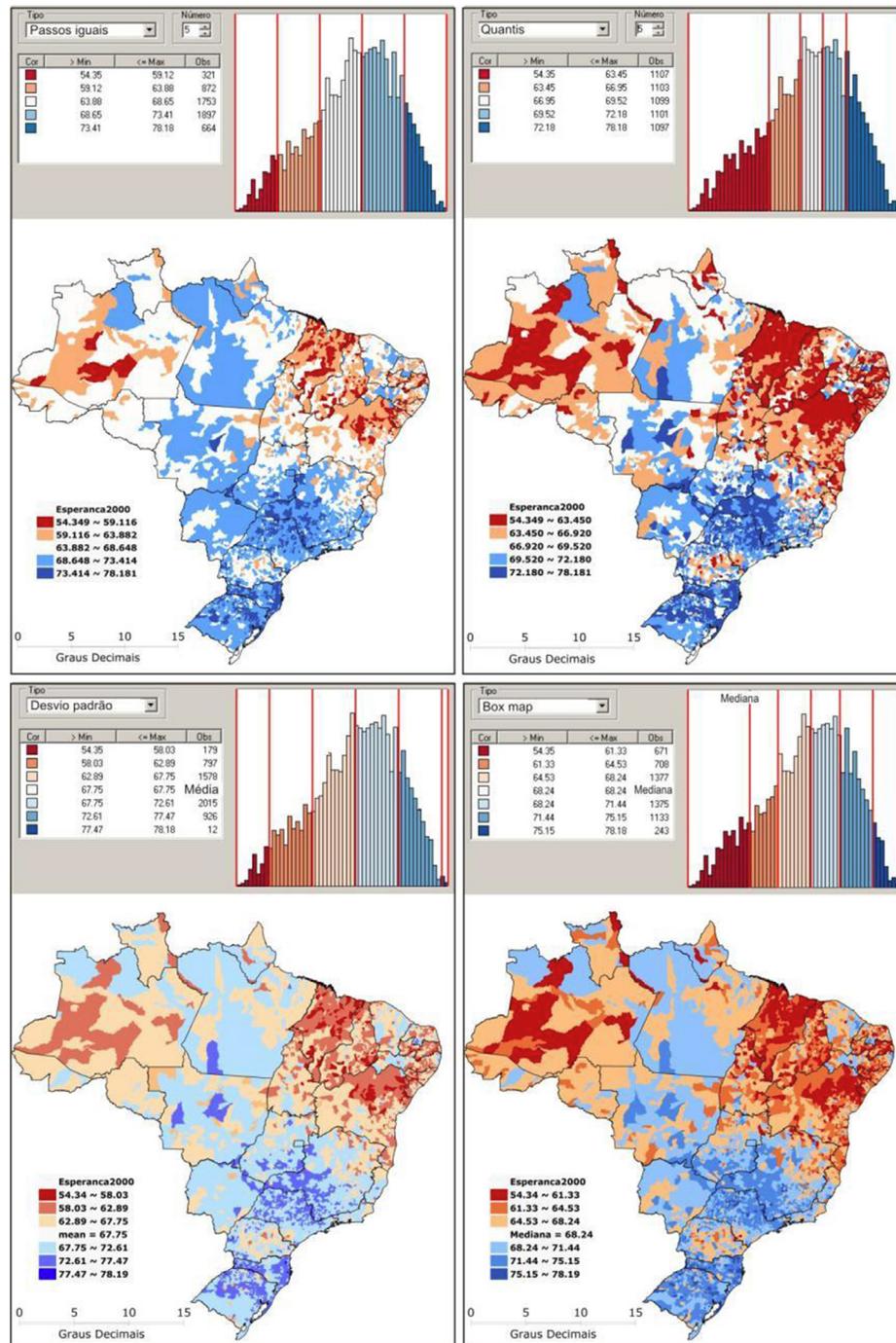
- Desvios-padrões: classificam os valores em torno da média, e as classes anteriores em valores de desvios-padrões (2, 1, ½, etc);

- Box map ou mapa de caixas: são construídos segundo os princípios dos gráficos de caixas ou *box plot*, e classificam os valores em torno da mediana, e as classes, em intervalos interquartílicos. Os grupos de valores são apresentados em seis categorias fixas: quatro quartis, mais duas categorias de valores extremos ou *outlier*, situados abaixo ou acima da distribuição. São classificados como *outliers*, tais como nos gráficos de caixas, os valores situados 1,5 vezes maior que o intervalo interquartílico (diferença entre o 3º. quartil e o 1º. quartil), para acima ou abaixo. Apesar de grande utilidade, ainda não estão completamente implementados na maioria dos SIGs, necessitando da intervenção do analista para a sua realização.

As figuras 6.14 e 6.15 mostram como a apresentação dos mapas temáticos podem transmitir informações distintas, dependendo do método utilizado para classificar o indicador. Na primeira figura é apresentada uma distribuição com formato próximo de uma gaussiana enquanto que na segunda, visualizamos forte desvio, como normalmente se apresentam os dados de Saúde. A escolha do método utilizado, portanto, está associado ao objetivo do investigador, que se pretende destacar os valores extremos, é possível utilizar os métodos de “Passos Iguais” e “Desvio-Padrão”, mas na maioria dos fenômenos de saúde, são preferíveis os *Quantis* e o *Box map* por serem métodos mais robustos.

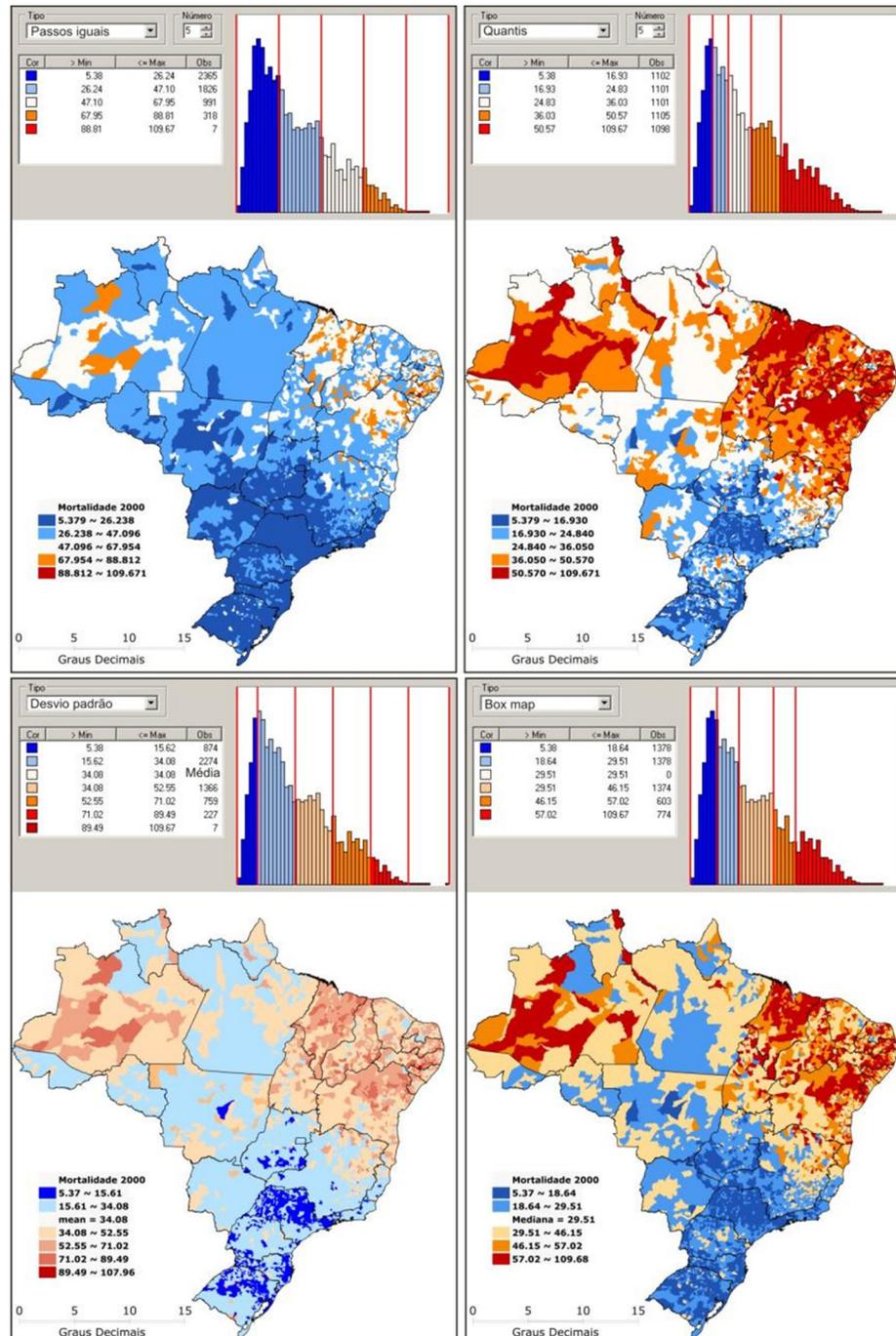
A opção por qual escala utilizar deverá fazer parte de processo parcimonioso e levar em consideração a forma de distribuição do indicador utilizado.

Figura 6.15: Distribuição da esperança de vida ao nascer, segundo município de residência – Brasil, 2000.



Fonte: PNUD, 2003¹¹.

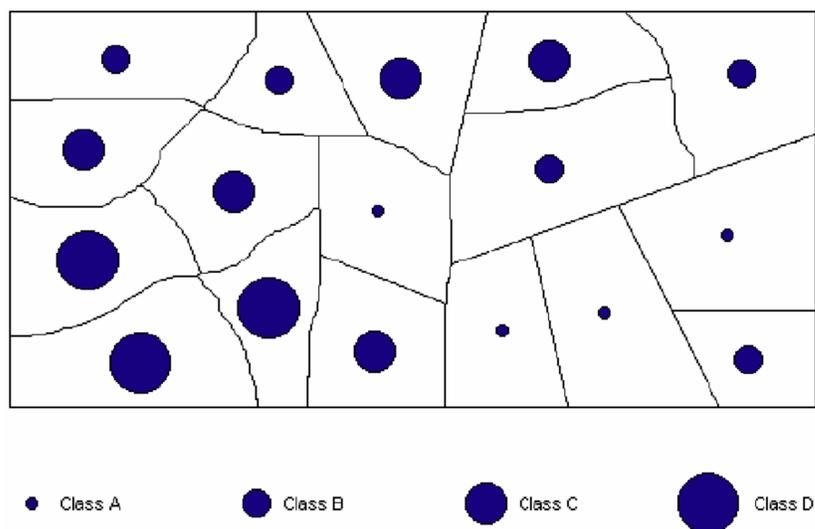
Figura 6.16: Distribuição do coeficiente de mortalidade infantil, segundo município de residência – Brasil, 2000.



Fonte: PNUD, 2003¹¹.

b) *Variáveis visuais*: As variações visuais são combinadas aos atributos das feições geográficas como forma de ressaltar as feições geográficas. As variáveis visuais promovem simplificação da percepção visual com foco no tema abordado, utilizando aspectos relacionados ao tamanho da forma gráfica, o tom da cor, a luminosidade ou valor e a forma como síntese do estudo.

Figura 7.17: Visualização baseado no tamanho.

