

# 13

## Qualidade dos Instrumentos Epidemiológicos

Michael Eduardo Reichenheim e Claudia Leite Moraes

### ► Introdução

Essencialmente, a qualidade da informação em estudos epidemiológicos pode ser sintetizada como resultante da apropriabilidade teórica e operativa de instrumentos de aferição quando utilizados à luz de um processo de mensuração adequado a cenários e contextos precípuos. Dentre vários aspectos que conferem qualidade a um estudo epidemiológico, dois são de particular interesse quando se discute a qualidade da informação (Bowling, 1997; Reichenheim & Moraes, 2002; Wilson, 2005; Rothman *et al.*, 2008; Streiner & Norman, 2008). Um aspecto se reporta ao processo de mensuração em si e a toda a cadeia de afazeres necessária para maximizar o potencial do(s) instrumento(s) em uso. O outro aspecto concerne à operacionalização de conceitos, *i. e.*, o traslado de significados de um elemento teórico aos elementos manifestos que possam ser efetivamente inquiridos e/ou observados. Em termos práticos, trata-se do procedimento de redução do conceito a representantes empíricos que se materializa no processo de desenvolvimento de um instrumento ou, em muitas situações, em um processo de adaptação transcultural (ATC) caso se almeje utilizar instrumentos preexistentes fora do contexto linguístico-sociocultural em que foi concebido.

Para melhor discernir estes dois aspectos, vale apontar que é perfeitamente cabível conceber uma situação em que ocorra um problema de classificação devido ao uso de uma escala (variável) construída a partir de indicadores (manifestos) inadequados, mesmo não havendo nenhum problema na execução da aferição. Por outro lado, mesmo diante de uma escala satisfatoriamente arquitetada e desenvolvida, nada impede que haja um problema de mensuração, levando a um problema na ordenação de indivíduos que potencialmente deveriam ser acertadamente ranqueados. Ambas as situações levam a uma má classificação dos sujeitos estudados, afetando a validade do estudo. Chama-se a atenção para a necessidade de explicitação destes dois aspectos importantes – a qualidade do instrumental e de sua aplicação – não só para garantir a validade interna de um estudo epidemiológico, mas também para permitir a comparação do próprio com achados obtidos em outras pesquisas.

Há sempre um dilema quando se está diante de um tema vasto e complexo, mas ainda pouco difundido entre um público leitor. Uma opção é selecionar alguns pontos identificados como sendo os mais importantes e, de preferência, relativamente fáceis de serem transmitidos e aprofundá-los à exaustão, cobrindo-se definições, métodos e operacionalização. Outra opção é buscar abrangência, na premissa de que, para iniciar o leitor, é profícuo oferecer um rico panorama sobre o estado d'arte. Foi esta a opção que fizemos, até mesmo porque a ampla lista de referências que daí decorre possibilita o encontro de mais informação sobre os vários pontos cobertos, inclusive sobre alternativas metodológicas, controvérsias vigentes e questões ainda em debate e investigação.

Esse texto, pois, não tem a intenção de apresentar conceitos e definições no âmbito de “aferições”, nem tampouco abarcar os vieses epidemiológicos que podem ocorrer nessa esfera. A proposta aqui é oferecer uma sistemática operacional para a montagem ou adaptação de instrumentos e, portanto, pressupõe algum conhecimento prévio sobre essas questões (p. ex., tipos de validade, tipos e facetas de confiabilidade, vieses de informação) e que pode ser obtido em excelentes textos (Rothman *et al.*, 2008; Streiner & Norman, 2008), inclusive em português (Medronho *et al.*, 2009). Também não nos ateremos às importantes questões envolvidas no processo de coleta de dados, mas estas podem ser localizadas em rica literatura afim (Moser & Kalton, 1984; Bowling, 1997; Reichenheim & Moraes, 2002).

Congregando e aprofundando alguns escritos anteriores (Reichenheim & Moraes, 2002; Reichenheim & Moraes, 2007a; Reichenheim & Moraes, 2007b), este capítulo, em duas partes, se debruça sobre os cuidados necessários no desenvolvimento de novos instrumentos de aferição e as particularidades envolvendo processos de ATC. Preliminarmente, em dois tópicos propedêuticos, procuramos distinguir o que titulamos de “instrumentos dimensionais” (e aos quais se aplicam primordialmente as propostas deste capítulo) dos chamados “pragmáticos” e avaliamos, ainda que de forma bastante sucinta, a questão da necessidade real de se investir em instrumentos novos. As etapas de operacionalização no âmbito do desenvolvimento de instrumentos novos e no que se refere ao processo de adaptação transcultural encontram-se nas Partes I e II, respectivamente, deste capítulo.



■ **Quadro 13.1** Etapas envolvidas com a elaboração de um instrumento novo

	<b>Etapas*</b>	<b>Estratégia de execução</b>
Especificação do mapa do constructo	Visita à teoria que subjaz à proposta do instrumento com vista à explicitação dos conceitos, identificação dos constructos e respectivas dimensões a considerar	<ul style="list-style-type: none"> <li>• Revisão bibliográfica</li> <li>• Apreciação do modelo teórico do estudo</li> <li>• Identificação das dimensões que compõem o constructo</li> <li>• Identificação do possível gradiente de intensidade do objeto teórico</li> </ul>
Especificação do desenho de itens	Proposição de itens que representem as dimensões a estudar	<ul style="list-style-type: none"> <li>• Revisão bibliográfica</li> <li>• Discussão envolvendo pesquisadores, outros especialistas e indivíduos da população alvo</li> </ul>
	Seleção dos itens que comporão as primeiras edições do instrumento (protótipos)  Redação das perguntas	<ul style="list-style-type: none"> <li>• Discussão envolvendo pesquisadores e outros especialistas</li> <li>• Pesquisadores</li> <li>• Discussão envolvendo pesquisadores e indivíduos da população alvo</li> </ul>
Especificação do espaço de desfecho	Discussão do sistema de escores/opções de respostas	<ul style="list-style-type: none"> <li>• Discussão envolvendo pesquisadores, outros especialistas</li> </ul>
Especificação do modelo de medida	Avaliação das características psicométricas dos protótipos	<ul style="list-style-type: none"> <li>• Avaliação de validade dimensional e adequação de itens componentes</li> <li>• Avaliação de confiabilidade (consistência interna, estabilidade temporal etc.)</li> <li>• Avaliação de validade de constructo e de critério</li> </ul>
Decisão	Seleção do instrumento final	<ul style="list-style-type: none"> <li>• Discussão envolvendo pesquisadores e outros especialistas</li> </ul>
	Estudos de corroboração	<ul style="list-style-type: none"> <li>• Utilização do instrumento em outros contextos de pesquisa</li> </ul>

\*Modelo e nomenclatura adaptados de Wilson (2005).

