

11227

36

24.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO.FACULTAD DE MEDICINA.  
DIVISION DE INVESTIGACION Y ESTUDIOS DE POSTGRADO.  
INSTITUTO NACIONAL DE LA NUTRICION SALVADOR ZUBIRAN.

LA VALORACION DE LA CALIDAD DE VIDA  
EN ESTUDIOS CLINICOS CONTROLADOS.


Tesis de post-grado que para obtener  
el Titulo de Especialista en Medicina  
Interna presenta..


Jaime Guzmán Ramirez.

Asesor de tesis:

TESIS CON  
FALLA DE ORIGEN

Dr. Sergio Ponce de León.

  
Dr. Rubén Lisker Y.  
Subdirector de Enseñanza.

  
Dr. Ezequiel López Amor  
Titular de la materia.

México D.F., Enero de 1991.



Universidad Nacional  
Autónoma de México



## **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 está 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.

INDICE.

Página.

4..	I.- Introducción.
5..	II.- Marco Teórico. A) De la Importancia de Valorar la Calidad de Vida. B) Diversas Maneras de Medir la Calidad de Vida. C) De cómo Juzgar la Calidad de los Instrumentos que Miden Calidad de Vida.
18..	III.- Objetivo del Estudio.
19..	IV.- Material y Métodos. A) El Diseño Básico del Estudio. B) Los Pacientes. C) Los Instrumentos y Parámetros de Comparación. E) El Análisis.
37..	V.- Resultados. A) Datos Generales. B) Instrumento No. 1. C) Instrumento No. 2. D) Comparación de Ambos Instrumentos.
75..	VI.- Discusión.
78..	VII.- Conclusiones.
80..	Referencias.
83..	Apéndices.

## I.- INTRODUCCION.

La valoración de la calidad de vida ha sido relegada en gran parte de los ensayos clínicos controlados de los últimos años, aduciendo su subjetividad. Sin embargo, es justamente dicha calidad de vida una de las dos metas fundamentales de la medicina --la otra es la de prolongar la vida.

En esta tesis pretendo discutir en primer lugar las bases teóricas que permiten integrar un marco de referencia para la confección y prueba de instrumentos para valorar la calidad de vida, o su cambio en el tiempo. Haré particular énfasis en la diferencia entre medir la calidad de vida y medir cambios en la calidad de vida; siendo este último rubro el de aplicación más directa en los ensayos clínicos de intervención terapéutica.

En segundo lugar expondré la investigación realizada; en la que se comparan dos instrumentos específicamente diseñados para medir cambio en la calidad de vida.

Para cerrar esta tesis integraré al marco teórico expuesto los resultados de la investigación, así como sus implicaciones en cuanto a la metodología para el diseño y experimentación de instrumentos que pretendan medir cambio en la calidad de vida.

## II.- MARCO TEORICO.

### A) De la Importancia de Valorar la Calidad de Vida.

Las metas ancestrales de la Medicina han sido alargar la vida y menguar el sufrimiento; parafraseado de otra manera: aumentar la CANTIDAD y la CALIDAD de vida. Con el desarrollo de la medicina moderna se han ampliado conceptos, así tenemos la muy conocida definición de salud de la OMS (39): "Salud es un completo estado de bienestar físico, psicológico y social"; es decir, se trata de un estado de inmejorable calidad de vida.

Existen algunas enfermedades e intervenciones terapéuticas donde la mortalidad es un índice adecuado de la eficacia del tratamiento. Sin embargo en enfermedades crónico-degenerativas, y/o cuando no se cuenta con maniobras capaces de alterar la mortalidad, se requieren otros índices de eficacia terapéutica. Es ahí donde cobra particular importancia considerar el alivio proporcionado, es decir la mejoría inducida en la calidad de vida.

