



**A Medição do Estado de Saúde:
Criação da Versão Portuguesa
do MOS SF-36**

Pedro Lopes Ferreira

Documento de Trabalho 2

1998

Prefácio	4
Agradecimentos	6
1. Introdução	7
2. Modelo conceptual de medição	11
2.1 Evolução do conceito de saúde	11
2.2 Alguns conceitos	16
2.3 Enquadramento conceptual da medição de resultados	17
3 Escalas do SF-36 e sistema de pontuação	19
3.1 Conceitos e escalas	19
3.2 Sistema de pontuação	23
3.2.1 Introdução dos dados.....	24
3.2.2 Transformação dos valores	24
3.2.3 Determinação das escalas	26
3.2.4 Verificação	26
3.3 Administração do questionário	27
4 Adaptação cultural e linguística para português	29
4.1 O problema	29
4.2 Metodologia para a validação cultural	30
4.3 Processo de tradução para português	32
5 Desenho do estudo empírico	34
5.1 Objectivo	34
5.2 Amostra	

6	Resultados	37
6.1	Descrição da amostra	37
6.2	Testes de escalas e consistência interna	38
6.3	Fiabilidade	40
6.4	Validade	42
6.5	Sensibilidade	45
6.6	Comparação com COOP Grávidas	49
7	Conclusão	51
	Apêndices	
A	Versão original do MOS SF-36	53
B	Versão Portuguesa do MOS SF-36	59
C	Questionário sobre o estado de saúde de grávidas	65
D	Estatística descritiva	78
E	Distribuição dos valores por escala	82
	Referências bibliográficas	91

PREFÁCIO

Ao contrário do que muitas vezes se possa pensar, não se pretende que as medidas genéricas de saúde sejam usadas como substitutas de outros procedimentos clínicos mais tradicionais. Aliás, toda a história da evolução deste tipo de medidas se tem baseado em comparações com medidas clínicas, sempre encaradas pelos investigadores como formas de validação do instrumento de medição de resultados e como forma de afirmação perante os clínicos, pela maneira como detectam variações no tempo nem sempre detectáveis pelas medidas usuais.

A motivação para realizarmos este trabalho foi o facto de não existir, em Portugal e em português, um instrumento genérico de medição de resultados em saúde suficientemente validado para poder ser utilizado não só para caracterizar o estado de saúde de populações e o impacto de determinadas medidas a nível estrutural e político, mas também para poder ser usado como apoio à tomada de decisão de prestadores de cuidados.

Estamos convictos de que este trabalho poderá ser um contributo importante para o estudo da saúde pública em Portugal, pois vai permitir munir os interessados de um instrumento de trabalho válido e fiável para a medição da saúde e da percepção que os indivíduos têm do seu próprio bem-estar e da sua qualidade de vida relacionada com a saúde.

O instrumento de medição escolhido por nós foi o chamado SF-36, resultado de um estudo de mais de dez anos levado a cabo por investigadores da *Rand Corporation* nos EUA e que foi denominado *Medical Outcomes Study*. Um dos objectivos deste investigadores era precisamente desenvolver instrumentos para serem usados em monitorização contínua dos resultados de saúde.

O Sistema de Saúde Português, na sua fase de definição e de ajuste às actuais restrições de natureza financeira e culturais, pode certamente ganhar com a utilização de um instrumento como este. Assim esperamos!

AGRADECIMENTOS

É meu dever agradecer a várias pessoas a possibilidade que tive de redigir este texto. Limitar-me-ei àquelas que contribuíram de uma forma mais directa para os trabalhos realizados.

Assim, devo agradecer primeiro à Dr.^a Catherine Acquadro, directora de projecto da sociedade MAPI de investigação médica em Lyon, França, por me ter envolvido num estudo europeu e multicêntrico com o objectivo de criar versões de instrumentos de medição de resultados nas várias línguas faladas na União Europeia. Uma vez iniciado este estudo, foi essencial a colaboração da Dr.^a Barbara Gandek, directora do projecto IQOLA (*International Quality of Life Assessment*) e investigadora do *New England Medical Center*, em Boston, EUA. Parte desta investigação foi financiada por eles.

Gostaria também de agradecer aos vários tradutores que nos apoiaram na fase da adaptação cultural do SF-36 para português e às doentes que tão pacientemente responderam aos vários questionários realizados.

Por fim, quero exprimir o meu agradecimento à Dr.^a Marta Rosete e ao Dr. Victor Raposo, que me apoiaram em diversas fases importantes deste estudo.

1

INTRODUÇÃO

Neste capítulo são apresentados alguns dos fundamentos e desenvolvimentos do MOS SF-36, assim como resultados de algumas aplicações específicas.

Segundo um crescente número de autores, estamos numa era em que os resultados em saúde estão cada vez mais a ser avaliados sob a perspectiva do doente e com base na forma como afectam a sua funcionalidade ou o seu bem-estar, e satisfazem as suas expectativas e necessidades (Geigle e Jones, 1990; Ware, 1992). Estas ideias não são, no entanto, novas como o demonstra, entre outros, Lembcke (1952) com a afirmação:

A melhor medida de qualidade não é a bondade ou a frequência com que um serviço médico é prestado, mas sim quão próximo o resultado está dos objectivos fundamentais de prolongar a vida, aliviar o *stress*, restaurar a função e prevenir a incapacidade.

Já desde o início deste século que existem textos a salientar a importância desta perspectiva (Codman, 1914) mas só mais recentemente é que o restabelecimento ou a manutenção de uma vida mais efectiva foi considerado o principal objectivo dos cuidados médicos (American College of Physicians, 1988; Cluff, 1981; Ellwood, 1988; McDermott, 1981; Schroeder, 1987; Tarlov, 1983) e só a partir da segunda metade deste século é que surgiam preocupações de natureza

metodológica (Katz *et al.*, 1963). No entanto, este tipo de informação não faz geralmente parte dos processos clínicos e ainda não é de acesso fácil por parte dos prestadores de cuidados.

Várias conferências têm sido recentemente dedicadas exclusivamente, ou quase, a este tema (Katz, 1987; Lohr, 1989, 1992; Lohr e Ware, 1987; Staquet 1995; Wenger *et al.*, 1984), o que tem levado a um maior esclarecimento em relação ao enquadramento conceptual com base no qual as medições são feitas. Segundo Ware (1993) têm sido vários os avanços neste domínio, podendo salientar-se uma maior compreensão das dimensões de saúde (Hays e Stewart, 1990; Liang, 1986; Ware *et al.*, 1981), uma maior inclusão deste tipo de medição em ensaios clínicos (Bombardier *et al.*, 1986; Croog *et al.*, 1986; Fowler *et al.*, 1988), em avaliações de política de saúde (Brook *et al.*, 1983; Ware *et al.*, 1986), em inquéritos às populações em geral (Bergner *et al.*, 1981; Stewart *et al.*, 1988, 1989; Ware *et al.*, 1986) e na prática clínica (Nelson e Berwick, 1989).

É interessante também notar que até meados deste século as medições em saúde se centravam na detecção da presença ou ausência de estados negativos de saúde, de limitações funcionais, de sintomas de doença e da existência de problemas agudos e crónicos (Kaplan, 1989). Só na segunda metade do século é que apareceram medidas de bem-estar, de desempenho e de funcionalidade (McDowell e Newell, 1987; Stewart *et al.*, 1978; Ware *et al.*, 1978, 1979; Wenger *et al.*, 1984) e só muito recentemente é que estas preocupações, inicialmente apenas existentes no domínio dos investigadores, passaram para a área da acção política prática e explícita com a aprovação, pelo Congresso dos EUA em 1987, da Lei de Investigação dos Resultados de Saúde dos Doentes que estabelece um programa de investigação em resultados de saúde centrado nas percepções dos doentes. Para além das medidas tradicionais de sobrevivência, dos indicadores clínicos e dos sintomas e problemas específicos de determinadas doenças e tratamentos,

esta lei prevê a utilização de medidas de estado funcional e bem-estar, assim como de satisfação dos doentes.

Como anteriormente referi, inicialmente o debate sobre a utilização e a interpretação deste tipo de avaliação foi levado a cabo principalmente por investigadores metodológicos (Codman, 1914), talvez reflexo da importância conceptual e conteúdo metodológico que têm os instrumentos da medição de estados de saúde (Bowling, 1991), muitos destes investigadores provenientes dos domínios da psicologia, de outras ciências sociais e, só depois, da medicina. Só mais recentemente este debate entrou nas portas das salas e gabinetes de gestores de cuidados de saúde e políticos da saúde que agora manifestam um interesse pelo conteúdo, pelo uso e pela interpretação de índices de saúde ao nível dos investigadores. Ter uma vida mais efectiva (McDermott, 1981) e preservar a função e o bem-estar (Ellwood, 1988; Schroeder, 1987; Tarlov, 1983) começam a ser as metas a atingir pelos vários actores da prestação de cuidados em saúde, isto apesar de relatórios sobre a experiência dos doentes com a doença e tratamento só raramente fazem parte dos processos clínicos.

Foi neste contexto que John Ware e a sua equipa, há cerca de 15 anos, planeou e executou um Estudo dos Resultados Médicos (MOS — Medical Outcomes Study). Um dos objectivos deste projecto era ver se as variações dos resultados nos doentes eram explicadas pelas diferenças nos sistemas de cuidados de saúde, formação e prática clínica. Além disso, era missão deste projecto desenvolver instrumentos práticos para uma monitorização contínua dos resultados (Ware e Sherbourne, 1992; Tarlov *et al.*, 1989; Ware *et al.*, 1993).

Um dos instrumentos desenvolvidos como resultado do projecto MOS foi precisamente o SF-36 (McHorney *et al.*, 1993, 1994; Ware e Sherbourne, 1992) a partir de uma extensa bateria de questionários que incluía quarenta conceitos relacionados com a saúde. Para se criar um novo questionário de uma forma eficiente, seleccionou-se o número mínimo de conceitos de modo a manter as

características do questionário inicial, não se perdendo muito da variância explicada. Pretendia-se um instrumento padrão e que fornecesse informação útil para a interpretação de resultados (Ware *et al.*, 1993), incluindo diferenças entre populações relativamente ao estado físico e mental, a carga para os respondentes, principalmente os possuidores de doenças crónicas e o efeito de tratamento no estado de saúde em geral.

O formato final deste instrumento genérico contém 36 itens cobrindo oito dimensões de estado de saúde e detectando tanto estados positivos como negativos de saúde (Alonso *et al.*, 1995). O SF-36 é considerado uma medida genérica de saúde uma vez que se destina a medir conceitos de saúde que representam valores humanos básicos relevantes à funcionalidade e bem-estar de cada um (Ware, 1987, 1990a). Além disso não é específico de qualquer nível etário, doença ou tratamento.

O seu conteúdo, tanto referente a saúde física como mental, a sua robustez psicométrica e a sua relativa simplicidade são factores que facilitam a sua utilização. Estes factores, associados à existência de um projecto internacional de adaptação do questionário original em diversos países (*International Quality of Life Assessment — IQOLA — Project*) (Aaronson *et al.*, 1992; Ware *et al.*, 1994), estão a transformar o SF-36 Health Survey num dos instrumentos genéricos de medição de estado de saúde com maior potencial na utilização internacional e na avaliação de resultados clínicos (Alonso *et al.*, 1995).

Entre nós, tanto quanto nos é dado saber, este trabalho descreve a primeira vez que este instrumento de medição de resultados em saúde foi adaptado culturalmente para português, validado e aplicado.

2

MODELO CONCEPTUAL DE MEDIÇÃO

Neste capítulo são apresentados alguns dos passos mais importantes da evolução do conceito de saúde, assim como algumas definições operacionais importantes para a elaboração de um enquadramento conceptual para a medição de resultados em saúde.

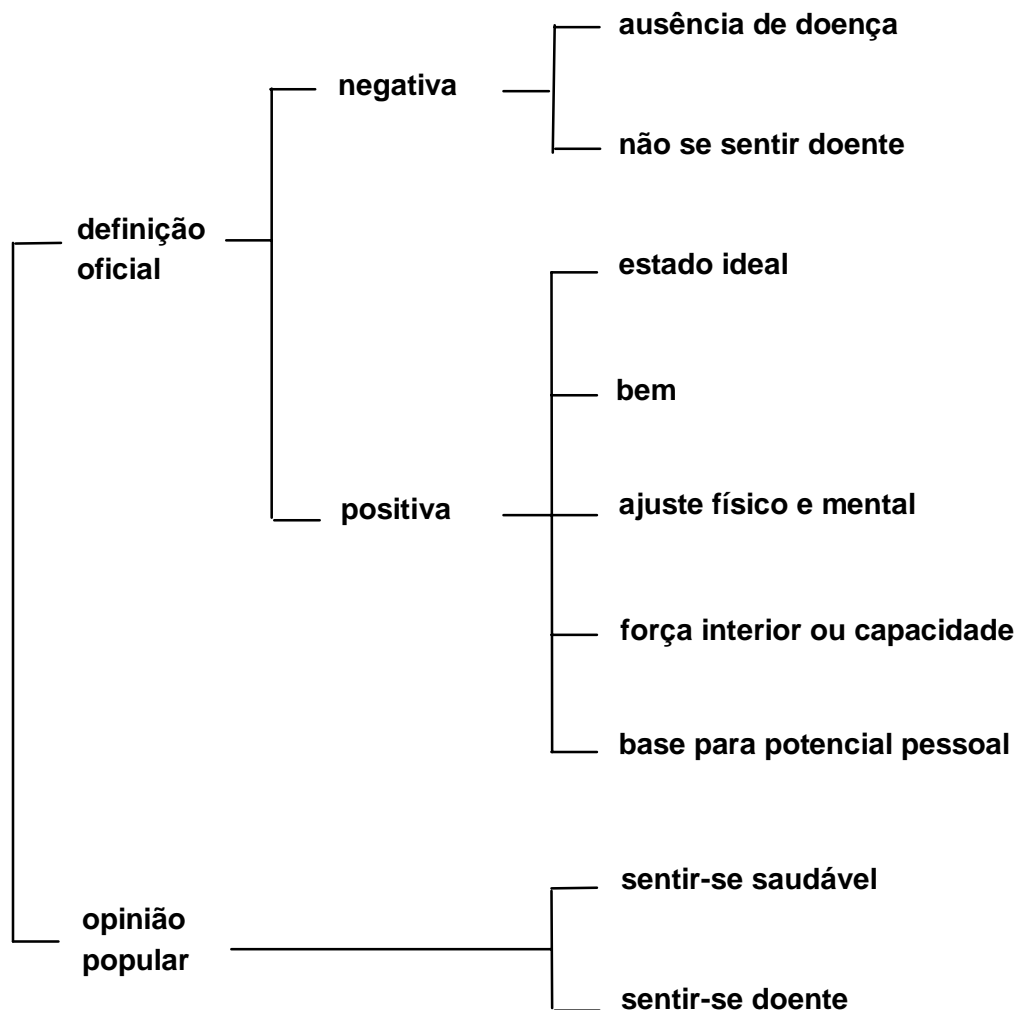
2.1 Evolução do conceito de saúde

Seguindo uma perspectiva histórica, McDowell e Newell (1987) apresentam a evolução dos conceitos relacionados com a saúde nos últimos 150 anos. De início, saúde era definida em termos de sobrevivência, mais tarde foi vista como a negação de doença e depois como algo de positivo e resultante de uma força interior existente em cada um de nós.