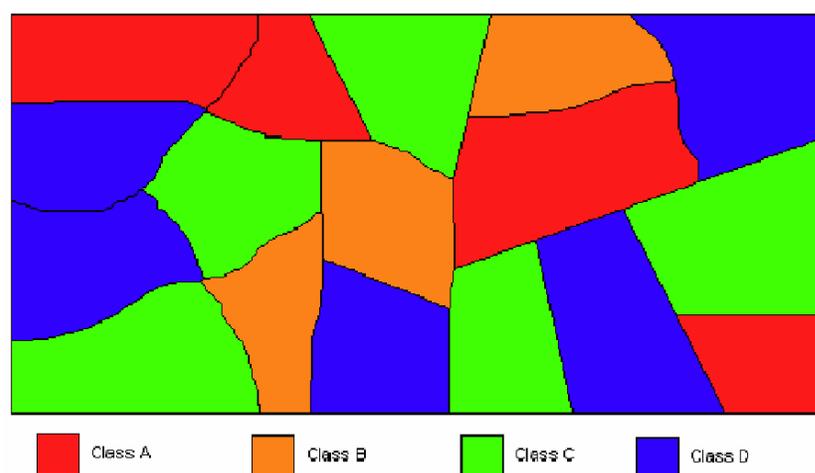


Fonte: Sluter, 2005¹².

- **Tamanho:** Os atributos estão representados por símbolos de diferentes tamanhos. Esta estratégia é utilizada tanto para realçar determinada feição, que por motivo de escala está imperceptível no mapa, quanto para estimular o raciocínio visual de ordem de grandeza um conjunto de observações. São adotadas as escalas ordinais e intervalares.

- **Tom da cor:** É a variação qualitativa de cores básicas que corresponde ao seu comprimento de onda no espectro visível. As feições são diferenciadas pela representação da cor como amarelo, verde, vermelho, azul etc., e devem ser utilizadas para as representações cujas características são definidas no nível de medida nominal, portanto não estabelecem uma hierarquia visual.

Figura 7.18: Visualização baseado no tom da cor.

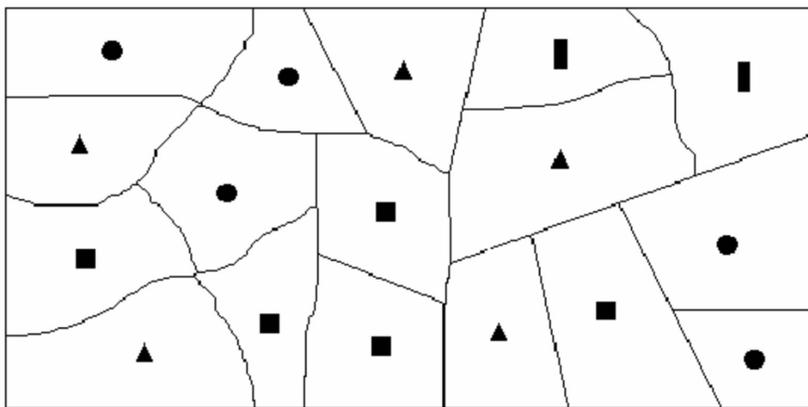


Fonte: Sluter, 2005¹².

- Luminosidade ou valor: As cores são diferenciadas pela quantidade de luz branca incidente, também chamada de gradiente de cor de uma única tonalidade. Esta técnica impõe ao usuário interpretação hierárquica da representação geográfica, portanto é adequada a atributos numéricos e adotada a escalas ordinais e intervalares.

- Forma: Trata de uma variável aplicada a símbolos pontuais para representação de feições de diferentes tipos, portanto mais adequada aos atributos de escala nominal.

Figura 7.20: Visualização baseado na forma.

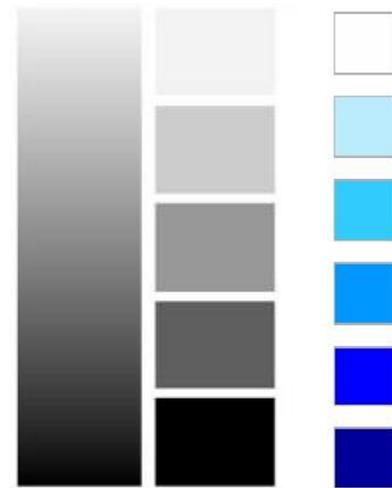


● Class A ■ Class B ▲ Class C ▭ Class D

Fonte: Sluter, 2005¹².

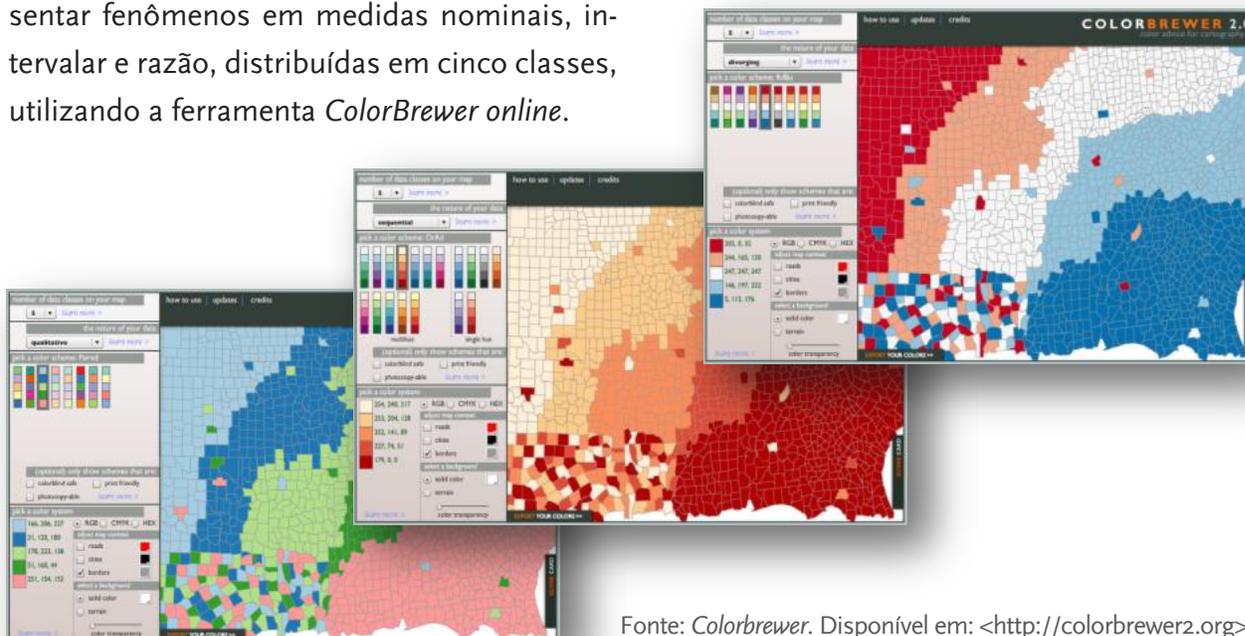
Para o auxílio na escolha das cores e formas nos mapas temáticos, alguns pesquisadores têm disponibilizado bibliotecas de paletas de cores, como resultado de estudos sobre a percepção visual de cores aplicadas na representação dos mais diversos fenômenos. Consultas *online* poderão ser realizadas no sítio do projeto *ColorBrewer* e ilustrado na Figura 7.21.

Figura 7.19: Visualização baseado na luminosidade.



Fonte: Sluter, 2005¹².

Figura 7.21: Combinação de cores para representar fenômenos em medidas nominais, intervalar e razão, distribuídas em cinco classes, utilizando a ferramenta *ColorBrewer online*.



Fonte: *Colorbrewer*. Disponível em: <<http://colorbrewer2.org>>.

Deve-se ressaltar que o número excessivo de classes pode prejudicar a objetividade da informação, já que o olho humano tem dificuldade para distinguir entre os elementos gráficos com mais de seis tons de mesma cor ou mais de doze cores distintas¹³.

3. Representação dos Fenômenos Espaciais

Denomina-se representação a forma de como serão repassadas as informações espaciais úteis na resolução de problemas e tomada de decisão, por meio da evidência dos dados e indicadores. A construção de modelos de visualização tem o poder de enquadrar os fenômenos de interesse segundo a natureza dos dados de atributo e objetivos, podendo ser estruturados em mapas qualitativos e discretos, quantitativos ou comparativos.

a) *Qualitativos e discretos*: Os métodos de mapeamento para os fenômenos qualitativos e os discretos utilizam as variáveis visuais seletivas baseadas na forma, orientação e cor, podendo ser utilizados nas feições geográficas de pontos, linhas e polígonos para representar o

indicador. A Figura 7.22 mostra um mapa discreto de símbolo pontual do tipo nominal, em que a presença da informação estudada é mostrada, cuja disposição dos pontos cria uma interpretação de regionalização no espaço.

Figura 7.22: Mapa discreto. Distribuição dos municípios com agência bancária da Caixa.



Fonte: Instituto Brasileiro de Geografia e Estatística (IBGE)¹⁴.

Mapas denominados de corocromáticos utilizam a coloração como ferramenta cognitiva e escala nominal, portanto não há ordem nem hierarquia. Para correta interpretação visual do usuário o padrão de cores estabelecido deverá ser o de tom de cor, como na Figura 7.23 que representa as zonas climáticas brasileiras. Podem também ser utilizados granulações e orientações.

Figura 7.23: Mapa qualitativo. Zonas climáticas brasileiras.

