

Tradução para português brasileiro e validação de um questionário de avaliação de produtividade

Patrícia Coelho de Soárez,¹ Clarissa Campos Guaragna Kowalski,¹
Marcos Bosi Ferraz¹ e Rozana Mesquita Ciconelli^{1,2}

Como citar Soárez PC, Kowalski CCG, Ferraz MB, Ciconelli RM. Tradução para português brasileiro e validação de um questionário de avaliação de produtividade. Rev Panam Salud Publica. 2007;22(1):21–8.

RESUMO

Objetivo. Traduzir para português brasileiro, adaptar culturalmente e avaliar as propriedades psicométricas, a confiabilidade e a validade do questionário sobre limitações no trabalho (Work Limitations Questionnaire, WLQ).

Métodos. O presente estudo transversal e observacional foi realizado no Hospital São Paulo/Escola Paulista de Medicina da Universidade Federal de São Paulo, Brasil. Os dados de 150 indivíduos empregados no momento do estudo foram obtidos utilizando o WLQ, o SF-36 (questionário genérico de avaliação de qualidade de vida) e o SRQ-20 (usado para triagem de transtornos mentais). O WLQ apresenta 25 itens agrupados em quatro domínios: gerência de tempo, demanda física, demanda mental-interpessoal e demanda de produção. Os questionários foram administrados por entrevista para indivíduos sem escolaridade superior completa e auto-administrados para indivíduos com escolaridade superior. A estatística descritiva foi utilizada para caracterizar a amostra. Para avaliar a confiabilidade (consistência interna, teste-reteste e consistência inter e intra-observador), foram calculados o coeficiente de correlação intraclasse e o alfa de Cronbach. O coeficiente de correlação de Pearson foi utilizado para aferir a validade da construção.

Resultados. A idade média dos entrevistados (64,7% do sexo feminino) foi de $37,6 \pm 9,6$ anos. O tempo médio na função ocupada no momento da entrevista foi de $8,6 \pm 8,3$ anos; 60,7% estavam satisfeitos com o emprego e 94,0% não tinham faltado ao trabalho nas 2 semanas anteriores à coleta. A perda média de produtividade registrada foi de 4,2%. A confiabilidade interobservador foi significativa e alta (entre 0,600 e 0,800) ou muito alta (0,800 a 1,000) em todos os domínios, com exceção de demanda física ($r = 0,497$, confiabilidade moderada). A confiabilidade intra-observador não foi significativa nos domínios gerência de tempo e demanda física. As correlações significativas encontradas para a confiabilidade intra-observador foram moderadas ($r = 0,400$ a $0,600$). A consistência interna foi muito alta (alfa de Cronbach de 0,800 a 1,000). Observou-se correlação significativa entre gerência de tempo, demanda mental-interpessoal, demanda de produção e o índice do WLQ e todos os domínios do SF-36. Também houve correlação significativa entre os domínios gerência de tempo, demanda mental-interpessoal, demanda de produção e resultado positivo (escore ≥ 7) no questionário SRQ-20.

¹ Universidade Federal de São Paulo (UNIFESP), Escola Paulista de Medicina, Centro Paulista de Economia da Saúde. Enviar correspondência para Patrícia Coelho de Soárez no seguinte endereço: Centro Paulista de Economia da Saúde, Escola

Paulista de Medicina, Universidade Federal de São Paulo (UNIFESP), Rua Botucatu 685, Vila Clementino, CEP 04023-062, São Paulo, SP, Brasil. E-mail: patricia.soarez@cpses.org.br

² Universidade Federal de São Paulo (UNIFESP), Escola Paulista de Medicina, Disciplina de Reumatologia.

Conclusão. *A versão do WLQ em português brasileiro é uma medida válida e confiável e pode ser útil para medir o impacto de problemas de saúde sobre a produtividade de trabalhadores brasileiros.*

Palavras-chave

Confiabilidade e validade, eficiência, estudos de validação, saúde do trabalhador, Work Limitations Questionnaire, Brasil.

A atual escassez de recursos no setor saúde e o aumento exponencial nos gastos em saúde resultam no encarecimento das doenças para a sociedade, assunto que tem adquirido grande destaque. O custo da doença é medido em termos econômicos através dos custos diretos e indiretos. Os custos diretos incluem, principalmente, os cuidados médicos, enquanto que os custos indiretos incluem absenteísmo, presenteísmo, perda de produtividade e morte prematura. O interesse crescente em mensurar o efeito dos problemas de saúde sobre a produtividade de trabalhadores tem fomentado o desenvolvimento científico dessa área e a criação de diversos instrumentos que quantifiquem essa relação.

Segundo o modelo conceitual de desempenho humano de O'Donnell, que estabelece vinculação entre saúde, produtividade e lucros, o desempenho dos empregados no trabalho é melhor quando eles estão física e emocionalmente capacitados a trabalhar (1). Entretanto, medidas objetivas e técnicas de mensuração de produtividade são difíceis de obter e não estão disponíveis uniformemente em todas as indústrias. Dessa forma, medidas subjetivas ou auto-informadas se tornaram um caminho promissor para reunir informações a respeito da percepção dos trabalhadores quanto a como a própria saúde influencia sua capacidade de desempenhar tarefas.

