
Métodos de Investigação Epidemiológica em Doenças Transmissíveis

Volume I



**Organização Pan-americana da Saúde
Fundação Nacional de Saúde/Centro Nacional de Epidemiologia**

Prefácio

Esta publicação é o resultado de um esforço coletivo de vários grupos de pesquisadores brasileiros para reunir um material que expresse a lógica, o delineamento e a racionalidade da análise de dados da investigação epidemiológica em saúde pública e a avaliação de procedimentos diagnósticos e de intervenção em doenças transmissíveis. É uma coletânea de textos, exercícios de análise estatística, rotinas de programação (em EPIINFO Versão 6) e roteiros de delineamentos de protocolos extraídos de projetos de pesquisa concluídos e publicados, para ser utilizado como instrução dirigida em oficinas de trabalho. O escopo do material permite a capacitação de recursos humanos em epidemiologia incluindo técnicos de programas de controle de doenças transmissíveis, e alunos de cursos de pós-graduação *sensu lato* e *sensu strictu*. Para um rendimento satisfatório o usuário dos módulos deverá ter conhecimentos básicos de epidemiologia, estatística e uso de computadores.

O primeiro volume está dividido em 6 módulos: **Estudos de prevalência, Avaliação de testes diagnósticos, Estudos de caso-controle, Estudos de coorte, Ensaio clínico, Análise epidemiológica - mortalidade e nascidos vivos**. O segundo volume corresponde ao Manual do Instrutor que contém todos os procedimentos de análise e os resultados dos exercícios produzidos pelo programa EPIINFO. O componente estatístico inclui procedimentos de análise descritiva e analítica, não incluindo, entretanto, métodos de análises multivariadas. O material vem acompanhado de um CD contendo: (1) cópias dos artigos científicos (formato PDF ou DOC) que foram utilizados nos exercícios; (2) bases de dados originais em formato REC; (3) cópia do EPIINFO ver. 6.04d; (4) cópia dos capítulos em formato PDF (Acrobat); (5) coletânea de slides dos módulos em formato PPS.

O módulo **Estudo de prevalência** (*determinando a frequência de eventos*) discute: (i) um inquérito sorológico para infecção pelo *T. cruzi* e (ii) a prevalência e fatores de risco para o vírus da hepatite B. **Avaliação de testes diagnósticos** (*Avaliando e selecionando o teste diagnóstico mais adequado*) traz como exemplos: (i) a concordância de resultados entre leitores de eletrocardiogramas; (ii) a avaliação do diagnóstico parasitológico na malária e (iii) a validação de técnicas sorológicas para doença de Chagas. Nos **Estudos de caso-controle** (*identificando fatores de risco para infecção e progressão de doenças*) são

avaliados fatores de risco para infecção chagásica e a efetividade da vacina BCG na hanseníase. Para os **Estudos de coorte** (*identificando os efeitos da exposição a fatores de risco*), utilizou-se a coorte de homossexuais e bissexuais masculinos HIV negativos e a análise de um ensaio controlado randomizado do impacto de redes impregnadas com permetrina no coeficiente de mortalidade infantil utilizando pessoas tempo de acompanhamento. Três exemplos são apresentados para **Ensaio clínico** (*avaliando a eficácia de medidas de intervenção*): (i) eficácia da cloroquina na malária por *P. falciparum* e (ii) eficácia do benzonidazol em crianças com infecção pelo *T. cruzi* e (iii) eficácia da vacina promastigota de *L. major* combinada com BCG na leishmaniose cutânea. O potencial de análise epidemiológica dos bancos de dados do **Sistema de Informação de Mortalidade (SIM) e Nascidos Vivos (SINASC)** do Ministério da Saúde (*avaliando fatores de risco para morte infantil*) é exemplificado através do delineamento de uma coorte retrospectiva.

O conteúdo teórico e adequação dos exercícios foram objeto de avaliação de duas oficinas de trabalho e Curso de Mestrado em Medicina Tropical do Instituto de Patologia Tropical e Saúde Pública da Universidade Federal de Goiás, incorporando também, as importantes sugestões advindas da reunião promovida pelo Ministério da Saúde/Fundação Nacional de Saúde/Cenepi em dezembro de 1996 em Recife. Nesta Segunda edição os exercícios dos Estudos de coorte e Ensaio clínico foram implementados, incluindo-se estudos sobre homossexuais e bissexuais masculinos HIV negativos, impacto de redes impregnadas com permetrina no coeficiente de mortalidade infantil e dose única da vacina de *L. major* morta associada com BCG contra leishmaniose cutânea. Sugestões advindas de diversas oficinas de trabalho realizadas no Brasil, Argentina e Venezuela e das oficinas realizadas em Nepal e Ghana (versão em inglês), promovidas pelo UNDP/WHO/TDR, foram também incorporadas nesta edição.

Os Editores

Organização e Edição

Ana Lúcia S. Sgambatti de Andrade – Professora titular e coordenadora do Mestrado em Medicina Tropical – área de concentração em Epidemiologia – Instituto de Patologia Tropical e Saúde Pública, Universidade Federal de Goiás.

Fábio Zicker – UNDP/World Bank/WHO Special Program for Research and Training in Tropical Diseases. Geneva, Switzerland.

Colaboradores

Ana Lúcia S. Sgambatti de Andrade *Ensaio clínico e Exercícios: Avaliação de testes diagnósticos, Estudos de Prevalência, Estudos de caso-controle, Ensaio clínico*
Departamento de Medicina tropical, Saúde Coletiva e Dermatologia
Instituto de Patologia Tropical e Saúde Pública – Universidade Federal de Goiás

Carlos Maurício F. Antunes *Estudos de coorte*
Departamento de Parasitologia – Instituto de Ciências Biológicas
Universidade Federal de Minas Gerais

Celina M. Turchi Martelli *Estudos de prevalência*
Departamento de Medicina Tropical, Saúde Coletiva e Dermatologia
Instituto de Patologia Tropical e Saúde Pública – Universidade Federal de Goiás

Isildinha Marques Reis *Revisão estatística*
Departamento de Epidemiologia – Faculdade de Saúde Pública
Universidade de São Paulo

José Maria P. Souza *Estudos de caso-controle*
Departamento de Epidemiologia – Faculdade de Saúde Pública
Universidade de São Paulo

Mariangela Carneiro *Estudos de coorte*
Departamento de Epidemiologia – Faculdade de Saúde Pública
Universidade de São Paulo

Maurício Gomes Pereira *Avaliação de testes diagnósticos*
Departamento de Saúde Coletiva
Universidade de Brasília

Otaliba L. Moraes Neto *Mortalidade e nascidos vivos – Análise epidemiológica*
Departamento de Medicina Tropical, Saúde Coletiva e Dermatologia
Instituto de Patologia Tropical e Saúde Pública – Universidade Federal de Goiás

Simonne Almeida e Silva *Processamento de dados, Manual do Instrutor*
Departamento de Medicina Tropical, Saúde Coletiva e Dermatologia
Instituto de Patologia Tropical e Saúde Pública – Universidade Federal de Goiás

Joaquim Gonçalves Valente *Revisor externo*
Instituto de Medicina Social
Universidade Estadual do Rio de Janeiro

Índice

Módulo 1	
Estudos de prevalência.....	01
Módulo 2	
Avaliação de Testes Diagnósticos.....	21
Módulo 3	
Estudos de Caso-Controle	41
Módulo 4	
Estudos de Coorte.....	62
Módulo 5	
Ensaio Clínico.....	91
Módulo 6	
Mortalidade e Nascidos vivos	
<i>Análise epidemiológica</i>	118
ARTIGOS	
ANDRADE, A.L.S.S., ZICKER, F., LUQUETTI, A.O., OLIVEIRA, R.M., SILVA, S.A., SOUZA, J.M.P. & MARTELLI, C.M.T. Surveillance of <i>Trypanosoma cruzi</i> transmission by serological screening of schoolchildren. WHO Bulletin, 70(5):625-9, 1992	
MARTELLI, C.M.T., ANDRADE, A.L.S.S., CARDOSO, D.D.P., SOUSA, L.C.S., SILVA, S.A., SOUSA, M.A. & ZICKER, F. Soroprevalência e fatores de risco para a infecção pelo vírus da Hepatite B pelos marcadores AgHBs e Anti-HBs em prisioneiros e primodoadores de sangue. Revista de Saúde Pública, 24(4):270-6, 1990	
ANDRADE, A.L.S.S., MARTELLI, C.M.T., LUQUETTI, A.O., OLIVEIRA, O.S.O., SILVA, S.A. & ZICKER, F. Serologic screening for <i>Trypanosoma cruzi</i> among blood donors in Central Brazil. Bulletin of the Pan American Health Organization, 26(2):157-164, 1992	
ANDRADE, A.L.S.S., MARTELLI, C.M.T., OLIVEIRA, R.M., ARIAS, J.R., ZICKER, F. & PANG, L. High prevalence of asymptomatic malaria in gold mining areas in Brazil. Clinical Infectious Disease, 20(2): 475, 1995.....	

ANDRADE, A.L.S.S., ZICKER, F., SILVA, I.G., SOUZA, J.M.P. & MARTELLI, C.M.T. Risk factors for *Trypanosoma cruzi* infection among children in Central Brazil: a case-control study in vector control settings. *American Journal of Tropical Medicine and Hygiene*, 52(2):183-187, 1995.....

RODRIGUES, M.L.O., SILVA, S.A., NETO, J.C.A., ANDRADE, A.L.S.S., MARTELLI, C.M.T. & ZICKER, F. Protective effect of intradermal BCG against leprosy: a case-control study in Central Brazil. *International Journal of Leprosy and Other Mycobacterium Disease*, 60(3):335-339, 1992

BINKA F.N., KUBAJE A, ADJUIK M. WILLIAMS L.A., LENGELER C. MAUDE G.H., ARMAH G.E., KAJIHARA B., ADIAMAH J.H. & SMITH P.G. Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. *Tropical Medicine and International Health*, 1(2): 147-154, 1996

CARNEIRO M., ANTUNES C.M.F., GRECO M., OLIVEIRA E., ANDRADE J., LIGNANI JR. L. & GRECO D.B. Design, Implementation, and Evaluation at Entry of a Prospective Cohort Study of Homosexual and Bisexual HIV-1 Negative Men in Belo Horizonte, Brazil: Project Horizonte. *Journal of Acquired Immune Deficiency Syndromes*, 25(2):182-187, 2000

ANDRADE, A.L.S.S.; ZICKER, F.; OLIVEIRA, R.M.; SILVA, S.A.; LUQUETTI, A.O.; ANDRADE, J.G.; ANDRADE, S.G.; ALMEIDA I.C.; TRAVASSOS, L.R.; MARTELLI, C.M.T. A randomised trial of efficacy of benznidazole in treatment of early *Trypanosoma cruzi* infection. *Lancet*, 348:1407-1413, 1996

ANDRADE JG; ANDRADE ALSS; ARAUJO ESO; OLIVEIRA RM; SILVA SA; MARTELLI CMT & ZICKER F. A randomized clinical trial with high dose of chloroquine for treatment of *Plasmodium falciparum* malaria in Brazil. *Revista do Instituto de Medicina Tropical de São Paulo*, 34(5):467-473, 1992

SHARIFI I., FEKRI A.R., AFLATONIAN M., KHAMESIPUOUR A, NADIM A. MOUSAVI M.A., MOMENI A.Z., DOWLATI Y., GODAL T., ZICKER F. & SMITH P.G. Randomised vaccine trial of single dose killed *Leishmania major* plus BCG against anthroponotic cutaneous leishmaniasis in Bam, Iran. *The Lancet*, 351:1540-1543, 1998.....

ESTUDOS DE PREVALÊNCIA

INTRODUÇÃO	
Definição.....	02
Utilização em Saúde Pública e Pesquisa.....	03
MEDIDAS DE PREVALÊNCIA	
Prevalência no ponto.....	03
Prevalência no período.....	03
EXEMPLOS DE ESTUDOS DE PREVALÊNCIA	
Estudos de soroprevalência.....	05
Estudos de prevalência repetidos.....	05
METODOLOGIA	
Amostragem.....	06
Tamanho da amostra.....	06
Fonte de dados Primária e Secundária.....	07
Tipos de viés.....	08
Viés de sobrevivente.....	08
Viés de seleção de participantes.....	08
Viés de observador.....	08
ANÁLISE DE DADOS	
Medida de prevalência.....	09
Estratificação.....	10
VANTAGENS E LIMITAÇÕES	10
ROTEIRO PARA DELINEAR UM ESTUDO DE PREVALÊNCIA	11
REFERÊNCIAS PARA LEITURA	12
EXERCÍCIOS	13
DICIONÁRIO DE BANCO DE DADOS	19

INTRODUÇÃO

Definição - Os estudos de prevalência são os estudos descritivos populacionais mais amplamente difundidos e publicados em epidemiologia. Neste tipo de delineamento, também conhecido como estudo transversal ou de corte-transversal, se obtém a frequência de ocorrência dos eventos de saúde numa população em um ponto no tempo ou em curto espaço de tempo. Os estudos transversais permitem, também, investigar associações entre fatores de risco e doença, embora não seja o delineamento mais eficiente para se estudar causalidade, e portanto, podem ser classificados como analíticos.

A taxa de prevalência (P) é estimada pela fórmula:

$$P = \frac{\text{Número de pessoas com o evento de interesse (infecção, doença, etc.)}}{\text{População sob risco de apresentar o evento de interesse em determinado tempo}}$$

Habitualmente, a população sob risco corresponde à população da área de estudo; ou é definida geograficamente, ou pelas características ocupacionais ou pelo uso dos serviços de saúde, entre outros. A taxa de prevalência frequentemente é expressa multiplicada por um fator. Por exemplo, a taxa de prevalência de hanseníase no Brasil, em 1995, foi de 4,8 por 100 000 habitantes.

A prevalência é uma medida de morbidade composta da incidência (I), correspondendo aos casos novos, e da duração média (D) da doença ou evento. É uma proporção, cujo numerador está incluído no denominador e portanto, não tem unidade de medida, podendo variar de 0 a 1,0. Quando a incidência e a dinâmica populacional são constantes, a prevalência (P) pode ser estimada pela fórmula:

$$P = \text{INCIDÊNCIA} \times \text{DURAÇÃO DA DOENÇA}$$

A duração da doença pode ser obtida quando a incidência e prevalência são conhecidas. Por exemplo, em 1995, foram notificados, em média, 33 casos novos mensais de Tuberculose no município de Goiânia, com taxa de incidência de 3,3 por 100 000 habitantes e, no mesmo período, a taxa média mensal de prevalência foi de 19,8 por 100 000 habitantes. Assim, a duração média do tratamento de Tuberculose foi de 6 meses, de acordo com a fórmula:

$$D = \frac{P}{I} = \frac{19,8}{3,3} = 6 \text{ meses}$$

A prevalência difere da incidência porque pode ser avaliada através de uma medida única

(seccional). Em contrapartida, a incidência requer pelo menos duas medidas do mesmo indivíduo ou da mesma população; a primeira, no início do estudo, para excluir os já doentes/infectados, e a segunda, para detectar o aparecimento do novo caso (doença/infecção). Em doenças infecciosas de evolução rápida, a prevalência não constitui medida de interesse, preferindo-se a incidência. Ao contrário, para eventos (infecções/doenças) de mais longa duração ou crônicos, a prevalência pode indicar o risco de exposição para os indivíduos suscetíveis.

Utilização em Saúde Pública e Pesquisa - Estudos de prevalência são frequentemente usados para planejamento em saúde pública e como etapa inicial (linha de base) para avaliar Programas de Controle. São também utilizados na seleção de participantes para outros estudos como caso-controle, coorte e ensaios clínicos. Por exemplo, uma triagem sorológica inicial para infecção pelo *Trypanosoma cruzi*, em aproximadamente 2000 escolares de área rural do Centro-Oeste/Brasil, detectou uma prevalência de 7,9% com intervalo de 95% de confiança de 6,8% a 9,1% (Andrade et al., 1992). Uma amostra das crianças soropositivas foram então selecionadas para participar de ensaio clínico para avaliar a eficácia do benzonidazol como tratamento específico. Ainda, os soropositivos (casos) e soronegativos (controles) foram comparados para determinar fatores de risco ambientais, familiares e nutricionais associados à infecção pelo *T. cruzi* na infância.

MEDIDAS DE PREVALÊNCIA

A prevalência é considerada uma medida de morbidade conceitualmente simples. No entanto, em doenças transmissíveis a sua interpretação requer o conhecimento dos mecanismos pelos quais a infecção é transmitida, da fase pré-clínica e dos diferentes graus de gravidade da doença, além do tempo de duração da infecção/doença e do percentual de casos letais. Os critérios de definição de infecção e doença e a duração provável do evento até a recuperação ou óbito devem ser especificados na etapa de delineamento dos estudos de prevalência.

Os tipos de coeficientes de prevalência mais comumente utilizados são: **Prevalência no ponto e Prevalência no período.**

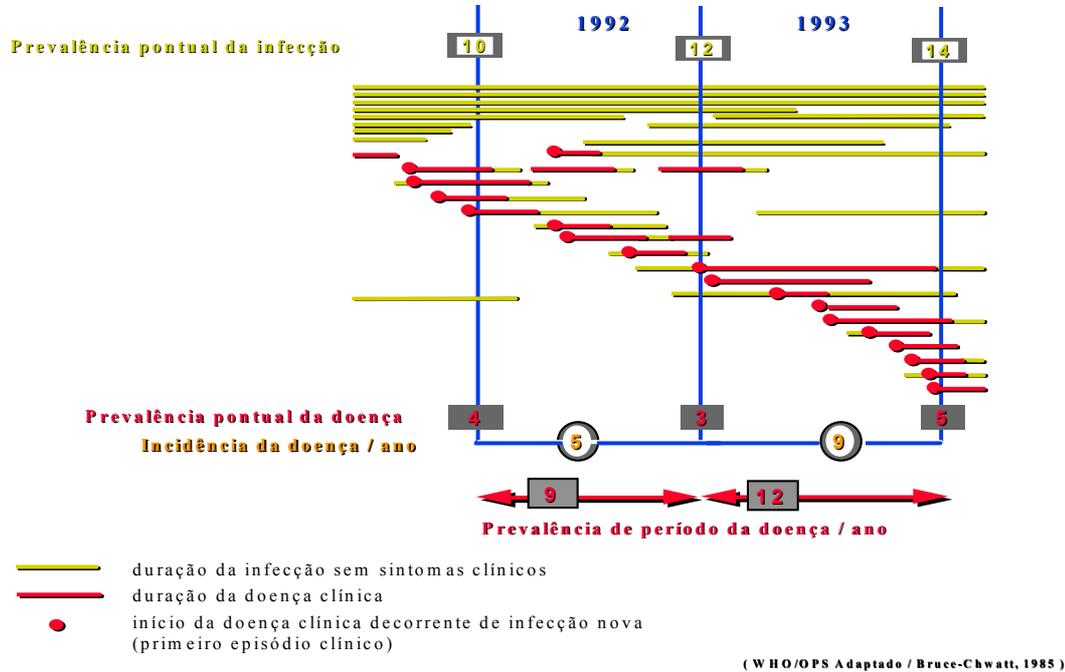
- **Prevalência no ponto** - é o número total de indivíduos na população com a doença ou atributo em um momento de tempo dividido pela população sob risco de ter o atributo ou doença no mesmo momento de tempo. O coeficiente de prevalência tem sido usado como sinônimo de coeficiente de prevalência pontual.
- **Prevalência no período** - refere-se ao número total de indivíduos com a doença ou atributo durante um determinado período de tempo dividido pela população sob risco de ter o atributo ou doença durante o mesmo período de tempo. Consiste na *prevalência do ponto* no início do período especificado, *mais todos os casos novos* que ocorreram durante o período.

Prevalência durante a vida ("*lifetime prevalence*") refere-se ao número total de indivíduos que tiveram a doença ou atributo pelo menos durante parte da vida para doenças sem cura; e a *Prevalência de episódios* diz respeito a eventos repetidos. Quando a prevalência é obtida através dos sistemas de vigilância é, por vezes, denominada de *Prevalência de registro*.

A Figura 1 exemplifica os conceitos de prevalência no ponto e no período, utilizando como exemplo a malária. A prevalência pontual, em áreas endêmicas de malária, pode ser obtida pela triagem parasitológica da população em curto espaço de tempo. As diferenças entre a prevalência de infecção e a incidência de casos clínicos estarão na dependência dos níveis de endemicidade da malária. De acordo com a Figura, no início de 1992, a **prevalência pontual** da malária sintomática foi de 4 casos e durante o ano foram diagnosticados 5 novos casos (incidência), resultando numa **prevalência de período** igual a 9. No início de 1993, a **prevalência pontual da infecção** foi de 12

casos e o número de casos clínicos de 3, o que mostra as diferenças entre as prevalências pontuais de infecção e da doença, respectivamente.

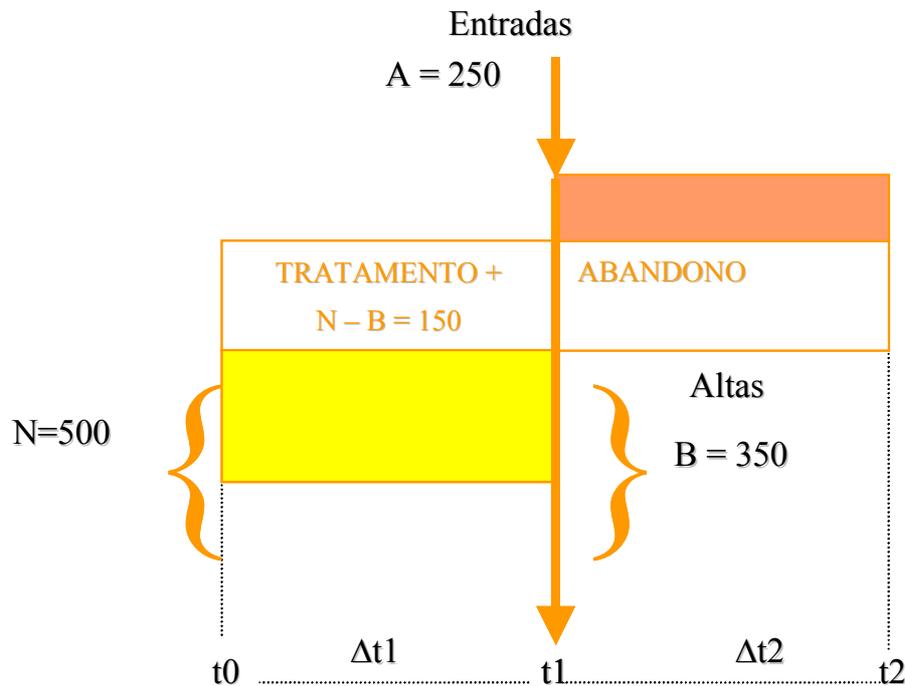
Figura 1. Malária - Prevalência no ponto e no período



As estimativas de prevalência utilizadas em Programas de Controle são influenciadas por normas administrativas, operacionais, mudanças na definição de caso, esquemas de tratamento e critérios de alta. Intervenções em massa podem interferir na transmissibilidade de uma infecção, reduzindo os casos novos, na duração e nas características da infecção/doença dos casos já instalados. Por exemplo, na hanseníase, de acordo com o Programa de Controle, níveis endêmicos, com taxa de prevalência inferior a 1 caso por 10 000 habitantes, definem área sob eliminação da doença.

A Figura 2 exemplifica o conceito de prevalência no ponto e no período para hanseníase. Considere 500 casos (N) no início do período (t_0) e que todos os novos casos ($A=250$), independente da forma clínica, ocorreram ao mesmo tempo, na metade do ano (t_1). A **prevalência no período** (Δt_1) será 750 casos (500 casos do início do período mais 250 casos novos). Supondo que no tempo t_1 ocorreram 350 altas ($B=350$), a **prevalência no ponto t_1** será o estoque de casos ($N-B=150$) mais os casos novos ($A=250$), totalizando 400 casos. Assim, em uma situação de incidência estável, a redução da prevalência no ponto estará na dependência do número de pacientes tratados (curados ou altas) e do percentual de abandono de tratamento.

Figura 2
Hanseníase – Prevalência no Ponto e no Período



EXEMPLOS DE ESTUDOS DE PREVALÊNCIA

Estudos de soroprevalência - são uma denominação comum para inquéritos que utilizam marcadores sorológicos, sendo particularmente úteis para infecções virais e bacterianas que induzem à formação de anticorpos ou outros marcadores biológicos específicos. Este tipo de estudo tem sido utilizado para determinar a distribuição geográfica de inúmeras doenças, como por exemplo, hepatites A, B, C; HIV e também em inquéritos pré e pós vacinação em massa para avaliar soroconversão de anticorpos. A prevalência é geralmente estimada por estratos de idade e sexo, possibilitando o entendimento da dinâmica de transmissão da infecção na comunidade. Este tipo de análise permite a identificação de áreas de alto risco na comunidade e os compartimentos populacionais de infectados, imunes, portadores e, pela ausência de marcadores, o grupo de suscetíveis. Desta forma, avalia-se a doença/infecção/imunidade no presente e passado, possibilitando prever o futuro risco de infecção em determinada população.

Estudos de prevalência repetidos - são estudos de prevalência realizados repetidamente em intervalos de tempo, geralmente anos ou décadas. Diferem do estudo de coorte por não estudarem o mesmo grupo de indivíduos. Avaliam amostras independentes de uma mesma população em diferentes períodos de tempo. Assim, é improvável que na amostragem probabilística os mesmos indivíduos sejam selecionados para mensurar os eventos de interesse. Estes estudos de prevalência repetidos são utilizados para avaliar o "status" de saúde/doença/infecção e tendência em grandes populações; são importantes para planejamento dos serviços de saúde e verificação de alterações de saúde ao longo do tempo. Alguns exemplos são os inquéritos nutricionais e de morbidade domiciliar nos censos.

METODOLOGIA

Amostragem

Amostragem aleatória - de forma geral, estudos de prevalência requerem *amostragem aleatória* de uma população. Uma amostragem probabilística aumenta substancialmente a chance de os participantes serem representativos da população-alvo, assegurando a validade interna do estudo e possibilitando extrapolar os resultados (validade externa) do estudo para outras comunidades.

Amostragem sistemática - pode ser utilizada em substituição à amostragem aleatória simples. Neste tipo de amostragem utiliza-se algum tipo de seqüência pré-estabelecida para selecionar participantes, por exemplo, de prontuários de fichário, casas numa rua e pacientes da demanda espontânea de ambulatório.

Amostragem estratificada – Neste tipo de *amostragem* a população é dividida em sub-grupos distintos de acordo com algumas características e seleciona-se uma amostra aleatória de cada sub-grupo. Se a proporção de amostras de cada estrato é a mesma que a proporção da população total, então todos os estratos estarão adequadamente representados em relação ao número de pessoas na amostra. Um exemplo de amostragem em dois estágios foi desenvolvido pela *WHO (EPI-WHO)* para avaliar cobertura vacinal e a qualidade dos serviços de saúde. Neste esquema, 30 aglomerados (vilas, bairros etc) são selecionados e 7 crianças da faixa etária são selecionadas de cada localidade (desenho 30x7). Variações sobre a técnica EPI-WHO, outros tipos de amostragens e fórmulas para correção de erro amostral são encontrados em livros de estatística mais avançados.

Tamanho da amostra

Enquanto a amostragem probabilística viabiliza a validade interna do estudo, a precisão na estimativa da prevalência obtida depende do tamanho da amostra. Assim, a amplitude do intervalo de confiança (estimada pelo intervalo da prevalência na população) reflete o grau de precisão decorrente do tamanho da amostra fixado.

O cálculo do tamanho da amostra para amostragem aleatória simples é:

$$n = Z * Z [P (1-P)] / (D*D)$$

onde:

Z valor da distribuição normal padrão correspondente ao nível de confiança desejado (Z= 1,96 para Intervalo de 95% de Confiança - IC 95%)

P prevalência esperada

D erro máximo aceitável na estimativa (semi-amplitude do IC - medida de precisão)

Por exemplo, para se estimar a soropositividade aos anticorpos do vírus da dengue em uma população de aproximadamente 1 000 000 hab. tendo uma prevalência esperada de 15% (P=0,15), com amplitude do IC de 95% de 12% (D=0,06), o número de pessoas a serem investigadas seria:

$$n = 1,96^2 (0,15 (1-0,15)) / (0,06^2)$$

n = 136 pessoas

Na maioria dos estudos, será necessário aumentar cerca de 10% o tamanho da amostra, para levar em consideração os não respondedores, isto é, indivíduos selecionados que se recusam a participar.

O tamanho da amostra deve dar uma idéia da ordem de grandeza da população necessária para o estudo, sem contudo ser rígido, uma vez que o cálculo baseia-se em estimativas de parâmetros (prevalência esperada). Geralmente, esta estimativa é obtida através de revisão da literatura. Tamanhos de amostras devem ser calculados levando-se em conta diferentes estimativas de prevalência e precisão de acordo com o objetivo do estudo. Deve-se balancear o que é desejável e o que é operacionalmente possível. De forma geral, pesquisas de opinião são realizadas com cerca de 1000 pessoas com o intuito de se obter uma boa precisão (por exemplo, IC 95% com amplitude no máximo de 6%). É bom enfatizar que estudos de prevalência não são adequados para eventos que apresentam baixa frequência de ocorrência, pois nestas circunstâncias seria necessário uma amostra de grande tamanho, o que acarretaria problemas operacionais e de custo.

Fonte de dados Primária e Secundária

Sistemas de Informação Oficiais.

Fonte de dados secundários podem ser úteis para estudos de prevalência e devem ser utilizados, sempre que possível, como primeira aproximação da questão a ser respondida. Entre as vantagens da utilização epidemiológica de bancos de dados existentes estão o baixo custo e disponibilidade de informações por um longo período de tempo (grandes séries históricas) que permitem avaliar a tendência secular e a facilidade de acesso da informação via sistemas computadorizados.

De forma geral, os programas de controle de doenças endêmicas são concebidos acoplados a sistemas de informação próprios para acompanhamento sistemático destes agravos, o que possibilita a construção de séries históricas. Outras fontes são os prontuários de hospitais gerais e de centros de referência ou sentinela para doenças infecciosas. Exemplo, prevalência de infecções oportunistas em pacientes com SIDA.

Vale mencionar o interesse recente no uso epidemiológico dos Formulários de Autorização de Internação Hospitalar (AIH), formulário padronizado exigido para o reembolso dos serviços hospitalares prestados pelos hospitais públicos e privados vinculados ao Sistema Único de Saúde. Enquanto as AIHs têm sido utilizadas na programação e gerência e avaliação dos serviços hospitalares, epidemiologistas têm explorado menos as potencialidades desta fonte de dados para investigações e vigilância em saúde. Neste tipo de sistema é difícil separar o caso novo (caso incidente) do prevalente. Este aspecto tem sido também descrito em relação ao uso epidemiológico dos dados oriundos do "Medicare" - programa de seguro médico americano.

Para a interpretação de dados secundários deve-se levar em conta a amplitude de cobertura e qualidade das informações, mudanças na definição de casos ao longo do tempo, ações administrativas como mudança de notificação voluntária para transitória e alteração das intervenções e/ou formulários implantados.

Reconhecer as limitações dos dados existentes, em particular as da qualidade da coleta original dos dados, viés associado à determinação da doença e seleção dos casos para tratamento faz parte da interpretação epidemiológica.

Coleta de dados primários

Questionários auto-administrados e entrevistas são formas habituais de se obter informações de morbidade, frequência de sintomas e variáveis de interesse em inquéritos de prevalência. Adicionalmente testes laboratoriais para detecção de marcadores biológicos podem ser utilizados para mensurar prevalências de infecção / doença e fatores de risco. A qualidade de dados primários é evidentemente superior em relação aos dados secundários. Estudos epidemiológicos devem ser

devidamente planejados para coletar dados de forma padronizada e que sejam válidos, com informações completas e cuidadosamente registradas para viabilizar uma base de dados consistente para análise e interpretação.

Tipos de viés

Viés de sobrevivente - Em qualquer investigação de corte-transversal, viés sistemático pode ocorrer uma vez que os casos prevalentes observados tendem a excluir os óbitos precoces, que não estarão presentes no momento da seleção, tendendo a super-representação dos casos de mais longa sobrevida. Como a probabilidade de sobrevivência da doença afeta a prevalência, estudos baseados em casos prevalentes geram associações que refletem determinantes da sobrevida dos casos.

Viés de seleção de participantes - A utilização de amostragem por critério de conveniência, como por exemplo acessibilidade da clientela a serviços públicos de saúde, clínicas especializadas e serviços de referência etc., pode introduzir, de forma consciente ou inconscientemente, viés de seleção de participantes. Portanto, amostragem não aleatória apresenta limitada validade (valor) para descrever a frequência da doença ou outras características da população. Participantes auto-selecionados ou voluntários tendem a ser mais saudáveis que a população geral.

Em qualquer circunstância onde exista a auto-seleção ou voluntariedade dos participantes, as generalizações para a população geral são difíceis. Por exemplo, amostragem procedente de subgrupos de comportamento de risco tendem a superestimar a prevalência de algumas Doenças Sexualmente Transmissíveis. Ao contrário, outros subgrupos populacionais como doadores de sangue e mulheres grávidas apresentariam taxas provavelmente mais próximas da população geral. Porém, estes estudos podem em conjunto, fornecer um cenário de exposição, dos grupos de maior risco ao menor risco, em diferentes subgrupos que compõem a população geral, mostrando as tendências mais recentes de epidemias/endemias na falta de pesquisas de base populacional ou de sistema de notificação confiável e atualizado.

A recusa de indivíduos em participar dos estudos de prevalência, tanto na entrevista como na doação de material biológico, pode, também, introduzir *viés de seleção dos participantes* e deve ser mantida ao mínimo para garantir a representatividade da amostra. Alternativas operacionais para diminuir o número de recusas em participar devem ser antecipadas no protocolo. As diferenças existentes entre participantes e não participantes do estudo devem ser avaliadas em relação às variáveis sexo, idade e algum indicador sócio-econômico.

É comum investigadores de área clínica e laboratorial utilizarem-se de coleta indiscriminada de sangue e/ou de uma soroteca ou espécimens biológicos já disponíveis em laboratório, para avaliar a prevalência de outro evento de saúde. Sorotecas ou bancos de material biológico sem a descrição da população de que são provenientes, ou do método de amostragem e das circunstâncias nas quais foram obtidas perdem o seu valor em estudos com objetivos epidemiológicos. Para que resultados obtidos através destas testagens representem a prevalência real, todos os requisitos de um delineamento de projeto devem ser seguidos: (a) objetivos claros; (b) representatividade da população, (c) tamanho de amostra; (d) conhecimento dos testes a serem usados, sensibilidade e especificidade, limites de acurácia e significado diagnóstico.

Viés de observador – Erros ou falhas na medida das variáveis de exposição ou de efeito/doença que resultem em perda diferencial da qualidade (acurácia) da informação podem induzir a um viés do observador. Como nos demais estudos, o pesquisador de campo deverá ser treinado para aplicar as perguntas de forma uniforme evitando distorções nas respostas por forçar obtenção de respostas positivas ou negativas.