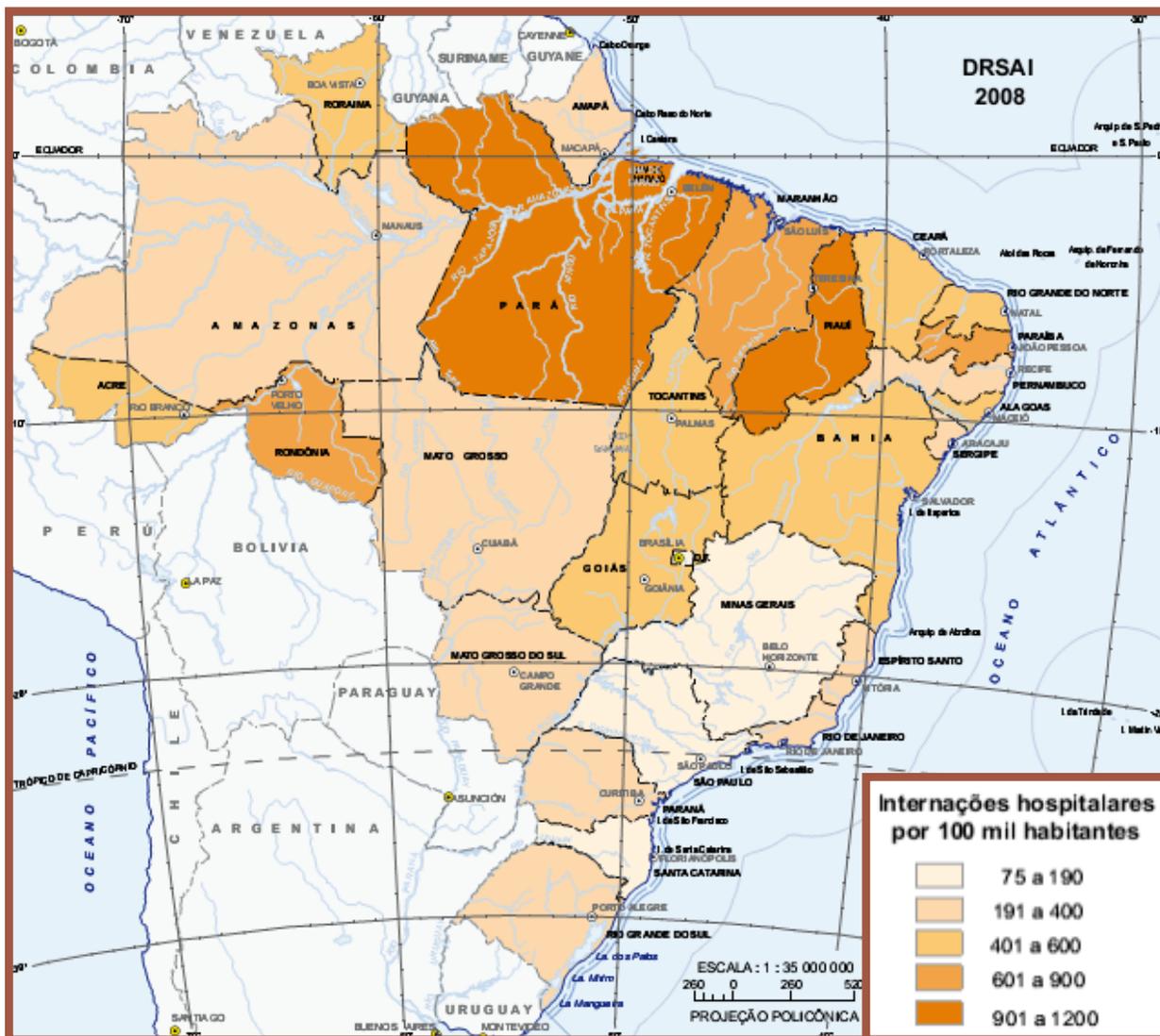


Fonte: Instituto Brasileiro de Geografia e Estatística (IBGE)¹⁴.

b) Quantitativo

- Padrões de áreas: Mapas temáticos coropléticos (do grego *choros*, valor e *plethos*, área, significando mapas com valores associados às áreas) são elaborados a partir de escala de intervalos estabelecida por dados quantitativos. As classes devem ser ordenadas a partir de estratos que valorizam a característica que se deseja examinar, como é o caso de divisão por intervalos iguais ou as de quartil.

Figura 7.24: Internações hospitalares por doenças relacionadas ao saneamento ambiental inadequado Brasil, 2008.

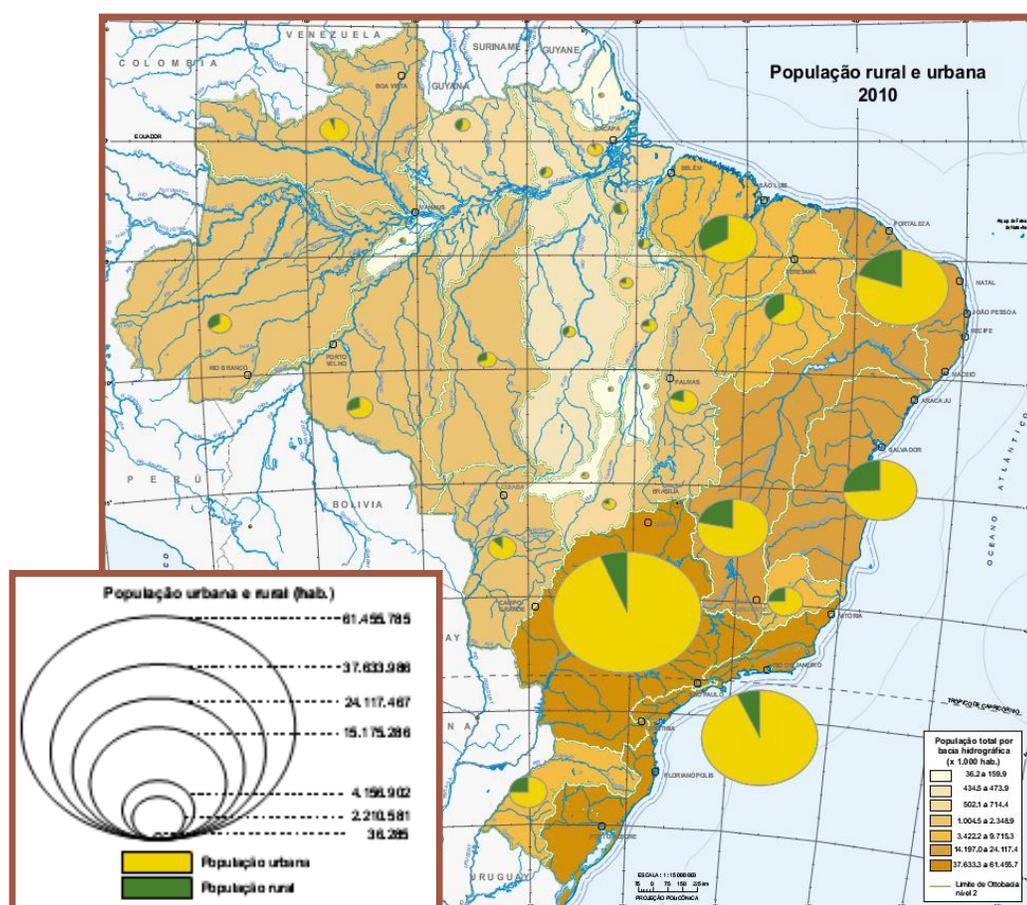


Nota: Doenças relacionadas com o saneamento ambiental inadequado (DRSAI): Diarreias, Febres Entéricas, Hepatite A, Dengue, Febre Amarela, Leishmanioses (tegumentar e visceral), Filariose Linfática, Malária, Doença de Chagas, Esquistossomose, Leptospirose, Tracoma, Conjuntivites, Micoses Superficiais, Helmintíases e Teníases.

Fonte: Instituto Brasileiro de Geografia e Estatística (IBGE)¹⁴.

- Símbolos graduados: Nem todo o mapa quantitativo é coroplético ou de áreas, mas podem ser aplicados também a outros elementos gráficos como pontos e linhas, desde que estabelecidas as regras aplicadas à simbologia. Na Figura 7.25 descreve a população urbana e rural segundo os símbolos graduados, em que a variação do tamanho do signo é proporcional à variável mensurada, indicada na legenda. Neste tipo de mapa, outros símbolos podem ser utilizados, como triângulos, quadrados, hexágonos etc.

Figura 7.25: Distribuição da população urbana e rural. Brasil, 2010.

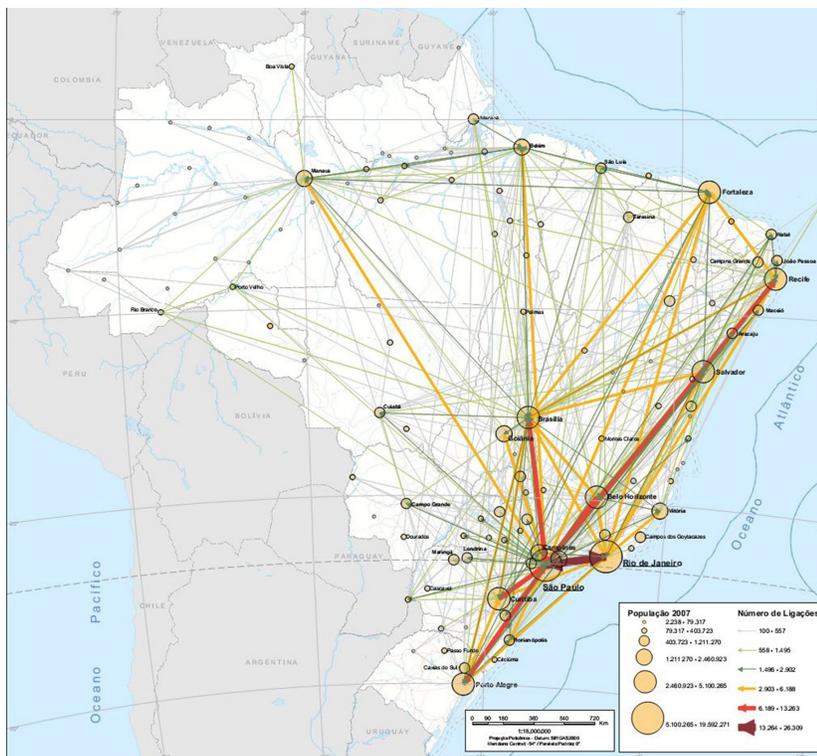


Fonte: Instituto Brasileiro de Geografia e Estatística (IBGE)¹⁴.

- Fluxo: Os mapas temáticos de fluxo também utilizam variáveis quantitativas com representações tipicamente de linhas, que simulam movimentos entre dois pontos ou dois polígonos,

nos quais estão presentes a direção do fluxo e a intensidade do indicador estudado é mostrado pela espessura, forma ou mesmo coloração das linhas. Sua aplicação é grande em diversas áreas que procuram analisar o fluxo ou influência entre unidades geográficas. Na Figura 7.26, a distribuição das conexões aéreas no Brasil está representada de forma bastante intuitiva.

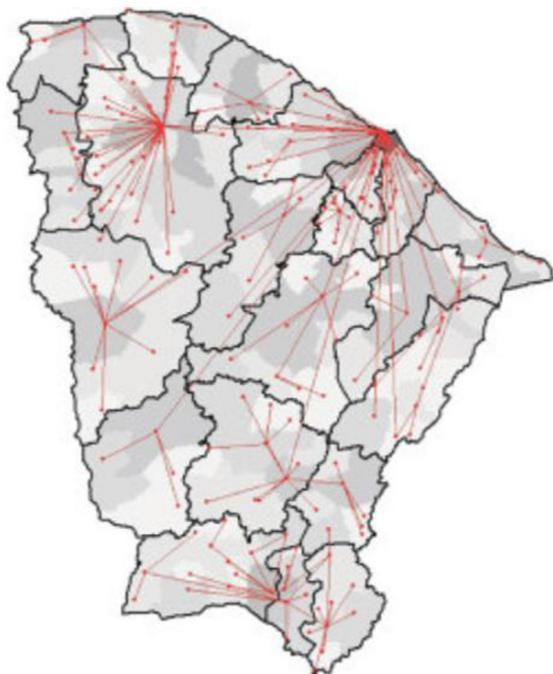
Figura 7.26: Conexões aéreas – Brasil, 2004.



Fonte: Instituto Brasileiro de Geografia e Estatística (IBGE)¹⁴.

O estudo dos fluxos é útil para tratar questões ligadas à regionalização da demanda e oferta de diversos serviços ligados à saúde, organização de pólos e referências e as distâncias envolvidas no atendimento. São utilizadas concomitantemente as informações de residência-ocorrência ou origem-destino e o resultado dos mapas de fluxo são setas cuja largura e tonalidade são proporcionais ao volume ou à distância de deslocamento. Na Figura 7.27 são mostrados os fluxos de internação, do local de residência ao município de atendimento no Estado do Ceará, onde são demonstradas algumas situações de extrapolação de polo regional por parte da população.

Figura 7.27: Fluxos dominantes de internação hospitalar, segundo municípios. Ceará, 2002.