As medidas auto-informadas, tais como absenteísmo e presenteísmo, podem ser obtidas através de questionários (2). Atualmente, estão disponíveis os seguintes instrumentos genéricos para estimar a produtividade: *Work Productivity and Activity Impairment Questionnaire (General Health) (WPAI-GH)* (questionário sobre produtividade no trabalho e comprometimento da atividade); *Work Limitations Ques-*

tionnaire (WLQ) (questionário sobre limitações no trabalho); *Health and Work Performance Questionnaire (HPQ)* (questionário sobre saúde e desempenho no trabalho); *Health and Work Questionnaire (HWQ)* (questionário sobre saúde e trabalho); *Endicott Work Productivity Scale (EWPS)* (escala de Endicott sobre produtividade no trabalho); e *Health and Labor Questionnaire (HLQ)* (questionário laboral e de saúde).

A maioria desses questionários foi desenvolvida na língua inglesa, e, para que possam ser utilizados no Brasil, devem passar por um processo de tradução, adaptação cultural e validação de suas propriedades psicométricas. Dos instrumentos disponíveis, o WLQ é o que teve suas propriedades psicométricas mais extensivamente testadas para avaliar o impacto da saúde geral e de condições específicas na produtividade dos indivíduos (3). Dessa forma, a validação do WLQ é relevante tanto para os aspectos científicos do desenvolvimento dessa área de pesquisa, quanto para as implicações sociais e econômicas que permeiam essa área do conhecimento.

O objetivo deste estudo foi formular a versão do WLQ em português brasileiro (PB) e testar suas propriedades de medida para que o mesmo possa ser utilizado como instrumento genérico para estimar o impacto de doenças sobre a produtividade de trabalhadores brasileiros. (O WLQ só pode ser utilizado com permissão. Para obter essa permissão, entre em contato com a Dra. Debra Lerner pelo e-mail wlg@tufts-nemc.org)

MATERIAIS E MÉTODOS

Esta pesquisa é um estudo observacional de corte transversal, desenvolvida no Hospital São Paulo da Escola

Paulista de Medicina, Universidade Federal de São Paulo (UNIFESP), entre 2005 e 2006. O estudo foi aprovado pelo Comitê de Ética e Pesquisa da UNIFESP e as informações foram coletadas após obtenção do consentimento informado dos indivíduos.

Participaram da pesquisa 150 indivíduos com idade maior ou igual a 18 anos, funcionários ou estudantes de pós-graduação do Hospital São Paulo, que estivessem trabalhando no mês anterior à realização da pesquisa e que não apresentassem déficit cognitivo para compreensão das questões. Foram excluídos da amostra pacientes em tratamento de doenças crônicas, tais como hipertensão, diabetes, insuficiência cardíaca congestiva, doença pulmonar obstrutiva crônica e doenças reumáticas.

Descrição do instrumento

O WLQ foi desenvolvido em um período de 4 anos, sendo disponibilizado para uso em 1999 após vários estudos terem fornecido evidência de sua validade, confiabilidade e precisão. Já foi validado em diferentes populações de pacientes e empregados e para várias condições de saúde. O WLQ é um questionário auto-administrável, com 25 itens, que mede o grau de interferência dos problemas de saúde na capacidade de desenvolver tarefas no trabalho e o impacto dessa interferência na produtividade. Esse questionário foi desenvolvido para avaliar indivíduos que estejam efetivamente empregados (4). Uma característica particular do WLQ é que ele pede ao entrevistado que avalie o seu próprio grau de dificuldade para realizar tarefas específicas exigidas no trabalho.

Os 25 itens são agrupados em quatro domínios de limitação de trabalho:

gerência de tempo, demanda física, demanda mental-interpessoal e demanda de produção. Em conjunto, eles abrangem o caráter multi-dimensional das funções desenvolvidas no trabalho. Essas escalas podem elucidar em quais domínios o indivíduo tem suas funções limitadas.

A escala de gerência de tempo é apresentada na questão 1, que contém cinco itens abordando dificuldades em cumprir horários e tarefas no tempo previsto. Na questão 2, a escala de demanda física é coberta por seis itens que avaliam a capacidade de realizar tarefas que exijam força corporal, resistência, movimento, coordenação e flexibilidade. A escala de demanda mental-interpessoal engloba as questões 3 e 4, somando nove itens. Seis itens avaliam a dificuldade em realizar tarefas cognitivas no trabalho e três itens abordam a dificuldade em interagir com pessoas no trabalho. A escala de demanda de produção compreende a questão 5, com cinco itens que se referem a decréscimos na capacidade da pessoa de concluir, em tempo hábil, a quantidade e qualidade necessárias de trabalho.

Cada escala apresenta um escore que varia de 0 (sem limitação) a 100 (todo o tempo com limitação). O escore indica a porcentagem de tempo, nas 2 últimas semanas, pela qual o indivíduo esteve limitado para realizar suas tarefas no trabalho. Um escore de 20 na escala de demanda de produção indica que o indivíduo esteve limitado em 20% do tempo dedicado a desenvolver esse tipo de tarefa.