(4) especificação do modelo de medida e (5) de decisão. O Quadro 13.1 sintetiza as etapas, indicando as respectivas estratégias de execução. Seguem algumas considerações.

### Especificação do mapa do constructo

A primeira etapa envolve necessariamente uma imersão na teoria que subjaz à proposta do instrumento a ser desenvolvido. Avaliam-se as dimensões que possivelmente mapeiam o constructo de interesse, bem como os conceitos que subjazem a estas dimensões (Wilson, 2005, Capítulo 2). Faz parte dessa etapa entender, debater e demarcar o que se pode chamar de *espaço de conteúdo* (Streiner & Norman, 2008). A rigor, esta etapa indica a *validade de face*, uma vez que delimita, ainda de forma conjectural e propositiva, a possível estrutura a ser corroborada ou não em etapas subsequentes (p. ex., de especificação do modelo de medida envolvendo estudos psicométricos formais).

### Especificação do desenho de itens

Uma vez especificado o constructo, passa-se para identificação, seleção e redação de seus indicadores, *i. e.*, dos itens que comporão o instrumento. É nessa etapa que os manifestos do constructo são efetivamente mapeados (Wilson, 2005, Capítulo 3).

### Identificação e seleção de itens

Mesmo em se tratando do desenvolvimento de uma ferramenta nova, é importante que o processo retome a busca bibliográfica pela qual se julgou insuficiente o histórico dos instrumentos e que, efetivamente, motivou o desenvolvimento de um novo. A crítica aos já existentes evita a repetição dos mesmos erros identificados no conjunto disponível. Um escrutínio exaustivo permite também identificar o que pode ser aproveitado das experiências anteriores, ainda que isto não deva ser feito sem aprofundamento. À tentação de simplesmente se criar um amálgama de indicadores oriundos de outros instrumentos se opõe

um processo rigoroso que visite todas as etapas necessárias, incluindo, a cadeia de testes formais, conforme descrito na seção “Especificação do modelo de mensuração” adiante. Não se trata, portanto, de simplesmente adicionar itens antigos. Merece ser lembrado que indicadores não têm um significado nominal (*i. e.*, não valem “por si sóis”), mas, formando um conjunto coerente, representam os espaços de conteúdo do constructo (dimensão) subjacente. Nessa perspectiva, itens não podem ser interpretados e utilizados de forma isolada. Esta prática não é incomum, mas, como bem nos alertam Nunnally & Bernstein (1995), podem ocorrer e levar a efetivos problemas de validade.

Na fase inicial de busca de itens é profícuo investir em estudos qualitativos, entre outros, os métodos de consenso pela técnica Delphi, processo de grupos nominais ou grupos focais (Dawson *et al.*, 1992; Denzin & Lincoln, 1994; Krueger, 1994; Bowling, 1997). Nas situações em que nada ou pouco se sabe sobre como certo constructo é percebido pela população-alvo, estudos qualitativos são mandatórios. A meta é reconhecer os indicadores que melhor representem os conceitos de interesse na população alvo. Várias opções devem ser propostas para que uma crítica subsequente avalie e selecione os mais interessantes.

O principal desafio é especificar um conjunto que seja suficientemente completo para garantir a validade de conteúdo, mas que não seja tão extenso a ponto de dificultar a aceitabilidade e aplicabilidade do instrumento. Neste sentido, a despeito de como o conjunto inicial tenha sido proposto – seja via abordagens envolvendo especialistas, seja através de métodos qualitativos focalizados na população-alvo – é capital proceder a uma escolha minuciosa visando diminuir o número de indicadores manifestos que necessitarão refinamento subsequente (p. ex., especificação semântica, análise psicométrica etc.). Para isso, várias técnicas têm sido propostas, dentre as quais, por exemplo, o método de intervalos equidistantes de *Thurstone* ou a técnica de comparação pareada. Foge ao escopo deste texto

entrevistador anotar separadamente se ela foi ou não compreendida. Esta é também uma boa oportunidade para avaliar se as opções de resposta dos itens se adaptam ou não à população-alvo, sendo que possíveis problemas devem ser explicitamente registrados em espaço próprio e/ou relatados em reuniões de seguimento. Tantas "séries" de  $n$  (p. ex.,  $n = 30$ ) entrevistas são realizadas até que um percentual preestabelecido de ajustamento ( $i$ , e., entendimento) em todos os itens seja alcançado (p. ex.,  $\geq 90\%$ ). As avaliações interinas podem ser realizadas pela própria equipe de campo ou pelo grupo de especialistas envolvido no processo desde o começo (etapas anteriores). A partir das evidências encontradas no pré-teste, são escolhidos os protótipos mais promissores e que são testados formalmente na etapa seguinte.

### Especificação do modelo de mensuração

No limite, corroborar e afirmar a adequação de um instrumento de aferição passa pela própria correção do objeto de este pressupõe mapear, tanto em termos de seus nexos internos – relações entre o construto e suas dimensões com os componentes manifestos –, como externos, com os seus contíguos. No que concerne aos nexos externos, estabelecer a validade (adequação) de um instrumento requer, em última instância, corroborar a própria teoria geral da qual faz parte o objeto (evento) mapeado pelo instrumento em desenvolvimento. Claramente que, assim sendo, a avaliação envolve uma enorme gama de procedimentos e métodos de análise, cujo escopo nos remete aos próprios modelos estatísticos apropriados à análise dos objetos em estudo e que, por sua vez, nos remete aos textos de Estatística e Epidemiologia, incluindo os encontrados nos outros capítulos deste livro.

Ainda assim, no escopo precípuo do processo de desenvolvimento de um instrumento, é possível propor alguns enfoques que permitam nortear a avaliação de sua adequação. Fazendo uma ponte com o exposto no parágrafo acima, dois grandes focos de avaliação podem ser propostos.

O primeiro se centra na adequação "interna" do instrumento, envolvendo a apreensão (corroboração) da estrutura dimensional dos indicadores componentes, seja escrutinando suas propriedades singulares, seja avaliando conjuntos de indicadores consolidados nas escalas resultantes. É aqui que a validade de face do espaço de conteúdo, postulada durante o mapeamento do construto, é ou não corroborada.