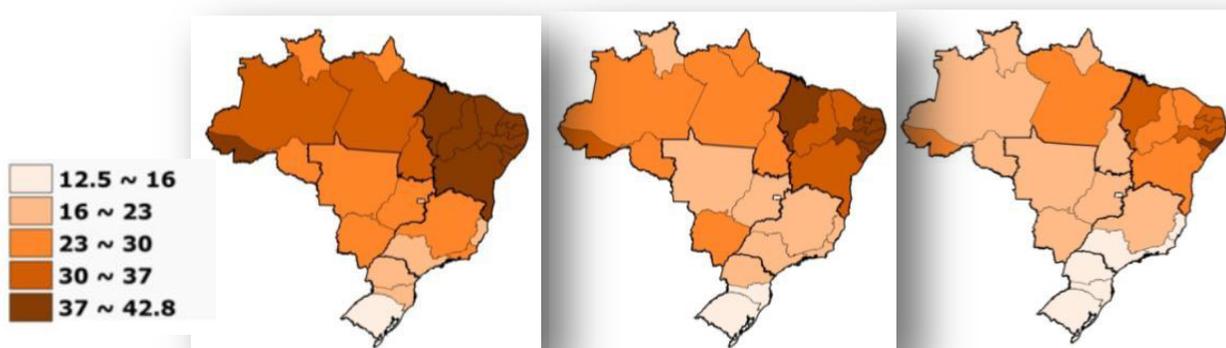


Fonte: Santos & Souza, 2006¹³.

c) Comparativos: A comparação dos mapas pode sugerir a dinâmica do indicador no tempo e no espaço, ou vários indicadores com conexão entre eles. Nesta metodologia, vários mapas são construídos utilizando os mesmos intervalos de classes e os mesmos padrões de cores, permitindo comparação visual direta, enfatizando os processos de tendência do indicador estudado.

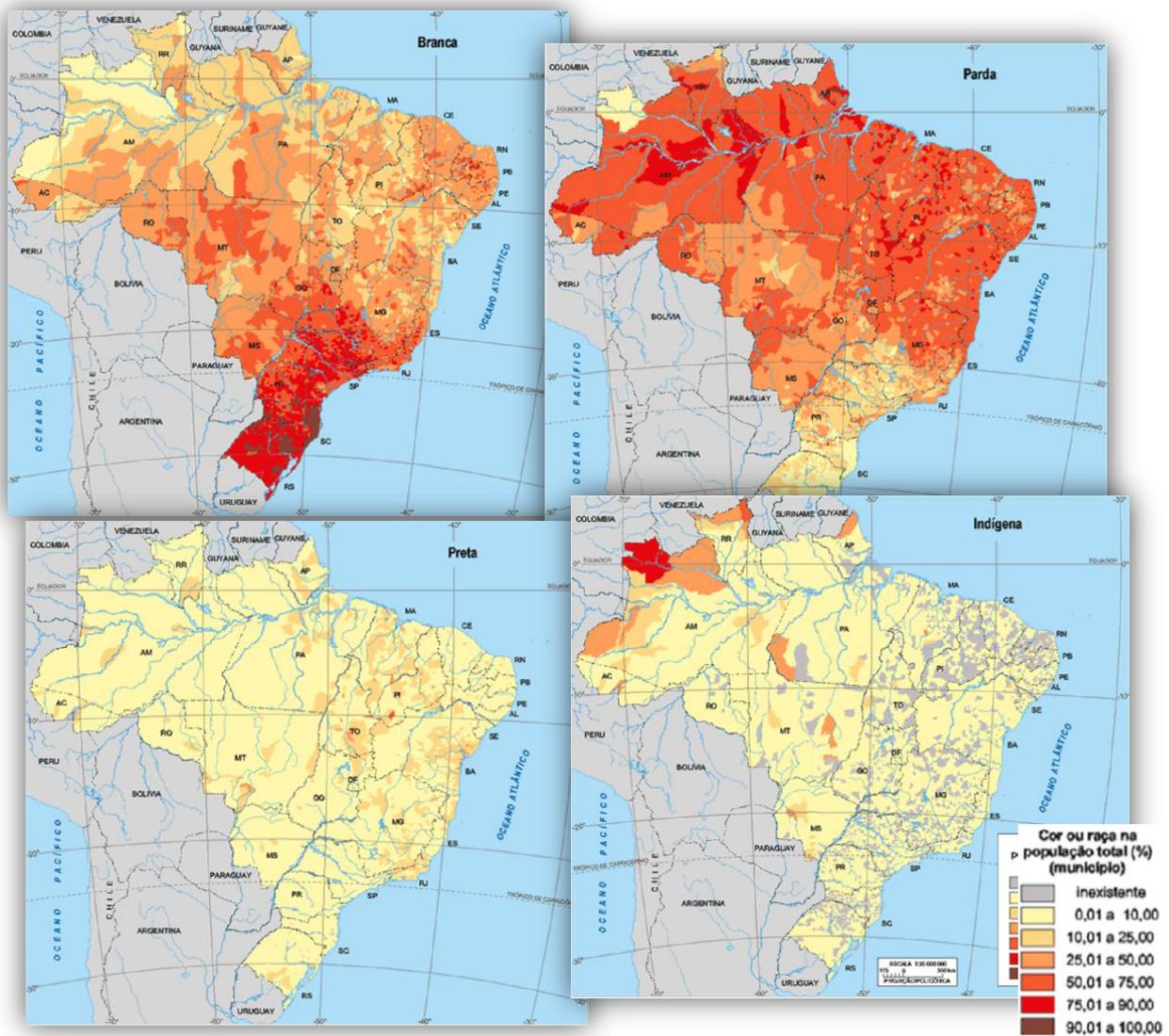
A Figura 7.28 mostra a involução dos coeficientes de mortalidade infantil em todas as unidades federadas brasileiras, porém, o Nordeste ainda concentra os maiores valores, comparados com as demais regiões no último ano estudado, enquanto que na Figura 7.29, a distribuição da população segundo cor ou raça, observada no censo 2000.

Figura 7.28: Distribuição do coeficiente de mortalidade infantil por unidades da Federação – Brasil, 1997, 2001 e 2006.



Fonte: RIPSAs (Rede Interagencial de Informações para a Saúde).

Figura 7.29: Distribuição da proporção população por cor e raça por município. Brasil, 2000.



Fonte: Instituto Brasileiro de Geografia e Estatística (IBGE)¹⁴.

Unidade 4

Análise explanatória de dados espaciais

Discutiremos um conjunto de abordagens comuns para a exploração de dados geográficos para o levantamento de hipóteses, sendo uma etapa posterior da visualização dos dados por mapas temáticos. Os conceitos da semiótica permanecem como procedimentos para apresentar um produto espacial intuitivo, porém agora o interesse é a aplicação de um conjunto de técnicas para tornar explícita e indicar características visualmente imperceptíveis no contexto espacial.

Como as abordagens dependem da primitiva gráfica (pontos, polígonos e linhas) que representa a variável no espaço, esta seção focará os polígonos (áreas) e pontos em separado.

1. A Análise Exploratória de Dados Espaciais

Trata-se de um conjunto de abordagens com o intuito de averiguar fenômenos de aglomerações com valores estatisticamente próximos, sinalizando possível existência de fatores espaciais significativos em contraposição de uma distribuição espacial aleatória, e que podem sugerir a forma e a dependência entre as observações.

Várias técnicas podem estar relacionadas, porém dois enfoques são importantes e complementares. O primeiro é a de identificação de situações atípicas ou *outliers* da distribuição, ou seja, trata-se de uma abordagem exploratória da estatística que já conhecemos. Outra abordagem é a busca de *clusters*, conhecido como observação de valores semelhantes em áreas conexas ou mesmo a ocorrência de eventos próximos entre si no espaço. Estas técnicas têm foco em conceito de autocorrelação espacial, e as formulações estatísticas julgam as ocorrências e influências entre os eventos.

O propósito da análise exploratória é a percepção de características relacionadas à distribuição espacial de eventos, agora de maneira explícita pelas distâncias obtidas por meio de uma matriz de vizinhança ou de relacionamento estabelecido previamente.

2. Áreas

Na Epidemiologia, o universo dos dados espaciais por área é uma abordagem ecológica, com vantagens no acesso aos indicadores disponibilizados por diversas instituições, já que o agregado das informações não permitem a identificação das pessoas.

Mas estudos como os ecológicos ainda trazem vantagens em alguns tipos de medidas de contexto e que são difíceis de associação (aqui tratando de exposição versus desfecho) com determinada situação de saúde a nível individual, como na influência da poluição, de renda, de acesso a serviços de água etc. A qualificação dos estudos ecológicos a partir da análise espacial é dada pela possibilidade de verificação da estrutura espacial como um todo, ou seja, pela oportunidade que o investigador tem de verificar se existe algum tipo de tendência a agrupamentos de risco na região estudada, a partir da representação da observação (cada uma das áreas) pelo valor do seu atributo (a variável de estudo).

Entre os cuidados a serem observados, alistam-se os relacionados aos estudos de agregados, como a falácia ecológica e os referentes à análise espacial, como o problema da unidade de área e a instabilidade decorrente dos pequenos números, que será discutido nos pontos seguintes.

a) Problema da unidade de área modificável: O termo “problema da unidade de área modificável”, conhecido como Modifiable areal unit problem (MAUP), descreve o fato da possibilidade de obtenção de resultados distintos para mesma população, por diferentes arranjos de agregação de unidades geográficas, desta forma, nas análises de dados agregados, os resultados podem estar associados apenas a efeitos de zoneamento¹⁵.

Pode ser difícil generalizar como um conjunto de dados em diferentes unidades espaciais afetados pelo MAUP e seus efeitos da escolha entre a escala de zonas não devem ser ignorados¹⁶. É possível o emprego de técnicas, como utilização de dados na menor desagregação

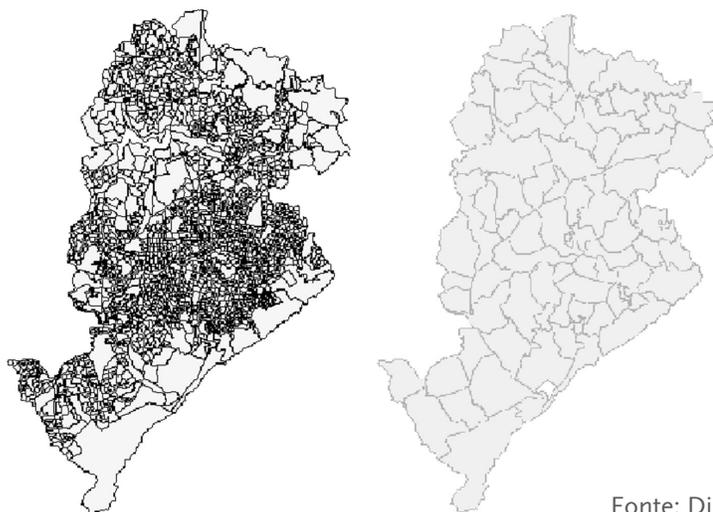
possível, metodologias de agrupamentos ou otimização combinatória para obter áreas mais agregadas, mas que preservem o fenômeno estudado da melhor forma possível¹⁷.

O MAUP é decorrente dos efeitos de escala e de zoneamento e condicionados pela forma e níveis de agregação dos indicadores. Os resultados são afetados pela magnitude de várias medidas associados às unidades, pelos coeficientes de autocorrelação espacial e os parâmetros nas modelagens estatísticas.

- Efeito de escala: É a variação nos resultados decorrentes do nível de agregação dos indicadores, com nítido aumento das correlações nos dados mais agregados, especialmente quando é observada autocorrelação no fenômeno estudado⁵. Na linguagem cartográfica, a passagem de uma escala grande (p.ex.: 1:1.000) para outra escala pequena (p.ex.: 1:100.000) perde-se conteúdo e especificidade devido a processos como generalização, simplificação ou eliminação de objetos geográficos, que facilita a visualização de informações de grandes áreas, mas surgem novas informações antes ignoradas¹⁸.