ANÁLISE DE DADOS

Medida de prevalência

O intervalo de confiança de uma prevalência é a técnica estatística mais comum para mostrar o grau de incerteza da estimativa. Fórmulas para o cálculo dos limites superiores e inferiores estão disponíveis em qualquer texto básico de estatística. Quanto maior o tamanho da amostra incluída para estudo, mais precisa será a proporção observada, assumindo que se trata de uma amostra aleatória da população de estudo. Geralmente calcula-se Intervalo de 95% de confiança para a prevalência populacional, e seus limites superior e inferior podem ser apresentados numericamente ou graficamente. A amplitude do intervalo de confiança indica a precisão inerente da estimativa da prevalência decorrente do tamanho da amostra. Intervalos de confiança terão limite superior e inferior próximos, isto é, de pequena amplitude, nos valores de proporções (prevalências) calculadas de grandes amostras. Quando dois intervalos de confiança não se sobrepõem, há evidência de uma diferença estatisticamente significativa entre as proporções dos dois grupos. No entanto, o procedimento mais adequado para comparar duas proporções é o teste de hipóteses (qui-quadrado ou de Fisher para amostras pequenas).

Em estudos de prevalência é possível avaliar a associação entre exposição e doença na população em um determinado momento. Uma estimativa do risco relativo ou do *odds ratio* pode ser obtida a partir do coeficiente do evento (ou doença) em expostos e não expostos, especialmente quando a frequência da doença/evento for baixa. Nestas circunstâncias pode-se utilizar a razão entre duas prevalências denominada de razão de prevalência (RP). Por exemplo, um estudo foi conduzido para avaliar a associação entre doença sexualmente transmissível (DST) e meninos sem vínculo familiar/de rua (Porto et al., 1994). Foram investigados 496 adolescentes de 9 a 20 anos, sendo 101 meninos de rua e 395 meninos trabalhadores de rua, porém, com vínculo familiar. A proporção de Doenças Sexualmente Transmissíveis foi de 24,8% (25/101) em meninos de rua e de 3,5 (14/395) em meninos na rua. A tabela 2 x 2 a seguir mostra os resultados da Razão de Prevalência (RP) e *odds ratio* (OR).

	Doença Sexualmente Transmissível		Total
	Sim	Não	
Meninos de rua	25(24,8%)	76(75,2%)	101(100%)
na rua	14 (3,5%)	381(96,5%)	395(100%)
Total	39 (7,9%)	457(92,1%)	496(100%)

$$RP = 0,248 / 0,035 = 7,1 \text{ (IC 95\% 3,7-12,9)}$$

$$OR = \frac{25 \times 381}{14 \times 76} = 8,9 \text{ (IC 95\% 4,2 - 19,1)}$$

Portanto, meninos de rua tiveram um risco pelo menos 7,1 vezes maior de apresentarem DST, quando comparados aos meninos na rua.

Estratificação - A principal técnica para avaliar *confusão* e para examinar *interações* (modificação de efeito) entre fatores de risco é a *estratificação*. As seguintes etapas são habitualmente realizadas na análise estratificada:

- Separar os indivíduos em estratos de acordo com a variável potencial de confusão/interação;
- Calcular para cada estrato-específico estimativas do efeito da exposição (*odds ratio* e intervalo de confiança, se tiver dados suficientes);
- Verificar se as magnitudes da diferença entre os OR dos diversos estratos sugerem interação ou confusão;
- Calcular a estimativa de risco agrupada. O teste de Mantel-Haenszel é utilizado para corrigir as estimativas de risco, levando em consideração a variável de confusão analisada.

É possível, também, ajustar a prevalência por mais de uma variável de confusão, através da estratificação. Porém, numerosos estratos tendem a produzir caselas com pequeno número de observações com perda de precisão dos cálculos. Esta limitação da estratificação nos ajustes simultâneos de diversas variáveis de confusão pode ser superada em alguma extensão pelo uso de técnicas de modelagem. Entre os modelos de regressão, o modelo de regressão logística pode ser aplicado aos estudos de prevalência, para avaliar o efeito de uma exposição na presença de outros fatores de risco adicionais da mesma forma que em estudos de coorte (quando se utiliza o risco ao invés de pessoa-tempo) e em estudos de caso-controle.

VANTAGENS E LIMITAÇÕES

Entre as **vantagens** citamos:

- (a) Rapidez, baixo custo e menor complexidade operacional dos estudos de prevalência, quando comparados aos estudos de caso-controle e de coorte
- (b) Os estudos de prevalência devem basear-se, sempre que possível, em amostras da população geral e não somente na clientela de serviços de saúde

Apresentam-se como **desvantagens**

- (a) Não são adequados para doenças raras ou de pequeno período de duração quando poucas pessoas estarão doentes em um ponto no tempo
- (b) É freqüentemente difícil separar causa e efeito, porque as medidas de exposição e doença são feitas ao mesmo tempo. De forma geral, estudos de prevalência não podem testar hipóteses etiológicas

ROTEIRO PARA DELINEAR UM ESTUDO DE PREVALÊNCIA

① *Defina a relevância da pergunta a ser respondida*

- . verifique a existência de estudos semelhantes já realizados na área
- . certifique que a pergunta gera conhecimento científico e intervenção em saúde pública

② *Formule com clareza as perguntas a serem respondidas*

- . esclareça nos objetivos o evento que se quer mensurar e a população que será investigada
- . certifique se as perguntas são passíveis de serem respondidas técnica e operacionalmente

③ *Descreva os procedimentos técnicos para avaliar o evento de interesse*

- . métodos laboratoriais, entrevista, questionários, exame clínico
- . interpretação e categorização dos parâmetros a serem avaliados

④ *Estabeleça os procedimentos de amostragem*

- . defina a população de referência e a população de estudo
- . estabeleça o método de amostragem a ser utilizado

⑤ *Calcule o tamanho da amostra*

- . estime o valor esperado de prevalência para o cálculo do tamanho amostral necessário
- . defina a precisão (erro) aceitável em sua estimativa de prevalência

⑥ *Esclareça as questões éticas*

- . riscos versus benefícios da detecção do evento
- . atenção médica aos participantes que apresentarem o evento
- . confidencialidade dos resultados
- . uso de amostras biológicas e sorotecas pré-existentes

⑦ *Descreva as etapas para a análise dos dados*

- . indique os parâmetros (proporção e IC95%, média), métodos estatísticos e subgrupos de comparação

REFERÊNCIAS PARA LEITURA

ANDRADE, A.L.S.S., ZICKER, F., LUQUETTI, A.O., OLIVEIRA, R.M., SILVA, S.A., SOUZA, J.M.P. & MARTELLI, C.M.T. Surveillance of *Trypanosoma cruzi* transmission by serological screening of schoolchildren. **WHO Bulletin**, **70(5)**:625-9, 1992.

BEAGLEHOLE, R., BONITA, R. & KJELLSTRÖM, T. *Basic Epidemiology*. World Health Organization, Geneva, 1993.

GIESECKE, J. *Modern Infectious Disease Epidemiology*, Boston: Little, Brown and Company, 1994.

HENNECKENS, H.C. & BURING, J.E. *Epidemiology in Medicine*, 5th ed. Boston: Toronto, Ed. Little, Brown and Company, 1987.

KLEINBAUM, D.G., KUPPER, L.L. & MORGENSTERN, H. *Epidemiologic Research*. New York, Ed. Van Nostrand Reinhold, 1982.

PAUL, J.R. & WHITE, C. *Serological epidemiology*. Academic Press New York and London, 1973.

PEREIRA, M.G. *Epidemiologia Teoria e Prática*. Rio de Janeiro. Guanabara Koogan, 1995.

PORTO, S.O.B., CARDOSO, D.D.P., QUEIROZ, D.A.O., ROSA, H., ANDRADE, A.L.S.S., ZICKER, F. & MARTELLI, C.M.T. Prevalence and risk factors for HBV infection among street youth in Central Brazil. **Journal of Adolescent Health**, **15**:577-81, 1994.

EXERCÍCIOS

Arquivos: 1. triagem.rec
2. preva.rec

1. Triagem sorológica para *Trypanosoma cruzi* em crianças. Um inquérito sorológico foi conduzido para mensurar a prevalência de infecção pelo *Trypanosoma cruzi* em escolares de 7 a 12 anos residentes em áreas rurais endêmicas do Brasil Central. Procedeu-se a coleta de sangue em papel de filtro de 1990 crianças para realização das reações de hemaglutinação indireta, imunofluorescência indireta e ELISA. Detalhes da área de estudo e metodologia estão na referência **Andrade et al, 1992**. A análise de dados foi planejada para: (a) comparar a soropositividade por técnica empregada, por município, sexo e por idade e (b) calcular a razão de prevalência (*odds ratio*) por sexo e faixa etária. Utilize o arquivo **triagem.rec** para responder as questões a seguir.

Questão 1. Calcule a prevalência e Intervalo de 95% de Confiança (IC 95%) da soropositividade por técnica empregada. Existe diferença estatística entre os resultados?

Notas 1: READ TRIAGEM.REC (para abrir o arquivo)
FREQ HAI
anote os resultados
FREQ IFI
anote os resultados
FREQ ELISA
anote os resultados
tecle F9 para ir ao DOS
digite EPITABLE para cálculo dos IC95% (comando DESCRIBE /
PROPORTION / SIMPLE RANDOM SAMPLING)
utilize os resultados das frequências
pressione F10 para sair do EPITABLE
digite EXIT para voltar ao ANALYSIS

Questão 2. Para responder as questões a seguir considere como infecção ao *T. cruzi* a positividade a pelo menos 2 testes sorológicos. Qual a prevalência (IC 95%) de crianças positivas à 2 testes sorológicos? Calcule a soroprevalência e IC 95% por município de residência; existe diferença entre as prevalências observadas? Calcule a soroprevalência e a Razão de Prevalência (OR) por sexo. O que você conclui sobre a associação sexo e infecção pelo *T. cruzi*?

Notas 2: criar a variável “RES” (positividade a pelo menos 2 testes)
 LET RES=2 (resultado negativo)
 IF HAI = “P “ and IFI = “P “ THEN RES=1
 (positividade a pelo menos 2 testes)
 IF HAI = “P “ and ELISA = “P “ THEN RES=1
 IF ELISA = “P “ and IFI = “P “ THEN RES=1
 FREQ RES
 SET PERCENTS=ON
 TABLES MUN RES
 anote os resultados e repita os comandos da questão anterior para ir
 ao DOS e calcular os IC95%
 TABLES SEX RES

Questão 3. Compare as prevalências por idade. O que voce conclui quanto a soropositividade com o aumento da idade? Crie uma nova variável GRAGE agrupando-a em 2 estratos etários (7-9 e 10-12 anos) e calcule a prevalência e OR por faixa etária.

Notas 3: TABLES AGE RES
 anote os resultados
 tecele F9 para ir ao DOS
 digite STATCALC e selecione CHI SQUARE FOR TREND
 utilize os dados produzidos pela tabela anterior
 pressione F10 para sair do STATCALC
 Digite EXIT para voltar ao ANALYSIS
 LET GRAGE = AGE
 RECODE GRAGE 7-9=2 10-12=1 (agrupamento por faixa etária)
 TABLES GRAGE RES
 CLOSE (para fechar o arquivo)

Questão 4. De acordo com os resultados das Questões 2 e 3 existe associação entre exposição ao *T.cruzi* na infância e sexo e/ou idade?

2. Prevalência e fatores de risco para o vírus da hepatite B. Um estudo de corte transversal foi delineado para mensurar a prevalência de marcadores sorológicos ao vírus da hepatite B (VHB) em primodoadores de sangue e em presidiários e para avaliar fatores de risco associados à soropositividade. O arquivo **preva.rec** inclui resultados de AgHBs e anti-HBs (ELISA) de 1033 primodoadores de sangue e de 201 prisioneiros e uma coleção de 14 variáveis consideradas potencialmente de risco para aquisição da infecção pelo vírus B. Detalhes da metodologia e população de estudo encontram-se na referência **Martelli et al, 1990**. Considerou-se como exposição ao vírus B (soroprevalência) a presença de qualquer dos marcadores sorológicos (AgHBs ou anti-HBs). A análise de dados foi planejada para: (1) avaliar a prevalência dos marcadores AgHBs e anti-HBs no grupo de doadores e prisioneiros; (2) comparar as características de sexo e idade e potenciais fatores de risco entre os grupos; (3) calcular a razão de prevalência (OR) da soropositividade em relação aos grupos e variáveis potenciais de risco.

Questão 5. Qual a prevalência de portadores do VHB (AGHBS), imunes (ANTIHBBSAG) e suscetíveis (ausência de marcador) nas populações estudadas? De acordo com os resultados obtidos, a área de estudo poderia ser classificada como de alta, média ou baixa endemicidade? Construa uma tabela comparando a prevalência dos marcadores sorológicos ao vírus da hepatite B (VHB) entre os grupos de doadores e prisioneiros. O que pode ser concluído quanto às proporções de exposição ao VHB entre os grupos?

Notas 5:

```

READ PREVA.REC (para abrir o arquivo)
para as variáveis HBSAG e ANTIHBBSAG excluir código -1 (Sem
informação)
SELECT HBSAG <>-1 AND GRUPO=2
FREQ HBSAG
SELECT (para desativar a seleção)
SELECT ANTIHBBSAG <>-1 AND GRUPO=2
FREQ ANTIHBBSAG
SELECT
siga os mesmos comandos para GRUPO=1
criar a variável EXP (exposição ao vírus da hepatite B)
SELECT HBSAG<>-1 AND ANTIHBBSAG<>-1
para a variável EXP excluir os sem informação
DEFINE EXP #
IF HBSAG=1 OR ANTIHBBSAG=1 THEN EXP=1 ELSE EXP=2
SET PERCENTS=ON
TABLES EXP GRUPO
SELECT
    
```

Questão 6. Compare a média de idade entre as populações estudadas e a frequência de masculinos e femininos em cada grupo. Estratifique por faixa etária (<=29 anos; 30-39 anos; >=40 anos) as populações de prisioneiros e doadores e compare a soroprevalência inter grupos em cada estrato de idade. Analise a frequência de soropositividade por faixa etária no grupo de doadores de sangue; o que pode ser observado em relação a tendência da soropositividade com a idade e quais as possíveis explicações?

Notas 6: MEANS AGE GRUPO/N
SELECT SEX <> 9 (para excluir os sem informação)
TABLES SEX GRUPO
SELECT
criar a variável GRAGE (faixas etárias)
DEFINE GRAGE #
IF AGE >=15 AND AGE<30 THEN GRAGE=1
IF AGE >=30 AND AGE<40 THEN GRAGE=2
IF AGE >=40 THEN GRAGE=3
TABLES GRUPO EXP GRAGE
anote os resultados
tecle F9 para ir ao DOS
digite STATCALC e selecione CHI SQUARE FOR TREND
utilize os dados produzidos pela tabela anterior (GRUPO=2)
pressione F10 para sair do STATCALC
digite EXIT para voltar ao ANALYSIS

Questão 7. Compare a frequência dos fatores de risco em potencial para a exposição ao VHB (TRANSF, MEDINJ, DROG, TATO, VDRL, DST) entre os prisioneiros e doadores de sangue. As diferenças observadas entre os grupos podem ter sido pelo acaso?

Notas 7: TABLES <nome da variável> GRUPO
Ex.: SELECT TRANSF <>-1
TABLES TRANSF GRUPO
SELECT
SET PERCENTS = OFF

Questão 8. Calcule a razão de prevalência (OR) e respectivo IC 95% do encarceramento (considerado como fator de risco) em relação à soropositividade ao VHB. Utilize os doadores de sangue como grupo de referência (OR=1). De acordo com os valores obtidos os prisioneiros podem ser considerados população de risco ao VHB? Justifique. Comente um possível fator de confusão nesta análise univariada. Considerando apenas o grupo de prisioneiros, o número de anos de encarceramento (YEXP) aumenta a chance de soropositividade ao VHB?

Notas 8: TABLES GRUPO EXP

```
SELECT GRUPO=1  
TABLES YEXP EXP  
anote os resultados  
tecle F9 para ir ao DOS  
digite STATCALC e selecione CHI SQUARE FOR TREND  
utilize os dados produzidos pela tabela anterior (GRUPO=1)  
pressione F10 para sair do STATCALC  
digite EXIT para voltar ao ANALYSIS  
SELECT
```

Questão 9. Assumindo a idade como variável de confusão, ajuste o OR do encarceramento por faixa etária. O que pode ser concluído em relação ao risco de exposição ao VHB entre os grupos após a estratificação por faixa etária?

Notas 9: TABLES GRUPO EXP GRAG

Questão 10. Calcule o OR e IC 95% das mesmas variáveis potencialmente de risco testadas na questão 7. Quais as variáveis que se mostraram associadas à infecção pelo VHB? Para estas variáveis assuma o encarceramento como possível fator de confusão para a exposição ao VHB e recalcule as razões de prevalência (OR) estratificadas por prisioneiros e doadores (GRUPO). Dentre estes fatores de risco testados, quais permaneceram estatisticamente significantes após o ajuste por GRUPO?

Notas 10: TABLES <nome da variável> EXP
Ex.: SELECT TRANSF <-1
TABLES TRANSF EXP
TABLES TRANSF EXP GRUPO
siga os mesmos comandos para as demais variáveis
SELECT
CLOSE (para fechar o arquivo)

Questão 11. Considerando a dinâmica de transmissão do VHB em áreas de baixa/média endemicidade, quais as principais considerações metodológicas em relação às populações estudadas? Responda levando em conta o viés de seleção e de sobrevivência. Discuta as intervenções adotadas em saúde pública para o controle da hepatite B (triagem sorológica de doadores de sangue, vacinação universal de crianças em áreas de baixa endemicidade, vacinação de grupos de risco).

REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, A.L.S.S., ZICKER, F., LUQUETTI, A.O., OLIVEIRA, R.M., SILVA, S.A., SOUZA, J.M.P. & MARTELLI, C.M.T. Surveillance of *Trypanosoma cruzi* transmission by serological screening of schoolchildren. **WHO Bulletin**, **70(5)**:625-9, 1992.

MARTELLI, C.M.T., ANDRADE, A.L.S.S., CARDOSO, D.D.P., SOUSA, L.C.S., SILVA, S.A., SOUSA, M.A. & ZICKER, F. Soroprevalência e fatores de risco para a infecção pelo vírus da Hepatite B pelos marcadores AgHBs e Anti-HBs em prisioneiros e primodoadores de sangue. **Revista de Saúde Pública**, **24(4)**:270-6, 1990.

Arquivo: **triagem.rec**

Variável	Descrição	Código	Descrição do código
NO	Número de identificação	1 a 1990	
MUN	Município	1 2 3	Posse Simolândia Guarani
SEX	Sexo	1 2	Masculino Feminino
AGE	Idade em anos	7 a 12	
HAI	Hemaglutinação Eluato	P N	Positivo Negativo
IFI	Imunofluorescência Eluato	P N	Positivo Negativo
ELISA	Elisa Eluato	P N	Positivo Negativo

Arquivo: **preva.rec**

Variável	Descrição	Código	Descrição do código
NO	Número de identificação		
AGE	Idade em anos	15 a 71	
SEX	Sexo	1 2 9	Feminino Masculino Sem informação
TRANSF	Antecedente de transfusão de sangue	-1 1 2	Sem informação Sim Não
MEDINJ	Uso de medicação injetável	-1 1 2	Sem informação Sim Não
DROG	Antecedente de droga injetável	-1 1 2	Sem informação Sim Não
TATO	Presença de tatuagem	-1 1 2	Sem informação Sim Não
ACP	Antecedente de acupuntura	-1 1 2	Sem informação Sim Não

Variável	Descrição	Código	Descrição do código
<i>HBSAG</i>	Resultado da sorologia para AgHBs	-1 1 2	Sem informação Positivo Negativo
<i>ANTIHBASAG</i>	Resultado da sorologia para anti-AgHBs	-1 1 2	Sem informação Positivo Negativo
<i>VDRL</i>	Resultado da sorologia para VDRL	-1 1 2	Sem informação Positivo Negativo
<i>YEXP</i>	Anos de encarceramento	-1 0 1 2	Doadores de sangue menos de 1 ano 1 ano 2 anos ou mais
<i>GRUPO</i>	População de estudo	1 2	Presidiários Primodoadores de sangue
<i>DST</i>	Relato de doença sexualmente transmissível	1 2	Sim Não

AVALIAÇÃO

DE TESTES DIAGNÓSTICOS

INTRODUÇÃO	
Reprodutibilidade ou repetibilidade	22
Validade ou acurácia	22
Relação entre precisão e acurácia	22
REPRODUTIBILIDADE	
Pesquisas de Laboratório, Pesquisa Clínica e Epidemiológica	23
Avaliação de Reprodutibilidade	23
Índice <i>Kappa</i> (<i>k</i>)	23
VALIDADE DE UM TESTE DIAGNÓSTICO	
Sensibilidade e Especificidade	25
Co-positividade e co-negatividade	26
Ponto de corte para delimitar resultados positivos	26
Valor preditivo do teste	27
Valor preditivo positivo	27
Valor preditivo negativo	27
Relação entre o valor preditivo e prevalência	27
ERRO SISTEMÁTICO E ERRO ALEATÓRIO NA DETERMINAÇÃO DA SENSIBILIDADE E ESPECIFICIDADE	
Erro Aleatório	27
Cálculo do tamanho da amostra para avaliar a sensibilidade e especificidade	28
Erro Sistemático	29
Viés de amostragem	29
Viés de mensuração	29
Viés de publicação	29
Princípios básicos para avaliar um teste diagnóstico/triagem	29
ROTEIRO PARA DETERMINAR A VALIDADE DE UM TESTE	31
REFERÊNCIAS PARA LEITURA	32
EXERCÍCIOS	33
DICIONÁRIO DE BANCO DE DADOS	38

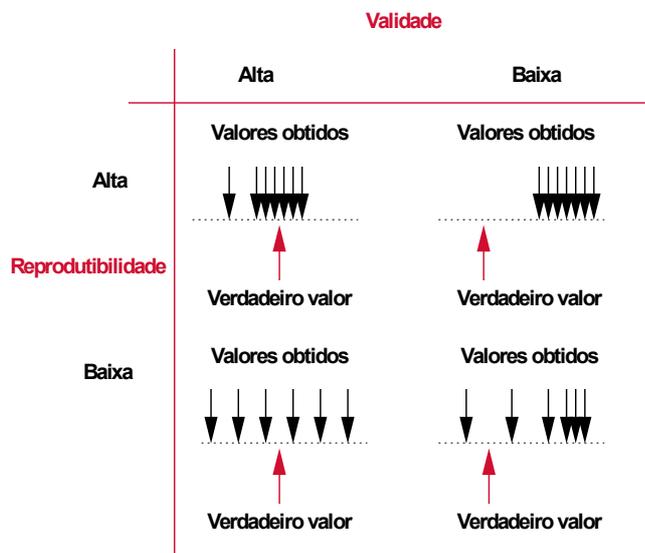
INTRODUÇÃO

A avaliação da qualidade de testes diagnósticos é um tema de interesse da investigação clínica e epidemiológica. Em pesquisa epidemiológica, "testes diagnósticos" são entendidos não apenas como exames laboratoriais, mas, também, referem-se a procedimentos diversos como interrogatório clínico, exame físico e métodos propedêuticos diversos. O desempenho de um teste diagnóstico depende da ausência de desvios da verdade (ausência de viés) e da precisão (o mesmo teste aplicado ao mesmo paciente ou amostra deve produzir os mesmos resultados): respectivamente da validade e da reprodutibilidade do "teste". Neste módulo são discutidos dois conceitos básicos da qualidade de um teste diagnóstico: reprodutibilidade e validade e os aspectos relativos ao delineamento e análise destes estudos.

. **Reprodutibilidade ou repetibilidade** é a consistência de resultados quando o exame se repete. Por exemplo, dois radiologistas que lêem de forma independente as mesmas radiografias e chegam ao mesmo diagnóstico alcançam o nível máximo de reprodutibilidade. Mas, os dois especialistas podem estar igualmente corretos ou igualmente errados em seus diagnósticos.

. **Validade ou acurácia** refere-se ao grau em que o teste ou uma estimativa baseada em um teste é capaz de determinar o verdadeiro valor do que está sendo medido. A validade informa se os resultados representam a "verdade" ou o quanto se afastam dela. Por exemplo, o ECG é um teste de maior validade, comparado à auscultação cardíaca feita com o estetoscópio, no intuito de detectar alterações cardiovasculares típicas da doença de Chagas. Um teste "dip-stick" para detecção de antígeno utilizado para diagnóstico de malária por *P. falciparum* pode ter 100% de acurácia quando for capaz de produzir resultados positivos para todas as amostras de pacientes infectados e produzir resultados negativos para os indivíduos negativos.

. **Relação entre precisão e acurácia.** A Figura abaixo mostra a relação entre o valor verdadeiro de uma medida quantitativa e o valor obtido pelo estudo em termos de baixa e alta validade e reprodutibilidade. Com baixa reprodutibilidade e estando a média dos valores obtidos pelo estudo próxima do verdadeiro valor, o teste poderá ter validade, mas, mesmo assim, terá pouca utilidade. Por outro lado, uma alta repetibilidade da medida (resultados idênticos ou próximos quando o teste diagnóstico é repetido) não assegura validade pois os valores obtidos podem estar distantes do valor verdadeiro, ou seja, podem estar errados. Como esse aspecto é fundamental para separar corretamente doentes de saudáveis, a validade e a reprodutibilidade têm de ser adequadamente mensuradas, no sentido de avaliar a qualidade de um exame diagnóstico e, conseqüentemente, a informação por ele produzida. É importante aferir ambos os parâmetros, tanto com referência a novos testes introduzidos no mercado, como testes já em uso mas, aplicados em outros contextos.



Adaptado de Beaglehole et al, 1993

REPRODUTIBILIDADE

. **Reprodutibilidade, repetibilidade ou precisão** - é a habilidade do teste em produzir resultados consistentes (quase os mesmos resultados) quando realizados independentemente e sob as mesmas condições. Por exemplo, um teste bioquímico é considerado de alta reprodutibilidade quando se obtém praticamente o mesmo resultado após várias testagens repetidas e de forma independente. Entretanto, se o aparelho eletrônico utilizado para realização do teste não estiver adequadamente calibrado, o teste pode ter alta reprodutibilidade, mas, produzir resultados consistentemente errados. O mesmo conceito de reprodutibilidade pode ser usado em situações mais gerais, como por exemplo, comparando-se os resultados de lâminas em diferentes ocasiões (variabilidade intra-observador).

. Pesquisas de Laboratório, Pesquisa Clínica e Epidemiológica

Melhores resultados de reprodutibilidade são geralmente obtidos no trabalho de laboratório, onde as condições de operação podem ser mais controladas (um só observador, aparelhos de alta precisão, calibrados, com pouco uso, uso de amostras controle, ambiente livre de maiores perturbações e horário apropriado). Por outro lado, em pesquisas clínicas e epidemiológicas, raramente obtém-se o nível de reprodutibilidade encontrado em investigações de laboratório. O diagnóstico clínico, por exemplo, é um processo subjetivo, e, por isto, suscetível a interpretações discordantes, mesmo entre clínicos competentes e experientes. Em geral, um nível baixo de reprodutibilidade tende a atenuar as verdadeiras correlações entre eventos. Isto limita a utilidade do diagnóstico clínico em pesquisas populacionais, pois prejudica a investigação de associações entre fatores de risco e danos à saúde.

. Avaliação de Reprodutibilidade

Há diversas maneiras de verificar a concordância de resultados entre leituras de um mesmo evento ou comparar métodos diagnósticos diferentes, e assim, estimar o erro cometido na sua aferição. Os resultados podem ser expressos sob forma de variável dicotômica (positivo / negativo), categórica (normal / anormal / níveis limítrofes), em medidas contínuas (miligramas, mililitros) ou títulos de sorologia. Este é um dos aspectos que influencia a forma de análise dos resultados. Geralmente, independente do tipo de dado produzido pelos testes diagnósticos, os médicos/epidemiologistas tendem a reduzi-lo à variáveis dicotômicas ou expressas em categorias para tornar a interpretação mais útil na prática. A comparação dos resultados pode ser apresentada através da taxa global de concordância entre os examinadores ou pelo indicador *Kappa*.

. **Índice Kappa (k)** - Uma maneira muito utilizada para expressar a confiabilidade de um teste é através do índice k que constitui um avanço em relação à taxa geral de concordância, por ser um indicador de concordância ajustada, pois leva em consideração, a concordância devida à chance. O k informa a proporção de concordância não aleatória (além da esperada pela chance) entre observadores ou medidas da mesma variável categórica, e seu valor varia de "menos 1" (completo desacordo) a "mais 1" (concordância total). Se a medida concorda mais frequentemente do que seria esperado pela chance, então o índice k é positivo; se a concordância é completa $k = 1$. Zero indica o mesmo que leituras feitas ao acaso. A Tabela 1 apresenta os valores do k e respectivas interpretações.

Tabela 1- Escala de concordância do *Kappa*

Kappa	Concordância
<0,00	Nenhuma
0,00-0,20	Fraca
0,21-0,40	Sofrível
0,41-0,60	Regular
0,61-0,80	Boa
0,81-0,99	Ótima
1,00	Perfeita

Adaptado de Landis & Koch, Biometrics, 1977

A Tabela 2 exemplifica o cálculo de *k*. Cento e vinte lâminas contendo esfregaços de gota espessa de sangue para pesquisa de hematozoários (malária) foram preparadas em condições uniformes e interpretadas por dois microscopistas independentes. O primeiro identificou 20 lâminas positivas e 100 negativas enquanto o segundo diagnosticou respectivamente 30 e 90 gerando 106 resultados concordantes (18+88) e 14 de discordantes (2+12). A taxa geral de concordância foi de 88,3% (106/120) e o valor de *k* = 65%.

Tabela 2 – Concordância entre dois observadores nas leituras de laminas para pesquisa de hematozoários.

	Microscopista 1		Total
	(+)	(-)	
Microscopista 2			
(+)	18 (a)	12 (b)	30
(-)	2 (c)	88 (d)	90
Total	20	100	120

O índice *kappa* é estimado como:

$$K = \frac{Po - Pe}{1 - Pe}$$

sendo:

- Po = Proporção de concordâncias observadas
 Pe = Proporção de concordâncias esperadas

$$P_o = \frac{a + d}{a + b + c + d}$$

$$P_e = \frac{[(a + b)(a + c)] + [(c + d)(b + d)]}{(a + b + c + d)^2}$$

Para interpretação do k deve-se levar em conta:

tipo de evento e outros fatores - o nível de concordância depende do tipo de evento, fatores relacionados ao examinador, ao procedimento sendo testado e ao ambiente onde as observações são realizadas. Também, a diminuição do número de categorias de resultados (valores positivos e negativos ao invés de valor alto, médio, baixo e muito baixo) tende a aumentar a concordância.

prevalência - a prevalência do diagnóstico ou evento na população, afeta o resultado final. Baixas prevalências tendem à estar associadas a baixos níveis de reprodutibilidade, pois o valor de k depende da concordância devida ao acaso. É possível encontrar-se baixos níveis de reprodutibilidade, devido à baixa prevalência do evento e não à erros relacionados ao procedimento diagnóstico empregado. Por este motivo, deve-se informar a prevalência juntamente com os resultados do k .

independência da avaliação - as avaliações devem ser independentes umas das outras, princípio também aplicável à verificação da validade/accurácia. Isto significa que quando um examinador repete o teste, deve ignorar resultados prévios, obtidos por ele ou por outro examinador, para evitar a possibilidade de ser influenciado por este conhecimento e prejudicar a avaliação, mesmo involuntariamente.

VALIDADE DE UM TESTE DIAGNÓSTICO

A validade de um teste refere-se à quanto, em termos quantitativos ou qualitativos, um teste é útil para diagnosticar um evento (validade simultânea ou concorrente) ou para predizê-lo (validade preditiva). Para determinar a validade, compara-se os resultados do teste com os de um padrão (padrão ouro): esse pode ser o verdadeiro estado do paciente, se a informação está disponível, um conjunto de exames julgados mais adequados, ou uma outra forma de diagnóstico que sirva de referência. O teste diagnóstico ideal deveria fornecer, sempre, a resposta correta, ou seja, um resultado positivo nos indivíduos com a doença e um resultado negativo nos indivíduos sem a doença. Além do que, deveria ser um teste rápido de ser executado, seguro, simples, inócuo, confiável e de baixo custo.

. Sensibilidade e Especificidade

Para definir os conceitos de sensibilidade e especificidade, serão utilizados como exemplos, testes com resultados dicotômicos, isto é, resultados expressos em duas categorias: positivos ou negativos.

A Tabela 3 mostra as relações entre os resultados de um teste e o diagnóstico verdadeiro. O teste é considerado positivo (anormal) ou negativo (normal), e a doença presente ou ausente. Assim, na avaliação de um teste diagnóstico existem 4 interpretações possíveis para o resultado do teste: duas em que o teste está correto e duas em que está incorreto. O teste está correto quando ele é positivo na presença da doença (resultados verdadeiros positivos), ou negativo na ausência da doença (resultados verdadeiros negativos). Por outro lado, o teste está incorreto quando ele é positivo na ausência da doença (falso positivo), ou negativo quando a doença está presente (falso negativo). Os melhores testes diagnósticos são aqueles com poucos resultados falso-positivos e falso-negativos.

Tabela 3. Validade de um Teste Diagnóstico

		Doença (padrão-ouro)		
		Presente	Ausente	
Teste	Positivo	Verdadeiro positivo a	Falso-positivo b	a + b
	Negativo	Falso-negativo c	Verdadeiro negativo d	c + d
		a + c	b + d	N(a+b+c+d)

As seguintes proposições/Indicadores podem ser calculados da comparação dos resultados da tabela:

Sensibilidade: $a/(a+c)$

Especificidade: $d/(b+d)$

Prevalência (real): $(a+c)/N$

Prevalência estimada (teste): $(a+b)/N$

Valor preditivo positivo: $a/(a+b)$

Valor preditivo negativo: $d/(c+d)$

Classificação correta (acurácia): $(a+d)/N$

Classificação incorreta: $(b+c)/N$

. **Sensibilidade** - é a capacidade que o teste diagnóstico/triagem apresenta de detectar os indivíduos verdadeiramente positivos, ou seja, de diagnosticar corretamente os doentes.

. **Especificidade** - é a capacidade que o teste diagnóstico/triagem tem de detectar os verdadeiros negativos, isto é, de diagnosticar corretamente os indivíduos saudáveis.

. **Co-positividade e co-negatividade** - são termos utilizados em substituição, respectivamente, à sensibilidade e especificidade, quando o padrão empregado é outro teste considerado de referência para a doença em questão e não os diagnósticos de certeza de presença ou ausência de doença. São também designados sensibilidade relativa e especificidade relativa.