Após o cálculo de todas as escalas, é possível definir o índice WLQ, que é interpretado com auxílio de uma tabela elaborada pela autora do instrumento. Por exemplo, um índice WLQ igual a 5 indica que o indivíduo apresenta uma redução na sua produtividade de 4,9% quando comparado a um indivíduo saudável; isso implica que esse indivíduo terá de aumentar em 5,1% o número de horas trabalhadas para compensar a sua produtividade perdida.

Embora abrangente, o WLQ é curto e de fácil administração. Leva em média de 5 a 10 minutos para ser respondido quando auto-administrado.

Tradução e adaptação cultural

A tradução e adaptação cultural do instrumento foram feitas com base em diretrizes existentes (5, 6), consistindo de tradução, retrotradução, revisão por painel de especialistas e adaptação cultural.

Os itens da versão do WLQ foram traduzidos inicialmente por dois profissionais de saúde, independentes, brasileiros, com grande conhecimento da língua inglesa, e que estavam cientes dos objetivos da pesquisa. Foi enfatizada a importância da tradução conceitual em detrimento de uma tradução literal. Foram, assim, obtidas as versões de número 1 e 2 em PB.

As traduções 1 e 2 foram vertidas para o inglês por dois professores de inglês, independentes, norte-americanos, que não participaram da etapa anterior e não tinham conhecimento dos objetivos nem dos conceitos apresentados no questionário. Em seguida, foi feita a comparação do instrumento original com as duas novas versões em inglês. Um painel de especialistas composto por cinco profissionais de saúde (duas médicas, duas dentistas e uma enfermeira, todas com experiência em tradução de questionários) documentou e analisou as discrepâncias encontradas. Alguns tempos verbais e sentenças em português foram reescritos até se obter um consenso, gerando, assim, a versão 3 em PB.

A versão 3 em PB do questionário foi aplicada a um grupo de 10 indivíduos para avaliar o grau de compreensão das questões, sendo selecionadas as que não apresentavam boa compreensão. O painel de especialistas avaliou esses itens e os substituiu por outros com o mesmo conceito, porém de mais fácil compreensão, utilizando inclusive sugestões dos próprios respondentes, procurando não alterar a estrutura e propriedades de avaliação dessas questões. Essa versão 4 foi então novamente aplicada a um outro grupo de 10 indivíduos, selecionados aleatoriamente no mesmo local, sendo sua equivalência cultural testada novamente até que todos os itens fossem compreendidos por 90% dos entrevistados.

Estudo psicométrico

Após finalização do processo de tradução e adaptação cultural do instrumento, foram avaliadas a confiabilidade e a validade do mesmo. Participaram desta fase 150 indivíduos.

Primeiramente, a confiabilidade da versão do WLQ para PB foi avaliada através de três entrevistas. Um grupo de 30 indivíduos foi selecionado e avaliado por dois entrevistadores (número 1 e número 2): o entrevistador 1 fez a primeira avaliação. Após um intervalo de 30 a 60 minutos, o entrevistador 2 fez a segunda avaliação. Num período máximo de 7 a 14 dias após a segunda avaliação, uma terceira avaliação foi feita pelo entrevistador 1.

A avaliação da validade de construção foi testada em todos os 150 indivíduos. O WLQ foi avaliado através de sua correlação com parâmetros sociodemográficos e parâmetros relacionados ao trabalho do indivíduo, como tempo na atual função, satisfação no trabalho, número de dias de falta ao trabalho e a nota dada pelo entrevistado à sua própria produtividade. Esses parâmetros foram colhidos através de uma entrevista inicial com o participante.

Além dos parâmetros acima mencionados, o instrumento foi comparado com o questionário genérico de avaliação de qualidade de vida SF-36 (*Medical Outcomes Study 36-Item Short-Form Health Survey*), que avalia os domínios de percepção geral de saúde, capacidade funcional, limitação por aspectos físicos e emocionais, dor, vitalidade, aspectos sociais e saúde mental. O escore para cada domínio varia de 0 a 100, sendo 0 o pior estado de saúde e 100 o melhor. Este instrumento já foi traduzido para o nosso idioma e suas propriedades psicométricas já foram testadas (7, 8).

O instrumento também foi comparado ao *Self-Reporting Questionnaire* (SRQ-20), um instrumento de triagem de transtornos mentais para serviços de atendimento primário, recomendado pela Organização Mundial da Saúde (OMS), já validado para a população brasileira (9). Esse questionário é composto de 20 questões. Se o indivi-

duo responder positivamente a sete ou mais questões está propenso a um estado depressivo. O escore varia de 0 a 20, sendo 0 o melhor estado de saúde e 20 o pior.

