

CONFIABILIDAD TEST-RETEST DEL CUESTIONARIO 'ÍNDICE DE CAPACIDAD DE TRABAJO' EN TRABAJADORES DE LA SALUD DE SERVICIOS HOSPITALARIOS DE ARROYO NARANJO, LA HABANA

TEST-RETEST RELIABILITY OF THE WORK ABILITY INDEX IN HEALTH WORKERS OF HOSPITAL INSTITUTIONS OF ARROYO NARANJO MUNICIPALITY, HAVANA PROVINCE

Graciela Rodríguez Mancebo ¹
Georgina Maritza López Pumar ²
María de Lourdes Marrero Santos ³

RESUMEN

Introducción: El Índice de capacidad de trabajo (ICT) se destaca como un instrumento útil para medir la capacidad de trabajo, basado en la percepción del trabajador acerca de su salud, las demandas de trabajo y sus recursos mentales. El conocimiento acerca de la calidad del proceso de medición mediante su uso es muy importante. El propósito de este estudio fue evaluar la confiabilidad en términos de estabilidad en el tiempo del ICT en su versión cubana. **Material y método:** En una muestra de 136 trabajadores de instituciones hospitalarias del municipio Arroyo Naranjo, La Habana, se empleó el procedimiento test-retest para evaluar la confiabilidad del ICT, con un intervalo de 2-3 semanas entre las aplicaciones del cuestionario. Se utilizó el coeficiente de correlación intraclass (CCI), la prueba *t* de comparación de medias pareadas para la puntuación total del ICT, y el coeficiente Kappa para su clasificación en cuatro categorías. **Resultados:** Se obtuvo un CCI de 0,923 [IC_{95%} (0,894-0,945)] y se aceptó la hipótesis de igualdad de las medias de la puntuación del ICT en las dos aplicaciones (*t* = 3,15; *p* = 0,75). La concordancia en la clasificación de los sujetos en las categorías del ICT en el test y el retest fue buena [Kappa 0,743; IC_{95%} (0,64-0,84)]. **Conclusiones:** Los resultados del estudio son indicativos de una adecuada confiabilidad test-retest en la clasificación de la capacidad de trabajo de los sujetos mediante el cuestionario ICT. Estos resultados son consistentes con la evidencia publicada en relación con la calidad del proceso de medición de la capacidad de trabajo utilizando el cuestionario ICT en poblaciones laborales de diferentes países, lo cual apoya la experiencia acerca de su aplicabilidad en el quehacer de la salud ocupacional.

Palabras clave: índice de capacidad de trabajo, confiabilidad test-retest, trabajadores de la salud

ABSTRACT

Introduction: The Work Ability Index (WAI) appears as a useful instrument for measuring work ability on the base of workers perception of own health, work demands and mental resources. Knowledge

about quality of work ability measurement process by using WAI questionnaire remains like an important research issue in occupational health field. This study aimed to assess test-retest reliability of the Cuban version of WAI questionnaire. **Material and method:** A test-retest procedure was used to evaluate reliability of WAI in a sample of 136 health workers from hospitals institutions located in Arroyo Naranjo municipality, Havana province. A 2-3 weeks interval between two applications was observed. Data analyses consisted in intra-class correlation coefficient (ICC), paired *t* test for the WAI score and Kappa coefficient for classification of WAI in four categories. **Results:** An ICC of 0,923 [IC_{95%} (0,894-0,945)] was observed for test retest reliability of the WAI. The paired *t*-test showed equality of two Means for the 1st and 2nd applications (*t* = 3,15; *p* = 0,75). It was observed a good agreement for the classification of subjects in one of the four WAI categories [Kappa 0,743; IC_{95%} (0,64-0,84)]. **Conclusions:** The study provided indication of an acceptable test retest reliability for classification of workers by means of the Cuban version of WAI questionnaire. Results are consistent with published evidence related to the quality of work ability measurement process by using WAI questionnaire for diverse workers populations of different countries, and add support about the applicability of this instrument in occupational health practice.

Keywords: work ability index, test-retest reliability, health workers

INTRODUCCIÓN

El estudio de la capacidad de trabajo en relación con la salud, los cambios que ocurren con la edad, la promoción y la prevención de incapacidades laborales son aspectos de gran interés para los investigadores a nivel mundial, el cual se ha visto reforzado por un nuevo escenario caracterizado por el envejecimiento poblacional y los bajos índices de natalidad,

¹ Licenciada en Enfermería, Máster en Salud de los Trabajadores. Vicedirección de Atención Médica, Instituto Nacional de Salud de los Trabajadores, La Habana, Cuba

² Médico especialista de I grado en Higiene, Máster en Salud de los Trabajadores, Investigadora Auxiliar, Profesora Asistente. Vicedirección de Investigaciones y Docencia, Instituto Nacional de Salud de los Trabajadores, La Habana, Cuba

³ Licenciada en Psicología de la Salud, Máster en Psicología de la Salud y en Salud de los Trabajadores, Investigadora Auxiliar, Profesora Asistente. Departamento de Psicología, Instituto Nacional de Salud de los Trabajadores, La Habana, Cuba

Correspondencia:

MSc Georgina Maritza López Pumar
Calzada de Bejucal km 7½, N° 3035, entre Heredia y 1ª, La Esperanza, Arroyo Naranjo,
La Habana, CP 10200, Cuba
Tel.: (537) 643 8107
E-mail: ginalp@infomed.sld.cu

que condicionan la disminución de las fuerzas productivas y comprometen el desarrollo económico de los países que enfrentan esta problemática.