Al Enfrentarse con la intangibilidad de la calidad de vida los investigadores han recurrido con mucha frecuencia a variables subrogadas que sean más tangibles y fáciles de medir al mismo tiempo que reflejen de alguna manera los cambios en la calidad de vida de los individuos. Así se recurre cada vez más a resultados de laboratorio y gabinete para juzgar la eficacia de diversos tratamientos, y se llega a olvidar que dichos datos paraclínicos sólo son útiles en el grado en que reflejan el alivio proporcionado a los enfermos. Esta tendencia ha llevado a absurdos tales como el "tratar" exámenes de laboratorio, o el proponer un determinado tratamiento por su capacidad de mejorar "x" estudio de gabinete.

Tal desplazamiento se intenta justificar muchas veces anteponiendo la "dureza" de los datos paraclínicos a la "blandura" de la medición de calidad de vida; pero ¿Cuál es la esencia detrás de estos términos?:

De acuerdo con Feinstein (15,16), la dureza de un dato hace referencia básicamente a su permanencia, y a la confianza con la que puede ser utilizado en manejos estadísticos; en último análisis un dato "duro" es aquel que es reproducible, es decir, que se repite consistentemente a lo largo de una serie de mediciones.

Por otro lado, un dato blando es aquel que se considera poco confiable y no factible de someterse a un análisis estadístico significativo. Es decir, que no siempre es el mismo a lo largo de una serie de mediciones repetidas

Ahora bien, la reproducibilidad de una medición con mucha frecuencia no depende de la naturaleza del fenómeno medido, sino del instrumento utilizado. Así, por ejemplo, existieron en la antigüedad muchas imprecisiones en la medición de longitud, hasta que por acuerdo entre los miembros de cierto grupo social se designó una unidad homogénea o patrón y la manera de utilizarlo (el metro, por ejemplo); dicho de otra manera: la longitud de los objetos siempre estuvo ahí, pero su medición fué poco reproducible durante muchos años, hasta que se convino en un patrón y un procedimiento de medición. Aplicado a nuestro caso se puede decir en términos generales que la calidad de vida de los sujetos siempre ha estado ahí, pero su medición será poco reproducible, mientras no se convenga en un patrón(es) y un procedimiento adecuados.

Sin embargo, es claro que el procedimiento a usar y las fuentes de error que se deben controlar son diferentes en uno y otro caso. Pongamos otro ejemplo: la definición de un kg, "es el peso de 1 decímetro cúbico de agua destilada a 4°C con una presión de una atmósfera" controla de manera muy estricta variables físicas tales como la temperatura, la presión, y la composición del líquido en cuestión. Análogamente la definición de una unidad patrón en calidad de vida debe controlar estrictamente para factores psicológicos, sociales, y las condiciones en que se hace la medición; y es altamente probable que si logramos tal grado de control, la reproducibilidad aumente y se pueda considerar a las mediciones así obtenidas como un dato duro, adecuado para análisis estadísticos y altamente confiable.

De lo expuesto se desprende que si un procedimiento que pretendemos utilizar para medir la calidad de vida nos da imprecisiones, no debemos por ello renunciar a la posibilidad de cuantificarla, antes bien, debemos modificarlo o cambiarlo por otro hasta obtener el máximo de consistencia y precisión.

Por otro lado. El surgimiento, en las últimas décadas, de la epidemiología clínica como una disciplina científica independiente ha llevado a una rigurosidad cada vez más depurada en la investigación médica; generándose ensayos controlados cada vez más cercanos a un planteamiento experimental. Sin embargo, en muchas ocasiones, el resultado de interés al comparar un tratamiento contra un placebo sigue siendo una medición de laboratorio o la mortalidad en los grupos comparados, lo que parece habernos hecho olvidar que al médico le importa en primerísimo lugar aliviar a su enfermo, es decir, le importa saber si tal o cual

tratamiento es capaz de mejorar la CALIDAD DE LA VIDA de los individuos sujetos al "experimento".

El planteamiento expuesto no sólo es aplicable a intervenciones terapéuticas concretas, sino en un sentido más amplio al diseño y aplicación de programas de atención a la salud; si es que se quiere que sus resultados sean evaluados de manera sensible y tengan un sentido significativo para las poblaciones en las que se apliquen.