. **Ponto de corte para delimitar resultados positivos** - O teste ideal, com 100% de sensibilidade e especificidade raramente existe na prática, pois a tentativa de melhorar a sensibilidade frequentemente tem o efeito de diminuir a especificidade. Em algumas situações clínicas os resultados são obtidos através de variáveis contínuas, não havendo uma separação clara e inquestionável entre o que é "normal" e "anormal". Para a definição do ponto de corte de positividade o investigador deverá levar em conta a importância relativa da sensibilidade e especificidade do teste diagnóstico, ponderando sobre as implicações dos dois possíveis erros. Em indicações de procedimentos de risco (certas cirurgias), por exemplo, deve-se evitar resultados

falso-positivos; nestes casos, o ponto de corte deve ser definido de tal forma que aumente a especificidade do teste. Por outro lado, em triagens sorológicas em bancos de sangue para prevenção de transmissão de infecções nas quais a não detecção de casos acarretará risco para a população, o ponto de corte deverá ser estabelecido tendo como objetivo alcançar 100% de sensibilidade do teste para que não ocorram resultados falso-negativos, em que pese o aumento da proporção de falso-positivos. Para aumentar a sensibilidade em uma triagem pode-se utilizar mais do que um teste diagnóstico, em paralelo considerando-se como positivo as amostras que apresentarem pelo menos uma reação positiva. Em inquéritos populacionais, testes com alta sensibilidade devem ser utilizados, quando a prevalência da infecção na população em geral for baixa. Por outro lado, em clínica, é comum realizarem-se testes em série. Testes adicionais são realizados para confirmar resultados positivos ou negativos previamente obtidos.

. **Valor preditivo do teste** - No contexto epidemiológico e clínico, a validade de um marcador sorológico diz respeito à extensão com que ele pode prever a ocorrência da doença / infecção. Nessas circunstâncias, devemos estar preparados para responder à seguinte questão: dado que o teste apresentou resultado positivo (ou negativo), qual a probabilidade do indivíduo ser realmente doente (ou sadio)? Esse atributo do teste é conhecido como Valor Preditivo (VP) podendo ser positivo (VPP) ou negativo (VPN), e é determinado pela interação de três variáveis: a sensibilidade e a especificidade do teste e a prevalência da doença no grupo de estudo.

Valor preditivo positivo - é a proporção de doentes entre os positivos pelo teste. No exemplo da Tabela 2 teríamos 60% (18/30), o que equivale a dizer que em cada 10 testes positivos, 6 indivíduos seriam realmente doentes.

Valor preditivo negativo - é a proporção de sadios (sem a doença) entre os negativos ao teste. Ainda em relação à Tabela 2, teríamos um VPN de 98% (88/90); a cada 100 testes negativos, 98 seriam sadios.

. **Relação entre valor preditivo e prevalência**

Enquanto a sensibilidade e especificidade de um teste são propriedades inerentes ao teste e não variam a não ser por erro técnico, os VPs dependem da prevalência da doença na população de estudo. O VPP aumenta com a prevalência enquanto os VPN diminuem. Assim, quando a doença é rara o VPP é baixo, pois a maior parte dos exames positivos pertencem a sadios, representando resultados falso-positivos. Por outro lado, O VPN é alto em baixas prevalências. Os resultados falso-positivos e falso-negativos podem ser minimizados utilizando-se a combinação de testes, em paralelo (dois ou mais testes realizados simultaneamente) ou em série (dois ou mais testes realizados em seqüência), para a definição de resultado positivo. Se a intenção é reduzir resultados falso positivos (e aumentar a especificidade), um diagnóstico positivo deverá ser confirmado somente quando pelo menos dois testes diferentes forem positivos. Por outro lado, para reduzir resultados falso negativos (e aumentar a sensibilidade), um único teste positivo seria suficiente para considerar um diagnóstico positivo. Por exemplo, o teste será positivo se os 2 testes forem positivos, ou negativo se os 2 forem negativos.

ERRO SISTEMÁTICO E ERRO ALEATÓRIO NA DETERMINAÇÃO DA SENSIBILIDADE E ESPECIFICIDADE

Erro Aleatório - Os estudos de avaliação de testes diagnósticos estão sujeitos a erros ao acaso; alguns pacientes com a doença apresentarão resultado normal do teste diagnóstico. Este tipo de erro pode ser avaliado calculando-se o intervalo de confiança para a sensibilidade e especificidade do novo teste. O intervalo de confiança indica o espectro de variação dos resultados obtidos para que se possa compará-los com os testes convencionais. Por exemplo, considere a comparação de dois testes A e B em uso na prática clínica; o teste A com sensibilidade de 80% e especificidade de 85%, calculadas após testagem em centenas de indivíduos e o novo teste diagnóstico (B) foi positivo em

10 indivíduos em um total de 10 pacientes com a doença, o que equivale a 100% de sensibilidade, e foi negativo em 9 de 10 indivíduos sem a doença (especificidade=90%). Apesar da sensibilidade e especificidade deste novo teste (b) serem maiores que as descritas para o teste (a), o intervalo de confiança de 95% tanto para a sensibilidade (61%-100%) como para a especificidade (55%-97%) mostrou um espectro grande de variação, com sobreposição destes intervalos com os do teste convencional; este fato decorre do pequeno número de indivíduos testados. Por este motivo, não é possível concluir que o novo teste (b) tenha um melhor desempenho do que o teste convencional (a).

Uma das estratégias para minimizar os erros aleatórios é estimar o tamanho da amostra para determinar a validade do teste diagnóstico, baseado na construção de intervalos de confiança, definindo-se espectros que incluam o valor que se deseja obter para a sensibilidade e especificidade do teste. Isto significa calcular dois tamanhos da amostra: um para a sensibilidade do teste e outro para a especificidade.

Cálculo do tamanho da amostra para avaliar a sensibilidade e especificidade

O cálculo do tamanho da amostra, para variáveis dicotômicas, segue os mesmos princípios estabelecidos em estudos descritivos/estudos de prevalência (Quadro abaixo), sendo necessária as seguintes informações:

- (1) estimativa da proporção esperada da positividade na população (quando maior de 50% utilize a proporção de pessoas com resultados negativos)
- (2) amplitude do intervalo de confiança que se deseja
- (3) definição do intervalo de confiança (geralmente 95%)

$$N = Z * Z (P (1-P)) / (D * D) \quad \text{Onde:}$$

P= proporção esperada
D= semi-amplitude do intervalo de confiança
Z = 1,96 (para $\alpha=0,05$ e IC 95%)

Por exemplo, em um estudo para determinar a sensibilidade de um novo teste diagnóstico para malária, espera-se que 80% dos pacientes com malária tenham teste positivo (resultado de estudo piloto). Quantos indivíduos com malária deverão ser testados para se estimar uma sensibilidade do teste de 80% com intervalo de 95% de confiança e precisão do teste de 0,04?. Considerando as 3 informações necessárias para o cálculo do tamanho da amostra, teríamos:

- (1) proporção esperada de casos com malária com teste positivo = 0,20 (80% é maior que 50%; portanto, a proporção de indivíduos com malária e teste negativo é 20%)
- (2) espectro do intervalo de confiança = 0,08 Utilize a semi-amplitude (0,04 acima ou 0,04 abaixo) como o erro máximo aceitável
- (3) intervalo de confiança = 95%

Utilizando-se a fórmula anexa, seria necessário aplicar o teste em 384 pacientes com malária para se estimar uma sensibilidade de 80% do teste com intervalo de 95% de confiança de 76%-84%:

$$n = 1,96^2 (0,20 (1-0,20)) / (0,04^2)$$

n = 384 pessoas

Os mesmos procedimentos são válidos para o cálculo do tamanho da amostra para determinar a especificidade do teste. Por exemplo, se o investigador espera que 90% dos indivíduos sem malária tenham teste negativo, 216 indivíduos sem malária deveriam ser incluídos no estudo para determinação de uma especificidade de 90% \pm 0,04 com um intervalo de 95% de confiança.

Erro Sistemático - De forma geral, os estudos de testes diagnósticos estão sujeitos aos mesmos vieses que os estudos observacionais; os mais comuns são os vieses de amostragem, de medida do teste e de relato dos resultados.

. **Viés de amostragem** - neste tipo de erro, a amostra de estudo não é representativa da população alvo na qual o teste deverá ser utilizado. Por exemplo, a seleção de indivíduos provenientes de serviços de referência tende a incluir pessoas com formas graves da doença ou pacientes nos quais os testes sejam mais anormais do que seriam em outras formas clínicas da doença. Isto faz com que o estudo forneça resultados superestimados da sensibilidade do teste, diferindo da sensibilidade do teste em condições de rotina. Da mesma forma, o estudo fornecerá uma especificidade aumentada do teste, se indivíduos sem a doença forem selecionados como voluntários, pois estes tendem a ser mais saudáveis do que indivíduos com sintomas, recrutados de ambulatórios, porém sem a doença. A estratégia utilizada para minimizar este tipo de erro é selecionar amostras de população semelhante a qual o teste deverá ser utilizado. A escolha de amostras de populações nas quais a prevalência da doença é maior do que a habitualmente detectada, trará como consequência valores preditivos positivos superestimados. Uma situação muito comum é investigar um número igual de indivíduos com a doença e sem a doença, o que equivale a 50% de prevalência da doença. Para lidar com este viés o estudo deveria fornecer, também, resultados dos valores preditivos do teste ajustados para outras probabilidades de doença, para que o leitor possa avaliar a utilidade do teste de acordo com sua realidade na prática clínica.

. **Viés de mensuração** - Sempre que possível o investigador deve desconhecer quais indivíduos têm a doença e quais não têm, para evitar vícios de interpretação de resultados, especialmente nas situações limítrofes. Da mesma forma, o investigador deve permanecer mascarado em relação à realização dos testes diagnósticos. O ponto de corte deve ser definido antes da realização do teste.

. **Viés de publicação** - Existe uma tendência em se publicar somente os estudos que mostrem "sucesso" dos testes diagnósticos o que acarreta um bias de literatura. Para minimizar este viés, os estudos devem ser planejados com número suficiente de indivíduos para que os resultados tenham credibilidade e sejam devidamente divulgados.

. **Princípios básicos para avaliar um teste diagnóstico/triagem**

O delineamento de estudos para avaliar/comparar a utilidade clínica ou populacional de testes diagnósticos deve incorporar dois aspectos. O primeiro deles diz respeito aos princípios da aleatorização e mascaramento. Se os pacientes são alocados aleatoriamente para receber o novo teste (versus aquele usado na rotina), os indivíduos que receberem este novo teste terão uma melhor evolução clínica? A comparação dos testes pode e deve ser feita nos mesmos indivíduos e amostras, para eliminar variações externas aos testes. O segundo aspecto a ser levado em conta refere-se à prática clínica vigente. O teste será aplicado nas mesmas condições de seu uso na clínica? O fato de um teste discriminar casos graves da doença não significa que será igualmente útil para distinguir pacientes portadores de doença leve dos demais pacientes com sintomas semelhantes.

Estudos conduzidos para determinar a validade de um teste diagnóstico apresentam estrutura semelhante aos estudos observacionais. Eles incluem a variável preditora (resultado do teste) e a variável de efeito (presença ou ausência da doença). A diferença entre eles reside nos seus objetivos. Na avaliação de testes diagnósticos, descreve-se a **intensidade da associação**, em termos de sensibilidade e especificidade (capacidade do teste em discriminar doentes de não doentes). Já, os estudos observacionais buscam determinar a presença de uma associação. Portanto, na análise da validade de um teste não basta apenas mostrar que existe uma associação entre o resultado do teste e a doença.

ROTEIRO PARA DETERMINAR A VALIDADE DE UM TESTE

① *Certifique-se da necessidade do teste*

- . vantagens do novo teste em relação aos existentes
- . benefícios para pacientes, com a introdução do novo teste
- . custos para aplicação do teste em nível individual e em saúde pública

② *Estabeleça o critério de amostragem*

- . defina a população de referência e a população de estudo
- . esclareça a fonte de seleção dos participantes
- . informe sobre a inclusão de casos graves, moderados e leves

③ *Descreva o teste e o padrão de referência*

- . produto químico, imunobiológico, antígeno, anticorpo, procedência
- . etapas para processamento das reações
- . interpretação e categorização dos parâmetros a serem avaliados

④ *Descreva os procedimentos para a aplicação do teste e do padrão de referência*

- . aplicação dos testes de forma mascarada
- . codificação das amostras para envio ao laboratório - processamento sem conhecimento do status doente/não doente

⑤ *Calcule o tamanho da amostra*

- . estipule o número mínimo e suficiente de participantes para se estimar a sensibilidade e especificidade do teste com intervalo de 95% de confiança
- . estime o número de casos disponíveis ou a serem detectados no futuro próximo no local de seleção dos participantes

⑥ *Esclareça as questões éticas*

- . riscos da aplicação do teste e benefício da detecção de indivíduos positivos
- . atenção médica aos indivíduos positivos
- . confidencialidade dos resultados

⑦ *Análise de dados*

- . apresente os resultados em termos de sensibilidade, especificidade e valores preditivos com respectivos IC95%

Adaptado de Hulley & Cummings, 1988

REFERÊNCIAS PARA LEITURA

BEAGLEHOLE, R., BONITA, R. & KJELLSTRÖM, T. *Basic Epidemiology*. World Health Organization, Geneva, 1993.

BUCK, A.A. & GART, J.J. Comparison of a screening test and a reference test in epidemiologic studies. I. Indices of agreement and their relation to prevalence. **American Journal of Epidemiology**,**83**:586-92, 1966.

FLEISS, J.L. *Statistical methods for rates and proportions*, 2nd ed. New York, John Wiley & Sons, 1981.

FLETCHER, R.M.; FLETCHER, S.W. & WAGNER, E.H. *Clinical Epidemiology, the essentials*. Baltimore - USA, Ed. Wawerly, 1983.

GALEN, R.S. & GAMBINO, S.R. *Beyond normality: the predictive value and efficiency of medical diagnosis*, New York:John Wiley & Sons ed., 1975.

HULLEY, S.B. & CUMMINGS, S.R. *Designing clinical research*. Williams & Wilkins Baltimore, 1988.

KRAEMER, H.C. & BLOCH, D.A. Kappa coefficients in epidemiology: an appraisal of a reappraisal. **Journal of Clinical**,**41**:959-68, 1988.

PEREIRA, M.G. *Epidemiologia: Teoria e Prática*. Rio de Janeiro. Guanabara Koogan, 1995.

EXERCÍCIOS

Arquivos:

1. ecg.rec
2. malmt.rec
3. bloodch.rec

1. Concordância entre leituras de eletrocardiogramas - o arquivo **ecg.rec** listado em anexo contém os resultados de 100 Eletrocardiogramas que foram enviados para leitura de forma independente para dois centros.

Questão 1. Calcule a reprodutibilidade (Kappa) do diagnóstico de alteração eletrocardiográfica feito por 2 observadores (A e B). Interprete os resultados.

Notas 1:

- READ ECG.REC (para abrir o arquivo)
- TABLES CENTROA CENTROB
- anote os resultados
- tecle F9 para ir ao DOS
- digite EPITABLE para cálculo do Kappa (comando COMPARE / PROPORTIONS)
- utilize os dados da tabela anterior
- pressione F10 para sair do EPITABLE
- digite EXIT para voltar ao ANALYSIS
- CLOSE (para fechar o arquivo)

2. Avaliação do diagnóstico parasitológico da malária - o arquivo **malmt.rec** listado em anexo contém resultados da parasitemia de 141 indivíduos com febre, da demanda espontânea de um dia de atendimento de 2 Postos de Notificação de malária (Fundação Nacional de Saúde). Todas as lâminas foram encaminhadas de forma independente para o Centro de Controle de Qualidade (Ministério da Saúde) e os resultados comparados aos fornecidos pelos Postos de Notificação (serviços locais de saúde). A análise de dados foi planejada para avaliar a reprodutibilidade do diagnóstico de malária entre o Centro de Controle de Qualidade e os serviços locais de saúde e avaliar a sensibilidade destes serviços no diagnóstico de malária. Responda as questões abaixo utilizando o arquivo malmt.rec.

Questão 2. Qual a prevalência global (IC 95%) de malária de acordo com os resultados fornecidos pelo Centro de Controle de Qualidade? Entre os casos com malária qual a frequência por espécie?

Notas 2: READ MALMT.REC (para abrir o arquivo)
FREQ LAMCQ /CI
selecione apenas os indivíduos com malária
SELECT LAMCQ=2
FREQ DIAGCQ
SELECT (para desativar a seleção)

Questão 3. Qual a prevalência global (IC 95%) de malária de acordo com os resultados fornecidos pelos Postos de Notificação da FNS? Qual a frequência de diagnóstico por espécie dentre os casos com malária?

Notas 3: FREQ LAMFNS /CI
selecione apenas os indivíduos com malária
SELECT LAMFNS=2
FREQ DIAGFNS
SELECT (para desativar a seleção)

Questão 4. Compare a proporção de malária por espécie utilizando os resultados fornecidos pelos Postos de Notificação (FNS) e o Centro de Controle de Qualidade (questões 2 e 3). Quais as implicações destes achados do ponto de vista do Programa de Controle da Malária?

Questão 5. Qual a reprodutibilidade (*Kappa*) do diagnóstico de malária entre os Postos de Notificação (FNS) e o Centro de Controle de Qualidade?

Notas 5: TABLES LAMFNS LAMCQ
anote os resultados
tecle F9 para ir ao DOS
digite EPITABLE para cálculo do *Kappa* (comando
COMPARE / PROPORTION)
pressione ESC para voltar ao menu do EPITABLE

Questão 6. Qual a sensibilidade (IC 95%) dos Postos de Notificação (FNS) no diagnóstico da malária? Qual o total de casos falso negativos por espécie? Assuma como padrão de referência os resultados do Controle de Qualidade.

Notas 6: utilize os dados da tabela produzida na questão 5
STUDY/ SCREENING
pressione F10 para sair do EPITABLE
digite EXIT para voltar ao ANALYSIS
TABLES DIAGFNS DIAGCQ

Questão 7. Qual o valor preditivo positivo (IC 95%) da febre no diagnóstico de malária nestes Postos de Notificação? Interprete o custo-benefício da indicação do tratamento presuntivo de malária nesta área endêmica.

Nota 7: FREQ LAMCQ
 CLOSE (para fechar o arquivo)

Questão 8. Discuta as implicações dos resultados desta investigação considerando a Estratégia Global de Controle da Malária - diagnóstico precoce e tratamento oportuno. Que estratégias você recomendaria para implementação da qualidade do diagnóstico de malária na área de estudo?

3. Validação da sorologia para Doença de Chagas em bancos de sangue - o arquivo **bloodch.rec** listado em anexo contém resultados de sorologia para infecção pelo *T.cruzi* de 1513 primodoadores de sangue triados nos 6 bancos de sangue da cidade de Goiânia (período 1988-1989) pelas técnicas de Hemaglutinação (HA) e Fixação de Complemento (FC). Amostras destes soros foram encaminhadas de forma independente a um dos Laboratórios de Referência para Doença de Chagas da OMS e os resultados comparados aos fornecidos pelos bancos de sangue (serviços locais de saúde). Detalhes da metodologia estão na referência **Andrade et al., 1992**. A análise teve como objetivo: (1) aferir a sensibilidade dos bancos de sangue na prevenção da transmissão transfusional da doença de Chagas e (2) avaliar a concordância do diagnóstico sorológico da infecção pelo *T.cruzi* em condições ideais (Laboratório de Referência) e na rotina dos bancos de sangue. Utilize o arquivo **bloodch.rec** para responder as questões abaixo.

Questão 9. Compare a prevalência de soropositividade ao *T.cruzi* pelas técnicas HA e IF realizadas pelo Laboratório de Referência. Avalie o benefício da utilização da triagem em paralelo comparada ao uso de apenas uma das técnicas.

Notas 9: READ BLOODCH.REC (para abrir o arquivo)
criar as variáveis GRHA ($HA \geq 16$) e GRIF ($IF \geq 40$)
DEFINE GRHA __ (grupo hemaglutinação)
IF HA=-1 THEN GRHA="SS" (sem sorologia)
IF HA>-1 AND HA<16 THEN GRHA="N" (negativo)
IF HA>=16 THEN GRHA="R" (reativo)
utilize o mesmo procedimento para criar a variável GRIF
utilize somente os registros com resultado de GRHA
SELECT GRHA <> "SS"
FREQ GRHA
SELECT (para desativar a seleção)
utilize somente os registros com resultado de GRIF
SELECT GRIF <> "SS"
FREQ GRIF
SELECT (para desativar a seleção)
TABLES GRHA GRIF

Questão 10. Calcule a concordância (*Kappa*) do diagnóstico de infecção pelo *T.cruzi* por 2 testes distintos (HA e IF), realizados no laboratório de referência. Interprete os resultados.

Notas 10: utilize somente os pares disponíveis para ambas as técnicas
SELECT GRHA <> "SS" and GRIF <> "SS"
TABLES GRHA GRIF
SELECT
anote os resultados
tecle F9 para ir ao DOS
digite EPITABLE para cálculo do *Kappa* (comando COMPARE / PROPORTION)
utilize os dados da tabela anterior
pressione F10 para sair do EPITABLE
digite EXIT para voltar ao ANALYSIS

Questão 11. Calcule a concordância do diagnóstico sorológico da infecção entre o laboratório de referência e os bancos de sangue. Para esta finalidade crie 2 novas variáveis com os resultados do laboratório de referência ("RES"=sorologia referência) e dos bancos de sangue ("POS"=resultado na triagem).

Notas 11: criar a variável RES (GRIF=R ou GRHA=R)
DEFINE RES __
IF GRHA="SS" AND GRIF="SS" THEN RES="SS" ELSE RES="N"
(recodificação para eliminar registros sem resultado)
IF GRHA="R" OR GRIF="R" THEN RES="R"
criar a variável POS (HABS=R ou FCBS=R)
DEFINE POS __
IF HABS="SS" AND FCBS="SS" THEN POS="SS" ELSE POS="N"
IF HABS="R" OR FCBS="R" THEN POS="R"
utilize somente os pares disponíveis para ambos, laboratório de
referência e bancos de sangue
SELECT RES <> "SS" AND POS <> "SS"
TABLES POS RES
SELECT
anote os resultados e organize a tabela
tecle F9 para ir ao DOS
digite EPITABLE para cálculo do Kappa (comando COMPARE /
PROPORTION)
utilize os dados da tabela anterior
pressione F10 para sair do EPITABLE
digite EXIT para voltar ao ANALYSIS

Questão 12. Calcule a sensibilidade e valores preditivos positivos e negativos do diagnóstico de infecção chagásica fornecido pelos Bancos de Sangue ("POS"). Para esta finalidade construa uma nova variável ("RESFIM") assumindo como padrão a positividade simultânea à HA e IF pelo Laboratório de Referência. Quantos indivíduos soropositivos deixariam de ser detectados pelos Bancos de Sangue na triagem de rotina de primodadores (falso negativos)? Dos doadores soropositivos encaminhados pelos Bancos de Sangue ao Ambulatório de Atenção ao Chagásico, qual o percentual de doadores que não teriam confirmação diagnóstica de doença de Chagas (responda considerando o VPP)?

Notas 12: criar a variável RESFIM (GRIF=R e GRHA=R)
DEFINE RESFIM __
IF GRIF="R" AND GRHA="R" THEN RESFIM="R"
IF GRIF="N" OR GRHA="N" THEN RESFIM="N"
IF GRIF="SS" AND GRHA="SS" THEN RESFIM="SS"
utilize somente os pares disponíveis para ambos, Laboratório
de Referência e Bancos de Sangue
SELECT POS <> "SS" AND RESFIM <> "SS"
TABLES POS RESFIM
SELECT
anote os resultados
tecle F9 para ir ao DOS
utilizar EPITABLE para cálculo da sensibilidade (comando
STUDY / SCREENING)
pressione F10 para sair do EPITABLE
digite EXIT para voltar ao ANALYSIS

Questão 13. Qual o número de indivíduos soropositivos detectados pela HA ("HABS") e pela FC ("FCBS") pelos Bancos de Sangue? Qual dos testes seria mais indicado para finalidade de triagem em Bancos de Sangue? Avalie o benefício da utilização da triagem em paralelo pelos Bancos de Sangue comparado ao uso de apenas uma técnica.

Notas 13: utilize somente os pares disponíveis para ambas as técnicas, HA e FC.
SELECT HABS <> "SS" AND FCBS <> "SS"
excluir BS6 pois os resultados estão disponíveis apenas para a HA
SELECT BS<>6
TABLES HABS FCBS
SELECT (para desativar a seleção)
CLOSE (para fechar o arquivo)

REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, A.L.S.S., MARTELLI, C.M.T., LUQUETTI, A.O., OLIVEIRA, O.S.O., SILVA, S.A. & ZICKER, F. Serologic screening for *Trypanosoma cruzi* among blood donors in Central Brazil. **Bulletin of the Pan American Health Organization**,26(2):157-164, 1992.

ANDRADE, A.L.S.S., MARTELLI, C.M.T., OLIVEIRA, R.M., ARIAS, J.R., ZICKER, F. & PANG, L. High prevalence of asymptomatic malaria in gold mining areas in Brazil. **Clinical Infectious Disease**,20(2): 475, 1995.

Arquivo: **ecg.rec**

Variável	Descrição	Código	Descrição do Código
IDENTIFICA	Número de identificação		
CENTROA	Leitura ECG pelo observador A	0 1	Sem alteração Com alteração
CENTROB	Leitura ECG pelo observador B	0 1	Sem alteração Com alteração

Arquivo: **malmt.rec**

Variável	Descrição	Código	Descrição do Código
NO	Número de identificação		
LAMFNS	Resultado Lâmina pela FNS	1 2	Negativo Positivo
LAMCQ	Resultado Lâmina pelo Controle Qualidade (MS)	1 2	Negativo Positivo
DIAGCQ	Diagnóstico pelo Controle Qualidade	1 2 3 4 5	<i>P.vivax</i> <i>P.falciparum</i> <i>P.malariae</i> mista Sem malária
DIAGFNS	Diagnóstico pela FNS	1 2 3 4 5	<i>P.vivax</i> <i>P.falciparum</i> <i>P.malariae</i> mista Sem malária

Arquivo: **bloodch.rec**

Variável	Descrição	Código	Descrição do Código
NO	Número de Identificação		
BS	Bancos de Sangue	1 2 3 4 5 6	Hemog Araújo Jorge Hemolabor Inst Hemoterapia Banco de Sangue Goiano Hospital das Clínicas
HABS	Hemaglutinação Banco de Sangue	R N SS	Reator Negativo Sem sorologia
FCBS	Fixação Complemento Banco de Sangue	R N SS	Reator Negativo Sem Sorologia
HA	Hemaglutinação Laboratório de Referência	-1 0 - 8 16 - 128	Sem Sorologia Negativo Positivo
IF	Imunofluorescência Laboratório de Referência	-1 0 40-1280	Sem Sorologia Negativo Positivo
ELISA	ELISA Laboratório de Referência	-1 0 - 0,9 1,0 - 3,5	Sem Sorologia Negativo Positivo

ESTUDOS DE CASO-CONTROLE

INTRODUÇÃO	42
SELEÇÃO DE PARTICIPANTES	
Seleção de casos	42
Seleção de controles	42
TIPOS DE ESTUDOS	
Caso-controle de base populacional	42
Caso-controle aninhado	43
MEDIDA DE ASSOCIAÇÃO	
Análise univariada	43
METODOLOGIA	
Amostragem	44
Tamanho da amostra	44
Medida da exposição	45
Tipos de viés	45
ANÁLISE ESTRATIFICADA	
Confusão	46
Estratificação	46
Modificação de efeito	47
EMPARELHAMENTO	49
VANTAGENS E LIMITAÇÕES	50
ROTEIRO PARA DELINEAR UM ESTUDO DE CASO-CONTROLE	52
REFERÊNCIAS PARA LEITURA	53
EXERCÍCIOS	54
DICIONÁRIO DE BANCO DE DADOS	60

INTRODUÇÃO

É uma pesquisa, de modelo retrospectivo, na qual os participantes são selecionados entre indivíduos que já têm a doença (**casos**) e entre indivíduos que não a têm (**controles**); em cada um desses dois grupos, verifica-se o número de indivíduos expostos, a algum fator de risco. O objetivo é verificar a possível existência de associação causal entre a exposição aos fatores de risco e a doença em estudo. Se o fator está associado à doença, a proporção do fator entre os casos será maior que a mesma proporção entre os controles. Este tipo de estudo tem grande aplicação para as situações em que a doença é relativamente pouco freqüente e o tempo decorrido entre a exposição ao risco e a evidenciação do seu efeito, é longo. Os estudos caso-controle têm limitada implicação ética uma vez que não existe intervenção nem observação prospectiva de exposições de risco. Os estudos tipo caso-controle, inicialmente propostos para o estudo de doenças crônico-degenerativas, principalmente câncer, também têm aplicação no estudo de doenças infecciosas.

SELEÇÃO DE PARTICIPANTES

Seleção de casos - O local de obtenção de casos e controles depende das características da doença em estudo. Casos podem ser identificados em hospitais, clínicas especializadas ou serviços de saúde (ex.: casos de hanseníase e de tuberculose). É possível fazer busca populacional de casos, mediante inquérito baseado em marcadores, como níveis de anticorpos.

Seleção de controles - A busca de controles deve seguir, como orientação geral, o princípio "*se o controle presente fosse um caso, ele seria encontrado onde os casos estão sendo encontrados*". Controles podem ser recrutados nos hospitais onde os casos foram obtidos, na vizinhança dos casos, nas mesmas escolas, entre amigos e colegas de trabalho dos casos, na população em geral sob esquema de amostra probabilística. Em qualquer situação haverá vantagens e desvantagens, sempre com possibilidade de resultados viesados. Controles obtidos por **sugestão dos próprios casos** podem ser muito semelhantes nos seus comportamentos e costumes, e, se o fator de risco estudado está relacionado a hábitos que podem ser comuns entre amigos, não será detectado. O custo e dificuldade operacional de obtenção de **controles populacionais** tornam esta abordagem pouco prática. No contexto das doenças infecciosas formas sub-clínicas e clínicas da doença podem ser detectadas. A estratégia a ser adotada para seleção do grupo controle depende do objetivo do estudo. Por exemplo, se a finalidade é avaliar fatores de risco para malária grave ou complicada (casos) o grupo controle deverá ser formado por indivíduos com parasitemia assintomática, ou formas leves de malária. Se o objetivo do estudo é determinar fatores de risco prognósticos para o desenvolvimento de cardiopatia chagásica, indivíduos soropositivos e com alterações ECG compatíveis com doença de Chagas serão selecionados como casos e dois tipos de controles, poderiam ser selecionados: a) indivíduos soronegativos e b) indivíduos soropositivos, ambos sem alterações ECG. Se indivíduos soronegativos são os controles (a), o estudo determinaria riscos de desenvolvimento de doença de Chagas enquanto para o segundo tipo de controles o estudo estaria avaliando fatores associados à evolução da infecção para doença.

TIPOS DE ESTUDOS

. **Caso-controle de base populacional** - Neste tipo de delineamento casos e controles são selecionados da população; casos podem ser detectados através de triagem populacional, em área geográfica definida durante um determinado período de tempo. Pode-se utilizar registros de hospitais para se identificar todos os possíveis casos da área de estudo ou uma amostra aleatória deles. Os controles são selecionados através de uma amostra probabilística dos indivíduos sem a doença, pertencentes à mesma área geográfica dos casos.

. **Caso-controle aninhado** - É um delineamento no qual casos e controles são selecionados no decorrer de uma coorte pré-definida, na qual algumas informações sobre exposições e fatores de risco já encontram-se disponíveis. Para cada caso, controles são selecionados aleatoriamente de indivíduos que encontram-se sob risco no momento do diagnóstico do caso, o que significa um emparelhamento pelo efeito de confusão do tempo. Informações adicionais são coletadas e analisadas no momento da seleção de casos incidentes e controles.

MEDIDA DE ASSOCIAÇÃO

. Análise univariada

A estatística usada como medida de associação é o *odds ratio* (**OR**), que, para doenças com tempo longo de latência, é uma estimativa da razão das forças de morbidade entre expostos e não expostos ao fator de risco. O *odds ratio* é uma aproximação da razão de riscos (risco relativo), quando as incidências são baixas e uma aproximação da razão de prevalência, quando as prevalências são baixas. Quando casos e controles são selecionados da população geral, a proporção de expostos ao fator de risco no grupo controle pode ser utilizada como uma estimativa da proporção de expostos na população em geral. Esta vantagem propicia que se calcule o risco atribuível percentual (populacional), que expressa a proporção da doença na população de estudo que é atribuível à exposição ou fator e que poderia ser eliminada se o fator fosse removido. Esta é uma informação útil pois indica quais exposições ou fatores são os mais relevantes e devem prioritariamente ser eliminados como medidas de saúde pública.

O resultado amostral de um estudo caso-controle pode ser expresso em uma tabela 2x2:

EXPOSIÇÃO	CONDIÇÃO		TOTAL
	caso	controle	
Presente	a	b	a + b
Ausente	c	d	c + d
TOTAL	a + c	b + d	T

Onde:

- a + c** = número de casos
- a** = número de casos com o fator de risco presente
- c** = número de casos com o fator de risco ausente
- b + d** = número de controles
- b** = número de controles com o fator de risco presente
- d** = número de controles com o fator de risco ausente
- a + b** = número total de indivíduos que estiveram expostos ao fator de risco
- c + d** = número que não esteve exposto ao fator de risco
- T** = total do conjunto das amostras de casos e controles

Odds é uma medida de probabilidade, definido como o quociente de duas probabilidades, complementares entre si. Na tabela, o *odds* de presença de exposição ao fator, entre os casos, é a/c ; o *odds* de presença de exposição ao fator, entre os controles, é b/d . O quociente entre estes dois *odds* é o *odds ratio* (OR): $(a/c) / (b/d) = (a.d) / (b.c)$; o *odds ratio*, sob a forma desta última expressão, é

também denominado “razão dos produtos cruzados”. Em português, *odds ratio* pode ser traduzido, literalmente, como "razão do quociente de probabilidades".

A associação entre o fator de risco e a doença pode ser especificada como positiva ou negativa, situação que leva a um teste estatístico monocausal; não sendo especificada, tem-se um teste bicaudal.

OR=1 indica que probabilidade de doença nos expostos ao fator de risco e nos não expostos são equivalentes. OR>1 indica que a exposição ao fator em estudo é de risco, podendo implicar em relação causa efeito. A decisão estatística é baseada na estatística χ^2 (qui quadrado), com um grau de liberdade. Se o nível de significância for $\alpha = 5\%$ e o teste for bicaudal, χ^2 é comparado com 3,84; se maior indica uma associação estatisticamente significativa.

Como exemplo, em um surto epidêmico de diarreia em participantes de uma oficina de trabalho na Fundação Nacional de Saúde, um estudo de caso-controle foi conduzido para investigar o risco de intoxicação alimentar associado com ingestão de maionese. Casos e controles foram identificados através de entrevistas com indivíduos que comeram no buffet servido durante os dias da oficina. Os dados estão dispostos na tabela abaixo com os resultados do *odds ratio* e o intervalo da estimativa da *odds ratio*, com 95% de confiança.

