

11227
39



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE MEDICINA

DIVISION DE ESTUDIOS DE POSGRADO

REPRODUCIBILIDAD INTRA E INTEROBSERVADOR
DE LA ESCALA DE KARNOFSKY

T E S I S

INSTITUTO NACIONAL DE LA QUE PARA OBTENER EL GRADO DE:
ESPECIALISTA EN MEDICINA INTERNA
SUBDIRECCION DE ENSEÑANZA
MEXICO, D. F.

PRESENTA:

DRA. CRISTINA MARTÍNEZ SIBAJA

TUTOR:

DRA. BLANCA HERNÁNDEZ CRUZ.

MÉXICO, D.F., FEBRERO DEL 2000

276089



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Gracias a mis padres y hermanos
por su amor y apoyo incondicional.

Y gracias a Blanca, por su ayuda
y su paciencia.

ÍNDICE

1 - INTRODUCCIÓN	
Antecedentes	1
Diferentes tipos de instrumentos clinimétricos	3
Reproducibilidad	4
¿Cómo se mide la reproducibilidad?	6
La escala de Karnofsky	8
2 - JUSTIFICACIÓN	9
3 - HIPÓTESIS	10
4 - ESCENARIO	10
5 - OBJETIVOS	10
6 - MÉTODOS	
Pacientes	11
Evaluadores	11
Diseño	12
Análisis estadístico	13
Consideraciones éticas	14
7 - RESULTADOS	
Características de los pacientes	15
Características de los observadores	15
Reproducibilidad interobservador	17
Reproducibilidad intraobservador	18
Representación de la reproducibilidad en la forma propuesta por Bland y Altman	19
8 - DISCUSIÓN	20
9 - CONCLUSIONES	25
10 - TABLAS	26
11 - FIGURAS	29
12 - APÉNDICES	46
13 - BIBLIOGRAFÍA	49

REPRODUCIBILIDAD INTRA E INTEROBSERVADOR DE LA ESCALA DE KARNOFSKY

1 - INTRODUCCIÓN.

A) ANTECEDENTES

Desde que se iniciaron los estudios clínicos para valorar el efecto de los diferentes agentes utilizados en el tratamiento del cáncer, se hizo evidente la necesidad de contar con algún método que sirviera para valorar el estado de salud de los pacientes antes y después de someterse al tratamiento

Después de la segunda guerra mundial, se inició la era moderna de la quimioterapia del cáncer. En parte, como resultado de los estudios observacionales que se realizaron durante ese periodo sobre las consecuencias de la detonación de explosivos en las personas expuestas. Por ejemplo, uno de los hallazgos que presentaron las personas normales expuestas a la explosión de mostaza nitrogenada, fue la aparición de leucopenia. Basados en ello, se empezó a utilizar a la mostaza nitrogenada y sus derivados en el tratamiento de las enfermedades linfoproliferativas. Un médico militar, David A. Karnofsky, después de recopilar la información publicada por otros médicos militares al respecto, se interesó en el estudio de los efectos de los posibles fármacos antitumorales y estableció junto con otros médicos el primer programa de quimioterapia organizado (1).

En 1948, David A. Karnofsky y Joseph H. Burchenal, describieron una escala numérica que tenía como propósito la valoración clínica de la respuesta a los agentes quimioterapéuticos en pacientes con cáncer (1-4)

A partir de su publicación en 1949, la escala numérica antes mencionada (llamada desde entonces escala de Karnofsky), ha sido el método más utilizado en el mundo para clasificar el estado funcional de los pacientes con cáncer (5)

En los ensayos clínicos sobre tratamiento del cáncer, el estado funcional ha mostrado ser un predictor importante de la supervivencia y de la respuesta al tratamiento (5,6)

La escala de Karnofsky se ha usado en la investigación en cáncer, para la estratificación de diferentes poblaciones de pacientes que serán expuestos a diferentes intervenciones (terapéuticas y / o paliativas) como una medida del desenlace, comparando las diferencias en las capacidades funcionales de los pacientes, antes y después de la exposición a la intervención (5,7-15)

No está demostrada la utilidad de la escala de Karnofsky en pacientes sin cáncer, sin embargo, es frecuente encontrar en la literatura el empleo de la escala en pacientes sin cáncer (16-23)

Por otro lado, existen informes de estudios en los que se ha utilizado la escala para evaluar calidad de vida (18,21,22), no obstante, la escala no fue diseñada para ello ni se ha validado su desempeño como instrumento para tal fin

A pesar del amplio uso de la escala de Karnofsky, su reproducibilidad, es decir, la capacidad del instrumento de medir siempre lo mismo cuando se aplica a un mismo sujeto en igualdad de circunstancias, fue asumida durante muchos años sin ser sometida a una investigación formal (2), y a la fecha, existe poca información disponible en la que se investigue este aspecto de la escala (2,4,5).

B) DIFERENTES TIPOS DE INSTRUMENTOS CLINIMÉTRICOS

Los índices clinimétricos, son los métodos cuantitativos que se utilizan en la colección y el análisis comparativo de los datos clínicos obtenidos por los médicos (24). En las últimas décadas, ha cobrado impulso el desarrollo de instrumentos clinimétricos que valoren adecuadamente el estado de salud (índices de salud), estos últimos inicialmente se referían solo a desempeño físico (como la escala de Karnofsky), pero en años recientes, se ha puesto particular interés en desarrollar índices que valoren calidad de vida (en la que necesariamente la evaluación del desempeño físico está incluida) (24-26)

De acuerdo con su aplicación potencial, los índices que valoran el estado de salud se pueden dividir en 3 categorías

1) Los índices discriminativos, que son los que se utilizan para distinguir entre individuos o grupos de individuos, cuando no existe un criterio externo o un estándar de oro disponible para validar estas mediciones. Por ejemplo las pruebas que se utilizan para medir inteligencia

2) Los índices predictivos, que se utilizan para clasificar a los individuos en categorías predefinidas cuando un estándar de oro o un desenlace inequívoco está disponible, ya sea en forma concurrente o prospectiva, para determinar si los individuos han sido clasificados en forma correcta

Este tipo de índices se utilizan generalmente como un instrumento de escrutinio o un instrumento de diagnóstico, para identificar cuáles individuos tienen o desarrollarán una condición o desenlace específico. La escala de Karnofsky corresponde a esta categoría

Esta escala se utiliza para evaluar el pronóstico de pacientes con cancer, y el desenlace inequívoco en estos casos es la muerte (para analizar la sobrevida)

La utilidad de los índices predictivos se basa en la capacidad que tienen para anticipar un desenlace (predecir), con mayor precisión de lo que se puede explicar únicamente por azar. Por lo tanto, la reproducibilidad de estos índices debe ser demostrada con un método para medir acuerdo corregido para el azar (6,28)

3) Los índices evaluativos, que se utilizan para valorar la magnitud de un cambio longitudinal en un individuo o un grupo de individuos

El desarrollo de los instrumentos evaluativos constituye el principal foco de atención para aquellos investigadores interesados en medir calidad de vida. Estos instrumentos se necesitan para cuantificar el beneficio del tratamiento en los ensayos clínicos y para medir calidad ajustada de los años de vida en análisis de costos (6)

Para que un índice clinimétrico resulte satisfactorio, debe contar con ciertas características

Debe ser

- a) Reproducible
- b) Válido, es decir, el instrumento debe medir lo que se quiere medir
- c) Sensible al cambio
- d) Fácil de usar (6,26,27)

C) REPRODUCIBILIDAD

Es uno de los requisitos estándar que debe cubrir un instrumento clinimétrico para que sea útil para propósitos clínicos.

La reproducibilidad es una propiedad básica que indica que se está midiendo algo en una forma consistente y se refiere al grado en el que se pueden obtener los mismos resultados bajo las mismas condiciones en pacientes estables (27) Dicho de otra forma, la reproducibilidad es la capacidad de un instrumento de medir siempre lo mismo cuando se aplica a un mismo sujeto en igualdad de circunstancias

Diferentes autores han utilizado algunos sinónimos para este término, lo que puede generar algún grado de confusión entre los lectores. Consistencia, repetibilidad, reproducibilidad y precisión, son los términos más frecuentemente utilizados

La reproducibilidad interobservador se refiere al acuerdo que existe entre las evaluaciones realizadas por diferentes observadores, a un mismo sujeto o a una misma variable. Los observadores valoran al sujeto o la variable en estudio en forma independiente y lo que se busca idealmente, es que se clasifique aquello que fue analizado siempre en la misma categoría (4-6,24,25,29)

La reproducibilidad intraobservador se refiere al grado en el que un mismo sujeto o una misma variable pueden ser clasificados en la misma forma por un mismo observador en dos o más ocasiones diferentes (25,29).

La reproducibilidad de un instrumento clínico, puede ser afectada por diferentes factores que están en relación al instrumento utilizado, al observador, al evaluado y a las condiciones en las que se llevan a cabo las valoraciones

Como fuentes de error en relación al instrumento utilizado podemos citar una pobre definición de los elementos que constituyen las categorías dentro de una escala y las instrucciones no estandarizadas, por ejemplo (4)

Los factores que pueden afectar la reproducibilidad en relación al observador u observadores son por ejemplo, el entrenamiento en el uso del instrumento, los años de entrenamiento clínico, el cansancio o estado de ánimo del observador al aplicar el instrumento y la subjetividad en la medición (27, 30).

Como fuentes de error en relación al evaluado podemos citar la colaboración del sujeto en estudio, las preferencias personales del mismo (al interrogar a un paciente, este puede dar mayor importancia a un punto que a su parecer es más valioso y minimizar otro que no considera tan importante para él), y la fluctuación en la condición del enfermo (25, 27)

Por último, en relación a las condiciones en las que se llevan a cabo las evaluaciones, la reproducibilidad puede ser afectada por la cantidad de luz, de ruido, y la ventilación en el sitio donde se realiza la valoración (27)

¿Cómo se mide la reproducibilidad?

El índice de elección para medir reproducibilidad en escalas nominales (en las que se expresan por ejemplo color del cabello o de piel, etc) es el estadístico kappa (k), que evalúa el acuerdo entre los observadores y lo corrige para el acuerdo por el azar

La kappa se usa ordinariamente para medir acuerdo entre dos observadores. Si más de dos observadores se están comparando, los índices de kappa se pueden calcular por el acuerdo entre diferentes pares, por ejemplo, para A vs. B, B vs. C, y A vs. C, etc (Después se calcula la concordancia que existe entre diferentes pares - k promedio -).

El índice de elección para medir acuerdo cuando se analizan datos ordinales (los que se expresan en categorías o grados, como es el caso de la escala de Karnofsky), es la kappa ponderada (k_w), la cual se deriva de kappa. Cuando se utiliza la kappa ponderada se asignan pesos que se basan en las magnitudes de los desacuerdos observados.

Al igual que en el caso de la kappa no ponderada, los valores de acuerdo van de -1 (máximo desacuerdo) a $+1$ (máximo acuerdo), pasando por el cero, que representa un acuerdo observado igual al que se puede esperar únicamente por azar (28).

Decidir un nivel cuantitativo de significancia para el valor de k y k_w es algo arbitrario.

Landis y Koch han sugerido las siguientes guías:

Valor de k y k_w	Fuerza del acuerdo
< 0	Pobre
0-0.20	Mínimo
0.21-0.40	Ligero
0.41-0.60	Moderado
0.61-0.80	Substancial
0.81-1.00	Casi perfecto

Si el nivel de entrenamiento entre los observadores es comparable, el acuerdo que es apenas mejor que el que se explica por azar, es inadecuado. La k_w debe probablemente ser mayor a 0.5 para que se pueda considerar un grado de acuerdo aceptable (28).

D) LA ESCALA DE KARNOFŠKY.