Ainda hoje, não há uma única maneira de olharmos e de definirmos o conceito de saúde. Aggleton (1990) dividiu as várias maneiras de definir saúde em dois tipos, conforme o apresentado na Figura 2.1 distinguindo as definições oficiais das chamadas opiniões populares. As primeiras correspondem às perspectivas dos clínicos e de outros profissionais de saúde; as segundas, às percepções populares de saúde dos que não estão envolvidos profissionalmente nas áreas da saúde. Ambas as perspectivas, embora competindo, coexistem e trocam informação entre

si. As próprias percepções populares são por vezes usadas pelos profissionais de saúde como apoio às tomadas de decisão. A saúde é, portanto, uma qualidade relativa.

Figura 2.1 — Tipos de definição de saúde



As definições que pertencem à categoria das oficiais podem ser apresentadas de uma forma negativa, a ausência de certas qualidades como a doença ou o sentir-se doente, ou de uma forma bem mais positiva. Ambas as formas são vitais para uma completa compreensão do significado do conceito de saúde.

Há duas maneiras de encarar a saúde sob uma forma negativa. Pode ser encarada como a ausência de doença ou anormalidade corporal; ou pode ser vista como a ausência da sensação de se sentir doente, de ansiedade, de dor ou de *stress*, sensações que normalmente acompanham uma doença. De acordo com a primeira abordagem negativa, as pessoas são saudáveis até ao ponto em que apresentam sinais de uma qualquer anormalidade corporal. Uma doença existe como resultado de um diagnóstico feito por um médico ou por um outro especialista. Alguns autores, no entanto, interrogam-se sobre a existência de normas universalmente aceites em relação às quais possamos falar em termos de normalidade ou anormalidade. Por outro lado, mesmo que existam, um desvio destas normas não é necessariamente sinónimo de uma sensação *stressante* ou dolorosa.

Este último argumento é o principal responsável pelo deslocamento do enfoque da análise do conceito de estar doente para o conceito de se sentir doente, definido como o conjunto de sentimentos indesejáveis que podem ou não acompanhar uma doença. A saúde pode então ser definida como não se sentir doente, algo que se experimenta. Estaremos saudáveis se não sentirmos dor, ansiedade ou *stress*. O relativismo desta abordagem é acusado por alguns, que argumentam tratar-se de uma definição *ex-post*, ou seja, que só actuamos face à existência de certos sentimentos; prevenção é uma palavra que não faz parte deste vocabulário.

Mas a saúde pode também ser caracterizada em termos positivos e definida pela presença de certas qualidades. Seedhouse (1986) distinguiu cinco teorias principais de definição de saúde em termos positivos. A primeira, considera saúde como um estado ideal. Nesta linha de raciocínio, a Organização Mundial de Saúde, ao definir saúde como um estado de completo bem-estar físico, mental e social e não apenas a ausência de doença ou enfermidade (WHO, 1946), constituiu um marco importante em toda esta discussão. Esta perspectiva denominada socrática

é conceptualmente importante mas, segundo alguns autores, não está associada a uma operacionalização clara. Por outro lado, considera não saudável todo o indivíduo que não atinja um bem-estar completo a nível físico, mental ou social (o que inclui, por exemplo, uma pessoa em estado de pobreza).

Uma segunda teoria vê a saúde como um bem, que é fornecido, dado (através de cirurgia e medicamentos) ou perdido (a seguir a um acidente ou a uma doença), comprado (através de investimento em cuidados de saúde privados) ou vendido (em lojas de comida saudável ou em centros de manutenção física) através de cuidados médicos. Está estreitamente relacionada com o progresso alcançado pela medicina e é algo que existe independentemente das pessoas, um bem último. Através desta perspectiva biomédica, as pessoas são naturalmente saudáveis e podem, sob certas circunstâncias, perder ou recuperar a saúde, procedendo a intervenções correctivas. É pouco mais do que uma matéria técnica, considerando saúde como uma série de qualidades definíveis e mensuráveis.

Uma terceira teoria considera saúde como um grau de ajuste, um estado de capacidade óptima de um indivíduo para o desempenho efectivo de tarefas e trabalho para as quais foi socializado (Parsons, 1972). A saúde já não é vista como um estado desejável *per se*, mas como uma medida no sentido do cumprimento de obrigações estratégicas de desempenho. Com base nesta perspectiva sociológica normativa, saúde significa o conjunto das funções físicas e mentais esperadas, é baseada nas normas e expectativas da sociedade e é definida como uma capacidade óptima, ignorando a existência de graus de saúde.

A quarta teoria considera saúde como uma capacidade ou força interior. Esta perspectiva humanista encara saúde como uma resposta positiva aos problemas e sofrimentos da vida (Dubos, 1959) e parte do pressuposto de que todas as pessoas possuem uma força humana não quantificada, uma reserva de energia, onde a auto-determinação desempenha um papel importante. O radicalismo dos seus defensores em relação à influência médica vai ao ponto de afirmarem que impede

as pessoas de se adaptarem, de uma forma autónoma, às circunstâncias da vida (Illich, 1977). Este tipo de perspectiva é de certo modo vaga e fornece pouca informação sobre a forma de medição destas forças ou capacidades.

Uma última definição oficial defende que saúde é composta por vários factores, a que Seedhouse (1986) chama bases para a realização, que permitem às pessoas atingir os seus potenciais pessoais máximos. Entre estes factores podemos incluir as necessidades básicas da vida (alimentação, água e abrigo), o acesso à informação e a habilidade para desempenhar tarefas. A realização é caracterizada pela alegria, relaxamento, sensação de força e de se sentir bem com os outros.

Concomitante com a definição oficial, surge a opinião popular. A investigação nesta área preocupa-se com a relação entre as crenças referentes à saúde e o contexto cultural alargado na qual ela se situa. Neste contexto, consideramos os fenómenos de se sentir saudável e de se sentir doente. Esta abordagem encara o corpo humano como um todo, o que, obviamente pode variar com o *status*, a integração social, o sexo, ou a raça de cada indivíduo, entre outros factores. As pessoas podem ver-se como saudáveis mesmo estando gravemente doentes. O que interessa é a integridade da pessoa e a sua capacidade de lidar com os problemas de saúde. Os defensores desta abordagem argumentam que algumas pessoas são também capazes de definir saúde em termos funcionais e em termos da sua capacidade de desempenhar alguns papéis e responsabilidades, principalmente as tarefas do quotidiano.

Por fim, existem também opiniões populares referentes às causas de nos sentirmos doentes. Podem ser causadas por forças externas (forças sobrenaturais, uma resposta ou retaliação de qualquer erro cometido ou uma dádiva que permite aos indivíduos aproximarem-se de Deus) ou por factores endógenos (hereditariedade, *stress* e defeitos genéticos).

Para finalizar esta parte, podemos dizer que não existe, no mundo em que vivemos, uma única fonte de informação sobre saúde. A maioria das técnicas usadas para medir saúde têm como base as definições negativas e, em vez de medirem saúde directamente como a presença de certas qualidades, medem-na em termos da detecção da doença ou de se sentirem doentes. Será porque gastamos mais tempo a tratar as doenças do que a preveni-las?

2.2 Alguns conceitos

Para sermos capazes de construir um modelo conceptual para a medição de resultados em saúde é importante clarificarmos alguns conceitos. Assim, resultado em saúde diz respeito à mudança temporal, no enquadramento usado para a saúde, como um resultado de um contacto ou comunicação entre um indivíduo e um profissional de cuidados de saúde.

Vejamos agora definições de alguns outros conceitos.

- Estado de saúde corresponde a uma descrição de saúde, num ou em mais pontos no tempo (abordagem transversal).
- Processo é o que acontece entre a entrada e os resultados.
- A qualidade de vida é a percepção, única e pessoal de vida, marcada por muitos factores interrelacionados, como a situação sócio-económica, o clima político, os factores ambientais, a habitação, a educação, o emprego, entre outros. Os cuidados de saúde não podem alterar os factores que contribuem para a qualidade de vida, mas estes inevitavelmente afectam a saúde e o estado de saúde.
- Qualidade de vida relacionada com a saúde é um subconjunto dos aspectos de qualidade de vida relacionados, na existência individual, com o domínio da saúde

- Qualidade de vida relacionada com os cuidados de saúde, ainda em sentido mais restrito, corresponde aos aspectos da qualidade de vida em que são os prestadores de cuidados os responsáveis pelos resultados das suas acções.

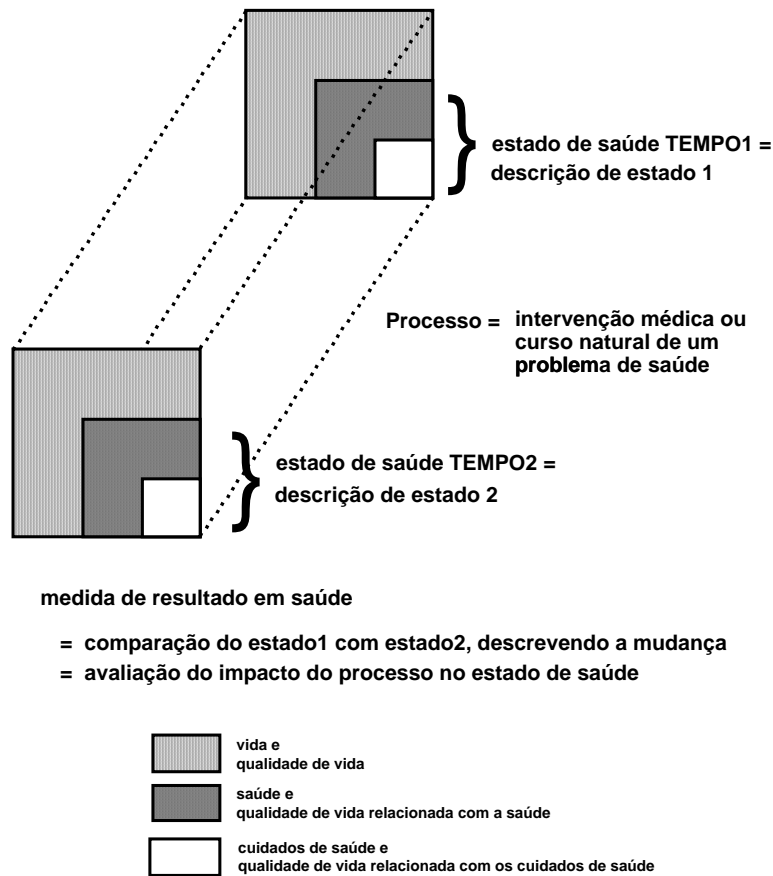
Os resultados de interesse, da perspectiva do prestador de cuidados de saúde e dos organizadores de cuidados, são os indicadores clínicos (sinais, sintomas, categorias nosológicas, condições bioquímicas e psicológicas), o estado funcional (desempenho físico, cognitivo, psicológico e social), a satisfação dos doentes com os vários aspectos de cuidados e os aspectos económicos da qualidade de vida relacionada com a saúde (o valor monetário da mortalidade, morbilidade, longevidade).

2.3 Enquadramento conceptual da medição de resultados

Os dois principais objectivos da medição de resultados na prática clínica são, por um lado, quantificar os progressos dos doentes e, implicitamente, a eficácia de um determinado tipo de cuidados e, por outro lado, comparar a eficácia de um tratamento com outro. No primeiro caso, a atenção vai para a pessoa do doente pretendendo-se detectar as alterações individuais em estados de saúde. Para alcançar o segundo objectivo, submetem-se populações a diferentes tratamentos e estudam-se as diferenças de resultados entre grupos.

Como também se pode extrair do diagrama da Figura 2.2, a medição de resultados em saúde diz respeito *strictu sensu*, ao passado (tempo 1) e ao processo de cuidados, resultando num novo estado de saúde (tempo 2).

Figura 2.2 — Modelo conceptual de medição de resultados em saúde



O efeito do processo (uma intervenção médica ou o desenvolvimento natural, no tempo, de um problema de saúde) no estado de saúde de uma pessoa é quantificado por uma medida de resultado em saúde que tem de ser sensível a alterações, detectando a ocorrência das que são clinicamente importantes. Contudo, é sempre possível a comparação entre duas populações semelhantes submetidas a tratamentos diferentes (por exemplo, um ensaio clínico). Neste caso, a comparação não é apenas feita entre dois estados de saúde, mas também entre dois resultados de saúde.

3

ESCALAS DO SF-36 E SISTEMA DE PONTUAÇÃO

Neste capítulo são apresentadas definições operacionais para cada uma das escalas e é descrito o modelo de medição SF-36, assim como a estrutura factorial com duas dimensões subjacentes às oito escalas iniciais. De seguida, é discutido o esquema utilizado para transformar os dados directamente obtidos dos respondentes em informação passível de ser utilizada posteriormente.

3.1 Conceitos e escalas

O SF-36 foi construído para representar oito dos conceitos mais importantes em saúde e representados no quadro 3.1. Estão também indicados neste quadro o número de perguntas que constituem cada conceito, assim como a respectiva precisão da medição.

A escala correspondente à função física pretende medir desde a limitação para executar actividades físicas menores (como tomar banho ou vestir-se) até às actividades mais exigentes, passando por actividades intermédias como levantar ou carregar as compras da mercearia, subir lanços de escadas, inclinar-se, ajoelhar-se, baixar-se ou mesmo andar uma determinada distância.

As escalas de desempenho medem a limitação em saúde em termos do tipo e da quantidade de trabalho executado. Inclui a limitação no tipo usual de tarefas executadas, a necessidade de redução da quantidade de trabalho e a dificuldade de realizar as tarefas. Ware (1993) considera que ambas as escalas de desempenho — físico e emocional — se aplicam mais a indivíduos reformados ou com mais do que um trabalho-tipo.

Quadro 3.1 — Escalas do SF-36

CONCEITOS	Nº DE ITEMS	Nº DE NÍVEIS
FF Função Física	10	21
DF Desempenho físico	4	5
DC Dor corporal	2	11
SG Saúde geral	5	21
VT Vitalidade	4	21
FS Função social	2	9
DE Desempenho emocional	3	4
SM Saúde mental	5	26
MS Mudança de saúde *	1	5

* não é considerado um conceito de saúde

As escalas da dor representam não só a intensidade e o desconforto causados pela dor mas também a extensão da forma como interfere nas actividades usuais.

A escala da saúde geral pretende medir o conceito de percepção holística da saúde, incluindo não só a saúde actual mas também a resistência à doença e a

aparência saudável, tornando-se assim, aos olhos dos respondentes e em relação às restantes perguntas, menos redundante.