A Figura 7.30 mostra uma avaliação dos efeitos de escalas dos indicadores organizados a partir de agregados distintos¹⁷. A partir de dados do censo de 1991, indicadores em sua versão original (1998 observações) foram correlacionados com unidades de planejamento (UP, 80 observações), construídos a partir de agregados dos setores censitários. Os resultados apontaram que 77% das correlações obtiveram maior intensidade quando a agregação era por UP, ao passo que 4% apresentaram menor intensidade, além das alterações de sinais, ou seja, o comportamento entre as variáveis era positivo em um agregado, em outro, negativo.

Figura 7.30: Malhas dos limites dos setores censitários do ano de 1991 (à esquerda) e unidades de Planejamento (à direita) para o município de Belo Horizonte.

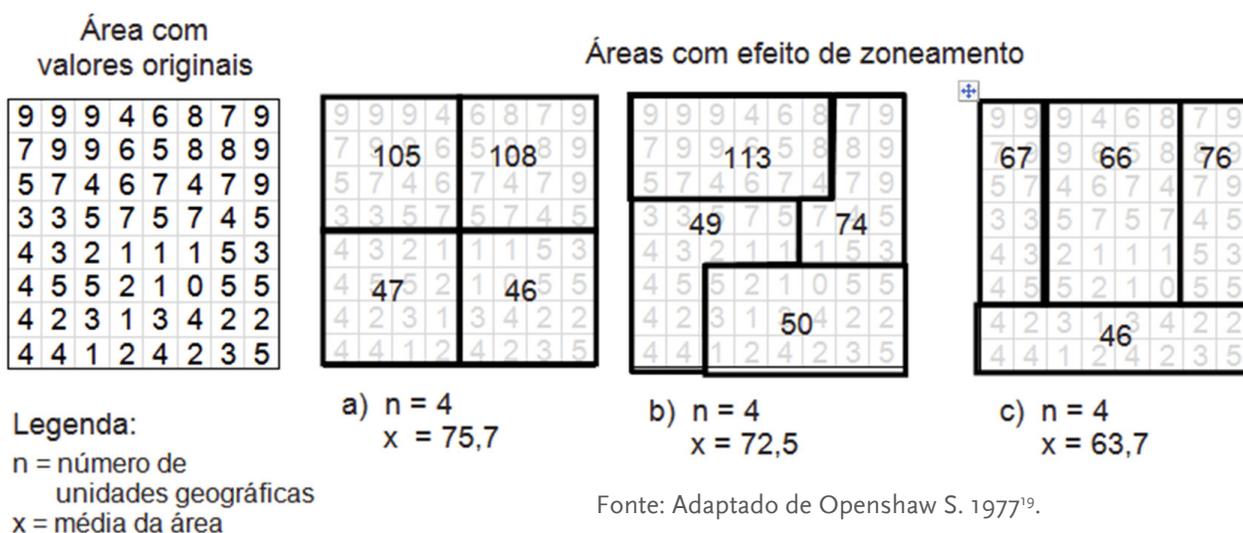


Fonte: Dias et al., 2002¹⁷.

- Efeito de zoneamento: É a variação nos resultados devido à forma como são definidas as fronteiras das áreas, ou seja, diferentes efeitos são observados apenas alterando as configurações entre os limites das zonas, diferente do efeito de escala, onde o efeito é observado com a agregação das áreas.

Um exercício para ilustrar o zoneamento está descrito na Figura 7.31, em que determinada região contendo um conjunto de valores, recebeu três recortes distintos, todos com quatro áreas. Observamos que todos os recortes possuem médias distintas.

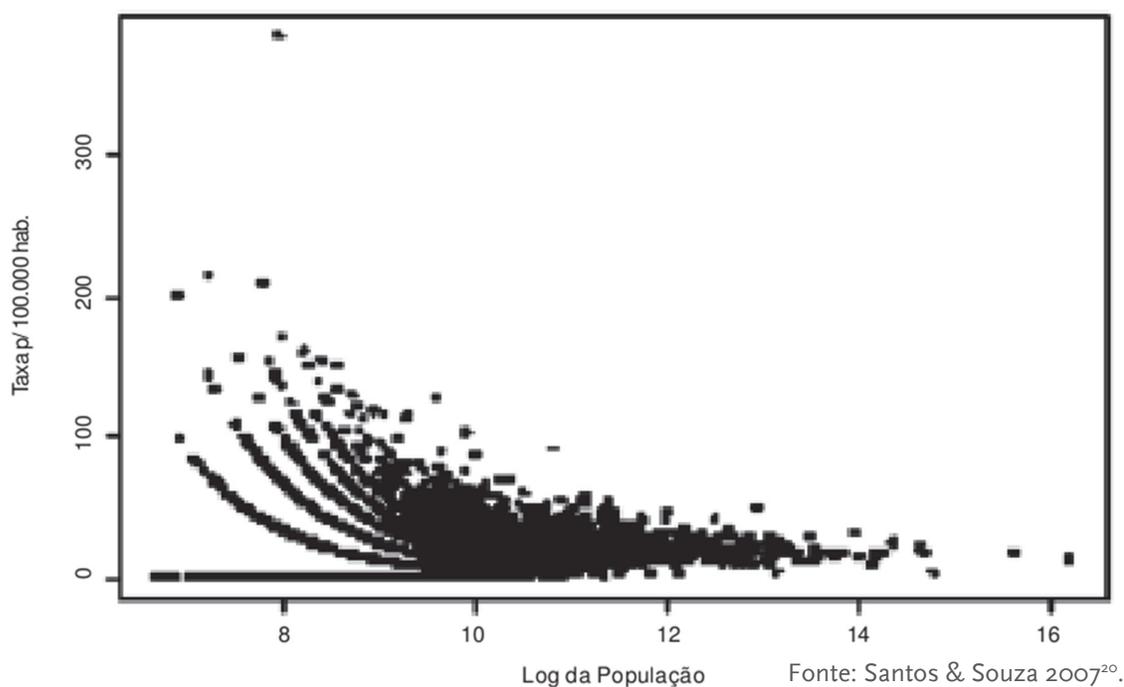
Figura 7.31: Efeito de zoneamento em uma área de estudo.



dia de dois óbitos infantis. A diferença entre a notificação de um óbito infantil a mais, em determinado ano por um fator circunstancial, acarreta uma diferença do Coeficiente de Mortalidade Infantil (CMI) em 50%.

Outro exemplo na Figura 7.32²⁰ apresenta o resultado da dispersão das taxas de acidentes de transportes e o logaritmo natural da população para os municípios do Brasil, no ano de 2004. A representação em forma de funil mostra a grande variabilidade encontrada nos municípios com populações em torno de 3 mil habitantes ($\log N=8$).

Figura 7.32: Taxa de mortalidade por Acidentes de Transporte e Logaritmo da População – Brasil, 2004.



Algumas técnicas podem ser aplicadas para lidar com problemas da instabilidade de dados agregados de forma direta, como a agregação temporal, a agregação de áreas, e de forma indireta, abordada adiante, a média móvel espacial e a suavização bayesiana.

c) *Técnicas de análise exploratória de dados espaciais:* As técnicas de análise exploratória por área são sensíveis ao tipo de distribuição, à presença de valores extremados e à ausência de es-

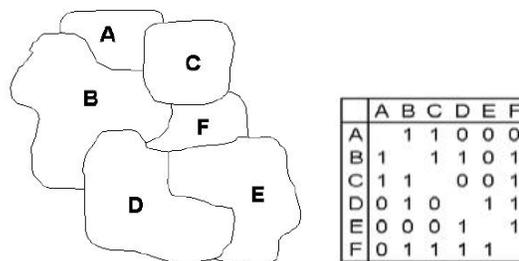
tacionariedade. Como em qualquer abordagem estatística, inicialmente o fenômeno estudado é descrito por medidas de tendência central e de dispersão, além de gráficos para averiguar a sua distribuição e além da criação de mapas temáticos para a visualização espacial.

Conhecendo a variável de estudo e satisfazendo o princípio de normalidade, no caso das taxas, ou de Poisson, por contagens de eventos, é possível evidenciar a forma de como os valores dos atributos estão correlacionados no espaço, estimando a magnitude da autocorrelação espacial entre as áreas.

- Dependência espacial: Trata-se de um conceito importante no entendimento do arranjo espacial, em que se procura evidências de que o valor da variável em uma unidade espacial seja parcialmente função do valor da mesma variável em unidades vizinhas¹⁶. A autocorrelação espacial é a da mensuração da dependência espacial, termo proveniente do conceito estatístico de correlação usado para medir a relação entre duas variáveis aleatórias, que no caso, o valor da variável do atributo estudado e a sua localização espacial, ou seja, o quanto é dependente o atributo da sua localização.

- Matriz de proximidade espacial: Os cálculos de exploração espacial são realizados a partir de uma tabela de dados denominada de matriz de proximidade espacial ou de vizinhança. Esta tabela armazena informações sobre características do arranjo espacial, segundo as investigações empíricas de configurações locais sobre conexões de influência entre as unidades geográficas. É possível montar uma matriz de diferentes maneiras, como os vizinhos (ou de primeira ordem), vizinhos dos vizinhos (segunda ordem), distâncias euclidianas, por acessibilidade, proporção de fronteiras etc.

Figura 7.33: Exemplo de divisão zonal com matriz de proximidade de primeira ordem.



Fonte: Adaptação de Carvalho et al., 2004¹⁶.

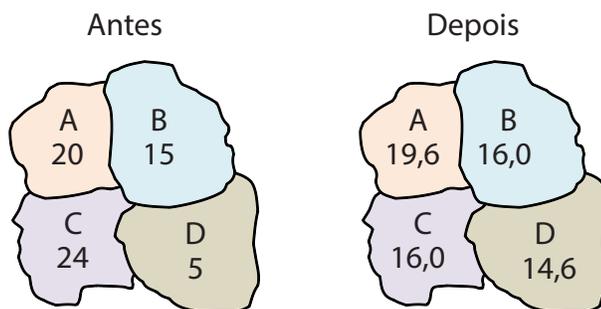
- Média móvel espacial: Uma forma prática de exploração da variação espacial em toda a área fornecendo uma visão geral da tendência espacial pode ser realizada a partir do cálculo da média dos valores dos vizinhos ou outro arranjo espacial dada pela matriz de proximidade. Para cada i -ésima área, é calculada a média móvel a partir do atributo das áreas vizinhas baseada na matriz de proximidade espacial, expressa na fórmula a seguir:

$$\hat{\mu}_i = \frac{\sum_{j=1}^n w_{ij} y_j}{\sum_{j=1}^n w_{ij}} \quad i = 1, 2, \dots, n$$

onde:

- W_{ij} é a matriz de proximidade.
- y_i é o valor do atributo em cada área.
- n é o número de polígonos (áreas).

Figura 7.34: Resultado do efeito de uma média móvel espacial.