Cada indicador é testado, não só para avaliar seu peso na formação de uma escala dimensional, mas também se e o quanto este contribui de forma exclusiva a uma das escalas formadoras do construto (dimensão). Também é nessa etapa do processo que se testa e se consolida o "escalonabilidade" do conjunto de itens, independentemente se a escala vier a ser constituída por um score calculado diretamente a partir das análises multivariadas que subjazem ao processo (análise de fatores, por exemplo); se por meio de um score "bruto" obtido através do somatório simples ou ponderado da pontuação dos itens com-pontentes; ou por transformações desses scores em percentis, scores padrão, scores padronizados ou scores normalizados (Streiner & Norman, 2008, Capítulo 7). Também é parte integral

aprofundadas, porém, detalhes podem ser obtidos em Streiner & Norman, 2008 (Capítulo 4) e nas referências encontradas nesse excelente livro.

Uma vez eleito o conjunto de itens que poderiam formar as dimensões do construto a ser representado empiricamente, passa-se ao aprimoramento e adequação semântica dos indicadores escolhidos, estabelecendo-se, para cada qual, uma ou mais alternativas de redação a serem depuradas (testadas) subsequentemente.

Interessa alcançar uma redação direta, curta, objetiva e clara. Frases ambíguas e com múltipla significação devem ser evitadas (Moser & Kalton, 1984; Converse & Presser, 1986; Streiner & Norman, 2008). Recomenda-se que termos empregados na redação estejam sintetizados às particularidades da população-alvo à qual o instrumento está voltado, com destaque às palavras ou expressões de fácil compreensão, harmônicas a cultura em questão e sem erudição desnecessária. Um bom texto (pergunta) deve evitar assertivas "positivas" e "negativas" inseridas no mesmo item, jargão profissional (p. ex., médico) e coloquialismo (gírias) indevido. Do ponto de vista da sequência de itens, recomenda-se que os mais delicados ou constrangedores sejam colocados no final do instrumento, ainda que exceções possam ser encontradas em certos casos.<sup>5</sup>

### Especificação do espaço de desfecho

O passo seguinte consiste em especificar o espaço de desfecho (Wilson, 2005, Capítulo 4), i. e., cuidar do escalonamento de cada indicador (item). Para que um instrumento tenha efetiva funcionalidade, é fundamental que o conjunto de itens seja capaz de posicionar as unidades de aferição (células, indivíduos, municípios etc.) no espaço de conteúdo do construto (dimensão) e lhes atribuir valores e/ou categorias que permitam a demarcação de distâncias e importância. Para isso, é necessário inicialmente sintonizar a métrica interna de cada manifesto ao que espúcia o *mapa do construto* subjacente e delimitado nas etapas anteriores. A literatura sobre o assunto oferece muitas técnicas e estratégias com vista à definição de opções de resposta (p. ex., escalas visuais analógicas, adjetivais, *Likert*, diferenciais semânticas). Para um aprofundamento, o leitor pode encontrar vários subsídios em Moser & Kalton (1984), Converse & Presser (1986), Streiner & Norman (2008) e Wilson (2005).

As etapas de *desenho de itens* e de especificação do espaço de desfecho contemplam uma primeira visita ao campo para que os primeiros lotes de protótipos (propostas alternativas do instrumento) sejam submetidos a uma avaliação de aceitabilidade, comprimento e impacto emocional. Conforme indica o Quadro 13.1, uma técnica interessante é pré-testar o instrumento. O pré-teste deve ser realizado em uma população que, mesmo não tendo sido escolhida de forma aleatória, seja semelhante àquela da população-alvo a qual o instrumento se propõe. Também ajuda executar-lo em um cenário operativo que se aproxime daquele em que se pretende aplicar o instrumento no futuro.

A aplicação formal do instrumento – e que, vale lembrar, nesse ponto do processo de desenvolvimento ainda contempla várias perguntas alternativas para um mesmo item – solicita-se aos respondentes que parafrazeiem cada pergunta, devendo o

<sup>5</sup> Por exemplo, no desenvolvimento do instrumento *Revised Conflict Tactics Scales* usado para avaliar violência entre parceiros íntimos, chegou-se a conclusão que intercalar itens de diversas intensidades (gravidades) seria a melhor forma de apresentá-los aos respondentes (Straus et al., 1996).

<sup>6</sup> Por falta de termos correspondente no nosso léxico, usamos uma tradução livre do inglês *scalability*.

da avaliação “interna” do instrumento estudar a sua confiabilidade segundo diversas facetas (Shavelson & Webb, 1991).

O outro foco de avaliação é “externo”, seja envolvendo uma comparação com um instrumento considerado de referência para definir/captar o evento de interesse central, seja na ausência deste, com outros pretensamente mapeando o tal constructo, ou ainda, explorando as relações com putativos correlatos pertencentes à teoria-geral da qual faz parte o objeto em pauta.

Em essência, a etapa de especificação do modelo de mensuração visa corroborar um ou vários instrumentos candidatos (protótipos) apresentados a partir das etapas anteriores e, como tal, envolve um estudo epidemiológico de porte e necessariamente minucioso do ponto de vista processual. É necessário projetar tamanhos amostrais consideráveis que permitam alcançar precisão satisfatória das estimativas previstas. Por implicarem muitas estimativas simultâneas, alguns métodos psicométricos tendem a utilizar muitos graus de liberdade [p. ex., análises de fatores confirmatórias (Brown, 2006)], levando a imprecisões importantes se a casuística for pequena.

Mas só tamanhos de amostra grandes não bastam. É essencial que todos os rigores de um estudo epidemiológico “tradicional” sejam atendidos (Reichenheim & Moraes, 1998). Vale apontar que a meta aqui é avaliar a adequação de um instrumento (protótipo) em seu potencial máximo, *i. e.*, à luz de uma operacionalização otimizada. Se porventura houver importantes deslizos de aferição levando à baixa confiabilidade – p. ex., inconsistência interna de itens, falta de estabilidade temporal e concordância interobservadores –, os resultados relativos a outros enfoques – p. ex., análises confirmatórias, correlações com outras variáveis na perspectiva de validade de constructo – estarão inevitavelmente afetados. Como uma análise psicométrica visa ser um “depoimento” sobre a qualidade do instrumento e suas “propriedades”, ter-se-á uma situação flagrantemente “falso-negativa” se, por exemplo, um estudo psicométrico inicial acabar descartando indevidamente um instrumento promissor, não pela suas características “imanescentes”, mas por meros percalços processuais.