A escala referente à vitalidade inclui os níveis de energia e de fadiga. Esta escala, segundo Ware (1993) permite captar melhor as diferenças de bem-estar.

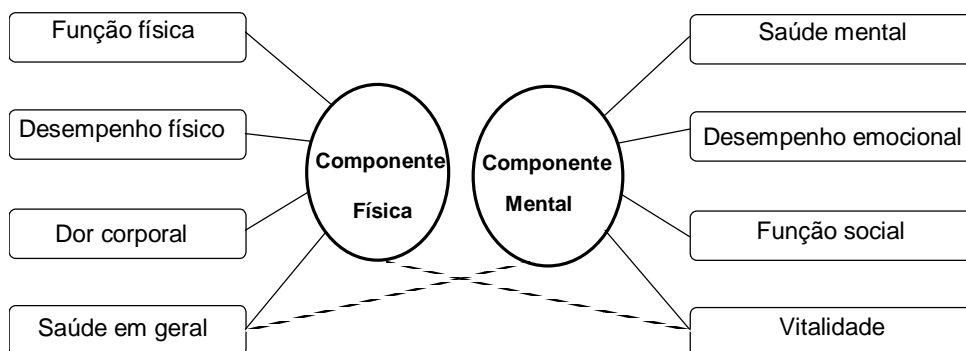
A escala da funcionalidade social estende a medição “para além da pele do indivíduo” (Ware, 1986), pretendendo captar a quantidade e a qualidade das actividades sociais, assim como o impacto dos problemas físicos e emocionais nas actividades sociais do respondente.

A escala da saúde mental inclui questões referentes a quatro das mais importantes dimensões da saúde mental. São elas a ansiedade, a depressão, a perda de controlo em termos comportamentais ou emocionais e o bem estar psicológico.

Por fim, a escala da transição ou mudança de saúde pede ao respondente que informe, com base na experiência previamente vivida, a quantidade de mudança em geral na sua saúde.

Estas oito escalas podem ser agrupadas em dois conceitos (saúde física e saúde mental) conforme se pode ver pelo diagrama da figura 3.1. Estas duas componentes foram obtidas a partir de análises factoriais de componentes principais depois de os dados terem sido sujeitos a rotações ortogonais (Ware et al., 1994).

Figura 3.1 — Modelo factorial SF-36 com duas componentes



As várias escalas contém de 2 a 10 itens e são pontuadas através do método de Likert (1932). Os conteúdos abreviados dos itens de cada escala estão apresentados no quadro 3.2 a seguir.

Quadro 3.2 — Conteúdos abreviados dos itens das escalas do SF-36

ESCALA	ITEM	CONTEÚDO ABREVIADO
FF Função Física	3a	Actividades violentas, tais como correr, levantar pesos, desportos violentos
	3b	Actividades moderadas, tais como deslocar uma mesa ou aspirar a casa
	3c	Levantar ou carregar as compras da mercearia
	3d	Subir vários lanços de escada
	3e	Subir um lanço de escadas
	3f	Inclinar-se, ajoelhar-se ou baixar-se
	3g	Andar mais de 1 Km
	3h	Andar vários quarteirões
	3i	Andar um quarteirão
	3j	Tomar banho ou vestir-se sozinho/a
DF Desempenho físico	4a	Diminuiu o tempo gasto a trabalhar ou noutras actividades
	4b	Fez menos do que queria
	4c	Limitado/a no tipo de trabalho ou outras actividades
	4d	Dificuldade em executar o trabalho ou outras actividades
DC Dor corporal	7	Intensidade das dores
	8	Interferência da dor no trabalho normal

Quadro 3.2 — Conteúdos abreviados dos itens das escalas do SF-36 (cont.)

SG	Saúde geral	1	A sua saúde é: óptima, muito boa, boa, razoável, fraca
		11a	Parece que adoecço mais facilmente do que os outros
		11b	Sou tão saudável como qualquer outra pessoa
		11c	Estou convencido/a que a minha saúde vai piorar
		11d	A minha saúde é óptima
VT	Vitalidade	9a	Cheio/a de vitalidade
		9e	Com muita energia
		9g	Sentiu-se estafado/a
		9i	Sentiu-se cansado/a
FS	Função social	6	Interferência dos problemas de saúde nas actividades sociais normais
		10	Número de casos em que a saúde física interferiu nas actividades sociais
DE	Desempenho emocional	5a	Diminuição do tempo gasto a trabalhar ou noutras actividades
		5b	Fez menos do que queria
		5c	Não trabalhou tão cuidadosamente como era costume
SM	Saúde mental	9b	Sentiu-se muito nervoso/a
		9c	Sentiu-se tão deprimido/a que nada o/a animava
		9d	Sentiu-se calmo/a e tranquilo
		9f	Sentiu-se triste e em baixo
		9h	Sentiu-se feliz
MS	Mudança de saúde	2	Classificação da saúde actual comparada com o que acontecia há um ano

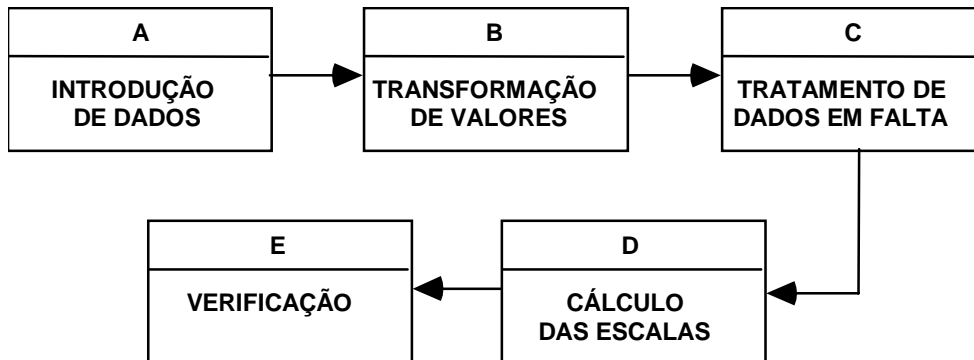
3.2 Sistema de Pontuação

Os dados em bruto provenientes da codificação das respostas ao questionário não estão, à partida e de imediato, em condições de serem usados. Há que os fazer passar por um procedimento, que os torne coerentes e interpretáveis em futuras manipulações.

O procedimento utilizado para o MOS SF-36 é o descrito no diagrama do quadro 3.3. O quadro 3.4 apresenta, para cada dimensão, quais as questões que a

compõem, assim como a correspondente amplitude aceitável para respostas, a transformação resultante da aplicação do sistema de codificação e os limites mínimo e máximo das escalas obtidas.

Quadro 3.3 — Diagrama do sistema de pontuação do SF-36



3.2.1 Introdução dos dados

Os dados são introduzidos segundo a codificação existente no questionário, mas, por vezes, há ambiguidade em relação à codificação mais apropriada. Assim, se o respondente marcou duas respostas adjacentes referentes à mesma questão, devemos escolher, de uma forma aleatória, qualquer uma delas. Se o respondente marcou duas respostas não adjacentes a uma mesma questão, este dado deve ser considerado em falta (*missing data*). Se o respondente, a uma pergunta do tipo "sim/não", responde com um texto em vez de assinalar com uma cruz a resposta certa, isto deve ser encarado como correcto e deve ser considerada a resposta que mais próxima estiver do texto em questão.

Quadro 3.4 — Informação para o sistema de pontuação

(1)	(2)	(3)	(4)			(5)	(6)
DIMENSÃO	PERG.S	VAL.S	TRANSFORMAÇÃO			MIN	MAX
FF Função Física	3a - 3j	1 - 3	—			10	30
DF Desempenho físico	4a - 4d	1 - 2	—			4	8
DC Dor corporal	7	1 - 6	1 → 6.0	4 → 3.1		2	12
			2 → 5.4	5 → 2.2			
			3 → 4.2	6 → 1.0			
	8	1 - 5	7 falta	7=1	7=2,...,6		
			1 → 6.0	1 → 6.0	6 - x		
			2 → 4.75	1 → 4.0			
			3 → 3.5	1 → 3.0			
			4 → 2.25	1 → 2.0			
			5 → 1.0	1 → 1.0			
SG Saúde geral	1	1 - 5	1 → 5.0	4 → 2.0		5	25
			2 → 4.4	5 → 1.0			
			3 → 3.4				
	11a, 11c	1 - 5	—				
	11b, 11d	1 - 5	x → 6 - x				
VT Vitalidade	9a, 9e	1 - 6	x → 7 - x			4	24
	9g, 9i	1 - 6	—				
FS Função social	6	1 - 5	x → 6 - x			2	10
	10	1 - 5	—				
DE Desempenho emocional	5a - 5c	1 - 2	—			3	6
SM Saúde mental	9b, 9c, 9f	1 - 6	—			5	30
	9d, 9h	1 - 6	x → 7 - x				
MS Mudança de saúde	2	1 - 5	—			—	—

— representa a transformação identidade

3.2.2 Transformação dos valores

Todas as possíveis 36 respostas devem ser verificadas. Se ocorrerem respostas para além dos limites definidos na coluna (3) do quadro 3.4, dever-se-á analisar o questionário original e tentar corrigir eventuais erros provocados na fase de entrada de dados. Se isso não for possível, dever-se-á considerar que o dado está em falta.

Além disso, deverá proceder-se à transformação apresentada na coluna (4) do quadro 3.4. Como se pode ver nesta tabela, as transformações são de dois tipos: (1) inversão de valores nos itens das dimensões DC, SG, VT, FS e SM; (2) recalibração nos itens das dimensões DC e SG.

Esta última transformação é o que essencialmente distingue este sistema de codificação do proposto pela Rand Corporation (Rand, 1992).

Uma das vantagens mais importantes de uma escala com vários itens é que podemos estimar a sua pontuação mesmo quando esta contenha informação em falta. Isto só é possível se estiverem, pelo menos, respondidos metade dos itens que compõem essa escala. A solução é dada pela média das respostas transformadas do mesmo indivíduo aos restantes itens da escala (Ware *et al.*, 1980).

3.2.3 Determinação das escalas

Nesta fase, calculamos uma pontuação para cada escala, através da simples soma das respostas aos itens que a compõem. Se o respondente não respondeu a, pelo menos, metade dos itens da escala, a este deve ser associado o código de dado em falta.

De seguida, os valores das escalas são transformados para uma escala de 0 a 100 através da fórmula

$$\text{Escala Transformada} = \left(\frac{\text{SOMA} - \text{MIN}}{\text{MAX} - \text{MIN}} \right) * 100$$

onde SOMA indica o valor da soma dos itens dessa escala depois de recodificados (se necessário) e MIN e MAX, respectivamente, os valores mínimo e máximo correspondentes a essa SOMA e apresentados nas colunas (5) e (6) do quadro 3.4. O valor obtido corresponde à percentagem da total pontuação possível.

3.2.4 Verificação

Os autores do MOS SF-36 aconselham que se faça uma verificação dos valores obtidos e das respectivas distribuições.

Assim, para além da inspecção das distribuições de frequência, devem ser percorridos os seguintes passos de verificação:

- a correlação ente cada escala e os itens que a compõem devem ser positivas e altas (0,30, pelo menos);
- as correlações entre a escala SG e as outras sete devem ser positivas e, com raras excepções, ser altas;
- as correlações entre as 8 escalas e o primeiro factor não rodado obtido pela análise factorial devem ser positivas e altas.

3.3 Administração do questionário

A Health Insurance Experiment (Brook *et al.*, 1983; Valdez *et al.*, 1989; Ware *et al.*, 1986) foi talvez a primeira grande demonstração da fiabilidade, da validade e do potencial dos inquéritos auto-administrados para avaliar o estado de saúde na população em geral. Apesar disso, e no entender de John Ware, no fim desta

experiência ainda não se tinha conseguido demonstrar a capacidade de instrumentos deste tipo para medir o estado de saúde em populações doentes ou idosas (Ware, 1993). Aliás, com excepção do Sickness Impact Profile (Bergner *et al.*, 1981), pouco se sabia, nessa altura, sobre a diferença de funcionalidade e bem-estar entre doentes crónicos ou mentais e a população em geral.

O SF-36 pode ser administrado a qualquer indivíduo com 14 ou mais anos de idade desde que possua a capacidade para ler o questionário. Pode ser auto-administrado, administrado através de uma entrevista, pelo correio ou por telefone. Para evitar um qualquer enviesamento causado pela interacção do doente com o prestador de cuidados, quando aplicado na área clínica, o SF-36 deve ser preenchido antes do sujeito responder a quaisquer outras questões sobre a sua saúde ou doença.

Ao ser uma forma reduzida de um extenso instrumento de medição de estado de saúde, o SF-36 pretende reduzir o peso do preenchimento e o custo da recolha de dados (praticabilidade), não sacrificando em demasia os parâmetros de natureza metodológica (elegância psicométrica). Por exemplo, na escala de funcionalidade física, são apenas necessários 10 itens para definir 21 níveis de funcionalidade (Stewart e Kamberg, 1992).

4

ADAPTAÇÃO CULTURAL E LINGUÍSTICA PARA PORTUGUÊS

Neste capítulo é apresentada a metodologia seguida por nós para proceder à adaptação cultural do MOS SF-36 e para a criação de uma versão portuguesa culturalmente equivalente.

4.1 O problema

Há, hoje em dia, uma grande procura de traduções de questionários de estado de saúde para serem usadas quer em ensaios clínicos ou em estudos multinacionais de eficácia de medicamentos, normalmente levados a cabo por companhias farmacêuticas, quer para comparação de estatísticas da saúde entre populações (Ware *et al.*, 1995). Compilando contribuições de vários especialistas na área, foi recentemente publicado o primeiro livro que foca um conjunto de assuntos e problemas que emergiram quando estes estudos começaram a ser aplicados (Shumaker e Berzon, 1995).

O problema cultural coloca-se essencialmente quando se pretende estabelecer comparações de resultados da aplicação de um mesmo instrumento em culturas diferentes. A necessidade de comparação surge quando são detectadas ou esperadas diferenças de comportamento nos doentes ou nos

prestadores de cuidados, face aos cuidados de saúde prestados (Ferreira e Rosete, 1995). O resultado destas comparações é importante quer para médicos na tomada de decisão clínica ou na avaliação dos procedimentos de intervenção, quer para políticos da saúde na avaliação da eficácia de programas de cuidados de saúde (Touw-Otten, 1994).

No entanto, apesar de existir uma grande necessidade de padronização inter-cultural dos questionários de medição do estado de saúde, são raros os esforços para os testar e os validar (Anderson *et al.*, 1993; Bullinger *et al.*, 1993; U.S. Congress, 1993) e alguns autores têm duvidado que medidas desenvolvidas num país possam ser traduzidas ou adaptadas para serem utilizadas, de uma maneira válida, noutros países (Hunt *et al.*, 1992). O desafio que se coloca à comunidade científica é como definir e usar medidas padronizadas entre países, enquanto simultaneamente se mantém a sensibilidade do instrumento de medição para com aspectos importantes e únicos de cada cultura (Berzon *et al.*, 1981; Figueiredo e Lemkau, 1980; Hunt, 1986a, 1986b; Kleinman *et al.*, 1978; Kuyken *et al.*, 1994).