Aplicação dos questionários

Os questionários foram aplicados individualmente. Os funcionários e estudantes foram abordados em seu local de atuação por duas estudantes de pós-graduação da UNIFESP que lhes explicavam os objetivos e o conteúdo do questionário e pediam o seu consentimento e colaboração para participar da pesquisa. Quando a pessoa preenchia os critérios de inclusão do estudo e tinha disponibilidade de tempo para responder imediatamente aos questionários, a aplicação era iniciada.

Para os indivíduos que tivessem escolaridade superior, o formulário de dados sociodemográficos e os questionários eram aplicados na forma de auto-administração. Para os indivíduos com escolaridade fundamental ou média, as pesquisadoras aplicavam inicialmente o formulário de dados sociodemográficos. Em seguida, sempre na forma de entrevista, aplicavam o WLQ, seguido pelo SRQ-20 e pelo SF-36. As pesquisadoras pediam, então, que os entrevistados aguardassem de 30 a 60 minutos para a segunda administração do WLQ, também na forma de entrevista. A terceira administração do WLQ era realizada num período de 7 a 14 dias, também na forma de entrevista. Não foi oferecida nenhuma forma de incentivo ou remuneração aos participantes do estudo.

Análise estatística

O programa utilizado para análise dos dados foi o *Intercooled Stata* versão 8.0. Foi realizada uma análise estatística descritiva para caracterização sociodemográfica da população de estudo. Os coeficientes de correlação intraclass e alfa de Cronbach foram utilizados para avaliação da confiabilidade e consistência interna do instrumento, e o coeficiente de correlação de

TABELA 1. Características sociodemográficas e de produtividade da população estudada, São Paulo (SP), Brasil, 2006^a

Variável	No. (%) ^b
Sexo	
Masculino	53 (35,3)
Feminino	97 (64,7)
Idade (anos) (média ± desvio padrão)	37,6 ± 9,6
Estado civil	
Casado	83 (55,3)
Solteiro	49 (32,7)
Separado	7 (4,7)
Divorciado	6 (4,0)
Viúvo	2 (1,3)
Outros	3 (2,0)
Grau de escolaridade	
Ensino fundamental incompleto	15 (10,0)
Ensino fundamental completo	31 (20,7)
Ensino médio incompleto	–
Ensino médio completo	–
Ensino superior completo	104 (69,3)
Tempo de trabalho na função atual (anos) (média ± desvio padrão)	8,6 ± 8,3
Grau de satisfação com o trabalho	
Muito satisfeito	27 (18,0)
Satisfeito	91 (60,7)
Nem satisfeito nem insatisfeito	23 (15,3)
Insatisfeito	9 (6,0)
Muito insatisfeito	–
Falta ao trabalho nas 2 últimas semanas	
Não	141 (94,0)
Sim	9 (6,0)
Dias de falta (média ± desvio padrão)	0,1 ± 0,4
Nota de produtividade ^c (média ± desvio padrão)	76,1 ± 16,2

^a Estudo para validação da versão em português brasileiro do *Work Limitations Questionnaire* (WLQ).

^b Ou média ± desvio padrão, conforme indicado. O total de respondentes foi 150.

^c Nota de produtividade = nota (0 a 100) dada pelo indivíduo à sua produtividade das 2 últimas semanas.

Pearson, para avaliação da validade. Assumimos para as correlações um nível de significância estatística de 0,05%.

RESULTADOS

Tradução e adaptação cultural

Vinte indivíduos participaram da fase de avaliação da adaptação cultural. Nenhuma questão foi considerada não aplicável. As questões 1, 3, 4 e 5 e seus subitens foram compreendidos por menos de 90% dos entrevistados. Foram modificadas e reaplicadas até que todos os itens apresentassem um nível de compreensão acima de 90%. A questão 2 também foi compreendida por menos de 90% dos entrevistados.

Nessa questão, a dificuldade foi resultado da construção do enunciado, que é diferente das demais questões.

Estudo psicométrico

A tabela 1 apresenta as características sociodemográficas e de produtividade dos 150 indivíduos incluídos na avaliação de confiabilidade e validade da versão em PB do WLQ. A tabela 2 apresenta os valores obtidos para cada domínio da versão em PB do WLQ. Os valores médios para cada domínio do questionário se situaram geralmente entre 12 e 15, com exceção do segundo domínio, demanda física, que avalia a capacidade de realizar tarefas que exijam força corporal, resistência, movimento, coordenação e flexibilidade.

TABELA 2. Escores obtidos para cada domínio do questionário WLQ e do questionário SF-36, São Paulo (SP), Brasil, 2006^a

Domínio	Média	Desvio padrão	Mínimo	Máximo
WLQ				
Gerência de tempo	12,7	16,9	0,0	75,0
Demanda física	25,2	35,3	0,0	100,0
Demanda mental-interpessoal	14,2	15,0	0,0	61,0
Demanda de produção	12,6	17,5	0,0	85,0
Índice	4,2	4,0	0,0	16,9
SF-36				
Capacidade funcional	86,9	16,2	20,0	100,0
Aspectos físicos	86,2	22,9	0,0	100,0
Dor	72,2	20,8	0,0	100,0
Estado geral da saúde	81,8	16,3	5,0	100,0
Vitalidade	64,8	20,7	0,0	100,0
Aspectos sociais	77,5	23,9	0,0	100,0
Aspectos emocionais	80,4	27,7	0,0	100,0
Saúde mental	71,6	21,0	4,0	100,0

^a Estudo para validação da versão em português brasileiro do *Work Limitations Questionnaire* (WLQ). No WLQ, cada domínio recebe do entrevistado um escore que varia de 0 (sem limitação) a 100 (todo o tempo com limitação). No SF-36, o escore para cada domínio varia de 0 a 100, sendo 0 o pior estado de saúde e 100 o melhor.

