

Confiabilidade teste-reteste do Índice de Capacidade para o Trabalho em trabalhadores metalúrgicos do Sul do Brasil*

Test-retest reliability of the Work Ability Index (WAI) in metallurgy workers from Southern Brazil

Alexandra Renosto^{I,II}

Patrícia Biz^{III}

Élida Azevedo Hennington^{IV}

Marcos Pascoal Pattussi^I

^I Programa de Pós-graduação em Saúde Coletiva da Universidade do Vale do Rio dos Sinos, São Leopoldo, RS, Brasil.

^{II} Faculdade da Serra Gaúcha, Caxias do Sul, RS.

^{III} Universidade de Caxias do Sul, Caxias do Sul, RS.

^{IV} Fundação Oswaldo Cruz, Rio de Janeiro, RJ.

*Parte da dissertação de mestrado apresentada por Alexandra Renosto no Programa de Pós-graduação em Saúde Coletiva da Universidade do Vale do Rio dos Sinos (UNISINOS), São Leopoldo, RS, 2006.

Agradecimentos: Aos funcionários e suas respectivas empresas, Metalbus e Manumetal de Caxias do Sul, pela disponibilidade e colaboração na realização da pesquisa; ao acadêmico Patriqui Moura Leite e ao professor Marcelo Marcon da Faculdade da Serra Gaúcha, pelo apoio na coleta de dados.

Colaboradores: AR originou o estudo, planejou, coletou e analisou os dados e escreveu o manuscrito. PB e EAH contribuíram na conceituação teórica e na discussão. MPP orientou o estudo e a análise dos dados. Todos os autores revisaram o manuscrito e auxiliaram na interpretação dos dados.

Correspondência: Marcos Pascoal Pattussi. PPG Saúde Coletiva, Universidade do Vale do Rio dos Sinos, Avenida Unisinos, 950, Caixa Postal 275 - São Leopoldo, RS CEP 93022-000. E-mail: mppattussi@unisinos.br

Resumo

Introdução: O Índice de Capacidade para o Trabalho (ICT) tem sido utilizado para mensurar a capacidade física, mental e social dos trabalhadores. Porém, não há relato de estudos que avaliem o grau de estabilidade do índice na língua portuguesa. O objetivo deste estudo foi testar a confiabilidade teste-reteste do ICT. **Métodos:** A população de estudo foi composta por trabalhadores de duas empresas (N= 153) do setor metal-mecânico do município de Caxias do Sul, RS. Os trabalhadores responderam questionários autoadministrados por duas vezes com intervalo de 4 semanas entre eles. A análise dos dados incluiu o coeficiente de correlação intra-classe (CCI), o t-teste pareado e o coeficiente Kappa. **Resultados:** A confiabilidade teste-reteste (ICT contínuo) apresentou um CCI igual a 0,84 (IC_{95%} 0,78 a 0,88, P<0,001) e o teste t-pareado não apresentou diferenças significativas entre as aplicações. O valor do coeficiente Kappa Ponderado para o índice categorizado foi igual a 0,72 (IC_{95%} 0,56 a 0,87). **Conclusões:** O ICT apresentou níveis de confiabilidade aceitáveis para a classificação da capacidade para o trabalho na população de estudo. Novos estudos são necessários para melhor avaliar a confiabilidade e validade do índice e a fim de complementar a avaliação e a qualidade das informações fornecidas pelo instrumento

Palavras-chave: Trabalhadores. Indústria metalúrgica. Capacidade para o trabalho. Teste-reteste. Confiabilidade. Índice de capacidade para o trabalho. ICT.

Abstract

Objective: Several studies have used the Work Ability Index (WAI) to assess physical, mental and social abilities of workers. Despite its common use, there are very few reports on the psychometric properties of the index. The objective of this study was to assess the test-retest reliability of WAI. **Methods:** A 4-week-interval test-retest study was carried out with 153 workers from the metal-mechanic industry in a Southern Brazilian city. Data analyses included intra-class correlation coefficient (ICC), paired t-test and Kappa coefficient. **Results:** An ICC of 0.84 (95%CI 0.78-0.88) was found for test-retest reliability of the WAI. The paired t-test did not show a significant difference between the 1st and the 2nd applications. Weighted Kappa was 0.72 (95%CI 0.56-0.87). **Conclusion:** Data analyses suggest an acceptable test-retest reliability of the WAI in the study population. New studies should be conducted to test the validity and reliability of the index.