A maioria das traduções feitas até agora tem sido realizada sem grande validação empírica (Anderson *et al.*, 1993), sendo o projecto *International Quality of Life Assessment (IQOLA)* um dos poucos projectos que seguiram uma metodologia sistemática de adaptação cultural.

4.2 Metodologia para a validação cultural

Este nosso projecto foi integrado nos trabalhos do projecto IQOLA (Aaronson *et al.*, 1992) e seguiu as regras de equivalência cultural apresentadas pelo *European Research Group on Health Outcomes (ERGHO)* (Meadows e Touw, 1995), um projecto financiado pela União Europeia no âmbito do programa Biomed-1/HSR (Ferreira, 1993). Também foram analisados os oito critérios exigidos pelo

Scientific Advisory Committee do Medical Outcomes Trust para que traduções dos seus instrumentos de medição de resultados em saúde possam ser reconhecidas (Medical Outcomes Trust, 1995).

A equivalência entre dois instrumentos de medição aplicados em culturas diferentes passa pelos seguintes critérios:

- Equivalência de conteúdo — Foi estudada a relevância do conteúdo da escala na outra cultura? Uma consistência interna dada pelo coeficiente de Cronbach com valor superior ou igual a 0,70 é considerada satisfatória. Qualquer eliminação ou adição de itens poderá comprometer a liberdade intercultural, ficando o instrumento muito mais dependente sob o ponto de vista cultural.
- Equivalência semântica — Foi a tradução feita por especialistas com o auxílio de traduções e retroversões? A técnica de retroversão descrita por Brisling (1970), usando um ou mais tradutores e um ou mais retrovertores, garante este tipo de equivalência.
- Equivalência técnica — Foi conseguido um acordo entre os dados obtidos pela medida traduzida e os dados obtidos por outras medidas referentes ao mesmo conceito? Foi feita alguma análise da validade concorrencial?
- Equivalência de critério — Foram feitas comparações inter-culturais entre ambas as medidas (a traduzida e a original) em termos dos resultados encontrados em ambas as culturas? A sensibilidade, a especificidade e o poder de resposta do instrumento de medição devem ser considerados nesta comparação. A análise da distribuição dos resultados fornece-nos alguma prova relativamente à capacidade do instrumento em distinguir entre grupos diferentes de doentes.
- Equivalência conceptual — Foi feita alguma comparação entre as estruturas e os domínios da medida? Um processo de análise associado a este tipo de equivalência é a análise das relações existentes entre as construções, através

da correlação ou da análise factorial, com vista à obtenção de grupos idênticos de variáveis em ambas as culturas.

Na secção seguinte apresentaremos os critérios usados para obter a versão portuguesa do MOS SF-36. Em capítulos subsequentes serão apresentados resultados de validação que garantem, em grande parte, as equivalências atrás descritas.

4.3 Processo de tradução para português

Para procedermos à validação cultural do MOS SF-36 e à elaboração da sua versão portuguesa recorreremos à versão publicada (Ware e Sherbourne, 1992) e apresentada em Apêndice A. A tradução para português do questionário seguiu o processo esquematicamente representado na Figura 4.1 a seguir.

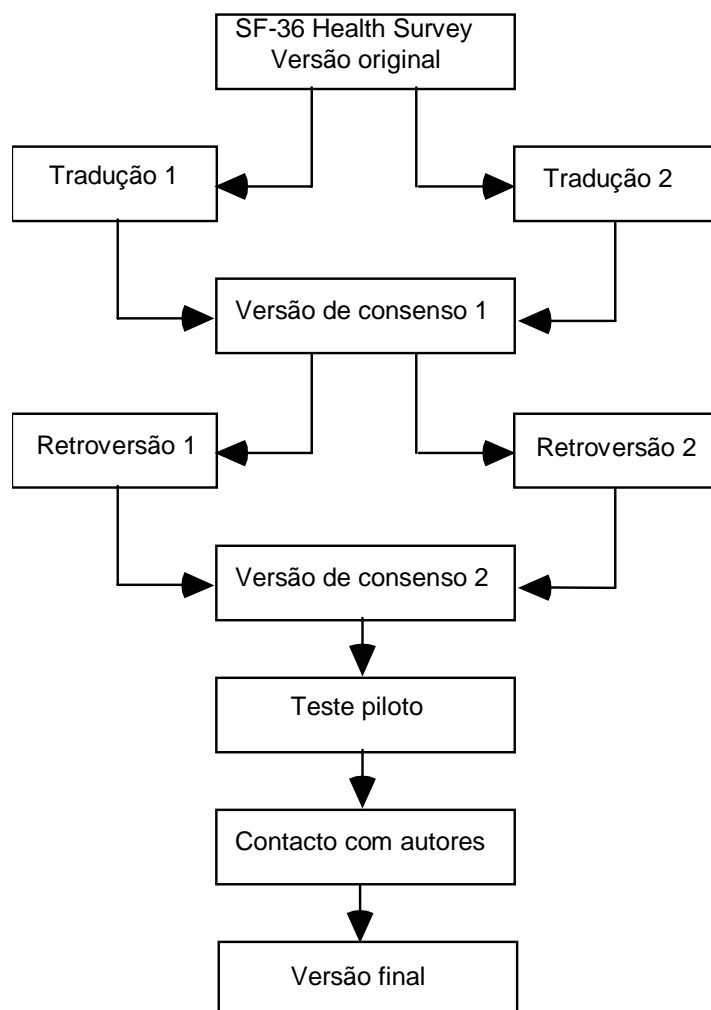
A tradução de inglês para português foi feita por dois tradutores profissionais portugueses fluentes em inglês que, independentemente um do outro, traduziram o original. Durante uma primeira reunião com ambos os tradutores foram discutidas as traduções e as razões para a existência de pequenas diferenças entre elas, obtendo-se assim uma primeira versão de consenso.

Esta primeira versão foi dada a outros dois indivíduos de língua inglesa fluentes em português para que procedessem à sua retroversão para inglês, independentemente um do outro. De igual modo estas duas versões foram analisadas e fundidas dando lugar a uma segunda versão consensual.

Para se atingir uma maior validade ao nível semântico, procedeu-se ainda a um pequeno teste piloto para analisar se as questões estavam a ser colocadas da forma correcta e sem ambiguidade. Assim, o questionário foi entregue a um grupo de pessoas e, após o seu preenchimento, foi conduzida uma pequena entrevista estruturada analisando cada questão segundo os seguintes parâmetros: resposta

difícil (por exemplo, devido às palavras utilizadas), questão confusa e utilidade da questão (tê-la-ia utilizado num questionário?).

Figura 4.1 — Descrição do processo de adaptação



Os resultados deste teste piloto e a versão pré-final dele resultante foram enviados aos autores do instrumento original para serem sujeitos a comentários.

Só então considerámos a tradução apta para ser sujeita a testes de validação e de fiabilidade (Ver Apêndice B).

5

DESENHO DE ESTUDO EMPÍRICO

Neste capítulo é descrito o estudo que fizemos para procedermos à validação da versão portuguesa do MOS SF-36.

5.1 Objectivo

O objectivo deste estudo foi proceder à validação do MOS SF-36. Está inserido num projecto denominado “A história natural da grávida” em curso na Maternidade Dr. Daniel de Matos dos Hospitais da Universidade de Coimbra, cuja coordenação cabe à Faculdade de Economia da Universidade de Coimbra e que conta com a colaboração de alguns Centros de Saúde da região.

Este projecto mais amplo visa detectar alguns amplificadores da história natural da gravidez. Para além dos aspectos físicos e biológicos da gravidez deixados à consideração e ao cuidado dos prestadores, os aspectos psicológicos e de funcionalidade da mulher carecem de maior atenção. São os seguintes os objectivos deste estudo: (1) obter maior conhecimento acerca da história natural da gravidez; e (2) provar a utilidade das medidas de estado funcional da grávida para apoiar as tomadas de decisão clínica.

5.2 Amostra

A amostra consiste em 930 mulheres grávidas ou no período pós-parto. Foram seleccionadas todas as grávidas que acorreram à Maternidade e aos Centros de Saúde. Assim, no prazo de um mês, os médicos distribuíram os questionários (em Apêndice C), tendo sido possível fazer-se uma estratificação referente aos trimestres da gravidez e ao período pós-parto. As grávidas não sentiram qualquer dificuldade em preencher os questionários, tendo-os considerados fáceis de entender.

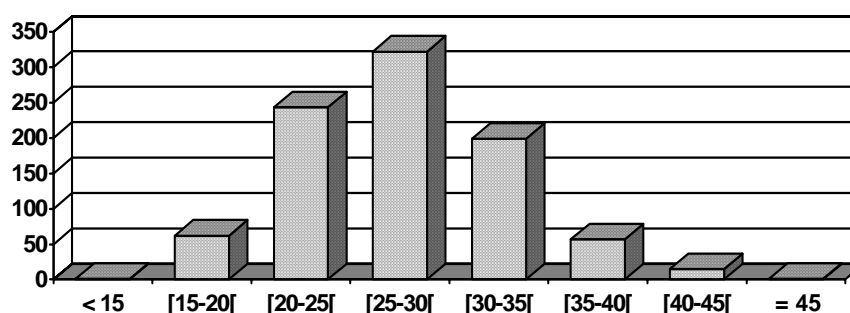
O quadro seguinte apresenta a distribuição da amostra pelos vários locais.

Quadro 5.1 — Distribuição dos locais de recolha de informação

Local	Nº de Grávidas
Maternidade Dr. Daniel de Matos - Consulta Externa	670
Maternidade Dr. Daniel de Matos - Internamento	158
Centros de Saúde de Arganil, Cantanhede e Lousã	102
TOTAL	930

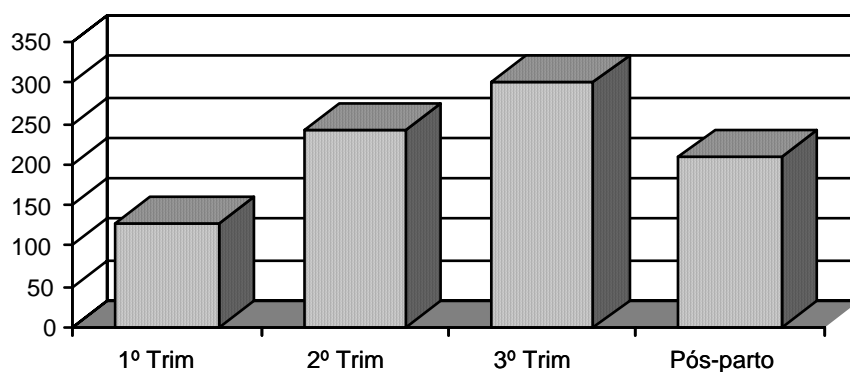
A distribuição etária foi normal com média 26.9 anos e com um desvio padrão de 5.3 anos (ver figura 5.1).

Figura 5.1 — Distribuição etária das grávidas



Destas 930 mulheres, 704 (76%) não tinham tido qualquer complicação durante a gravidez, 126 (14%) estavam no primeiro trimestre da gravidez, 241 (27%) no segundo trimestre da gravidez, 302 (35%) no fim da gravidez e 208 (24%) estavam no período pós-parto. A figura 5.2 apresenta graficamente a distribuição das grávidas por período de gravidez.

Figura 5.2 — Distribuição das grávidas por período de gravidez



Foram também recolhidas outras variáveis socio-demográficas e as relacionadas com a condição de gravidez (variáveis de contexto). De entre estas seleccionámos as variáveis referentes ao número de gravidez e partos anteriores. No caso da mulher ter sido internada foi também recolhida informação referente às datas de admissão e de alta, à razão da admissão, aos diagnósticos, intervenções efectuadas e ao GDH correspondente.

6

RESULTADOS

Neste capítulo são apresentados os resultados da aplicação da versão portuguesa do MOS SF-36 a uma amostra de 930 mulheres grávidas. São indicados os resultados de testes das escalas, incluindo a determinação dos valores de consistência interna e de fiabilidade. No entanto, como um instrumento fiável não é necessariamente válido, por fim, são também apresentados resultados referentes a vários testes de validade.

6.1 Descrição da amostra

Para além dos dados correspondentes às características sócio-demográficas da amostra, procedemos também a uma descrição das respostas ao instrumento SF-36. Em Apêndice D estão apresentadas as frequências absolutas e as percentagens referentes às respostas a cada um dos itens do questionário.

Assim, quase 50% das respondentes consideraram a sua saúde boa, embora esta percentagem seja ligeiramente inferior um anos atrás. Em relação à saúde e actividades diárias, as percentagens de percepções de grandes limitações variam de 12,7% a 34,5%. Em termos de sentimentos (de vitalidade e de saúde mental) o perfil das doentes é francamente positivo.

Enquanto que, em termos de desempenho físico, as respondentes se distribuíram igualmente, em relação ao desempenho emocional houve menos pessoas debilitadas. Só menos de 10% das respondentes afirmaram que tiveram, durante as quatro semanas anteriores, dores fortes ou muito fortes e que as limitaram bastante ou imenso.

Em relação à função social, para cerca de 50% das respondentes a saúde não teve qualquer interferência nem as limitou no seu relacionamento normal com a família, amigos, vizinhos ou outras pessoas.

6.2 Testes das escalas e consistência interna

O primeiro grupo de testes que efectuámos nos dados começou por um teste às escalas que englobou não só o estudo das distribuições dos vários itens dentro de uma mesma escala e a correspondente dispersão, mas também a análise da relação linear entre os valores de cada item e os da escala a que pertencem (consistência interna), assim como o facto de cada item ser mais bem utilizado para medir uma determinada escala do que qualquer outra (validade discriminante). Estes últimos testes foram executados após termos recorrido às técnicas *multitrait* de escalas.

A consistência interna é, portanto, determinada analisando as correlações entre cada item e a escala hipotética, considerando-se, em princípio, uma consistência interna grande a que apresenta valores de correlação para além de 0,40 (Howard e Forehand, 1962). Para além disso, a taxa geral de êxito para uma determinada escala é igual ao número de êxitos de escalonamento dividido pelo número total de testes de escalonamento. Por exemplo, para a escala FF foram executados 10 testes, um por cada item (McHorney *et al.*, 1994; Ware, 1993).

A validação do agrupamento feito passa pelo teste da validade discriminante, isto é, passa pela análise das diferenças entre as correlações de

cada item com a sua escala e as correlações do mesmo item com outras escalas; as primeiras deverão ser, em princípio, maiores do que as segundas. Também aqui a taxa da validade discriminante foi calculada dividindo o número total de êxitos pelo número de testes realizados. Por exemplo, para a escala FF, foram realizados 80 testes, oito por cada um dos dez itens.

O quadro 6.1 apresenta as médias e os desvios padrão dos itens, assim como as correlações entre cada um deles e as respectivas escalas hipotéticas. Daí se pode ver que as variâncias dos itens podem ser comparadas.