Uma vez coletados e processados os dados, passa-se às análises propriamente ditas. As possibilidades são muitas e quaisquer tentativas de impor uma sistemática (sequência) de análise soariam arbitrárias. Até mesmo pela limitação de espaço editorial, nos limitamos a oferecer apenas alguns nortes, indicando alguns aspectos psicométricos que mereceriam atenção. Visando aprofundamento futuro, várias referências são oferecidas.

## • Avaliação “interna” de adequação

### *Corroborando dimensionalidade e adequação de indicadores*

À primeira tarefa concerne corroborar a validade dimensional<sup>7</sup> do instrumento e a adequação dos itens componentes. O Quadro 13.2 oferece alguns nortes para que se possam julgar satisfatórias as escalas (e respectivos indicadores) de um instrumento. Métodos multivariados estão no âmago do processo. Este processo pode começar com uma análise de fatores exploratória (AFE) (Gorsuch, 1983; Kline, 1994; Pett *et al.*, 2003;

Loehlin, 2004; Skrondal & Rabe-Hesketh, 2004), ainda que no contexto do desenvolvimento de instrumentos já se tenha alguma estrutura postulada *a priori* quanto à dimensionalidade e itens participantes. Mesmo que a conotação de exploração seja um tanto nebulosa aqui, para que se possa implementar uma análise de fatores confirmatória (AFC) com bases firmes (Maruyama, 1998; Loehlin, 2004; Skrondal & Rabe-Hesketh, 2004; Kline, 2005), no início do processo de análise de um instrumento novo, é boa prática realizar uma AFE prévia. Primeiro, para explorar se efetivamente existe a estrutura multidimensional conjecturada, e segundo, para explorar o comportamento dos indicadores. Evidenciada uma inadequação, nada impede que já nesse ponto da sequência se “volte para a prancheta”, *i. e.*, para fases anteriores com vista ao encontro de novos e melhores itens. O processo iterativo de todo o desenvolvimento é bem nítido.

Ainda que não explícito no Quadro 13.2, o método de teoria de resposta ao item (TRI) (Hambleton *et al.*, 1991; Van der Linden & Hambleton, 1996; Cella & Chang, 2000; Embretson & Reise, 2000; Sijtsma & Molenaar, 2002; De Boeck & Wilson, 2004; Skrondal & Rabe-Hesketh, 2004; Wilson, 2005; Streiner & Norman, 2008) é, de fato, uma AFC baseada em modelos não lineares, apropriado para escalas formadas por itens dicotômicos ou ordinais. Além de se alcançar uma boa especificação do modelo estatístico, uma análise via TRI permite também apreciar algumas propriedades psicométricas atraentes e proveitosas para uma escolha conscienciosa de indicadores. Usando a TRI é possível corroborar a presença de escalonamento conjunto dos indicadores; a capacidade discriminante de cada item; o posicionamento absoluto e relativo dos itens ao longo do contínuo da variável latente (dimensão) subjacente; a abrangência da informatividade dos itens ao longo da escala; e a precisão da informação ao longo do espectro (contínuo) da variável latente (Reichenheim *et al.*, 2007).

Por mais que uma análise via TRI seja proveitosa quando os indicadores são dicotômicos ou ordinais, vale comentar que existe alternativa para acomodá-los em análises de fatores do tipo AFE ou AFC, uma vez que, em princípio, estas utilizam matrizes de correlação que assumem distribuições gaussianas. Uma opção para contornar o real problema da má especificação de modelo ao se aplicarem análises de fatores “tradicionais” a dados discretos (Gorsuch, 1983; Rummel, 1988; Jöreskog & Sörbom, 1996) é utilizar matrizes de correlação tetracóricas ou policóricas obtidas por transformações prévias à submissão à análise (Divgi, 1979; Finney & DiStefano, 2006; Uebersax, 2006). Em alguns aplicativos como, por exemplo, [R] (Fox, 2006) e Stata (StataCorp, 2005; Kolenikov, 2006) essas transformações são atualmente optativas e necessitam, pois, ser ativamente implementadas. Em outros, como Lisrel (Jöreskog & Sörbom, 2006) ou Mplus (Muthén & Muthén, 1998-2007), as transformações são mandatórias quando os indicadores não são intervalares e/ou gaussianos (normais) e, logo, são implementadas automaticamente. Visando melhores estimativas, ajustes e ponderações complexas destas matrizes transformadas têm sido propostas (Finney & DiStefano, 2006) e embutidas nos métodos de estimação implementados em diversos *softwares* específicos (Muthén & Muthén, 1998-2007; Bentler, 2004; Arbuckle, 2006; Jöreskog & Sörbom, 2006).

### *Avaliações formais de confiabilidade das escalas*

O enfoque psicométrico subsequente envolve avaliações formais de confiabilidade das escalas obtidas após a “deapuração” dos itens e satisfatória evidência de dimensionalidade. O objetivo é avaliar em que medida os escores de um instrumento (*i. e.*, das escalas componentes) estão livres de erro aleatório

<sup>7</sup> Alguns autores entendem as avaliações da estrutura dimensional e das propriedades dos indicadores como uma forma de avaliar validade de constructo, mas outros as confinam aos estudos comparando o instrumento com eventos (variáveis) reconhecidamente fazendo ou não parte da teoria geral na qual se aninha objeto de estudo referente ao instrumento. A perspectiva (e respectiva nomenclatura) que damos aqui procura contornar essa cisma.