Keywords: Workers. Metal-mechanic industry, Work Ability, Test-Retest. Reliability, Work Ability Index, WAI.

Introdução

As mudanças no perfil demográfico do país apontam para o crescimento da população idosa, ocasionando uma diminuição da razão entre a população ativa e inativa. Aliado a isso, o quadro político e socioeconômico e a consequente precariedade do mercado têm obrigado os trabalhadores a permanecerem por mais tempo no mercado de trabalho e sob piores condições, contribuindo para a perda ou redução da capacidade funcional dessa população^{1,2}.

O envelhecimento funcional influencia diretamente na capacidade para o trabalho. Esse conceito vem se destacando em pesquisas de saúde ocupacional, tanto no Brasil como no mundo³⁻⁸. A capacidade para o trabalho engloba as capacidades físicas, mentais e sociais do indivíduo em relação às exigências do trabalho, levando em consideração também aspectos da comunidade de trabalho, organização e ambiente de trabalho^{5,9}.

Em meados da década de 80, pesquisadores finlandeses desenvolveram o Índice de Capacidade para o Trabalho (ICT), resultado de uma autoavaliação do trabalhador sobre a sua capacidade para o trabalho.

É um instrumento capaz de auxiliar o profissional de Saúde Ocupacional a detectar precocemente possíveis alterações que os trabalhadores possam apresentar em sua funcionalidade no ambiente de trabalho, prevenindo assim o risco de incapacidade. O índice busca indicar quão bem está, ou estará, um trabalhador no presente ou num futuro próximo, e com qual capacidade ele poderá executar o seu trabalho em função das exigências de seu estado de saúde e capacidades física e mental⁸⁻¹¹.

A utilização do índice Brasil teve início na década de 90 e desde então vários estudos com trabalhadores brasileiros têm utilizado o ICT tanto para avaliar a capacidade funcional para o trabalho bem como os fatores de risco associados a esta capacidade^{4-6,12-14}.

Como forma de aumentar a aplicabilidade de um instrumento em pesquisas

científicas e na prática diária, é de grande relevância obter informações adicionais a respeito da sua qualidade. Para tanto, são necessárias diversas avaliações das propriedades psicométricas do instrumento, dentre elas a confiabilidade.

A confiabilidade de um instrumento é o grau de estabilidade que uma medida possui quando a mesma é repetida em condições ideais. Em outras palavras, refere-se ao grau com o qual os resultados obtidos por uma aferição ou procedimento podem ser replicados¹⁵. A confiabilidade do ICT já foi testada em um estudo realizado com 97 trabalhadores da construção civil na Holanda, sugerindo uma confiabilidade aceitável no teste-reteste para aquela população⁸. Entretanto não existem estudos publicados que forneçam informações sobre a confiabilidade do índice no país. Neste sentido, o objetivo do presente estudo foi verificar a confiabilidade do ICT através do método teste-reteste.

Métodos

O estudo foi realizado com uma amostra de conveniência incluindo todos os trabalhadores (N=165) de duas empresas do setor metal-mecânico do município de Caxias do Sul, RS. O uso de amostra não-probabilística se deveu ao fato de que a realização de estudos dentro de empresas privadas não é frequentemente autorizada por diretores e em geral não é bem aceita pelos trabalhadores, pois estes temem que possam se constituir numa forma de avaliação e posterior seleção de funcionários. O tamanho da amostra adotado supera o tamanho proposto por Pasquali de cerca de 10 pessoas por questão¹⁶. Os critérios de inclusão eram possuir idade igual ou superior a 18 anos, ter pelo menos 4 anos de estudo, carga horária de 44 horas semanais, trabalhar no turno diurno e exercer funções técnicas ou de produção. Os trabalhadores participantes responderam o ICT 2 vezes, com intervalo de 4 semanas da aplicação inicial..