El estudio de la capacidad de trabajo constituye un gran reto para la investigación y la prevención en salud ocupacional¹. El desarrollo de métodos e instrumentos para su estudio transita por el reconocimiento de su complejidad como concepto. Disponer de instrumentos sencillos para obtener información acerca de la capacidad de trabajo no resulta una tarea sencilla, pero sí necesaria para garantizar la sostenibilidad y el avance de las estrategias para su promoción y la prevención de su temprano deterioro, como importantes objetivos de los servicios de salud ocupacional.

La capacidad de trabajo puede ser examinada a través de la percepción de los individuos, pero es necesario aquí remarcar que esta no es una cuestión que se circunscribe únicamente a características individuales. La capacidad de trabajo percibida está íntimamente vinculada con el trabajador y también con la labor que este realiza. Las dimensiones de la capacidad de trabajo que se relacionan con la tarea, la organización y el contexto social amplio, están representadas en el concepto de capacidad de trabajo percibida a través de la experiencia y las valoraciones del trabajador².

El Índice de capacidad de trabajo (ICT) se destaca como un instrumento útil para medir la capacidad de trabajo, basado en un conjunto de preguntas que exploran la percepción del trabajador sobre diferentes aspectos que tributan al constructo *capacidad de trabajo*. El ICT constituye la operacionalización de un concepto que refleja la interacción entre los recursos individuales del trabajador y las demandas del trabajo en las condiciones sociales y organizativas en que este se desenvuelve¹³.

El ICT fue desarrollado por un equipo multidisciplinario integrado por psicólogos, médicos, bioestadísticos, epidemiólogos y encuestadores del área de Salud Ocupacional del Instituto Finlandés de Salud Ocupacional. La definición conceptual que sustenta el ICT se expresa en la pregunta "¿qué tan bien se encuentra el trabajador en el presente y prevé estará en el futuro cercano, para realizar su labor en correspondencia con las demandas del trabajo, su salud y sus recursos mentales?". El modelo propuesto por los autores del ICT describe la capacidad de trabajo como una resultante de la interacción entre dos componentes fundamentales; por una parte, el trabajo en sí mismo, el ambiente de trabajo, la organización y la comunidad de trabajo, y por otra parte, el trabajador, su salud, competencias y valores. En un enfoque integral, el modelo reconoce que la capacidad de trabajo refleja el impacto de factores que se colocan fuera del contexto laboral, como la familia, los amigos y la sociedad toda³.

Con independencia del reconocimiento de la utilidad de un instrumento a los fines para los cuales se ha creado, la extensión de su uso fuera de las fronteras y el contexto en el cual se ha probado requiere en ocasiones de su traducción a otros idiomas, de la adaptación lingüística atendiendo a las particularidades de la cultura y la práctica profesional de cada país o región, y a la obtención de la evidencia que avale la calidad de

las mediciones que derivan de su aplicación.

La validez y confiabilidad del ICT han sido evaluadas por varios autores. Los primeros estudios se realizaron en trabajadores municipales de Finlandia^{4,7}. Eskelinen et al⁸, en un estudio de seguimiento en trabajadores envejecidos, mostró que las variables de salud presentan una fuerte correlación con las puntuaciones del ICT a nivel grupal, indicando una adecuada validez de constructo de este instrumento. La puntuación del ICT y de todos sus ítems predice la discapacidad para el trabajo, el retiro anticipado y la mortalidad^{9,10}. Radkiewicz y Widderszal-Bazyl¹¹ realizaron un estudio de las propiedades psicométricas del ICT en una muestra de 38 000 trabajadores de enfermería en 10 países. Estos autores identificaron por medio del método de componentes principales, dos factores que explican la estructura dimensional del ICT en 8 países, y apenas un factor en dos países. El coeficiente alfa de Cronbach en este estudio fue de 0,72 para toda la muestra, y varió de 0,54 a 0,79 en los diferentes países. Los resultados de otros estudios¹² han mostrado una asociación fuerte y significativa entre la puntuación del ICT y la puntuación del estado de salud medido por el Índice de salud general (SF-36).

Martínez et al¹³, en un estudio con 475 trabajadores del sector eléctrico, identificaron, por medio del análisis de componentes principales, tres factores con autovalores mayores que 1, explicando el 57,9 % de la varianza total acumulada. En este estudio los autores encontraron asociación positiva y estadísticamente significativa de la puntuación del ICT con las tres dimensiones de salud general del Cuestionario de salud y calidad de vida (SF-36). La dimensión de salud física presentó una correlación más fuerte con el ICT en comparación con las dimensiones de salud general y salud mental.

Además de la evidencia obtenida sobre la validez de constructo del ICT, las investigaciones realizadas por Alavinia et al¹⁴ demostraron que el ICT es un fuerte predictor del retiro por discapacidad. En su estudio de seguimiento realizado a trabajadores de 40 años y más, este investigador obtuvo que las dimensiones del ICT: capacidad de trabajo actual, problemas de salud, deterioro funcional y ausentismo por enfermedad, así como los recursos mentales, constituyen importantes factores pronóstico para la discapacidad futura.

La confiabilidad test-retest del cuestionario ICT fue estudiada por Zwart et al¹⁵ en una muestra de 859 trabajadores de la construcción; los resultados evidenciaron la estabilidad de los mismos después de un intervalo de tiempo de 4 semanas. Los estudios realizados por Renosto et al¹⁶ y por Almeida da Silva et al¹⁷, obtuvieron una confiabilidad test-retest aceptable en aplicaciones sucesivas del cuestionario ICT. En la tabla 1 se recogen los resultados obtenidos en estudios realizados para comprobar la confiabilidad del cuestionario ICT, correspondiente a versiones en diferentes idiomas, realizadas en muestras de trabajadores pertenecientes a diversos sectores de la producción y los servicios. De forma general, los resultados denotan una adecuada consistencia y estabilidad del cuestionario ICT en diferentes versiones y países.