La escala clasifica a los pacientes en 3 grandes grupos designados con letras, de acuerdo a la capacidad física de los pacientes para: A) trabajar y llevar una vida normal, B) para el cuidado de su persona (con un grado variable de asistencia) y en el grupo C se encuentran los pacientes gravemente incapacitados. Estos 3 grupos se dividen a su vez en once categorías que cubren los posibles niveles de función, desde el paciente completamente normal (al que se le asigna una calificación de 100 puntos), hasta la muerte (con una calificación de 0 puntos), tabla 1

En 1979 Hutchinson y colaboradores publicaron un estudio (4), en el que investigaron la reproducibilidad interobservador entre dos médicos asignados a un servicio de urgencias que valoraron con base en la escala de Karnofsky a 29 pacientes consecutivos que llegaron al servicio (grupo 1), y entre 2 nefrólogos que valoraron a 31 pacientes que llegaron en forma consecutiva programados para hemodiálisis (grupo 2), en el mismo hospital. Los pacientes fueron evaluados en forma independiente, el mismo día. En este estudio se obtuvo un acuerdo del 35% en el grupo 1 y 29% en el grupo 2, cifras que no fueron corregidas para el azar. Con este hallazgo de acuerdo tan bajo, los investigadores decidieron analizar el acuerdo con base en el grupo en el que quedaban incluidos los pacientes en relación al puntaje que se les había asignado (grupos A, B y C de la escala de Karnofsky). Se utilizó para el análisis el método de kappa. El acuerdo en el grupo 1 fue 72%, con un valor de kappa de 0.50, en el grupo 2 el acuerdo fue de 71%, con un valor de kappa de 0.46.

En 1980 Yates y colaboradores publicaron un estudio en el que investigaron la reproducibilidad interobservador entre una enfermera y una trabajadora social entrenadas en el uso de la escala de Karnofsky. Las dos observadoras valoraron a 50 pacientes con cáncer que se encontraban hospitalizados. Los pacientes fueron evaluados en forma independiente y con un tiempo máximo de 1 semana en la valoración de cada paciente por ambos observadores. En este estudio el análisis se llevó a cabo mediante el coeficiente de correlación de Pearson, se obtuvo un resultado de 0.69 que fue interpretado como un grado moderado de reproducibilidad interobservador (2).

En otro estudio publicado por Mor y colaboradores en 1984, se investigó la reproducibilidad interobservador entre 47 evaluadores (no médicos) entrenados en el uso de la escala de Karnofsky, quienes asignaron un puntaje con base en este instrumento a 17 ejemplos hipotéticos de pacientes que cubrían todo el espectro de la escala. Los resultados se analizaron utilizando el coeficiente alfa de Cronbach. El coeficiente de reproducibilidad interobservador fue 0.97, resultado que se encontró estadísticamente significativo (5).

2 - JUSTIFICACIÓN

En el Instituto Nacional de la Nutrición Salvador Zubirán (INNSZ), el uso de la escala de Karnofsky es muy frecuente. No contamos con estudios donde se investigue la reproducibilidad de esta escala de clasificación del estado funcional y los informes en la literatura mundial son escasos y sus resultados no son comparables (2,4,5).

Como ya se mencionó, la reproducibilidad de un índice clinimétrico es uno de los elementos clave en la evaluación de su utilidad. Todo lo anterior nos motivó a investigar la reproducibilidad de la escala en nuestro medio.

3 - HIPOTESIS ALTERNA

La escala de Karnofsky no es un instrumento reproducible.

4 - ESCENARIO

El INNSZ cuenta con un centro de atención médica de tercer nivel, un hospital en el que se imparte enseñanza de posgrado en diferentes especialidades médicas y en subespecialidades derivadas de la medicina interna (por ejemplo oncología, hematología, gastroenterología, nefrología, endocrinología, reumatología y geriatría).

Acuden al Instituto muchos enfermos con diferentes padecimientos oncológicos, en 1997 hubo un total de 943 ingresos de pacientes con diagnóstico de tumores malignos (incluyendo neoplasias hematológicas) en los servicios de consulta externa, sectores de hospitalización, urgencias, estancia corta y terapia intensiva. En estos pacientes se puede aplicar (y frecuentemente se utiliza) la escala de Karnofsky, para evaluar el estado funcional de los enfermos que serán sometidos a una intervención terapéutica o paliativa, o como orientación a cerca de la sobrevivencia en un caso particular.

5 -OBJETIVOS

a - Evaluar la reproducibilidad intra e interobservador de la escala de Karnofsky en el Instituto Nacional de la Nutrición Salvador Zubirán.

Como ya se mencionó, la reproducibilidad de un índice clinimétrico es uno de los elementos clave en la evaluación de su utilidad. Todo lo anterior nos motivó a investigar la reproducibilidad de la escala en nuestro medio.

3 - HIPOTESIS ALTERNA

La escala de Karnofsky no es un instrumento reproducible

4 - ESCENARIO

El INNSZ cuenta con un centro de atención médica de tercer nivel, un hospital en el que se imparte enseñanza de posgrado en diferentes especialidades médicas y en subespecialidades derivadas de la medicina interna (por ejemplo oncología, hematología, gastroenterología, nefrología, endocrinología, reumatología y geriatría)

Acuden al Instituto muchos enfermos con diferentes padecimientos oncológicos, en 1997 hubo un total de 943 ingresos de pacientes con diagnóstico de tumores malignos (incluyendo neoplasias hematológicas) en los servicios de consulta externa, sectores de hospitalización, urgencias, estancia corta y terapia intensiva. En estos pacientes se puede aplicar (y frecuentemente se utiliza) la escala de Karnofsky, para evaluar el estado funcional de los enfermos que serán sometidos a una intervención terapéutica o paliativa, o como orientación a cerca de la sobrevida en un caso particular

5 -OBJETIVOS

a - Evaluar la reproducibilidad intra e interobservador de la escala de Karnofsky en el Instituto Nacional de la Nutrición Salvador Zubirán

Como ya se mencionó, la reproducibilidad de un índice clinimétrico es uno de los elementos clave en la evaluación de su utilidad. Todo lo anterior nos motivó a investigar la reproducibilidad de la escala en nuestro medio.

3 - HIPOTESIS ALTERNA

La escala de Karnofsky no es un instrumento reproducible.

4 - ESCENARIO

El INNSZ cuenta con un centro de atención médica de tercer nivel, un hospital en el que se imparte enseñanza de posgrado en diferentes especialidades médicas y en subespecialidades derivadas de la medicina interna (por ejemplo oncología, hematología, gastroenterología, nefrología, endocrinología, reumatología y geriatría).

Acuden al Instituto muchos enfermos con diferentes padecimientos oncológicos, en 1997 hubo un total de 943 ingresos de pacientes con diagnóstico de tumores malignos (incluyendo neoplasias hematológicas) en los servicios de consulta externa, sectores de hospitalización, urgencias, estancia corta y terapia intensiva. En estos pacientes se puede aplicar (y frecuentemente se utiliza) la escala de Karnofsky, para evaluar el estado funcional de los enfermos que serán sometidos a una intervención terapéutica o paliativa, o como orientación a cerca de la supervivencia en un caso particular.

5 -OBJETIVOS

a - Evaluar la reproducibilidad intra e interobservador de la escala de Karnofsky en el Instituto Nacional de la Nutrición Salvador Zubirán.

b- Conocer las diferencias que puedan existir en la reproducibilidad intra e interobservador, cuando se comparan médicos con diferente nivel de entrenamiento, en especialidades relacionadas con padecimientos oncológicos (oncología, hematología, medicina interna, geriatría) en el INNSZ

6 -METODOS

A) PACIENTES

Se incluyeron en el estudio todos los pacientes adultos con diagnóstico de algún padecimiento oncológico o hematológico, que se encontraban hospitalizados en sectores de internamiento (piso), en el Instituto Nacional de la Nutrición Salvador Zubirán en la ciudad de México, los días 10 y 23 de septiembre de 1997.

Se excluyeron aquellos pacientes que tenían programados estudios invasores o cirugía el día de la evaluación, los pacientes que se encontraban en posoperatorio el día del estudio (que habían sido operados en el mismo internamiento), los pacientes en las primeras 48 horas en tratamiento con quimioterapia, o los que tenían varios días de tratamiento con quimioterapia y presentaron muchos efectos secundarios (ya que esto podía hacer variar su condición en las horas que los pacientes serían evaluados por los diferentes observadores), y los pacientes con neutropenia grave que se encontraban en cuartos aislados

B) EVALUADORES.

Participaron en total 10 evaluadores: una médica de base de oncología quien egresó de la subespecialidad de hemato-oncología en marzo de 1995, una hematóloga que egresó de la subespecialidad en marzo de 1997 y que cursaba al participar en el estudio un año de

medicina transfusional, una médica de base de hematología que había egresado de la subespecialidad en marzo de 1997, un residente de primer año de hematología; tres residentes del cuarto año de medicina interna; dos residentes de geriatría, una de primer año y la otra de segundo año de la subespecialidad y un residente del segundo año de oncología

Con excepción del observador que cursaba medicina transfusional, quien realizó solo un año de medicina interna antes de ingresar a la subespecialidad de hematología, todos los demás evaluadores completaron un mínimo de tres años de medicina interna.

Todos los observadores participaron en forma preliminar (varios días antes), en una valoración del conocimiento del instrumento a utilizar, que consistió en asignar un puntaje con base en la escala de Karnofsky a una serie de pacientes hipotéticos descritos en un cuestionario. Después, se analizaron las respuestas, se proporcionó a cada médico una copia de la escala de Karnofsky traducida al español, y se aclararon las dudas que surgieron al respecto

C) DISEÑO

Estudio observacional con prueba repetida

D) EVALUACION DE LA REPRODUCIBILIDAD.

Evaluamos la reproducibilidad de la escala de Karnofsky, determinando el grado de acuerdo intra e interobservador.

Para investigar el acuerdo intraobservador, solicitamos la evaluación de cada paciente con base en la escala de Karnofsky en dos ocasiones diferentes por un mismo observador, permitiendo un lapso de tiempo de 2 horas como mínimo y 12 horas como máximo entre la primera y la segunda evaluación. Ambos periodos de tiempo fueron escogidos en forma arbitraria, tomando en consideración que en un tiempo mayor a 12 horas la posibilidad de que ocurriera una variación en el estado clínico de un paciente hospitalizado era alta, y esto afectaría la reproducibilidad. Por otra parte, al escoger el tiempo mínimo entre la primera y la segunda evaluación de cada enfermo, nos vimos limitados en gran medida por el horario de trabajo de los diferentes observadores. Idealmente el tiempo mínimo debería haber sido mayor pero al estudiar los diferentes horarios de los participantes esto no fue posible.

Para investigar el acuerdo interobservador, solicitamos la evaluación de cada paciente con base en la escala de Karnofsky por los 10 observadores, permitiéndose como máximo un lapso de 12 horas de diferencia en la evaluación de cada paciente por todos los observadores.

Para realizar la valoración se proporcionaron a cada observador unas hojas diseñadas especialmente para ese propósito (apéndice 1).

E) ANÁLISIS ESTADÍSTICO

Se utilizaron promedios, desviación estándar y límites para describir a los pacientes.

Para analizar la reproducibilidad se empleó el índice de acuerdo de kappa ponderada con pesos cuadráticos.

El cálculo de kappa ponderada se basa en la proporción de desacuerdo, su fórmula es la siguiente

$$k_w = 1 - (q'o / q'c)$$

Donde $q'o$ = a la proporción de desacuerdo ponderado observado, y $q'c$ = a la proporción de desacuerdo ponderado explicado por azar

Como se explicó en otro apartado, cuando se comparan más de dos observadores, los índices de kappa se pueden calcular por el acuerdo entre diferentes pares, por ejemplo para A vs B, A vs C, A vs D, B vs C, B vs D, etc., y después se calcula la concordancia que existe entre los diferentes pares (kappa promedio)

También se construyeron gráficas para presentar la reproducibilidad en la forma descrita por Bland y Altman (31-33) En estas gráficas de dispersión se marca el valor promedio de las dos mediciones a analizar (representado sobre el eje X) contra la diferencia que existe entre dichas mediciones (mostrada sobre el eje Y), además se muestran líneas que marcan el promedio de las diferencias y otras que señalan los límites de acuerdo, estos últimos se calculan con el promedio de las diferencias ± 2 desviaciones estándar.

F) CONSIDERACIONES ÉTICAS

Se informó a cada paciente en forma oral y por escrito las características del estudio y se explicó en que consistiría la participación solicitada. Después de haber sido informados, se solicitó consentimiento por escrito de todos los pacientes que aceptaron participar (apéndice 2)

7 - RESULTADOS

A) CARACTERÍSTICAS DE LOS PACIENTES.