## II.- Marco Teórico.

### B) Diversas Maneras de Medir la Calidad de Vida

El interés por medir la calidad de vida y conceptos relacionados --como funcionalidad, sensación de bienestar, estado de salud, etc.-- empezó a manifestarse en las publicaciones médicas desde mediados de este siglo con los trabajos pioneros de Karnofski, Katz y Sokolow (24,25,34), y ha sido objeto de algunas revisiones recientes que pretenden clarificar nuestros conceptos al respecto (5,13,28.35); sin embargo estamos muy lejos de llegar a un consenso y actualmente coexisten múltiples enfoques que intentan cuantificar la calidad de vida de los individuos.

Retomemos el hilo de nuestro planteamiento teórico ahí donde afirmamos que la valoración de la calidad de vida ha sido soslayada por su "blandura".

La blandura de la medición de la calidad de vida parte de lo abstracto y personal que es este concepto; de ahí que no se tenga una definición universalmente aceptada y que se hayan hecho múltiples intentos para obtener mediciones confiables. En este punto surgen dos cuestiones: El de las "múltiples maneras de medir la calidad de vida" --que desarrollaré en este capítulo. Y el de la "confiabilidad en la medición" --que será objeto del capítulo siguiente.

Existen toda una serie de preguntas y puntos por analizar para poder contestar a la pregunta ¿Cómo medir la calidad de vida?. En los siguientes párrafos intentaré exponer los aspectos que son a mi parecer más importantes.

#### a) ¿Observar o preguntar la calidad de vida?

Es creencia generalizada que un dato de observación es más consistente que un dato derivado de una encuesta; sin embargo, ¿Es esto aplicable a la calidad de vida?. Pongamos un ejemplo:

" Para un locutor de radio el no poder flexionar el índice de la mano derecha muy probablemente no deteriore gravemente su calidad de vida, pero ¿Qué pasa si esto le sucede a un pianista?"



Es obvio que la observación de una incapacidad no es suficiente y que difícilmente podemos "ver" la calidad de vida. Parece entonces razonable recurrir a los métodos de encuesta e individualizar la medición. A fin de incrementar la confiabilidad de las respuestas se debe intentar con diversos subterfugios el aumentar su consistencia; lo cual significa en la mayoría de los casos controlar ESTRICTAMENTE las condiciones de la entrevista (19).

b) ¿A quién se debe preguntar sobre la calidad de vida?

Los lectores de esta tesis, estarán de acuerdo conmigo en que la calidad de vida es un fenómeno eminentemente subjetivo; y que si bien múltiples personas pueden dar una apreciación de la calidad de vida de un individuo, es esperable que dichos individuos ajenos al sujeto juzgen el particular en el marco de su propio concepto de calidad de vida, y no en el marco conceptual del individuo en estudio. Esto es válido también para el personal médico o paramédico, quién no por estar entrenado para la atención a la salud, está más autorizado a juzgar la calidad de vida de un individuo.

Del planteamiento enunciado se desprende una sencilla respuesta a la pregunta: No importa cuán ignorante o diferente culturalmente a nosotros sea el individuo cuya calidad de vida queremos medir; él mismo es quién debe evaluar su calidad de vida.

En todo caso, cuando se considere incapáz (por alteración de sus funciones mentales o motivos similares) al sujeto de evaluar su calidad de vida, quien podría hacerlo sería una persona que la juzgara desde un marco vivencial lo más cercano posible al del enfermo; y aquí la primera opción son los convivientes cercanos, habitualmente sus familiares.

c) ¿Para qué se mide la calidad de vida?

Quando se diseña un instrumento o un procedimiento cualquiera, una de las principales consideraciones es el uso que se pretenda dar a él. De la misma manera, es conveniente tener presente para qué se diseña un instrumento de medición de calidad de vida, antes de su confección; según Kirshner (26) se pueden tener 3 objetivos básicos:

- 1.- Predecir .
- 2.- Discriminar.
- 3.- Evaluar.

#### **Instrumento predictivo.**

Un instrumento predictivo es aquel que se diseña para permitirnos anticipar el valor de cierta variable de interés que es difícil medir directamente. Dicha variable puede ser equiparada entonces a un patrón de oro, o criterio externo de comparación.