Nesse domínio, o valor médio obtido foi de 25,2, sendo o desvio padrão de 35,3. Isso significa que, nas 2 semanas anteriores à aplicação do questionário, o tempo desses indivíduos para realizar tarefas físicas no trabalho foi limitado em 25%. O escore do índice representa o percentual de produtividade perdida em relação a pessoas saudáveis, ou seja: nesta amostra, as pessoas apresentaram, em média, um decréscimo de 4,2% na produtividade quando comparadas a pessoas completamente sadias. A tabela 2 mostra ainda os valores obtidos para cada do-

mínio do SF-36. Os valores médios para cada domínio do questionário se situaram geralmente entre 70 e 90, com exceção do domínio vitalidade, que apresentou escore mais baixo, de 64,8.

Com relação aos valores obtidos no questionário SRQ-20, 41 (27,3%) indivíduos foram identificados como portadores de transtornos mentais. Em média os indivíduos responderam positivamente a 4,3 questões, com desvio padrão de 3,8 questões, variando de zero a 17 questões.

Na tabela 3 são apresentados os resultados relativos à confiabilidade

intra e interobservador e à consistência interna de cada domínio da versão em PB do WLQ, determinados utilizando-se os coeficientes de correlação intraclass e alfa de Cronbach. A confiabilidade interobservador foi considerada estatisticamente significativa para todos os domínios do WLQ e foi alta ($r = 0,600$ a $0,800$) ou muito alta ($r = 0,800$ a $1,000$) em todos os casos, com exceção do domínio demanda física ($r = 0,497$), para o qual a confiabilidade interobservador foi moderada. Por outro lado, a confiabilidade intra-observador não foi considerada estatisticamente significativa em dois domínios (gerência de tempo e demanda física). As correlações estatisticamente significativas encontradas para a confiabilidade intra-observador foram consideradas moderadas ($r = 0,400$ a $0,600$).

Com relação à consistência interna, as correlações obtidas nas três aplicações foram positivas e estatisticamente significativas. A correlação foi considerada muito alta ($r = 0,800$ a $1,000$) conforme o alfa de Cronbach: $r = 0,808$ ($P < 0,05$) para a primeira entrevista; $r = 0,904$ ($P < 0,01$) para a segunda entrevista; e $r = 0,852$ ($P < 0,001$) para a terceira entrevista.

A tabela 4 apresenta os coeficientes de correlação de Pearson (10) entre cada domínio da versão do WLQ em PB e as características sociodemográficas e de produtividade. Foram observados coeficientes de correlação positivos e estatisticamente significativos para a correlação entre grau de satisfação, dias de falta, nota de produtividade e os domínios gerência de tempo, demanda mental-interpessoal e demanda de produção. Houve correlação negativa e estatisticamente significativa entre idade e o domínio demanda mental-interpessoal e entre nota de produtividade e os domínios gerência de tempo, demanda mental-interpessoal e demanda de produção. Observou-se também uma correlação positiva e estatisticamente significativa entre o índice do WLQ e a variável grau de satisfação, e uma correlação negativa e estatisticamente significativa entre o índice do WLQ e a variável nota de produtividade.

TABELA 3. Confiabilidade intra e interobservador de cada domínio da versão em português brasileiro do WLQ avaliada pelo coeficiente de correlação intraclass, São Paulo (SP), Brasil, 2006^a

Domínio	Coeficiente de correlação intraclass	
	Intra-observador	Interobservador
Gerência de tempo	0,137	0,767 ^b
Demanda física	0,389	0,497 ^c
Demanda mental-interpessoal	0,551 ^c	0,815 ^b
Demanda de produção	0,522 ^d	0,648 ^b
Índice	0,408	0,815 ^b

^a Estudo para validação da versão em português brasileiro do *Work Limitations Questionnaire* (WLQ). Para esta análise, 30 indivíduos foram entrevistados três vezes por dois pesquisadores com intervalo de 30 a 60 minutos entre a primeira e a segunda entrevistas e de 7 a 14 dias entre a segunda e a terceira entrevistas. Classificação dos coeficientes: confiabilidade moderada = 0,400 a 0,600; alta = 0,600 a 0,800; muito alta = 0,800 a 1,000.

^b $P < 0,001$.

^c $P < 0,01$.

^d $P < 0,05$.