INGESTÃO DE MAIONESE	DIARRÉIA		TOTAL
	Sim	Não	
Sim	75	152	227
Não	10	140	150
TOTAL	85	292	377

OR=6,91 (IC 95% 3,4-13,9) $\chi^2=35,85$

METODOLOGIA

. Amostragem

Apesar de não se ter que utilizar formalmente nenhuma técnica de amostragem, deve-se verificar em que extensão o grupo escolhido representa todos os indivíduos com a doença, não só para interpretação dos resultados para os indivíduos incluídos no estudo (**validade interna do estudo**), como também para se poder fazer inferências, extrapolações para indivíduos não incluídos no estudo (**validade externa do estudo**)

. Tamanho da amostra

O número de casos e controles a ser selecionado, depende do **tamanho da amostra** necessário para testar uma hipótese. De maneira geral, o tamanho da amostra é inversamente proporcional ao risco que se pretende detectar. Para se detectar um pequeno risco relativo (ex.: 1,2 ou seja, 20% de aumento no risco comparado ao grupo controle) é necessário incluir um grande número de

participantes. Estudos conduzidos com poucos casos apresentam baixo poder estatístico para detectar riscos.

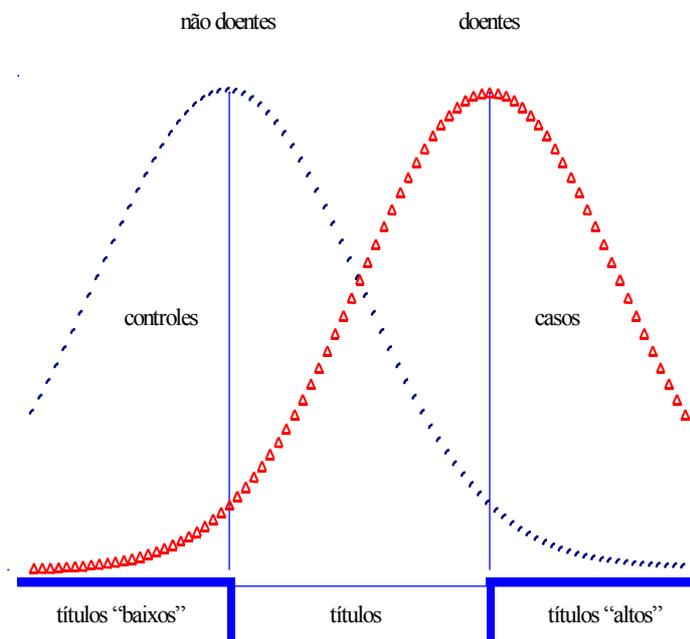
Para o cálculo do tamanho da amostra as seguintes informações são necessárias:

- (1) nível de significância do teste (geralmente $\alpha=5\%$)
- (2) o poder do teste (geralmente $1-\beta=80\%$)
- (3) a proporção de pessoas expostas ao fator de risco na população geral
- (4) o valor de *odds ratio* mínimo a ser detectado
- (5) razão entre o número de controles e casos

Por exemplo, para $\alpha = 5\%$, poder do teste = 80% ($1-\beta$), 10% de expostos ao fator de risco na população geral e $OR = 2$, um total de 307 casos e 307 controles devem ser selecionados (ver exercícios em EPIINFO).

. Medida da exposição

Para avaliar as medidas de exposição pode-se utilizar entrevistas, questionários padronizados, informações de parentes/vizinhos ou, marcadores biológicos. Os procedimentos devem ser iguais tanto para casos como para os controles. Os entrevistadores devem desconhecer o “status” caso-controle do indivíduo para assegurar o caráter de mascaramento, minimizando-se o **viés do observador**.



Tipos de viés

Viés de classificação - Estudos de caso-controle estão sujeitos à diferentes fontes de bias. A classificação de indivíduos como casos, por terem a doença de estudo, e como controles, por não terem a doença deve seguir critérios bem definidos, a fim de evitar viés de classificação. Este é um

erro sistemático pelo qual doentes são selecionados como controles e indivíduos sem a doença são selecionados como caso. Exames laboratoriais, altamente sensíveis e específicos, são desejáveis, em complementação a diagnósticos clínicos; se o nível de anticorpos for usado, pode-se estabelecer dois pontos de corte na escala de títulos, o menor deles como limite superior para seleção de controles e o maior deles como limite inferior para seleção de casos. Este procedimento visa a minimizar a classificação de não doentes como casos (falsos casos) e a de doentes como controles (falsos controles); a figura acima mostra esquematicamente a situação.

Viés de seleção - Outra possibilidade de vícios em estudos caso-controle refere-se ao viés de seleção por incorreções ou limitações no delineamento do estudo, afetando a comparabilidade de casos e controles. Um dos motivos deve-se ao fato de que os princípios da comparabilidade na seleção de controles geralmente são influenciados pelo princípio da eficiência, ou seja, disponibilidade de recursos e tempo.

Viés do observador - Ao selecionar os controles deve-se assegurar que as observações sejam realizadas nos dois grupos sob as mesmas condições. O ideal seria que o investigador não tivesse o conhecimento de quem tem a doença e quem não tem, ou seja, quem é caso e quem é controle, para evitar tendenciosidades na coleta de informações (ou seja, o viés do observador).

Casos prevalentes - Outra fonte potencial de bias é a seleção de casos prevalentes ao invés de casos incidentes. A prevalência é afetada pela duração da doença que é influenciada pelo tratamento e cura e também pela mortalidade associada à doença. Quando se inclui casos prevalentes os fatores podem estar estatisticamente associados à doença devido a um "efeito de sobrevivência" e duração da doença e não devido à uma associação causal. Na infecção pelo *T. cruzi*, por exemplo, um percentual de óbitos ocorre na fase aguda de infecção, e portanto, fatores de risco avaliados para casos selecionados na fase crônica da infecção (casos prevalentes) estariam associados à sobreviventes da fase aguda.

ANÁLISE ESTRATIFICADA

Confusão - Quando um fator está associado simultaneamente à exposição e à doença é chamado de variável de confusão. Confusão é uma distorção, causada por outra variável C, no resultado numérico que mede a associação entre uma variável E (exposição) e a condição D (doença), estando a variável C associada à variável E e à variável D. A confusão é um viés que deve ser controlado, o que pode ser feito na análise, mediante análise estratificada.

Estratificação - Se uma possível variável de confusão C tiver dois estratos, com C presente e C ausente, a associação da variável E com a condição D deverá ser examinada segundo estes estratos (análise estratificada), além da análise conduzida sem considerar C (análise bruta); o resultado está esquematizado nas tabelas a seguir:

Análise bruta

EXPOSIÇÃO AO FATOR	CONDIÇÃO (D)		TOTAL
	caso	controle	
Presente	a	b	a + b
Ausente	c	d	c + d
TOTAL	a + c	b + d	T

Análise Estratificada

ESTRATO1 (C presente)

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso	controle	
Presente	a ₁	b ₁	a ₁ + b ₁
Ausente	c ₁	d ₁	c ₁ + d ₁
TOTAL	a₁ + c₁	b₁ + d₁	T₁

ESTRATO 2 (C ausente)

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso	controle	
Presente	a ₂	b ₂	a ₂ + b ₂
Ausente	c ₂	d ₂	c ₂ + d ₂
TOTAL	a₂ + c₂	b₂ + d₂	T₂

Na análise bruta $OR_{bruto} = (a.d) / (b.c)$. Na análise estratificada, o subscrito indica o estrato; o *odds ratio* do primeiro estrato é:

$OR_1 = (a_1.d_1) / (b_1.c_1)$ e o do segundo estrato é:

$OR_2 = (a_2.d_2) / (b_2.c_2)$

A medida de associação comum, *odds ratio* sintético, é:

$$OR_{MH} = \frac{((a_1.d_1)/T_1 + (a_2.d_2)/T_2)}{((b_1.c_1)/T_1 + (b_2.d_2)/T_2)}$$

Esta é a média ponderada dos *odds ratios* dos dois estratos, com pesos proporcionais aos logaritmos naturais das variâncias. O subscrito MH refere-se aos autores Mantel e Haenszel, que desenvolveram este estimador da associação comum. Se $OR_{bruto} \neq OR_{MH}$ existe confusão, podendo, ainda, haver modificação de efeito. Se $OR_{bruto} = OR_{MH}$ não há confusão, podendo, porém, haver modificação de efeito. A decisão sobre a diferença entre os *odds ratio* indicadora de confusão é arbitrária, não devendo ser feito teste de hipóteses para isto; uma conduta possível é estabelecer-se uma porcentagem limite para a diferença entre as medidas.

Modificação de efeito - Diz-se que há modificação de efeito quando as medidas de associação entre a variável E e a variável D nos estratos de C são diferentes, indicando processos desiguais de causalidade, segundo as características da variável C. Ao contrário da situação de confusão, no caso de modificação de efeito não se deve ter uma medida resumo de associação, mas sim, uma comparação das medidas de associação (OR) para cada estrato. A decisão se há ou não modificação

de efeito é tomada estatisticamente mediante o teste de heterogeneidade ou interação estatística. As hipóteses a serem testadas são:

H_0 : há homogeneidade entre os estratos

H_A : há heterogeneidade entre os estratos

O teste de hipótese, na situação de análise estratificada, assumindo H_0 : homogeneidade entre os estratos (ou seja, não haver modificação de efeito), é conduzido para o conjunto de todos os estratos, utilizando-se uma estatística com distribuição χ^2 , cujo grau de liberdade corresponde ao número de estratos menos 1.

Exemplo

Continuando a análise do estudo da associação entre ingestão de maionese e diarreia, incluindo uma variável eventual de confusão, com duas possibilidades, < 30 anos e \geq 30 anos, teríamos:

ESTRATO1 (idade < 30 anos)

INGESTÃO DE MAIONESE	DIARRÉIA		TOTAL
	Sim (caso)	Não (controle)	
Sim	74	120	194
Não	5	54	59
TOTAL	79	174	253

OR=6,66 (exato IC95% 2,50-22,18)

ESTRATO 2 (idade \geq 30 anos)

INGESTÃO DE MAIONESE	DIARRÉIA		TOTAL
	Sim (caso)	Não (controle)	
Sim	1	32	33
Não	5	86	91
TOTAL	6	118	124

OR=0,54 (exato IC95% 0,01-5,10)

$OR_{MH} = 4,50$ (IC 95% 1,84-9,08)

$OR_{bruto} = 6,91$ (IC 95% 3,30-14,84)

Os resultados indicam associação entre consumo de maionese e diarreia. O teste de heterogeneidade foi significativo ($p=0,00018$), o que significa que < 30 anos e \geq 30 anos devem ser trabalhados separadamente, pois eles tem riscos diferentes.

EMPARELHAMENTO

Há estudos em que, para cada caso, é escolhido um controle (às vezes dois, ou três, etc), como um irmão, um vizinho, um companheiro de trabalho, com a finalidade de garantir uma maior comparabilidade entre os grupos de casos e de controles ou controlar fatores comuns não facilmente identificáveis. Idade e sexo são, em geral, consideradas variáveis que estão intimamente associadas à possibilidade de exposição e ao desenvolvimento da doença. Por este motivo, casos e controles são usualmente selecionados dentro da mesma faixa etária e sexo. A estratégia de levar em conta a variável de confusão já no delineamento do estudo, é denominada **emparelhamento**. Este emparelhamento dos casos e controles por sexo e idade torna os grupos mais comparáveis e minimiza potenciais distorções dos resultados na avaliação de riscos. O emparelhamento leva a ter-se um número de estratos igual ao número de casos, cada estrato composto de 1 caso e seu respectivo controle (ou controles). O princípio da análise é o mesmo, mas a variável que serviu como base do emparelhamento não pode ser analisada.

Para a situação 1 caso/1 controle, tem-se as seguintes possíveis configurações, para cada par caso/controle:

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso(1)	controle(0)	
Presente(1)	1	1	2
Ausente(0)	0	0	0
TOTAL	1	1	2

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso(1)	controle(0)	
Presente(1)	1	0	1
Ausente(0)	0	1	1
TOTAL	1	1	2

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso(1)	controle(0)	
Presente(1)	0	1	1
Ausente(0)	1	0	1
TOTAL	1	1	2

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso(1)	controle(0)	
Presente(1)	0	0	0
Ausente(0)	1	1	2
TOTAL	1	1	2

O resultado para k pares pode ser expresso na seguinte tabela resumo:

CASO	CONTROLE	
	Exposto (1)	Não exposto (0)
Exposto(1)	v_{11}	v_{10}
Não Exposto(0)	v_{01}	v_{00}

onde:

v_{11} = o número de pares em que o caso e o respectivo controle estiveram expostos ao fator de risco

v_{10} = o número de pares em que o caso esteve exposto e o controle não

v_{01} = o número de pares em que o caso não esteve exposto e o controle esteve

v_{00} = o número de pares em que nem o caso nem o controle sofreram exposição ao fator de risco.

O OR é calculado como a razão entre os pares discordantes:

$$OR = \frac{v_{10}}{v_{01}}$$

A situação em que o número de controles varia para cada caso pode ser tratada mais facilmente como na situação de estratificação, embora o número de estratos naturalmente seja bem grande.

VANTAGENS E LIMITAÇÕES

.Vantagens - Estudos de caso-controle são particularmente adequados para avaliação de doenças raras constituindo-se em método rápido, prático e de baixo custo para testar o efeito e interação de um grande número de fatores que se relacionam com o evento estudado. Ao contrário dos estudos de coorte, pelo fato de não necessitar de um grande número de participantes, pode-se empregar exames/testes mais dispendiosos e laboriosos. Também, evita perdas de seguimento de casos.

Estudos de caso-controle e os programas de saúde pública - Em saúde pública a metodologia caso-controle tem sido utilizada principalmente para avaliação de eficácia de vacinas, em investigações de epidemias, para avaliar testes de triagens populacionais, avaliação de impacto de medidas preventivas como programas de saúde pública, e eficácia de tratamentos. Na eficácia de vacinas

após a implantação em programas de rotinas a informação sobre o estado vacinal dos indivíduos é o fator de exposição a ser avaliado como associado ou não à doença. A eficácia da vacina (EV) pode ser estimada pela fórmula $EV=1-OR$.

Limitações - As principais limitações de estudos de caso-controle são: a) não adequados para investigações de exposições raras, a menos que o risco atribuído à exposição na população seja muito alto; b) não estima a incidência das doenças/infecções estudadas e, c) a informação sobre a exposição ou fator é obtida após a ocorrência da doença e portanto, não há como se distinguir uma cronologia nítida entre a exposição e o aparecimento da doença, diferente do que ocorre nos estudos de coorte onde a exposição é o ponto de partida. Por exemplo, para avaliação da associação entre estado nutricional e crianças infectadas por *T. cruzi*, se a soropositividade e o grau de desnutrição são medidos ao mesmo tempo, fica difícil determinar se a desnutrição levou a uma maior susceptibilidade à infecção pelo *T. cruzi*, ou se as crianças infectadas desenvolvem mais desnutrição.

ROTEIRO PARA DELINEAR UM ESTUDO DE CASO-CONTROLE

① *Defina de forma clara a questão que o estudo pretende responder*

- . esclareça nos objetivos a doença/ evento, a população que será investigada e os fatores que serão testados como associados ao evento

② *Estabeleça a definição de caso e de controle*

- . faixa etária, sexo
- . testes laboratoriais, exames clínicos, parâmetros a serem avaliados, interpretação
- . critérios de inclusão (casos graves, leves) e exclusão
- . número de controles necessários por caso
- . necessidade de emparelhamento de casos e controles

③ *Especifique o critério de amostragem*

- . defina a população de referência
- . esclareça a fonte de seleção dos casos e dos controles: serviços de saúde, população
- . forneça informações sobre a validade interna e externa do estudo

④ *Defina as variáveis de exposição e as variáveis potenciais de confusão que serão controladas na análise*

- . descreva os métodos laboratoriais, parâmetros a serem avaliados, interpretação, exames clínicos, entrevistas
- . esclareça os procedimentos utilizados para minimizar os possíveis vieses do observador - mascaramento

⑤ *Calcule o tamanho da amostra necessário*

- . defina a proporção aproximada de exposição ao fator de risco na população geral (controles)
- . calcule o número mínimo e suficiente de casos e de controles necessário para estimar o risco relativo definido pelo estudo
- . defina o nível de significância e o poder estatístico do estudo com o número de participantes que serão estudados
- . certifique-se da factibilidade do estudo, em termos logístico e de tempo de recrutamento dos casos e controles

⑥ *Descreva as etapas da análise de dados*

- . indique os parâmetros (proporções, médias), métodos estatísticos e sub-grupos para avaliação das características básicas dos casos e controles e avaliação dos fatores de risco (odds ratio e IC95%)
- . descreva o tipo de análise - univariada, estratificada, multivariada

REFERÊNCIAS PARA LEITURA

ARMENIAN HK (Ed.) Applications of the Case-Control Method. **Epidemiologic Review**,16(1). 1994.

BRESLOW, N.E. & DAY, N.E. *Statistical methods in cancer research*. Switzerland, International Agency for Research on Cancer, Lyon, V.1. Scientific publication N°. 32, 1980.

HENNECKENS, H.C. & BURING, J.E. *Epidemiology in Medicine*, 5th ed. Boston: Toronto, Ed. Little, Brown and Company, 1987.

KLEINBAUM, D.G., KUPPER, L.L. & MORGENSTERN, H. *Epidemiologic Research*. New York, Ed. Van Nostrand Reinhold, 1982.

KOPEC, J.A. & ESDAILE, J.M. Bias in case-control studies. A review. **Journal of Epidemiology and Community Health**,44:179-86, 1990.

MANTEL, N. & HAENSZEL, W. Statistical aspects of the analysis of data from retrospective studies of disease. **Journal Natl Cancer Inst**,22:719-48, 1959.

PEREIRA, M.G. *Epidemiologia - Teoria e Prática*. Guanabara Koogan, Rio de Janeiro, 1995.

RODRIGUES, L. & KIRKWOOD, B.R. Case-control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. **International Journal of Epidemiology**,19:205-13, 1990.

SCHLESSELMAN, J.J. *Case-control studies - Design, conduct and analysis*, New York:Oxford University Press, 1982.

SMITH, P.G. & DAY, N.E. The design of case-control studies: The Influence of confounding and interaction effects. **Int. J. Epidemiol.**,13:356-65, 1984.

WACHOLDER, S., MCLAUGHLIN, J.K., SILVERMAN, D.T. & MANDEL, J.S. Selection of controls in case-control studies. **American Journal of Epidemiology**, 135(9):1019-50, 1992.

EXERCÍCIOS

Arquivos:

1. **casoch.rec**
2. **casach.rec**
3. **casohan.rec**

1. **Fatores de risco para infecção chagásica** - um estudo de caso-controle de base populacional foi delineado para avaliar fatores de risco associados à infecção pelo *T.cruzi* na infância, em municípios do nordeste do Estado de Goiás/Brasil. Um mil novecentos e noventa (1990) crianças de 7-12 anos foram triadas sorologicamente para seleção de 149 crianças soropositivas (casos) e 298 soronegativas (controles). Casos e controles foram emparelhados por sexo, frequência de idade e localidade. Detalhes da metodologia estão na referência **Andrade et al, 1995**. A variável "NOPAR" indica o número do triplo, incluindo caso, controle 1 e controle 2. A análise adotada teve como objetivos: (1) verificar a comparabilidade entre casos e controles em relação às suas características básicas; (2) avaliar o risco associado à presença de moradia infestada por triatomíneos para infecção, e (3) avaliar a associação entre filho soropositivo e soropositividade dos pais. Utilize o arquivo **casoch.rec** para responder as questões abaixo.

Questão 1. **Construa uma tabela comparando as características básicas dos casos e controles, incluindo sexo, grupo etário "GRAGE" 7-9 e 10-12 anos, média de idade da criança, município de residência, antecedentes mórbidos e hospitalização. De acordo com os resultados da tabela o que pode ser concluído em relação a comparabilidade entre casos e controles?**

Notas 1:

```
READ CASOCH.REC (para abrir o arquivo)
TABLES SEX CACO
criar a variável GRAGE (grupo etário)
DEFINE GRAGE #
IF AGE >= 7 AND AGE <= 9 THEN GRAGE=1 ELSE
GRAGE=2
TABLES GRAGE CACO
MEANS AGE CACO /N
TABLES MUN CACO
para as variáveis MORB (antecedentes mórbidos) e HOSP
(hospitalização) excluir código 9 (Sem informação)
SELECT MORB <>9
TABLES MORB CACO
SELECT (para desativar a seleção)
SELECT HOSP <>9
TABLES HOSP CACO
SELECT
```

Questão 2. Calcule o *odds ratio* (OR) emparelhado e respectivo intervalo de 95% de confiança (IC 95%) para "moradia infestada por triatomíneo" ("BARB"), captura de triatomíneo no domicílio ("DCAP"), sorologia da mãe e sorologia do pai. Crie a variável "BARB" correspondente à "moradia infestada por triatomíneo". Utilize as variáveis "DCAP" e "DVTR" para obter "BARB"=1 (DCAP=1 OU DVTR=1) e "BARB"=2 (DCAP=2 E DVTR=2). Crie uma nova variável "GRNPES" ("numero de pessoas"). Sugestão: de 1-4 pessoas, de 5-8 e ≥ 9 pessoas por moradia. Para a variável "GRNPES" escolha o primeiro estrato (1-4 pessoas) como referência e calcule o OR (não emparelhado) e IC95% dos demais estratos comparando com a categoria de referência.

Notas 2:

```
criar a variável "BARB"
DEFINE BARB #
IF DCAP=1 OR DVTR=1 THEN BARB=1 ELSE BARB=2
comando MATCH para o cálculo do OR emparelhado.
MATCH NOPAR CACO BARB
MATCH NOPAR CACO DCAP
para as variáveis SORMAE (sorologia da mãe) e SORPAI (sorologia do pai)
excluir código 9 (Não realizada)
SELECT SORMAE <>9
MATCH NOPAR CACO SORMAE
SELECT
SELECT SORPAI <> 9
MATCH NOPAR CACO SORPAI
SELECT
criar variável GRNPES (categorias do número de pessoas na família)
LET GRNPES = NPES
RECODE GRNPES 1-4=1 5-8=2 9-HIGH=3
TABLES GRNPES CACO
anote os resultados
tecle F9 para ir ao DOS
digite STATCALC para cálculo do OR não emparelhado (selecione
Tables 2 x 2, 2 x n)
utilize os dados produzidos pela tabela anterior
pressione F10 para sair do STATCALC
digite EXIT para voltar ao ANALYSIS
CLOSE (feche o arquivo)
```

Questão 3. Quais as possíveis explicações para as diferenças dos OR obtidos para "captura de triatomíneo" (DCAP) e "moradia infestada" (BARB)? Utilize o IC 95% e a descrição das variáveis para responder.

Questão 4. Qual a relação entre número de pessoas na família e risco de infecção pelo *T.cruzi* em crianças? Relacione a significância estatística ao significado clínico.

Questão 5. A associação entre soropositividade da mãe e filho soropositivo pode ter ocorrido ao acaso? Quais as prováveis interpretações para esta associação? Qual a interpretação do resultado do OR da sorologia do pai e soropositividade do filho?

2. Fatores de risco para infecção /infestação chagásica - o arquivo **casach.rec** extraído da base de dados **casoch.rec** inclui somente uma criança (soropositiva-caso ou soronegativa-controle) por moradia, excluindo a possibilidade de casos e controles coabitando a mesma moradia. A unidade de estudo passa a ser a "moradia de criança soropositiva"/caso (n=89) e "moradia de criança soronegativa"/controle (n=178) e a variável caso/controle é CASAPOS. Detalhes da metodologia estão na referência **Andrade et al, 1995**.

Questão 6. Qual a relação entre aumento do número de cômodos e risco de infecção pelo *T.cruzi* em crianças? Construa categorias para a variável "número de cômodos" (NCOM), criando uma nova variável "GRNCOM". Sugestão: 1-3 cômodos, 4-6 e ≥ 7 cômodos. Escolha a categoria 1-3 como referência e calcule o OR não emparelhado para as demais categorias comparando com a categoria de referência.

Notas 6:

READ CASACH.REC (para abrir o novo arquivo)
Para criar categorias para a variável "número de cômodos" ("GRNCOM") repita os comandos utilizados para categorias do "número de pessoas na família" (questão 2).
TABLES GRNCOM CASAPOS
anote os resultados
tecle F9 para ir ao DOS
Utilize o STATCALC para o cálculo do OR não emparelhado; siga os mesmos comandos da questão 2

Questão 7. Calcule a soroprevalência para as mães e para os pais no grupo controle. Que inferências podem ser feitas para a região de estudo em relação à endemicidade da infecção? Que implicações têm estes resultados do ponto de vista da possibilidade de transmissão congênita na área?

Notas 7:

```
SET PERCENTS=ON
SELECT SORMAE <> 9
TABLES SORMAE CASAPOS
SELECT
SELECT SORPAI <> 9
TABLES SORPAI CASAPOS
SELECT
SET PERCENTS=OFF
```

Questão 8. Calcule para cada base de dados (*casoch.rec* e *casach.rec*) o poder do estudo de acordo com o aumento do número de controles por caso. Considere $\alpha=0,05$, número de casos=149 e 89, OR 2,0 e 2,2 respectivamente e prevalência da variável de exposição=15%. Qual a relação entre número de controles por caso e o poder do estudo? Discuta o custo-benefício e aspectos operacionais referentes a seleção de mais de um controle por caso, levando em conta a estratégia para recrutamento de casos e controles adotada nesta investigação. Qual a relação entre o número de indivíduos (tamanho do estudo) e o poder do estudo? (responda comparando o número de participantes dos dois bancos de dados).

Notas 8:

```
tecle F9 para ir ao DOS
digite EPITABLE para cálculo do poder estatístico (comando
SAMPLE / POWER CALCULATION / CASE-CONTROL
STUDY)
pressione F10 para sair do EPITABLE
digite EXIT para voltar ao ANALYSIS
CLOSE (para fechar o arquivo)
```

3. Efetividade da vacina BCG na Hanseníase - um estudo de caso-control foi delineado para avaliar a efetividade da vacina BCG na hanseníase. Detalhes do estudo estão na referência **Rodrigues et al.,1992**. 62 crianças/escolares menores de 16 anos recém-diagnosticadas com hanseníase pelo ambulatório de referência da cidade, no período de 1989-1990, foram incluídas como **Casos**. **Controles** (3:1) foram selecionados de escolas localizadas na mesma área geográfica de procedência dos casos. Casos e Controles foram emparelhados por frequência de sexo e grupo etário. A exposição ao BCG foi avaliada pela presença de cicatriz vacinal de BCG no antebraço direito. A análise de dados foi planejada tendo como objetivos: (1) avaliar a comparabilidade entre casos e controles em relação às suas características básicas; (2) mensurar a proteção (efetividade) do BCG em relação a hanseníase e (3) determinar a associação entre BCG e formas clínicas de hanseníase. Utilize o arquivo **casohan.rec** para responder as questões abaixo.

Questão 9. Faça uma tabela incluindo as características gerais dos casos e controles - sexo e idade. O que pode ser concluído em relação à comparabilidade entre os grupos?

Notas 9:

```
READ CASOHAN.REC (para abrir o arquivo)
criar a variável GRAGE (grupo etário)
LET GRAGE = AGE
RECODE GRAGE low- 5=1 6-10=2 11-15=3
SET PERCENTS=ON
TABLES GRAGE TIPO
TABLES SEX TIPO
SET PERCENTS=OFF
MEANS AGE TIPO /N
```

Questão 10. Analise a média de idades entre indivíduos vacinados e não vacinados com BCG. Quais as prováveis interpretações destes resultados? Responda considerando que a vacina BCG foi implementada pelo Programa Nacional de Imunização a partir de 1975.

Notas 10: MEANS AGE CIC /N

Questão 11. Calcule o *odds ratio* e IC 95% da vacina BCG. Qual a efetividade da vacina (IC 95%) na proteção da hanseníase?

Notas 11:

```
TABLES CIC TIPO
anote a tabela produzida
tecle F9 para ir ao DOS
digite EPITABLE para cálculo da eficácia (comando
STUDY /VACCINE EFFICACY / CASE-CONTROL
STUDY)
utilize os dados da tabela anterior
pressione F10 para sair do EPITABLE
digite EXIT para voltar ao ANALYSIS
```

Questão 12: Avalie a associação entre forma clínica de hanseníase e exposição à vacina BCG. Esses resultados poderiam ser explicados pelo acaso? O que você conclui sobre a implementação do BCG como uma das estratégias de Eliminação da Hanseníase? Responda levando em conta a cobertura atual da vacina BCG.

Notas 12:

```
SELECT TIPO=1 (para selecionar apenas os casos)
SET PERCENTS=ON
TABLES CIC FC2
SELECT
SET PERCENTS=OFF
CLOSE (para fechar o arquivo)
```

REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, A.L.S.S., ZICKER, F., SILVA, I.G., SOUZA, J.M.P. & MARTELLI, C.M.T. Risk factors for *Trypanosoma cruzi* infection among children in Central Brazil: a case-control study in vector control settings. **American Journal of Tropical Medicine and Hygiene**, **52(2)**:183-187, 1995.

RODRIGUES, M.L.O., SILVA, S.A., NETO, J.C.A., ANDRADE, A.L.S.S., MARTELLI, C.M.T. & ZICKER, F. Protective effect of intradermal BCG against leprosy: a case-control study in Central Brazil. **International Journal of Leprosy and Other Mycobacterium Disease**, **60(3)**:335-339, 1992.

Arquivo: **casoch.rec**

Nome	Descrição	Código	Descrição do Código
NOID	Número de identificação	1 a 463	
NOPAR	Identificador dos triplos	1 a 154	
CACO	Caso/Controle	1 2	Caso Controle
AGE	Idade em anos completos	7 a 12	
SEX	Sexo	1 2	Masculino Feminino
MUN	Município	1 2 3	Posse Guarani de Goiás Simolândia
NCOM	Número de cômodos na moradia atual	1 a 13	
NPES	Número de moradores	2 a 14	
CP	Casa própria	1 2	Sim Não
DCAP	Captura triatomíneo no domicílio	1 2	Sim Não
DVTR	Vestígios triatomíneo no domicílio	1 2	Sim Não
MORB	Antecedentes morbidos	1 2 9	Sim Não Sem informação
HOSP	Antecedentes de hospitalização	1 2 9	Sim Não Sem informação
SORPAI	Sorologia do pai	1 2 9	Positivo Negativo Não realizada
SORMAE	Sorologia da mãe	1 2 9	Positivo Negativo Não realizada

Arquivo: **casach.rec**

Variável	Descrição	Código	Descrição do Código
NOID	Número de identificação	1 a 463	
CASAPOS	“Caso”/ “Controle”	1 2	Casa caso Casa controle
NCOM	Número de cômodos na moradia atual	1 a 13	
SORPAI	Sorologia do Pai	1 2 9	Positivo Negativo Não realizada
SORMAE	Sorologia da Mãe	1 2 9	Positivo Negativo Não realizada
NOPAR2	Identificador dos triplos	1 a 89	

Arquivo: **casohan.rec**

Variável	Descrição	Código	Descrição do Código
NO	Número de identificação	1 a 301	
AGE	Idade em anos completos	2 a 15	
SEX	Sexo	1 2	Masculino Feminino
CIC	Presença de cicatriz vacinal - BCG	1 2	Sim Não
TIPO	Identificação caso/controle	1 2	Caso Controle
FC2	Forma Clínica	P M	Paucibacilar Multibacilar

ESTUDOS DE COORTE

INTRODUÇÃO	
Definição	63
Modelo Conceitual	63
TIPOS DE ESTUDO DE COORTE	
Estudo Concorrente ou de Coorte Prospectiva	65
Estudo Não-Concorrente ou de Coorte Retrospectiva	66
ASPECTOS METODOLÓGICOS	
Cálculo do Tamanho da Amostra	67
Amostragem: Recrutamento e Acompanhamento	68
Avaliação da Exposição	69
Avaliação do Evento de Interesse	69
ANÁLISE DOS RESULTADOS	
Medidas de Associação Fator de Risco-Doença	70
Pessoas tempo	71
Tábua de vida	72
Estratificação	72
VANTAGENS E LIMITAÇÕES	75
ROTEIRO PARA DELINEAR UM ESTUDO DE COORTE	77
REFERÊNCIAS PARA LEITURA	79
EXERCÍCIOS	80
DICIONÁRIO DE BANCO DE DADOS	87

INTRODUÇÃO

Definição

Estudo de coorte é um tipo de estudo em que o investigador limita-se a observar e analisar a relação existente entre a presença de fatores de riscos ou características e o desenvolvimento de enfermidades, em grupos da população. Este delineamento é também conhecido como prospectivo, longitudinal, de incidência, ou de seguimento. Entretanto, o termo coorte é o mais utilizado. É um estudo longitudinal, onde os participantes são observados por um período de tempo pré-estabelecido, cuja duração é dependente da enfermidade de interesse, para que sejam verificadas mudanças na frequência da ocorrência da enfermidade associadas à presença do fator de risco. Sendo a unidade de observação o indivíduo, o acompanhamento permite detectar as mudanças que ocorreram em todos os participantes do estudo.

O termo coorte vem do latim "cohors", que identificava um décimo das legiões de soldados romanos que marchavam juntos nas campanhas bélicas do Império. Em epidemiologia, coorte identifica um grupo de pessoas com uma experiência em comum.

Modelo Conceitual

O modelo conceitual de um estudo de coorte é relativamente simples: uma amostra representativa da população a ser estudada é selecionada e informações a respeito de algum fator de risco ou característica de interesse são obtidas. Esta amostra é dividida em dois grupos distintos:

(1) os expostos (ou que possuam) ao fator de risco ou característica de interesse

(2) os não expostos (ou que não possuam) ao fator de risco ou característica de interesse. Estes grupos são acompanhados, no tempo, com objetivo de se verificar quais dos seus membros irão desenvolver o evento a ser estudado (infecção, enfermidade, morte ou outro problema de saúde) e, se a exposição prévia se relaciona ou não a ocorrência deste evento. A questão básica a ser respondida em um estudo de coorte é, portanto, a identificação dos efeitos da exposição na incidência do evento de interesse (Figura 1).

Figura 1 - Esquema de um estudo de coorte



Adaptado de Beaghole et al, 1993

Nos delineamentos mais simples formam-se pelo menos dois grupos, os expostos e os não expostos, de modo que os resultados a serem obtidos (ocorrência da enfermidade) possam ser comparados. A exposição, ao contrário de um estudo experimental, não é controlada; ao investigador cabe apenas observar e mensurar os potenciais fatores de risco (exposição). Nos delineamentos mais complexos, podem ser formados diversos grupos, de acordo com o grau e o tempo de exposição.

Nos estudos de coorte, os registros de ocorrência do evento (infecção, doença ou morte) nos grupos acompanhados, permitem o cálculo das taxas de incidência ou de mortalidade, que são as medidas básicas do risco de adoecer ou de morrer. Estas taxas podem ser calculadas:

- (1) entre aqueles expostos e não expostos ao fator de risco de maior interesse
- (2) entre aqueles expostos a diferentes níveis e por diferentes períodos de tempo ao fator de risco
- (3) entre aqueles expostos a uma combinação destes fatores.

É possível também determinar se as mudanças que possam ter ocorrido nos níveis de exposição durante o acompanhamento, modificaram as taxas sendo calculadas. Desta maneira, por permitir uma medida básica de risco associada à diferentes níveis e tipos de exposição, os estudos de coorte são de grande importância em epidemiologia.