Para investigar as características da população em estudo, foi aplicado um ques-

tionário padronizado, pré-codificado e pré-testado contendo variáveis demográficas, ocupacionais e comportamentais.

O questionário do ICT é composto por sete itens, sendo que os escores são adquiridos através de uma determinada pontuação das questões, totalizando dez questões. Os itens são os seguintes:

- Avaliação da capacidade atual para o trabalho comparada com o melhor de toda a vida. Este item compreende uma questão com escore de zero a dez;
- Capacidade para o trabalho relacionado às determinações físicas, mentais ou ambas. A pontuação é dada utilizando-se apenas uma das fórmulas abaixo de acordo com o tipo de exigência do trabalho (física, mental ou com ambas);
 - ✓ Para o trabalho com demandas físicas, por exemplo, auxiliar, instalação ou trabalho doméstico: $(\text{escore físico} \times 1,5) + (\text{escore mental} \times 0,5) = \text{total}$
 - ✓ Para o trabalho com demandas mentais como escritório, magistério ou trabalho administrativo: $(\text{escore físico} \times 0,5) + (\text{escore mental} \times 1,5) = \text{total}$
 - ✓ Já para trabalhos com ambas as exigências, por exemplo, cuidados de enfermagem, odontologia, transporte e trabalho de supervisão em cozinha, a quantidade de pontos permanece inalterada: $\text{escore físico} + \text{escore mental} = \text{total}$
- Este item apresenta uma lista de 51 doenças, onde o trabalhador deve assinalar as patologias diagnosticadas pelo médico e aquelas que em sua opinião ele possui. Não são utilizadas na pontuação aquelas de opinião do indivíduo, somente as diagnosticadas clinicamente. O escore é atribuído da seguinte forma : 1 ponto se pelo menos tem diagnóstico de 5 doenças, 2 pontos se 4 doenças, 3 pontos se 3 doenças, 4 pontos se 2 doenças, cinco se 1 doença e 7 se não apresentar nenhuma doença;
- Estimativa de perda do trabalho devido

a doenças. Este item é avaliado através de uma questão, com escore variando entre 1 e 6 pontos, escolhendo-se o de pior valor assinalado;

- Faltas ao trabalho durante os últimos 12 meses. Este item é avaliado através de uma questão, com escore variando de 1 à 5 pontos;
- Prognóstico próprio sobre a capacidade para o trabalho daqui a 2 anos. Este item é avaliado através de uma questão, com escore de 1, 4 ou 7 pontos, considerando-se o valor circulado no questionário;
- Habilidades mentais, avaliadas através de 3 questões. Os pontos da questão são somados e o resultado é contado da seguinte forma: soma 0-3 = 1 ponto; soma 4-6 = 2 pontos; soma 7-9 = 3 pontos; e soma 10-12 = 4 pontos.

As pontuações de cada item são somadas totalizando um escore mínimo de 7 (pior índice possível) e um máximo de 49 (melhor índice possível). A partir do escore, o sujeito é classificado quanto à sua capacidade para o trabalho em: baixa capacidade (7-27), moderada capacidade (28-36); boa capacidade (37-43) e ótima capacidade (44-49)¹⁰.

A coleta de dados ocorreu entre os meses de setembro a novembro de 2005, sendo realizada dentro das próprias empresas de forma assistida pela pesquisadora principal. Ou seja, cada item do questionário era lido em voz alta, concomitantemente com o autopreenchimento do questionário pelos participantes.

A entrada dos dados no banco foi realizada com o Programa Epi Info, versão 6.0, em dupla entrada, por dois digitadores independentes, para posterior comparação da consistência dos dados e possíveis erros de digitação. A análise dos dados foi realizada com os pacotes estatísticos SPSS versão 11.0 e 'Stata-7.0'.

Foi realizada a análise do estudo teste-reteste, considerando-se o desfecho contínuo e categórico. A normalidade do ICT enquanto variável contínua foi testada através de histogramas e do teste Kolmogorov-Smirnov.