TABELA 4. Coeficientes de correlação de Pearson entre as características sociodemográficas e de produtividade e os domínios do questionário WLQ, São Paulo (SP), Brasil, 2006^a

Característica	Gerência de tempo	Demanda física	Demanda mental-interpessoal	Demanda de produção	Índice
Idade					
Correlação de Pearson	0,054 ^b	0,036 ^b	-0,159 ^b	-0,047 ^b	-0,054 ^b
P	0,508	0,655	0,051	0,568	0,506
Grau de escolaridade					
Correlação de Pearson	0,066 ^b	0,175 ^b	0,148 ^b	0,048 ^b	0,141 ^b
P	0,417	0,031	0,070	0,555	0,083
Tempo de trabalho					
Correlação de Pearson	0,081 ^b	-0,051 ^b	-0,090 ^b	-0,015 ^b	-0,037 ^b
P	0,323	0,528	0,271	0,853	0,649
Grau de satisfação					
Correlação de Pearson	0,301 ^c	0,066 ^b	0,337 ^c	0,277 ^c	0,326 ^c
P	0,000	0,416	0,000	0,000	0,000
Dias de falta					
Correlação de Pearson	0,210 ^c	-0,010 ^b	-0,022 ^b	-0,009 ^b	0,028 ^b
P	0,010	0,898	0,784	0,907	0,732
Nota de produtividade ^d					
Correlação de Pearson	-0,258 ^c	-0,082 ^b	-0,337 ^c	-0,424 ^e	-0,387 ^c
P	0,001	0,315	0,000	0,000	0,000

^a Estudo para validação da versão em português brasileiro do *Work Limitations Questionnaire* (WLQ).

^b Correlação muito baixa ($r = 0,000$ a $0,200$).

^c Correlação baixa ($r = 0,200$ a $0,400$).

^d Nota de produtividade = nota (0 a 100) dada pelo indivíduo à sua produtividade nas 2 últimas semanas.

^e Correlação moderada ($r = 0,400$ a $0,600$).

A tabela 5 apresenta os coeficientes de correlação de Pearson para os quatro domínios do WLQ em PB e os questionários SF-36 e SRQ-20. Coeficientes de correlação negativos e estatisticamente significativos foram encontrados quando se avaliou a correlação entre os domínios gerência de tempo, demanda mental-interpessoal e demanda de produção e todos os domínios do SF-36. Observou-se também uma correlação negativa e estatisticamente significativa entre o índice do WLQ e todos os domínios do SF-36. A correlação mais expressiva ocorreu entre o domínio gerência de tempo e o domínio vitalidade do SF-36. Uma outra correlação bastante expressiva ocorreu entre o domínio demanda mental-interpessoal e o domínio aspectos sociais do SF-36.

Os escores do WLQ variam de um valor menor para maior, de acordo com o aumento de perda de produtividade. Ao contrário, o escore do SF-36 varia de um valor menor para um valor maior, de acordo com a melhora do estado geral do paciente. Isso explica o fato de as comparações entre os

questionários apresentarem coeficientes de correlação negativos.

Como mostra a tabela 5, coeficientes de correlação positivos e estatisticamente significativos foram encontrados quando se avaliou a correlação entre os domínios gerência de tempo, demanda mental-interpessoal, demanda de produção e o resultado positivo (escore ≥ 7) do questionário SRQ-20. A correlação mais expressiva ocorreu com o domínio demanda mental-interpessoal.

DISCUSSÃO

Medir a perda de produtividade no ambiente de trabalho em virtude de problemas de saúde ainda é uma tarefa difícil. Apesar de o assunto ter recebido muita atenção nos últimos anos, e de já existirem vários instrumentos genéricos e específicos para realizar tal tarefa, tais instrumentos, na sua grande maioria, estão disponíveis apenas na língua inglesa. Recentemente, o questionário WPAI foi validado para o português brasileiro (11).

Embora o WLQ tenha sido validado em populações de doentes (12–18), este trabalho testou o instrumento em uma população “saudável”, e procurou avaliar a sua capacidade de captar pequenas alterações no estado de saúde desses indivíduos. Incluímos na amostra indivíduos com escolaridade alta e baixa, e os nossos resultados foram satisfatórios, mesmo com os indivíduos de baixa escolaridade.

Os escores obtidos nos domínios do WLQ (12,6 a 25,2) demonstraram que, apesar de ser considerada saudável, a população apresentou um certo grau de comprometimento na sua produtividade. Esse tipo de constatação é de fundamental importância para empresas ou empregadores que queiram monitorar mais de perto a saúde dos seus empregados, ou de governos que queiram avaliar o impacto de alguns programas na saúde da população como um todo.