## ■ Quadro 13.3 Avaliação de confiabilidade

Objetivos	Método	Aspectos a observar
<ul style="list-style-type: none"> <li>• Avaliar a consistência interna das escalas identificadas anteriormente.</li> </ul>	<ul style="list-style-type: none"> <li>• Estimador de confiabilidade composta (<math>\rho_c</math>) obtido via Análise de Fatores Confirmatória (Fornell &amp; Larcker, 1981; Raykov, 1997b; Hair <i>et al.</i>, 1998; Raykov, 2002; Raykov &amp; Shrout, 2002).</li> <li>• Coeficiente <math>\alpha</math> de Cronbach para o caso de variáveis contínuas (Cronbach, 1951; Nunnally &amp; Bernstein, 1995; Osburn, 2000) ou coeficiente de Kuder-Richardson, Fórmula 20 (kr-20) no caso de variáveis discretas (Kuder &amp; Richardson, 1937; Streiner &amp; Norman, 2008). Estimadores alternativos são descritos em Osburn (2000).</li> <li>• Estes métodos só fornecem estimativas consistentes em restritas condições de tau-equivalência e paralelismo (vide texto para detalhes) (Raykov, 1997a).</li> <li>• Correlação entre cada item e o escore total sem o mesmo (item-resto) (Nunnally &amp; Bernstein, 1995).</li> <li>• Percentual de aumento ou redução do coeficiente <math>\alpha</math> ou kr-20 à retirada de cada item da escala (Reichenheim &amp; Moraes, 2006).</li> </ul>	<ul style="list-style-type: none"> <li>• Pontos de corte de decisão (adequação) são discutidos em Nunnally &amp; Bernstein (1995). Proposta dos autores para avaliar adequação de consistência interna (aplicável a <math>\rho_c</math>, <math>\alpha</math> ou kr-20): <math>\geq 0,70</math>.</li> <li>• Correlações item-resto devem ser baixas (p. ex., <math>&lt; 0,20</math>)</li> <li>• Percentual de aumento ou redução devem ser substantivos (p. ex., <math>&gt;</math> ou <math>&lt; 10\%</math>).</li> </ul>
<ul style="list-style-type: none"> <li>• Avaliar a estabilidade temporal (reprodutibilidade intraobservador e teste-reteste) das escalas identificadas anteriormente.</li> <li>• Avaliar a reprodutibilidade interobservador das escalas identificadas anteriormente.</li> </ul>	<ul style="list-style-type: none"> <li>• Para o caso de variáveis contínuas: análise via correlações intraclassa (Shrout, 1998; Streiner &amp; Norman, 2008), sendo a correlação de Pearson (<math>\rho</math>) e o coeficiente de concordância (<math>\rho_I</math>) de Lin (1989) tipos especiais.</li> <li>• Para o caso de variáveis discretas (dicôtomos ou policôtomos): análises de concordância via estimador <math>kappa</math> simples ou ponderado (Fleiss, 1981; Donner &amp; Eliasziw, 1992); alternativamente, coeficiente <math>kappa</math> ajustado para viés e prevalência (Byrt <i>et al.</i>, 1993).</li> <li>• Para o caso de variáveis discretas (dicôtomos ou policôtomos): análises de concordância via estimador <math>kappa</math> simples ou ponderado (Fleiss, 1981; Donner &amp; Eliasziw, 1992); ou, alternativamente, coeficiente <math>kappa</math> ajustado para viés e prevalência (Byrt <i>et al.</i>, 1993).</li> <li>• Estimadores alternativos são descritos em Cicchetti &amp; Feinstein (1990) e em uma revisão de Elmore &amp; Feinstein (1992).</li> </ul>	<ul style="list-style-type: none"> <li>• Pontos de corte de decisão (adequação) são discutidos em Landis &amp; Koch (1977) e Shrout (1998).</li> <li>• Proposta de Shrout (1998) (aplicável a <math>\rho</math>, <math>\rho_I</math> ou <math>\kappa</math>): <ul style="list-style-type: none"> <li>• <math>&gt; 0,1 \rightarrow</math> ausente;</li> <li>• <math>&gt; 0,1-0,4 \rightarrow</math> fraca;</li> <li>• <math>&gt; 0,4-0,6 \rightarrow</math> discreta;</li> <li>• <math>&gt; 0,6-0,8 \rightarrow</math> moderada;</li> <li>• <math>&gt; 0,8-1,0 \rightarrow</math> substantiva (forte).</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>• Avaliar a equivalência (de formas) das escalas identificadas anteriormente.</li> </ul>	<ul style="list-style-type: none"> <li>• Análise pelo método de <i>half-split</i>, que consiste em estimar de forma sistemática (exaustiva) as correlações entre escores de pares de subescalas (formas paralelas) formadas pela metade dos itens constituintes da escala sob escrutínio (Pett <i>et al.</i>, 2003; Streiner &amp; Norman, 2008).</li> </ul>	<ul style="list-style-type: none"> <li>• Como acima.</li> </ul>

(Pedhazur & Schmelkin, 1991). Vale frisar que estas análises servem não apenas para robustecer a evidência sobre a qualidade do estudo relacionado com o desenvolvimento do instrumento em si, mas como uma instância de adequação processual. A longo prazo, uma série de estudos usando certo instrumento e revelando consistentemente uma boa confiabilidade da mensuração (informação) acaba também atestando sobre a sua própria qualidade. Evidências como essas acrescentam ao histórico do instrumento e podem subsidiar uma decisão sobre qual instrumento utilizar em uma pesquisa epidemiológica.

O Quadro 13.3 oferece várias referências que o leitor poderá consultar para obter detalhes sobre a finalidade, mérito e procedimentos concernentes a cada tipo de confiabilidade (consistência interna; estabilidade/reprodutibilidade intra<sup>8</sup> ou interobservador; ou equivalência de formulários).

Cabe aqui um comentário sobre o uso de certas correlações intraclassa, como é o caso do coeficiente  $\alpha$  de Cronbach (1951), reconhecidamente o estimador mais usado para avaliar consis-

tência interna. Este método só fornece estimativas não enviesadas em situações em que as cargas fatoriais e/ou erros são iguais para todos os indicadores de uma escala (situação de *tau-equivalência* e *paralelismo*, respectivamente) (Raykov, 1997a). Trata-se de uma condição muito restrita e incomum, cuja violação leva consistentemente a uma subestimação. A situação corriqueira é conhecida como de *congeneridade* (cargas fatoriais e/ou erros desiguais) e para a qual é preciso utilizar informações obtidas a partir de AFC (cargas e resíduos) afim de se obterem estimativas não enviesadas. Um aprofundamento sobre *confiabilidade composta* obtida via AFC pode ser encontrado em Raykov (1997b; 2002), Raykov & Shrout (2002), Hair *et al.* (1998) e Fornell & Larcker (1981).