Así, instrumentos predictivos en el área de medición de la calidad de vida serían aquellos que nos permitieran predecir la muerte de los sujetos, la necesidad de requerimientos especiales de manejo, la respuesta a un tratamiento, la incapacidad o invalidez, etc.

#### **Instrumento discriminativo.**

Con frecuencia se recurre a instrumentos discriminativos cuando se pretende distinguir entre individuos o grupos en base a una variable de interés, para la cual no se cuente con un estándar de oro o criterio externo definitivo.

Así, podemos estar interesados en saber si cierta enfermedad deteriora la calidad de vida de los sujetos en mayor grado que otra; o si un grupo humano atendido por un determinado sistema de atención a la salud tiene mejor calidad de vida que otro atendido con diferente sistema. Un ejemplo, sólidamente validado, de este tipo de instrumentos es el llamado Perfil de Impacto de la Enfermedad (SIP) (1,2,12).

#### **Instrumento evaluativo.**

Instrumento evaluativo es el que pretende reflejar los cambios que ocurren al transcurrir el tiempo en un sujeto dado, en cierta variable de interés.

Este es el tipo de instrumento de mayor aplicación en los ensayos clínicos controlados, en los que se desea detectar los cambios inducidos por una maniobra terapéutica en un grupo de pacientes bajo estudio. El acento en este caso no está en la calidad de vida como valor absoluto, sino en el cambio en la calidad de la misma.

Se puede hablar de dos enfoques generales en la confección de los instrumentos evaluativos (20,29,31,32). El primero consiste en aplicar un mismo instrumento antes y después de la intervención, midiendo el cambio como la diferencia entre los dos valores obtenidos. El segundo es abordar directamente el cambio, utilizando un índice basal para definir las condiciones previas a la intervención, y un índice de transición después del tratamiento para averiguar específicamente los cambios ocurridos con respecto al índice basal (17,30).

Kirshner y Guyatt (26) han señalado e ilustrado convenientemente que el diseñar un instrumento con alguno de estos fines no asegura, y de hecho muchas veces se opone a, que funcione bien al utilizarlo con los otros objetivos.

Conviene asentar, que dado que el objetivo primordial en esta tesis es analizar instrumentos útiles para su uso en ensayos terapéuticos, nuestro interés primario será en instrumentos EVALUATIVOS de CAMBIO en la calidad de vida.

d) ¿Hay una o varias calidades de vida?

Con respecto a sus áreas de aplicación, se pueden dividir los instrumentos para medición de la calidad de vida en generales y especiales (19,20).

Los instrumentos especiales --que son aplicables solo a cierta enfermedad o grupo de enfermedades-- tienen la ventaja de que pueden enfocarse directamente a las áreas de la calidad de vida del sujeto más afectadas por la enfermedad, o a aquellas con mayores posibilidades de mostrar cierta modificación con el tratamiento instituido. Esta particularidad los haría más sensibles al cambio.

Sin embargo, el uso de tales instrumentos implica disponer de uno para cada enfermedad, o incluso el diseño de uno diferente para cada investigación, y validarlos cada vez. Además, el afirmar que dichos instrumentos tienen mayor sensibilidad al cambio parte de consideraciones teóricas, y antes de aceptarlo como un hecho se debe comprobar bajo comparación directa con otros instrumentos.

Por otro lado existen los instrumentos generales, de aplicación a enfermos o población general sin importar el diagnóstico o tipo del padecimiento. De ellos existen muy buenos ejemplos para estudio de poblaciones y para valorar riesgos de enfermedad y políticas de salud (1,2,33), pero realmente pocos que tengan aplicación directa a pequeños grupos de pacientes en ensayos terapéuticos (30).

## II.- MARCO TEORICO.

C) De cómo Juzgar la Calidad de los Instrumentos que Miden la Calidad de Vida.

A partir de las contribuciones de numerosos autores (3,7,14,15,19,33) podemos afirmar que una buena manera de estimar la calidad de tales instrumentos es analizarlos en base a los siguientes cinco puntos:

- 1.- Reproducibilidad.
- 2.- Consistencia Interna.
- 3.- Validez.
- 4.- Sensibilidad al Cambio.
- 5.- Factibilidad.