Se incluyeron 25 enfermos 4 de ellos tenían hepatocarcinoma (16%), 3 tenían cáncer de pancreas (12%), 3 cáncer de mama (12%), 2 cáncer de ovario (8%), 2 cáncer renal (8%), 2 sarcoma (8%), 2 cáncer de estómago (8%), 1 cáncer de colédoco (4%), 1 adenocarcinoma de yeyuno (4%), 1 cáncer de colon (4%), 1 melanoma (4%), 1 tumor carcinoide duodenal (4%), 1 carcinoma de pulmón (4%), y 1 paciente con linfoma no Hodgkin (4%), figura 1

Solamente hubo un paciente con diagnóstico hematológico, ya que los demás pacientes con diagnóstico de neoplasia hematológica hospitalizados en los días que se llevó a cabo el estudio, estaban en cuarto aislado con neutropenia grave y fiebre (uno de los criterios de exclusión)

La edad promedio de los pacientes fue de 52.4 años (desviación estándar -SD- 16.98, límites 18-73), el tiempo promedio de la evolución de la enfermedad neoplásica fue 16.1 meses (SD 23, límites 1- 84meses), tabla 2 Trece pacientes eran varones (52%), figura 2

B) CARACTERÍSTICAS DE LOS OBSERVADORES

Los diez observadores eran médicos del Instituto Nacional de la Nutrición, y su categoría se describió en el apartado de material y métodos. Cinco observadores eran mujeres En total nueve observadores participaron los dos días que se llevó a cabo el estudio (el residente de segundo año de oncología -observador diez- no pudo participar en la segunda etapa por motivos de salud, y evaluó solamente a los 14 pacientes incluídos en la primera etapa).

Seis observadores la médica de base de oncología (observador uno), la hematóloga que cursaba medicina transfusional (observador dos), 2 residentes de cuarto año de medicina interna (observadores cinco y seis) y las dos geriatras (observadores ocho y nueve), realizaron dos evaluaciones en cada paciente para medir la reproducibilidad intraobservador. El tiempo mínimo transcurrido entre la primera y la segunda evaluación fue de dos horas y el máximo de siete horas. Veintiun pacientes fueron valorados dos veces por los seis observadores, en los cuatro pacientes restantes (16%) faltó la segunda valoración por uno o más de los 6 observadores .

Los nueve observadores que participaron los dos días que se realizó el estudio evaluaron una vez a los pacientes para investigar la reproducibilidad interobservador. Veintitres pacientes fueron valorados por los nueve observadores, de los otros dos pacientes (8%) uno fue valorado por ocho observadores (no fue evaluado por el residente de medicina interna que valoró una sola vez a los pacientes -observador 7-), el otro enfermo fue valorado por siete observadores (no fue evaluado por el residente de primer año de hematología -observador 4-, ni por la residente de segundo año de geriatría -observador 8-). El tiempo máximo promedio transcurrido entre la evaluación de cada paciente por los diferentes observadores fue de 6 horas con 52 minutos (SD 0 49, límites 4 horas con 30 minutos- 7 horas con 50 minutos)

C) RESULTADOS DE LAS EVALUACIONES

Los puntajes asignados a los pacientes con base a la escala de Karnofsky variaron de 30 a 100, con un valor promedio de 75 2 (SD 18,9) Figuras 3 - 11.

Para analizar si existieron tendencias entre los diferentes observadores en la forma de asignar los puntajes (es decir, para investigar si hubo observadores que calificaron sistemáticamente con puntajes más altos o más bajos que otros), elaboramos la tabla 3, en la que se muestran las medianas de las calificaciones asignadas por cada observador, en la primera y segunda evaluación de los 25 pacientes incluidos en el estudio, y la tabla cuatro, en la que se muestra la mediana de las calificaciones por grupos (hematólogos, internistas, geriatras), en la primera y segunda valoración de los enfermos. Como puede verse en las tablas, no se encontraron tendencias que fueran significativas entre los observadores ni entre los grupos.

1 - REPRODUCIBILIDAD INTEROBSERVADOR

La kappa ponderada promedio de los 9 observadores que participaron los 2 días del estudio fue 0.52, con una mediana de 0.51 y límites de 0.37 a 0.66 (Dado que fueron 9 observadores, se formaron 36 posibles pares de calificaciones por cada paciente evaluado, en total obtuvimos 877 pares de 900 teóricamente posibles). Para el análisis de la reproducibilidad interobservador, se utilizó el resultado de la primera evaluación en el caso de los observadores que calificaron 2 veces a los pacientes (observadores 1, 2, 5, 6, 8 y 9) así como la evaluación única de los observadores 3, 4 y 7.

La kappa ponderada entre los hemato-oncólogos considerados como grupo (la hemato-oncóloga, las dos hematólogas y el residente de hematología, observadores 1, 2, 3 y 4), fue de 0.59, con una mediana de 0.59 y límites de 0.48 a 0.66 (En este caso, con 4 observadores, se forman 6 posibles pares de calificaciones para cálculo de kappas, por cada

paciente evaluado. En total obtuvimos 147 pares de 150 teóricamente posibles). Si se incluye en el análisis el residente de oncología que participó únicamente el primer día del estudio (observador 10), la kappa ponderada del grupo de hemato-oncólogos es de 0.56, con una mediana de 0.58 y límites de 0.38 a 0.66.

La kappa ponderada entre los 3 residentes de medicina interna (observadores 5, 6 y 7) fue de 0.42, con una mediana de 0.42 y límites de 0.37 a 0.46. (Con los internistas -3-, conseguimos formar 73 pares de calificaciones, de 75 teóricamente posibles).

La kappa ponderada (de 24 pares de calificaciones) entre las dos residentes de geriatría (observadores 8 y 9) fue de 0.41, con límites de 0.37 a 0.46. Figura 12.

2 - REPRODUCIBILIDAD INTRA-OBSERVADOR

El acuerdo entre las dos evaluaciones realizadas a cada paciente por la médica de base de oncología dio una kappa ponderada de 0.73, para la residente de medicina transfusional la kappa ponderada fue de 0.73, para los residentes de medicina interna (observadores 5 y 6) fue de 0.49 y 0.55 para cada uno. La kappa ponderada de las dos evaluaciones realizadas a cada paciente por las residentes de geriatría fue de 0.78 y 0.69 para cada una. Figura 13.

Si consideramos a los observadores por grupos, el acuerdo intraobservador de los dos residentes de medicina interna dio un valor de kappa ponderada de 0.62, el acuerdo entre las 2 residentes de geriatría dio un valor de kappa ponderada de 0.73 y el acuerdo entre la hemato-oncóloga y la residente de medicina transfusional dio un valor de kappa ponderada de 0.80. Figura 14.

3 - REPRESENTACIÓN GRÁFICA DE LA REPRODUCIBILIDAD CON BASE EN LA FORMA PROPUESTA POR BLAND Y ALTMAN

Se realizaron gráficas en la forma propuesta por Bland y Altman de los observadores que consiguieron el peor y el mejor acuerdo intraobservador (la kappa ponderada mas baja y la mas alta, observadores 5 y 8 respectivamente) En estas gráficas (figuras 15 y 16) se representa sobre el eje X el promedio de las dos mediciones realizadas por un mismo observador marcado contra la diferencia de dichas mediciones (representada sobre el eje Y) También se muestra en las figuras el promedio de las diferencias (línea punteada) y los límites de acuerdo (líneas continuas). En estas figuras se observa que solo uno de 24 pares de mediciones del observador 5 y uno de 21 pares de mediciones del observador 8 queda fuera de los límites de acuerdo (promedio de las diferencias ± 2 SD), pero los límites de acuerdo para el observador 5 (+20.42 y -31.24) fueron mayores comparados con aquellos del observador 8 (+14.03 y -14.97), lo que muestra una mayor variabilidad intraobservador del observador 5. Lo anterior a pesar de que el número de casos en los que no hubo diferencia (diferencia= cero, diferencia ideal) fue similar para ambos observadores (el observador 5 no mostró diferencia en 12 pares de mediciones y el observador 8 en 13)

En el caso de la reproducibilidad interobservador se escogió como estándar a la hemato-oncóloga, observador que contaba con mas tiempo de entrenamiento clínico, y se comparó su primera medición contra la de uno de las geriatras y uno de los internistas escogidos al azar. Se construyeron gráficas en la forma descrita anteriormente que se muestran en las figuras 17 y 18. En la construcción de la figura 17 se incluye la primera medición de los observadores 1 y 9 (la hemato-oncóloga y una de las geriatras) y se puede observar que

solamente uno de 25 pares de mediciones queda fuera de los límites de acuerdo, que van de +21 a -14.2 y no se encontró diferencia en siete de 25 pares de mediciones (una medición de cada observador). En la elaboración de la figura 18 se utiliza la primera medición de los observadores 1 y 6, y se observa también solamente uno de 25 pares de mediciones fuera de los límites de acuerdo que van de +29.33 a -25.57 y no se encontró diferencia en 6 pares de mediciones

DISCUSIÓN

La reproducibilidad intraobservador de la hemato-oncóloga, la residente de medicina transfusional y las geriatras consideradas en forma individual fue substancial en este estudio, de acuerdo a lo estipulado por Landis y Koch, mientras que la reproducibilidad intraobservador de los médicos internistas considerados en forma individual fue moderada

Si consideramos a los observadores por grupos, la reproducibilidad intraobservador de las dos hematólogas (una de ellas hemato-oncóloga), las dos geriatras y los dos internistas fue substancial, pero el acuerdo intraobservador de las hematólogas fue mayor (κ ponderada = 0.80) que el de los internistas (κ ponderada = 0.62) y el de las geriatras (κ ponderada = 0.73)

No contamos con informes de la literatura donde se investigue la reproducibilidad intraobservador de la escala de Karnofsky para poder comparar nuestros resultados, sin embargo con base en lo estipulado por Landis y Koch, podemos considerar el acuerdo intraobservador en nuestro medio como favorable.

Sin embargo es importante mencionar, que no se puede descartar que la reproducibilidad intraobservador haya sido afectada en nuestro estudio, por la influencia que la memoria del observador puede haber tenido sobre la segunda medición. Un tiempo máximo de diferencia de 7 horas entre las dos valoraciones realizadas a cada paciente por un mismo observador disminuye por un lado la posibilidad de variación en las condiciones del paciente, pero por otro lado, una diferencia mínima de dos horas entre la primera y la segunda evaluación aumenta la posibilidad de que exista sesgo de memoria

Hubiera sido conveniente que transcurriera mas tiempo entre ambas valoraciones, pero como ya se mencionó, esto no fue posible por los horarios de las diferentes actividades de los observadores. Otra posible solución a este problema sería la valoración de pacientes externos, que puede suponerse tienen una mayor estabilidad en su condición clínica que los pacientes hospitalizados y por lo tanto, pudieran ser valorados en días diferentes. Se pueden encontrar algunos inconvenientes para llevar a cabo lo anterior, por ejemplo, es más difícil cubrir todo el espectro de la escala de Karnofsky con pacientes externos que con pacientes hospitalizados, además, el paciente tiene que acudir al hospital a una segunda valoración o bien el observador tiene que ir al domicilio del paciente en dos días diferentes para poder realizar las dos evaluaciones

La reproducibilidad interobservador en este estudio fue moderada. Los hemato-oncólogos considerados como grupo tuvieron una reproducibilidad interobservador mayor (κ ponderada = 0.59) que la de los geriatras (κ ponderada = 0.41) y la de los internistas (κ ponderada = 0.42)

Lo anterior puede estar relacionado al mayor entrenamiento clínico en la valoración de pacientes con cáncer de los especialistas en hematología y en oncología, además de que ellos cuentan con una mayor experiencia en el uso de la escala de Karnofsky. Los especialistas en geriatría y medicina interna están menos familiarizados con el uso de la escala y, aunque todos los observadores contaban con una copia del instrumento al momento de realizar las valoraciones, posiblemente la dificultad estriba en la clasificación de los pacientes en cada categoría; esto último podría ser no solamente consecuencia de las diferencias en el entrenamiento clínico o en el uso de la escala de Karnofsky, sino que también puede deberse a un defecto del instrumento para definir claramente las diferentes categorías dentro de la escala.