Para contornar o problema da normalidade da distribuição, procedeu-se à transformação logarítmica do ICT enquanto variável contínua. Em seguida, foi realizado o T-teste pareado para comparar as médias dos escores totais do ICT na condição de teste-reteste e o Coeficiente de Correlação Intra-Classe (CCI). Também foi realizada análise não paramétrica com o teste de Wilcoxon. Além destes, foram realizados gráficos de dispersão para os escores do teste e do reteste e o Gráfico Bland-Altman (gráfico de dispersão, com a diferença entre a 1ª e a 2ª medida contra a média das duas medidas)¹⁷.

Para o ICT em 4 categorias, as medidas de concordância entre as duas aferições (teste-reteste) incluíram o Percentual Total de Concordância, o coeficiente Kappa (κ) e o coeficiente Kappa ponderado com pesos pré-definidos pela função *wgt(w)* no programa Stata. Também foi realizada uma análise item a item de todos os componentes do ICT.

O projeto desta pesquisa foi submetido e aprovado pelo Comitê de Ética e Pesquisa da Universidade do Vale do Rio dos Sinos, RS. Todos os participantes assinaram um Termo de Consentimento Livre e Esclarecido.

Resultados

Um total de 153 trabalhadores (93%) concordou em participar da pesquisa. Não houve perdas entre o teste e o reteste. A grande maioria dos participantes (83%) eram homens, cor da pele branca (69%), casados (74%) e com idade entre 20 e 40 anos (75%). A maioria trabalhava no setor de produção (72%), com escolaridade igual ou acima do ensino médio (77%) e com renda entre 2 e 4 salários mínimos (75%).

A Tabela 1 apresenta a análise do teste-reteste realizado com 153 trabalhadores do estudo. O Coeficiente de Correlação Intra-Classe apresentou correlação forte (CCI=0,84; IC_{95%} 0,78 a 0,88, P<0,001). Com a normalização dos dados e a aplicação do T-Teste, as médias entre a 1ª e a 2ª coleta não apresentaram diferenças estatística-

mente significativas ou efeito importante (a diferença entre a média do ICT1 e o ICT2 foi de -0,33 [IC_{95%} -0,66 a 0,01]), porém o limite superior do intervalo de confiança se aproximou da nulidade. O teste não-paramétrico Wilcoxon indicou uma diferença estatisticamente significativa.

Cerca de 37% (N=56) dos participantes obtiveram o mesmo valor do ICT na 1ª e na 2ª aplicações (Figura 1). O gráfico de dispersão de Bland Altman apresentou mais de 95% das diferenças individuais entre a 1ª e a 2ª aplicação em menos de 7,1 pontos (2 DP) acima e abaixo da média (Figura 2).

O percentual de concordância entre as duas aplicações para o ICT em 4 categorias foi igual a 79,7% (IC_{95%} 72,3% a 85,6%) e o valor do Kappa ponderado igual a 0,72 (IC_{95%} 0,56 a 0,87) (Tabela 1).

Quando analisadas individualmente as concordâncias de cada item do questionário foram encontradas concordâncias substanciais ou quase perfeitas (coeficiente Kappa \geq 0,6) naqueles itens que se referem ao número de doenças, se as doenças são impedimento para o trabalho ou ausência do trabalho, e a autoavaliação da capacidade futura para realizar o seu trabalho atual. Os itens que se referem a autopercepção da capacidade física e mental para o trabalho e realização das atividades diárias com energia, prazer e esperança apresentaram

concordâncias moderadas (coeficiente Kappa \geq 0,4) (Tabela 2).