Outro ponto a tocar concerne à teoria da generalização (TG) desenvolvida por Cronbach *et al.* (1972), cujo objetivo principal é oferecer uma elaborada sistemática para a redução das fontes de erros aleatórios de mensuração. No caso específico de estudos de desenvolvimento de instrumentos em que diferentes tipos de confiabilidade devem ser buscados, é possível obter uma análise unificada, na qual os componentes de erros são decompostos e cada aspecto (*faceta*, no jargão da TG) avaliado à luz da contribuição dos outros (Cronbach *et al.*, 1972; Shavelson & Webb, 1991; Nunnally & Bernstein, 1995). Por extensão, é também possível

<sup>8</sup> No contexto de instrumentos de autoquestionário ou laboratoriais, a confiabilidade intraobservador tem sido denominada *teste-reteste*.

Nessa direção, os Quadros 13.4 e 13.5 explicitamente discernem duas situações. A primeira, coberta no Quadro 13.4, concerne a objetos de pesquisa em que não há consenso sobre o que seria a referência de aferição (ou padrão-ouro) para o fenômeno de interesse ou quando não é possível defini-la de forma inequívoca. Construtos como *autoestima* e *resiliência* são bons exemplos. Nessa situação, avaliam-se as relações entre as dimensões supostamente captadas pelas diferentes escalas do instrumento, bem como as relações com outros conceitos, atributos e características ligadas à teoria geral na qual se insere o construto sob escrutínio. O encontro de associações previstas ou afinadas com evidências pregressas corrobora e reforça a validade substantiva do instrumento. Avaliar o inverso também é relevante, pois constatar a inexistência de relações entre os conceitos teóricos manifestos pelas escalas em pauta e certos construtos (escalas) reconhecidamente fora do escopo da teoria geral envolvendo o fenômeno de interesse também fortalece a ideia de validade. Constatar a validade de construto é, portanto, a epítome de validade teórica.

**Avaliação de validade de critério quando há um instrumento de referência**

Ainda que não seja impeditivo buscar a validade de construto por meio de comparações com outros componentes de uma teoria geral da qual faz parte o construto de interesse, quando existe um instrumento, exame ou teste de referência para contrastar o "novo" instrumento em desenvolvimento, é próprio se avaliar a validade de critério. Sob esse cunho, Streiner & Norman (2008) distinguem a validade concorrente da validade preditiva. A classificação se baseia na finalidade da prova e depende da cronologia de realização dos testes. A validade concorrente é admissível quando já se tem o resultado de um instrumento de referência na ocasião da aplicação do instrumento em teste e permite a apreciação da validade paralela das informações obtidas através do instrumento de referência são obtidas tempos depois da aplicação do instrumento em teste. Comumente, estudos de validade de critério são utilizados no contexto de instrumentos "pragmáticos" quando é de interesse maximizar custo-benefício, prevenir e planejar ações sanitárias, seja explicitamente reduzindo o próprio instrumento considerado de referência ou ao se propor um completamente diferente, mas que retenha a capacidade de classificação deste.

**Quadro 13.4** Avaliação de validade de construto via comparações com outros componentes da teoria geral

Aspectos a observar	Método	Objetivos
<ul style="list-style-type: none"> <li>• Se e como os conceitos teóricos manifestos pelas escalas dimensionais do construto se relacionam entre si.</li> <li>• Se e como os conceitos teóricos manifestos pelas escalas dimensionais do construto em pauta se relacionam com os outros conceitos prescritos ou postulados pela teoria (validade convergente).</li> <li>• Se os conceitos teóricos manifestos pelas escalas dimensionais do construto em pauta apropriadamente não se relacionam a conceitos que a teoria da qual fazem parte não prescreve (ou postula (validade divergente)).</li> </ul>	<ul style="list-style-type: none"> <li>• Análise exploratória de associações via tabulações envolvendo duas ou três variáveis (estratificada) e usando razão de risco/prevalência ou razão de produtos cruzados (<i>odds ratio</i>) como estimador.</li> <li>• Análise exploratória de associações via coeficiente de correlações de Pearson para variáveis contínuas (Armitage &amp; Berry, 1994) ou coeficientes de correlação não paramétricos (teste de posição de Spearman ou tau-b de Kendall) para variáveis ordinais (Blalock Jr., 1985).</li> <li>• Análise epidemiológica multivariável complexa encerrando o o quadro teórico-conceitual do qual faz parte o construto (e suas respectivas dimensões) sob escrutínio (Rabe-Hesketh &amp; Skrondal, 2005; Rothman <i>et al.</i>, 2008).</li> </ul>	<ul style="list-style-type: none"> <li>• Avaliar a validade de construto quando não há instrumento de referência (padrão-ouro) para o contraste.</li> </ul>

obter um coeficiente de generalização que resuma a fração de erro decorrente do conjunto de abordagens.

Avaliação "externa" de adequação Mesmo que tenha sido possível identificar dimensionalidade, adequação de itens (em termos de variância compartilhada, de, como requer uma análise de fatores) e confiabilidade, a validade em termos de conteúdo (significado) de uma escala necessária ser avaliada explicitamente. Afinal, se um pesquisador visto centralmente captar um construto  $C_1$  (p. ex., *apito social*) inadvertidamente articular uma gama de itens consistentemente alinhada a outro construto  $C_2$  (p. ex., *resiliência*), é bem plausível que os resultados obtidos em análises psicométricas como as descritas anteriormente sejam bastante satisfatórios. Mas nem por isso o instrumento traz "embutida" automaticamente a validade sobre o construto  $C_1$  em foco. Se as situações no dia a dia das pesquisas epidemiológicas são certamente bem menos claras, o exemplo lembra que escrutinar a validade de um instrumento vai além das avaliações dos componentes "interiores" de variância, requerendo um escrutínio adicional das covariâncias das escalas (dimensões) com outros elementos pertencentes à teoria geral subjacente. Como já mencionado, assumir validade de face (dos itens) importa nas fases iniciais do programa de investigação para guiar as discussões e decisões de escolha dos protótipos de instrumentos a serem mais trabalhados. No entanto, diferente do que muitos creem, a validade de face pode não estar condizente com a validade de fato, sendo necessários estudos aprofundados para alcançar uma real sintonia.

**Avaliações de validade (de construto) mediante comparações com outros componentes da teoria geral**

Além da validade de face, vários outros tipos de validade têm sido definidos, propostos, utilizados e, até certo ponto, criticados (Streiner & Norman, 2008). Entretanto, no âmbito do desenvolvimento de instrumentos que buscam conceptualizações dimensionais talvez seja de interesse enfatizar a perspectiva dada por Streiner & Norman (2008), na qual estabelecer a validade de um instrumento, em última instância, é estabelecer a adequação da teoria que a suporta. Estudar a validade de um instrumento é estudar a própria teoria que a embasa, em ciclos de conjecturas e refutações/corroborações. É um processo contínuo pelo qual se determina o grau de credibilidade a ser dado a uma inferência a partir da "leitura" de uma escala (Landy, 1986; Streiner & Norman, 2008).