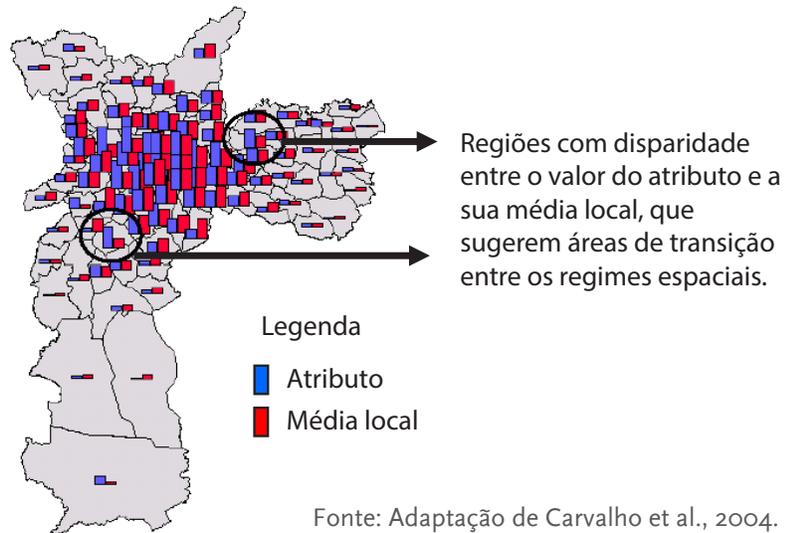


Fonte: Adaptado de Carvalho et al., 2004.

Outra forma prática de apresentar a média móvel espacial é por meio de um gráfico de barras, de forma a evidenciar padrões resultantes da comparação entre os valores do atributo e a sua média local, como apresentada na Figura 7.35.

• Indicadores globais de autocorrelação espacial: Possuem a função de caracterizar a dependência espacial em toda a região de estudo, fornecendo uma medida de resumo que estima o quanto é dependente um valor de atributo em dada localização com os seus vizinhos, ou seja, a autocorrelação espacial. Um coeficiente de autocorrelação espacial descreve um conjunto de dados que está ordenado em uma sequência espacial, na qual os testes mais utilizados nesta categoria são o Índice Global de Moran e o Índice de Geary.

Figura 7.35: Indicador de segregação espacial e valores da média móvel espacial. São Paulo.



Fonte: Adaptação de Carvalho et al., 2004.

– Índice Global de Moran: O Índice de Moran fornece uma medição genérica da correlação espacial que o conjunto de dados na região possui, viabilizando a percepção da estrutura da dependência espacial do fenômeno estudado. Como teste de hipótese, a hipótese nula é de independência espacial, na qual a estatística apresenta valores próximos a zero, variando de -1 a 1 nos quais os valores extremos significam respectivamente autocorrelação negativa e autocorrelação positiva. O índice I de Moran global é dado por:

$$C = \frac{n - 1}{2 \sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (y_i - y_j)^2}{\sum_i (y_i - \bar{y})^2}$$

Onde:

- n: é a quantidade de áreas;
- y_i : o valor do atributo na área i;
- y_j : o valor do atributo na área j;
- \bar{y} : o valor médio do atributo na área de estudo;
- w_{ij} : peso da a matriz de proximidade espacial atribuído conforme a relação topológica entre os locais i e j.

– Índice Global de Geary: A estatística c de Geary é outra medida de autocorrelação espacial global. Da mesma forma que o I de Moran testa a aleatoriedade espacial, porém possui interpretação diferente, pois o valor resultante de sua estatística situa-se entre 0 e 2, ao passo que sua média teórica é 1. Valores menores que seu valor esperado, isto é, entre 0 e 1 indicam autocorrelação espacial positiva, enquanto que os maiores que 1 indicam autocorrelação espacial negativa. O índice c de Geary global é dado por:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Onde:

- n : é a quantidade de áreas;
- y_i : o valor do atributo na área i ;
- y_j : o valor do atributo na área j ;
- \bar{y} : o valor médio do atributo na área de estudo;
- w_{ij} : peso da matriz de proximidade espacial atribuído conforme a relação topológica entre os locais i e j .

• Indicadores de autocorrelação espacial local (Lisa): Objetivam captar padrões de associação local úteis para identificar áreas de diferentes regimes de associação, permitindo a investigação com maior detalhe da região de estudo, pois os resultados do teste local e sua significância são atribuídos a cada uma das observações, e a visualização gráfica por meio de mapas temáticos. Os Lisas são decompostos dos indicadores globais, porém com as contribuições individualizadas, e ao contrário dos testes globais, indicam aglomerados ou *clusters* e regiões de não estacionariedade.

Os testes mais difundidos são o Índice Local de Moran (I_i) e as estatísticas G_i e G^*i , cujos resultados apresentam interpretações distintas. No Moran, a significância do teste em cada observação indica alta probabilidade de haver associação espacial local, tanto em regiões com altos valores quanto as observações que apresentam baixa significância, podendo ser interpretadas como áreas de não estacionariedade, com grande flutuação espacial do atributo. As estatísticas G_i e G^*i , os valores altos e significativos indicam locais de associação espacial em regiões de alto valor do atributo e valores baixos e significativos indicam associação em regiões de baixo valor.

Índice Local de Moran (I_i)

$$I_i = \frac{(y_i - \bar{y}) \sum_{j=1}^n w_{ij} (y_j - \bar{y})}{\frac{\sum_{i=1}^n (y_j - \bar{y})^2}{n}}$$

Onde:

n : é a quantidade de áreas;

y_i : valor do atributo de interesse;

y_j : o valor do atributo vizinha a_j ;

\bar{y} : o valor médio do atributo na área de estudo;

w_{ij} : matriz de proximidade espacial atribuído conforme a topologia dos locais i e j .

Estatística G_i

$$G_i(d) = \frac{\sum_{j=1}^n w_{ij}(d) x_j}{\sum_{j=1}^n x_j}, j \neq i$$

$$G_i^*(d) = \frac{\sum_{j=1}^n w_{ij}(d) x_j}{\sum_{i=1}^n x_j}, j \neq i$$

Onde:

n : é a quantidade de áreas;

x_i : valores dos atributos considerados na área i ;

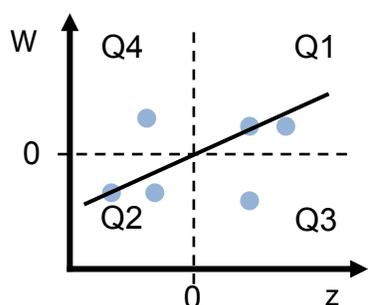
x_j : valores dos atributos considerados na área j ;

d : distância entre pontos

w_{ij} : matriz de proximidade espacial atribuído conforme a topologia dos locais i e j .

Além dos mapas de estatísticas locais e suas significâncias, pode ser utilizada outra ferramenta de exploração de padrões espaciais, conhecida como diagrama de espalhamento de Moran Moran Scatterplot Map, que não afere a autocorrelação espacial diretamente, mas oferece uma percepção gráfica dos padrões da distribuição espacial. O grau de similaridade entre

Figura 7.36: Diagrama de Espalhamento de Moran.



Fonte: Adaptado de Santos & Souza 2007.

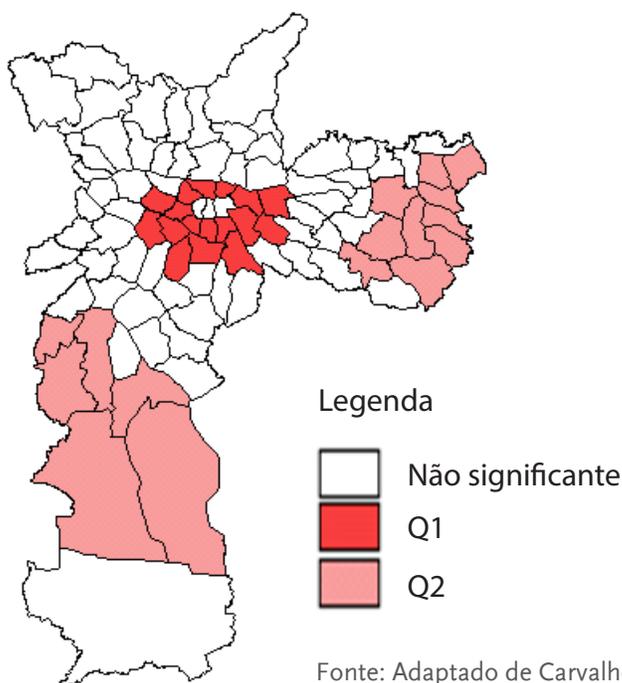
vizinhos é dado pela dispersão entre o indicador normalizado e a média dos seus vizinhos, e analisado segundo quadrantes de regimes espaciais:

- Regiões com associação espacial positiva: Q1 (valor positivo e médias positivas ou alto-alto) e Q2 (valor negativo e médias negativas ou baixo-baixo), e indicam padrões espaciais.

- Regiões com associação espacial negativa: Q3 (valor positivo e médias negativas ou alto-baixo) e Q4 (valor negativo e médias positivas ou baixo-alto), e indicam zonas de transição ou de não estacionariedade.

O diagrama de espalhamento de Moran pode ser apresentado no formato de mapa temático, denominado de *Moran map*, nas quais as áreas consideradas significantes (>95%) assumem as posições dos quadrantes, sendo normalmente caracterizado com intensidade de cores fortes nos quadrantes Q1 e Q2, e mais amenas para o Q3 e Q4, como mostrado na Figura 7.36.

Figura 7.37: Moran map do indicador de segregação espacial – São Paulo.



Fonte: Adaptado de Carvalho et al., 2004.

Leitura recomendada

Almeida et al. Análise espacial da dengue e o contexto socioeconômico no município do Rio de Janeiro. Rev. Saúde Pública, São Paulo, 2009;43(4),666-73. (Almeida et al. 2009)

Gomes et al. Landscape risk factors for attacks of vampire bats on cattle in Sao Paulo, Brazil. Prev Vet Med, 2010;93(2-3),139-46. (Gomes et al. 2010)

Sales et al . Análise espacial da tuberculose infantil no estado do Espírito Santo, 2000 a 2007. Rev. Soc. Bras. Med. Trop, 2010;43(4),435-9. (Sales et al. 2010)

Martins-Melo et al. Mortality of Chagas' disease in Brazil: spatial patterns and definition of high-risk areas. Trop Med Int Health, 2012;17(9),1066-75. (Martins-Melo et al. 2012)

d) *Técnicas de estimação espacial de indicadores*: Observamos que a utilização de estimadores de risco tradicional em Epidemiologia, como os coeficientes brutos, estão muito suscetíveis a flutuações nas regiões com pequenas populações, por ocorrência casual de eventos aleatórios ou por sua subnotificação. Análises baseadas diretamente nessas estimativas brutas são de interpretação direta e, em muitos casos, criam falsas conclusões porque a diferença de apenas um caso poderá criar um coeficiente bastante expressivo em relação às demais observações. Esta situação é corriqueiramente encontrada nos estudos de distribuição espacial por regiões políticas ou administrativas, cujas delimitações apresentam áreas de tamanhos variados, e aquelas de maior área geralmente possuem menor contingente populacional, e por conseguinte, quando apresentam taxas elevadas, desviam a atenção nos mapas temáticos.

As taxas são menos instáveis, mais interpretativas e informativas. Várias técnicas podem ser utilizadas para minimizar a flutuação aleatória. Entre as mais utilizadas estão as médias trienais, a agregação de áreas e as médias móveis espaciais, descritas anteriormente. Mas se o intuito é assumir que o risco de determinada área é semelhante ao dos seus vizinhos, ou seja, estão autocorrelacionados, é possível utilizar a Estimativa Bayesiana Empírica para criar uma nova estimativa.