As inferências sobre associação causa-efeito que podem ser derivadas de um estudo de coorte são mais consistentes que as inferências resultantes de um estudo caso-controle. Neste último as conclusões são muitas vezes limitadas pelo modelo conceitual, por erros na seleção de casos e controles, pela representatividade da amostra estudada e por erros na coleta de informações. Portanto, quando possível, é desejável que uma associação doença-fator de risco determinada em um estudo caso-controle seja confirmada através de um estudo de coorte.

Estudos de coorte, muitas vezes chamados de “estudos prospectivos”, têm sido realizados com objetivo de investigar a história natural (aspectos subclínicos e clínicos) de enfermidades e estudar várias características relacionadas à transmissão e à manutenção destas doenças. Para doenças tropicais estes estudos são geralmente realizados em regiões rurais e conduzidos durante um longo período de tempo, em pequenas comunidades que apresentem alta frequência da doença em questão, normalmente doenças endêmicas de evolução crônica (esquistossomose mansoni, doença de Chagas, leishmaniose). Inicialmente é realizado um inquérito epidemiológico através de um estudo transversal para identificar os possíveis fatores de risco (sócio-econômicos, demográficos, biológicos) associados às enfermidades entre os moradores da área. A seguir realiza-se um estudo de coorte, com característica “longitudinal prospectivo”, para acompanhamento dos pacientes assintomáticos e sintomáticos e, em alguns estudos, da coorte de participantes não infectados. Os pacientes são avaliados através de exames periódicos, podendo receber terapêutica disponível; serão acompanhados por longo período de tempo, com objetivo de determinar os marcadores clínicos prognósticos da evolução da enfermidade. Além da avaliação clínica dos pacientes, estes estudos, poderão estimar a prevalência da doença na área, analisar os aspectos descritivos da doença e, dependendo do tempo de acompanhamento e da metodologia empregada, estimar a incidência e identificar fatores de risco para infecção e para o adoecimento.

TIPOS DE ESTUDO DE COORTE

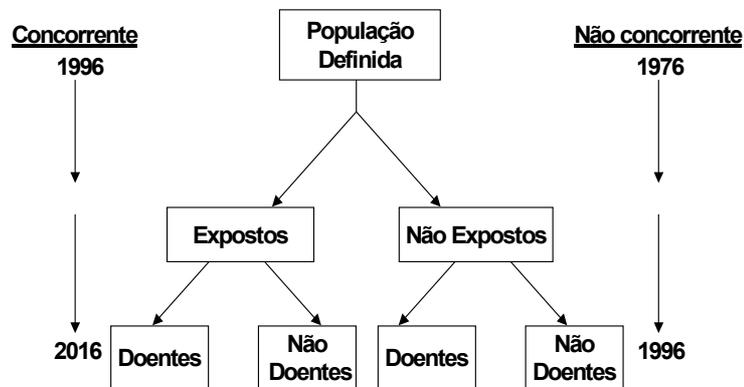
Os estudos de coorte podem ser conduzidos de dois modos distintos:

- (1) estudos concorrentes ou de coorte prospectiva
- (2) estudos não-concorrentes ou de coorte histórica, retrospectiva

Nos estudos concorrentes, indivíduos com e sem exposição ao fator de risco sendo investigado são selecionados no início do estudo e acompanhados por um período especificado de tempo. Nos estudos não-concorrentes, o investigador volta ao passado, onde seleciona os grupos de comparação (com base na exposição ao fator de risco) e “segue” ou “acompanha” estes grupos, através do tempo, geralmente até o presente, por uma variedade de métodos. Como pode ser visto na Figura 2, estes dois tipos de estudos são distintos e envolvem diferentes aspectos metodológicos.

Estudo Concorrente ou de Coorte Prospectiva

Figura 2 - Diferença entre um estudo hipotético de coorte concorrente e não-concorrente em relação ao tempo de seleção de expostos e não expostos. Ano de início do estudo: 1996



Adaptado de Gordis, 1996

Nos estudos concorrentes ou de coorte prospectiva, o epidemiologista identifica e seleciona os grupos expostos e não expostos no momento do início da investigação e os acompanha por um determinado período de tempo.

Exemplo 1: uma investigação para estimar o risco associado com a intensidade de contato com água e transmissão da esquistossomose e desenvolvimento da doença.

Exemplo 2: um estudo conduzido com objetivo de determinar o risco relacionado à incidência do HIV/AIDS; a identificação dos grupos expostos pode ser feita de acordo com os diversos fatores de risco para a infecção pelo HIV (usuários de drogas injetáveis, homossexuais e/ou bissexuais masculinos, etc) e estes grupos serão acompanhados e comparados à grupos não expostos (ou grupos que apresentem diferentes níveis de exposição), quanto à infecção pelo HIV e posterior desenvolvimento da AIDS.

Em alguns estudos, a exposição já pode estar ocorrendo antes do início da investigação, sendo que a duração e intensidade da exposição podem influenciar o evento de interesse. Nestes casos pode-se estimar a duração e a intensidade desta exposição através da utilização de instrumentos como questionários ou entrevistas, prontuários médicos ou testes diagnósticos. Deve ficar claro, entretanto, que em muitas investigações, será impossível determinar o início exato da exposição.

Exemplo 1: desnutrição crônica como fator de risco para o desenvolvimento de infecção pulmonar.

Exemplo 2: infecção por leishmaniose como fator de risco para o desenvolvimento de formas

clínicas da doença. Em ambos os exemplos os participantes expostos podem ter sido infectados há muito tempo, sendo impossível precisar a data da infecção.

Dependendo da enfermidade e dos objetivos do estudo, o tempo de seguimento dos grupos expostos e não expostos pode variar de horas a décadas. Nas enfermidades infecciosas com curto período de incubação, os resultados são obtidos rapidamente; naquelas com longo período de incubação ou de latência, o tempo de seguimento costuma ser bastante longo. Nestes estudos, as dificuldades operacionais são grandes, principalmente quando são considerados o custo, as modificações que podem ocorrer na fase de coleta dos dados e as perdas que seguramente ocorrerão durante o acompanhamento dos participantes.

Exemplo: o estudo de coorte concorrente realizado em Pelotas, Rio Grande do Sul, com objetivo de avaliar a influência de uma série de fatores (perinatais, demográficos, ambientais, alimentares e assistenciais) sobre a saúde infantil. Foram identificadas em 1982, todas as crianças nascidas naquela cidade, que foram acompanhadas prospectivamente durante os primeiros anos de suas vidas para monitorar a ocorrência de vários problemas de saúde.

Estudo Não-Concorrente ou de Coorte Retrospectiva

Nos estudos não-concorrentes ou de coorte retrospectiva, a identificação dos grupos expostos e não expostos é feita em algum momento do passado e estes grupos são então “seguidos” ou “acompanhados”, usualmente até o passado recente ou presente (ou ocasionalmente até o futuro), com objetivo de se identificar as experiências de morbidade ou mortalidade que tenham ocorrido nos participantes destes grupos. É uma investigação que se inicia no passado, no momento em que a exposição ocorreu, mas conserva o princípio básico dos estudos de coorte: exposição em direção ao evento.

Os estudos não-concorrentes raramente podem ser conduzidos em uma amostra geral da população; para que isto ocorra, é necessário que o investigador tenha acesso às informações sobre a exposição ao fator de risco no passado, ou seja, no início da investigação. Somente se estas informações estiverem disponíveis será possível identificar amostras de participantes expostos e não expostos e acompanhar estes indivíduos selecionados. Exemplo: para investigar os efeitos da exposição à intoxicação alimentar os participantes do estudo seriam indivíduos que comeram e os que não comeram de um determinado *buffet*. Para exposições que ocorreram há muito tempo no passado a disponibilidade de informações pode ser limitada havendo necessidade de se recorrer à informações existentes em prontuários médicos, dados sobre a intensidade e duração da exposição, etc. Quando o evento de interesse for o óbito, as informações poderão ser obtidas mais facilmente, através das declarações de óbito, que é um documento legal.

Em consequência do tempo e de recursos financeiros necessários à condução de um estudo de coorte prospectiva, muitas vezes a escolha de um delineamento epidemiológico se faz entre um estudo de coorte histórica ou um estudo de caso-controle. Estes estudos têm em comum a vantagem de não ser necessário esperar um longo período de tempo para o aparecimento de casos da doença. Para doenças raras, o estudo caso-controle é o delineamento ideal por identificar casos em número suficiente para realização de análises consistentes. Para exposições de baixa prevalência, a coorte histórica poderá ser o único delineamento capaz de assegurar um número adequado de participantes expostos.

ASPECTOS METODOLÓGICOS

Cálculo do Tamanho da Amostra

O tamanho da amostra necessária para realizar um estudo de coorte pode ser estimado através de fórmulas encontradas em livros de epidemiologia ou estatística, ou através de pacotes de programas de computador. Para o cálculo do tamanho da amostra, é necessário que o investigador defina alguns parâmetros ainda na fase de planejamento do estudo; o programa EpiInfo, utiliza os seguintes parâmetros:

- (1) Razão entre participantes expostos e não expostos (geralmente esta razão é de um (1) participante exposto para um (1) participante não exposto).
- (2) Valor mínimo do Risco Relativo (RR) a ser detectado: um valor aceitável do RR, que se detectado, permitirá concluir que existe associação entre o fator de risco e a doença.
- (3) Frequência do evento de interesse no grupo de não exposto (p_0): a revisão de literatura permitirá estimar este valor.
- (4) Nível de significância (α): geralmente é especificado o valor de 0,05 que significa uma chance em 20 de se cometer o erro tipo I, isto é, de concluir que a exposição é relacionada ao evento de interesse quando na realidade isto não ocorre.
- (5) Poder do teste ($1-\beta$): geralmente é especificado um valor entre 80% e 90%. Quando especificado em 80% teríamos 20% de probabilidade de se cometer o erro tipo II (β), ou seja, de concluir que a exposição não é relacionada ao evento de interesse, quando isto ocorre na realidade.

Exemplo: cálculo do tamanho da amostra para um estudo de coorte prospectiva de um ano, planejado com o objetivo de investigar a incidência de tuberculose em indivíduos HIV positivos e que são PPD positivo e PPD negativo.

- (1) Razão entre expostos e não expostos: 1:1
- (2) Valor do Risco a ser detectado: 2
- (3) Frequência do evento nos não expostos: 8,0%
- (4) Valor de Alfa: 0,05
- (5) Poder estatístico: 0,80

A amostra necessária para condução deste estudo será de 566 indivíduos HIV(+) sendo 283 PPD(+) e de 283 PPD(-).

Dois aspectos importantes: (1) quanto maior o valor de RR, menor será o tamanho da amostra a ser estudada; um RR menor, exigirá amostra de maior tamanho; (2) os valores especificados para os parâmetros (erros alfa e beta, razão expostos/não expostos, frequência do evento entre os não expostos e, principalmente o risco relativo mínimo a ser detectado), quando modificados, alteram o tamanho da amostra.

O estudo poderá ser inviabilizado pelo número de participantes necessário à sua realização. Caberá ao investigador decidir a viabilidade de sua execução, considerando os recursos financeiros, e o tempo disponível e os riscos mínimos a serem detectados, que realmente tenham significado clínico ou biológico.

Amostragem: Recrutamento e Acompanhamento

Os critérios de elegibilidade para inclusão dos participantes na investigação devem ser estabelecidos no protocolo. É importante certificar que, na ocasião da seleção, o participante não esteja com a doença de interesse ou com outras enfermidades conseqüentes da exposição. Nos estudos concorrentes, a exposição é determinada por observação e/ou exame da população (ou de uma amostra representativa desta população). É sempre desejável que se determine o tempo e a intensidade da exposição a que estiveram (ou estão) submetidos os participantes.

A escolha do grupo de comparação (não expostos) deve obedecer os mesmos critérios empregados para seleção do grupo de expostos. Se os participantes são selecionados através de uma amostra da população, classificados de acordo com a presença ou não da exposição, então o estudo é de comparação interna. O uso de grupos de comparação interna (1) aumenta a probabilidade dos membros de ambas as coortes pertencerem a subgrupos similares da população e (2) implica que eles serão submetidos aos mesmos procedimentos durante o tempo de seguimento, tendo portanto a mesma chance de ter a doença detectada.

Em alguns estudos, quando não houver a disponibilidade de grupos de comparação interna, só a comparação externa poderá ser utilizada. Nestes casos, o grupo não exposto é selecionado como uma amostra da população geral de onde o grupo exposto foi identificado e os dois grupos podem diferir em relação a outros fatores de risco e quanto à qualidade do monitoramento.

É importante que os participantes sejam informados sobre os objetivos da investigação e que concordem em participar, através de consentimento pós-informado. Isto é essencial do ponto de vista ético e poderá garantir uma maior adesão ao estudo.

O maior problema metodológico associado à um estudo de coorte é o de manter o seguimento (aderência) da população selecionada. Métodos para contatos periódicos com a amostra de participantes devem ser estabelecidos com objetivo de minimizar as perdas (taxa de atrito). No início do estudo, deve ser programada e informada ao participante a frequência das visitas domiciliares, dos contatos telefônicos, das cartas ou telegramas. Devem também ser obtidos nomes e endereços de parentes ou de amigos, de maneira que o contato possa ser estabelecido mesmo com aqueles que mudarem da área de estudo. Deve ser ressaltada a importância do acompanhamento e de que todo esforço será feito no sentido de garantir sua participação, respeitando os limites estabelecidos no seu recrutamento.

Apesar de todos estes esforços, em um estudo de coorte uma percentagem dos participantes será sempre perdida durante o seguimento. A possibilidade de introdução de vícios devem ser sempre investigada quando o seguimento não atingir 95% da amostra selecionada. Em estudos quando o evento de interesse é o óbito, será quase sempre possível obter informações sobre os participantes perdidos (vivos ou mortos) através das declarações de óbito. Quando o interesse for adoecimento, é possível comparar os participantes que permaneceram no estudos com aqueles perdidos no acompanhamento, com relação à algumas características coletadas no início do estudo. Se as características forem semelhantes, a possibilidade de que vícios não tenham sido introduzidos no estudo é fortalecida, ou seja, não foram detectadas evidências de perda diferencial.

Em estudos de coorte não-concorrente, principalmente quando as informações sobre exposição e acompanhamento são de longo tempo (vinte ou trinta anos atrás), o maior problema é o da qualidade das informações. Métodos diagnósticos (enfermidade) ou de identificação e mensuração (exposição) podem ter mudado através do tempo, tornando as comparações impossíveis de serem realizadas.

Avaliação da Exposição

Um dos aspectos importantes nas investigações de coorte é a classificação dos indivíduos nos grupos expostos e não-expostos. Se os participantes não forem corretamente classificados, o estudo fica seriamente comprometido. As técnicas usadas para medir a exposição podem variar consideravelmente em diferentes estudos e para diferentes fatores de risco. A exposição é obtida através de entrevistas ou questionários, testes de laboratórios, avaliações clínicas ou biológicas, prontuários médicos; os membros da coorte podem então ser classificados em várias categorias da exposição, incluindo a ausência à exposição.

Em estudos prospectivos para doenças infecciosas, outras características devem ser observadas na identificação da exposição. A presença, duração e intensidade da exposição ao agente infeccioso depende da fonte de infecção e dos meios de transmissibilidade do agente. Quando a fonte de infecção e o período de transmissão são bem definidos e de um só tipo, a divisão em grupos expostos e não expostos poderá ser simples. Exemplo: na epidemia de cólera em Londres, estudada por Snow, pôde ser facilmente identificado aqueles que utilizavam e que não utilizavam água de uma determinada fonte de abastecimento (exposição).

Entretanto, quando existem múltiplas fontes de exposição ou diferentes meios de transmissão, a classificação de expostos e não expostos pode ser difícil. Quando as investigações sobre doenças infecciosas são conduzidas com o propósito de se estabelecer a natureza da exposição e os meios de transmissão, o tempo médio entre a exposição e o aparecimento da doença entre os expostos é normalmente utilizado. As investigações de toxi-infecções alimentares são um bom exemplo: se altas taxas de doenças ocorrem entre os indivíduos que comeram um alimento específico em um determinado tempo e lugar, então a provável fonte de infecção e modo de transmissão pode ser identificados.

Durante o desenvolvimento do estudo podem ocorrer mudanças na exposição aos fatores de riscos: (1) os participantes podem trocar de atividade; (2) mulheres podem modificar os métodos anticoncepcionais utilizados; (3) homo ou bissexuais masculinos podem modificar hábitos (ou intensidade) sexuais, etc. A própria participação no estudo pode levar os membros da coorte a alterarem sua exposição em consequência das informações recebidas: os participantes de um estudo sobre nutrição, questionados sobre hábitos alimentares, podem alterar sua alimentação. Quando isto ocorre, o investigador não tem controle em relação à estas mudanças; os participantes são livres para escolher e em muitos casos as modificações são desejáveis. Entretanto, é importante registrar durante o estudo todas as mudanças que ocorrerem, suas razões e os diferentes períodos de exposição, para consideração na análise. Em alguns casos, estas modificações podem ocorrer de maneira alheia ao estudo: modificação em um fator ambiental que pode alterar sensivelmente o grau de exposição.

Nos estudos de coorte não-concorrente, quando a exposição ocorreu muitos anos antes do estudo ser iniciado, as informações podem ser insuficientes para classificação em nível e duração e, muitas vezes, somente uma classificação de expostos e não expostos pode ser utilizada. A inexistência destas informações impossibilita avaliar a existência de uma relação dose-resposta entre a exposição e a doença. A utilização de informações sobre exposição, coletadas unicamente de fontes disponíveis, apresenta várias limitações: na sua grande maioria, os dados foram coletados com outras finalidades que não aquelas propostas no estudo.

Avaliação do Evento de Interesse

Em estudos concorrentes, informações para identificar a enfermidade ou morte são coletadas em exames clínicos, exames de laboratório, entrevistas ou questionário, revisão de prontuários médicos, revisão da declaração de óbito, realizadas periodicamente durante o desenvolvimento do estudo. A periodicidade depende primariamente da enfermidade que está sendo investigada. É também

importante obter informações de outras características dos grupos de estudo, tais como idade, sexo, ocupação e outros fatores de interesse com a finalidade de identificar outras variáveis que possam estar relacionados à doença ou à exposição.

Em muitos estudos, o evento de interesse é a mortalidade por uma enfermidade específica. Suas limitações serão aquelas inerentes à utilização da declaração de óbitos como fonte de dados: qualidade no preenchimento da declaração, erro na classificação da causa básica de morte, erro na codificação, etc. O estudo será também limitado às informações contidas na declaração de óbito.

A determinação da incidência depende particularmente da doença sendo estudada e das fontes de informação que estão sendo utilizadas. Independente dos métodos utilizados, os procedimentos para identificação de doenças deverão ser os mesmos para os expostos e os não expostos. Sempre que possível, para evitar vícios, o diagnóstico clínico ou laboratorial da enfermidade deve ser realizado por profissionais que desconheçam a categoria de exposição do participante.

Quando são utilizados prontuários médicos para determinar a ocorrência de doenças, é necessário atenção especial, pois as informações coletadas de diferentes hospitais e clínicas podem não ser padronizadas, os critérios diagnósticos utilizados podem variar de um clínico para outro e alguns prontuários podem ser mais completos do que outros.

Os expostos podem buscar mais atenção médica comparados aos não expostos: neste caso a ocorrência da doença pode ser superestimada, criando uma falsa associação (espúria).

A definição de caso (evento de interesse) deve basear-se em testes sensíveis e específicos, principalmente para aquelas doenças com alto percentual de formas subclínicas ou assintomáticas.

Em doenças com baixas taxas de letalidade, a mortalidade não deve ser escolhida como evento de interesse: doenças parasitárias endêmicas não são fatais e sinais e sintomas clínicos são mais adequados como evento de interesse.

Nos estudos de coorte não-concorrente, os problemas são semelhantes aqueles já discutidos quanto à exposição. Entretanto, se as perdas não forem diferenciais, ou seja, iguais entre os expostos e os não expostos, a relação entre as taxas dos eventos será válida (não viciada), embora as taxas calculadas possam sub ou super-estimar estes eventos.

ANÁLISE DOS RESULTADOS

Medidas de Associação Fator de Risco-Doença

Como já discutido anteriormente, em um estudo de coorte os dados coletados referem-se (1) à informações sobre a exposição dos participantes à um determinado fator de risco e (2) e ao desenvolvimento da doença específica em decorrência da exposição. Taxas de incidência (ou mortalidade) serão então calculadas para (1) o grupo de participantes expostos e (2) o grupo de participantes não expostos ao fator de risco sendo investigado. O objetivo será o de verificar se a taxa de incidência entre os expostos, $a/(a+b)$ é maior do que a taxa de incidência entre aqueles não expostos, $c/(c+d)$ (Tabela 1); se isto for observado, aceita-se a existência de uma associação entre o fator de risco investigado e o subsequente desenvolvimento da doença.

Tabela 1. Estrutura dos Estudos de Coorte

Exposição ou Característica	Desenvolverão doença		Total
	Sim	Não	
Presente (Expostos)	a	b	a+b
Ausentes (Não expostos)	c	d	c+d

O próximo passo será determinar a “força” desta associação através do cálculo do **RISCO RELATIVO (RR)**, definido como “a razão (ou o quociente) entre a taxa de incidência da doença no grupo exposto e a taxa de incidência da doença no grupo não exposto”.

$$RR = \frac{\text{Tx Incidência Grupo Exposto}}{\text{Tx Incidência Grupo Não Exposto}} = \frac{a/(a+b)}{c/(c+d)}$$

Um RR = 1 indica não haver associação entre o fator de risco e a enfermidade; a medida em que o valor do RR aumenta (afasta-se de 1), aumenta também a “força” desta associação. A magnitude do RR reflete, portanto, a força da associação fator de risco-doença. Fórmulas para cálculo do intervalo de confiança do RR estão disponíveis para testar sua significância estatística. Este procedimento é indicado para aqueles estudos onde o tempo de seguimento é uniforme e constante para todos os participantes no estudo.

Vários estudos de coorte, entretanto, sejam prospectivos ou históricos, requerem observações por períodos de tempo variáveis e prolongados. Durante a investigação, alguns participantes podem ser “perdidos” do estudo (recusas, abandonos, etc) ou morrerem em épocas diferentes; em consequência, a amostra sendo estudada estaria sob observação durante intervalos de tempo não uniformes. Em outros estudos, os participantes podem ser recrutados ou entram na investigação em datas diferentes (durante um período de tempo estabelecido) e se o estudo for encerrado em uma data pré-fixada, eles terão sido observados por períodos variáveis de tempo. Nestas situações, os resultados devem ser analisados utilizando-se:

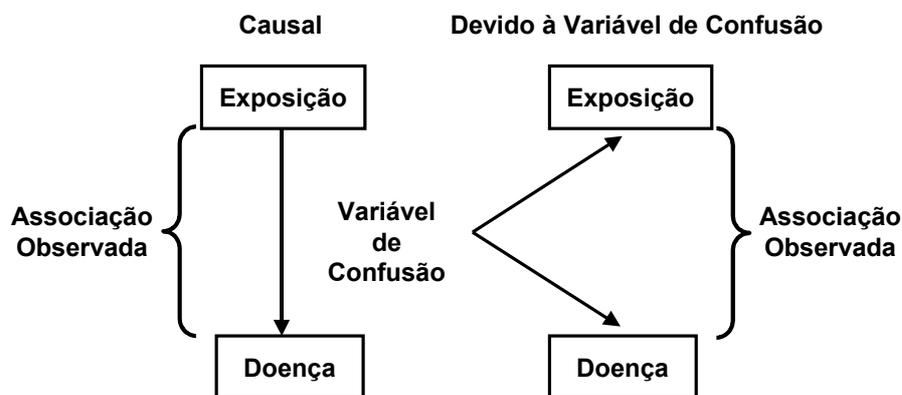
- (1) pessoas/tempo de observação como denominador para cálculo das taxas de incidência ou mortalidade ou
- (2) tábua de vida (análise de sobrevivência), para cálculo da incidência ou mortalidade acumulada.

Pessoas tempo - O conceito de pessoas/tempo de observação é freqüentemente usado como denominador para determinação de taxas em estudos de coorte, principalmente quando a existência de várias características (idade, sexo, grupo étnico, etc) e diferentes tempos de seguimento dificultam, ou mesmo impossibilitam, os cálculos necessários para a elaboração de uma tábua de vida. Leva em consideração tanto o número de pessoas que estão sendo observadas, quanto a duração dos períodos de observação, permitindo expressar em uma única estatística - eventos por pessoas/tempo de observação - o período de tempo em que um número variável de pessoas são expostas ao risco deste evento. A principal limitação para uso de pessoas/tempo é a premissa de que o risco da ocorrência de um evento por unidade de tempo é constante durante o período de observação; entretanto, este pressuposto é aceitável na maioria das investigações.

Tábua de vida - A análise de sobrevivência (tábuas de vida) é considerada por vários epidemiologistas como o método de escolha para análise de estudos de coorte, permitindo o cálculo da probabilidade de um evento por período de tempo especificado. O RR pode ser computado como a razão entre estas probabilidades. Este método deve ser utilizado quando as premissas para o uso de pessoas/tempo não puderem ser satisfeitas.

O RR calculado através de qualquer dos métodos acima discutidos é conhecido com RR “não ajustado”; não leva em consideração um possível efeito de outras variáveis, conhecidas como variáveis “de confusão”, que podem se intrometer na associação fator de risco-enfermidade. Uma variável de confusão é uma variável externa ou fator secundário que “confunde” a associação entre o fator de risco (exposição) e a enfermidade, super ou sub estimando o valor do RR encontrado (Figura 3).

Figura 3. Variável de Confusão



Para que uma variável seja de confusão deve preencher três critérios:

- (1) estar associada à exposição de interesse
- (2) ser um fator de risco para a enfermidade, independente da exposição de interesse e
- (3) não constituir elo de ligação entre a exposição e a enfermidade.

Estratificação - Quando não for indicado ajustar o efeito da variável de confusão no planejamento do estudo (emparelhamento pela variável de confusão), utilizam-se métodos na análise de dados para neutralizar seu efeito. O ajuste pode ser realizado através da estratificação dos dados pelas possíveis variáveis de confusão em múltiplas tabelas 2x2, para o cálculo do RR nos estratos e do RR ajustado (técnica de Mantel-Haenszel), que é comparado ao RR não ajustado. Uma diferença estatisticamente significativa indica a presença da variável de confusão. Modelos estatísticos, como o modelo múltiplo logístico de Cox, permitem ajustar simultaneamente para o efeito de vários fatores de confusão.

O exemplo abaixo mostra um estudo de coorte (hipotético) para investigar em indivíduos com infecção pelo *M. tuberculosis* (exposição) com e sem infecção pelo HIV (variável de confusão) o desenvolvimento de tuberculose pulmonar.

Fator de Risco Principal (exposição): Infecção pelo *M. tuberculosis*

Enfermidade (evento): Tuberculose pulmonar

Variável de Confusão: Infecção pelo HIV

- (1) **RR NÃO AJUSTADO:** Fator de Risco Principal (Infecção pelo *M. tuberculosis*) x Tuberculose pulmonar

Infecção <i>M. tuberculosis</i>	Tuberculose Pulmonar		Total
	Sim	Não	
Presente	39	701	740
Ausente	27	1244	1271

$$RR = \frac{39/740}{27/1271} = 2,48$$

RR não ajustado = 2,48

Intervalo de 95% de Confiança = 1,53 - 4,02

- (2) Analisar se a infecção pelo HIV (possível variável de confusão) está associada à doença e à exposição:

- a) Associada à Tuberculose Pulmonar?

Infecção pelo HIV	Tuberculose Pulmonar		Total
	Sim	Não	
Sim	23	180	203
Não	43	1765	1808

$$RR = \frac{22/203}{43/1808} = 4,76$$

Intervalo de 95% de Confiança = 2,93 - 7,74

b) Associada à infecção pelo *M. tuberculosis*?

Infecção pelo HIV	<i>M. tuberculosis</i>		Total
	Sim	Não	
Sim	124	79	203
Não	616	1192	1808

$$RR = \frac{124/203}{616/1808} = 1,79$$

Intervalo de 95% de Confiança = 1,58 - 2,04

(3) Ajustamento para variável de confusão: Infecção pelo HIV

a) Primeiro estrato = HIV(+)

Infecção <i>M. tuberculosis</i>	Tuberculose Pulmonar		Total
	Sim	Não	
Presente	17	107	124
Ausente	6	73	79

$$RR = \frac{17/124}{6/79} = 1,81$$

Intervalo de 95% de Confiança = 0,74 - 4,38

b) Segundo estrato = HIV(-)

Infecção <i>M. tuberculosis</i>	Tuberculose Pulmonar		Total
	Sim	Não	
Presente	22	594	616
Ausente	21	1171	1192

$$RR = \frac{22/616}{21/1192} = 2,03$$

Intervalo de 95% de Confiança = 1,12 - 3,66

(4) *Cálculo do RR ajustado* - O RR ajustado pelo método de Mantel-Haenszel é uma média ponderada dos valores dos estratos. A fórmula para cálculo pode ser encontrada nos textos avançados em Epidemiologia. Esta estimativa pode ser calculada diretamente no programa EpiInfo.

RR ajustado = 1,95

Intervalo de 95% de Confiança = 1,19 - 3,19

(5) *Interpretação*: O valor do RR não ajustado foi 2,48 e após o ajustamento e remoção do efeito da variável de confusão, o valor do RR decresceu para 1,95. A infecção pelo HIV neste estudo é uma variável de confusão; super-estimou o valor de RR.

VANTAGENS E LIMITAÇÕES

Vantagens

- 1) Os estudos de coorte permitem o cálculo do risco de desenvolver (ou falecer) uma determinada enfermidade em indivíduos expostos a um fator de risco específico em relação aqueles não expostos à este mesmo fator de risco.
- 2) O estabelecimento de critérios e procedimentos para condução do estudo antes que este se inicie, faz com que os dados sobre a exposição e a doença possam ser de excelente qualidade; os dados são coletados durante a realização do estudo. Decresce, portanto, a possibilidade de introdução de vícios ao se obter informações.
- 3) Relações entre o fator de risco de interesse e outras enfermidades podem ser investigadas; todo espectro de morbidade (gravidade) e/ou mortalidade pode ser estudado. Durante o planejamento é importante definir quais serão os eventos a serem avaliados durante o acompanhamento e o evento final do estudo. Ao contrário, nos estudos de caso-controle, uma única doença é selecionada.
- 4) Dependendo da característica e da enfermidade sendo investigadas, informações sobre aqueles participantes nos quais houve mudança de exposição ao fator de risco podem ser obtidas.
- 5) Não apresentam problemas éticos quanto à decisão em expor os participantes a fatores de risco ou tratamento como ocorre nos estudos experimentais
- 6) A seleção dos controles ao contrário dos estudos de caso-controle é relativamente simples.

Limitações

- 1) Além do alto custo, são mais difíceis de serem conduzidos, principalmente em estudos de longa duração. Mudanças administrativas e dificuldades de financiamento podem comprometer a realização do estudo.
- 2) Como requerem exames periódicos de toda a amostra, o fato de participar em um estudo pode influenciar o comportamento do exposto ou do não exposto em relação ao fator de risco de interesse e, conseqüentemente, ao desenvolvimento da enfermidade estudada.
- 3) São ineficientes para doenças raras, pois necessitam de grandes amostras para que Riscos Relativos significantes possam ser calculados.
- 4) As perdas durante o seguimento podem ser grandes, principalmente em estudo de longa duração.
- 5) Ausência de informações sobre exposição e sobre morbidade para a realização de estudos de coorte histórica.
- 6) Mudanças na categoria de exposição podem levar a erros de classificação.
- 7) Um problema comum aos estudos de observação é o de separar os efeitos da exposição principal daqueles produzidos pelos demais fatores ou variáveis extrínsecas. As variáveis extrínsecas ou de confusão podem mascarar uma possível associação entre o fator de exposição e enfermidade, super ou sub estimando os resultados.

ROTEIRO PARA DELINEAR UM ESTUDO DE COORTE

① *Defina com clareza a pergunta a ser respondida pela investigação*

- . formule a hipótese a ser testada, o objetivo geral e os objetivos específicos que serão investigados

② *Esclareça os critérios para definição de expostos e não expostos*

- . exames laboratoriais, clínicos ou informações que deverão ser coletadas
- . certeza da existência de informações sobre a exposição, principalmente para as coortes históricas

③ *Defina a população de referência a ser amostrada*

- . grupo populacional do qual serão recrutados os indivíduos expostos e não expostos
- . esclareça a fonte de seleção dos expostos e não expostos: serviços de saúde, grupos da população, indústrias, etc
- . para coortes retrospectivas certifique-se da existência de grupos de não expostos adequados para comparações

④ *Calcule o tamanho da amostra*

- . defina a relação entre expostos e não expostos
- . estabeleça o nível de significância e o poder estatístico do estudo
- . informe o valor mínimo do Risco Relativo que se detectado permitirá inferir que existe associação entre a exposição e o evento
- . estime a frequência aproximada do evento na população geral (não expostos)
- . certifique-se da factibilidade do estudo em termos logísticos e de tempo de recrutamento dos participantes
- . estime o tempo de acompanhamento versus “ponto final” do estudo

⑤ *Elabore os instrumentos para coleta de informações*

- . defina as informações que serão coletadas e as escalas de medidas das variáveis
- . elabore formulários para coleta de dados através de entrevistas, prontuários ou registros especiais
- . elabore manuais de instruções com informações e definições das questões para coleta de informações

⑥ *Defina os métodos diagnósticos para identificação do evento*

- . defina os eventos que serão avaliados durante o acompanhamento e o evento final de interesse
- . defina a interpretação dos métodos laboratoriais, exames clínicos
- . esclareça a periodicidade que serão realizados os exames durante o acompanhamento

ROTEIRO PARA DELINEAR UM ESTUDO DE COORTE

⑦ *Especifique os critérios de elegibilidade e as questões éticas*

- . faixa etária, sexo, condições biológicas e físicas
- . formas de recrutamento dos participantes
- . procedimentos para obtenção do consentimento pós-informado
- . estratégias de garantir a confiabilidade das informações coletadas
- . assistência média aos participantes que apresentarem o evento durante o acompanhamento

⑧ *Descreva a metodologia de acompanhamento dos participantes*

- . tempo de acompanhamento
- . formas e periodicidade de contato com os participantes
- . estratégias para minimizar as perdas de seguimento

⑨ *Descreva as etapas da análise de dados*

- . especifique as comparações que serão realizadas (proporções, médias) entre os grupos expostos e não expostos
- . calcule a incidência do evento de interesse entre expostos e não expostos e o Risco Relativo
- . escreva a análise que será adotada: univariada, multivariada, pessoas/tempo, análise de sobrevida

REFERÊNCIAS PARA LEITURA

- Beaglehole R, Bonita R & Kjellstrom. *Basic Epidemiology*. World Health Organization, Geneva, 1993.
- Breslow NE & Day NE. *Statistical Methods in Cancer Research*. Volume II - The Design and Analysis of Cohort Studies. IARC Scientific Publications No. 82, 1987.
- Gordis L. *Epidemiology*. Philadelphia, Editora W.B. Saunders Company, 1996.
- Khan HA & Christopher TS. *Statistical Methods in Epidemiology*. New York, Editora Oxford University Press, 1989.
- Kelsey JL, Thompson WD & Evans AS. *Methods in Observational Epidemiology*. New York, Editora Oxford University Press, 1986.
- Last JM. *A Dictionary of Epidemiology*. 3 ed., New York, Editora Oxford University Press, 1995.
- Lilienfeld DE & Stolley P. *Foundations of Epidemiology*. 3 ed. New York, Editora Oxford University Press, 1994.
- Pereira MG. *Epidemiologia: Teoria e Prática*. Rio de Janeiro, Editora Guanabara Koogan SA, 1995.
- Victora CG, Barros FC & Vaughan JP. *Epidemiologia da desigualdade: um estudo longitudinal de 6.000 crianças brasileiras*. São Paulo, Editora Hulcitech, 1988.