### Reproducibilidad.

Es la propiedad de un índice que le permite asignar valores iguales (o muy similares) cuando se aplica en repetidas ocasiones a los mismos sujetos en las mismas condiciones. Se trata de la propiedad principal para conferir "dureza" a los valores del índice (sección II.-A).

El término reproducibilidad es preferible a términos como confiabilidad y consistencia, que se utilizan frecuentemente para referirse a este concepto (15). Ya que en español el término confiabilidad señala a la mayoría de las personas un concepto mucho más amplio que el que nos interesa, y como se usa frecuentemente sin definir estrictamente su significado puede causar confusiones. Por otra parte, el término consistencia no sólo incluye el concepto de reproducibilidad, sino también el de consistencia interna, agregando complicaciones innecesarias.

Independientemente de la palabra que se utilice, el concepto debe definirse con precisión, y aplicarse de manera estricta.

En nuestro caso de valoración de cambio en la calidad de vida, es importante demostrar que factores tales como la personalidad del entrevistador, el lugar donde se aplica el índice, o el tono en que se hace, no interfieran significativamente con la medición.

Esta característica se estima habitualmente calculando la variación en las mediciones obtenidas con el instrumento por diferentes observadores al mismo tiempo (variación inter-observador) o por un mismo observador en diferentes ocasiones (variación intra-observador).

A fin de obtener estimaciones adecuadas de la reproducibilidad es esencial garantizar que las condiciones de los sujetos son efectivamente las mismas en cada una de las ocasiones que se aplica el índice.

Ahora bien, la calidad de vida puede variar de un momento a otro por diversos factores difíciles de controlar, por lo que parece conveniente obtener las diferentes mediciones a un mismo tiempo. De ahí que en términos generales sea preferible medir la reproducibilidad del tipo de índices que nos interesa como la variación inter-observador, y de manera secundaria considerar la variación intra-observador.

#### **Consistencia interna.**

Es la propiedad de un índice que le otorga congruencia a las diferentes partes en un todo armónico. En las ciencias psico-sociales (15) con frecuencia se exige la mayor consistencia interna posible; sin embargo, en nuestro caso no parece ser lo más conveniente.

Así, podemos afirmar que si un índice de calidad de vida está compuesto por 5 diferentes áreas (mental, física, actividades diarias, etc.), es deseable que las mediciones en cada una de ellas muestren congruencia en la dirección de los cambios, pero se espera asimismo que cada parte agregue información no proporcionada por las otras áreas. De hecho si alguna de ellas resultara totalmente repetitiva buscaríamos eliminarla.

Un concepto relacionado con la consistencia interna, es el de explorar la confianza que se puede tener en ciertas respuestas repitiéndolas en varias partes del índice, o incluyendo preguntas corroboratorias.

#### **Validez.**

Este término se refiere al grado de apego de los valores del índice a la realidad del fenómeno estudiado. Es decir, se trata de la propiedad que nos permite aseverar que el índice mide en realidad lo que pretende medir.

Se puede lograr una estimación fácil de la validez cuando se cuenta con una "regla de oro", es decir, con una manera infalible de averiguar la verdad en cuanto al fenómeno que nos interesa. A esta verificación contra un patrón establecido se le llama validación por criterio (5,13,38), siendo de aplicación más directa en los instrumentos predictivos, en los que la ocurrencia del fenómeno que se pretende predecir es claramente el patrón de referencia para medir la validez del índice.

En algunas ocasiones en que no se cuenta con una manera infalible de determinar la situación real del fenómeno que nos interesa, se puede recurrir al subterfugio de confeccionar un patrón de referencia tentativo (15). Así, con frecuencia se recurre al acuerdo entre expertos, intentando de este modo validar instrumentos discriminativos o evaluativos que carecen con frecuencia de un patrón de oro universalmente aceptado.

Sin embargo en el área de valoración de calidad de vida se debe recurrir frecuentemente a métodos alternativos de validación: así, hablamos de validez aparente ("face validity"), validez de contenido (10), y validez de constructo (4,7,14).