O estimador Bayesiano empírico calcula uma nova estimativa para cada área que consiste da média ponderada que leva em conta os eventos e a população da correspondente área e a dos seus vizinhos. Para a ponderação, os pesos são inversamente proporcionais à população de cada uma das áreas. A estimativa pode ser aplicada em toda a região de estudo, sendo

denominada de *Estimativa bayesiana empírica global*, ou de forma desagregada em pequenas áreas, denominada de *Estimativa bayesiana empírica local*, calculada da seguinte forma:

$$\hat{\theta}_i = w_i r_i + (1 - w_i) \mu_i$$

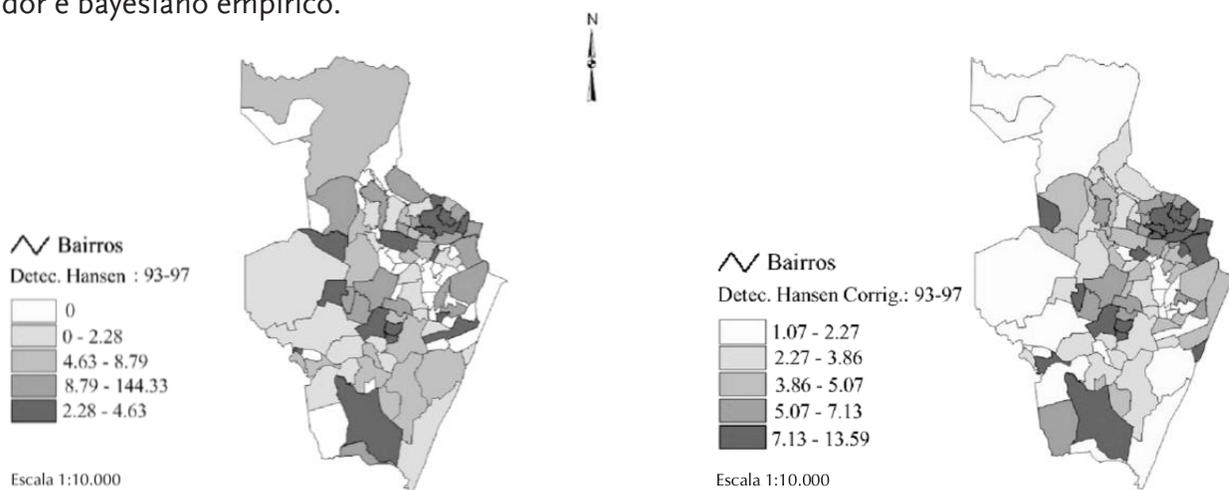
Onde:

r_i : Coeficiente observado

μ_i : média

w_i : fator de ponderação

Figura 7.38: Exemplo de estimador bayesiano empírico. À esquerda a representação de coeficientes brutos de detecção de hanseníase, e à direita, o mesmo coeficiente suavizado por estimador bayesiano empírico.



Fonte: Souza et al., 2001.

Leitura recomendada

Souza et al. Aplicação de modelo bayesiano empírico na análise espacial da ocorrência de hanseníase. *Rev Saúde Pública*, 2001;35(5),474-80. (Souza et al. 2001)

Imbiriba et al. Desigualdade social, crescimento urbano e hanseníase em Manaus: abordagem espacial. *Rev. Saúde Pública*, 2009;43(4),656-65. (Imbiriba et al. 2009)

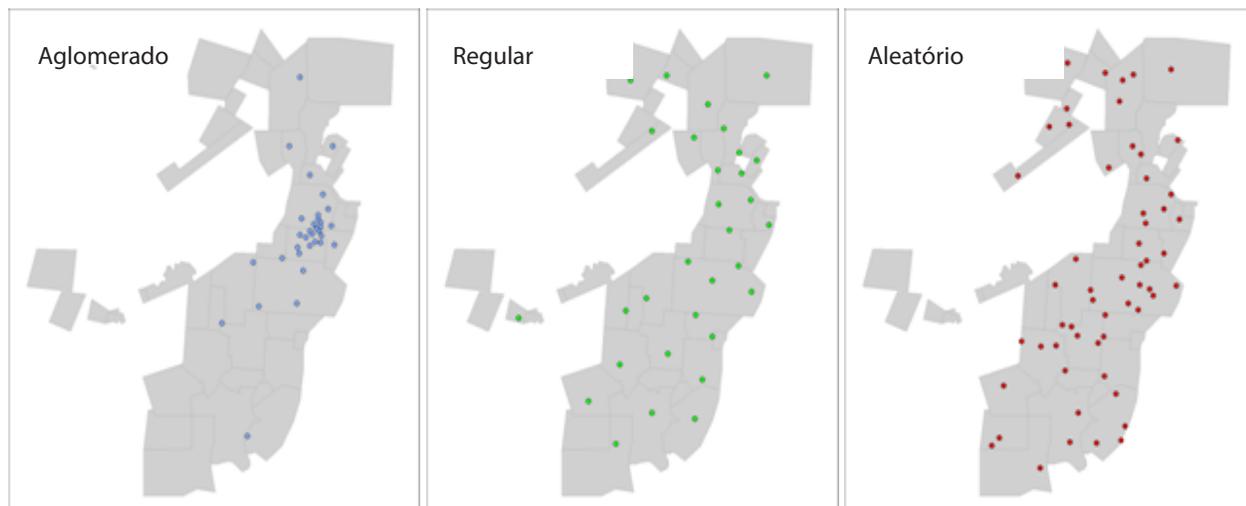
Caram et al. Distribuição espaço-temporal dos candidatos à doação de sangue da Fundação Hemominas, Belo Horizonte, Minas Gerais, Brasil, nos anos de 1994 e 2004. *Cad. Saúde Pública*, 2010;26,229-39.

3. Pontos

Uma das formas mais tradicionais de se visualizar e analisar problemas de saúde em mapas é a representação de eventos de saúde em forma de pontos, denominados de padrões pontuais. Nestes casos, cada evento é associado a um par de coordenadas, cuja disposição pode sugerir um padrão espacial sistemático¹⁵.

A medida implícita que se está utilizando para avaliar a distribuição de pontos é a distância entre eles em relação à área de estudo, ou seja, o padrão de distribuição de pontos é diferente ao aleatório ou regular. O padrão de pontos aglomerados clusters apresenta menor distância média entre os pontos, o que indica uma concentração de pontos no espaço, mas um valor de desvio-padrão maior. Os padrões regular e aleatório apresentam grande distância média entre os pontos, porém um desvio-padrão pequeno indica homogeneidade da distribuição no regular, e no aleatório, desvio-padrão grande.

Figura 7.39: Padrões espaciais de dados na forma de pontos.



Fonte: Santos & Souza, 2007.

Em geral, se estes eventos estiverem concentrados no espaço é bem provável que exista uma determinação espacial deste evento e que encontrando os locais com maior concentração, temos uma pista para identificar fatores ambientais, sociais ou ligados à própria assistência à saúde que podem ser objeto de ação da vigilância em Saúde. Existe grande número de procedimentos de exploração espacial de processos pontuais baseados em gráficos, podemos citar a

função k , a função cumulativa ($G(w)$) também conhecida como método do vizinho mais próximo, além de análise baseada na distância entre os pontos como a de aglomerados hierárquicos, que é o método de aglomerados hierárquicos de vizinhos mais próximos *Nearest Neighbor Hierarchical Clustering*. Porém vamos descrever a técnica de Kernel, por ser uma das mais utilizadas para exploração de eventos pontuais na Saúde.

a) Estimador de intensidade (*Kernel estimation*): O estimador de intensidade é muito útil para mostrar a distribuição e o comportamento dos pontos em toda a região de estudo e indicar a ocorrência de *clusters* (como sugestão de alguma dependência espacial), porém não é um teste comprovatório e nem indica as suas localizações. Permite explorar e mostrar, de forma não parametrizada, o padrão formado em uma superfície de densidade para a identificação visual de áreas quentes (*hot spot*) de maior concentração de pontos em relação a outras áreas de estudo¹⁶. A sua técnica é baseada em uma função bidimensional para interpolação dos eventos especializados, resultando em uma superfície proporcional ao das amostras por unidade de área a partir da contagem dos pontos dentro de uma região de influência, ponderada pela distância. Para a aplicação, é necessário que o analista defina dois parâmetros: a função de densidade k e o raio de influência a ser aplicada em:

$$\lambda(s) = \frac{1}{\delta_r(s)} \sum_{i=1}^n \frac{1}{\tau^2} k \left(\frac{s - s_j}{\tau} \right)$$

Onde:

$\lambda(s)$: intensidade em s ;

s : localização arbitrária;

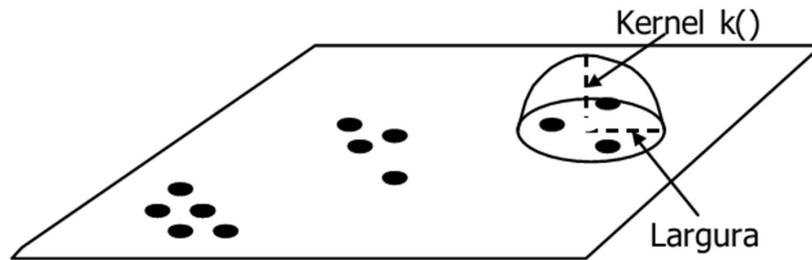
τ : raio de influência;

k : função densidade de probabilidade bivariada;

$\delta_r(s)$: volume sob o *kernel* centrado em s .

A função de densidade possui propriedades de suavização espacial do fenômeno estudado ao tempo em que promove interpolação dos valores em subáreas que não possui eventos observados. As funções mais comuns são a quártica, a gaussiana (ou normal), a triangular, a exponencial negativa e a uniforme. As quatro primeiras funções promovem uma ponderação de distâncias entre os eventos com maior peso naqueles mais próximos do centro do raio

Figura 7.40: Estimador de intensidade de distribuição dos pontos.

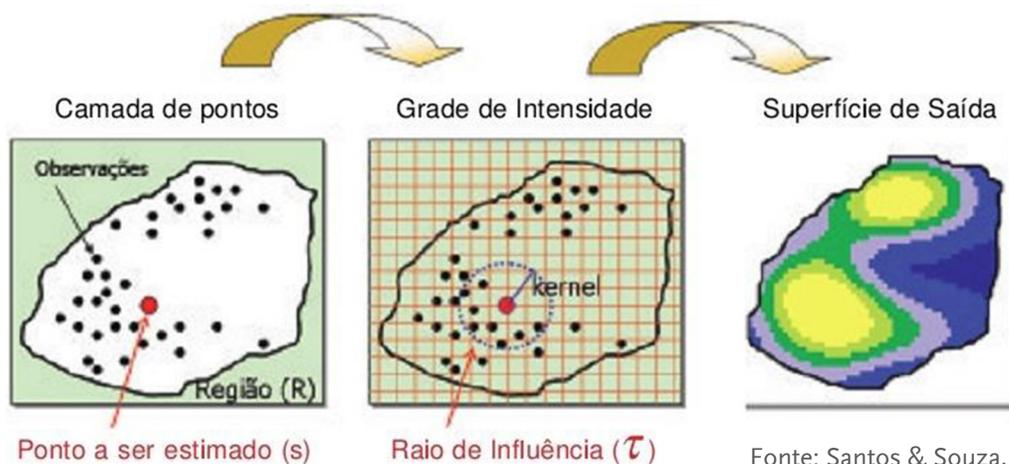


Fonte: Carvalho & Câmara, 2002.

de busca que os mais afastados, porém com diferenças na ponderação, sendo mais gradual na quártica e mais acentuada na exponencial negativa, enquanto que na uniforme, os pontos dentro da largura da banda são ponderados igualmente, sem depender da distância. O *kernel* quártico apresenta melhores soluções como suavização, devido a sua simetria à origem.