EXERCÍCIOS

Arquivos: coortebh.rec
ghanabl.rec
ghanafcl.rec
ghanapre.rec
ghanapos.rec
ghanasmc.rec
ghanasm.rec

1. Projeto Horizonte: Coorte de homossexuais e bissexuais masculinos HIV negativos: Um estudo de coorte aberta de homossexuais e bissexuais masculinos, HIV negativos, está sendo conduzido com os seguintes objetivos: (1) avaliar a incidência da infecção pelo HIV, (2) avaliar a possibilidade de acompanhamento desta população por um período de tempo, (3) avaliar o impacto das medidas de intervenção e (4) avaliar a possibilidade de realizar futuros ensaios clínicos. Entre 1994 e 1999, foram recrutados 470 voluntários para o estudo, acompanhados semestralmente através de entrevistas (epidemiológica e comportamental) exames clínicos e testes sorológicos para HIV e outras doenças sexualmente transmissíveis. Detalhes da metodologia estão na referência Carneiro et al, JAIDS, 2000. A análise de dados foi planejada para: (1) estudar as características dos participantes ao ingressarem na coorte, (2) analisar as perdas ocorridas durante a fase inicial do seguimento, (3) avaliar os fatores de risco associados a infecção pelo HIV e (4) calcular a incidência da infecção pelo HIV, utilizando pessoa-tempo de acompanhamento.

Questão 1. Faça uma análise descritiva dos participantes admitidos no estudo durante o período de tempo utilizando as variáveis: idade, escola, trabalha, etnia, renda, passivo, parfixo, parocas, vacinas.

Notas 1:

```
READ COORTEBH.REC
SELECT IDADE <> 99
MEANS IDADE /N
SELECT
SELECT ESCOLA <> 9
FREQ ESCOLA
SELECT
SELECT TRABALHA <> 9
FREQ TRABALHA
SELECT
SELECT ETNIA <> 9
FREQ ETNIA
SELECT
SELECT RENDA <> 9
FREQ RENDA
```

```

SELECT
SELECT PASSIVO <> 9
FREQ PASSIVO
SELECT
SELECT PARFIXO <> 8 (para excluir os códigos de não se aplica)
FREQ PARFIXO
SELECT
SELECT PAROCAS <> 8
FREQ PAROCAS
SELECT
SELECT VACINAS <> 9
FREQ VACINAS
SELECT
    
```

Questão 2. Durante o período de seguimento ocorreram perdas de acompanhamento por mudanças, morte por outras causas, desinteresse em participar do estudo, etc. As perdas ocorreram principalmente na fase inicial de acompanhamento. Construa tabelas utilizando as variáveis incluídas no estudo, comparando os participantes que foram perdidos na fase inicial do estudo (vieram somente uma visita) com os que permaneceram no estudo. De acordo com os resultados, as diferenças entre os participantes que saíram do estudo e os que permaneceram, têm significado estatístico? Interprete.

Notas 2:

```

SELECT IDADE <> 99
MEANS IDADE SITUACAO /N
SELECT
LET GPRENDA = RENDA
RECODE GPRENDA 2=1
RECODE GPRENDA 3=2
RECODE GPRENDA 4=2
RECODE GPRENDA 5=2
SELECT GPRENDA <> 9
TABLES GPRENDA SITUACAO
SELECT
SELECT TRABALHA <> 9
TABLES TRABALHA SITUACAO
SELECT
SELECT TEMSRVSD <> 9
TABLES TEMSRVSD SITUACAO
SELECT
SELECT COMPSEXO <> 9
TABLES COMPSEXO SITUACAO
SELECT
SELECT VACINAS <> 9
TABLES VACINAS SITUACAO
SELECT
SELECT VOLUNTAR <> 9
TABLES VOLUNTAR SITUACAO
SELECT
    
```

Anote os resultados da tabela voluntar/situacao.

Tecla F9 para ir ao DOS.

Digite EPITABLE opção: STUDY/CASE-CONTROL/UNMATCHED.

Calcule "Odds Ratio" considerando a categoria "Seria Voluntário" (código=1) como ausência de exposição (baseline).

Tecla F10 para sair do EPITABLE; digite EXIT para voltar ao ANALYSIS.

Questão 3. Supondo que os participantes foram acompanhados por igual período de tempo, identifique os fatores de risco associados à infecção pelo HIV nesta coorte. Compare voluntários infectados e não infectados, com relação a algumas variáveis incluídas no estudo que possam estar associadas a soroconversão. Para a variável NPAROCAS crie uma nova variável “GPPAROCAS” considerando as seguintes categorias ≤ 9 e ≥ 10 a < 888 . Analise os RR e seus intervalos de confiança. Quais são os fatores de risco associados a infecção pelo HIV?

Notas 3:

```

SELECT PASSIVO < 9
TABLES PASSIVO SOROCONV
SELECT
SELECT PARFIXO < 8
TABLES PARFIXO SOROCONV
SELECT
SELECT SEMCAMPF < 8
TABLES SEMCAMPF SOROCONV
SELECT
SELECT PAROCAS < 8
TABLES PAROCAS SOROCONV
SELECT
SELECT SEMCAMPO < 8
TABLES SEMCAMPO SOROCONV
SELECT
Transforme NPAROCAS em uma variável discreta, agrupando os valores.
DEFINE GPAROCAS #
IF NPAROCAS <= 9 THEN GPAROCAS=2
IF NPAROCAS >= 10 AND NPAROCAS < 888 THEN GPAROCAS=1
TABLES GPAROCAS SOROCONV

```

Questão 4. Identifique a existência de variáveis potenciais de confusão utilizando as variáveis: uso de drogas (drogas não injetáveis), uso de álcool e idade. Uma variável de confusão preenche as seguintes condições: (1) está associada à doença (infecção pelo HIV) (2) está associada a exposição. Considere como exposição a variável “PASSIVO” (fazer sexo anal passivo), por ter sido este o maior valor de RR encontrado. Agrupe os valores de IDADE numa nova variável “GRIDADE” considerando as seguintes categorias: ≤ 28 anos e ≥ 29 a < 99 anos.

Notas 4:

```

TABLES DROGAS SOROCONV
SELECT PASSIVO < 9
TABLES DROGAS PASSIVO
SELECT
TABLES ALCOOL SOROCONV
SELECT PASSIVO < 9
TABLES ALCOOL PASSIVO
SELECT
DEFINE GRIDADE #
IF IDADE<=28 THEN GRIDADE=1
IF IDADE>=29 AND IDADE < 99 THEN GRIDADE=2
TABLES GRIDADE SOROCONV
SELECT PASSIVO < 9
TABLES GRIDADE PASSIVO
SELECT

```

Questão 5. Ajuste suas análises para as variáveis de confusão. Compare os RR não ajustados e ajustados. Houve alteração nos valores dos RR? Quais são as variáveis de confusão?

Notas 5: SELECT PASSIVO <> 9
TABLES PASSIVO SOROCONV DROGAS
SELECT

Repita os comando anteriores para as variáveis alcool e gridade, caso você julge necessário.

Questão 6. As taxas de incidência em estudos de coorte aberta são estimadas considerando o tempo de acompanhamento de cada participante. O tempo pode ser calculado em meses, semestres ou anos de acompanhamento. Calcule as taxas de incidência para 18, 36 e 48 meses de acompanhamento, conforme os dados fornecidos na tabela abaixo.

Meses	Numerador (n1 de infectados pelo HIV)	Denominador (pessoa-ano)	Taxa de incidência 100 pessoas-ano
18	8	597,3	
36	15	855,8	
48	18	905,5	

Notas 6: Selecione DESCRIBE / PROPORTION / SIMPLE RANDOM SAMPLING.
Forneça os dados da tabela acima para o cálculo da densidade de incidência.
Pressione F10 para sair do EPITABLE
Digite EXIT para voltar para ANALYSIS
CLOSE (para fechar o arquivo)

2. Impacto de redes impregnadas com permetrina no coeficiente de mortalidade infantil no distrito Kasena, Nankana, Ghana: Um ensaio controlado randomizado. Um ensaio de comunidade controlado, randomizado, foi conduzido no norte de Ghana para avaliar o uso de mosquiteiros impregnados com permetrina no impacto da mortalidade infantil em área rural de intensa transmissão de malária sem tradição de uso de mosquiteiros. O distrito foi dividido em 96 áreas geográficas e 48 “clusters” domiciliares aleatoriamente selecionados receberam mosquiteiros impregnados com permetrina. Um sistema de vigilância foi usado para registrar nascimentos, óbitos e migração para avaliar aderência e medir a mortalidade infantil entre julho de 1993 e junho de 1995. Detalhes do estudo e metodologia estão em **Binka et al, 1996.**

Questão 7. O ensaio foi delineado para obter 90% de poder estatístico para detectar uma redução de 30% em todas as causas de mortalidade associadas com uso de mosquiteiros impregnados em crianças de 6 meses a 4 anos de idade. Cerca de 14000 crianças foram alocadas a cada um dos 2 grupos de intervenção por um período de seguimento de 2 anos. Assumindo uma taxa de mortalidade geral de 2,3 óbitos/100 crianças/ano, o número de participantes no estudo foi adequado?

Notas 7: Tecele F9 para ir ao DOS
Digite EPITABLE
Select SAMPLE/ SAMPLE SIZE / TWO PROPORTIONS

Questão 8. Utilize os dados de mortalidade fornecidos para o período pré e pós-intervenção. Calcule as taxas de mortalidade por faixa etária para cada período. Calcule os riscos relativos e intervalos de 95% de confiança para mortalidade infantil associados com uso de redes impregnadas para o período pós-intervenção. Qual sua conclusão com relação ao uso das redes impregnadas? Qual o impacto estimado nas taxas de mortalidade geral no período pós-intervenção?

Notas 8: Inicie construindo uma tabela com rótulos de linha e coluna
Comandos para Tabela 1

```
READ GHANABCL.REC
```

Calcule os anos de seguimento por grupo (tratado ou não), denominadores para o período pré-intervenção

```
MEANS FOLLYR BEDNET /N  
[ANOTE OS RESULTADOS DA COLUNA TOTAL]
```

para o somatório do seguimento por grupo de estudo e faixa etária

```
SELECT BEDNET = 1  
MEANS FOLLYR AGEGR /N  
[ANOTE OS RESULTADOS DA COLUNA TOTAL]  
SELECT  
SELECT BEDNET = 2  
MEANS FOLLYR AGEGR /N  
[ANOTE OS RESULTADOS DA COLUNA TOTAL]  
SELECT
```

```
READ GHANAFCL.REC
```

Somatório de anos de seguimento para grupos tratados ou não, denominadores para o período de pós-intervenção

```
MEANS FOLLYR BEDNET /N
```

```

Para o somatório do seguimento por grupo e faixa etária
SELECT BEDNET = 1
MEANS FOLLYR AGEGR /N
[ANOTE OS RESULTADOS DA COLUNA TOTAL]
SELECT
SELECT BEDNET = 2
MEANS FOLLYR AGEGR /N
[ANOTE OS RESULTADOS DA COLUNA TOTAL]
SELECT

```

Cálculo dos numeradores para o baseline (pré-intervenção)
 READ GHANAPRE.REC

Para a lista de total

```
TABLES OUTCOME BEDNET
```

Para cada faixa etária

```
LET AGEEXIT = TRUNC (AGEMN+FOLLMN)
```

```
DEFINE AGEGR ##
```

```
IF AGEEXIT >= 6 AND AGEEXIT < 12 THEN AGEGR = 6
```

```
IF AGEEXIT >= 12 AND AGEEXIT < 24 THEN AGEGR = 12
```

```
IF AGEEXIT >= 24 AND AGEEXIT < 36 THEN AGEGR = 24
```

```
IF AGEEXIT >= 36 AND AGEEXIT < 48 THEN AGEGR = 36
```

```
IF AGEEXIT >= 48 THEN AGEGR = 48
```

Mostre o efeito final em cada faixa etária de acordo com o grupo de tratamento

```
TABLES OUTCOME AGEGR BEDNET
```

Calcule os numeradores para o período de pós-intervenção

```
READ GHANAPOS.REC
```

Para a linha de total

```
TABLES OUTCOME BEDNET
```

Para cada faixa etária

```
LET AGEEXIT = TRUNC (AGEMN+FOLLMN)
```

```
DEFINE AGEGR ##
```

```
IF AGEEXIT >= 6 AND AGEEXIT < 12 THEN AGEGR = 6
```

```
IF AGEEXIT >= 12 AND AGEEXIT < 24 THEN AGEGR = 12
```

```
IF AGEEXIT >= 24 AND AGEEXIT < 36 THEN AGEGR = 24
```

```
IF AGEEXIT >= 36 AND AGEEXIT < 48 THEN AGEGR = 36
```

```
IF AGEEXIT >= 48 THEN AGEGR = 48
```

Mostre o efeito final em cada faixa etária de acordo com o grupo de tratamento

```
TABLES OUTCOME AGEGR BEDNET
```

Para calcular o RR e intervalo de confiança de 95% tecele F9 para ir ao DOS

Digite EPITABLE – selecione

STUDY, então selecione COHORT/CROSS SECTIONAL, então selecione

INCIDENCE DENSITY

[Use os resultados obtidos]

Tecele F10 para fechar EPITABLE

Digite EXIT para retornar ao ANALYSIS

Questão 9. Calcule o coeficiente de mortalidade específico por causa (malária, infecção respiratória aguda, gastroenterite, acidente, outras causas e causa indeterminada) para toda a população estudada no período de pós-intervenção. Como a intervenção afetou o coeficiente de mortalidade específico para malária? Compare estes resultados com os da mortalidade geral no grupo pós-intervenção obtidos na questão 2.

Notas 9: Comandos para a tabela 2 – Mortalidade específica por causa

Denominadores
READ GHANASMC.REC
MEANS FOLLYR BEDNET /N
[ANOTE OS RESULTADOS DA COLUNA TOTAL]

Numeradores
READ GHANASM.REC

TABLES BEDNET MAL
TABLES BEDNET ARI
TABLES BEDNET GASTR
TABLES BEDNET ACCIDENT
TABLES BEDNET OTHER
TABLES BEDNET UNDETMIS

Para calcular o RR e intervalo de confiança de 95% tecele F9 para ir ao DOS
Digite EPITABLE – selecione
STUDY, então selecione COHORT/CROSS SECTIONAL, então selecione
INCIDENCE DENSITY
[Use os resultados obtidos]
Tecele F10 para fechar EPITABLE
Digite EXIT para retornar ao ANALYSIS

CLOSE (para fechar o arquivo)

REFERÊNCIAS BIBLIOGRÁFICAS

- BINKA F.N., KUBAJE A, ADJUIK M. WILLIAMS L.A., LENGELER C. MAUDE G.H., ARMAH G.E., KAJIHARA B., ADIAMAH J.H. & SMITH P.G. Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. **Tropical Medicine and International Health**, **1(2)**: 147-154, 1996.
- CARNEIRO M., ANTUNES C.M.F., GRECO M., OLIVEIRA E., ANDRADE J., LIGNANI JR. L. & GRECO D.B. Design, Implementation, and Evaluation at Entry of a Prospective Cohort Study of Homosexual and Bisexual HIV-1 Negative Men in Belo Horizonte, Brazil: Project Horizonte. **Journal of Acquired Immune Deficiency Syndromes**, **25(2)**:182-187, 2000.

Arquivo: **coortebh.rec**

Variável	Descrição	Código	Descrição do código
REGISTRO	Número de Registro		
SOROCONV	Soroconversão (infecção pelo HIV)	1	Sim
		2	Não
IDADE	Idade (em anos)	99	Sem informação
ESCOLA	Grau de Escolaridade	1	Primeiro Grau
		2	Segundo Grau
		3	Terceiro Grau
TRABALHA	Atualmente está trabalhando	1	Sim
		2	Não
		9	Sem informação
RENDA	Renda mensal	1	Menos de 1 salário mínimo
		2	De 1 a 3 salários mínimos
		3	De 3 a 6 salários mínimos
		4	De 6 a 10 salários mínimos
		5	10 salários mínimos ou mais
		9	Não Respondeu
TEMSRVSD	Tem seguro saúde	1	Sim
		2	Não
		9	Sem informação
ETNIA	Cor	1	Branco
		2	Mulato
		3	Preto
		9	Sem informação
COMPSEXO	Comportamento Sexual	1	Homossexual
		2	Bissexual
		9	Sem informação
PASSIVO	Sexo anal passivo	1	Sim
		2	Não
		9	Não Respondeu
PARFIXO	Teve parceiros fixos nos últimos seis meses	1	Sim
		2	Não
		8	Não Aplica
NPARFIXO	Número de parceiros fixos	888	Não Aplica
SEMCAMPF	Fez sexo anal passivo sem camisinha nos últimos seis meses com parceiros fixos	1	Sim
		2	Não
		8	Não Aplica

Variável	Descrição	Código	Descrição do código
COMCAMPF	Fez sexo anal passivo com camisinha nos últimos seis meses com parceiros fixos	1	Sim
		2	Não
		8	Não Aplica
PAROCAS	Teve parceiros ocasionais nos últimos seis meses	1	Sim
		2	Não
NPAROCAS	Quantos parceiros ocasionais	888	Não aplica
SEMCAMPO	Fez sexo anal passivo sem camisinha nos últimos seis meses com parceiros ocasionais	1	Sim
		2	Não
		8	Não Aplica
COMCAMPO	Fez sexo anal passivo com camisinha nos últimos seis meses com parceiros ocasionais	1	Sim
		2	Não
		8	Não Aplica
ALCOOL	Utilizou álcool nos últimos seis meses	1	Sim
		2	Não
DROGAS	Utilizou drogas nos últimos seis meses	1	Sim
		2	Não
VACINAS	Já ouviu falar de vacinas	1	Sim
		2	Não
		9	Sem informação
VOLUNTAR	Seria voluntário para testes de vacinas anti-HIV	1	Sim
		2	Não
		3	Depende
		9	Não respondeu
SITUACAO		1	Voluntários que vieram somente uma vez no Projeto
		2	Voluntários que permaneceram no Projeto

Arquivo: **ghanabel.rec**

Variável	Descrição	Código	Descrição do código
BNC	Identificação	1-96	
BEDNET	Grupo de tratamento	1	Tratado
		2	Não tratado
AGEGR	Faixa etária em meses	6	6-11 meses
		12	12-23 meses
		24	24-35 meses
		36	36-47 meses
		48	48-59 meses
FOLLYR	Crianças/ano de seguimento		

Arquivo: **ghanafcl.rec**

Variável	Descrição	Código	Descrição do código
BNC	Identificação	1-96	
BEDNET	Grupo de tratamento	1 2	Tratado Não tratado
AGEGR	Faixa etária em meses	6 12 24 36 48	6-11 meses 12-23 meses 24-35 meses 36-47 meses 48-59 meses
FOLLYR	Crianças/ano de seguimento		

Arquivo: **ghanapre.rec**

Variável	Descrição	Código	Descrição do código
ID	Identificação		
BEDNET	Grupo de tratamento	1 2	Tratado Não tratado
FOLLMN	Meses de seguimento		
AGEMN	Idade em meses no final do estudo		
OUTCOME	Situação ao final do período de seguimento	0 1	Vivo Morto

Arquivo: **ghanapos.rec**

Variável	Descrição	Código	Descrição do código
ID	Identificação		
BEDNET	Grupo de tratamento	1 2	Tratado Não tratado
FOLLMN	Meses de seguimento		
AGEMN	Idade em meses no final do estudo		
OUTCOME	Situação ao final do período de seguimento	0 1	Vivo Morto

Arquivo: **ghanasmc.rec**

Variável	Descrição	Código	Descrição do código
BEDNET	Grupo de tratamento	1 2	Tratado Não tratado
FOLLYR	Crianças/ano seguimento		

Arquivo: **ghanasm.rec**

Variável	Descrição	Código	Descrição do código
ID	Identificação		
BEDNET	Grupo alocado	1 2	Tratado Não tratado
MAL	Óbito por malária	0 1	Não Sim
ARI	Óbito por infecção respiratória aguda	0 1	Não Sim
GASTR	Óbito por gastroenterite	0 1	Não Sim
ACCIDENT	Óbito por acidente	0 1	Não Sim
OTHER	Outras causas conhecidas de óbito	0 1	Não Sim
UNDETMIS	Causa de óbito desconhecida/indeterminada	0 1	Não Sim

ENSAIO

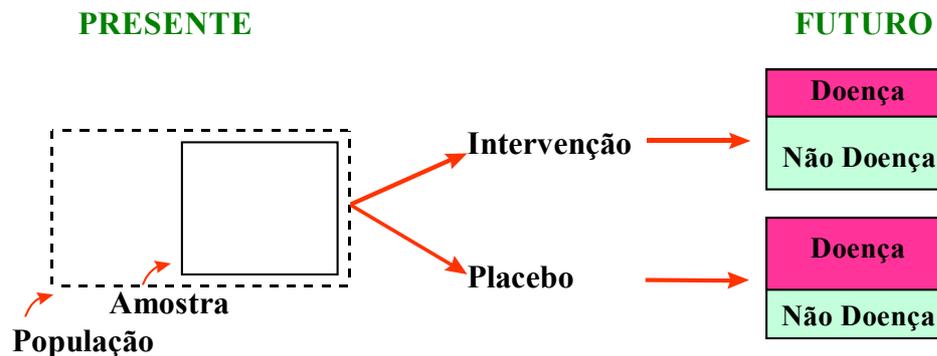
DEFINIÇÃO	92
ASPECTOS ÉTICOS	92
OBJETIVO DE UM ENSAIO CLÍNICO	93
GUIA DE QUALIDADE PARA ENSAIOS CLÍNICOS	94
CLASSIFICAÇÃO DOS ENSAIOS CLÍNICOS	
Fase I	94
Fase II.....	94
Fase III.....	94
Fase IV.....	95
DELINEAMENTO DO ENSAIO CLÍNICO	95
SELEÇÃO DOS PARTICIPANTES	
Critérios de elegibilidade.....	95
Critérios de exclusão	95
ALOCAÇÃO DA INTERVENÇÃO	
Aleatorização	96
Mascaramento	96
MANUAL DE OPERAÇÃO	96
TAMANHO DO ESTUDO	97
DOCUMENTAÇÃO	
Informação sobre o produto.....	98
Protocolo do estudo	98
Manual de operação	98
Registros do monitor.....	98
Formulário de registro de caso	98
ANÁLISE ESTATÍSTICA	
Comparação de proporções.....	99
Pessoas-tempo de seguimento.....	99
Análise Interina	100
Intenção de tratamento	100
COMITÊ INDEPENDENTE DE SEGUIMENTO	101
ROTEIRO PARA DELINEAR UM ENSAIO CLÍNICO	102
REFERÊNCIAS PARA LEITURA	104
EXERCÍCIOS	105
DICIONÁRIO DE BANCO DE DADOS	114

DEFINIÇÃO

Os ensaios clínicos são indicados para avaliar a segurança e eficácia de: (i) um novo produto; (ii) uma nova formulação de um mesmo produto ou associação de produtos já em uso e (iii) uma nova indicação clínica de um produto já aprovado. Os ensaios podem avaliar o efeito terapêutico (drogas) ou profilático (vacinas).

Toda substância para uso médico deve ter uma indicação específica, em função de seu efeito biológico desejado para o qual se elabora um ensaio clínico. O desenho do protocolo e documentação clínica dos estudos devem seguir as recomendações dos órgãos normativos e de vigilância de medicamentos do país, para que os resultados possam ser considerados válidos para aprovação do produto. Um novo produto só é levado à experimentação em seres humanos depois de conhecido seus aspectos químicos, farmacológicos, mecanismos de ação e toxicidade em provas pré-clínicas, *in vitro* ou em modelos experimentais quando disponíveis.

Como modelo de desenho epidemiológico os ensaios clínicos são sempre de natureza prospectiva. A figura abaixo mostra o fluxograma básico de um estudo de intervenção. O investigador define segundo os critérios de interesse dois ou mais grupos de comparação e administra uma intervenção-teste e uma intervenção de comparação. O seguimento é realizado baseado em parâmetros clínicos e laboratoriais previamente definidos. Os grupos de comparação devem ser similares em todos os aspectos, com exceção do tipo de intervenção recebida - as características biológicas e clínicas dos indivíduos selecionados e alocados a cada grupo, assim como as observações clínicas de seguimento, devem ser independentes dos produtos administrados.



Passos:

1. Selecionar amostra da população de referência
2. Medir as variáveis antes do início do estudo
3. Proceder a aleatorização dos participantes
4. Aplicar as intervenções de comparação
5. Acompanhar os efeitos terapêuticos ou profiláticos em ambos os grupos
6. Medir as variáveis de efeito

Adaptado de Hulley & Cummings, 1988

ASPECTOS ÉTICOS

A experimentação em seres humanos envolve aspectos de natureza ética que requerem uma avaliação cuidadosa em cada caso específico. A avaliação dos riscos da intervenção ou da não intervenção no grupo placebo, e os benefícios potenciais do estudo devem ser considerados. Os princípios da voluntariedade e confidencialidade da informação são fundamentais. Os participantes devem estar informados da natureza da investigação, da metodologia, dos exames que serão

realizados, do aspecto voluntário e da possibilidade de interromper e sair do estudo quando quiserem. Estas informações e o consentimento em participar devem ser documentados por escrito.

Sempre que possível uma intervenção de efeito conhecido deve ser utilizada como grupo de comparação. Usualmente, utiliza-se o tratamento convencional ou uma vacina nos casos onde não exista imunizante para a doença em estudo. Em situações nas quais se comprove a superioridade de uma intervenção sobre outra, todos os indivíduos não curados ou não imunizados devem ter a oportunidade de receber a medicação mais eficaz.

Os protocolos dos ensaios clínicos devem ser revisados e aprovados por um Comitê de Ética Institucional que tem por objetivo avaliar a justificativa científica para a realização do estudo, a qualificação dos investigadores, a adequação dos protocolos e documentação, os critérios de recrutamento e segurança dos participantes. Os princípios éticos estão indicados na Declaração de Helsinki, adotada originalmente em 1964, por ocasião da 18ª Assembléia Médica Mundial, e posteriormente atualizada. Estes princípios encontram-se revisados no documento “*International Ethical Guidelines for Biomedical Research Involving Human Subjects*” de 1993, editado pelo Conselho de Organizações Internacionais de Ciências Médicas (CIOMS).

OBJETIVO DE UM ENSAIO CLÍNICO

O objetivo do ensaio clínico deve ser claramente determinado a priori, especificando o produto, dose, forma de administração, tipo de paciente a que se destina, efeito esperado, parâmetros a serem medidos (toxicidade, alteração de provas bioquímicas, resposta imunológica, efeito terapêutico ou preventivo). O delineamento da investigação, o cálculo do tamanho da amostra, procedimentos de monitoramento dos participantes e a interpretação dos resultados finais do estudo dependem de uma definição precisa do objetivo do estudo. Em alguns casos é possível definir um objetivo principal (baseado em um efeito principal a ser avaliado) e objetivos secundários (baseado em um efeito secundário a ser estudado).

Exemplos de objetivos :

Ensaio para avaliação de drogas - Determinar se a administração da droga A na dose de 20mg/kg/dia x 7 dias via IM em pacientes adultos não tratados previamente, com lesão única por leishmaniose cutânea apresenta um percentual de cura (definido como 100% de reepitelização e negatificação parasitológica de lesões) pelo menos 30% superior ao tratamento convencional com droga B na dose de 50mg/kg/dia x 20 dias via IM, durante um período de observação clínica de 4 meses.

Ensaio para avaliação de vacina - Determinar se o produto QT01-97, administrado por via subcutânea em duas doses com intervalo de 30 dias em crianças de 1-5 anos, residentes em áreas endêmicas de malária, é capaz de reduzir em pelo menos 50% o número de episódios clínicos de malária diagnosticados através de vigilância ativa e passiva durante um período de 12 meses de seguimento após a vacinação.

Note que nos ensaios clínicos é sempre necessário estabelecer uma **definição de caso** e **definição do efeito final** a observar. A padronização do critério diagnóstico de caso é utilizada no recrutamento de pacientes para ensaios terapêuticos e na vigilância de casos incidentes em estudos profiláticos. A definição de caso tem implicações na validade externa dos resultados possibilitando a extrapolação das conclusões.

GUIA DE QUALIDADE PARA ENSAIOS CLÍNICOS

A Organização Mundial da Saúde tomando como referência a experiência de vários países publicou o documento “*Guideline for good clinical practice for trials on pharmaceutical products - GCP (WHO Technical Report Series no.859, 1995, pgs 97-137)*” que constitui um guia de delineamento, condução e relatório de estudos clínicos. Este documento apresenta normas para a condução de ensaios clínicos que permitem que os resultados sejam reconhecidos internacionalmente e adequados para o registro do produto. O guia inclui recomendações relacionadas aos aspectos éticos e proteção dos participantes, responsabilidades do investigador em todo o processo de condução e análise da investigação, responsabilidades do patrocinador, responsabilidades do monitor, seguimento clínico de toxicidade, manutenção de registros, plano de análise e papel das autoridades de regulamentação de medicamentos.

CLASSIFICAÇÃO DOS ENSAIOS CLÍNICOS

Os ensaios clínicos são classificados em quatro fases I-IV segundo seu nível de complexidade, estágio de desenvolvimento do produto a ser testado e objetivo da avaliação.

Fase I

Constitui a primeira etapa de avaliação de um produto químico/biológico em seres humanos. Os ensaios de Fase I são geralmente precedidos de provas em modelos experimentais em animais para avaliar toxicidade e eficácia. Quando estes não são disponíveis ou os resultados não podem ser extrapolados para seres humanos é necessário iniciar experimentação clínica sem conhecimento da eficácia dos produtos. Os estudos de Fase I devem ser conduzidos no país de produção da droga ou vacina (mesmo quando o produto não se destina à esta população). São realizados sob rigorosa supervisão médica, usualmente em hospitais, e envolvem um número limitado de voluntários adultos sadios geralmente do sexo masculino. O objetivo principal nesta fase é avaliar toxicidade e farmacocinética do produto.

Fase II

São ensaios clínicos pilotos limitado à um pequeno número de participantes ou pacientes com o objetivo de demonstrar atividade terapêutica (droga) ou atividade imunogênica (vacina). Os ensaios de Fase II requerem um grupo de comparação. Avalia-se também a toxicidade do produto nos indivíduos (ou pacientes) para os quais o produto está sendo desenvolvido. Nesta fase se realizam os estudos de dose-resposta com o propósito de encontrar a dose e esquema ótimo de administração do produto. Na avaliação de antígenos candidatos a vacinas se estuda a produção e cinética de anticorpos e resposta imunológica celular. Na avaliação de vacinas se reconhece uma Fase IIa onde se realizam inóculos padronizados do agente infeccioso (desafio artificial) após imunização para se verificar o efeito protetor da vacina candidata. Estes estudos permitem otimizar em tempo e custo a avaliação de eficácia de vacinas. A Fase IIb se refere à avaliação de eficácia em situação de desafio natural, ou seja, pela exposição natural à infecção em áreas onde há transmissão.

Fase III

Ensaio de Fase III são considerados críticos para o registro e aprovação de um produto farmacêutico. Envolvem um grande número de participantes, eventualmente em estudos multicêntricos, quando se inclui vários grupos de pacientes tratados em serviços distintos - sempre utilizando o mesmo protocolo de investigação. O objetivo principal é demonstrar eficácia e inocuidade a curto e longo prazo. Na avaliação da eficácia de vacinas estes estudos devem ser conduzidos na população em geral, selecionando aqueles indivíduos que serão alvo de vacinação futura (por exemplo, crianças durante o primeiro ano de vida). Estes ensaios devem ser realizados em condições semelhantes ao uso rotineiro futuro da intervenção. É fundamental que sejam delineados como estudos aleatorizados, duplo cego, controlados.

Fase IV

Esta Fase refere-se a ensaios clínicos realizados após aprovação, registro e comercialização do produto farmacêutico. Estes estudos se destinam principalmente a avaliar a ocorrência de efeitos adversos raros ou desconhecidos. Em intervenções em saúde pública, como no caso de vacinas, os estudos de Fase IV permitem: (i) avaliar estratégias operacionais alternativas para administrar a intervenção; (ii) conhecer a duração do efeito (imunidade); (iii) avaliar o efeito da intervenção em situações epidemiológicas distintas; e (iv) avaliar o impacto epidemiológico da intervenção na transmissão da doença. Após o produto estar disponível no mercado, os ensaios clínicos delineados com o objetivo de explorar uma nova indicação, novas combinações de drogas ou vias alternativas de administração devem ser considerados como um ensaio de um produto farmacêutico novo.

DELINEAMENTO DO ENSAIO CLÍNICO

Os ensaios clínicos em geral, e especialmente aqueles para avaliação de eficácia, devem ser delineados como estudos de intervenção comparativa incluindo pelo menos dois grupos de estudo - Grupo de Intervenção e Grupo Controle. No delineamento clássico em paralelo os participantes são alocados aos dois grupos de estudo de forma aleatória e às cegas, para assegurar que os participantes apresentem características semelhantes e que os resultados sejam analisados de forma comparativa e imparcial. O grupo controle usualmente recebe uma substância sem atividade farmacológica (placebo) ou outro produto farmacológico com atividade terapêutica ou profilática conhecida. Este tipo de desenho de investigação é descrito como ensaio clínico controlado, duplo-cego, aleatorizado. A denominação duplo-cego se refere ao mascaramento que se realiza na aplicação das intervenções e seguimento dos participantes - os profissionais responsáveis pelo seguimento clínico-epidemiológico e laboratorial e os próprios participantes voluntários desconhecem qual dos produtos em avaliação foi administrado (substância ativa ou placebo). Outros tipos de delineamento para ensaio clínico como estudo sequencial e "**cross-over**" podem ser usados em situações especiais. Maiores informações podem ser obtidas em publicações especializadas.

SELEÇÃO DOS PARTICIPANTES

Critérios de elegibilidade - Os critérios para seleção dos participantes do ensaio dependem fundamentalmente do objetivo do estudo. Estes critérios devem estar claramente enunciados no protocolo. De modo geral os participantes devem representar o grupo de indivíduos/população para os quais o produto foi desenvolvido e aqueles que mais poderiam se beneficiar da intervenção. Em ensaios terapêuticos se incluem pacientes utilizando um critério diagnóstico definido. Com a finalidade de reduzir o número de fatores que possam dificultar a realização do estudo, modificar o efeito da intervenção ou confundir a interpretação dos resultados é comum restringir o estudo a indivíduos segundo algumas características gerais, por exemplo, idade, local de residência, possibilidade de comparecer aos exames de seguimento, ausência de tratamento anterior.