La validez aparente se refiere al análisis directo del índice en sí, juzgando si a la luz de los conocimientos del investigador mide lo que pretende medir.

La validación por contenido implica un análisis más profundo de las diferentes partes del índice, considerando si cubre todos los dominios importantes de la variable en estudio; por ejemplo: podría decirse que un índice que pretenda medir salud debe incluir, en base a la definición de la OMS (39), las áreas biológica, psicológica y social del individuo, dando a cada una de ellas atención específica.

La validación por constructo constituye un método más estricto que la validación aparente y la validación por contenido, aplicable en aquellas áreas en las que no se cuenta con un patrón de oro. Recibe este nombre ya que se basa en la enunciación previa de hipótesis (constructos) que se confirman o descartan con un experimento realizado ex-profeso.

Las hipótesis postuladas pueden ser de diferentes tipos: Se pueden comparar los valores del índice con otros parámetros que se sabe (o se supone) varían paralelamente a la variable en estudio. También se puede analizar el desempeño del índice en diferentes situaciones que modifican la variable de interés. Así, por ejemplo, parece razonable hipotetizar que un instrumento que mida capacidad física detecte empeoramiento en el post-operatorio inmediato, o que

exista una correlación positiva entre dicha medición de capacidad física y el desempeño de una ocupación económicamente activa.

#### Sensibilidad al cambio.

Cuando un instrumento se utiliza para discriminar entre grupos o para predecir un determinado suceso, su sensibilidad al cambio es de importancia secundaria; sin embargo, en el caso de instrumentos evaluativos (ver II.-B) esta es una de las propiedades principales.

Por sensibilidad al cambio entendemos, de acuerdo con Guyatt (21), la capacidad de un instrumento para reflejar cambios pequeños en la variable de interés con el transcurso del tiempo. Como cuando se comparan las condiciones antes y después de un tratamiento.

Existen varias formas de evaluar la sensibilidad al cambio, pero todas ellas se basan en el grado de cambio en el instrumento, cuando se piensa (de acuerdo con algún criterio externo) que ha ocurrido un cambio clínicamente significativo. Esta propiedad es fácil de cuantificar cuando se cuenta con una manera infalible para detectar el momento en que ha ocurrido un cambio y su cuantía. Sin embargo, en el área de calidad de vida habitualmente se carece de dicho patrón de referencia y deberemos recurrir a utilizar varios conceptos y variables relacionados con el fenómeno de interés a fin de obtener estimaciones de la sensibilidad al cambio del instrumento en cuestión.

El cálculo más simple de la sensibilidad al cambio probablemente sea el promedio del valor absoluto de cambio en el instrumento, cuando de acuerdo con alguno de los criterios comentados existió efectivamente un cambio. Sin embargo esta estimación ignora la variación propia del instrumento, y habitualmente resulta en valores muy optimistas.

Para corregir esta sobre-estimación se puede hacer un cociente entre dicho promedio absoluto, dividido por la variación interobservador (ver sección de reproducibilidad); o de manera más directa (21) dividido por el cambio promedio en el instrumento cuando según el criterio externo no existe "cambio clínicamente significativo". También se puede utilizar, en lugar del valor absoluto de cambio, a la desviación estándar, varianza, o los cuadrados medios de un análisis de varianza.



Un método alterno, de interpretación más directa y sencilla, es el calcular el valor predictivo positivo y la especificidad del instrumento (11). Es decir, utilizar el instrumento como una prueba diagnóstica, y el criterio externo como el patrón de oro; sin embargo, como ya mencioné, existen inconvenientes teóricos a este respecto, ya que se está considerando como patrón de oro a una variable que se sabe de antemano que no es tal.

### **Factibilidad.**

Una última consideración para juzgar la calidad de un instrumento es la facilidad o dificultad relativas de su aplicación: ¿Se requiere personal con entrenamiento especializado?, ¿Es indispensable un determinado ambiente para su aplicación?, ¿Requiere un grado extremo de colaboración o atención por parte del enfermo?, ¿Cuánto tiempo consume su aplicación?, ¿Cuál es el costo de su uso?