Tabla 1
Resultados de la evaluación de la confiabilidad del cuestionario ICT en diferentes estudios

Autor, año y país	Población de estudio y ocupación	Clasificación del ICT (escala)	Método de evaluación de la confiabilidad	Resultados
Sjögren-Rönkä ¹⁹ (2002) Finlandia	88 trabajadores de oficina	Puntuación total (cuantitativa)	Consistencia interna	CCI = 0,94
Zwart et al ¹⁵ (2002) Holanda	859 empleados del sector industrial.	4 categorías, puntuación total (continua)	Confiabilidad test-retest	Índice de concordancia = 0,66 Gráfico de Bland-Altman: El 95% de las diferencias en la puntuación del test y retest se ubican en el intervalo de 2DE (-14 a 9)
Radkiewicz P et al ²⁰ (2005) Polonia	38 000 enfermero(a)s de 10 países de Europa	Puntuación total (cuantitativa)	Consistencia interna	Alfa de Cronbach en los 10 países fue de 0,72
MA La ji et al ²¹ (2000) China	30 trabajadores profesionales	Puntuación total (cuantitativa)	Confiabilidad test-retest, consistencia interna	CCI = 0,84 (test retest) R de Pearson = 0,503 (Dos mitades)
Zhang L et al ²² (2008) China	245 trabajadores de diferentes ocupaciones	Puntuación total (cuantitativa)	Confiabilidad test-retest, consistencia interna	CCI (rango de 0,597 a 0,897) r de Pearson = 0,30 Alpha de Cronbach (rango de 0,701 a 0,808)
Fischer et al ²³ (2007) Brasil	996 trabajadores de enfermería	2 categorías (nominal)	Consistencia interna:	Alfa de Cronbach = 0,90
Martinez et al ¹³ (2008) Brasil	582 trabajadores del sector eléctrico	Puntuación total (cuantitativa)	Consistencia interna	Alfa de Cronbach = 0,72
Martinez et al ²⁴ (2009) Brasil	475 trabajadores del sector energético	Puntuación total (cuantitativa)	Consistencia interna	Alfa de Cronbach = 0,72
Renosto et al ¹⁶ (2009) Brasil	153 trabajadores del sector metalmeccánico	Puntuación total (cuantitativa) 4 categorías (ordinal)	Confiabilidad test-retest	CCI = 0,8 Prueba <i>t</i> pareada no observó diferencias significativas entre las dos aplicaciones. Kappa ponderado de 0,72 (IC _{95%} 0,56 - 0,87)
Almeida da Silva H et al ¹⁷ (2010) Brasil	80 enfermeras	Puntuación total (cuantitativa) 4 categorías (ordinal)	Confiabilidad test-retest	CCI = 0,79 [IC _{95%} (0,67-0,86)], Kappa ponderado para el ICT categórico 0,69 [IC _{95%} (0,50-0,80)] o Kappa ponderado para todos los ítems entre 0,39 y 0,82

Actualmente disponemos de una versión traducida y adaptada del cuestionario ICT para Cuba¹⁸. En el marco del estudio de sus propiedades psicométricas, la presente investigación se aplica a evaluar la confiabilidad del cuestionario ICT en su versión cubana.

MATERIAL Y MÉTODO

La aplicabilidad de un instrumento en la práctica depende de la calidad de la información que este puede generar. La confiabilidad, como principal variable de nuestro estudio, se define como una propiedad psicomé-

trica que hace referencia a la ausencia de errores de medida, lo que es lo mismo, al grado de consistencia y estabilidad de las puntuaciones obtenidas a lo largo de sucesivos procesos de medición con un mismo instrumento²⁵. La confiabilidad constituye una medida de la calidad del dato, en el sentido del error que se puede cometer al realizar mediciones individuales. La evaluación de la confiabilidad mediante el procedimiento test-retest implica la aplicación del mismo test dos veces a los mismos sujetos. La aplicaciones se pueden realizar inmediatamente o dejando un intervalo de tiempo entre el test y el retest. Este procedimiento nos permitió eva-

luar el grado de estabilidad del test en el tiempo.

Población de estudio

El estudio se llevó a cabo en el periodo comprendido entre septiembre de 2009 y octubre de 2010, en tres instituciones del II y III nivel de atención de salud enclavadas en el municipio Arroyo Naranjo. Se estudiaron los sujetos con un mínimo de un año laborando en el puesto de trabajo actual, edades de 45 años y más y voluntariedad expresada mediante consentimiento informado de participar en el estudio. Se excluyeron aquellos casos que no completaron la segunda aplicación del cuestionario. El tamaño de la muestra recolectada (n=136) supera el tamaño propuesto por Pascuali de cerca de 10 personas por pregunta contenida en el cuestionario para este

tipo de estudio²⁵. Se observó un intervalo de 2 a 3 semanas entre la primera y segunda aplicación del cuestionario ICT, el que fue aplicado en el horario laboral diurno en las propias áreas de trabajo, utilizando la técnica de la entrevista, a fin de incrementar la claridad en la comprensión de términos y conceptos²⁶.

Cuestionario Índice de capacidad de trabajo

El cuestionario ICT representa la operacionalización del concepto de capacidad de trabajo, definido como la facultad o habilidad de un trabajador para realizar su trabajo, atendiendo a su salud, las exigencias del trabajo y sus reservas mentales¹. El cuestionario está integrado por 10 preguntas (tabla 2).