Discussão

O ICT é um questionário auto-aplicável, requerendo assim a capacidade de compreensão e interpretação dos itens. Para o preenchimento do mesmo, tem sido recomendada uma escolaridade mínima de 4 anos dos participantes¹⁰. Todavia, esta regra refere-se a uma população distinta (finlandesa) da população brasileira. Apesar de todos os participantes terem ensino médio ou mais, optou-se pela aplicação assistida do instrumento, visando uma maior compreensão de termos técnicos e científicos, como, por exemplo, morbidades. Este procedimento evitou informações ausentes e as taxas de resposta em ambas aplicações foram totais. Acredita-se que tal procedimento não tenha alterado os resultados da confiabilidade teste-reteste, pois o mesmo método foi adotado em ambas aplicações. Embora vários estudos brasileiros tenham utilizado o ICT, há de se considerar que os parâmetros psicométricos deste instrumento ainda não haviam sido avaliados. Neste sentido o objetivo deste estudo foi o de testar a confiabilidades teste-reteste do índice. Os resultados do CCI demonstraram correlações fortes e estatisticamente

Tabela 1 - Análise de concordância teste-reteste do ICT contínuo e categórico em trabalhadores da indústria metalúrgica de Caxias do Sul 2005. (N=153)

Table 1 - Agreement analyses of test and re-test of the Work Ability Index (ICT) as a continuous and categorical variable in Brazilian metallurgy workers (n=153). Caxias do Sul, 2005.

ICT contínuo	
Coefficiente de Correlação Intra-Classe	0,84 (IC _{95%} 0,78 - 0,88)
T-Teste*	Diferença média= -0,3 (IC _{95%} -0,66 a -0,01) Média teste=41,5/ Média reteste=41,8 P=0,056
ICT categórico	
Concordância	79,7% (IC _{95%} 72,3% - 85,6%)
Kappa	0,61 (IC _{95%} 0,49 - 0,73)
Kappa ponderado	0,72 (IC _{95%} 0,56 - 0,87)

* Teste realizado com o logaritmo da distribuição. O teste de Wilcoxon apresentou resultado significativo P=0,039

* Test conducted with log distribution. Wilcoxon test showed a significant P-value (P=0.039)

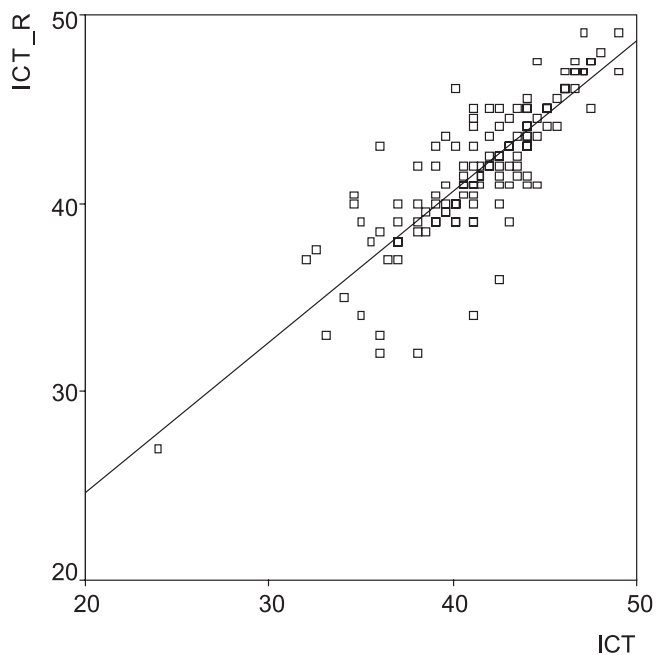


Figura 1 - Gráfico da concordância do Índice de Capacidade para o Trabalho (ICT) entre o teste e o reteste com trabalhadores da indústria metal mecânica de Caxias do Sul, RS. 2005 (n=153).
Figure 1 – Scatter plot of the agreement between test and re-test for the Work Ability Index (ICT) in Brazilian metallurgy workers (n=153).