No podemos descartar en nuestro estudio que el cambio o fluctuación en las condiciones de los pacientes evaluados pudiera influir en el resultado de la reproducibilidad interobservador, sin embargo consideramos que la contribución de este aspecto en las diferencias encontradas entre los observadores es mínima, porque el tiempo máximo transcurrido entre la valoración de cada paciente por los diferentes observadores fue de siete horas con treinta minutos, un tiempo en el que es poco probable que ocurra una variación importante en el estado clínico de un enfermo, además, se excluyeron los pacientes que podían presentar cambios marcados en sus condiciones en un periodo breve de tiempo, en relación por ejemplo a estudios invasores, cirugía, inicio de administración de quimioterapia, etc.

En cuanto a otros factores que pueden afectar el resultado de la reproducibilidad interobservador, podemos mencionar que en este caso, las valoraciones se realizaron

durante el día, siempre en el mismo sitio (la cama del paciente), por lo que las condiciones de luz, ruido y ventilación fueron similares para todos los observadores. Los efectos de algunos otros factores son más difíciles de valorar, por ejemplo la subjetividad del observador y la cooperación del paciente con cada observador en particular

Existen algunos informes de estudios en los que se ha investigado la reproducibilidad interobservador de la escala de Karnofsky como ya se comentó en los antecedentes. Ninguno de ellos es comparable con este trabajo, ya que en los demás estudios no se analizaron los resultados con el método estadístico de kappa ponderada (como ya se mencionó, la kappa ponderada es el instrumento de elección en el análisis de la reproducibilidad interobservador de las escalas ordinales). Sin embargo, en uno de los estudios mencionados (2), se informó una reproducibilidad interobservador moderada, en ese caso se utilizó como método estadístico el coeficiente de correlación de Pearson.

La reproducibilidad interobservador entre los hemato-oncólogos en este estudio puede considerarse favorable, no así la reproducibilidad interobservador entre los internistas y los geriatras, que se encuentra muy cerca del límite del valor de kappa ponderada para ser considerado como un grado de acuerdo ligero ($kappa$ ponderada = 0.40), con base en lo estipulado por Landis y Koch.

Si analizamos las gráficas que se elaboraron con base en lo propuesto por Bland y Altman, podemos observar que en todas las gráficas que presentamos, el 95% de pares de mediciones queda incluido dentro de los límites de acuerdo, un dato a favor de una buena reproducibilidad según lo que mencionan estos autores.

En las gráficas en las que se muestran los observadores que resultaron con la kappa ponderada más alta y la más baja cuando se analizó la reproducibilidad intraobservador, se aprecia una mayor variabilidad en las mediciones del observador con la kappa ponderada más baja, resultado que era esperado

Este tipo de gráficas le permiten al clínico una estimación visual de la variabilidad en los pares de mediciones, lo que por una parte nos da una buena idea de la reproducibilidad del instrumento y por otra parte, nos permite conocer si la diferencia en las mediciones puede tener un significado clínico o no (ya que por ejemplo, una diferencia de 10 puntos en dos mediciones basadas en la escala de Karnofsky puede no tener ninguna importancia clínica y una diferencia de 30 ó 40 puntos sí)

Es necesario establecer medidas pertinentes para incrementar la reproducibilidad de la escala de Karnofsky, considerando que es un instrumento que se utiliza con frecuencia. Proponemos las siguientes medidas para tratar de incrementar la reproducibilidad de esta escala

- 1) Que los observadores tengan una experiencia clínica similar en la valoración de pacientes con cáncer
- 2) Que antes de incluir la escala de Karnofsky como un método para valorar estado funcional de los pacientes incluidos en un estudio, se busque que los evaluadores tengan un grado de entrenamiento similar en el uso del instrumento. Es necesario llevar a cabo un período de entrenamiento adecuado en el uso de la escala, en el que se incluyan

pacientes reales y en el que se realice una evaluación al terminar el entrenamiento para valorar la uniformidad del grupo de observadores

- 3) Que el tiempo permitido para la valoración de cada uno de los pacientes por todos los observadores (cuando se investiga la reproducibilidad interobservador) sea siempre el mínimo posible, para disminuir al máximo el efecto que puede tener la fluctuación en las condiciones del paciente sobre la valoración por los diferentes observadores.
- 4) Que el tiempo entre la primera y la segunda valoración de cada paciente por cada uno de los observadores (cuando se investiga la reproducibilidad intraobservador) sea suficiente para disminuir al máximo posible el sesgo de memoria.

9 - CONCLUSIONES

- a) La reproducibilidad intraobservador de la escala de Karnofsky en este estudio fue substancial con base en lo estipulado por Landis y Koch.
- b) La reproducibilidad interobservador fue moderada.
- c) En ambos casos, la reproducibilidad fue mayor entre los hemato-oncólogos que entre los internistas y los geriatras, lo que probablemente refleja un mayor entrenamiento clínico en la valoración de pacientes con cáncer y en el uso de la escala de Karnofsky.
- d) Es necesario establecer medidas pertinentes para incrementar la reproducibilidad de la escala de Karnofsky, ya que se trata de un instrumento que se utiliza con frecuencia

Tabla 1 La escala de Karnofsky

CONDICIÓN	CALIFICACIÓN	COMENTARIOS
A) Capaz de realizar actividad normal y trabajar. No necesita cuidado especial	100	Actividad normal, no hay evidencia de enfermedad.
	90	Capaz de realizar actividad normal, hay leves signos o síntomas de enfermedad.
	80	Realiza actividad normal con esfuerzo, hay signos o síntomas de enfermedad
B) No es capaz de trabajar, atiende el cuidado personal con un grado variable de asistencia	70	Atiende su cuidado personal sin ayuda
	60	Requiere asistencia en forma ocasional, pero es capaz de atender la mayoría de las actividades del cuidado personal
	50	Requiere asistencia considerable en su cuidado personal, requiere asistencia médica frecuente.
C) Incapaz de cuidar de sí mismo Requiere cuidado especializado, la enfermedad puede ser rápidamente progresiva.	40	Incapacitado, requiere cuidado especial.
	30	Gravemente incapacitado La muerte no es inminente.
	20	Muy enfermo. Necesita hospitalización y tratamiento con soporte especializado.
	10	Moribundo.
	0	Muerte.

Tabla 2.- Características de los pacientes.

n= 25

Edad Promedio en años	Desviación estándar	Límites
52.4	16.98	18-73
Tiempo promedio de evolución (en meses).	Desviación estándar	Límites
16.1	23	1-84

Tabla 3 Medianas de las calificaciones asignadas por cada uno de los observadores a los 25 pacientes (primera y segunda valoración)

Observador	1ª valoración		2ª valoración	
	Mediana-Límites		Mediana-Límites	
1- MB Oncología	80	40-100	80	40-100
2- MR Med Transfusional	80	40-100	80	40-100
3- MB Hematología	70	30-100	--	
4- MR1 Hematología	75	40-100	--	
5- MR4 Med Interna	80	50-100	90	50-100
6- MR4 Med Interna	80	50-100	80	50-100
7- MR4 Med Interna	80	50-100	--	
8- MR2 Geriatria	75	40-100	80	40-100
9- MR1 Geriatria	80	45-100	80	50-100

MB Médico de base MR. Médico residente R1.R2.R4. Residente de primero, segundo, cuarto año

Tabla 4 Medianas y límites de las calificaciones por especialidad

Especialidad	Primera evaluación		Segunda evaluación	
	Mediana - Límites		Mediana - Límites	
Hemato-oncólogos	80	40-100	80	40-100
Internistas	80	50-100	85	50-100
Geriatras	80	40-100	80	40-100

Figura 1.-Tipos de neoplasias en los pacientes estudiados
n = 25

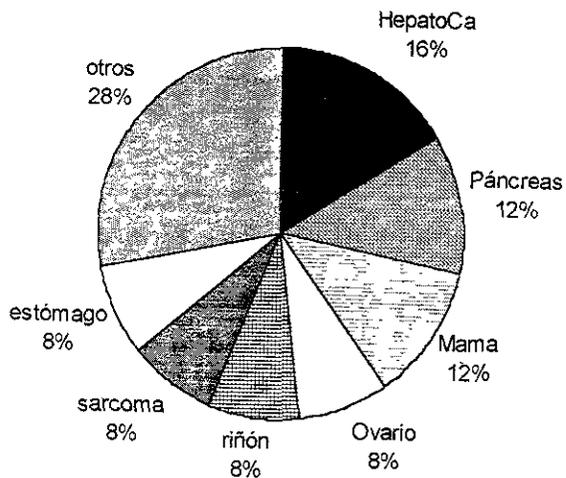


Figura 2.- Sexo de los pacientes estudiados

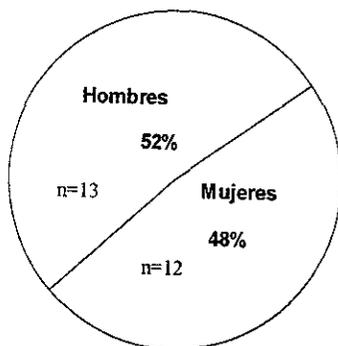


Figura 3.- Puntajes asignados a los pacientes por los hemato-oncólogos. Primera valoración.

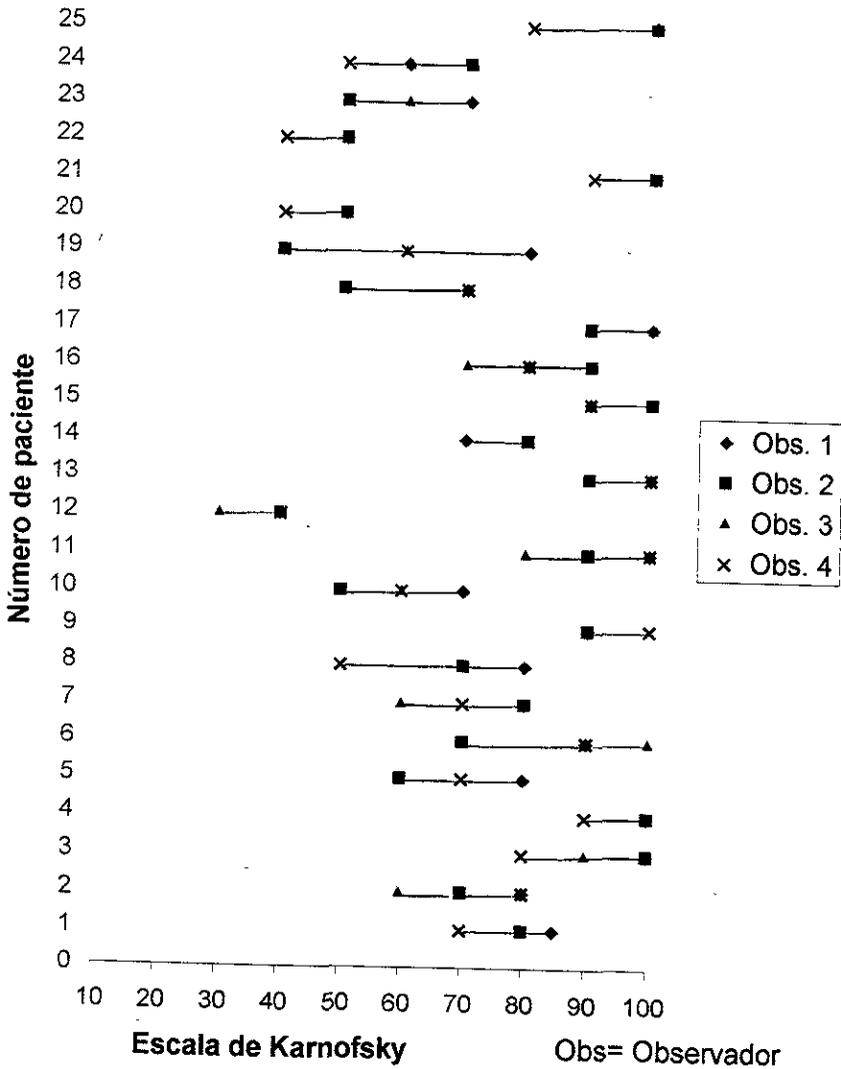


Figura 4.- Puntajes asignados a los pacientes por los internistas. Primera valoración.