No entanto, visto sob a ótica precípua do contexto de desenvolvimento de um instrumento de conceptualização dimensional – e que não necessariamente pretende ser uma redução de outro maior, nem uma ferramenta de finalidade pragmática – avaliações da capacidade discriminante das escalas de um instrumento podem ser esclarecedoras. Saber que um instrumento de aplicação em estudos epidemiológicos não só capta o contínuo da variável latente subjacente, mas também está substancialmente “colado” ao que um exame ou instrumento de referência encontraria, é claramente proveitoso e atraente. Os procedimentos apresentados no Quadro 13.5 são exemplos a serem contemplados.

### ▪ Decisão

Uma questão importante, mas nem sempre transparente, concerne ao momento em que é admissível afirmar que um instrumento está “pronto” para ser divulgado e usado. Provavelmente, esta decisão deve ser feita após novas consultas envolvendo os pesquisadores diretamente vinculados ao estudo, bem como outros especialistas da área de interesse.

Como sugere o Quadro 13.1, este também é o momento em que ocorre a seleção e o lançamento do instrumento “final”. Mas as aspas nos lembram que o processo de avaliação da qua-

lidade de um novo instrumento não se esgota no primeiro estudo que o utiliza, mesmo que este seja de porte. Uma primeira edição necessita ser continuamente posta à crítica pelos profissionais interessados. A vasta gama de detalhes e opções, muitas intrinsecamente subjetivas, demanda que o aprimoramento do novo instrumento dependa de debates e negociações contínuas entre pares. Mesmo que as evidências iniciais tenham sugerido validade, é capital que se conheça seu desempenho em outros contextos, inclusive linguístico-socioculturais.

## ► Parte II: Adaptação transcultural de instrumentos de aferição

### ▪ Processo de adaptação transcultural

Para alguns, a adaptação de instrumentos elaborados em outra cultura e/ou idioma se restringe à simples tradução do original ou, por vezes, à comparação literal de sua retrotradução com o instrumento no idioma original. Entretanto, já há algum tempo, pesquisadores de diferentes áreas temáticas vêm sugerindo que este componente de avaliação semântico deva constituir apenas um dos passos necessários ao processo de ATC e

■ Quadro 13.5 Avaliação de validade de critério

Objetivos	Método	Aspectos a observar
<ul style="list-style-type: none"> <li>• Avaliar a validade de critério (concorrente e preditiva) na situação em que há um instrumento “novo” para o qual se usa o escore completo (ordinal ou intervalar), comparado a um instrumento ou exame de referência também de métrica contínua.</li> </ul>	<ul style="list-style-type: none"> <li>• Análise via correlações intraclasse (Bartko, 1976; Shrout &amp; Fleiss, 1979; Shrout, 1998; Streiner &amp; Norman, 2008).</li> <li>• Análises via correlação de Pearson (seu uso nesse contexto requer alguma reserva [vide Bartko (1976), Bland &amp; Altman (1986) ou Streiner &amp; Norman (2008) para detalhes]).</li> </ul>	<ul style="list-style-type: none"> <li>• Grau de concordância/ discriminação do instrumento “novo” sob escrutínio com uma medida de referência</li> </ul>
<ul style="list-style-type: none"> <li>• Avaliar a validade de critério (concorrente e preditiva) na situação em que há um instrumento “novo” para o qual se usa o escore completo (ordinal ou intervalar), comparado a um instrumento ou exame de referência de métrica dicotoma.</li> </ul>	<ul style="list-style-type: none"> <li>• Análise via ROC (Receiver Operating Characteristic analysis) (Tanner &amp; Swets, 1954; Hanley &amp; McNeil, 1982; Streiner &amp; Norman, 2008), observando-se a área abaixo da curva, que indica o grau de discriminação da escala em teste em relação ao instrumento de referência.</li> </ul>	
<ul style="list-style-type: none"> <li>• Avaliar a validade de critério (concorrente e preditiva) na situação em que há um instrumento “novo” dicotomo, comparado a um instrumento ou exame de referência de métrica também em dois níveis.</li> </ul>	<ul style="list-style-type: none"> <li>• Índices de sensibilidade e especificidade (Streiner &amp; Norman, 2008).</li> </ul>	
<ul style="list-style-type: none"> <li>• Avaliar a validade de critério (concorrente e preditiva) na situação em que há um instrumento “novo” para o qual se usa o escore completo, comparado a um instrumento ou exame de referência de métrica em mais de dois níveis.</li> </ul>	<ul style="list-style-type: none"> <li>• Análise via ROC (Tanner &amp; Swets, 1954; Hanley &amp; McNeil, 1982; Streiner &amp; Norman, 2008) entre níveis crescentes (em gradação, p. ex., nível 1 vs. 2+3 e 1+2 vs. 3) do instrumento de referência.</li> <li>• Índices de sensibilidade e especificidade segundo subgrupos/estratos populacionais usando modelagem multivariável (Coughlin <i>et al.</i>, 1992).</li> <li>• Alternativas: (a) análises via índices de sensibilidade e especificidade corrigidos por concordância aleatória (Coughlin &amp; Pickle, 1992); (b) método de razão de verossimilhança (Sackett <i>et al.</i>, 1991); (c) coeficiente Phi de concordância (Streiner &amp; Norman, 2008) entre o escore do instrumento “novo” sob escrutínio e uma medida de referência.</li> </ul>	

como fazer", o que é uma real lacuna diante da recente e crescente presença de estudos desse tipo na literatura brasileira em saúde coletiva.

### • Etapas de operacionalização de um processo de adaptação transcultural

Uma síntese do processo de avaliação de ATC encontra-se no Quadro 13.6. Cada uma das etapas necessária à apreciação dos diferentes aspectos da equivalência é aprofundada a seguir. A equivalência funcional não é elaborada uma vez que, como os próprios proponentes do modelo definem, esta representa uma síntese das pregressas, captando a extensão com que um instrumento realiza o que deste se espera de forma igualmente satisfatória em duas ou mais culturas (Herdman *et al.*, 1998).