Quadro 6.1 — Médias e desvios padrão dos itens e correlações com as escalas

ITEM	MÉDIA	DesvPd	FF	DF	DC	SG	VT	FS	DE	SM
3a	1,87	,74	,09	,21	,15	,03	,23	,03	,01	,13
3b	2,22	,67	,67	,24	,14	,17	,20	,25	,05	,11
3c	2,18	,69	,66	,22	,17	,17	,20	,23	,09	,15
3d	2,08	,71	,45	,23	,15	,09	,20	,10	,12	,13
3e	2,43	,68	,68	,18	,09	,19	,14	,28	,22	,13
3f	2,11	,69	,61	,18	,19	,17	,24	,21	,17	,16
3g	2,30	,74	,69	,25	,15	,19	,16	,26	,19	,09
3h	2,45	,74	,70	,13	,03	,24	,08	,29	,22	,09
3i	2,47	,80	,60	,09	,01	,21	,03	,28	,23	,07
3j	2,47	,80	,60	,09	,01	,21	,03	,28	,23	,07
4a	1,52	,50	,21	,48	,23	,11	,21	,10	,40	,10
4b	1,43	,50	,18	,58	,31	,15	,35	,16	,43	,22
4c	1,39	,49	,25	,59	,31	,17	,35	,16	,28	,19
4d	1,38	,48	,18	,49	,32	,15	,39	,06	,27	,22
7	4,33	1,22	,11	,34	,76	,29	,48	,15	,19	,36
8	4,95	,96	,18	,41	,76	,30	,48	,23	,25	,33
1	3,17	,84	,14	,17	,32	,54	,37	,12	,21	,37
11a	3,89	,91	,28	,15	,21	,56	,25	,26	,21	,28
11b	3,75	,90	,14	,12	,17	,59	,25	,18	,18	,25
11c	3,88	,82	,24	,11	,22	,51	,23	,16	,21	,29
11d	3,46	,85	,12	,17	,24	,62	,37	,14	,21	,33
9a	3,67	1,32	,15	,34	,40	,32	,57	,20	,23	,45
9e	3,38	1,32	,11	,37	,40	,34	,59	,11	,25	,49
9g	3,92	1,16	,21	,31	,39	,31	,62	,22	,19	,40
9i	3,61	1,18	,24	,32	,41	,26	,57	,26	,19	,38
6	4,29	,86	,15	,14	,25	,25	,27	,19	,20	,41
10	4,17	1,29	,32	,11	,09	,14	,14	,19	,20	,09
5a	1,63	,48	,25	,43	,18	,20	,21	,18	,51	,24
5b	1,55	,50	,22	,46	,24	,22	,26	,22	,62	,32
5c	1,56	,50	,15	,19	,13	,23	,17	,19	,40	,25
9b	4,20	1,23	,17	,16	,29	,31	,41	,23	,29	,63
9c	5,15	1,07	,17	,19	,31	,35	,41	,26	,32	,69
9d	3,90	1,32	,12	,23	,35	,36	,52	,21	,28	,67
9f	4,87	1,11	,15	,16	,27	,30	,42	,25	,26	,64
9h	4,63	1,24	,06	,18	,22	,27	,38	,12	,17	,51
2	3,23	,86	,03	,11	,21	,20	,23	,06	,09	,24

Como também se pode ver neste quadro, quase todas as correlações entre cada item e a sua escala igualam ou excedem o ponto de corte 0,40, o que nos conduz a umas taxas gerais de êxito de consistência interna quase perfeitas (quadro 6.2). Também os testes de validade discriminante resultaram em grande êxito por todos os itens da versão portuguesa do SF-36.

Quadro 6.2 — Testes de consistência interna e validade discriminante

Escala	Nº de Items	Amplitude de Correlações		T e s t e	Testes de Consistência	Testes de Discriminação			
		Consistência Interna	Validade Discriminante			Nº Êxitos/ Total	Taxa de Êxito (%)	Nº Êxitos/ Total	Taxa de Êxito (%)
FF	10	0,45 – 0,70	0,01 – 0,29			9/10	90	75/80	94
DF	4	0,48 - 0,59	0,06 – 0,43			4/4	100	32/32	100
DC	2	0,76	0,11 – 0,48			2/2	100	16/16	100
SG	5	0,51 – 0,62	0,12 – 0,37			5/5	100	40/40	100
VT	4	0,57 - 0,62	0,11 – 0,49			4/4	100	32/32	100
FS	2	0,19	0,09 – 0,41			0/2	0	9/16	56
DE	3	0,39 - 0,62	0,13 – 0,46			3/3	100	24/24	100
SM	5	0,51 - 0,69	0,06 - 0,52			5/5	100	40/40	100

6.3 Fiabilidade

Fiabilidade é a medida segundo a qual um instrumento de medição fornece resultados de uma forma consistente e precisa. De uma outra maneira, consiste em determinar quanto da variação em pontos é verdadeira ou apenas devida a erros de natureza aleatória. (> 0,90).

Para testarmos a fiabilidade da versão portuguesa do MOS SF-36 usámos o procedimento do teste-reteste, isto é, administrámos o mesmo questionário duas vezes aos mesmos sujeitos.

Uma outra forma de testar a fiabilidade é calcular a correlação da divisão em metade. Parte-se do pressuposto de que, se se está a medir uma determinada característica, ambas as metades do teste deverão fornecer resultados sensivelmente equivalentes.

Uma outra técnica análoga para a determinação da fiabilidade é o coeficiente alfa que consiste na média de todas as possíveis fiabilidades de divisão em metade ajustadas para o número dos itens (Cronbach e Warrington, 1951).

Os valores de fiabilidade encontrados para as escalas do SF-36 estão apresentados no Quadro 6.3. Todos os estimadores excedem os padrões aceitáveis usados. Para cada escala, a mediana dos coeficientes de fiabilidade iguala ou excede 0,80, com excepção da função social (a média para esta escala com dois itens é 0.76).

Quadro 6.3 — Estimadores de fiabilidade para as escalas do SF-36

Escala	Cronbach	Teste-Reteste	Divisão em metade
FF	0,8731	0,676	0,6881
DF	0,7511	0,786	0,6944
DC	0,8441	0,452	0,8441
SG	0,8745	0,781	0,4501
VT	0,8264	0,722	0,7664
FS	0,6031	0,557	0,6031

DE	0,7104	0,481	0,7543
SM	0,6446	0,779	0,4468

Os elementos da amostra usada (n=930) diferem segundo características socio-demográficas e segundo o período de gravidez em que se encontram.

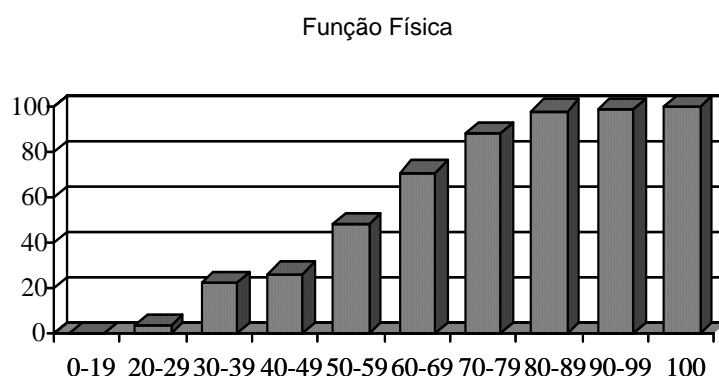
6.4 Validade

Validade é a medida segundo a qual um instrumento de medição mede aquilo que se pretende que meça. No entanto, embora esta frase seja entendível pela grande parte das pessoas, existem diversas estratégias para se testar a validade. Estes estudos sobre a validade permitem aumentar a compreensão do que constituem as diferenças ou as alterações nas pontuações médias.

No campo da saúde, um dos primeiros testes de validade que normalmente é feito é o da validade de conteúdo, isto é, a determinação se o instrumento de medição cobre extensivamente o universo dos objectos ou domínio de medição. A distribuição das principais escalas é analisada com um nível de um dos itens previamente fixado.

A Figura 6.1 representa a distribuição dos pontos da escala de Função Física (FF) correspondente à respostas das mulheres que afirmaram serem capazes de percorrer um quarteirão sem qualquer limitação devida à sua saúde.

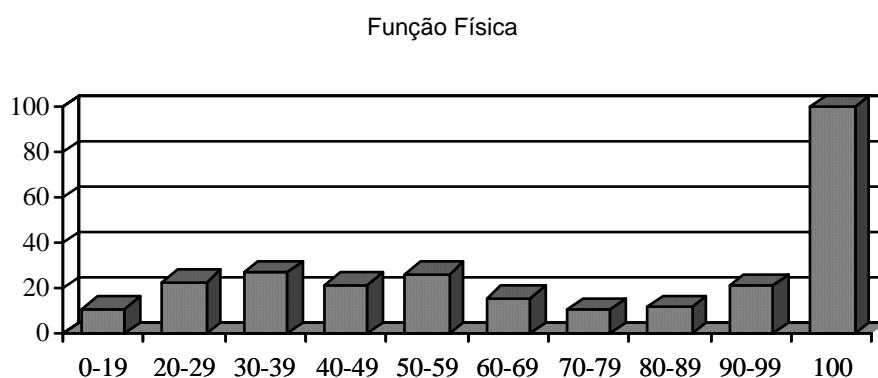
Figura 6.1 — % das mulheres que conseguem andar um ou mais quarteirões



Como se pode ver, nos últimos três níveis da escala FF, praticamente todas as mulheres (98% a 100%) não sentiam qualquer limitação ao andar, enquanto que nos primeiros dois níveis só cerca de 0% a 3% andavam sem limitação. Nos cinco níveis intermédios 22%, 26%, 48%, 71% e 88% indicaram serem capazes de percorrer um quarteirão. Assim, um salto de 45 para 55 na escala FF representa uma diferença de cerca 22 pontos percentuais (48%-26%) na capacidade de percorrer um quarteirão sem limitação.

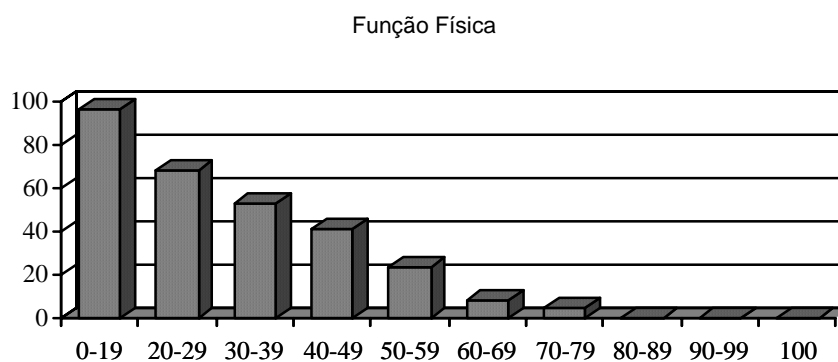
Todos os restantes nove itens que fazem parte desta escala foram analisados. Realçaremos aqui apenas duas situações. A primeira é a correspondente ao item 3a, relativo às actividades violentas, onde todas as mulheres da amostra (100%) no mais alto nível de Função Física não sentiam qualquer limitação (ver Figura 6.2). As percentagens de respostas para os restantes nove níveis da escala variaram de 10,6% a 26,6%. Isto demonstra que este item 3a nos informa se o respondente pontua no mais alto nível; em relação aos outros níveis não há grandes variações.

Figura 6.2 — % das mulheres que conseguem realizar actividades violentas



Tal como também se esperava, as limitações em tomar banho ou vestir-se (item 3j) só se fazem sentir em sujeitos com fracos scores de Função Física (ver Figura 6.3). De facto, os últimos seis níveis da escala FF têm percentagens de limitações que vão de 0% a 23,5%, o que nos permite concluir que este item nos define a parte mais baixa da escala.

Figura 6.3 — % das mulheres que têm limitações a tomar banho ou vestir-se



Mas também analisámos a escala Saúde em Geral. O quadro 6.4 apresenta a distribuição dos níveis da escala SG para os indivíduos que responderam, respectivamente, “óptima” (resposta 1), “boa” (resposta 3) e “razoável ou fraca” (respostas 4 ou 5) à pergunta 1.

Quadro 6.4 — Valores da escala SG segundo a opinião dos respondentes

Níveis	Escala SG		T e s t e	Percentagem		
	Média	N		Ótima	Boa	Razoável/Fraca
91 - 100	94,8	68		26,5	10,3	0,0
81 - 90	85,3	122		13,2	34,7	0,0
71 - 80	74,1	170		6,5	61,5	0,0
61 - 70	64,9	209		1,9	78,7	6,8
51 - 60	55,5	131		3,1	60,8	23,0
41 - 50	47,2	95		0,0	22,0	78,0
31 - 40	38,5	40		5,1	2,6	92,3
21 - 30	25,7	7		0,0	0,0	100,0
11 - 20	20,0	1		0,0	0,0	100,0
0 - 10	7,5	2		0,0	0,0	100,0

Por fim, ao estudar a escala Vitalidade, comparámos a sua distribuição em relação às grávidas que se sentiram sempre ou a maior parte do tempo (respostas 1 ou 2) cansadas (item 9i) ou com muita energia (item 9e). O Quadro 6.5 apresenta estas distribuições.

Quadro 6.5 — Valores da escala VT segundo a opinião dos respondentes

Níveis	Escala VT		T e s t e	Percentagem	
	Média	N		Cansada	Muita energia
91 - 100	96,5	10		0,0	100,0
81 - 90	86,3	42		0,0	97,6
71 - 80	77,7	90		1,1	82,0
61 - 70	67,3	128		6,3	38,3
51 - 60	57,4	148		4,7	11,0
41 - 50	47,8	182		11,0	6,7

31 - 40	37,9	113	25,7	2,7
21 - 30	28,3	83	65,1	2,4
11 - 20	18,8	37	94,6	0,0
0 - 10	6,5	18	100,0	0,0

6.5 Sensibilidade

As medidas em saúde, para além de terem de ser fiáveis e válidas, necessitam também de ser sensíveis. E para testarmos isso, fomos ver como é que as várias dimensões de saúde medidas pelo SF-36 se comportavam quando passávamos pelas sub-amostras dos vários períodos da gravidez e pós-parto.

A validade da construção é testada através da especificação do domínio das variáveis, estabelecendo a estrutura interna das variáveis observadas e verificando relações teóricas entre os scores das escalas e critérios externos. Isto consegue-se através de comparações entre grupos de sujeitos para os quais é de esperar que existam diferenças. Por exemplo, restringindo-nos somente ao *cohort* das mulheres grávidas, é de esperar que mulheres no último trimestre da sua gravidez apresentem valores mais baixos de funcionalidade física devido, essencialmente, ao peso.

Esta mesma validade de construção, muitas vezes é testada através dos testes da validade convergente e da validade discriminante. Obtém-se uma *validade convergente* quando métodos diferentes para medir a mesma construção fornecem resultados semelhantes. A validade discriminante consegue-se quando uma medida de uma construção fornece valores diferentes da medida de outra construção.