Na avaliação da confiabilidade, observamos escores satisfatórios de concordância entre as três aplicações e escores excelentes para a consistência interna do instrumento. A baixa corre-

TABELA 5. Coeficientes de correlação de Pearson entre os domínios dos questionários SF-36 e SRQ-20 e os domínios do questionário WLQ, São Paulo (SP), Brasil, 2006^a

Questionário	WLQ				Índice
	Gerência de tempo	Demanda física	Demanda mental-interpessoal	Demanda de produção	
SF-36					
Capacidade funcional					
Correlação de Pearson	-0,381 ^b	-0,143 ^c	-0,262 ^b	-0,352 ^b	-0,374 ^b
P	0,000	0,080	0,001	0,000	0,000
Aspectos físicos					
Correlação de Pearson	-0,464 ^d	-0,174 ^c	-0,449 ^d	-0,392 ^b	-0,478 ^b
P	0,000	0,033	0,000	0,000	0,000
Dor					
Correlação de Pearson	-0,366 ^b	-0,104 ^c	-0,195 ^c	-0,251 ^b	-0,292 ^b
P	0,000	0,205	0,016	0,002	0,000
Estado geral de saúde					
Correlação de Pearson	-0,258 ^b	-0,149 ^c	-0,177 ^c	-0,165 ^c	-0,233 ^b
P	0,001	0,068	0,030	0,043	0,004
Vitalidade					
Correlação de Pearson	-0,599 ^d	-0,153 ^c	-0,519 ^d	-0,456 ^d	-0,563 ^d
P	0,000	0,061	0,000	0,000	0,000
Aspectos sociais					
Correlação de Pearson	-0,533 ^d	-0,178 ^c	-0,531 ^d	-0,416 ^d	-0,541 ^d
P	0,000	0,029	0,000	0,000	0,000
Aspectos emocionais					
Correlação de Pearson	-0,417 ^d	-0,176 ^c	-0,464 ^d	-0,393 ^b	-0,483 ^d
P	0,000	0,031	0,000	0,000	0,000
Saúde mental					
Correlação de Pearson	-0,547 ^d	-0,079 ^c	-0,552 ^d	-0,438 ^d	-0,533 ^d
P	0,000	0,338	0,000	0,000	0,000
SRQ-20					
Escore ≥ 7					
Correlação de Pearson	0,347 ^b	-0,075 ^c	0,418 ^d	0,339 ^b	0,348 ^b
P	0,000	0,361	0,000	0,000	0,000
Número de questões positivas					
Correlação de Pearson	0,529 ^d	0,000 ^c	0,592 ^d	0,444 ^d	0,520 ^d
P	0,000	0,999	0,000	0,000	0,000

^a Estudo para validação da versão em português brasileiro do *Work Limitations Questionnaire* (WLQ).

^b Correlação baixa ($r = 0,200$ a $0,400$).

^c Correlação muito baixa ($r = 0,000$ a $0,200$).

^d Correlação moderada ($r = 0,400$ a $0,600$).

lação encontrada para a reprodutibilidade intra-observador do domínio gerência de tempo se deveu provavelmente aos resultados de um único indivíduo, que teve escores muito diferentes entre a primeira e a terceira entrevista. Como esse intervalo foi de aproximadamente 14 dias, pode ter ocorrido uma mudança no estado de saúde desse entrevistado. Por outro lado a consistência interna avaliada pelo alfa de Cronbach apresentou correlações excelentes (0,808 a 0,904), semelhantes aos resultados ($\geq 0,840$) encontrados em outro estudo de validação (13).

Devido à ausência de questionários em PB para estimar a perda de produtividade, para avaliar a validade de construção do WLQ, tivemos de compará-lo a algumas características sociodemográficas, a um instrumento genérico de qualidade de vida, o SF-36, e ao SRQ-20. A avaliação da perda de produtividade permite incorporar os custos indiretos com maior precisão à estimativa do custo de uma doença. Em certas situações, esses custos indiretos superam inclusive os custos diretos. Por isso, é importante desenvolver e validar outros instrumentos que possam avaliar o impacto

de determinadas doenças na perda de produtividade dos trabalhadores, para que seja possível comparar diversos instrumentos de produtividade entre si e confirmar a validade dos resultados.

Em conclusão, a versão do WLQ traduzida e adaptada culturalmente para o português do Brasil é um instrumento prático, de fácil aplicação, que apresenta boa confiabilidade e validade, podendo ser utilizada na nossa população. São necessários estudos adicionais para testar a sensibilidade do instrumento em diversas doenças crônicas e tipos de ocupação.