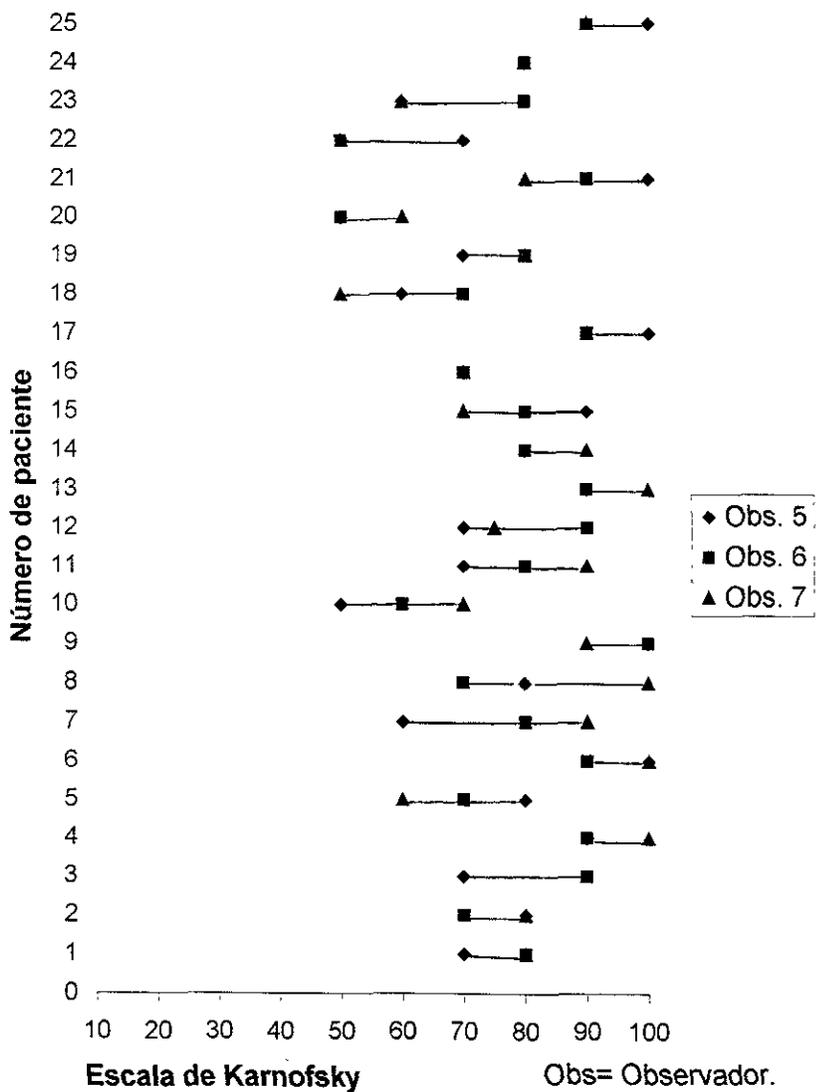


Figura 5.- Puntajes asignados a los pacientes por los geriatras. Primera valoración.

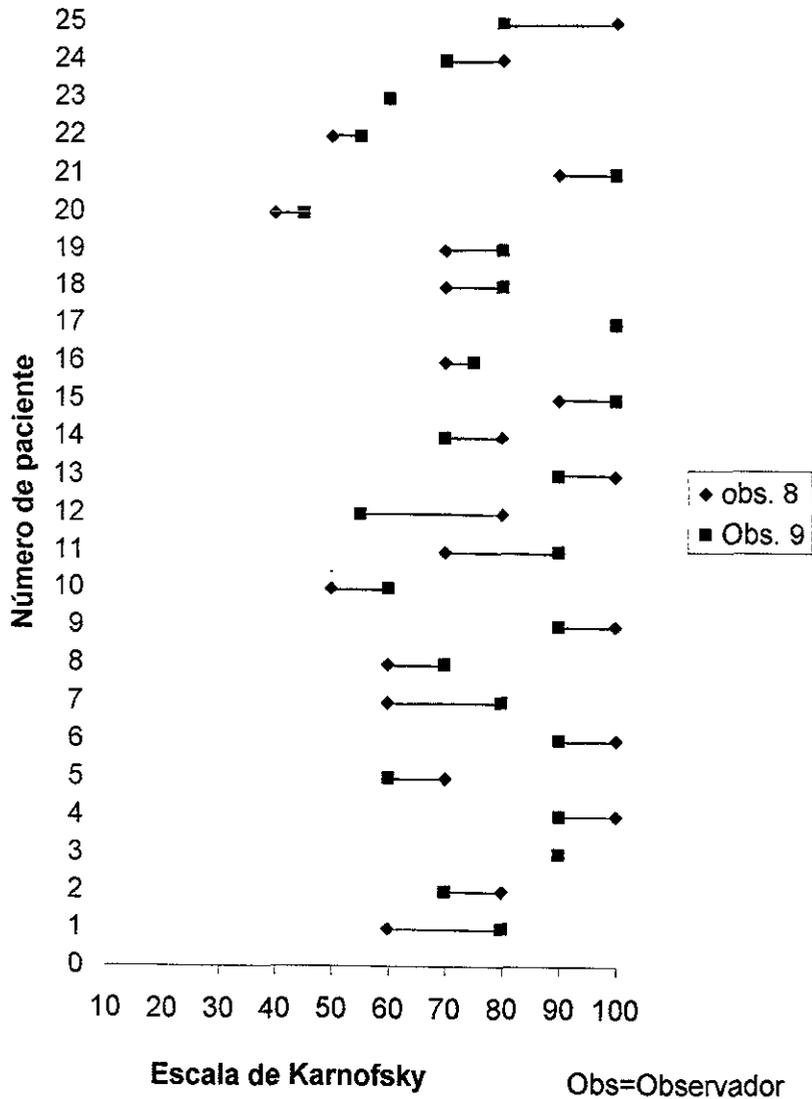


Figura 6.- Puntajes asignados a los pacientes en la primera y segunda valoración por el observador 1.

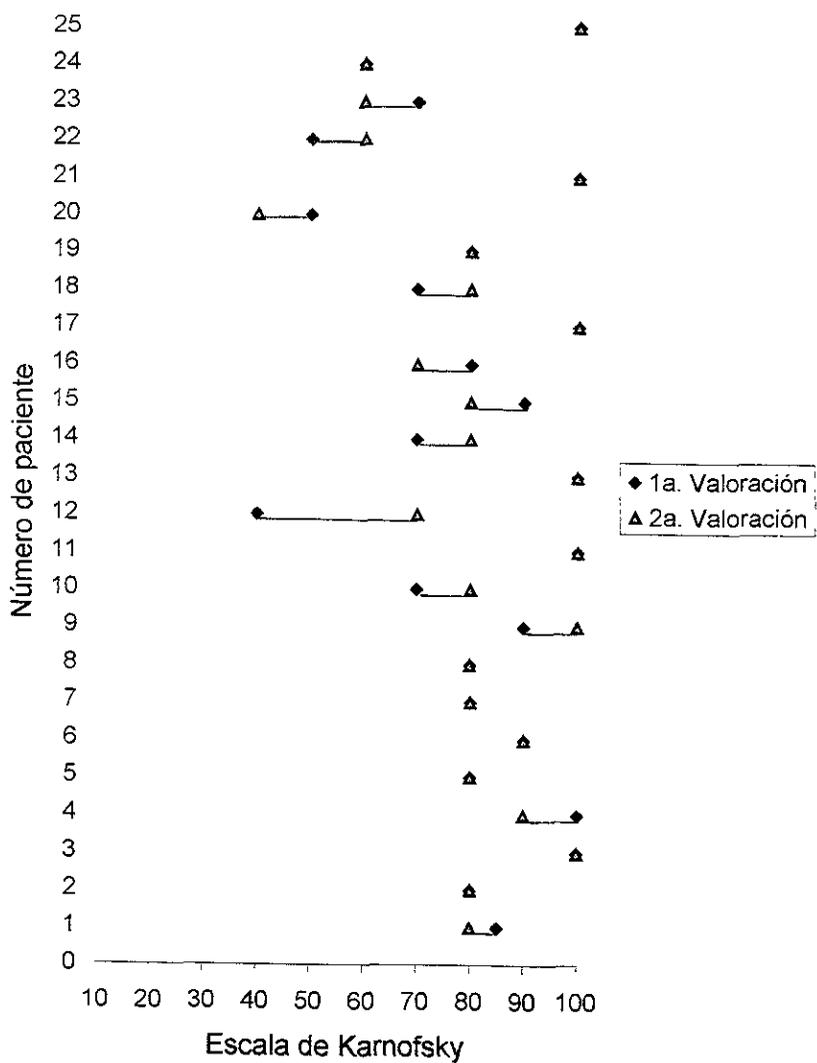


Figura 7.- Puntajes asignados a los pacientes en la primera y segunda valoración por el observador 2.

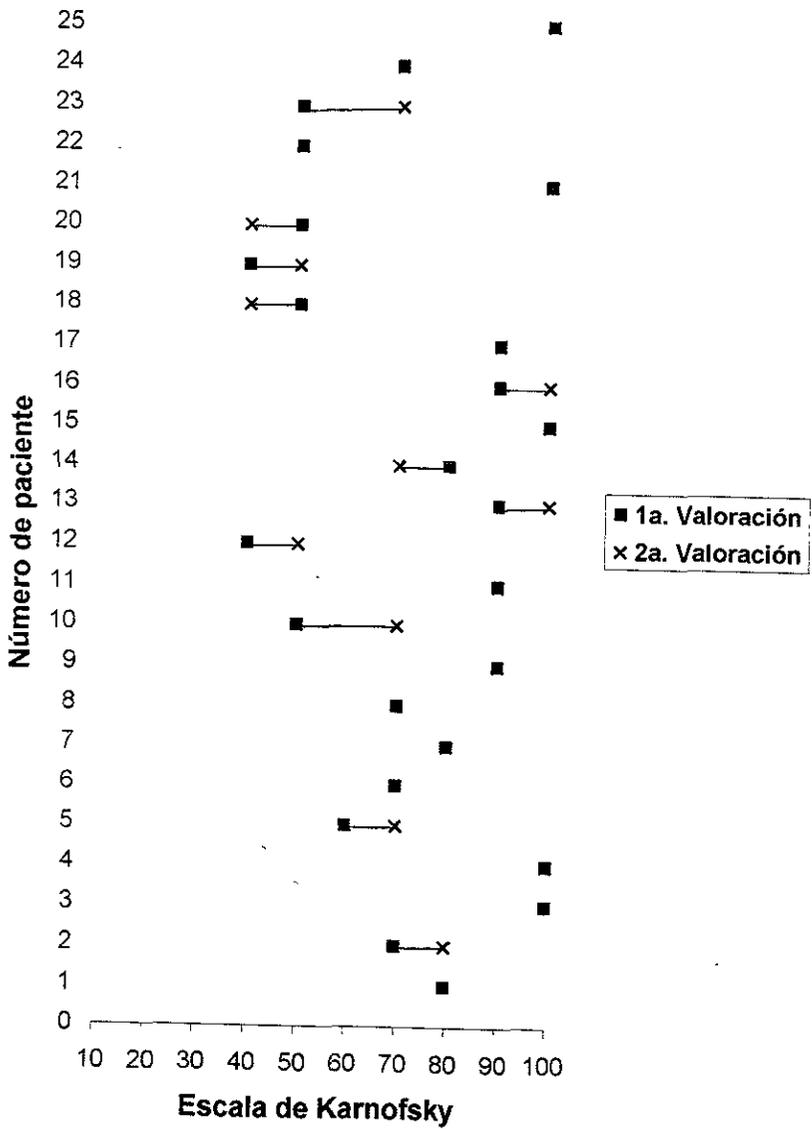


Figura 8.- Puntajes asignados a los pacientes en la primera y segunda valoración por el observador 5.