O raio de influência define a área centrada no ponto de estimação s que indica quantos eventos s_j contribuem para a intensidade λ , portanto define a vizinhança a ser interpolada e controla a superfície suavizada a ser criada, de forma que um raio muito pequeno suaviza pouco e produz uma área descontinuada, em oposição ao raio muito grande, que produzirá grande suavização. Uma alternativa é a adoção do kernel adaptativo, de forma que o raio de influência sempre tenha uma quantidade mínima de pontos necessários ao algoritmo, com vantagem na adaptação constante para a estimativa de toda a área de estudo, ou seja, o raio passa a ser variável, inversamente a concentração de pontos nas subáreas de estudo.

Figura 7.41: Passos para o cálculo de intensidade de pontos segundo a técnica de Kernel.



Fonte: Santos & Souza, 2007.

b) *Razão de intensidade (kernel dual)*: Os métodos com base na presença de casos como o mapa do estimador de intensidade do fenômeno poderá indicar de forma errônea as áreas de interesse, devido à falsa impressão da real exposição ao fenômeno estudado. Em situações críticas, o estimador é bastante influenciado pela distribuição heterogênea da população delimitando áreas de adensamento e dispersão populacional, em que se espera maior e menor ocorrência do evento estudado, respectivamente. Uma alternativa para relativizar o fator populacional é a utilização da razão da intensidade, estimando duas densidades, formas pela população e pelos eventos, criando uma superfície de risco populacional, tal como mostrado a seguir:

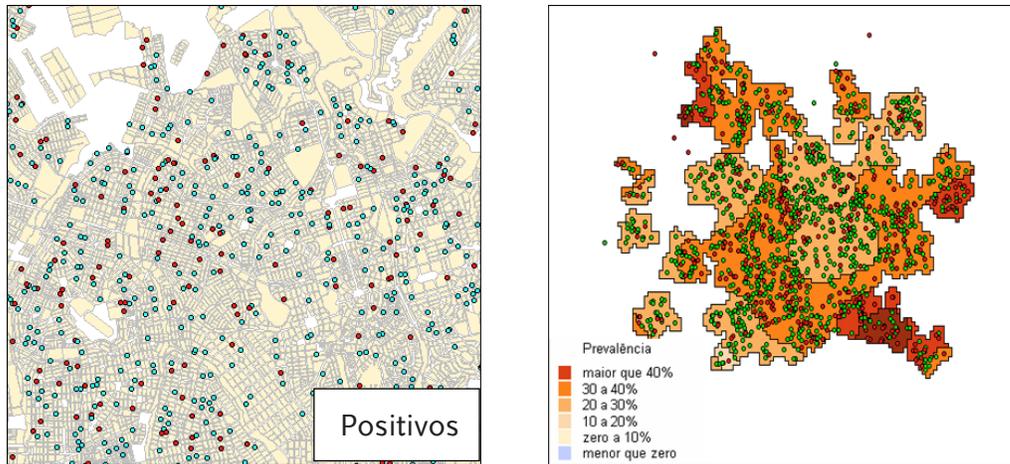
$$\frac{\text{Densidade estimada para os eventos}}{\text{Densidade estimada para a população}}$$

A densidade de população é obtida a partir de dados de área, como setores censitários, municípios etc., portanto trata-se de uma estimativa obtida de forma diferente dos eventos no formato padrões pontuais. Nestes casos, o estimador para dados de área calcula a intensidade do evento para o centroide da área de interesse usando o atributo numérico para ponderar a superfície populacional a partir do raio de busca.

É possível também utilizar um atributo de chance de um evento ocorrer como em uma pesquisa de base territorial na qual a resposta é binomial, como nos casos e controles ou inquéritos. A Figura 7.42 mostra o resultado de uma pesquisa de soroprevalência para dengue em Goiânia, com as categorias: positivos e negativos.

Com a técnica de *kernel dual*, na Figura 7.42, é possível a obtenção de um gradiente de risco, mostrando áreas com menor e maior prevalência, valor correspondente à taxa de prevalência suavizada no espaço, resultado da razão entre o número de casos (pontos vermelhos) pelo número total de amostras (pontos vermelhos e verdes). Nas piores áreas, a prevalência chegou a mais de 40%.

Figura 7.42: Distribuição do resultado de inquérito de soroprevalência de dengue – Goiânia, 2004. À esquerda, pontos amostrais e o resultado da técnica *kernel dual*.



Fonte: Siqueira et al., 2004.

Leitura recomendada

Siqueira et al. Household survey of dengue infection in Central Brazil: spatial point pattern analysis and risk factors assessment. *Am J Trop Med Hyg*, 2004;71(5),646-51.

Cabral APS, Souza WV. Serviço de Atendimento Móvel de Urgência (SAMU): análise da demanda e sua distribuição espacial em uma cidade do Nordeste brasileiro. *Rev Bras Epidemiol*, 2008;11(4),530-40.

Rodrigues PCO et al. Distribuição espacial das internações por asma em idosos na Amazônia Brasileira. *Rev Bras Epidemiol*, 2010;13(3),1-10.

Silva MBTD et al. Esporotricose urbana: epidemia negligenciada no Rio de Janeiro, Brasil. *Cad Saúde Pública*, 2012;28(10),1867-80.

Referências

- 1 – Hippocrates (2006) On airs, waters, and places (classics revisited – 400 BCE), *Hygeia*, 2(3), 1-12.
- 2 – Werneck, G. L. (2008) Georeferenced data in epidemiologic research, *Ciência & Saúde Coletiva*, 13(6), 1753-1766.
- 3 – Ferreira, M. U. (1991) Epidemiologia e geografia: o complexo patogênico de Max. Sorre, *Cad Saude Publica*, 7(3), 301-309.
- 4 – Snow, J. (1999) *Sobre a maneira de transmissão do cólera*. São Paulo. Rio de Janeiro. Editora Hucitec. ABRASCO.
- 5 – Bailey, T. C. & Gatrell, A. C. (1995) *Interactive spatial data analysis*. Longman Scientific & Technical Essex.
- 6 – Ramalho, W. M., Skaba, D. A., Barcellos, C., Pina, M. F. & Magalhães, M. (2007) Conceitos e arquitetura de sistemas de informações geográficas, in: M. d. Saúde (Ed) *Sistemas de informações geográficas e análise espacial na saúde pública. Capacitação e Atualização em Geoprocessamento em Saúde*. Brasília, 148.
- 7 – Carvalho, M. S., Pina, M. d. F. d. & Santos, S. M. d. (2000) *Conceitos básicos de sistemas de informação geográfica e cartografia aplicados à saúde*. Brasília. Organização Panamericana da Saúde. Ministério da Saúde. Rede Internacional de Informações para a Saúde, 124.
- 8 – Santos S, Barcellos C. (2006) *Abordagens espaciais na saúde pública*. 1. ed. Brasília: Ministério da Saúde. v. 1. 136pgs.
- 9 – Francisco, C., Guimarães, L., Silva, L. & Vieira, L. (2007) *Estudo dirigido em SIG*. Disponível em: <http://www.professores.uff.br/cristiane/Estudodirigido/Index.htm> (acessado em 10 Out 2012).
- 10 – NOAA (2013) *Geographic Information System*. Disponível em: <http://www.ncddc.noaa.gov/> (acessado em 01 Fev 2013).

- 11 – Pina MF, Magalhães M, Oliveira E, Skaba D, Barcellos C. (2006) Análise de dados espaciais, In: Santos S, Barcellos C, (2006) Abordagens espaciais na Saúde Pública, Série Capacitação e actualização em geoprocessamento em saúde, Ministério da Saúde do Brasil. Brasília.
- 12 – MacEachren, A. M. (1986) Map use and map making education: attention to sources of geographic information, *Cartogr J*, 23(2), 115-122.
- 13 – PNUD (2003) Atlas do desenvolvimento humano no Brasil. Disponível em: http://www.pnud.org.br/IDH/Atlas2003.aspx?indiceAccordion=1&li=li_Atlas2003.
- 14 – Sluter, C. R. (2005) Cartografia geral, Apostila didática. (Curitiba, Departamento de Geomática. Universidade Federal do Paraná).
- 15 – Santos, S. M. & Souza, W. V. (Eds) (2006) Abordagens espaciais na saúde pública Brasília. Ministério da Saúde.
- 16 – Instituto Brasileiro de Geografia e Estatística, IBGE (2012) Disponível em: <http://www.ibge.gov.br/home>. (Acessado em 07-08-2013)
- 17 – Openshaw, S. & Taylor, P. J. (1979) A million or so correlation coefficients: three experiments on the modifiable areal unit problem, in: N. Wrigley (Ed) *Statistical Applications in Spatial Sciences*. London. Pion, 127–144.
- 18 – Carvalho, M., Câmara, G., Cruz, O., Correa, V., Druck, S. & Monteiro, A. (2004) Análise de dados de área, Druck S, Carvalho MS, Câmara G, Monteiro AVM, organizadores. Análise especial de dados geográficos. Brasília: Empresa Brasileira de Pesquisa Agropecuária, 157-209.
- 19 – Dias, T. d. L., Oliveira, M. P. G., Câmara, G. & Carvalho, M. S. (2002) Problemas de escala e a relação área-indivíduo em análise espacial de dados censitários, *iP*, 4(1), 89-104.

- 20 – Barcellos, C., Lammerhirt, C. B., Almeida, M. A. B. d. & Santos, E. d. (2003) Distribuição espacial da leptospirose no Rio Grande do Sul, Brasil: recuperando a ecologia dos estudos ecológicos, *Cad Saude Publica*, 19, 1283-1292.
- 21 – Openshaw S. (1977) Optimal zoning systems for spatial interaction models, *Environment and Planning A*, 9, pp. 169-184.
- 22 – Santos, S. M. & Souza, W. V. (Eds) (2007) *Introdução à estatística espacial para a saúde pública Brasília*. Ministério da Saúde. Fundação Oswaldo Cruz.
- 23 – Souza, W. V., Barcellos, C. C., Brito, A. M., Carvalho, M. S., Cruz, O. G., Albuquerque, M. F. M., Alves, K. R. & Lapa, T. M. (2001) Aplicação de modelo bayesiano empírico na análise espacial da ocorrência de hanseníase, *Rev Saude Publica*, 35(5), 474-480.
- 24 – Carvalho, M. & Câmara, G. (2002) Análise de eventos pontuais, in: S. Druck, C. Carvalho, G. Câmara & A. Monteiro (Eds) *Análise espacial de dados geográficos*. Brasília. EMBRAPA.
- 25 – Cromley, E. & McLafferty, L. (2002) *GIS and public health*. New York. The Guilford Press.
- 26 – Siqueira, J. B., Martelli, C. M., Maciel, I. J., Oliveira, R. M., Ribeiro, M. G., Amorim, F. P., Moreira, B. C., Cardoso, D. D., Souza, W. V. & Andrade, A. L. (2004) Household survey of dengue infection in central Brazil: spatial point pattern analysis and risk factors assessment, *Am J Trop Med Hyg*, 71(5), 646-651.