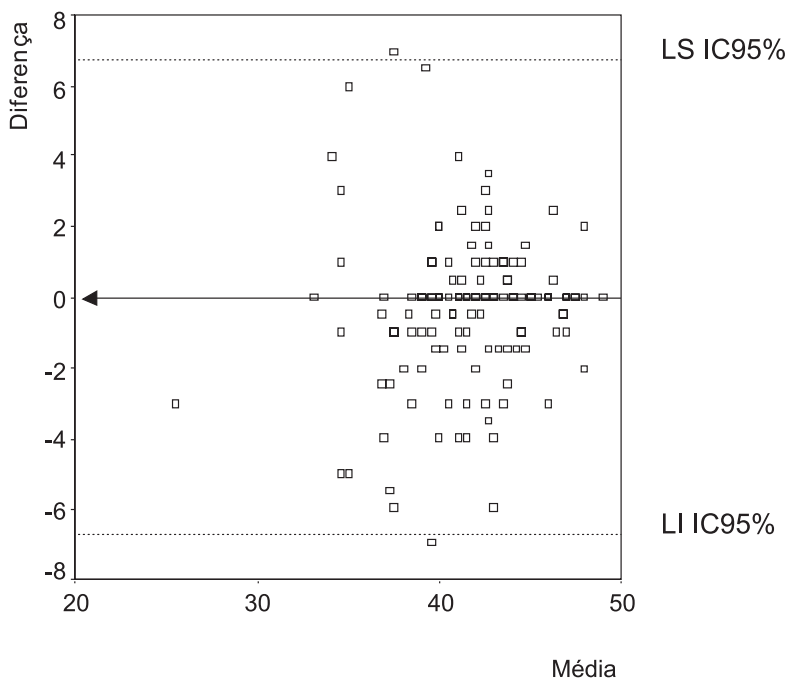


Figura 2 - Gráfico de Bland e Altman, diferenças entre o teste e o reteste contra as médias do teste e reteste e limites dos intervalos de confiança 95% em trabalhadores da indústria metal mecânica de Caxias do Sul, RS.2005 (n=153).
Figure 2 – Bland and Altman scatter plot of test-retest differences (“diferença”) against test-retest means (“media”) and 95% confidence interval limits for the Work Ability Index (ICT) in Brazilian metallurgy workers (n=153).

Tabela 2 - Percentual de concordância, Coeficiente Kappa e Kappa Ponderado de acordo com os itens do Índice de Capacidade para o Trabalho em trabalhadores da indústria metal mecânica de Caxias do Sul, 2005. (N=153).

Table 2 - Percent of agreement, Kappa and Weighted Kappa according to items of the Work Ability Index in metal-mechanic industry workers (N=153). Caxias do Sul 2005.

Item do ICT	% concordância	Kappa (IC _{95%})	Kappa Ponderado (IC _{95%})
1. Na escala de zero a dez, quantos pontos você daria para sua capacidade para o trabalho atual?	59,5	0,43 (0,34 – 0,56)	0,54 (0,39 - 0,69)
2.1. Como você classificaria sua capacidade atual para o trabalho em relação às exigências físicas do seu trabalho?	78,4	0,59 (0,46 – 0,72)	0,66 (0,50 - 0,81)
2.2. Como você classificaria sua capacidade atual para o trabalho em relação às exigências mentais do seu trabalho?	79,1	0,52 (0,40 – 0,65)	0,59 (0,47a 0,78)
2.3. Somatório dos escores das capacidades física e mental.	66,0	0,51 (0,37 – 0,68)	0,67 (0,455 - 0,87)
3.1. Número de doenças.	75,2	0,69 (0,61 – 0,76)	0,90 (0,74 - 1,0)
3.2. Escore das doenças.	75,8	0,70 (0,61 – 0,69)	0,90 (0,72 - 1,0)
4. Sua lesão ou doença é um impedimento para seu trabalho atual?	81,7	0,70 (0,58 – 0,81)	0,80 (0,64 - 0,96)
5. Quantos dias inteiros você esteve fora do trabalho durante os últimos 12 meses?	92,2	0,82 (0,69 – 0,94)	0,89 (0,07 - 1,0)
6. Considerando sua saúde, você acha que será capaz de daqui a 2 anos fazer seu trabalho atual?	95,4	0,70 (0,58 – 0,82)	0,88 (0,72 - 1,03)
7.1. Recentemente você tem conseguido apreciar suas atividades diárias?	75,2	0,57 (0,45 – 0,69)	0,63 (0,47 - 0,78)
7.2. Recentemente você tem se sentido ativo ou alerta?	76,5	0,56 (0,42 – 0,70)	0,63 (0,48 - 0,79)
7.3. Recentemente você tem se sentido cheio de esperança para o futuro?	71,9	0,46 (0,35 – 0,57)	0,59 (0,43 - 0,74)
7.4. Escore dos três itens da questão 7.	79,7	0,52 (0,40 – 0,65)	0,51 (0,43 - 0,77)