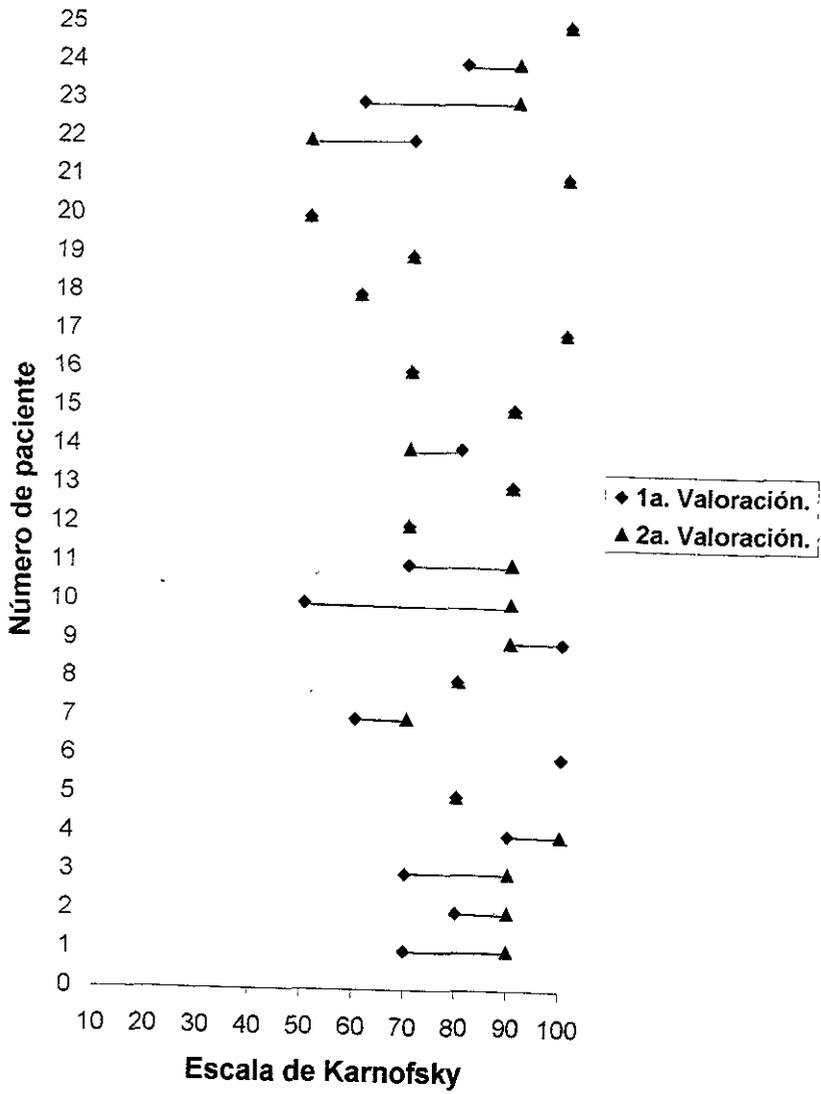


Figura 9.- Puntajes asignados a los pacientes en la primera y segunda valoración por el observador 6.

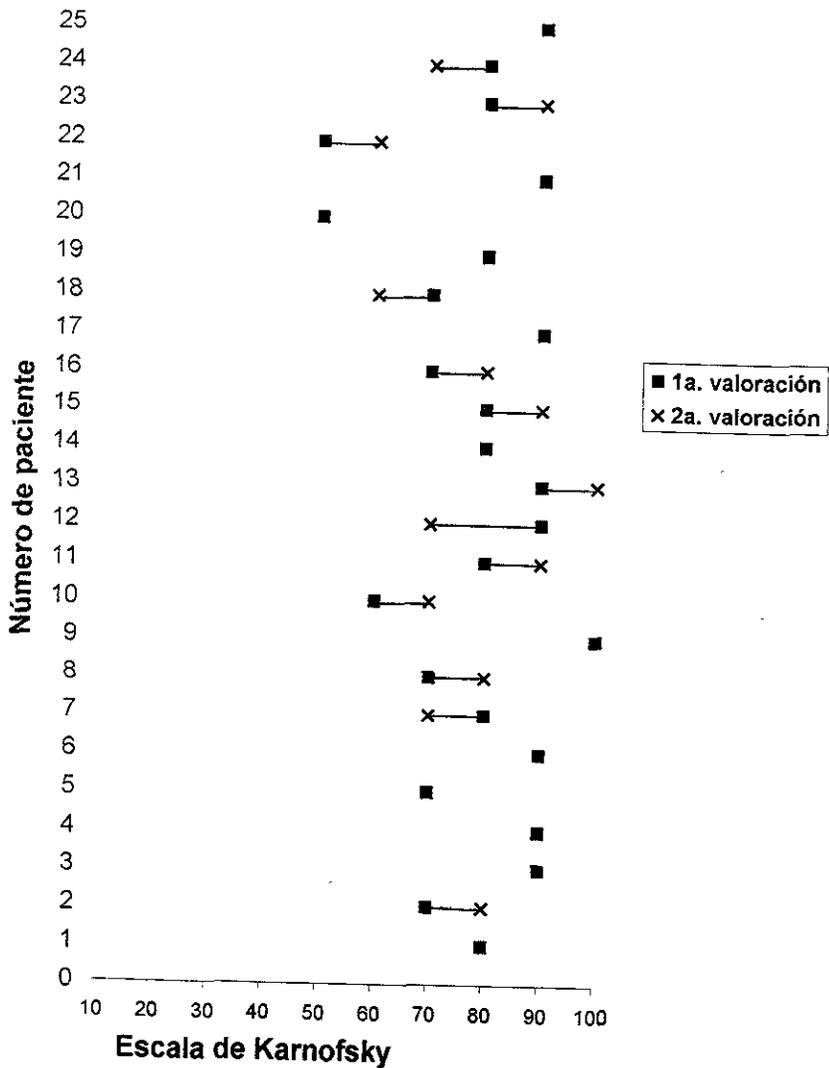


Figura 10.- Puntajes asignados a los pacientes en la primera y segunda valoración por el observador 8.

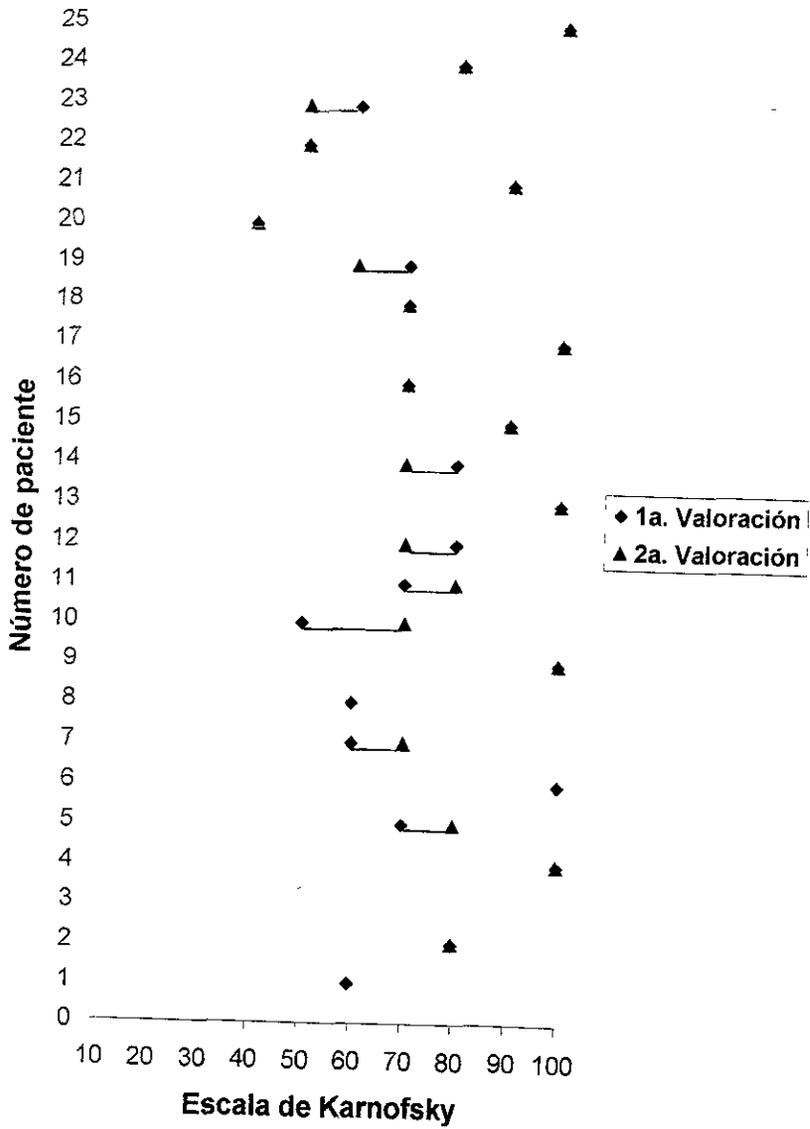


Figura 11.- Puntajes asignados a los pacientes en la primera y segunda valoración por el observador 9.

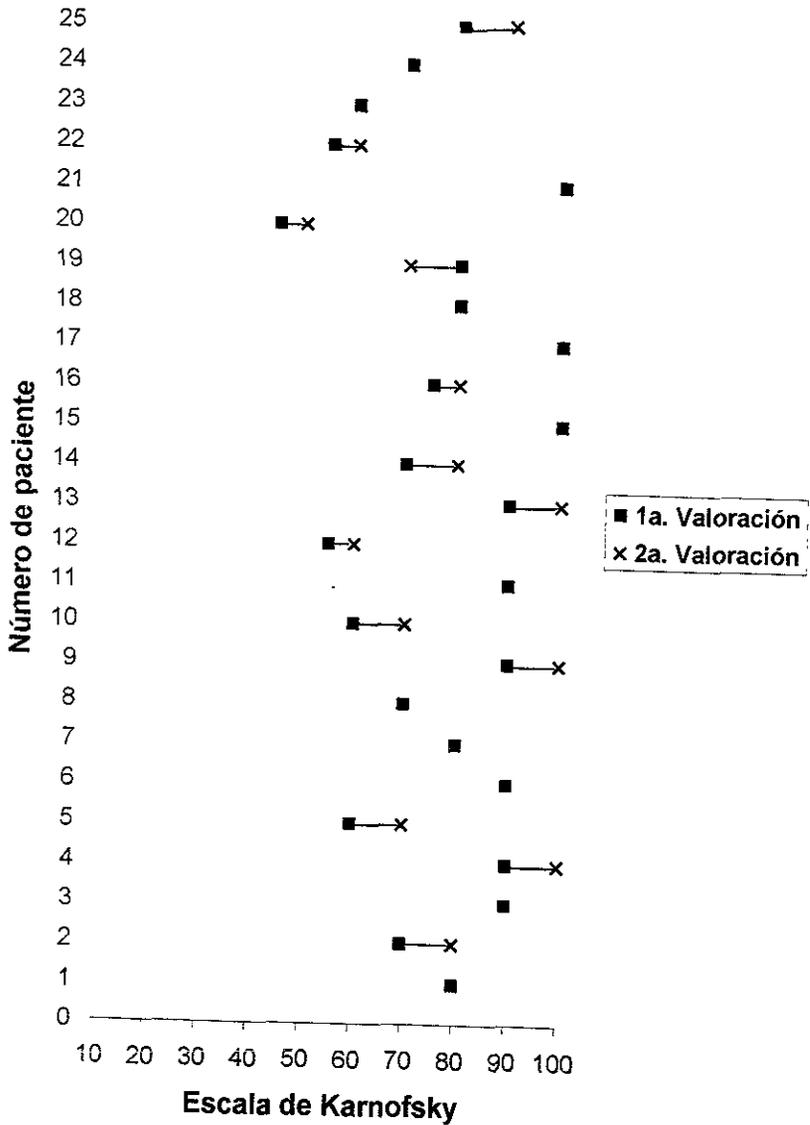


Figura 12.- Reproducibilidad interobservador por grupos.