#### *Equivalência conceitual e de indicadores (itens)*

A avaliação da equivalência conceitual está intimamente ligada à questão da especificação do *mapa do constructo* abordada na seção de mesmo nome na Parte I do capítulo. Para dar início a um processo de adaptação é preciso, primeiro, visitar os conceitos que subjazem às dimensões formadoras do constructo de interesse (Wilson, 2005, Capítulo 2). A avaliação de equivalência consiste, portanto, na exploração deste constructo e dos pesos dados aos seus diferentes domínios constituintes no local de origem (país, região, cidade etc.) e na população-alvo à qual o instrumento pretende ser aplicado.

Conforme mostra o Quadro 13.6, esta etapa encerra diversas atividades. É boa prática iniciar com uma discussão envolvendo um grupo de especialistas. A finalidade é explorar se os diferentes domínios abarcados pelo instrumento original na definição dos conceitos de interesse são também relevantes e pertinentes ao novo contexto ao qual está sendo adaptado. No processo, avalia-se também a pertinência dos indicadores (itens) propostos para a captação de cada um desses domínios. Estas discussões devem ocorrer à luz de uma revisão bibliográfica que priorize as publicações sobre os processos envolvidos na construção do instrumento-fonte, sem deixar de lado, no entanto, o material bibliográfico disponível no contexto local.

Nessa etapa do processo, interessa também envolver membros da população-alvo, quer através de entrevistas abertas com setos e representativos indivíduos da comunidade (Bowling, 1997), quer mediante atividades coletivas como, por exemplo, os grupos focais (Dawson *et al.*, 1992; Krueger, 1994).

#### *Equivalência operacional*

A equivalência operacional se refere aos aspectos de utilização de um instrumento na população-alvo, de tal sorte que haja uma eficácia semelhante à do original, mesmo que os *modi operandi* não sejam os mesmos. Importa escrutinar as possíveis influências de certas características do instrumento, tais como o veículo e o formato das questões/instruções (p. ex., se em papel impresso ou em forma eletrônica); o cenário de administração (p. ex., se intra-hospitalar ou domiciliar); ou ainda, o modo de aplicação (p. ex., se por meio de entrevista face a face ou autoperenchimento). Também está em pauta debater e decidir o escalonamento de cada indicador, o que coloca esta etapa em estreita sintonia com o que é denominado *espaco do desfecho* (Wilson, 2005, Capítulo 4) e apresentado em seção específica na Parte I. Aqui, no entanto, interessa focalizar especificamente a questão da equivalência, importando, assim, observar o modo de categorização e as possíveis repercussões de se optar por alguma modificação. Como exemplo, uma modificação perfeitamente cabível em situações

não ser tomado como o próprio (Berkanovich, 1980; Patrick *et al.*, 1985; Bucquet *et al.*, 1990; Bravo *et al.*, 1991; Guillemín *et al.*, 1993; Badia & Alonso, 1995; Herdman *et al.*, 1997; Beaton *et al.*, 2000; Dana, 2000; Merenda, 2006; Prince, 2008). Recomendam que o processo seja uma combinação entre um conteúdo de tradução literal de palavras e frases de um idioma ao outro e um processo metódico de sintonização que contemple o contexto cultural e o estilo de vida da população-alvo à qual se aplicará a versão (Guillemín *et al.*, 1993; Herdman *et al.*, 1998; Behling & Law, 2000).

Existem vários artigos na literatura contendo excelentes aplicações de abordagens teóricas e propostas processuais que, *grasso modo*, têm essa visão ampliada (Bullinger *et al.*, 1993; Guillemín *et al.*, 1993; Herdman *et al.*, 1998; Perneger *et al.*, 1999; Beaton *et al.*, 2000; Behling & Law, 2000; Manesstrivongul & Dixon, 2004; Sperber, 2004; Eremenco *et al.*, 2005; Prince, 2008). Todavia, não há consenso quanto às estratégias de execução, o que necessariamente torna uma síntese operacional um mosaico de procedimentos oriundos de diversas fontes. Ainda assim, dirigida pela própria prática dos autores, faz-se uma escolha aqui usando como norte uma destas propostas (Herdman *et al.*, 1997; Herdman *et al.*, 1998).

Lapidada na área de desenvolvimento de instrumentos de aferição sobre qualidade de vida, a proposta de Herdman *et al.* (1998) se alicerça em uma interessante revisão sobre o tema, na qual os autores identificam a piora terminológica encontrada na literatura e a confusão que a consequente superposição gera entre os pesquisadores da área (Herdman *et al.*, 1997). No primeiro de dois importantes artigos também apontam quatro perspectivas que tendem a reger os programas de investigação de ATC (Herdman *et al.*, 1997).

A primeira, denominada "ingênua" (*naive*), se baseia apenas no processo de tradução simples e informal do instrumento original. A segunda, denominada "relativista" afirma a impossibilidade do uso de instrumentos padronizados em diferentes culturas e propõe que somente aqueles concebidos localmente devam ser utilizados. Nesse caso, a noção de equivalência não é pertinente, e por extensão, a possibilidade de interlocução. A terceira perspectiva, cunhada de "absolutista", assume que a cultura tem um impacto mínimo nos constructos a serem mensurados e que estes são invariantes em diferentes contextos. Em termos metodológicos, a ênfase é toda no processo de tradução e retrotradução do instrumento. A última perspectiva, denominada "universalista", não assume *a priori* que os constructos são os mesmos em diferentes culturas. Nesse sentido, é necessário primeiro investigar se um conceito efetivamente existe e/ou se é interpretado similarmente na nova cultura, para depois se estabelecer sua equivalência transcultural através de metodologia própria.

No artigo de sequência publicado em 1998, Herdman *et al.* (1998) propõem um roteiro básico a ser adotado. Assumindo a postura "universalista", apresentam um guia de avaliação do processo de ATC que abrange a apreciação de equivalência entre o instrumento original e aquele a ser adaptado. Definições e detalhes são oferecidos a respeito de seis tipos, a saber, equivalência conceitual, de item, semântica, operacional, de mensuração e funcional.

O objetivo desta Parte II é oferecer uma sistemática operacional aos pesquisadores interessados em utilizar instrumentos oriundos de outros contextos linguístico-socioculturais. Motiva esta perspectiva o interesse de se compararem perfis epidemiológicos e achados de pesquisas realizadas em diferentes localidades e culturas. Outra razão é a relativa carencia de textos estruturados em português sobre "o que é