Tabla 2
Cuestionario Índice de capacidad de trabajo

Ítem	Preguntas	Puntuación
1	Capacidad de trabajo actual comparada con la mejor de por vida	0-10
2	Capacidad de trabajo actual relación con las exigencias de trabajo (físicas y mentales)	2-10
3	Enfermedades actuales diagnosticadas por el facultativo	1-7
4	Estimado del deterioro del desempeño laboral por enfermedad	1-6
5	Ausentismo por enfermedad (últimos 12 meses)	1-5
6	Pronóstico de la capacidad de trabajo para los próximos dos años	1, 4 o 7
7	Recursos psicológicos (disfrute cotidiano, actividad y buen ánimo, optimismo)	1-4
	Cuestionario ICT	7-49

La sumatoria de las calificaciones de las preguntas se mueve en un rango de 7 a 49 puntos. Se utilizó, además de la escala continua para la puntuación total del cuestionario, su categorización en escala ordinal y nominal

dicotómica, utilizando los intervalos y puntos de corte recomendados en la bibliografía consultada correspondientes a los percentiles 15 y 85 y la mediana de la distribución de la puntuación del ICT¹.

Escala ordinal	Capacidad de trabajo
7-27 puntos	Pobre
28-36 puntos	Moderada
37-43 puntos	Buena
44-49 puntos	Excelente

Diseño estadístico

Se confeccionó una base de datos utilizando el paquete estadístico SPSS versión 11.5 para Windows, y el programa Análisis epidemiológico de datos tabulados EPIDAT versión 3.1. El análisis de la confiabilidad test-retest se realizó considerando la variable *capacidad de trabajo* medida mediante el cuestionario ICT de forma continua (puntuación total del cuestionario), y de forma categórica (utilizando los intervalos de la puntuación obtenida descritos con anterioridad). Para el caso de la variable continua del ICT, se comprobó el ajuste de los datos a la distribución normal mediante la prueba de Kolmo-

gorov-Smirnov. Para comparar las medias de las puntuaciones totales del ICT obtenidas en las dos aplicaciones del cuestionario, se utilizó la prueba *t* para muestras relacionadas y se calculó el coeficiente de correlación intraclase. Se realizaron gráficos de dispersión para las puntuaciones obtenidas en el test y el retest. Para evaluar el acuerdo entre las puntuaciones de las dos mediciones, se utilizó el método de Bland & Altman²⁸. Para el ICT categórico, se tomó como medida de la concordancia entre las dos mediciones el coeficiente Kappa (κ) y el coeficiente Kappa ponderado por pesos de Cicchetti, basado en las distancias relativas entre las categorías de clasificación. Esta ponderación da un peso similar a los acuerdos y a los des-

acuerdos observados^{29,30}. Se consideró un valor kappa de 0,40 o mayor como fiabilidad buena, y por encima de 0,75 como fiabilidad excelente³¹. Para los análisis que fueron planteados se fijó una significación del 95 %.

RESULTADOS Y DISCUSIÓN

Análisis de la confiabilidad test-retest de la versión cubana del ICT

Se observó un resultado muy similar en la puntuación obtenida en las preguntas que integran el cuestionario ICT en las

dos aplicaciones realizadas (tabla 3). La mayor fluctuación en las respuestas se puede observar en la pregunta 1 del cuestionario, que explora la valoración de la capacidad de trabajo actual comparada con la mejor de por vida. La prueba de los rangos con signos de Wilcoxon denota que no existen diferencias entre las puntuaciones de esta pregunta en las dos pruebas ($Z=545$; $p=0,586$). En cuanto la calificación del ICT, se observa que el rango de las puntuaciones alcanzadas por los sujetos coincide para los ítems 2 al 7.

Tabla 3
Descriptivos de las puntuaciones de las preguntas y del cuestionario ICT en el test y el retest

Estadísticos	Ítems del cuestionario ICT							Puntuación total ICT
	1	2	3	4	5	6	7	
TEST								
Mediana	10	9	3	6	5	7	4	41
Media	8,92	8,76	3,19	5,35	3,98	6,54	3,58	40,32
Desviación estándar	1,53	1,29	1,91	1,06	1,55	1,23	,61	5,04
Mínimo	3	6	1	2	1	1	2	26
Máximo	10	10	7	6	5	7	4	49
RETEST								
Mediana	9	9	3	6	5	7	4	41
Media	9,13	8,82	3,15	5,32	3,91	6,43	3,62	40,37
Desviación estándar	1,06	1,28	1,90	1,00	1,57	1,44	0,56	4,68
Mínimo	5	6	1	2	1	1	2	26
Máximo	10	10	7	6	5	7	4	49

Se obtuvo el coeficiente de correlación intraclass, como medida del acuerdo de los valores de puntuación del ICT entre los propios casos, el cual muestra una correlación fuerte entre las dos mediciones (0,923). Su intervalo de confianza se mueve en un rango estrecho, con un límite inferior que supera el valor de 0,85 como expresión de una buena correlación entre las dos puntuaciones (tabla 4). Se comprobaron los presupuestos requeridos para los datos en cuanto su apego a la distribución normal ($Z_{\text{Kolmogorov-Smirnov}} \text{ test}=1,262$; $p=0,083$) y ($Z_{\text{Kolmogorov-Smirnov}} \text{ retest}=1,154$ $p=0,139$), y se corrió la prueba t para

muestras relacionadas a fin de determinar si las medias de las dos puntuaciones del ICT diferían entre sí. La diferencia observada entre las medias resultó ser no significativa estadísticamente ($p=0,753$), por lo que se acepta la hipótesis de la igualdad de las medias de la puntuación del ICT en las dos aplicaciones (test y retest). La evidencia que aporta el intervalo de confianza de la diferencia de medias corrobora esta hipótesis, al contener al cero.