significativas entre as duas aplicações¹⁸. A concordância exata entre os escores para o teste e o reteste em 37% dos estudados é superior ao valor encontrado por Zwart et al⁸ (2002), em que 25% dos estudados apresentaram o mesmo escore. No entanto, com esse gráfico não é possível julgar se a concordância varia entre as amplitudes dos valores das medidas. Diante disso, Bland e Altman¹⁷ (1986) sugerem comparar a diferença das medidas contra as médias das medidas, sendo assim possível a visualização dos valores de discordância. Nesse estudo, os resultados foram compatíveis com os encontrados por Zwart et al⁸ (2002), o que sugere uma boa confiabilidade do ICT. Tal achado foi confirmado pelo T-teste pareado,

o qual não apresentou diferença estatisticamente significativa entre as médias do ICT na 1ª e 2ª aplicações, e seu intervalo de confiança incluiu a nulidade¹⁷. Entretanto, cautela deve ser exercida, pois o teste não paramétrico de Wilcoxon mostrou-se significativo (P=0,039) e os valores P do teste t encontram-se próximos da significância estatística. Tais resultados apontam para a necessidade de condução de novos estudos em outros ambientes de trabalho e outras ocupações, de forma a confirmar a confiabilidade do índice.

Quando realizado o cálculo do percentual de concordância para o ICT em 4 categorias, este percentual é maior que o esperado; entretanto, tal resultado dá uma

falsa boa impressão da concordância entre as medidas. Isto posto, certo grau de concordância é esperado ao acaso¹⁸. Para evitar tal problema, utilizou-se o teste Kappa. Destaca-se que os valores do Kappa ponderado diferem e são sempre mais favoráveis que o da estatística Kappa. Ao serem utilizadas três ou mais categorias para se expressar os resultados, o índice admite discordâncias entre a primeira e a segunda aplicações, porém confere uma maior importância à concordância em categorias de resposta mais próximas quando procede o cálculo da reprodutibilidade. Quando o índice foi analisado nas categorias propostas pelo autor original, seus resultados sugerem boa concordância. Entretanto, em alguns dos itens do questionário essa concordância é apenas moderada. Por exemplo, os mesmos valores para a capacidade referida em uma escala de zero a dez pontos foram encontrados em apenas 60% dos respondentes (Tabela 2) Isto talvez seja explicado pelo grau de generalidade ou subjetividade da pergunta ou por sua relação com a situação ou sentimento no momento do teste. Observa-se que o percentual de concordância aumenta quanto mais específica é a pergunta ou quanto mais se relaciona, por exemplo, a capacidade para o trabalho com algum outro aspecto como exigência física ou mental, ou nas questões que envolvem dados mais objetivos, tais como número de

dias de afastamento e incapacidade para a realização de tarefas.

Há de se ressaltar que o intervalo de tempo entre os testes pode influenciar a confiabilidade teste-reteste. Por exemplo, se o teste é repetido em um curto intervalo de tempo, a memória da 1ª aplicação do ICT pode influenciar os resultados da 2ª aplicação do instrumento e, assim, falsamente inflar a confiabilidade teste-reteste. Opostamente, as mudanças no estado de saúde que ocorrem entre os testes pode influenciar as respostas dos entrevistados e diminuir a confiabilidade teste-reteste.⁸ O tempo adotado neste estudo, de 4 semanas entre as duas medidas, pode ter evitado a variação da percepção dos sintomas positivos ou negativos e os problemas de memória.