A validade de critério demonstra se os scores do teste estão sistematicamente relacionados com um ou mais critérios de resultados. Por

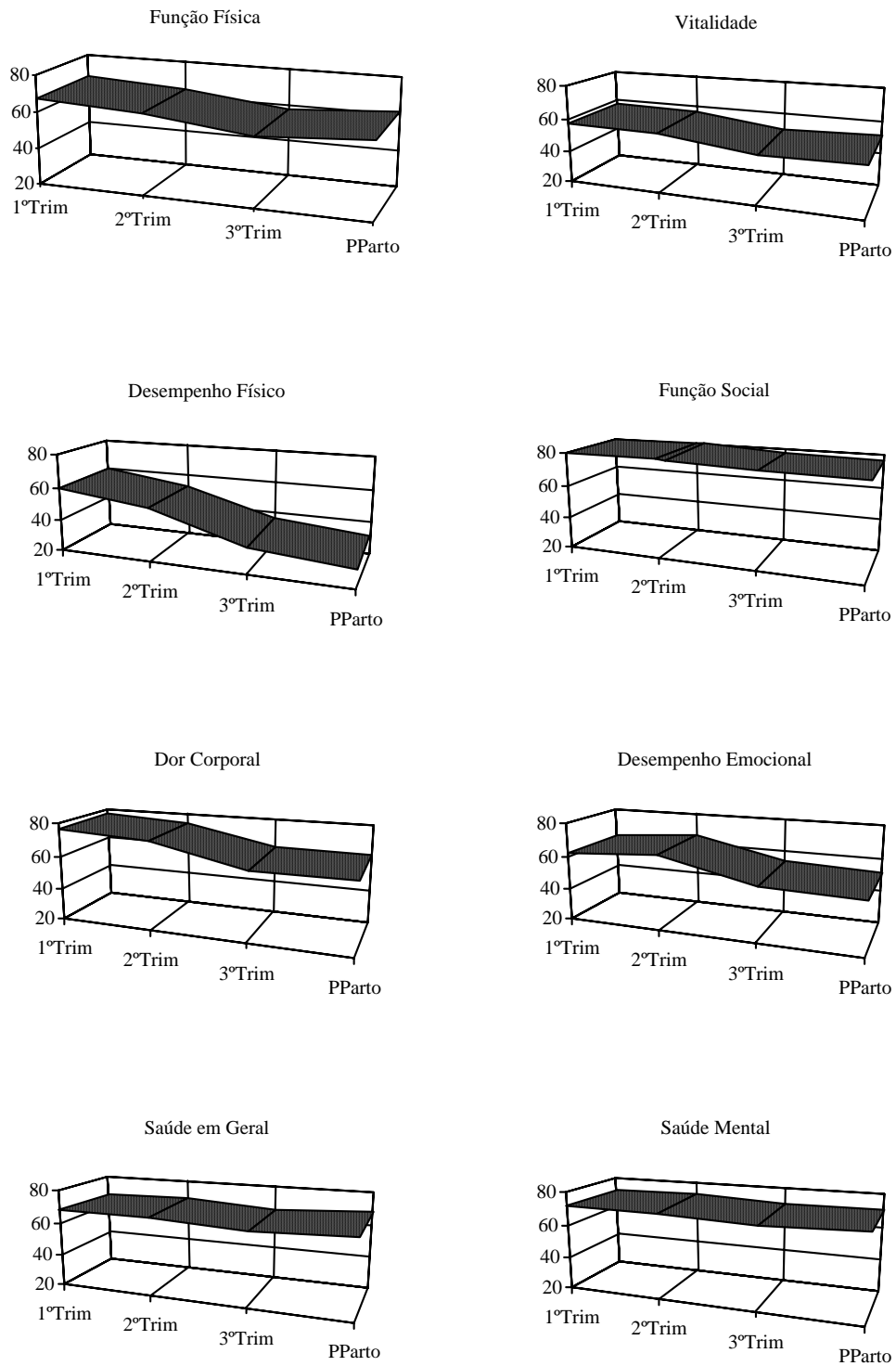
exemplo, é de esperar que as mulheres no 3º trimestre de gravidez tenham alguns indicadores de saúde mais baixos das dos outros períodos.

O quadro 6.6 a seguir apresenta, para cada um dos períodos da gravidez, os valores da média e do desvio padrão para cada uma das dimensões do SF-36.

Quadro 6.6 — Valores das escalas do SF-36 nos diversos períodos da gravidez

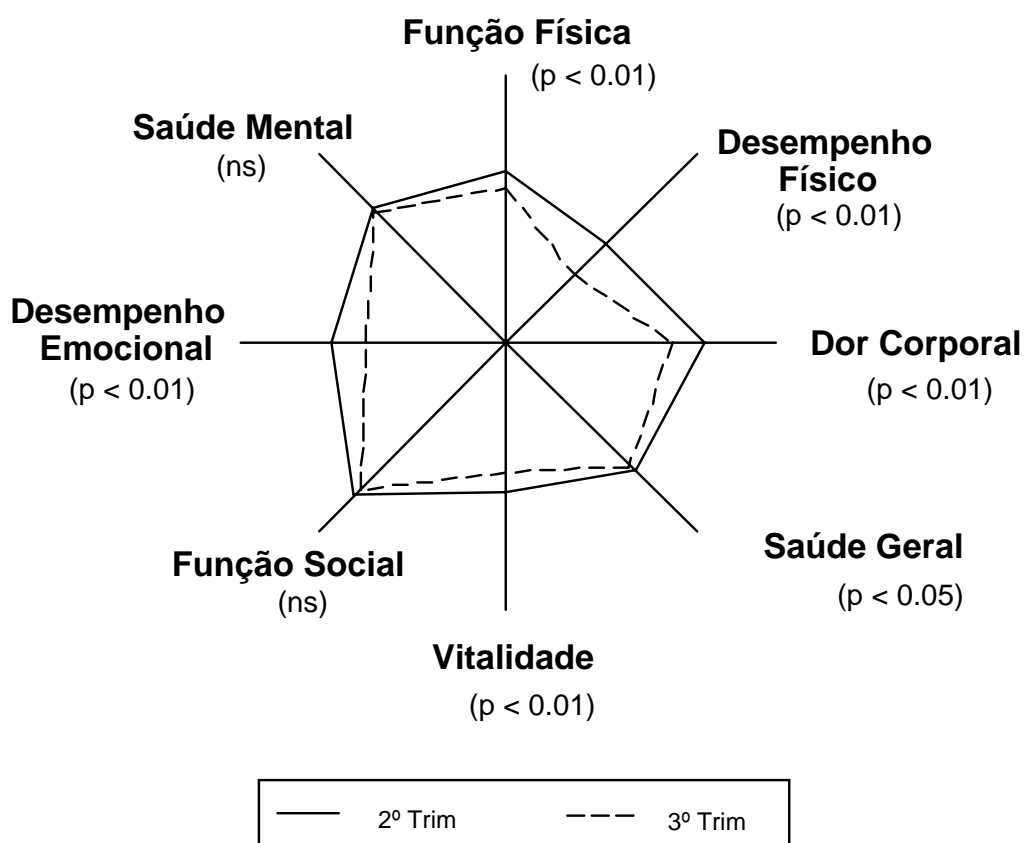
Período		FF	DF	DC	SG	VT	FS	DE	SM
1º Trimestre	n	118	116	117	115	112	117	111	112
	média	67,4	59,5	76,3	67,5	57,5	81,3	61,9	71,7
	desvpd	56,3	37,8	20,2	15,1	19,5	19,7	38,1	18,8
2º Trimestre	n	218	218	227	220	219	224	210	216
	média	64,2	52,0	73,2	68,2	56,1	80,2	65,4	71,6
	desvpd	22,5	37,1	22,8	16,9	18,4	21,0	37,2	19,4
3º Trimestre	n	292	289	295	284	291	294	279	289
	média	57,7	35,7	61,5	65,2	49,1	77,8	53,0	69,7
	desvpd	22,5	34,1	21,1	15,7	19,1	21,1	39,5	17,0
Pós-parto	n	202	203	205	202	206	206	198	205
	média	61,4	30,3	61,4	67,4	50,7	76,8	51,0	70,8
	desvpd	28,5	31,7	24,7	16,2	19,1	23,4	38,5	19,1

Quadro 6.4 — Escalas do SF-36 nos diversos períodos da gravidez



Entre o primeiro e segundo trimestre nada foi possível ser detectado apenas pelo questionário SF-36. Todas as diferenças encontradas eram estatisticamente não significativas. No entanto, entre o segundo e o último trimestre de gravidez, encontrou-se, à excepção das dimensões de saúde mental e função social, uma diferença significativa entre as médias dos vários valores. Isto é apresentado na figura 6.5 a seguir. De novo, entre o último período da gravidez e o período pós-parto, nada foi encontrado.

Quadro 6.5 — Sensibilidade das escalas SF-36 nos diversos períodos da gravidez



Em resumo, excepto na função social e na saúde mental, onde os valores das médias nos vários períodos são quase constantes, assiste-se a uma descida no valor médio de saúde em relação a todos os restantes, com especial incidência no desempenho físico e na dor corporal.

6.6 Comparação com COOP Grávidas

Por fim comparámos os resultados obtidos neste instrumento com os obtidos, para as mesmas pessoas, com o COOP Grávidas, um instrumento especialmente desenvolvido pela Faculdade de Medicina de Dartmouth, EUA especialmente para este grupo de doentes.

O objectivo deste instrumento de medição é detectar problemas de saúde e problemas sociais durante a gravidez e o período pós-parto, um pré-requisito para a gestão de qualquer problema psicossocial. Ao analisar os factores hipotéticos subjacentes aos dez COOP Grávidas podemos ver que os COOPs *Dor/Mal-estar* e *Actividades Diárias* medem o impacto da gravidez na percepção do bem-estar físico. Os COOPs *Sentimentos*, *Stress*, *Estado Emocional* e *Amor Próprio* medem dimensões da saúde mental. O COOP de *Apoio Social* revela se a mulher se apercebe se alguém está disponível para a ajudar. Os COOP *Hábitos de Saúde* medem o grau de envolvimento em comportamentos que podem ser perigosos ou benéficos para a sua saúde ou para a do feto. Finalmente, o COOP *Preparação* examina se será capaz de lidar com as exigências de um recém-nascido, e o COOP *Qualidade de Vida* permite-lhe integrar todas as dimensões de bem-estar numa única escala. Cada COOP tem cinco opções de resposta: uma pontuação de 5 corresponde a uma má pontuação.

A equivalência de critério existe quando o instrumento de medição permite avaliar a mesma variável em ambas as culturas e a interpretação dos resultados (concorrente ou predictiva) é a mesma nas duas culturas. Assim recorreremos ao

instrumento de medição de estado de saúde SF-36, independente do instrumento em estudo, mas que é considerado critério do mesmo fenómeno.

O quadro 6.7 apresenta os resultados das correlações da comparação do SF-36 com o COOP Grávidas. A análise destes resultados fornece-nos alguma evidência relativamente à capacidade do instrumento em distinguir entre grupos diferentes de doentes.

Quadro 6.7— COOP Grávidas ⇔ MOS SF-36

MOS SF-36	COOP Gráficas				
	Mental	Físico	Social	Preparação	Qualidade
Função física	.1959	.2420	.0778	-.0710	.1729
Saúde mental	.6645	.4202	.3858	.1658	.4814
Desempenho	.3426	.4439	.1713	.0195	.2624
Saúde geral	.3406	.3402	.2190	.0633	.4590
Vitalidade	.4286	.5019	.3292	.1888	.4174
Dor corporal	.4156	.6937	.2292	.0981	.3566

Reparar que ambos os instrumentos têm comportamentos semelhantes. No entanto,

- a dimensão COOP Social não corresponde a nenhuma dimensão no SF-36; contudo, está altamente correlacionada com a saúde mental e a vitalidade;
- O COOP Preparação é, de facto, uma dimensão nova; no entanto é interessante verificarmos que está negativamente correlacionada com a dimensão física;
- surpreendentemente, a dimensão física do SF-36 não está correlacionada com nenhuma dimensão COOP, o mesmo acontecendo com a dimensão Alteração de Saúde. Isto parece indicar que, pelo menos para as mulheres desta amostra,

o bem-estar físico está mais relacionado com aquilo que são capazes de fazer no âmbito das suas actividades diárias do que em termos absolutos;

- As limitações devidas a problemas de saúde (física ou emocional) do SF-36 apareceram correlacionadas com os COOP físico e mental;
- A dimensão Saúde em geral do SF-36 correlaciona-se com o COOP Qualidade de Vida.

7

CONCLUSÕES

Neste capítulo conclui-se o trabalho com a recomendação de utilização da versão portuguesa do instrumento de medição de estado de saúde SF-36.

Este estudo permitiu obter a versão portuguesa do questionário de saúde SF-36. O quadro 7.1 resume o que atrás foi dito relativamente à interpretação possível do MOS SF-36. As primeiras escalas medem a dimensão Função Física, as últimas, a dimensão Saúde Mental e as intermédias são as apresentaram uma validade moderada em relação às duas componentes do estado de saúde.

Quadro 7.1 — Resumo da informação sobre o SF-36

ESCALA	Nº DE ITEMS	Nº DE NÍVEIS	MÉDIA	DESVIO PADRÃO	FIABILIDADE
FF Função Física	10	21	61,4	24,9	,86
DF Desempenho físico	4	5	42,2	36,8	,74
DC Dor corporal	2	11	66,8	23,1	,85
SG Saúde geral	5	21	66,9	16,0	,79
VT Vitalidade	4	21	52,7	19,4	,78
FS Função social	2	9	78,7	21,5	,29
DE Desempenho emocional	3	4	57,0	38,8	,69
SM Saúde mental	5	26	70,9	18,5	,83

O número de itens e o número de níveis definidos por cada escala permitem distinguir as escalas mais precisas (FF, SG, VT e SM) com vinte ou mais níveis das menos precisas (DF e DE) com apenas quatro ou cinco níveis.

Embora três das escalas (FF, DF e DC) apresentem validade substancial como medidas de estado de saúde físico é possível distingui-las entre si. Assim, FF mede as limitações em desempenho comportamental nas actividades físicas quotidianas, DF mede a incapacidade em executar tarefas diárias devido a problemas físicos e DC mede essencialmente a gravidade da dor e as limitações resultantes.

Por outro lado, as escalas SM, DE e FS são as que melhor representam a componente mental do estado de saúde. Aqui, enquanto FS e DE medem as limitações ou as incapacidades sentidas pelo sujeito e atribuíveis a problemas pessoais ou emocionais, a escala SM é uma escala bipolar com um *chão* representando sempre nervosismo e depressão e um *tecto* sempre com alegria, paz e calma.