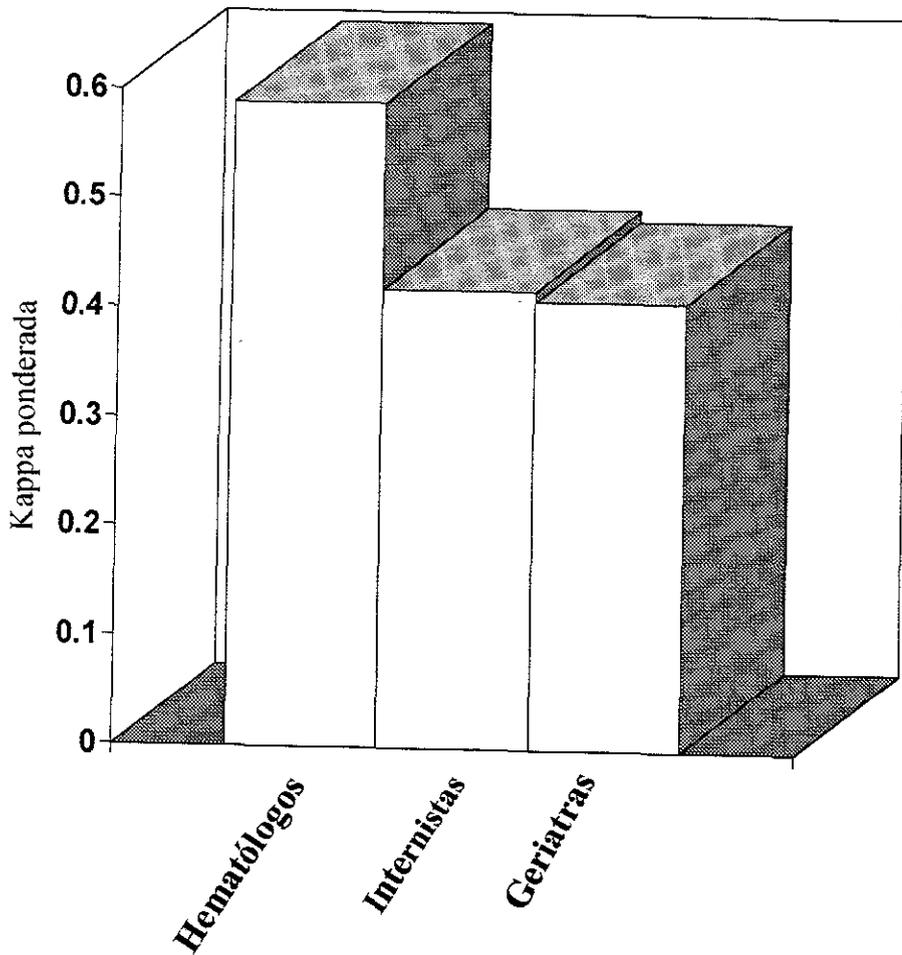


Figura 13.- Reproducibilidad intraobservador.

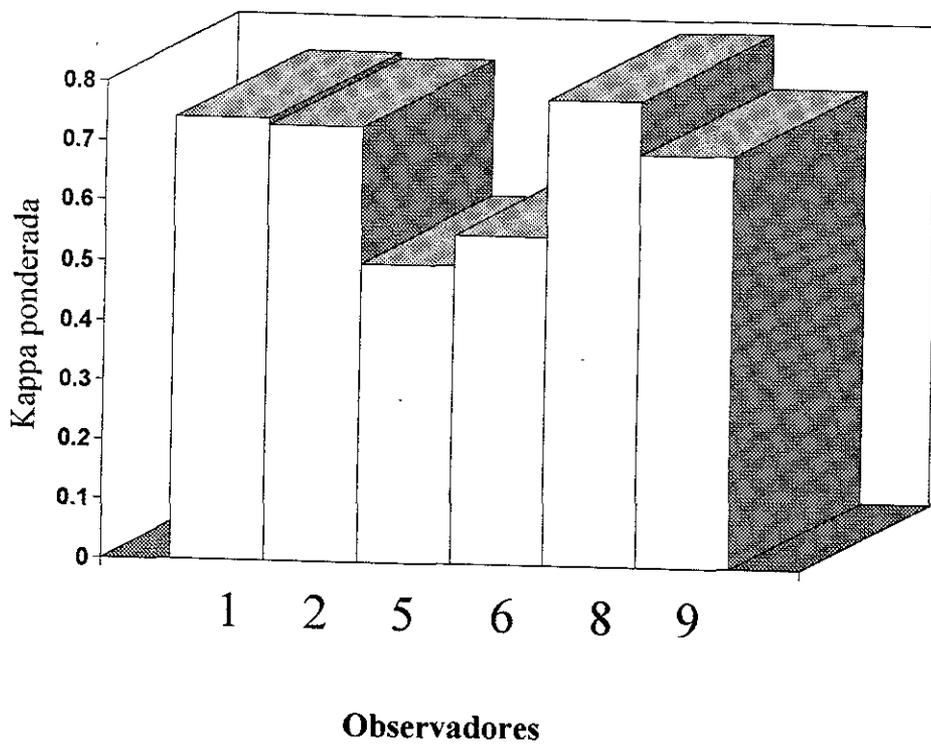


Figura 14.- Reproducibilidad intraobservador por grupos

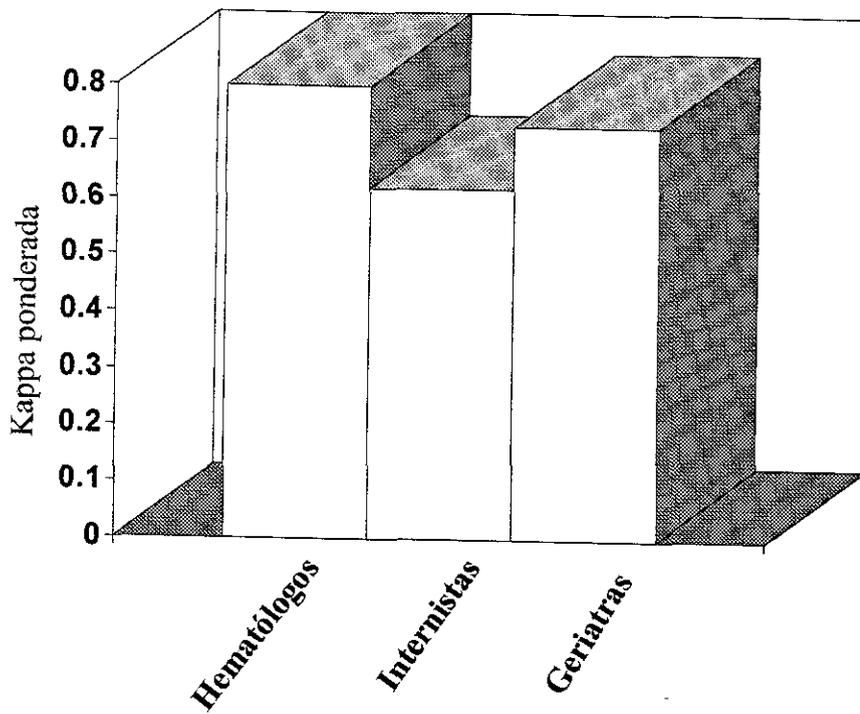
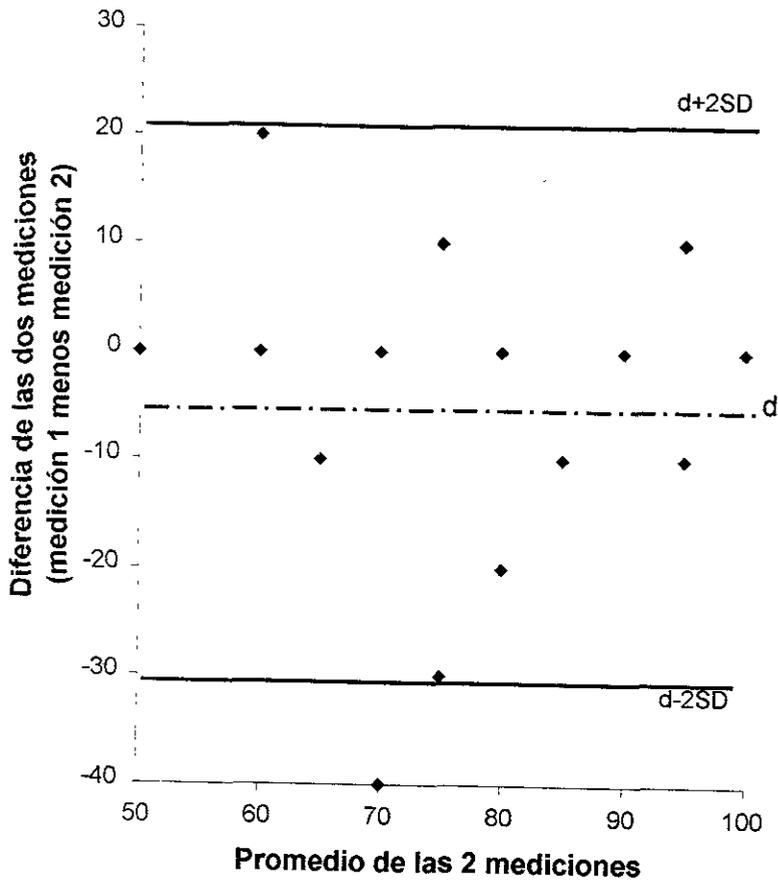
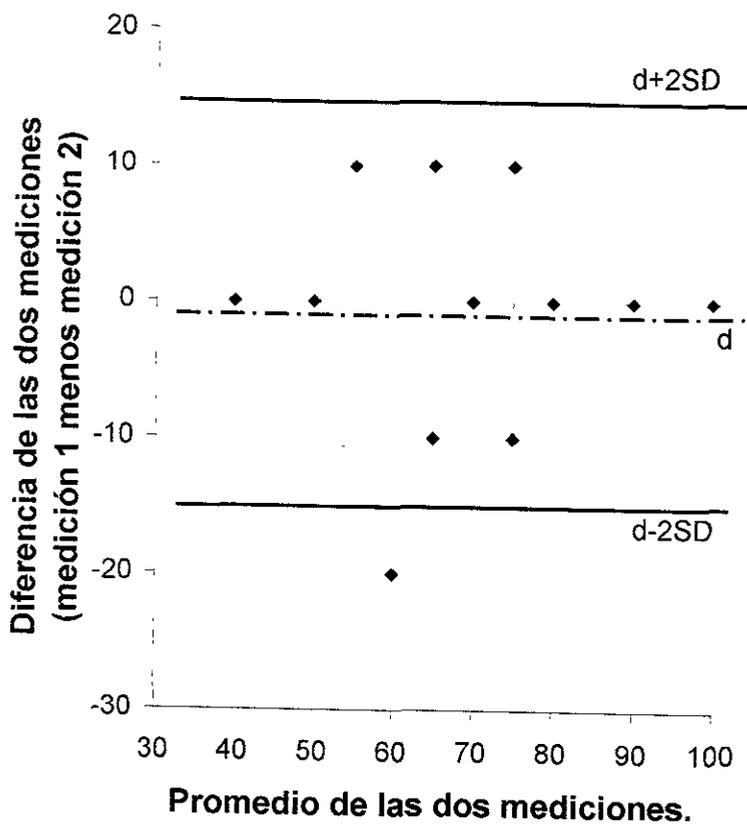


Figura 15.- Diferencia contra el promedio de las 2 mediciones del observador 5.



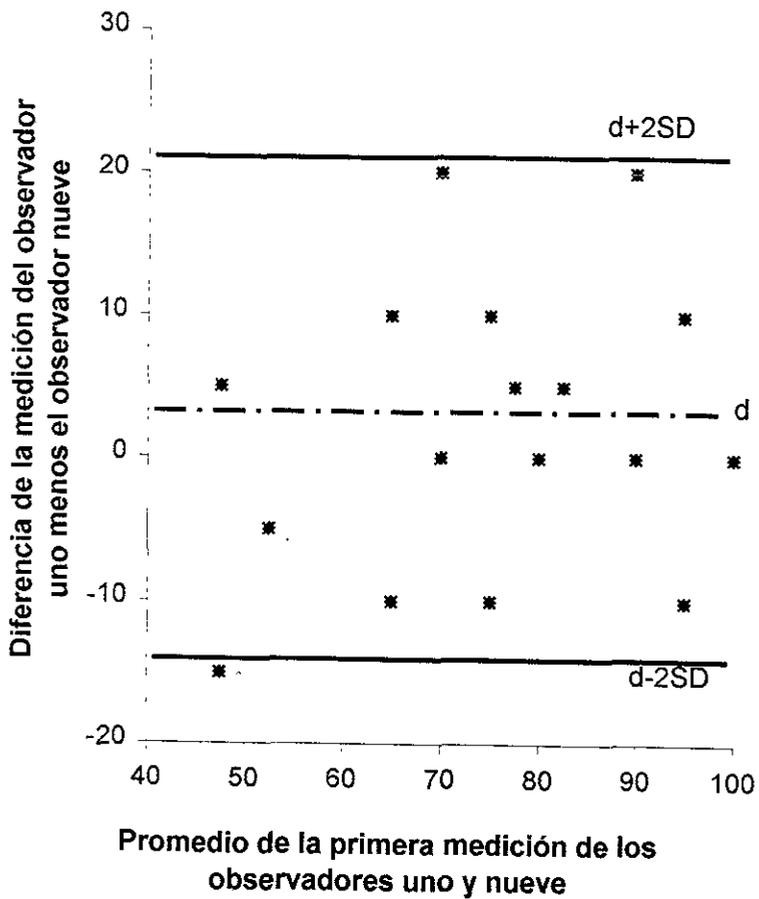
d= Diferencia promedio d+2SD= d mas 2 desviaciones estándar
 estándar d- 2SD= d menos 2 desviaciones estándar.