Em estudos que dependem do recrutamento de casos incidentes, por exemplo, ensaios de vacinas, é necessário selecionar áreas de maior transmissão onde o número esperado de casos é suficiente para a realização do estudo e interpretação dos resultados. Populações estáveis, com baixas taxas de migração devem ser escolhidas para estudos que requerem seguimento epidemiológico de longa duração com avaliações sucessivas.

Critérios de exclusão - São características especiais que colocam os indivíduos em uma situação de risco se incluídos nos estudos. Por exemplo, gestantes, crianças desnutridas, indivíduos em uso de outros medicamentos com potencial de interação química e doenças crônicas. Estes critérios devem estar definidos no protocolo e fazer parte da rotina e do guia de operação do ensaio clínico.

Em caso de se observar manifestações clínicas graves decorrentes ou não das intervenções, os pacientes devem ser retirados do estudo. O código de tratamento deve ser aberto para investigar causas. Assistência médica deve ser oferecida e o caso documentado e incluído no relatório final.

ALOCAÇÃO DA INTERVENÇÃO

Dois princípios gerais estão envolvidos na alocação dos participantes aos grupos de estudo - aleatorização e mascaramento.

Aleatorização

Aleatorização é o processo de distribuição ao acaso de pacientes ou voluntários sadios às diferentes intervenções em comparação. A aleatorização constitui um dos elementos fundamentais do desenho do ensaio clínico. Seu objetivo é garantir uma distribuição balanceada de participantes aos grupos de comparação reduzindo ou eliminando o viés de alocação de pacientes a uma determinada intervenção e a avaliação tendenciosa dos efeitos da intervenção. A unidade de aleatorização pode ser o indivíduo ou um grupo de indivíduos por ocasião do recrutamento (utilizando-se, por exemplo, envelopes selados com códigos dos tratamentos); ou pode ser aplicada à própria intervenção (codificando adequadamente vidros de comprimidos ou seringas) e alocando os participantes em seqüência.

Aleatorização em bloco – (por exemplo, a cada grupo de 10 indivíduos) utiliza-se para assegurar uma distribuição balanceada de participantes quando o número de participantes do estudo é pequeno. A aleatorização pode ser também **estratificada**, por exemplo, por local de recrutamento quando existe interesse em realizar análises por centro de recrutamento ou quando existem razões para considerar uma resposta diferente de acordo com o local de tratamento ou recrutamento.

Em ensaios populacionais, se recomenda preparar previamente uma lista de voluntários que atendam os critérios de inclusão e aleatorizar os indivíduos por número de identificação. Este procedimento tem a vantagem de permitir a repetição do processo de aleatorização várias vezes até se conseguir um balanceamento das variáveis críticas na avaliação do efeito da intervenção entre os dois grupos, por exemplo: idade, distância do serviço de saúde, etc.

Mascaramento

Mascaramento é o processo utilizado para que o paciente e o investigador não tenham conhecimento sobre a alocação individual dos participantes aos grupos de tratamento. O mascaramento em um ensaio clínico é necessário para evitar viés de observação, durante o seguimento clínico-laboratorial, mantendo-se assim completa imparcialidade na avaliação dos efeitos. Para garantir um mascaramento perfeito deve haver completa independência entre os investigadores que aplicam a intervenção (droga ou vacina) e aqueles responsáveis pelo seguimento clínico. Os produtos utilizados nos grupos intervenção e controle devem, na medida do possível, ter aspecto externo semelhante (cor, tamanho e forma) e esquema posológico idêntico (número de comprimidos, freqüência) para que não seja possível identificar diferenças entre os indivíduos que recebem um ou outro medicamento. É necessário estabelecer um sistema de codificação específico utilizando-se letras ou números, que deve ser guardado pelo monitor ou comitê supervisor.

MANUAL DE OPERAÇÃO (SOPs)

São instruções precisas, escritas em forma de um manual, para a realização de todos os procedimentos previstos no ensaio clínico, incluindo recrutamento de participantes, alocação aos

grupos de estudo, administração da intervenção, sistemas de registros, critérios para interrupção da intervenção, etc. Todas as atividades a serem realizadas devem estar previamente estabelecidas em forma de um roteiro de tarefas distribuídas à equipe de investigação. Este guia permitirá ao monitor externo supervisionar a qualidade e aderência do estudo ao protocolo.

TAMANHO DO ESTUDO

Nos ensaios clínicos clássicos nos quais se compara em paralelo o efeito de duas drogas/vacinas, a análise de dados baseia-se usualmente na comparação de duas proporções/taxas de cura/proteção: grupo-intervenção vs. grupo controle. O cálculo do tamanho da amostra nesta situação é relativamente simples e requer as seguintes informações:

- (i) valor mínimo da diferença a ser detectada entre os grupos
- (ii) razão entre o número de participantes no grupo 1 e 2
- (iii) nível de significância; geralmente é especificado o valor $\alpha = 5\%$
- (iv) poder do teste ($1-\beta$); geralmente é especificado o valor $1-\beta = 80\%$

Por exemplo, em um estudo para avaliar a eficácia da combinação atovaquone-proguanil no tratamento da malária em comparação com amodiaquina como grupo controle, um tamanho da amostra foi calculado para se estimar uma diferença entre uma possível taxa de parasitemia negativa de 90% no grupo experimental versus 65% no grupo amodiaquina ao final de 28 dias. Nesta situação, seriam necessários cerca de 65 pacientes em cada um dos grupos de tratamento, para que o estudo tivesse um poder estatístico de 90% para detectar esta diferença de 25% de cura entre os grupos de tratamento. Para se obter o tamanho da amostra total de 130 pacientes, utilizou-se as seguintes informações (ver Epiinfo-Epi-table-sample opção-duas proporções):

- (i) razão entre o número de participantes no grupo 1 e 2 = 1
- (ii) proporção de cura esperada no grupo 2 = 90%
- (iii) proporção de cura esperada no grupo 1 = 65%
- (iv) valor $\alpha = 5\%$
- (v) poder do teste = 90%

Total de participantes por grupo = 65

Entretanto, este cálculo pode se tornar mais complicado quando é necessário ajustar para outros fatores como, por exemplo, um número desigual de participantes em cada grupo, interesse em fazer análises estratificadas por subgrupos de pacientes, interesse em fazer análise interina antes de terminar o estudo. Cálculos especiais são necessários para desenhos mais complicados e quando se comparam múltiplos efeitos.

Para estudos de avaliação de vacinas nos quais se comparam taxas de incidência baseadas em pessoas-tempo de exposição, o cálculo de tamanho de amostra é semelhante ao de estudos de coorte.

Para estimativas mais elaboradas, como por exemplo, aquelas necessárias para cálculo de amostras para obter uma determinada precisão ou para pré-estabelecer o poder de detecção de uma diferença sugerimos consultar Smith & Morrow (1996). A assessoria de um estatístico é importante nesta etapa de delineamento do projeto para analisar os requerimentos de tamanho de amostra em relação à proposta de análise dos dados.

DOCUMENTAÇÃO

Um ensaio clínico bem organizado, e especialmente aqueles que preparam documentação para registro de drogas, devem ter um conjunto de documentos e sistema de registro que permitam uma análise exaustiva da qualidade do estudo.

Informação sobre o produto - usualmente fornecida pelo laboratório produtor, deve incluir a descrição do produto, características de produção, controle de qualidade, resultados de estudos pré-clínicos, e em desenvolvimento.

Protocolo do estudo - descrição detalhada dos fundamentos, desenho, operação e plano de análise do estudo.

Manual de operação - descrição detalhada de todos os procedimentos utilizados no ensaio clínico.

Registros do monitor - formulários de observação/supervisão das visitas inicial/viabilidade, de iniciação, de seguimento periódico e de término do estudo.

Formulário de registro de caso - registro individual dos participantes (ficha clínica) contendo todos os dados clínicos, laboratoriais e de seguimento, que permitirão a supervisão oportuna da execução do estudo.

ANÁLISE ESTATÍSTICA

Os procedimentos analíticos de um ensaio clínico vão depender da natureza do estudo, dos parâmetros estimados, e dos efeitos que foram medidos. O plano de análise do ensaio clínico deve estar incluído no protocolo. Neste plano se identificam os efeitos principal e secundários que se quer avaliar e a forma de comparação. Critérios para definição de variáveis de análise, definição de variáveis de confusão e subgrupos de comparação devem estar definidos segundo uma lógica científica. O estabelecimento *a priori* do plano de análise evita a análise estatística exploratória, com múltiplas comparações, que eventualmente levarão a encontrar uma diferença significativa sem fundamentação científica.

O relatório final conclusivo do estudo deve estar limitado às comparações propostas no plano de análise pré-estabelecido. A manipulação exaustiva dos dados, buscando comparações/diferenças não previstas no delineamento do estudo é motivo para perda de credibilidade no resultado final.

Comparação das características básicas entre os grupos de intervenção - Como regra geral, a análise de dados deve inicialmente verificar a comparabilidade (decorrente da aleatorização) dos grupos incluídos no estudo em relação às características eventualmente associadas à resposta ao tratamento ou risco de infecção. Em estudos terapêuticos são usualmente a idade, sexo, tempo de doença, gravidade do quadro clínico. Nos estudos de avaliação de vacinas pode-se incluir ainda a avaliação de níveis de anticorpos prévios à vacinação, local de residência, condições sócio-econômicas, etc.

Comparação de proporções - Para comparação de proporções utiliza-se o teste qui-quadrado, e estima-se intervalos de 95% de confiança para a diferença encontrada entre os tratamentos. Este intervalo indica que, tomando em conta o número de participantes no estudo, o benefício adicional real da intervenção estaria situado dentro daqueles limites calculados.

A **EFICÁCIA** de uma INTERVENÇÃO é calculada pela fórmula:

$$\frac{\% \text{ falha terapêutica no grupo controle} - \% \text{ falha terapêutica no grupo intervenção}}{\% \text{ falha terapêutica no grupo controle}}$$

Exemplo – A tabela abaixo mostra o resultado de um ensaio clínico que recrutou 50 pacientes em cada um dos grupos. O tratamento experimental não curou 5/50 (10%) dos indivíduos e o tratamento placebo falhou em 37/50 (74%) dos indivíduos, resultando uma eficácia terapêutica de 86,5% do tratamento experimental.

GRUPO	FALHA	SUCESSO	TOTAL
Experimental	5	45	50
Placebo	37	13	50
TOTAL	42	58	100

$$EF = \frac{(37/50) - (5/50)}{37/50} = 86,5\%$$

Este mesmo resultado pode ser obtido utilizando-se a fórmula:

Eficácia = 1-RR	Onde:
------------------------	--------------

RR = 0,14

Eficácia = 1-0,14 = 86,5% (IC 95% 68,5% - 94,2%)

Para estudos tipo caso-controle pode-se utilizar 1-OR.

Pessoas-tempo de seguimento - Nos estudos que envolvem observação de grupos de participantes por um tempo prolongado, deve-se tomar em conta a possibilidade de perdas de seguimento. Neste

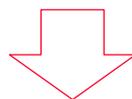
caso a estimativa mais adequada seria a densidade de incidência, que considera pessoas-tempo de seguimento para estimar taxas. Vários métodos estatísticos podem ser usados para este tipo de análise. A construção de tábuas de vida com estimativa do teste log-rank compara o tempo de ocorrência dos eventos de interesse do estudo. Curvas de sobrevivência permitem analisar a velocidade ou ritmo de ocorrência dos eventos. O cálculo de risco relativo e limites de confiança comparando taxas de incidência ou de cura podem ser obtidos através de regressões logísticas multivariadas (Poisson), permitindo o cálculo de eficácia terapêutica ou preventiva.

Análise “Interina” - A realização de uma análise interina/intermediária de dados é recomendada em alguns estudos de longa duração nos quais um nível pré-estabelecido de eficácia ou de suscetibilidade/intensificação da doença pode ser testado antes de se concluir o seguimento do estudo. O monitor independente do estudo fornece os códigos para um estatístico independente que analisa os dados em função de critérios definidos no protocolo. Por exemplo, uma análise interina em um estudo de avaliação de vacinas com 2 anos de duração poderia investigar se existe uma evidência significativa de proteção ou aumento de suscetibilidade à infecção ao final do primeiro ano de seguimento. Em qualquer das duas situações (proteção ou aumento de suscetibilidade) o estudo deveria ser interrompido por razões éticas.

Intenção de tratamento - Uma análise de dados mais rigorosa é chamada de intenção de tratamento e inclui todos os indivíduos que iniciaram o ensaio clínico, independente se eles concluíram a intervenção e período de seguimento. Considera-se como falha terapêutica todos os casos que foram retirados do estudo por razões de efeitos colaterais e aqueles nos quais não foi possível completar o seguimento de avaliação.

Por exemplo, a Tabela a seguir apresenta os resultados de um estudo que comparou a eficácia da combinação artemisina/mefloquina versus mefloquina no tratamento da malária por *P. falciparum*. A eficácia calculada incluindo somente os participantes que concluíram o seguimento seria 71,1%.

Tratamento	Falha	Cura	Perdas	TOTAL avaliados	TOTAL randomizados
Artemisina + mefloquina	26	194	20	220	240
mefloquina	94	136	10	230	240
TOTAL	120	330	30	450	480

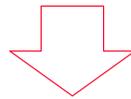


Eficácia considerando somente os pacientes que completaram o seguimento

$$= \frac{(94/230)-(26/220)}{(94/230)} = 71,1\% \text{ (IC95\% 57,2-80,5)}$$

No entanto, o cálculo da eficácia por intenção de tratamento, considerando todas as perdas de seguimento como falha terapêutica (Tabela abaixo), produziria um resultado de eficácia mais conservador:

Tratamento	Falha	Cura
Artemisina + mefloquina	46 (26+20)	194
Mefloquina	104 (94+10)	136
TOTAL	150	330



$$\text{Eficácia intenção de tratamento} = \frac{(104/240)-(46/240)}{(104/240)} = 55,8\% \text{ (IC95\% 40,5-67,1)}$$

Observe que os resultados da eficácia devem ser apresentados com os respectivos intervalos de 95% de confiança. No exemplo acima, embora as taxas de eficácia sejam diferentes, não houve diferença estatisticamente significativa entre a eficácia calculada por intenção de tratamento ou considerando somente os pacientes que concluíram o seguimento clínico, pois os IC95% apresentam sobreposição.

COMITÊ INDEPENDENTE DE SEGUIMENTO

Em alguns estudos, além de um monitor clínico independente, usualmente contratado pela companhia farmacêutica que supervisiona todas as etapas do ensaio clínico e a aderência ao protocolo, estabelece-se um comitê técnico de revisão independente, responsável por aprovar o desenho da investigação, proceder à análise dos aspectos éticos, e seguir as operações e registros do estudo dando um aval de qualidade à investigação. Este comitê deve ter a autoridade de terminar o ensaio clínico em função de eventos inesperados que prejudiquem a conclusão satisfatória do estudo ou pela ocorrência de efeitos colaterais importantes.

ROTEIRO PARA DELINEAR UM ENSAIO CLÍNICO

① *Formule com clareza a pergunta a ser respondida pelo estudo*

- especifique nos objetivos o tipo de intervenção (dose, via de administração, duração), as características dos indivíduos que receberão a intervenção (doença de base), tempo e método de seguimento, a variável que será utilizada para avaliar o efeito da intervenção, o tamanho da diferença que o estudo pretende detectar entre os grupos de intervenção e o poder estatístico do estudo

② *Esclareça as questões éticas*

- descreva os procedimentos para obtenção do consentimento informado
- defina quando serão tratados os indivíduos do grupo controle nos casos em que a intervenção se mostrar eficaz
- estabeleça as regras de decisão para retirada dos participantes do estudo e para interromper o estudo
- esclareça os riscos e benefícios da intervenção e o tipo de assistência médica que será garantida aos participantes
- informe a instância que o protocolo será submetido para aprovação ética

③ *Defina os critérios de elegibilidade e a definição de caso*

- características demográficas (sexo, idade, etc)
- testes laboratoriais, exame clínico
- interpretação e categorização dos parâmetros a serem avaliados

④ *Descreva as intervenções que serão aplicadas nos grupos de estudo (intervenção/controle)*

- produto químico, procedência
- via de administração, dose, tempo
- procedimentos para avaliar a aderência ao protocolo - intervenção sob supervisão, dosagens laboratoriais da droga

⑤ *Descreva os métodos para recrutamento e alocação dos participantes à intervenção*

- fonte de seleção dos participantes
- alocação dos participantes à intervenção: aleatorização simples, estratificada, em bloco
- mascaramento: duplo-cego, triplo-cego

⑥ *Descreva os métodos para avaliar o efeito da intervenção*

- métodos laboratoriais, parâmetros a serem avaliados, interpretação, exame clínico
- efeitos adversos da intervenção e como serão avaliados – testes laboratoriais, exames clínicos, periodicidade e interpretação

ROTEIRO PARA DELINEAR UM ENSAIO CLÍNICO

⑦ *Descreva a metodologia para seguimento dos participantes*

- . estratégias que serão utilizadas para minimizar as perdas de seguimento durante o estudo
- . guarda de códigos, critérios para quebra de códigos

⑧ *Calcule o tamanho da amostra*

- . especifique o tamanho mínimo da diferença que se espera detectar entre os grupos de intervenção
- . defina o poder estatístico do estudo para detectar a diferença entre os grupos e o nível de significância estatística (alpha)
- . certifique-se da factibilidade do estudo em termos logísticos e de tempo de recrutamento dos participantes
- . estime o tempo de follow-up versus “ponto final” do estudo

⑨ *Descreva as etapas para análise de dados*

- . indique os parâmetros (proporções, médias), métodos estatísticos e sub-grupos para avaliação das características básicas dos grupos de estudo e do efeito da intervenção
- . esclareça o tipo de análise que será adotada - “intenção de tratamento”, pessoas/tempo, análise de sobrevida, análise intermediária

REFERÊNCIAS PARA LEITURA

- BERQUÓ, E.S., SOUZA, J.M.P. & GOTLIEB, S.L.D. *Bioestatística*, São Paulo: Editora Pedagógica e Universitária Ltda., 1981.
- BRASIL. MINISTÉRIO DA SAÚDE. Diretrizes e Normas Regulamentadoras de Pesquisas Envolvendo Seres Humanos. Brasília, 1997, 20 pg.
- CLAYTON, D. & HILLS, M. *Statistical Models in Epidemiology*, New York:Oxford University Press, 1993.
- COUNCIL FOR INTERNATIONAL ORGANIZATIONS OF MEDICAL SCIENCES. International ethical guidelines for biomedical research involving human subjects. Geneva, CIOMS, 1993, Annex 1.
- HULLEY, S.B. & CUMMINGS, S.R. *Designing clinical research*. Williams & Wilkins Baltimore, 1988.
- INFORME EPIDEMIOLÓGICO DO SUS. Diretrizes Nacionais e Internacionais para pesquisas em seres humanos. Fundação Nacional de Saúde, ano IV, 1995.
- MEINERT, C. *Clinical trials*. New York, Oxford University Press, 1986.
- PEREIRA, M.G. *Epidemiologia Teoria e Prática*. Rio de Janeiro. Guanabara Koogan, 1995.
- POCOCK, S.J. *Clinical trials - a practical approach*. A Wiley Medical Publication, 1983.
- SMITH, P.G. & MORROW, R.H. A "tool-box" for field trials of intervention against tropical diseases. *UNDP/World bank/ WHO Special Programme for research and training in tropical diseases(TDR)*. Geneva, 1988.
- WORLD HEALTH ORGANIZATION - WHO Technical Report Series, No 850. *Guidelines for good clinical practice (GCP) for trials on pharmaceutical products*. Geneva, 1995, pgs 97-13

EXERCÍCIOS

Arquivos: 1. ensaioml.rec
2. ensaio.rec
3. bam.rec

1. **Eficácia da cloroquina na malária por *P. falciparum*.** Um ensaio clínico duplo-cego foi delineado para comparar a eficácia do tratamento da malária por *Plasmodium falciparum* com cloroquina oral nas doses de 50 mg/kg e 25 mg/kg. Os níveis de parasitemia foram avaliados diariamente até o 7º dia e no 14º, 21º e 30º dias. O arquivo **ensaioml.rec** contém informações de 124 pacientes alocados aleatoriamente a ambos os grupos e atendidos em ambulatório da Fundação Nacional de Saúde na cidade de Goiânia (área não endêmica para malária). Detalhes da metodologia estão na referência **Andrade et al, 1992**. Para este exercício, a estratégia de análise foi elaborada com os seguintes objetivos: (1) comparar as características básicas dos participantes no início do ensaio; (2) avaliar os níveis de parasitemia nos dias 7 e 30 após o tratamento.

Questão 1. Qual o número de pacientes alocados ao tratamento de 50 mg/kg e 25 mg/kg?

Notas 1: READ ENSAIOML.REC (para abrir o arquivo)
FREQ GP

Questão 2. Construa uma tabela com a linha de base dos participantes comparando a idade (“AGE”), sexo (“SEX”), ocupação (“OCUP”), medicação utilizada no último episódio de malária (“MED”), e parasitemia (“PARD0”) entre os grupos. Para comparar o número de episódios de malária nos últimos 5 anos (“XMAL”) entre os grupos, criar nova variável (“GRXMAL”) e agrupar de 1-3; 4-6 e ≥ 7 episódios. Faça o mesmo para o tempo de início dos sintomas (“SIND”), criando uma nova variável (“GRSIND”) agrupando de 1-5; 5-10 e ≥ 11 dias. O que pode ser concluído em relação à comparabilidade entre os grupos?

Notas 2:

MEANS AGE GP /N

para comparar utilizando percentuais (%) utilize o comando

SET PERCENTS=ON

TABLES SEX GP

TABLES OCUP GP

TABLES MED GP

SET PERCENTS=OFF (para desativar a %)

para comparar a parasitemia utilize a média geométrica da densidade de parasitemia criando nova variável (“LPARD0”)

DEFINE LPARD0 ##.### (média geométrica da densidade de parasitas)

LET LPARD0=LOG(PARD0) (logaritmo da parasitemia antes do tratamento)

MEANS LPARD0 GP /N

anote as médias, desvios padrões e tamanho de cada grupo

tecle F9 para ir ao DOS

digite EPITABLE para cálculo do IC (95%) (comando DESCRIBE / MEANS)

utilize os valores anotados

anote IC calculados

para transformar as medias e IC 95% obtidos com o logaritmo das parasitemias

utilize a base 10 com o auxílio de uma calculadora

Exemplo: $10^{2,993}$ ($10^{\text{média}}$; média = 2,993); $10^{\text{limite inferior do IC 95\%}}$; limite inferior do IC 95% = 2,89); $10^{3,10}$ ($10^{\text{limite superior do IC 95\%}}$; limite superior do IC 95% = 3,10)

Faça a transformação para os 2 grupos de tratamento

pressione F10 para sair do EPITABLE

digite EXIT para voltar ao ANALYSIS

LET GRXMAL=XMAL (criando agrupamento dos episódios de malária)

RECODE GRXMAL 1-3=1 4-6=2 7-20=3

SET PERCENTS=ON

TABLES GRXMAL GP

LET GRSIND=SIND (para criar agrupamento do tempo do início dos sintomas)

RECODE GRSIND 1-5=1 6-10=2 11-40=3

TABLES GRSIND GP

Questão 3. Utilize a estratégia de “intenção de tratamento” e calcule a taxa de “cura” da cloroquina no 7º e 30º dias, em relação ao “clearance” da parasitemia, para as doses de 50 mg/kg e 25 mg/kg. O que pode ser concluído em relação à “eficácia” da cloroquina para tratamento da malária por *P. falciparum* na área de estudo?

Notas 3: para o cálculo da taxa de “cura” da cloroquina por “intenção de tratamento” considere como falha terapêutica todas as perdas de seguimento.
 RECODE SITD7 3=2 (para recodificar a perda de seguimento como parasitemia positiva no Dia 7)
 TABLES SITD7 GP
 RECODE SITD30 3=2 (para recodificar a perda de seguimento como parasitemia positiva no Dia 30)
 TABLES SITD30 GP
 SET PERCENTS=OFF
 CLOSE (para fechar o arquivo)

2. Eficácia do benzonidazol em crianças com infecção pelo *T. cruzi*. Um ensaio de campo aleatório, duplo cego, foi conduzido para avaliar a eficácia e toxicidade do benzonidazol (7,5 mg/kg/dia; 2 x dia/ 2 meses) nas formas crônicas recentes da infecção pelo *T. cruzi*. A seleção de participantes foi por triagem sorológica de 1990 escolares de 7-12 anos residentes em área rural. 129 crianças soropositivas por 3 testes (ELISA, HAI, IFI) e pela técnica de ELISA quimioluminescente (CL-ELISA, antígeno A&T) foram alocadas aleatoriamente para receber benzonidazol (n=64) e placebo (n=65). Testes sorológicos foram realizados após 2 meses do tratamento e com 3, 6, 12 e 36 meses. Efeitos colaterais foram monitorados clínica e laboratorialmente durante a fase de intervenção. A estratégia de análise foi planejada levando em conta a “intenção de tratamento” e teve como objetivos: (1) avaliar a comparabilidade entre os grupos benzonidazol e placebo no início do ensaio; (2) comparar a média geométrica de títulos de anticorpos entre os grupos antes do tratamento e após os 3 anos de seguimento; (3) comparar a incidência de crianças soronegativas pela técnica CL-ELISA entre os grupos, ao final do ensaio; (4) avaliar diferenças entre a queda de títulos entre os grupos e (5) comparar a frequência de efeitos colaterais entre o grupo benzonidazol e placebo. Detalhes da metodologia estão na referência **Andrade et al, 1996**. O arquivo **ensaio.rec** contém informações sobre as 129 crianças participantes do ensaio e uma coleção de 24 variáveis selecionadas para responder as questões abaixo; a lista das variáveis e códigos encontram-se em anexo.

Questão 4. Considerando que o ensaio foi delineado para detectar uma incidência de 20% de soronegatividade no grupo benzonidazol, com erro alpha de 5% e poder estatístico de 80%, você considera suficiente o número de participantes incluídos no estudo? Quais as possíveis explicações para a diferença entre o número de participantes necessário e o número de crianças incluídas no ensaio?

Notas 4: READ ENSAIO.REC (para abrir o arquivo)
 tecla F9 para ir ao DOS
 digite EPITABLE para cálculo do tamanho da amostra (comando SAMPLE / SAMPLE SIZE/ TWO PROPORTION)
 pressione F10 para sair do EPITABLE
 digite EXIT para voltar ao ANALYSIS

Questão 5. Construa uma Tabela com a linha de base do grupo benzonidazol e placebo incluindo sexo ("SEX"), idade ("GRAGE"), município de residência ("MUN"), e eletrocardiograma ("ECG"). Você considera que o processo de aleatorização foi eficiente? Por que?

Notas 5: TABLES SEX DROGA
repetir o comando para GRAGE, MUN e ECG

Questão 6. Calcule a média geométrica de títulos (MGT) e intervalo de 95% de confiança (IC 95%) das técnicas HAI e IFI para o grupo benzonidazol e placebo na linha de base ("HAT" e "IFT") e ao final do ensaio ("HA6" e "IF6"). Calcule a média e IC 95% de ELISA ("ELT" e "EL6") e CL-ELISA ("ATTITULO_T" e "ATTITULO_F"). O que pode ser concluído em relação à comparabilidade das MGT e médias de ELISA e CL-ELISA na linha de base? Os valores obtidos para o grupo benzonidazol e placebo, ao final do ensaio, podem ser considerados estatisticamente diferentes? Justifique?

Notas 6: para calcular a MGT de cada técnica criar variável correspondente ao logaritmo na base 2 do título. Sugestão: para o cálculo da MGT crie as seguintes variáveis - "LOGHAT", "LOGIFT", "LOGHA6", "LOGIF6" correspondendo respectivamente à "HAT", "IFT", "HA6" e "IF6".
Exemplo utilizando HAT:
DEFINE LOGHAT ###.### (logaritmo da hemaglutinação antes do tratamento)
LET LOGHAT=LN (HAT)/LN (2)
MEANS LOGHAT DROGA /N
anote a média, desvio padrão e tamanho dos grupos benzonidazol e placebo
tecle F9 para ir ao DOS
digite EPITABLE para cálculo do IC 95% para benzonidazol e placebo (comando DESCRIBE / MEAN)
utilize os valores anotados
anote o IC calculado
pressione F10 para sair do EPITABLE
digite EXIT para voltar ao ANALYSIS
para transformar as médias e IC 95% obtidos em Média Geométrica de Títulos (MGT) utilize a base 2 com o auxílio de uma calculadora.
Exemplo: $2^{7,656}$ ($2^{\text{média}}$; média=7,656); $2^{7,19}$ ($2^{\text{limite inferior do IC95\%}}$; limite inferior do IC95%=7,19); $2^{8,12}$ ($2^{\text{limite superior do IC95\%}}$; limite superior do IC95%=8,12)
siga os mesmos comandos para a MGT da IFT
para HA6 e IF6 utilize o comando SELECT para excluir códigos <-1 e o comando SELECT (para desativar a seleção)
para a média e IC 95% das técnicas de ELISA (ELT e ATTITULO_T) utilize o comando MEANS. Exemplo para ELT:
MEANS ELT DROGA /N
para o cálculo do IC 95% siga os mesmos comandos utilizados para LOGHAT
para a média de EL6 e ATTITULO_F comando SELECT <EL6> ou <ATTITULO_F> <-1 (para excluir códigos sem informação) utilize o comando SELECT para desativar a seleção ao final.

Questão 7. Defina 2 novas variáveis para avaliar a queda de 3 títulos de anticorpos para IFI (“QUEDAIF”) e HAI (“QUEDAHA”) ao final do ensaio. Levando em conta a “intenção de tratamento”, construa uma tabela para cada variável comparando a proporção de queda de títulos entre os grupos. Analise a consistência do resultado obtido com a resposta da questão anterior; o que pode ser concluído em relação ao efeito do benzonidazol sobre os níveis de anticorpos em crianças com infecção pelo *T.cruzi*?

Notas 7:

```

DEFINE QUEDAIF # (queda de 3 títulos da imunofluorescência – IF6 <= IFT/8)
IF IF6 <=IFT/8 THEN QUEDAIF = 2 ELSE QUEDAIF= 1
TABLES QUEDAIF DROGA
DEFINE QUEDAHA # (queda de 3 títulos da hemaglutinação – HA6 <= HAT/8)
IF HA6 <=HAT/8 THEN QUEDAHA = 2 ELSE QUEDAHA=1
TABLES QUEDAHA DROGA
    
```

Questão 8. Considere a negatividade à CL-ELISA ao final do ensaio como indicador de “cura” da infecção. Utilizando a estratégia de “intenção de tratamento” construa uma tabela para avaliar a “incidência” de soronegatividade à CL-ELISA (“ATFIM_F”) nos grupos benzonidazol e placebo. As diferenças de resultado entre os grupos poderiam ser explicadas pelo acaso? Qual o risco relativo e IC 95% do benzonidazol para soropositividade à CL-ELISA? Qual a eficácia e respectivo IC 95% do benzonidazol?

Notas 8:

para a análise tipo “intenção de tratamento “ considere como falha terapêutica todas as perdas de seguimento ao final do ensaio

```

RECODE ATFIM_F -1=2 1=3 (para recodificar as perdas/ -1 como sorologia positiva).
SET PERCENTS = ON
TABLES DROGA ATFIM_F
SET PERCENTS = OFF
    
```

anote os resultados da tabela

tecle F9 para ir ao DOS

digite EPITABLE para cálculo da eficácia (comando STUDY /VACCINE EFFICACY/ COHORT STUDY)

utilize os dados da tabela anterior

pressione F10 para sair do EPITABLE

digite EXIT para voltar ao ANALYSIS

Questão 9. Avalie a toxicidade do benzonidazol. Compare a média e desvio padrão dos leucócitos e os valores do hematócrito entre os grupos, na linha de base (“HT1” e “LEUC1”) e durante o tratamento (“HT3” e “LEUC2”). Construa uma tabela comparando a incidência de lesões cutâneas (“LESAO”) no grupo benzonidazol e no grupo placebo durante o tratamento. O que pode ser concluído em relação à toxicidade do benzonidazol em crianças?

Notas 9:

```
MEANS HT1 DROGA /N
MEANS LEUC1 DROGA /N
SELECT HT3 <> -1
MEANS HT3 DROGA /N
SELECT (para desativar a seleção)
SELECT LEUC2 <> -1
MEANS LEUC2 DROGA /N
SELECT
TABLES LESAO DROGA
CLOSE (para fechar o arquivo)
```

Questão 10. De acordo com os resultados das questões 8 e 9 você recomendaria o uso do benzonidazol em escolares como estratégia para implementação do Programa de Controle da Doença de Chagas?

3. Ensaio randomizado para dose única de *L. major* morta associada com BCG contra leishmaniose cutânea em Bam, Iram. Sharifi et al, 1998. Um ensaio randomizado, controlado, duplo-cego foi conduzido em Bam, Iram, para avaliar a segurança e eficácia de dose única de vacina de promastigota de *L. major* autoclavada (ALM) combinada com BCG versus BCG como grupo de comparação contra leishmaniose cutânea (LC). Foram recrutadas 3633 crianças de 6 a 15 anos de idade de 49 escolas primárias. As crianças foram examinadas nos dias 1, 7, 30 e 80 depois da vacinação para avaliar a presença de efeitos colaterais locais e sistêmicos. O teste cutâneo de leishmanina (TCL) foi realizado no dia 80. Incidência de LC foi avaliada por vigilância passiva e por visitas de seguimento nas escolas a cada 2 meses num período de 2 anos. O arquivo **bam.rec** contém os registros de 3633 participantes, e a lista de variáveis e códigos encontra-se em anexo.

Questão 11. O estudo foi desenhado para detectar uma redução de 50% na incidência de LC, com nível de significância de 5%, poder estatístico de 80%, assumindo uma incidência anual de LC de 2% e perda de seguimento de 10% num período de 2 anos de acompanhamento. O número de participantes recrutados foi suficiente?

Notas 11: READ BAM.REC (para abrir o arquivo)
Tecla F9 para ir ao DOS
Digite EPITABLE para calcular o tamanho da amostra
Selecione SAMPLE, SAMPLE SIZE e TWO PROPORTIONS
Tecla F10 para fechar EPITABLE
Digite EXIT para retornar ao ANALYSIS

Questão 12. Compare os 2 grupos com relação a efeitos colaterais aos 7 e 30 dias após a vacinação. Calcule a frequência de eritema (RED7, RED30), ulceração (ULCER7, ULCER30), linfadenoma (LYMPH7, LYMPH30), prurido (ITCH7, ITCH30) dor (PAIN7, PAIN30) e induração (INDUR7, INDUR30). Recodifique as variáveis: 0 = sem efeito colateral e 1 = efeito colateral em qualquer grau. Existe diferença na frequência de reações adversas ao comparar os dois grupos?