Tabla 4
Análisis de concordancia test-retest para la puntuación total del ICT (n=136)

ICT (puntuación total) – (escala continua)	
Coeficiente de correlación intra-clase	0,923 [IC _{95%} (0,894-0,945)]
Prueba t	Diferencia de medias = -0,051 [IC _{95%} (-0,374-0,271)] Media test = 40,32 / Media retest = 40,37 Valor de t = -0,315; p = 0,753

El diagrama de Bland & Altman (figura y tabla 5) nos permitió observar la dispersión de las diferencias individuales entre la primera y segunda aplicación, y la media de las puntuaciones obtenidas en el test y el retest. Podemos apreciar la extensión del desacuerdo entre las dos mediciones, al tomar el rango de 2DE (3,8 puntos)

por encima y por debajo de la media de la diferencia de las puntuaciones del ICT en las dos aplicaciones. Aproximadamente el 95 % de las discrepancias observadas se ubica en este intervalo (+3,8 y -3,8 puntos), y solo 9 casos (5,6%) se colocan fuera del límite de la media más 2DE.

Figura
Diferencias entre las puntuaciones del ICT en el test y el retest contra la media de las dos mediciones (diagrama de Bland&Altman)

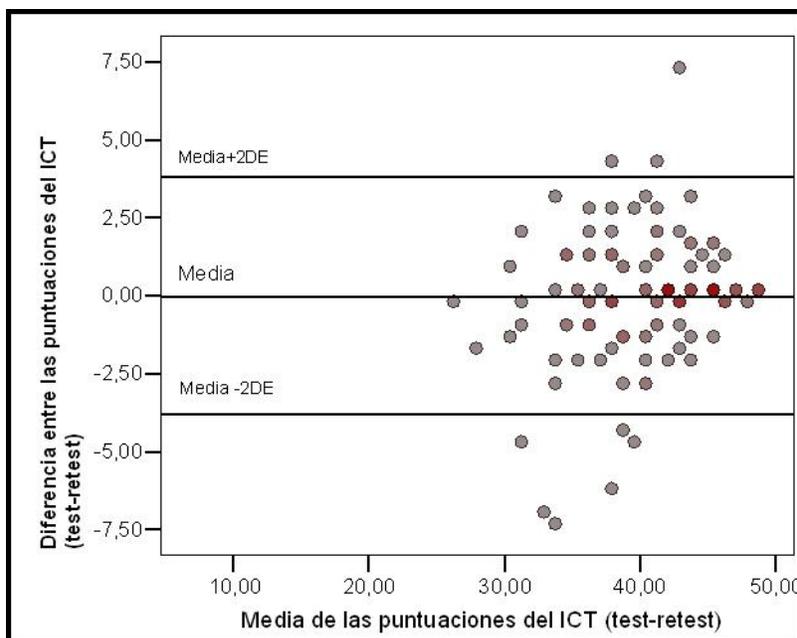


Tabla 5
Media de la diferencia de las puntuaciones del ICT y su DE

	N	Media	DE	2DE
Diferencia en la puntuación del ICT entre el test y el retest	136	-0,051	1,906	3,813

Se analizó también el acuerdo observado para la clasificación de los sujetos en las diferentes categorías del ICT (tabla 6). El índice de concordancia observado fue del 83,8 %. En la segunda aplicación del cuestionario, 10 sujetos (7,35 %) fueron clasificados en una categoría superior, y 12 sujetos (8,82 %) en una categoría inferior a la alcanzada en la primera aplicación.

El valor del coeficiente de correlación por rangos de Spearman para datos ordinales es indicativo de una fuerte correlación ($\geq 0,85$) en la clasificación de los sujetos en las dos aplicaciones (tabla 7). El índice de concordancia observado fue de 83,8 % y superó el esperado. El índice de Kappa muestra un resultado que califica de

muy buena la concordancia entre las clasificaciones, lo que descarta un posible efecto del azar en el acuerdo existente entre las categorías del ICT en las dos aplicaciones del cuestionario. El índice Kappa ponderado, que tiene en cuenta la distancia entre las discrepancias observadas, denota un acuerdo excelente en los resultados obtenidos en el test y el retest.

El análisis de las puntuaciones obtenidas en las diferentes preguntas que integran el cuestionario ICT (tabla 8) muestra un acuerdo que califica como muy bueno (Índice Kappa $\geq 0,6$) entre las dos aplicaciones, excepto para la pregunta 1, en que la concordancia fue moderada (Índice Kappa $\geq 0,4$).

Tabla 6
Distribución de los sujetos en las categorías del ICT en el test y el retest

Categorías del ICT	Retest				Total (%)
	Pobre	Moderada	Buena	Excelente	
Test					
Pobre	1	1	0	0	2 (0,14)
Moderada	0	22	6	0	28 (20,59)
Buena	0	6	57	3	66 (48,53)
Excelente	0	0	6	34	40 (29,41)
Total (%)	1 (0,07)	29 (21,32)	69 (50,73)	37 (27,21)	136 (100,00)

Tabla 7
Análisis de concordancia test-retest para las categorías del ICT (n=136)

ICT (categorías - escala ordinal)	
Índice de concordancia	83,8 %
Coefficiente de correlación (Rho de Spearman)*	0,851
Kappa	0,7431 IC _{95%} (0,6444-0,8419)
Kappa ponderado	0,7904 IC _{95%} (0,7081-0,8727)

* La correlación es significativa al nivel 0,01 (bilateral).