La factibilidad es particularmente importante cuando se pretende la aplicación del instrumento en las condiciones habituales del ejercicio clínico, o cuando la medición de cambio en la calidad de vida forma parte de una extensa batería de pruebas.

### III.- OBJETIVO.

Del marco teórico expuesto se desprende que para medir cambios en la calidad de vida en ensayos terapéuticos controlados, serían de gran utilidad instrumentos generales diseñados específicamente para medir cambio; que además estén debidamente validados, y sean fáciles de aplicar.

Mackenzie y colaboradores (30) han propuesto un índice general de esta naturaleza, y han mostrado evidencia de su reproducibilidad y validez en algunos grupos de enfermos. Sin embargo dichos estudios son aplicables en idioma inglés y población estadounidense, por lo que su uso en nuestra población debe ser precedido (7,8,10) por nueva validación en idioma español y en nuestro grupo cultural.

Por otro lado, nosotros decidimos implementar un instrumento general para medición de la calidad de vida, basado en el uso de escalas visuales análogas, con el fin de estudiar su utilidad.

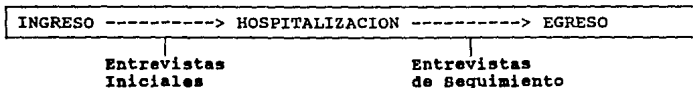
De tal manera, era necesaria una comparación directa de los instrumentos en una misma población latinoamericana. Por lo tanto, el objetivo de esta investigación es:

**"COMPARAR DOS ESCALAS RAPIDAS DE MEDICION DE CAMBIO EN LA CALIDAD DE VIDA EN BASE A SU REPRODUCIBILIDAD, CONSISTENCIA INTERNA, VALIDEZ, SENSIBILIDAD AL CAMBIO Y FACTIBILIDAD".**

## IV.- MATERIAL Y METODOS.

## A) El Diseño Básico.

La investigación realizada es un estudio prospectivo consistente en varias entrevistas, durante las cuales se contestaron un número determinado de cuestionarios. El diseño básico se puede esquematizar de la siguiente manera:



Al ingreso se captaron los enfermos que cumplieran con los criterios de inclusión y aceptaban participar en el estudio, anotando sus datos en una forma especial (ver apéndice II) que incluía ficha de identificación y diagnósticos del paciente. Para cada caso se determinaron, antes de su inclusión, los entrevistadores y el orden de entrevista, a fin de contar con todas las combinaciones posibles; así cada nuevo paciente fué asignado a los entrevistadores predestinados de acuerdo con la secuencia de admisión al estudio.

Las entrevistas iniciales fueron 5 en total, realizadas en el primero y segundo días de hospitalización como sigue:

	Día 1	Día 2
*Primer entrevistador	INSTRUMENTO A	INSTRUMENTO B
*Segundo        "	INSTRUMENTO B	INSTRUMENTO A
‡Tercer         "	KARNOFSKI	

\* Médico o trabajadora social.  
‡ Médico.

Los entrevistadores para cada enfermo siempre fueron los mismos ; el primer y segundo entrevistador desconocían el valor de la escala de Karnofski (22,24) asignado al enfermo; y a su vez, el médico que valoraba el Karnofski no participaba en la aplicación de los instrumentos a dicho enfermo, desconociendo los resultados obtenidos durante su aplicación.

En los días previos al alta y después de múltiples intervenciones terapéuticas decididas por el médico tratante del paciente (no relacionado con los participantes en la investigación), se realizaban las entrevistas de seguimiento:

	Día previo al alta	Día del alta
*Primer entrevistador	INSTRUMENTO B	INSTRUMENTO A
*Segundo "	INSTRUMENTO A	INSTRUMENTO B
‡Tercer "	KARNOFSKI Opinión del médico tratante. Opinión del familiar.	

\* Médico o trabajadora social.  
‡ Médico.