Notas 12: SET PERCENTS = ON

RECODE RED7 2-3=1
RECODE RED30 2-3=1
TABLES RED7 GROUP
TABLES RED30 GROUP

RECODE ULCER7 2-3=1
RECODE ULCER30 2-3=1
TABLES ULCER7 GROUP
TABLES ULCER30 GROUP

RECODE LYMPH7 2-3=1
RECODE LYMPH30 2-3=1
TABLES LYMPH7 GROUP
TABLES LYMPH30 GROUP

RECODE ITCH7 2-3=1
RECODE ITCH30 2-3=1
TABLES ITCH7 GROUP
TABLES ITCH30 GROUP

RECODE PAIN7 2-3=1
RECODE PAIN30 2-3=1
TABLES PAIN7 GROUP
TABLES PAIN30 GROUP

RECODE INDUR7 2-3=1
RECODE INDUR30 2-3=1
TABLES INDUR7 GROUP
TABLES INDUR30 GROUP

Questão 13. Compare as taxas de conversão do teste cutâneo de leishmanina (TCL \geq 5 mm) entre os 2 grupos. Recodifique LST80 como uma variável dicotômica (LST80GR), conversão = 1, ausência de conversão = 2. Existe diferença significativa na resposta do teste cutâneo ao se comparar os 2 grupos?

Notas 13: Crie nova variável LST80GR
 DEFINE LST80GR #
 IF LST80 \geq 5 THEN LST80GR=1
 IF LST80 \geq 0 AND LST80 $<$ 5 THEN LST80GR=2
 TABLES LST80GR GROUP

Questão 14. Calcule a incidência acumulada de LC para 2 anos nos 2 grupos. Qual foi a eficácia da vacina ALM+BCG? Calcule a eficácia da vacina estratificada por sexo. Calcule a eficácia da vacina considerando apenas os casos diagnosticados depois de 9 meses (270 dias) de vacinação.

Notas 14: TABLES GROUP LEISH

Tecla F9 para ir ao DOS
 Digite EPITABLE para calcular a eficácia da vacina
 Selecione STUDY, VACCINE EFFICAY e COHORT STUDY
 Tecla F10 para fechar EPITABLE
 Digite EXIT para retornar ao ANALYSIS

Para análise estratificada por sexo
 TABLES GROUP LEISH SEX

Tecla F9 para ir ao DOS
 Digite EPITABLE para calcular a eficácia da vacina
 Selecione STUDY, VACCINE EFFICAY e COHORT STUDY
 Tecla F10 para fechar EPITABLE
 Digite EXIT para retornar ao ANALYSIS

Crie nova variável DAYS
 LET DAYS = DATEINF-VACDATE

Defina nova variável LEISH9M considerando casos apenas os casos incidentes de leishmaniose que ocorreram depois de 9 meses de vacinação

DEFINE LEISH9M #
 IF LEISH=1 AND DAYS \geq 270 THEN LEISH9M=1
 IF LEISH=1 AND DAYS $<$ 270 THEN LEISH9M=2
 IF LEISH=2 THEN LEISH9M=2
 TABLES GROUP LEISH9M

Tecla F9 para ir ao DOS
Digite EPITABLE para calcular a eficácia da vacina
Selecione STUDY, VACCINE EFFICAY e COHORT STUDY
Tecla F10 para fechar EPITABLE
Digite EXIT para retornar ao ANALYSIS

SET PERCENTS=OFF
CLOSE [para fechar o arquivo]

REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, A.L.S.S.; ZICKER, F.; OLIVEIRA, R.M.; SILVA, S.A.; LUQUETTI, A.O.; ANDRADE, J.G.; ANDRADE, S.G.; ALMEIDA I.C.; TRAVASSOS, L.R.; MARTELLI, C.M.T. A randomised trial of efficacy of benznidazole in treatment of early *Trypanosoma cruzi* infection. **Lancet**, **348**:1407-1413, 1996.

ANDRADE JG; ANDRADE ALSS; ARAUJO ESO; OLIVEIRA RM; SILVA SA; MARTELLI CMT & ZICKER F. A randomized clinical trial with high dose of chloroquine for treatment of *Plasmodium falciparum* malaria in Brazil. **Revista do Instituto de Medicina Tropical de São Paulo**, **34(5)**:467-473, 1992.

SHARIFI I., FEKRI A.R., AFLATONIAN M., KHAMESIPUOUR A, NADIM A. MOUSAVI M.A., MOMENI A.Z., DOWLATI Y., GODAL T., ZICKER F. & SMITH P.G. Randomised vaccine trial of single dose killed *Leishmania major* plus BCG against anthroponotic cutaneous leishmaniasis in Bam, Iran. **The Lancet**, **351**:1540-1543, 1998.

Arquivo: ensaioml.rec

Variável	Descrição	Código	Descrição do Código
NO	Número de identificação		
GP	Grupo de tratamento	1 2	50 mg/Kg 25 mg /Kg
AGE	Idade em anos completos	15 a 62	
SEX	Sexo	1 2	Masculino Feminino
OCUP	Ocupação	1 2 3 4 5	Garimpo Pesca Lavoura Comércio Outros
XMAL	Número de episódios de malária nos últimos 5 anos	1 a 20	
MED	Medicamento utilizado no último episódio de malária	1 2 3 4	cloroquina quinino mais de 1 medicamento outro
SIND	Dias decorridos desde o início dos sintomas	1 a 40	
SITD7	Parasitemia com 7 dias de seguimento	1 2 3	Sem parasitemia Com parasitemia Perda de seguimento
SITD30	Parasitemia com 30 dias de seguimento	1 2 3	Sem parasitemia Com parasitemia Perda de seguimento
PARD0	Parasitemia antes do início do tratamento	10 a 37800	trofozoítas / campo

Arquivo: **ensaio.rec**

Variável	Descrição	Código	Descrição do Código
NOEC	Número de identificação		
SEX	Sexo	1 2	Masculino Feminino
GRAGE	Faixa Etária	1 2	7-9 anos 10-12 anos
MUN	Município de residência	1 2 3	Guarani Posse Simolândia
ECG	Eletrocardiograma antes da intervenção	0 1	Normal Com alteração codificável
HAT	Reverso do título da hemaglutinação antes da intervenção	16 a 8192	
IFT	Reverso do título da Imunofluorescência antes da intervenção	320 a 5120	
ELT	Índice de ELISA antes da intervenção	1,2 a 4,3	
HA6	Reverso da hemaglutinação após 3 anos	-1 4 a 4096	Sem informação
IF6	Reverso do título da imunofluorescência após 3 anos	-1 20 a 2560	Sem informação
EL6	Índice de ELISA após 3 anos	-1 0,2 a 3,8	Sem informação
ATFIM_F	CL-ELISA após 3 anos	-1 1 2	Sem informação Negativo Positivo
ATTITULO_T	Título CL-ELISA antes da intervenção	1,001 a 22,363	
ATTITULO_F	Título CL-ELISA após 3 anos	-1 0 a 22,380	Sem informação
DROGA	Grupo de tratamento	0 1	Benzonidazol Placebo
HT1	Hematócrito antes da intervenção	30 a 48	
LEUC1	Leucócitos antes da intervenção	3200 a 39400	
HT3	Hematócrito durante o tratamento	-1 28-44	Sem informação
LEUC2	Lecucócitos durante o tratamento	2900 a 16800	
LESAO	Lesão cutânea durante o tratamento	1 2	Não Sim

File: bam.rec

Variável	Descrição	Código	Descrição do código
NUMBER	Número de identificação		
SEX	Sexo	F M	Feminino Masculino
RESPPD	Resposta ao PPD	0-10.0	
VACDATE	Data de vacinação		
RED7	Eritema no dia 7	0	Ausente
		1	Fraco
		2	Moderado
		3	Severo
ULCER7	Ulceração no dia 7	0	Ausente
		1	Fraco
		2	Moderado
		3	Severo
LYMPH7	Linfoadenopatia no dia 7	0	Ausente
		1	Fraco
ITCH7	Prurido no dia 7	0	Ausente
		1	Fraco
		2	Moderado
		3	Severo
PAIN7	Dor local no dia 7	0	Ausente
		1	Fraco
		2	Moderado
		3	Severo
INDUR7	Induração no dia 7	0	Ausente
		1	Fraco
		2	Moderado
		3	Severo
RED30	Eritema no dia 30	0	Ausente
		1	Fraco
		2	Moderado
		3	Severo
ULCER30	Ulceração no dia 30	0	Ausente
		1	Fraco
		2	Moderado
		3	Severo
LYMPH30	Linfoadenopatia no dia 30	0	Ausente
		1	Fraco
		2	Moderado
		3	Severo
ITCH30	Prurido no dia 30	0	Ausente
		1	Fraco
		2	Moderado
		3	Severo
PAIN30	Dor local no dia 30	0	Ausente
		1	Fraco
		2	Moderado
INDUR30	Induração no dia 30	0	Ausente
		1	Fraco
		2	Moderado
		3	Severo
LST80	Teste cutâneo de leishmania no dia 80	0.0-15.0	

Variável	Descrição	Código	Descrição do código
DATEINF	Data de aparecimento da primeira lesão		
LEISH	Definição de caso	1 2	Sim Não
GROUP	Grupo de alocação	1 2	ALM+BCG BCG

MORTALIDADE E NASCIDOS VIVOS

ANÁLISE EPIDEMIOLÓGICA

INTRODUÇÃO	119
O SISTEMA DE INFORMAÇÃO SOBRE MORTALIDADE (SIM)	119
SISTEMA DE INFORMAÇÃO SOBRE NASCIDOS VIVOS (SINASC)	120
VINCULAÇÃO ENTRE O SIM E SINASC	121
REFERÊNCIAS PARA LEITURA	123
EXERCÍCIOS	124
DICIONÁRIO DE BANCO DE DADOS	129

INTRODUÇÃO

No Brasil, nas últimas décadas, foram desenvolvidos e implementados diversos sistemas de informações para eventos específicos que constituíram grandes bancos de dados de abrangência nacional. São exemplos de alguns desses grandes bancos de dados o Sistema de Informações sobre Mortalidade (SIM), O Sistema de Informações Hospitalares (SIH/AIH) e o Sistema de Informações sobre Nascidos Vivos (SINASC). A estruturação de cada sistema deu-se em momentos diferentes de forma vertical e centralizada, com lógicas e objetivos diferenciados. A lógica de estruturação do SIH/AIH, por exemplo, é o controle da autorização de internação hospitalar (AIH) e que apresenta como subprodutos algumas informações epidemiológicas importantes como o diagnóstico principal da internação hospitalar.

Um desafio na utilização da epidemiologia nos serviços de saúde é a exploração de todo o potencial dos bancos de dados, não se restringindo apenas à algumas variáveis clássicas de identificação como sexo e idade. Para realizar análises mais complexas algumas dificuldades devem ser superadas:

- a) a qualidade em termos de cobertura e fidedignidade das informações provenientes dos bancos de dados;
- b) a falta de padronização e compatibilização das informações nos diversos bancos de dados;
- c) a ausência de normas operacionais que facilitem a utilização dos bancos de dados;
- d) a dificuldade de acesso às informações no âmbito dos estados e municípios, seja devido à grande defasagem de tempo entre a produção da informação e a disponibilidade da mesma, ou devido à disponibilização das informações em níveis de desagregação insuficientes ou não compatíveis com a utilização municipal ou local, e;
- e) a falta de integração das informações produzidas em cada subsistema de informações.

Neste módulo serão abordados o SIM e o SINASC, e as possibilidades de integração entre esses sistemas de informação com vistas à análise epidemiológica sobre fatores de risco para os componentes neonatal e pós-neonatal da mortalidade infantil.

O SISTEMA DE INFORMAÇÃO SOBRE MORTALIDADE (SIM)

O SIM foi implantado no Brasil em 1975 como um subsistema de informações do Ministério da Saúde com o objetivo de:

- 1) aprimorar as informações de mortalidade a partir da definição de uma Declaração de Óbito (DO) padronizada para todo o país;
- 2) estabelecimento do fluxo da DO;
- 3) centralização do processamento;
- 4) divulgação dos dados através de uma publicação anual *Estatísticas de mortalidade* e atualmente através da Internet e CD-ROOM.

As informações de mortalidade são freqüentemente criticadas pelo fato de utilizarem dados provenientes das estatísticas vitais, reconhecidamente de baixa qualidade, com problemas na

definição dos eventos vitais como nascido vivo e nascido morto, sub-registro de nascimentos e óbitos, invasão e evasão de óbitos, declaração errada da idade.

Do ponto de vista quantitativo, ainda existem falhas na cobertura dos dados e na coleta das DOs, principalmente nas regiões Norte e Nordeste e em alguns estados da região Centro-Oeste. Do ponto de vista qualitativo, os principais problemas referem-se à qualidade do preenchimento da DO, notadamente na parte do atestado das causas que levaram ao óbito.

A validação da causa básica de óbitos, utilizando diversas estratégias como investigação de prontuários hospitalares, entrevistas com pediatras e com familiares das crianças, constata índices de concordância diferenciados para os óbitos neonatais e pós-neonatais. Os primeiros apresentam maior grau de concordância com pouca alteração nos grupos de causa e sim nos códigos de três algarismos da classificação internacional de doenças (CID). Como estratégia para contornar os problemas decorrentes da baixa qualidade do preenchimento da DO pode-se relacionar melhorias com relação ao preenchimento em si, tais como, incentivar e ensinar os profissionais de saúde para um correto preenchimento do documento, implementação de sistemas de investigação de causas mal definidas ou incompletas. Do ponto de vista da análise esses problemas podem ser minimizados através da análise dos dados a partir dos grandes grupos da CID e através da análise de causas múltiplas de óbito, incluindo, além da causa básica, as causas associadas e contribuintes.

SISTEMA DE INFORMAÇÃO SOBRE NASCIDOS VIVOS (SINASC)

O Sistema de Informação sobre Nascidos Vivos (SINASC) foi implantado a partir de 1990, inicialmente nas capitais do país. No ano de 1992, apresentava uma cobertura parcial em todos os Estados, exceto em Santa Catarina.

O objetivo fundamental do SINASC é implementar um sistema de informações sobre os nascimentos, com dados individualizados, que contenha informações coletadas em hospitais, no momento do nascimento, relacionadas à gestação, ao parto, às condições de nascimento da criança e às características da mãe.

As informações oriundas desse sistema apresentam base populacional e diferem das estatísticas do registro civil divulgadas pelo IBGE em dois aspectos fundamentais: primeiro, porque dispõem de uma gama maior informações sobre as crianças nascidas vivas, como pode ser visto pelo exame da declaração de nascido vivo (DN), e segundo, porque o preenchimento da DN em três vias e o seu fluxo originado nos hospitais de nascimento e nos cartórios possibilitam o conhecimento de todos os nascidos vivos de uma localidade e não apenas dos registrados em cartórios, possibilitando uma correção do sub-registro de nascimentos.

O instrumento de coleta de dados utilizado pelo SINASC é a declaração de nascido vivo (DN) que deve ser preenchida nos serviços de saúde em que a criança nasce ou pelo cartório de registro civil no caso de nascimento no domicílio. A DN é preenchida em três vias contendo informações distribuídas em oito blocos:

- a) identificação da DN;
- b) informações referentes ao cartório onde o nascimento foi registrado;
- c) informações sobre o local de ocorrência do nascimento;
- d) informações sobre a mãe (nome, idade, estado civil, escolaridade, ocupação, número de filhos tidos em gestações anteriores e endereço de residência);

- e) informações sobre a gestação e parto (duração da gestação, tipo de gravidez e tipo de parto e número de consultas pré-natal);
- f) dados sobre o recém-nascido (data de nascimento, sexo, peso ao nascer, índice de Apgar, raça/cor e se for detectada alguma malformação congênita e/ou anomalia cromossômica);
- g) informações sobre o responsável pelo preenchimento da DN. Impressão digital da mãe e plantar do RN.

A avaliação do SINASC para o país mostra um subregistro de 30,5% sendo de 66,7% na Região Norte e de 10,3% na Região Sudeste.

A avaliação do SINASC em cinco municípios do Estado de São Paulo - Itararé, Marília, Pariquera-Açu, Santo André, e São José dos Campos - mostrou os seguintes resultados:

- a) o SINASC apresenta uma cobertura próxima de 100% para os nascimentos hospitalares
- b) a duplicidade de DN encontrada foi desprezível e,
- c) a validação das informações preenchidas na DN encontrou índices de concordância elevados – acima de 90% – para a maioria das variáveis, exceto o índice de Apgar no primeiro e quinto minutos, duração da gestação, idade da mãe, grau de instrução da mãe, número de filhos tidos e nome do pai.

Apesar do SINASC estar implantado em aproximadamente 85% dos municípios do país existem problemas no fluxo das informações dos municípios para os estados e destes para o Ministério da Saúde. A cobertura do SINASC para as capitais brasileiras é bastante diferenciada: para o ano de 1992 as capitais Porto Velho, Rio Branco e Teresina apresentaram um número de nascidos vivos obtidos via SINASC superior ao esperado segundo a estimativa do coeficiente geral de natalidade de cada cidade; as capitais Manaus, Boa Vista, Belém, Aracaju e Campo Grande apresentaram cobertura do SINASC igual ou superior a 90%; as capitais Palmas, Fortaleza, Belo Horizonte, Vitória e Porto Alegre apresentaram cobertura do SINASC entre 50 e 90% e as capitais João Pessoa e Goiânia apresentaram coberturas do SINASC inferior a 50%. A qualidade do preenchimento da DN avaliada tomando como indicador a proporção de itens não preenchidos ou ignorados mostra que a qualidade do preenchimento da DN está melhorando gradativamente. As variáveis com menor proporção de não preenchimento/ignorado são: sexo, peso ao nascer, duração da gestação, tipo de gravidez, tipo de parto e idade da mãe, enquanto as variáveis com maiores problemas de preenchimento são: índice de Apgar (primeiro e quinto minutos), grau de instrução da mãe e filhos tidos (nascidos vivos e nascidos mortos).

VINCULAÇÃO ENTRE O SIM E SINASC

A análise de duas bases de dados de dois subsistemas de informações independentes, concebidos em momentos diferentes e por lógicas diferentes apresenta algumas dificuldades tais como a ausência de uma padronização e compatibilização entre as diversas variáveis. Uma vez promovida uma compatibilização mínima entre as bases de dados o seu uso simultâneo pode-se dar de duas formas: através de uma **INTEGRAÇÃO** ou de uma **VINCULAÇÃO**. Na primeira, utiliza-se de um sistema para se avaliar ou complementar informações de outro como por exemplo a notificação de casos de AIDS para o sistema de informação de agravos notificáveis (SINAN) ou a avaliação da cobertura do SINASC utilizando-se o número de nascidos vivos obtidos no SIH/AIH. Na vinculação existe um processo de *linkage* entre os dois subsistemas através de campos comuns às duas bases de dados como por exemplo o estudo das variáveis presentes na DN como potenciais fatores de risco para a mortalidade infantil que é o objeto de exercício desse módulo e que detalharemos a seguir.

A análise da mortalidade infantil a partir de coortes de nascimento não é utilizada de forma rotineira em nosso país. A implementação do SINASC tornou possível esta metodologia através da vinculação entre os dois bancos de dados – SIM e SINASC. O conjunto de nascidos vivos residentes em uma localidade em um determinado ano identificados através de sua respectiva DN constituem uma coorte de nascidos vivos. O monitoramento do SIM possibilita identificar a ocorrência de óbitos infantis de nascidos vivos desta mesma coorte. Uma vez identificado o óbito, a sua DO é emparelhada com a respectiva DN promovendo-se um processo de vinculação entre os dois sistemas de informação.

A partir desses procedimentos, define-se uma **coorte retrospectiva** na qual o **efeito** – os **óbitos infantis** – correspondeu ao conjunto de DN/DO emparelhadas e o conjunto das DN não emparelhadas constituiu o grupo dos sobreviventes. Esta metodologia tem-se mostrado eficaz na obtenção da vinculação dos dois subsistemas.

As principais vantagens da vinculação entre os dois sistemas de informações no que se refere à análise da mortalidade infantil em um determinado município ou localidade são:

- a) estudo da mortalidade infantil a partir de coortes de nascimento;
- b) cálculo direto da probabilidade de morte antes de completar um ano de vida;
- c) estudo de potenciais fatores de risco ou de proteção presentes na DN e coletados no momento do nascimento da criança;
- d) quantificação da magnitude da associação entre os fatores estudados e a probabilidade de morte infantil;
- e) realização de estudos epidemiológicos analíticos utilizando-se dados secundários disponíveis por sistemas de informações existentes nos serviços de saúde, sem a necessidade de levantamentos especiais ou inquéritos de base populacional.

REFERÊNCIAS PARA LEITURA

- ALMEIDA, M.F. *Mortalidade neonatal em Santo André*, São Paulo, 1994. Tese de Doutorado em Saúde Pública (Epidemiologia), Faculdade de Saúde Pública da USP.
- FONSECA, L.A.M. & LAURENTI, R. A qualidade da certificação médica da causa de morte em São Paulo, Brasil. **Revista de Saúde Pública**,**8**: 21-29,1974.
- LAURENTI, R., MELLO, JORGE, M.H.P., LEBRÃO, M.L., GOTLIEB, S.L.D. *Estatísticas de Saúde*. 2a Edição, São Paulo, EPU,1987.
- MELO JORGE, M.H.P, GOTLIEB, S. L. D., SOBOLL, M. L. M. S., BALDIJÃO, M. F. A. & LATORRE, M. R. D. O. *O sistema de informação sobre nascidos vivos - SINASC*. Centro Brasileiro para Classificação de Doenças (MS/USP/OPAS-OMS), Série Divulgação (N.º 7) São Paulo,1992.
- MELO JORGE, M.H.P, GOTLIEB, S. L. D., SOBOLL, M. L. M. S., BALDIJÃO, M. F. A. & LATORRE, M. R. D. O. Avaliação do sistema de informação sobre nascidos vivos e o uso de seus dados em epidemiologia e estatísticas de saúde. **Revista de Saúde Pública**,**27(6 suplemento)**,1993.
- MELO JORGE, M.H.P., GOTLIEB, S.D. & OLIVEIRA, H. O sistema de informação sobre nascidos vivos: primeira avaliação dos dados brasileiros. **Informe Epidemiológico do SUS**, Ano V, (2):15-48, 1996.
- MORAIS NETO, O. L. & BARROS, M. B. A. Fatores de risco para mortalidade neonatal e pós-neonatal na Região Centro-Oeste do Brasil: *Linkage* entre bancos de dados de nascidos vivos e óbitos infantis. **Cadernos de Saúde Pública**, **16(2)**: 477-485, 2000.
- PAULA, A.M.C., EVANGELISTA FILHO, D., PEREIRA, I.P.A., ALBANO, A.H.B.L. & FERNANDES, R.M. Avaliação dos dados de mortalidade, Brasil - 1979 a 1989. **Informe Epidemiológico do SUS**, Ano III, (1): 21-44,1994.

EXERCÍCIOS

Realizou-se um estudo epidemiológico com os objetivos de: (a) determinar a probabilidade de morte infantil nos períodos neonatal e pós-neonatal no município de Goiânia; (b) estudar o papel das variáveis presentes na declaração de nascido vivo (DN) como fatores de risco para os componentes neonatal e pós-neonatal da mortalidade infantil. O estudo foi realizado a partir do procedimento de *linkage* entre os arquivos de óbitos infantis do Sistema de Informações sobre Mortalidade (SIM) e o arquivo de nascidos vivos do Sistema de Informações sobre Nascidos Vivos (SINASC) o que possibilitou avaliar potenciais fatores de risco para mortalidade neonatal e pós-neonatal em uma **coorte retrospectiva** de nascidos vivos no ano de 1992. Para determinar a probabilidade de morte infantil e avaliar os potenciais fatores de risco para a mortalidade infantil deve-se seguir alguns passos na organização dos arquivos de dados:

- (a) Vincular os arquivos de Nascidos vivos e óbitos infantis;
- (b) Criar um campo identificador do efeito distinguindo-se os nascidos vivos óbitos dos sobreviventes;
- (c) Separar um arquivo para o componente neonatal e outro para o componente pós-neonatal da mortalidade infantil;
- (d) Analisar a associação entre as variáveis de exposição e a mortalidade nos períodos neonatal e pós-neonatal;
- (e) Avaliar o papel de algumas variáveis enquanto potenciais fatores de confusão.

Questão 1. Utilizando os arquivos NASCVIV.REC e OBITOS.REC, realize o processo de *linkage* dos arquivos de nascidos vivos e óbitos infantis, salve o arquivo vinculado em um novo arquivo chamado COORTE1.REC e inclua na estrutura do novo arquivo um campo identificador COORTE para distinguir os nascidos vivos que evoluíram para óbito dos sobreviventes.

Notas 1:

```
Vincular os arquivos OBITOS.REC e NASCVIV.REC
Selecione o módulo ANALYSIS
READ NASCVIV.REC
RELATE CODIGO OBITOS.REC (O comando relate relaciona o
campo CODIGO – número da DN comum aos dois arquivos criando
uma vinculação entre os arquivos de óbitos e nascidos vivos)
O próximo passo é definir uma variável indicadora do efeito no arquivo
vinculado
ROUTE COORTE1.REC (Direciona a saída para o arquivo
COORTE1.REC)
WRITE RECFILE (Salva o arquivo vinculado no novo arquivo
COORTE1.REC)
CLOSE (Fecha o arquivo NASCVIV.REC)
READ COORTE1.REC (Abre o arquivo vinculado COORTE1.REC)
DEFINE COORTE # (criar a variável COORTE que será a
identificadora do efeito)
IF DO<>. THEN COORTE=1 ELSE COORTE=2 (O registro que
apresentar o campo DO preenchido é um recém-nascido que evoluiu
para óbito no primeiro ano de vida. COORTE=1 são os óbitos e
coorte=2 são os recém-nascidos sobreviventes)
```

Questão 2. Utilizando o arquivo MICOORTE.REC, crie dois arquivos: o

NEONAT.REC para análise dos fatores de risco no período neonatal e o arquivo POSNEO.REC para análise do componente pós-neonatal.

Notas 2:

Ainda no módulo ANALYSIS mantenha o arquivo COORTE.REC aberto
SELECT AGECA1="1" OR AGECA1=. (Selecionar os nascidos vivos que morreram no período neonatal – idade inferior a 28 dias de vida – AGECA1=1 – e os nascidos vivos sobreviventes – AGECA1=.)
ROUTE NEONAT.REC (direciona a saída para o arquivo NEONAT.REC)
WRITE RECFILE (Salva o arquivo apenas com os óbitos neonatais e nascidos vivos sobreviventes, ignorando os óbitos no período pós-neonatal)
SELECT (desativa a seleção)

Para criar o arquivo de óbitos pós-neonatais
SELECT AGECA1="2" OR AGECA1=. (Selecionar os nascidos vivos que morreram no período pós-neonatal – 28 dias de vida até um ano – AGECA1=2 – e os nascidos vivos sobreviventes – AGECA1=.)
ROUTE POSNEO.REC (direciona a saída para o arquivo POSNEO.REC)
WRITE RECFILE (Salva o arquivo apenas com os óbitos pós-neonatais e nascidos vivos sobreviventes)
SELECT (desativa a seleção)
CLOSE (Fecha o arquivo COORTE1.REC)

OBS.: As variáveis tipo caracter devem ter os seus valores digitados entre aspas durante o processo de transformação de variáveis

Questão 3.

Para o período NEONATAL e PÓS-NEONATAL, construa tabelas 2x2 e compare os grupos de óbitos e sobreviventes segundo as variáveis de exposição PESOCAT, GESTA, TGRAV, TPARTO, AGEMAE, CATHOSP, INSTMAE e RADM. Considere como expostos os nascidos vivos (NV) das seguintes categorias em cada variável de exposição: PESOCAT (< 2500g), GESTA (< 37 sem.), TPARTO (operatório), AGEMAE (< 20 anos), CATHOSP (público-estatal), INSTMAE (< I grau completo), TGRAV (gravidez múltipla) e RADM (Região Noroeste como exposto e a Região Central como não exposto).

Notas 3:

Análise para o período Neonatal
READ NEONAT.REC (abre o arquivo NEONAT.REC)
SET STATISTICS=OFF (desativa as estatísticas)
SELECT PESOCAT<>9 (elimina os registros sem informação do peso)
TABLES PESOCAT COORTE (anote os resultados)
SELECT (desativa a seleção anterior)
SELECT GESTA<>9
TABLES GESTA COORTE (anote os resultados)
SELECT
SELECT TPARTO<>9
TABLES TPARTO COORTE (anote os resultados)
SELECT
SELECT AGEMAE<>9
TABLES AGEMAE COORTE (anote os resultados)
SELECT
SELECT CATHOSP<>9

```

TABLES CATHOSP COORTE (anote os resultados)
SELECT
SELECT INSTMAE<>9
TABLES INSTMAE COORTE (anote os resultados)
SELECT
SELECT TGRAV<>9
TABLES TGRAV COORTE (anote os resultados)
SELECT
SELECT RADM<>0
TABLES RADM COORTE (anote os resultados)
SELECT

Para o período Pós-neonatal
READ POSNEO.REC (abre o arquivo POSNEO.REC)
SET STATISTICS=OFF (desativa as estatísticas)
SELECT PESOCAT<>9 (elimina os registros sem informação do peso)
TABLES PESOCAT COORTE (anote os resultados)
SELECT (desativa a seleção anterior)
SELECT GESTA<>9
TABLES GESTA COORTE (anote os resultados)
SELECT
SELECT TPARTO<>9
TABLES TPARTO COORTE (anote os resultados)
SELECT
SELECT AGEMAE<>9
TABLES AGEMAE COORTE (anote os resultados)
SELECT
SELECT CATHOSP<>9
TABLES CATHOSP COORTE (anote os resultados)
SELECT
SELECT INSTMAE<>9
TABLES INSTMAE COORTE (anote os resultados)
SELECT
SELECT TGRAV<>9
TABLES TGRAV COORTE (anote os resultados)
SELECT
SELECT RADM<>0
TABLES RADM COORTE (anote os resultados)
SELECT

```

Questão 4. Considerando os resultados das tabelas anteriores para os componentes neonatal e pós-neonatal, responda: (a) qual a probabilidade de morte antes de completar um ano de vida nos expostos e não expostos às variáveis estudadas; (b) quais os riscos relativos e os seus respectivos intervalos de 95% de confiança e (c).avalie a existência de associação entre a exposição e a mortalidade neonatal e pós-neonatal utilizando o teste do qui-quadrado para testar a hipótese nula de ausência de associação entre a exposição e a mortalidade nos períodos neo e pós-neonatal. Relacione quais os fatores associados e entre eles quais são os fatores de risco e de proteção. Compare os resultados entre o período neonatal e pós-neonatal

Notas 4: F9 (tecle F9 para ir ao DOS)
EPITABLE

STUDY / COHORT / CUMULATIVE INCIDENCE (selecione essas opções)
Digite os dados de cada uma das tabelas
Pressione CALCULATE (executa os cálculos da probabilidade, RR, IC 95% e qui-quadrado)
Pressione EDIT (permite visualizar e editar todos os resultados)
Pressione F10 (sai do epitável)
Digite EXIT (volta ao módulo ANALYSIS)

Questão 5. **Avalie o potencial efeito de confusão das variáveis PESOCAT e CATHOSP nas seguintes situações:**

(a) Variável PESOCAT na associação entre o TPARTO e COORTE considerando como expostos os nascidos vivos por parto operatório.

(b) Variável CATHOSP na associação entre o TPARTO e COORTE considerando como expostos os nascidos vivos por parto operatório.

Compare os riscos relativos ajustados e não ajustados identificando se a variável atuou como fator de confusão ou modificadora de efeito em cada uma das associações testadas.

Notas 5:

Para o período Neonatal
No módulo ANALYSIS
READ NEONAT.REC
SELECT TPARTO<>9 AND PESOCAT<>9
TABLES TPARTO COORTE PESOCAT
Avalie para cada estrato o valor do RR, os IC95%, a significância do teste de homogeneidade entre os estratos, o valor das medidas ajustadas – RR de Mantel-Haenszel e Qui-quadrado de Mantel-Haenszel. Compare as medidas ajustadas e não ajustadas e avalie se a variável PESOCAT atua como variável de interação, confusão ou se não interfere na associação estudada.
SELECT (desativa a seleção)
SELECT TPARTO<>9 AND CATHOSP<>9
TABLES TPARTO COORTE CATHOSP
Avalie para cada estrato o valor do RR, os IC95%, a significância do teste de homogeneidade entre os estratos, o valor das medidas ajustadas – RR de Mantel-Haenszel e Qui-quadrado de Mantel-Haenszel. Compare as medidas ajustadas e não ajustadas e avalie se a variável CATHOSP atua como variável de interação, confusão ou se não interfere na associação estudada.
SELECT

Para o período Pós-Neonatal
No módulo ANALYSIS
READ POSNEO.REC
SELECT TPARTO<>9 AND PESOCAT<>9
TABLES TPARTO COORTE PESOCAT
Avalie para cada estrato o valor do RR, os IC95%, a significância do teste de homogeneidade entre os estratos, o valor das medidas ajustadas – RR de Mantel-Haenszel e Qui-quadrado de Mantel-Haenszel. Compare as medidas ajustadas e não ajustadas e avalie se a variável PESOCAT atua como variável de interação, confusão ou se não interfere na associação estudada.
SELECT (desativa a seleção)

```
SELECT TPARTO<>9 AND CATHOSP<>9  
TABLES TPARTO COORTE CATHOSP
```

Avalie para cada estrato o valor do RR, os IC95%, a significância do teste de homogeneidade entre os estratos, o valor das medidas ajustadas – RR de Mantel-Haenszel e Qui-quadrado de Mantel-Haenszel. Compare as medidas ajustadas e não ajustadas e avalie se a variável CATHOSP atua como variável de interação, confusão ou se não interfere na associação estudada.

```
SELECT
```

Arquivo: **neonat.rec** e **posneo.rec**

Nome	Descrição	Código	Descrição do código
CODIGO	Número de identificação da DN		
COORTE	Variável identificadora do efeito	1 2	Óbito Sobrevivente
DO	Número identificador da DO	000039 a 020038	
SEXO	Sexo da criança	1 2 9	Masculino Feminino Sem informação
TGRAV	Tipo de gravidez	1 2 9	Múltipla Única Sem informação
AGECAT	Grupo de idade do nascido vivo	0 1 2	Sobrevivente Óbito no período neonatal Óbito no período pós-neonatal
AGEMAE	Grupo de idade da mãe	1 2 9	< 20 anos ≥ 20 anos Sem informação
CATHOSP	Categoria do hospital de nascimento	1 2 9	Público-estatal Privado SUS e não SUS Sem informação
GESTA	Categoria de duração da gestação	1 2 9	< 37 semanas ≥ 37 semanas Sem informação
INSTMAE	Grupos de instrução da mãe	1 2 9	< I grau completo ≥ I grau completo Sem informação
PESOCAT	Grupos de peso ao nascer	1 2 9	< 2500 g ≥ 2500 g Sem informação
TPARTO	Tipo de parto	1 2 9	Operatório Espontâneo Sem informação
RADM	Região de Residência da Mãe	1 2 3 4 5 6 7 8 9 0	Central Sul Sudoeste Oeste Mendanha Noroeste Norte Meia Ponte Leste Ignorada