Tabla 8
Porcentaje de concordancia, Coeficiente Kappa y Kappa ponderado para las preguntas del ICT (test y retest)

Preguntas del cuestionario	Índice de capacidad de trabajo	IC	Kappa (IC _{95%})	Kappa ponderado (IC _{95%})
1.	Capacidad de trabajo actual comparada con la mejor de por vida	64,71	0,472 (0,366-0,573)	0,552 (0,453-0,649)
2.	Capacidad de trabajo en relación con las exigencias del trabajo	72,06	0,606 (0,501-0,711)	0,701 (0,606-0,796)
a.	Capacidad de trabajo actual en relación con las exigencias físicas del trabajo	78,52	0,657 (0,546-0,767)	0,700 (0,595-0,805)
b.	Capacidad de trabajo actual en relación con las exigencias mentales del trabajo	81,48	0,658 (0,537-0,778)	0,695 (0,580-0,810)
3.	Número de enfermedades diagnosticadas por el médico	94,07	0,927 (0,878-0,976)	0,958 (0,926-0,990)
4.	Deterioro del desempeño laboral por enfermedad	85,93	0,747 (0,644-0,851)	0,958 (0,260-0,990)
5.	Días de ausencia por enfermedad (último año)	91,85	0,844 (0,757-0,930)	0,888 (0,815-0,960)
6.	Pronóstico de la capacidad de trabajo en dos años	90,40	0,603 (0,407-0,798)	0,587 (0,377-0,798)
7.	Recursos mentales	86,76	0,727 (0,612-0,842)	0,747 (0,638-0,858)
a.	Disfrute cotidiano	83,70	0,724 (0,623-0,826)	0,783 (0,695-0,871)
b.	Actividad y buen ánimo	85,19	0,699 (0,584-0,813)	0,693 (0,572-0,813)
c.	Optimismo sobre el futuro	84,40	0,766 (0,676-0,856)	0,840 (0,772-0,907)

El cuestionario ICT fue elaborado para ser un instrumento autoaplicable, por lo que requiere de una determinada capacidad de comprensión e interpretación de

las preguntas que lo componen. Se recomienda por los autores del cuestionario una escolaridad mínima de nivel primario para los participantes ¹, según resultados obte-

nidos en población finlandesa. En el presente estudio todos los participantes tienen vencido el noveno grado (nivel de enseñanza secundaria), mas se optó por una aplicación del cuestionario asistida por el investigador para asegurar una mayor comprensión de los términos de las preguntas. Este procedimiento evitó respuestas incompletas y preguntas sin respuesta, y se considera que no afectó el resultado de la confiabilidad, en tanto el mismo método fue adoptado en las dos aplicaciones. Tanto el procedimiento autoaplicado^{15,17} como el asistido¹⁶, han sido empleados al evaluar la confiabilidad del cuestionario ICT.

La tasa de completamiento de las etapas del estudio fue del 100%. Todos los trabajadores que realizaron la primera aplicación completaron también la segunda aplicación del cuestionario, por lo que los resultados obtenidos son robustos a los errores de pérdida selectiva de sujetos que pueden afectar este tipo de evaluación. Sí, es necesario tener en cuenta la negatividad de los sujetos de participar en estudios acerca del comportamiento de la capacidad de trabajo en colectivos laborales y trabajadores de distintos sectores, ya que puede verse afectada la comparabilidad de los resultados del ICT con los niveles y población de referencia.

La población estudiada se corresponde con personal que labora en instituciones hospitalarias del II y III nivel de atención de salud. El sector de la salud se caracteriza por un contenido de trabajo con predominio de exigencias de tipo mental, aunque en algunas profesiones y ocupaciones propias de este sector se observa una paridad en las exigencias físicas y mentales de trabajo. La mayoría de los sujetos calificó con una buena o excelente capacidad de trabajo (n=106; 77,9 %) en las dos aplicaciones del cuestionario ICT, un 22,05 % (n=30) se ubicó en las categorías de pobre y moderada capacidad de trabajo.

Los estudios acerca de la confiabilidad del cuestionario ICT se han realizado en poblaciones de trabajadores que laboran en diferentes sectores de la producción y los servicios. Las poblaciones estudiadas se corresponden con trabajadores de oficina¹⁹, metalmecánicos¹⁶, eléctricos¹³, de diferentes sectores productivos²¹ y trabajadores de la salud^{17,23}. En estos estudios se han utilizado diferentes métodos de evaluación de la confiabilidad, como la consistencia interna, estabilidad en el tiempo o bien la combinación de estos procedimientos.

Los resultados de nuestro estudio sugieren que hay estabilidad en la puntuación total del ICT que obtuvieron los sujetos y en la clasificación de los mismos en una categoría determinada de la capacidad de trabajo. Esto lo indica la ausencia de diferencias significativas en la puntuación obtenida en las preguntas que integran el cuestionario entre el test y el retest, y el hecho de que fueron buenos o excelentes los resultados del análisis de concordancia al examinar la clasificación de los sujetos en las categorías del ICT en las dos aplicaciones. Estos

resultados son similares a los encontrados por Renosto¹⁶, Zwart¹⁵ y Ma La ji²¹, en sus respectivos estudios.

En este estudio se observó que el 45,5 % de los participantes obtuvo igual calificación en las dos mediciones, resultado superior al valor encontrado por Renosto¹⁶, que fue del 37 %, y también al observado por de Zwart¹⁵ en el que cerca del 25 % de los sujetos estudiados presentó la misma puntuación en el test y el retest. El diagrama de Bland & Altman no mostró un patrón sistemático en la dispersión de las medidas del test y el retest. Al comparar la diferencia de las puntuaciones del ICT en el test y el retest contra la media de estas dos mediciones, se obtuvo que el 5,4 % (n=9) de los casos se ubicó fuera del intervalo de la media \pm 2DE (-3,8 y +3,8). Este resultado fue semejante al encontrado por Zwart¹⁵, donde el 5 % de los puntos se ubicaron fuera de este intervalo (-6,86 a +6,86) y de Almeida¹⁷, que observó también al 5 % de los puntos fuera del límite de la media \pm 2DE (-7,1 a +7,1).