Figura 16.- Diferencia contra el promedio de las dos mediciones del observador 8.



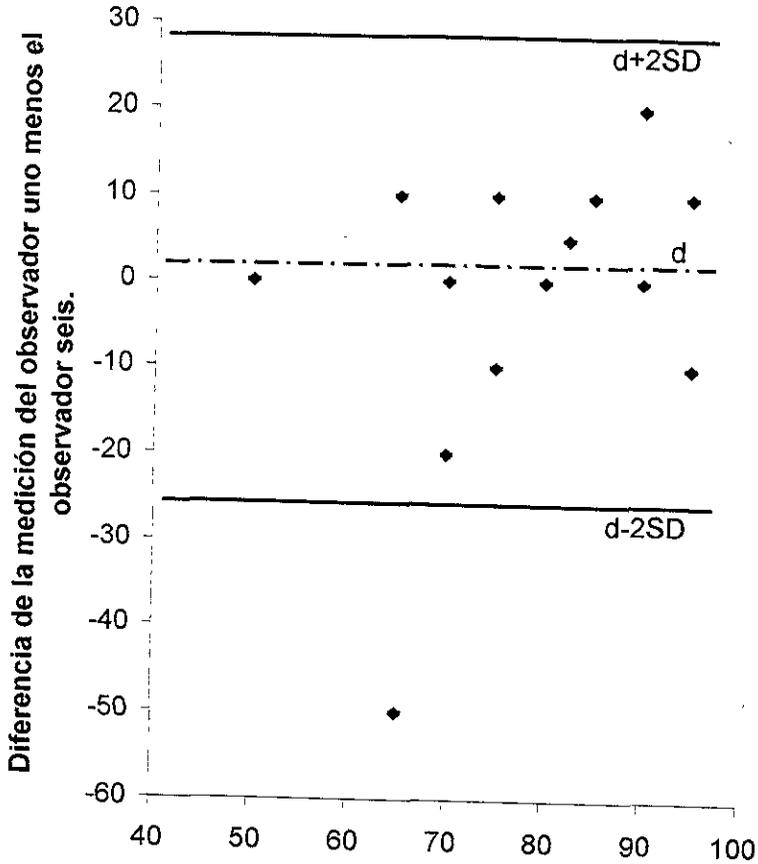
d= Diferencia promedio d+2SD= d mas 2 desviaciones estándar
 d-2SD= d menos 2 desviaciones estándar

Figura 17.- Diferencia contra el promedio de la primera medición de los observadores uno y nueve.



d= Diferencia promedio d+2SD= d mas 2 desviaciones estándar d-2SD= d menos 2 desviaciones estándar.

Figura 18.- Diferencia contra el promedio de la primera medición de los observadores uno y seis.



Promedio de la primera medición de los dos observadores.

d= Diferencia promedio d+2SD= d mas 2 desviaciones estándar d-2SD= d menos 2 desviaciones estándar.

APÉNDICE 1

REPRODUCIBILIDAD DE LA ESCALA DE KARNOFSKY

EVALUACIÓN: 1ª () 2ª ()

Nombre del paciente:

Registro: _____

Cama: _____

Sexo: _____

Edad: _____

Karnofsky: _____

Evaluó

Grado y especialidad: _____

Fecha y hora: _____

APÉNDICE 2

México, D F 09 de Septiembre de 1997

Estimado paciente:

Estamos evaluando la utilidad de un método que sirve para analizar el desempeño físico de los pacientes y solicitamos su valiosa colaboración para llevar a cabo este proyecto.

De permitirlo usted, su colaboración consistirá en ser entrevistado por 10 personas diferentes en un lapso de 12 horas. Los diferentes entrevistadores son médicos de este instituto y le harán preguntas a cerca de su capacidad para realizar diferentes actividades, tomando en cuenta las limitaciones que pueda usted presentar como consecuencia de su enfermedad actual

Si usted está de acuerdo en participar con nosotros, le rogamos nos conceda una firma en este documento

Agradecemos de antemano su colaboración

ATTE: Dra Cristina Martínez Sibaja.

Estoy de acuerdo en participar:

Nombrey firma: _____

Testigo

Nombre y firma: _____

Apéndice 3

REPRODUCIBILIDAD DE LA ESCALA DE KARNOFSKY

(Hoja de colección de datos de los pacientes evaluados)

Fecha: _____

Nombre. _____ Edad _____

Registro _____ Cama. _____ Sexo: _____

Diagnóstico oncológico o hematológico:

Estadio _____ Fecha del diagnóstico: _____

Tipo de tratamiento.

1.- Quimioterapia Sí () No ()

a) Esquema. _____

b) Ciclos recibidos. _____ c) Fecha del último: _____

2 - Cirugía Sí () No ()

a) Tipo de cirugía. _____

b) Fecha de cirugía: _____ c) ¿Fue curativa? Sí () No ()

Otros diagnósticos:

a) _____

b) _____

c) _____

BIBLIOGRAFÍA.

- 1 - Krakoff H Irwin. Progress and Prospects in Cancer Treatment: The Karnofsky Legacy. J Clin Oncology 1994;12 (2).432-438
- 2.-Yates J W, Chalmer B, McKegney P: Evaluation of Patientes with Advanced Cancer Using the Karnofsky Performance Status. Cancer 1980; 45, 2220-2224.
- 3 -Karnofsky D A , Abelmann W.H , Craver L.F, Burchenal J.H. The Use of the Nitrogen Mustards in the Palliative Treatment of Carcinoma. Cancer 1948, 1(4) 634-656.
- 4 -Hutchinson T A , Boyd N.F., Feinstein R.A: Scientific Problems in Clinical Scales, as Demonstrated in the Karnofsky Index of Performance Status. J Chron Dis 1979; 32:661-6
- 5 -Mor V, Laliberte L, Morris J N, Wiemann M The Karnofsky Performance Satatus Scale, An examination of its Reliability and Validity in a Research Setting Cancer 1984; 53: 2002-2007.
- 6 -Kirhsner B, Guyatt G A Methodological Framework For Assessing Health Indices. J Chron Dis 1985, 38 (1) 27-36.
- 7 -O'toole D M, Golden M A. Evaluating Cancer Patientes for Rehabilitation Potencial West J Med 1991, 155 384-387
- 8 -Houwelingen J C, Bokkel W W, Van der Burg M.E, Oosterom A T, Neijt J P: Predictability of the Survival of Patients With Advanced Ovarian Cancer. J Clin Oncology 1989; 7 (6) 769-773
- 9.-Pollack Ian, Lunsford D, Flickinger J.C, Dameshek L: Prognosis Factors in The Diagnosis and Treatment of Primary Central Nervous System Lymphoma Cancer 1989, 63 939-947
- 10 -Buccheri G, Ferrigno D, Tamburini M. Karnofsky and ECOG Performance Status Scoring in Lung Cancer Eur J Cancer 1996; 32A (7) 1135-41
- 11 -Vecht C.J , Hovestadt A, Verbiest H B, Van-Vliet J.J, Van Putten W.L Dose-effect relationship of dexamethasone on Karnofsky performance in metastatic brain tumors: a randomized study of doses of 4, 8, and 16 mg per day. Neurology 1994; 44 (4) 675-80.
- 12.-Quoix E , Hedelin G, Popin E, Charloux A. Can we predict very short term survival in small cell lung cancer? Lung Cancer 1993; 10 (3-4). 229-38.
- 13 -Yaneva M, Arolsky I, Iliev E. Factors influencing prognosis of patients with CNS supratentorial astroglial tumors who underwent posoperative radiotherapy. Folia-Med-Plovdiv 1993, 35 (3-4) 49-54

- 14 -Chandler K.L, Prados M, Malec M, Wilson C B. Long-term survival in patients with glioblastoma multiforme *Neurosurgery*. 1993; 32 (5): 716-20.
- 15 -Devaux B C, O'Fallon J.R, Kelly P J. Resection, biopsy, and survival in malignant glial neoplasms A retrospective study of clinical parameters, therapy and outcome. *J Neurosurg* 1993; 78 (5). 767-75.
- 16 -Payne J.L. and cols. Outcomes analysis for 50 liver transplant recipients: the Vanderbilt experience *American Surgeon*. 1996, 62 (4) 320-325.
- 17 -Gutman R.A, Stead W.W, Robinson R. Physical activity and employment status of patients on maintenance dialysis *New Eng J Med*. 1981, 304 (6). 309-313
- 18 -Moreno F, López-Gómez J M, Sanz-Guajardo D. Quality of life in dialysis patientes. A Spanish cooperative renal patientes quality of life study group. *Nephrol-Dial-Transplant* 1996, 11 suppl 2: 125-29
- 19 -Lazard T, Retel O, Guidet B, Maury E. AIDS in a medical intensive care unit immediate prognosis and long-term survival. *JAMA* 1996; 276 (15): 1240-45
- 20 -Ifudu O, Brezsnayak W.F, Reydel C. Pathobiology and functional status of long-term hemodialysis patients *Am J Nephrol* 1995; 15 (5). 379-85.
- 21 -Ganz P.A, Coscarelli-Schag C, Kahn B. Describing the health-related quality of life impact of HIV infection: findings from a study using the HIV overview of problems evaluation system (HOPES) *Qual-Life-Res* 1993, 2 (2): 109-19.
- 22.-O'Brien B.J, Buxton M.J, Patterson D.L. Relationship between functional status and health-related quality of life after myocardial infarction *Med care*. 1993; 31 (10): 950-55.
- 23.-Gish R.G, Lee A.H, Keeffe E.B. Liver transplantation for patients with alcoholism and end-stage liver disease. *Am J Gastroenterol* 1993; 88 (9): 1337-42.
- 24.-Feinstein A R. An additional basic science for clinical medicine IV. The development of clinimetrics. *Ann Int Med* 1983; 99 (6): 843-48
- 25 -Feinstein A.R, Josephy B.R, Wells C. Scientific and clinical problems in indexes of functional disability. *Ann Int Med*. 1986; 105 (3): 413-20
- 26.-Guzmán J, Ponce de León S, Pita R.L. El cambio en la calidad de vida como indicador de curso clínico de la enfermedad. Comparación de dos índices *Rev Invest Clín*. 1993; 45 (5): 439-52
- 27.-Cardiel-Ríos M.H. Design of an index to measure health-related quality of life in mexican patients with systemic lupus erithematosus in a tertiary care center. Thesis.

Mc-Master University September 1989

- 28 -Kramer M S, feinstein A.R Clinical biostatistics. LIV The biostatistics of concordance Clin. Pharmacol. Ther. 1981, 29 (1): 111-123.
- 29 -Koran L.M The reliability of clinical methods, data and judgments. New Eng J Med. 1975, 293 (13): 642-646.
- 30 -Koran L.M The reliability of clinical methods, data and judgments (second of two parts). New Eng J Med 1975, 295 (14): 695-701.
- 31.-Bland J.M, Altman D.G. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 8:307-310.
- 32 -Bland J.M, Altman D G Comparing methods of measurement: why plotting difference standard method is misleading Lancet 1995, 346: 1085-1087
- 33.-Hernández-Cruz B, Cardiel M H. Intra-observer reliability of commonly used outcome measures in rheumatoid arthritis. Clinical and experimental Rheumatology 1998, 16: 320-23