As duas escalas restantes apresentam-se concomitantemente sensíveis a resultados físicos e mentais. A escala VT representa a energia e a fadiga sentidas pelo sujeito e a escala bipolar SG fornece uma pontuação média a quem não avalie desfavoravelmente a sua saúde em termos gerais.

Para tornar a validação da versão portuguesa do MOS SF-36 ainda mais completa, estamos neste momento a implementar um estudo tendente a estudar a sensibilidade a alterações no tempo e um outro tendente à obtenção de normas nacionais.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aaronson NK, Acquadro C, Alonso J, Apolone G, Bucquet D, Bullinger M, Bungay K, Fukuhara S, Gandek B, Keller S, Razavi R, Sanson-Fisher M, Sullivan S, Wood-Dauphinee S, Wagner A, Ware JE. International quality of life assessment (IQOLA) project. *Quality of Life Research* 1992; 1: 349-351.
- Anderson RT, Aaronson NK, Wilkin D. Critical review of the international assessments of health-related quality of life. *Quality of Life Research* 1993; 2: 369-95.
- Alonso J, Prieto L, Antó JM. La versión española del SF-36 Health Survey (Cuestionário de Salud SF-36): un instrumento para la medida de los resultados clínicos. *Medicina Clínica* 1995; 104: 771-776.
- American College of Physicians. Comprehensive functional assessment for elderly patients. *Annals of Internal Medicine* 1988;109: 70-72.
- Aggleton P. *Health*. London: Routledge, 1990.
- Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: Development and final revision of a health status measure. *Medical Care* 1981; 19: 787-805.
- Berzon R, Hays RD, Shumaker SA. International use, application and performance of health-related quality of life instruments. *Quality of Life Research* 1993; 2: 367-68.
- Bombardier C, Ware JE, Russell IJ, Larson M, Chalmers A, Read JL. *Auranofin therapy and quality of life in patients with rheumatoid arthritis: Result of a multicenter trial*. *American Journal of Medicine* 1986; 81: 565-578.
- Bowling A. *Measuring health. A review of quality of life measurement scales*: Buckingham, Reino Unido: Open University Press, 1991.

- Brisling RW. Back-translation for cross-cultural research. *Journal of Cross-cultural Psychology* 1970; 1: 185-216.
- Brook RH, Ware JE, Rogers WH, Keeler EB, Davies AR, Donald CA, Golsberg GA, Lohr KN, Masthay PC, Newhouse JP. Does free care improve adults' health? Results from a randomized controlled trial. *New England Journal of Medicine* 1983; 309: 1426-1434.
- Bullinger M, Anderson R, Cella D, Aaronson N. Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. *Quality of Life Research* 1993; 2: 451-459.
- Cluff LE. Chronic diseases, function and the quality of care. *Journal of Chronic Diseases* 1981; 34: 299-304.
- Codman EA. The product of a hospital. *Surgery, Gynecology and Obstetrics* 1914; 18: 491-496.
- Cronbach LJ, Warrington WG. Time-limit tests: estimating their reliability and degree of speeding. *Psychometrika* 1951; 16: 167-188.
- Croog SH, Levine S, Testa MA, Brown B, Bulpitt CJ, Jenkins D, Klerman GL, Williams GH. The effects of antihypertensive therapy on the quality of life. *New England Journal of Medicine* 1986; 214: 1657-1664.
- Dubos R. *The mirage of health*. New York: Harper and Row, 1959.
- Ellwood PM. Outcomes measurement: A technology of patient experience [Shattuck Lecture]. *New England Journal of Medicine* 1988; 318: 1549-1556.
- Ferreira PL. Abordagem europeia de harmonização da medição do estado de saúde [Notícias]. *Notas Económicas* 1993;1: 105.
- Ferreira PL, Rosete ML. Cross-cultural validation of health outcomes measures: A Portuguese experience. *Quality of Life Research* 1995; 4 (5): 425.
- Figueiredo JM, Lemkau PV. Psychiatric interviewing across cultures: Some problems and prospects. *Social Psychiatry* 1980; 15: 117-21.

- Fowler FJ, Wennberger JE, Timothy RP, Barry MJ, Mulley AG, Henley D. Symptom status and quality of life following prostatectomy, *Journal of the American Medical Association* 1988; 259: 3018-3022.
- Geigle R, Jones SB. Outcomes measurement: A report from the front. *Inquiry* 1990; 27: 7-13.
- Hays RD, Stewart AL. The structure of self-reported health in chronic disease patients. *Journal of Consulting and Clinical Psychology* 1990; 2: 22-30.
- Howard KI, Forehand GG. A method for correcting item-total correlations for the effect of relevant item inclusion. *Educational and Psychological Measurement* 1962; 22: 731-735.
- Hunt SM. Cross-cultural issues in the use of socio-medical indicators. *Health Policy* 1986a; 6: 149-58.
- Hunt Sm, McEwen J, McKenna SP. *Measuring health status*. Dover, NH: Croom Helm, 1986b.
- Hunt SM, McKenna S. Cross-cultural comparability of quality of life measures. *British Journal of Medical Economics* 1992; 4: 17-23.
- Illich I. *Limits to medicine*. London: Pelican Books, 1977.
- Kaplan RM. Health outcome models for policy analysis. *Health Psychology* 1989; 8: 723-735.
- Katz S, editor. The Portugal conference: Measuring quality of life and functional status in clinical and epidemiological research. *Journal of Chronic Diseases* 1987; 40 (Special issue).
- Katz S, Ford AB, Moskowitz RW, Jacobsen BA, Jaffe MW. Studies of illness in the aged: The index of ADL: A standardized measure of biological and psychosocial function. *Journal of the American Medical Association* 1963; 185: 914-919.

- Kleinman A, Eisenberg L, Good B. Culture, illness, and care: Clinical lessons from anthropologic and cross-cultural research. *Annals of Internal Medicine* 1978; 88: 251-58.
- Kuyken W, Orley J, Hudelson P, Sartorius N. Quality of life assessment across cultures. *International Journal of Mental Health* 1994; 23: 5-27.
- Lembcke PA. Measuring the quality of medical care through vital statistics based on hospital service areas: 1. Comparative study of appendectomy rates. *American Journal of Public Health* 1952; 42: 276-86.
- Liang J. Self-reported physical health among aged adults. *Journal of Gerontology* 1986; 41: 248-260.
- Likert R. A technique for the measurement of attitudes. *Archives of Psychology* 1932; 140: 5-55.
- Lohr KN. Advances in health status assessment: Overview of the conference. *Medical Care* 1989; 27: S1-S11.
- Lohr KN. Applications of health status assessment measures in clinical practice. Overview of the third conference on advances in health status assessment. *Medical Care* 1992; 30 (Sup.): MS1-MS14.
- Lohr KN, Ware, JE, editors. Proceedings of the advances in health assessment conference, *Journal of Chronic Disease* 1987; 40 (Sup. 1).
- McDermott W. Absence of indicators of the influence of its physicians on a society's health: Impact of physician care on society. *American Journal of Medicine* 1981; 70: 833-843.
- McDowell I, Newell C. *Measuring Health: A guide to rating scales and questionnaires*. New York: Oxford University Press, 1987.
- McHorney CA, Ware JE, Lu JFR, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care* 1994; 32: 40-66.

- McHorney CA, Ware JE, Raczek AE. The MOS 36-item short-form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care* 1993; 31: 247-263.
- Meadows K, Touw F. Cross-cultural issues in outcome measurement. *In: A, Hutchinson et al. editors. Health outcomes for primary and outpatient care* 1995 (no prelo).
- Medical Outcomes Trust Scientific Advisory Committee. Instrument review criteria. *Medical Outcomes Trust Bulletin* 1995; 3 (4): I-IV.
- Nelson EC, Berwick DM. The measurement of health status in clinical practice. *Medical Care* 1989; 27: S77-S90.
- Parsons T. Definition of health and illness in the light of American values and social structure. *In: Jaco E, Gartley E (editors). Patients, physicians and illness: A sourcebook on behavioral science and health.* London: Collier-MacMillan, 1972.
- Rand Corporation. *Rand 36-item health survey 1.0. Rand health sciences program.* Santa Monica, CA: Rand, 1992.
- Schroeder SA. Outcome assessment 70 years later: Are we ready? *New England Journal of Medicine* 1987; 316: 160-162.
- Seedhouse D. *Health: The foundations for achievement.* Chichester: John Wiley, 1986.
- Shumaker SA e Berzon R (editors). The international assessment of health-related quality of life: Theory, translation measurement & analysis. Oxford: Rapid Communications, 1995.
- Staquet M (editor). Abstracts and programme of the Second Annual Meeting of the International Society for Quality of Life Research (ISOQOL) 14-17 October 1995 Montreal, Canada. *Quality of Life Research* 1995; 4 (5): 373-512.
- Stewart AL, Kamberg CJ. Physical functioning measures. *In Stewart AL, Ware JE, editors. Measuring functioning and well-being: The Medical Outcomes Study approach.* Durham, NC: Duke University Press, 1992: 86-101.

- Stewart AL, Greenfield S, Hays RD, Wells KB, Rogers WH, Berry SD, McGlynn EA, Ware JE. Functional status and well-being of patients with chronic conditions: Results from the Medical Outcomes Study. *Journal of the American Medical Association* 1989; 262: 907-913.
- Stewart AL, Hays RD, Ware JE. The MOS Short-Form General Health Survey: Reliability and validity in a patient population. *Medical Care* 1988; 26: 724-735.
- Stewart AL, Ware JE, Brook RH, Davies-Avery A. *Conceptualization and measurement of health for adults in the Health Insurance Study. Volume II: Physical health in terms of functioning*. Santa Monica, CA: The RAND Corporation (publicação nº R-1987/2-HEW), 1978
- Tarlov AR. The increasing supply of physicians, the changing structure of the health-services system, and the future practice of medicine [Shattuck Lecture]. *New England Journal of Medicine* 1983; 398: 1235-1244.
- Tarlov AR, Ware JE, Greenfield S, Nelson EC, Perrin E, Zubkoff M. The Medical Outcomes Study: An application of methods for monitoring the results of medical care. *Journal of the American Medical Association* 1989; 262: 925-930.
- Touw-Otten F. Cross-cultural issues in health outcome measurement. *SIMG Spring Meeting Estoril, 25-28 May 1994*: 12.
- U.S. Congress, Office of Technology Assessment. *International health statistics: What the number mean for the United States — Background paper*, Washington, DC: U.S. Government Printing Office, 1993, OTA-BP-H-116.
- Valdez RB, Ware JE, Manning WG, Brook RH, Rogers WH, Goldberg GA, Newhouse JP. Prepaid group practice effects on the utilization of medical services and health outcomes for children: Results from a controlled trial. *Pediatrics* 1989; 83: 168-180.
- Ware JE. The assessment of health status. In: Aiken LH, Mechanic D, editors. *Applications of social sciences to clinical medicine and health policy*. New Brunswick, NJ: Rutgers University Press, 1986: 204-228.

- Ware JE. Standards for validating health measures: Definition and content. *Journal of Chronic Diseases* 1987; 40: 473-480.
- Ware JE. Measuring patient function and well-being: Some lessons from the Medical Outcomes Study. In Heitgoff KA, Lohr KN, editors. *Effectiveness and outcomes in health care: Proceedings of an invitational conference by the Institute of Medicine, Division of Health Care Services* . Washington, DC: National Academy Press, 1990a: 107-119.
- Ware JE. Measures for a new era of health assessment. In Stewart AL, Ware JE, editors. *Measuring functioning and well-being* . Durham, NC: Duke University Press, 1992: 1-11.
- Ware JE, Brook RH, Davies-Avery A, Lohr KN. Choosing measures of health status for individuals in general populations. *American Journal of Public Health* 1981; 71: 620-625. (Também RAND publicação nº N-1642-HHS).
- Ware JE, Brook RH, Rogers WH, Keeler EB, Davies AR, Sherbourne CD, Goldberg GA, Camp P, Newhouse JP. Comparison of health outcomes at a health maintenance organization with those of fee-for-service care. *Lancet* 1986; 1: 1017-1022.
- Ware JE, Davies-Avery A, Donald C. *Conceptualization and measurement of health for adults in the Health Insurance Study. Volume V: General health perceptions*. Santa Monica, CA: The RAND Corporation, 1978 (publicação nº R-1987/5-HEW).
- Ware JE, Davies-Avery A, Brook RH. *Conceptualization and measurement of health for adults in the Health Insurance Study. Volume VI: Analysis of relationship among health status measures*, Santa Monica, CA: The RAND Corporation, 1980 (publicação nº R-1987/6-HEW).
- Ware JE, Gandek B, and the IQOLA Project Group. The SF-36 Health Survey: development and use in mental health research and the IQOLA project. *International Journal of Mental Health* 1990; 80: 704-708.
- Ware JE, Johnston SA, Davies-Avery A, Brook RH. *Conceptualization and measurement of health for adults in the Health Insurance Study. Volume III:*

Mental health. Santa Monica, CA: The RAND Corporation, 1979 (publicação nº R-1987/3-HEW).

Ware JE, Keller SD, Gandek B, Brazier JE, Sullivan M. Evaluating translations of health status questionnaires. Methods from the IQOLA Project. *International Journal of Technology Assessment in Health Care* 1995; 11 (3): 525-551.

Ware JE, Kosinski M, Keller SD. *SF-36 physical and mental health summary scales: A user's manual*. Boston, MA: The Health Institute, 1994.

Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 1992; 30: 473-483.

Ware JE, Snow KK, Kosinski M, Gandek B. *SF-36 Health Survey: Manual & Interpretation Guide*. Boston, MA: The Health Institute, New England Medical Center, 1993.

Wenger NK, Mattson ME, Furberg CD, Elinson J. *Assessment of quality of life in clinical trials of cardiovascular therapies*. New York: Le Jacq Publishing Company, 1984.

World Health Organization. *Constitution of the World Health Organization*. Geneva: WHO, 1946.

Apêndice A — Versão original do MOS SF-36

Apêndice B — Versão Portuguesa do MOS SF-36

Apêndice C — Questionário sobre o estado de saúde de grávidas

Apêndice D — Estatística Descritiva

Para termos uma informação mais completa sobre o seu estado de saúde, agradecemos que respondesse às perguntas que se seguem.

ESTADO GERAL.

Por favor coloque um círculo no número que melhor descreve a sua saúde.

- **Em geral**, a sua saúde é:

Ótima	Muito boa	Boa	Razoável	Fraca
58	200	437	175	6
(6,6%)	(22,8%)	(49,9%)	(20,0%)	(0,7%)

- Comparando com o que acontecia há um ano, como descreve o seu presente estado geral:

Muito melhor	Com algumas melhoras	Aproximadamente igual	Um pouco pior	Muito pior
102	123	515	107	13
(11,9%)	(14,3%)	(59,9%)	(12,4%)	(1,5%)

SAÚDE E ACTIVIDADES DIÁRIAS

- As perguntas que se seguem são relativas às actividades do seu dia-a-dia. Será que a sua saúde a limita nestas actividades? Por favor assinale com um círculo um número em cada linha.