Estudos de confiabilidade do ICT não foram identificados no Brasil, não sendo possível estabelecer comparações da confiabilidade teste-reteste do índice no nível nacional. As análises realizadas sugerem níveis aceitáveis de confiabilidade da avaliação da capacidade para o trabalho¹⁹. Porém, para que o mesmo possa ser utilizado em pesquisas de saúde ocupacional e na prática diária, reforça-se a necessidade da condução de novos estudos sobre as mais diversas formas de validade, incluindo a transcultural, e confiabilidade, a fim de complementar a avaliação e a qualidade das informações fornecidas pelo instrumento.

Referências

1. Fundação Instituto Brasileiro de Geografia e Estatística (IBGE). *Perfil dos idosos responsáveis pelo domicílio no Brasil*. Estudos e Pesquisa: Rio de Janeiro; 2002.
2. Giatti L, Barreto SM. Saúde, trabalho e envelhecimento no Brasil. *Cad Saúde Pública* 2003; 19: 759-71.
3. Bellusci SM, Fischer FM. Envelhecimento funcional e condições de trabalho em servidores forenses. *Rev Saúde Pública* 1999; 33(6): 602-9.
4. Duran ECM, Cocco MIM. Capacidade para o trabalho entre trabalhadores de enfermagem do pronto-socorro de um hospital universitário. *Rev Latino-am Enfermagem* 2004; (1): 43-9.
5. Ilmarinen J, Tuomi K, Eskelinen L, Nygard C-H, Huuhtanen P, Klockars M. Summary and recommendations of a project involving cross-sectional and follow-up studies on the aging worker in Finnish municipal occupations (1981-1985). *Scand J Work Environ Health* 1991; 17 (S1): 135-41.
6. Monteiro MS. *Envelhecimento e capacidade para o trabalho entre trabalhadores brasileiros* [tese de doutorado]. São Paulo: Faculdade de Saúde Pública da Universidade de São Paulo; 1999.
7. Raffone AM, Hennington EA. Avaliação da capacidade funcional dos trabalhadores de enfermagem. *Rev Saúde Pública* 2005; 39(4): 669-76.

8. Zwart BCH, Frings-Dresen, MHW, Duivenbooden, JC. Test-retest reliability of the work ability index questionnaire. *Occup Med*, 2002; 52(S 4): 177-81.
9. Bellusci SM, Fischer FM. Envelhecimento funcional e condições de trabalho em servidores forenses. *Rev Saúde Pública* 1999; 33(6): 602-9.
10. Tuomi K, Ilmarinen J, Jahkola A, Katajarinne L, Tulkki A. *Work Ability Index*. Helsinki: Finnish Institute of Occupational Health; 1998.
11. Pohjonen T. Perceived work ability of home care workers in relation to individual and work-related factors in different age groups. *Occup Med* 2001; 51(3): 209-17.
12. Andrade CB. *Envelhecimento e capacidade para o trabalho dos trabalhadores do serviço de higiene e limpeza de um hospital universitário* [dissertação de mestrado]. Campinas: Universidade Estadual de Campinas; 2002.
13. Metzner R, Fischer F. Fadiga e capacidade para o trabalho em turnos fixos de doze horas. *Rev Saúde Pública* 2001; 35(6): 548-53.
14. Walsh IAP, Corral S, Franco RN, Canetti EEF, Alem MER, Coury HJCG. Capacidade para o trabalho em indivíduos com lesões músculo-esqueléticas crônicas. *Rev Saúde Pública* 2004; 38(2): 149-56.
15. Last JM. *A dictionary of epidemiology*. Oxford: University Press; 2001. pp. 156-7.
16. Pasquali L. *Psicometria: teoria dos testes na psicologia e na educação*. Petrópolis: Vozes; 2003.
17. Bland JM, Altman, DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307-10.
18. Rothwell PM. Analysis of agreement between measurements of continuous variables: general principles and lessons from studies of imaging of carotid stenosis. *J Neurol* 2000; 435: 825-34.
19. Cohen J. A coefficient of agreement for nominal scales. *Educational Psychology Measurement* 1960; 20: 37-46.

Recebido em: 03/06/08
Versão final reapresentada em: 19/02/09
Aprovado em: 26/03/09