La media de la diferencia entre las puntuaciones del ICT en las dos aplicaciones del cuestionario fue muy pequeña (-0,051), lo cual es indicativo de que no existe error de tipo sistemático en estos resultados que puedan atribuirse al instrumento *per se*. La desviación estándar de esta diferencia es de 1,9 puntos, que también es pequeña comparada con los valores de las puntuaciones obtenidas (rango de 26 hasta 49 puntos).

En el análisis de la estabilidad de la respuesta a cada pregunta del cuestionario, se aprecia que la mayor discrepancia en la respuesta recae en la pregunta 1, la cual pudiera ser explicada por el grado de subjetividad y generalidad que ésta encierra, y la influencia de determinadas situaciones o sentimientos que pudieron rodear a las personas al momento de ser entrevistadas. Renosto¹⁶, en una población de enfermeras, obtuvo resultados calificados como moderados a bajos del índice kappa ponderado en esta pregunta (0,39) [IC_{95%} (0,16-0,64)] y también para la pregunta 6, sobre el pronóstico de la capacidad de trabajo en dos años con valor de Kappa (0,39) [IC_{95%} (0,07-0,74)].

Se observa que cuanto más objetiva sea una pregunta, menor será el margen de error en la respuesta y, por tanto, mayor puede ser la concordancia entre las respuestas en mediciones repetidas. Por otra parte, resulta interesante que las preguntas que indagan acerca del número de enfermedades o los días de ausencia por enfermedad (preguntas 3 y 5) obtuvieran los mayores valores de acuerdo (Kappa) entre las dos mediciones. Los autores consultados refieren acuerdos más bajos para estas preguntas, ya que los resultados pueden verse afectados por el sesgo de memoria¹⁵⁻¹⁷.

En el análisis de la clasificación de los sujetos en las categorías del ICT en las dos aplicaciones, no se observaron cambios bruscos, es decir, en el rango de más de una categoría. Los encuestados se movieron solo en el rango de una categoría, ya fuera hacia una inferior o

superior a la observada en la primera aplicación. Otros autores refieren cambios más bruscos en sus estudios, aunque en un número muy reducido de sujetos^{15,17}. Los datos fueron obtenidos con un intervalo de 2 a 3 semanas, considerado un tiempo suficientemente prolongado como para que no se recordaran las respuestas a las preguntas del cuestionario de la primera aplicación. En este período pudieron ocurrir eventos que afectaran de forma positiva o negativa la capacidad de trabajo de los sujetos. Estos cambios pueden estar relacionados con la verdadera dinámica que sufre la capacidad de trabajo, la que puede verse afectada por eventos vitales, la percepción de salud, así como factores que actúan en el medio laboral^{14,32-34}.

Estos resultados son indicativos de que la versión cubana del cuestionario ICT que se prueba en el presente estudio, califica como un instrumento confiable para los fines de evaluación de la capacidad de trabajo.

El estudio realizado indica que la versión cubana del Índice de capacidad de trabajo posee una adecuada confiabilidad en términos de la estabilidad de la información que deriva de su aplicación en una población de trabajadores de la salud. Los resultados del análisis de confiabilidad test-retest pueden ser considerados aceptables para la medición del ICT en escala cuantitativa, así como para su clasificación en escala ordinal. Las respuestas a las preguntas que integran el ICT también mostraron una adecuada confiabilidad. Los resultados obtenidos son consistentes con la evidencia publicada en relación con la calidad del proceso de medición de la capacidad de trabajo mediante el cuestionario ICT, en poblaciones laborales de diferentes países, lo cual respalda la experiencia acerca de su aplicabilidad en el quehacer investigativo y la práctica cotidiana de la salud ocupacional.