REFERÊNCIAS

- O'Donnel MP. Health and productivity management: the concept, impact, and opportunity. Commentary to Goetzel and Ozminowski. *Am J Health Promot.* 2000;14(4):215–17.
- Loeppke R, Hymel PA, Lofland JH, Pizzi LT, Konicki DL, Anstadt GW, et al. Health-related workplace productivity measurement: general and migraine-specific recommendations from the ACOEM Expert Panel. *J Occup Environ Med.* 2003;45(4):349–59.
- Prasad M, Wahlqvist P, Shikier R, Shih YC. A review of self-report instruments measuring health-related work productivity: a patient-reported outcomes perspective. *Pharmacoeconomics.* 2004;22(4):225–44.
- Lerner D, Amick BC, Rogers WH, Malspeis S, Bungay K, Cynn D. The Work Limitations Questionnaire. *Med Care.* 2001;39(1):72–85.
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol.* 1993;46(12):1417–32.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* 2000;25(24):3186–91.
- Ciconelli RM. Tradução para o português e validação do questionário genérico de avaliação de qualidade de vida Medical Outcomes Study 36-Item Short-form Health Survey (SF-36) [tese]. São Paulo: Universidade Federal de São Paulo; 1997.
- Ciconelli RM, Ferraz MB, Santos W, Meinão I, Quaresma MR. Tradução para a língua portuguesa e validação do questionário genérico de avaliação de qualidade de vida SF-36 (Brasil SF-36). *Rev Bras Reumatol.* 1999;39(3):143–50.
- Mari JJ, Williams P. A validity study of a psychiatric screening questionnaire (SRQ-20) in primary care in the city of Sao Paulo. *Br J Psychiatry.* 1986;148(jan):23–6.
- Bisqueria R, Sarriera JC, Martinez F. Introdução à estatística: enfoque informático com o pacote SPSS. Porto Alegre: Artmed; 2004
- Ciconelli RM, Soárez PC, Kowalski CCG, Ferraz MB. The Brazilian Portuguese version of Work Productivity and Activity Impairment: General Health (WPAI-GH) Questionnaire. *Sao Paulo Med J.* 2006;124(6):325–32.
- Lerner D, Reed JI, Massarotti E, Wester LM, Burke TA. The work Limitations Questionnaire's validity and reliability among patients with osteoarthritis. *J Clin Epidemiol.* 2002;55(2):197–208.
- Lerner D, Amick BC, Lee JC, Rooney T, Rogers WH, Chang H, et al. Relationship of employee-reported work limitations to work productivity. *Med Care.* 2003;41(5):649–59.
- Walker N, Michaud K, Wolfe F. Work limitations among working persons with rheumatoid arthritis: results, reliability, and validity of the work limitations questionnaire in 836 patients. *J Rheumatol.* 2005;32(6):1006–12.
- Beaton DE, Kennedy CA. Beyond return to work: testing a measure of at-work disability in workers with musculoskeletal pain. *Qual Life Res.* 2005;14(8):1869–79.
- Wolfe F, Michaud K, Choi HK, Williams R. Household income and earnings losses among 6,396 persons with rheumatoid arthritis. *J Rheumatol.* 2005;32(10):1875–83.
- Burton WN, Chen CY, Conti DJ, Schultz AB, Edington DW. The association between health risk change and presenteeism change. *J Occup Environ Med.* 2006;48(3):252–63.
- Burton WN, Chen CY, Schultz AB, Conti DJ, Pransky G, Edington DW. Worker productivity loss associated with arthritis. *Dis Manag.* 2006;9(3):131–43.

Manuscrito recebido em 20 de novembro de 2006. Aceito em versão revisada em 24 de maio de 2007.

ABSTRACT

Translation into Brazilian Portuguese and validation of the Work Limitations Questionnaire

Objectives. To translate into Brazilian Portuguese, cross-culturally adapt, and evaluate the psychometric properties, reliability, and validity of the Work Limitations Questionnaire (WLQ).

Methods. This cross-sectional observational study was performed in 2005 and 2006 at the Hospital São Paulo/Escola Paulista de Medicina of the Federal University of São Paulo, Brazil. Data from 150 individuals who were employed at the time of the study were obtained using the WLQ, the SF-36 (a generic quality of life questionnaire), and the SRQ-20 (used to screen for mental disorders). The WLQ has 25 items, which are divided into four domains: time management, physical demands, mental-interpersonal demands, and output demands. The questionnaires were administered as interviews to individuals without a university education, and were self-administered to individuals with a university degree. Descriptive statistics were used to characterize the sample. The intraclass correlation coefficient and Cronbach's alpha were used to assess reliability (internal consistency, test-retest, interobserver, and intraobserver agreement). The Pearson's correlation coefficient was used to assess construct validity.

Results. The mean age of the subjects was 37.6 years (standard deviation, ± 9.6 years), and 64.7% of the subjects were female. The mean number of years at the job held at the time of the interview was 8.6 (± 8.3 years); 60.7% of the subjects were satisfied with their job, and 94.0% had not missed work in the preceding two weeks. In comparison to completely healthy persons, the mean productivity loss reported by the study subjects was 4.2%. Interobserver agreement was significant and high (0.600 to 0.800) or very high (0.800 to 1.000) in all domains, except physical demands ($r = 0.497$, moderate agreement). Intraobserver agreement was not significant for the time management and physical demands domains. The correlations associated with intraobserver agreement were moderately significant (0.400 to 0.600). Internal consistency was very high (Cronbach's alpha = 0.800 to 1.000). There was a significant correlation between all the SF-36 questionnaire domains and the WLQ domains of time management, mental-interpersonal demands, and output demands as well as the WLQ index score. A positive correlation was also found between the WLQ domains of time management, mental-interpersonal demands, and output demands and a positive result (score ≥ 7) on the SRQ-20.

Conclusion. The Brazilian Portuguese version of the WLQ is a reliable and valid scale to assess the impact of health problems on the productivity of Brazilian workers.

Key words

Occupational health, questionnaires, validation studies, reproducibility of results, efficiency, Brazil.