	Sim, muito limitada	Sim, um pouco limitada	Não, nada limitada
a Actividades violentas, tais como correr, levantar pesos, participar em desportos violentos	290 (34,5%)	362 (43,0%)	189 (22,5%)
b Actividades moderadas, tais como deslocar uma mesa ou aspirar a casa.	129 (15,2%)	423 (49,8%)	297 (35,0%)
c. Levantar e carregar as compras da mercearia	146 (17,3%)	413 (48,9%)	286 (33,8%)
d Subir vários lanços de escada	187 (22,1%)	402 (47,6%)	256 (30,3%)
e Subir um lanço de escadas	106 (12,7%)	292 (34,9%)	438 (52,4%)
f. Inclinar-se, ajoelhar-se ou baixar-se	169 (19,9%)	437 (51,4%)	245 (28,7%)
g Andar mais de 1 Km	158 (18,9%)	299 (35,7%)	380 (45,4%)
h Andar vários quarteirões	137 (16,7%)	193 (23,7%)	486 (59,6%)
i. Andar um quarteirão	188 (22,0%)	128 (15,0%)	540 (63,0%)
j. Tomar banho ou vestir-se sozinho/a	188 (22,0%)	128 (15,0%)	540 (63,0%)

- Durante as últimas 4 semanas teve, no seu trabalho ou actividades diárias, algum dos problemas apresentados a seguir em consequência do seu estado de saúde? Por favor ponha um círculo à volta do número 1, se a sua resposta for Sim, e à volta do número 2, se a resposta for Não.

	Sim	Não
a. Diminuiu o tempo gasto a trabalhar, ou noutras actividades	407 (48,6%)	430 (51,4%)
b. Fez menos do que queria	494 (58,2%)	355 (41,8%)
c. Sentiu-se limitada no tipo de trabalho ou de actividades	514 (62,1%)	314 (37,9%)
d. Teve dificuldade em executar o seu trabalho ou outras actividades (por exemplo, foi preciso mais esforço)	530 (62,3%)	321 (37,7%)

- Durante as últimas 4 semanas, teve com o seu trabalho ou com as suas actividades diárias, algum dos problemas apresentados seguidamente por causa de qualquer problema emocional (tal como sentir-se deprimida ou ansiosa).
Por favor ponha um círculo à volta do número 1, se a sua resposta for Sim, e à volta do número 2, se a resposta for Não.

	Sim	Não
a. Diminuiu o tempo gasto a trabalhar, ou noutras actividades	319 (38,7%)	506 (61,3%)
b. Fez menos do que queria	383 (46,5%)	440 (53,5%)
c. Não executou o trabalho ou outras actividades tão cuidadosamente como de costume	373 (44,7%)	461 (55,3%)

Para as perguntas que se seguem, por favor ponha um círculo no número que melhor descreve a sua saúde.

- Durante as últimas 4 semanas, em que medida é que a sua saúde física ou problemas emocionais interferiram com o seu relacionamento normal com a família, amigos, vizinhos ou outras pessoas?

Absolutamente nada	Pouco	Moderadamente	Bastante	Imenso
428	291	103	32	3
(49,9%)	(34,0%)	(12,0%)	(3,7%)	(0,4%)

- Durante as últimas 4 semanas teve dores?

Nenhumas	Muito fracas	Ligeiras	Moderadas	Fortes	Muito fortes
189	195	257	170	45	9
(21,9%)	(22,5%)	(29,7%)	(19,7%)	(5,2%)	(1,0%)

- Durante as últimas 4 semanas, de que forma é que a dor interferiu com o seu trabalho normal (tanto o trabalho fora de casa como o trabalho doméstico)?

Absolutamente nada	Um pouco	Moderadamente	Bastante	Imenso
262	366	144	67	11
(30,8%)	(43,1%)	(16,9%)	(7,9%)	(1,3%)

OS SEUS SENTIMENTOS

- As perguntas que se seguem pretendem avaliar a forma como se sentiu durante o último mês. Para cada pergunta, coloque por favor um círculo à volta do número que melhor descreva a forma como se sentiu. Certifique-se que coloca um círculo em cada linha.

Quanto tempo, nas últimas quatro semanas ...	Sempre	A maior parte do tempo	Bastante tempo	Algum tempo	Pouco tempo	Nenhum tempo
a. Se sentiu cheia de vitalidade?	70 (8,6%)	203 (24,9%)	131 (16,1%)	249 (30,6%)	130 (16,0%)	31 (3,8%)
b. Se sentiu muito nervosa?	29 (3,4%)	71 (8,4%)	96 (11,3%)	261 (30,8%)	286 (33,8%)	104 (12,3%)
c. Se sentiu tão deprimida que nada a animava?	10 (1,2%)	19 (2,3%)	22 (2,6%)	134 (16,0%)	248 (29,6%)	404 (48,3%)
d. Se sentiu calma e tranquila?	97 (11,5%)	232 (27,6%)	154 (18,3%)	217 (25,7%)	121 (14,4%)	21 (2,5%)
e. Se sentiu com muita energia?	54 (6,4%)	152 (18,2%)	126 (15,0%)	271 (32,3%)	193 (23,0%)	43 (5,1%)
f. Se sentiu triste?	15 (1,7%)	25 (3,0%)	37 (4,4%)	184 (21,9%)	301 (35,8%)	279 (33,2%)
g. Se sentiu estafada?	30 (3,6%)	84 (10,0%)	133 (15,7%)	339 (40,3%)	204 (24,2%)	52 (6,2%)
h. Se sentiu feliz?	266 (31,5%)	229 (27,1%)	152 (18,1%)	158 (18,7%)	31 (3,7%)	8 (0,9%)
i. Se sentiu cansada?	49 (5,8%)	125 (14,8%)	149 (17,7%)	345 (40,9%)	152 (18,0%)	24 (2,8%)

- Durante as últimas 4 semanas, até que ponto é que a sua saúde física ou problemas emocionais limitaram a sua actividade social (tal como visitar amigos ou familiares próximos)? trabalho normal (tanto o trabalho fora de casa como o trabalho doméstico)?

Sempre	A maior parte do tempo	Algum tempo	Pouco tempo	Nunca
73 (8,9%)	56 (6,8%)	46 (5,6%)	143 (17,4%)	504 (61,3%)

SAÚDE EM GERAL

- Por favor, diga em que medida são verdadeiras ou falsas as seguintes afirmações.
Ponha um círculo para cada linha.

	Absolutamente verdade	Verdade	Não sei	Falso	Absolutamente falso
a. Parece que adoço mais facilmente do que os outros.	13 (1,5%)	45 (5,3%)	203 (24,1%)	358 (42,4%)	225 (26,7%)
b. Sou tão saudável como qualquer outra pessoa.	174 (20,3%)	375 (43,9%)	231 (27,0%)	64 (7,5%)	11 (1,3%)
c. Estou convencida que a minha saúde vai piorar.	5 (0,6%)	16 (1,9%)	279 (33,1%)	337 (40,0%)	206 (24,4%)
d. A minha saúde é excelente	77 (9,2%)	349 (41,5%)	317 (37,7%)	85 (10,1%)	13 (1,5%)

Apêndice E — Distribuição dos valores por escala

Quadro E.1 — Escala de Função Física (FF)

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
0,00	13	1,5	1,5
5,00	5	0,6	2,1
5,56	3	0,4	2,5
6,25	1	0,1	2,6
7,14	2	0,2	2,8
10,00	20	2,3	5,2
11,11	1	0,1	5,3
15,00	8	0,9	6,2
16,67	1	0,1	6,3
20,00	16	1,9	8,2
25,00	12	1,4	9,6
27,78	2	0,2	9,8
28,57	1	0,1	10,0
30,00	26	3,0	13,0
31,25	2	0,2	13,2
33,33	3	0,3	13,6
25,00	29	3,4	17,0
37,50	2	0,2	17,2
38,89	3	0,4	17,6
40,00	45	5,3	22,8
44,44	2	0,2	23,1
45,00	41	4,8	27,9
50,00	54	6,3	34,2
55,00	59	6,9	41,1
55,56	4	0,5	41,6
56,25	1	0,1	41,7
57,14	1	0,1	41,8
60,00	59	6,9	48,7
61,11	1	0,1	48,8
65,50	1	0,1	48,9
65,00	51	6,0	54,9

Quadro E.1 — Escala de Função Física (FF) — cont.

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
66,67	8	1,0	55,9
68,75	2	0,2	56,1
70,00	58	6,8	62,9
75,00	73	8,5	71,4
78,57	1	0,1	71,5
80,00	48	5,6	77,2
81,25	1	0,1	77,3
83,33	1	0,1	77,4
85,00	47	5,5	82,9
88,89	1	0,1	83,0
90,00	44	5,2	88,2
93,75	1	0,1	88,3
94,44	1	0,1	88,4
95,00	53	6,2	94,6
100,00	46	5,4	100,0

854

Número de casos em falta = 76

Quadro E.2 — Escala de Desempenho Físico (DF)

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
0,00	246	29,0	29,0
25,00	186	21,9	50,9
33,33	8	0,9	51,9
50,00	151	17,8	69,7
66,67	4	0,5	70,2
75,00	90	10,6	80,8
100,00	163	19,2	100,0

848

Número de casos em falta = 82

Quadro E.3 — Escala de Dor Corporal (DC)

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
0,00	3	0,3	0,3
10,00	3	0,3	0,7
12,00	3	0,3	1,0
20,00	2	0,2	1,3
21,00	2	0,2	1,5
22,00	28	3,2	4,7
24,00	1	0,1	4,8
30,00	1	0,1	5,0
31,00	28	3,2	5,2
32,00	8	0,9	9,1
41,00	89	10,3	19,4
42,00	14	1,6	21,0
51,00	46	5,3	26,3
52,00	45	5,2	31,5
54,00	1	0,1	31,6
60,00	1	0,1	31,7
61,00	4	0,5	32,2
62,00	173	20,0	52,1
64,00	7	0,8	52,9
72,00	28	3,2	56,2
74,00	128	14,8	70,9
75,00	1	0,1	71,0
80,00	10	1,2	72,2
84,00	61	7,0	79,2
88,00	1	0,1	79,4
100,00	179	20,6	100,0

867

Número de casos em falta = 63

Quadro E.4 — Escala de Saúde em Geral (SG)

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
5,00	1	0,1	0,1
10,00	1	0,1	0,2
20,00	1	0,1	0,4
25,00	6	0,7	1,1
30,00	1	0,1	1,2
31,25	1	0,1	1,3
35,00	9	1,1	2,4
37,00	1	0,1	2,5
37,50	1	0,1	2,6
40,00	28	3,3	5,9
42,00	8	0,9	6,9
43,75	1	0,1	7,0
45,00	32	3,8	10,8
47,00	12	1,4	12,2
50,00	42	5,0	17,2
52,00	39	4,6	21,8
52,50	1	0,1	21,9
55,00	24	2,8	24,7
56,25	2	0,2	25,0
57,00	47	5,6	30,5
60,00	18	2,1	32,7
62,00	83	9,8	42,5
62,50	1	0,1	42,6
65,00	15	1,8	44,4
67,00	103	12,2	56,6
68,75	1	0,1	56,7
70,00	6	0,7	57,4
71,25	6	0,7	58,1
72,00	92	10,9	69,0
75,00	6	0,7	69,7
77,00	59	7,0	76,7

Quadro E.4 — Escala de Saúde em Geral (SG) — cont.

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
78,33	1	0,1	76,8
80,00	6	0,7	77,5
81,25	1	0,1	77,6
82,00	44	5,2	82,8
83,75	1	0,1	83,0
85,00	5	0,6	83,6
87,00	60	7,1	90,7
90,00	11	1,3	92,0
92,00	36	4,3	96,2
95,00	5	0,6	96,8
97,00	14	1,7	98,5
100,00	13	1,5	100,0

845

Número de casos em falta = 85

Quadro E.5 — Escala de Vitalidade (VT)

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
0,00	5	0,6	0,6
5,00	2	0,2	0,8
6,67	1	0,1	0,9
10,00	10	1,2	2,1
15,00	9	1,1	3,2
20,00	28	3,3	6,5
25,00	28	3,3	9,8
30,00	55	6,5	16,2
33,33	7	0,8	17,0
35,00	38	4,5	21,5
40,00	68	8,0	29,5
45,00	73	8,6	38,1
46,67	10	1,2	39,2
50,00	99	11,6	50,9
53,33	7	0,8	51,7
55,00	68	8,0	59,7
60,00	73	8,6	68,3
65,00	65	7,6	75,9
66,67	6	0,7	76,6
70,00	57	6,7	83,3
73,33	4	0,5	83,8
75,00	36	4,2	88,0
80,00	50	5,9	93,9
85,00	30	3,5	97,4
86,67	1	0,1	97,5
90,00	11	1,3	98,8
95,00	7	0,8	99,6
100,00	3	0,4	100,0

851

Número de casos em falta = 79

Quadro E.6 — Escala de Função Social (FS)

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
0,00	3	0,3	0,3
12,50	5	0,6	0,9
25,00	23	2,7	3,6
37,50	28	3,2	6,8
50,00	88	10,2	17,0
62,50	83	9,6	26,7
75,00	202	23,4	50,1
87,50	126	14,6	64,7
100,00	305	35,3	100,0

863

Número de casos em falta = 67

Quadro E.7 — Escala de Desempenho Emocional (DE)

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
0,00	172	21,0	21,0
33,33	191	23,3	44,3
50,00	1	0,1	44,4
66,67	156	19,0	63,5
100,00	299	36,5	100,0

819

Número de casos em falta = 111

Quadro E.8 — Escala de Saúde Mental (SM)

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
0,00	2	0,2	0,2
6,67	1	0,1	0,4
12,00	1	0,1	0,5
16,00	4	0,5	0,9
20,00	2	0,2	1,2
24,00	3	0,4	1,5
28,00	4	0,5	2,0
32,00	10	1,2	3,2
36,00	10	1,2	4,4
40,00	20	2,4	6,8
44,00	23	2,7	9,5
46,67	1	0,1	9,6
48,00	39	4,6	14,2
50,00	2	0,2	14,5
52,00	40	4,7	19,2
53,33	1	0,1	19,3
55,00	1	0,1	19,4
56,00	53	6,3	25,7
60,00	39	4,6	30,3
64,00	54	6,4	36,7
65,00	1	0,1	36,8
68,00	60	7,1	44,0
70,00	2	0,2	44,2
72,00	67	7,9	52,1
73,33	1	0,1	52,3
76,00	60	7,1	59,4
80,00	61	7,2	66,6
84,00	73	8,6	75,2
86,67	2	0,2	75,5
88,00	45	8,9	84,4
90,00	3	0,4	84,7

Quadro E.8 — Escala de Saúde Mental (SM) — cont.

Pontuação Transformada	Frequência	Porcentagem	Porcentagem Acumulados
92,00	62	7,3	92,1
95,00	1	0,1	92,2
96,00	33	3,9	96,1
100,00	33	3,9	100,0

844

Número de casos em falta = 86