BIBLIOGRAFÍA

1. Tuomi K, Ilmarinen J, Jahkola M, Katajarinne L, Tulkki A. Work ability index. First reprint of the 2nd revised edition. Helsinki: Finnish Institute of Occupational Health; 2006.
2. Valkonen J, Työkyky T. Työkykyko kemuksen narratiivista tarkastelua [Work ability as stories. Narrative analysis of the perception of work ability]. Helsinki Rehabilitation Foundation, Working papers; 1997. Disponible en: http://www.ttl.fi/en/organization/work_organizations/innovations_management_and_knowledge_creation/pages/default.aspx.
3. Dubois C-A, McKee M. Cross-national comparisons of human resources for health: What can we learn? Health Economics Policy and Law. 2006;59-78.
4. Tuomi K, Ilmarinen J, Seitsamo J, Huuhtanen P, Martikainen R, Nygard C-H, et al. Summary of the Finnish research project (1981-1992) to promote the health and work ability of aging workers. Scand J Work Environ Health. 1997;23(1):66-71.
5. Tuomi K, Ilmarinen J, Eskelinen I, Jarvinen E, Toikkanen J, Klockars M. Prevalence and incidence of diseases and work ability in different work categories of municipal occupations. Scand J Work Environ Health. 1991;17(1): 67-74.
6. Tuomi K, Wagarv G, Eskelinen L. Health, work capacity and work conditions in municipal occupations. Scand J Work Environ Health. 1985;3: 95-132.
7. Seitsamo J, Ilmarinen J, Life style, aging and work ability among active Finnish workers in 1981-1992. Scand J Work Environ Health. 1997;23(1):20-6.
8. Eskelinen L, Kohvakka A, Marisalo T, Hurri H, Wagar G. Relationship between the self-assessment and clinical assessment of health status and work ability. Scand J Work Environ Health. 1991;17:40-7.
9. Tuomi K, Huuhtanen P, Nykyri E, Ilmarinen J. Promotion of work ability, the quality of work and retirement. Occup Med. 2001;51(5):318-24.
10. Mikaela B, Bonsdorff P, Seitsamo P, Ilmarinen J, Nygard C-H, von Bondorff ME, et al. Work ability in midlife as predictor of mortality and disability in later life: a 28-year prospective follow-up study. CMAJ. 2011;183(4):235-42.
11. Radkiewicz P, Widerszal-Bazy TML. Psychometric properties of Work Ability Index in the light of comparative survey study. International Congress Series. 2005;1280:304-9.
12. Harkonmaki K, Rahkonen O, Martikainen P, Silventoinen K, Lahelma E. Associations of SF-36 mental health functioning and work and family related factors with intentions to retire early among employees. Occup Environ Med. 2006;63:558-63.
13. Martinez MC, Latorre MRDO. Saúde e capacidade para o trabalho de eletricitários do Estado de São Paulo. Ciência & Saúde Coletiva. 2008;13(3):1061-73.
14. Alavinia SM, de Boer AG, van Duivenbooden C, Frings-Dresen MH, Burdorf A. Determinants of work ability and its predictive value for disability. Occup Med. 2009;59(1):32-7.
15. Zwart M, Frings W. Test-retest reliability of the Work Ability Index questionnaire. Occup Med. 2002;52:177-81.
16. Renosto, A, Biz P, Hennington EA, Pattussi MC. Confiabilidade teste-reteste do Índice de Capacidade para o Trabalho em trabalhadores metalúrgicos do Sul do Brasil. Rer Bras Epidemiol. 2009;12(2):217-25.
17. Almeida da Silva SH, Godoi AG, Harter R, Rotenberg L. Validade e confiabilidade do índice de capacidade para o trabalho (ICT) em trabalhadores de enfermagem. Cad. Saúde Pública. 2011;27(6):1077-87.
18. López GM, del Castillo NP, Oramas A. Validez de

- contenido y confiabilidad del cuestionario Índice de Capacidad de Trabajo en su versión cubana. *Revista Cubana de Salud y Trabajo*. 2011;12(1):29-34.
19. Sjögren-Rönkä T, Ojanen M, Leskinen E, Mustalampi S, Mälkiä E. Physical and psychosocial prerequisites of functioning in relation to work ability and general subjective well-being among office workers. *Scand J Work Environ Health*. 2002;28(3):184-90.
 20. Radkiewicz P, Widerszal-Bazy TML. Psychometric properties of Work Ability Index in the light of comparative survey study. *International Congress Series*. 2005;1280:304-9.
 21. Ma La ji, Zhou Tong, Jin Tai yi, Shen Guang zhu, Ji Xi peng. [Reliability and Validity of Work Ability Index in Chinese Version]. Shanghai. *Journal of Labour Medicine*. 2000;02 (abstract in English).
 22. Zhang L, Wang ZM, Wang MZ. [Reliability and validity of Comprehensive Work Ability Index scales]. *Zhonghua*.
 23. Fischer FM, Borges FN, Rotenberg L, Latorre MR, Soares NS, Rosa PL, et al. Workability of health care shift workers: what matters? *Chronobiol Int*. 2006; 23(6):1165-79.
 24. Martinez MC, Latorre MRDO, Fischer FM. Validade e confiabilidade da versão brasileira do Índice de Capacidade para o Trabalho. *Rev. Saúde Pública*. 2009;43(3):525-32.
 25. Pasquali L. *Psicometria: Teoria dos testes na Psicologia e na Educação*. Rio de Janeiro: Vozes, 2009.
 26. Shover MF, Frederick GC. Does conversational interviewing reduce survey measurement error? *Public Opinion Quarterly*. 1997;61:576-602 [citado 1 Nov 2010]. Disponible en: <http://www.poq.oxfordjournals.org>.
 27. Gould R, Ilmarinen J, Järvisalo J, Koskinen S, eds. *Dimensions of work ability: results of the Health 2000 Survey*. Vaasa: Waasa Graphics Oy; 2008.
 28. Bland JM, Altman DG. Validating scales and indexes. *Br Med J*. 2002;324:606-7.
 29. Cohen J. A coefficient of agreement for nominal scales. *Educational Psychology Measurement*. 1960;20:46-73.
 30. Cerda J, Villarroel L. Evaluación de la concordancia inter-observador en investigación pediátrica: Coeficiente de Kappa. *Rev Chil Pediatr*. 2008;79(1): 54-8.
 31. Fleiss JL, Cohen J. The equivalence of weighted Kappa and the intraclass correlation coefficient as measures of reliability. *Education Psychology Measuring*. 1973;33:613-9.
 32. Alavinia SM, Duivenbooden JC, Burdorf A. Influence of work-related factors and individual characteristics on work ability among Dutch construction workers. *Scand J Work Environ Health*. 2007;33: 351-7.
 33. Makowiec – Dabrowska T, Rossada W, Bortkizwicz A, et al. Occupational and non occupational determinants of work ability. *Med Pr*. 2008;59(1):9-24.
 34. van den Berg TIJ, Elders TM, de Zwart BCH, Burdorf A. The effects of work-related and individual factors on the Work Ability Index: a systematic review. *Occup Environ Med*. 2009;66:211-220.

Recibido: 7 de junio de 2012

Aprobado: 16 de febrero de 2013