Tanto el médico tratante como el familiar desconocían los resultados de los instrumentos y la calificación de Karnofski. El familiar interrogado en cuanto al cambio en la calidad de vida era el que se encontraba con el paciente al momento del alta; y en caso de existir varios, el que conviviera más estrechamente con el enfermo.

#### IV.- MATERIAL Y METODOS.

##### B) Los Pacientes.

Los individuos estudiados se eligieron entre los enfermos hospitalizados en el Instituto Nacional de la Nutrición Salvador Zubirán durante los meses de Septiembre y Octubre de 1988.

El Instituto Nacional de la Nutrición Salvador Zubirán es un Instituto Nacional de Salud Descentralizado, dependiente del Gobierno de los Estados Unidos Mexicanos, que cuenta entre otras instalaciones, con un centro hospitalario de referencia de cobertura nacional; destinado a la atención de problemas médicos y quirúrgicos especializados, particularmente padecimientos gastro-intestinales, endocrinopatías, nefropatías, trastornos hematológicos, enfermedades auto-inmunes, e infecciones graves.

Dentro del grupo de enfermos admitidos al hospital, se consideraron candidatos al estudio a los pacientes con residencia en el área metropolitana de la Ciudad de México, que podían leer y escribir, y que se encontraban en condiciones de ser entrevistados. Se propuso diariamente el estudio a los pacientes al momento de su ingreso, hasta obtener un máximo de 4 casos diarios y se recabó el consentimiento por escrito.

Se excluyeron los enfermos residentes fuera de la ciudad de México, y aquellos cuyas condiciones mentales o de extrema gravedad no permitían la aplicación de cuestionarios. Asimismo se excluyeron las personas que declinaron participar en el estudio.

#### IV.- MATERIALES Y METODOS.

##### C) De los Instrumentos y Parámetros de Comparación.

###### Instrumento No. 1.

En adelante se llamará Instrumento No. 1 a la traducción de la escala propuesta por MacKenzie (30), tal como se muestra en la figura IV-1.

Este instrumento está integrado por dos partes, aplicadas en diferentes entrevistas: la primera parte es la valoración inicial o basal, que se utiliza exclusivamente para elegir una actividad concreta, sobre la cual juzgar el cambio cuando se aplica la segunda parte (valoración subsecuente o de seguimiento).

La medición de cambio se hace para cada una de las áreas integrantes del instrumento (física, mental y emocional) en una escala ordinal de cinco puntos que va de -2 a +2 (mucho peor, ligeramente peor, igual, ligeramente mejor, mucho mejor). Se calculó un índice global con la suma algebraica de las tres áreas (escala de 13 grados; desde -6 hasta +6, pasando por cero). Cabe hacer notar que tal índice global no está contemplado en la versión original del índice, y se utilizó en nuestro caso para facilitar cálculos y comparaciones.

En el apéndice I se consigna la forma original del instrumento en inglés, y en el apéndice II un grupo de instrucciones precisas para su aplicación; estas últimas las consideramos fundamentales para disminuir la variación explicable debido al entrevistador. Cabe comentar que la forma final en español que se muestra fué el resultado de varias modificaciones realizadas en base a la traducción inversa del instrumento (traducción del español al inglés por un médico de habla materna inglesa no relacionado con el estudio) (8), y también con base en su aplicación a un grupo "piloto" de pacientes antes de iniciada la investigación que nos ocupa (9). Los entrevistadores aplicaron varias veces el instrumento y acordaron sobre los malentendidos y dificultades suscitados, y la manera de evitarlos.



B) Concentración -- Mis siguientes preguntas tienen que ver con cosas que usted hace que requieren esfuerzo mental o concentración. De nuevo, estas pueden ser actividades relacionadas con su trabajo, los quehaceres domésticos, o alguna otra cosa que le interese. Le voy a preguntar qué actividad de las que usted hace requiere la mayor concentración o esfuerzo mental.

1. De las cosas que usted hace habitualmente ¿Qué actividad mental requiere la mayor concentración o esfuerzo mental? \_\_\_\_\_ (B1).

2. En sus días buenos ¿Cuánto sería lo más que puede... (Actividad mental de la pregunta 1)? \_\_\_\_\_ (B2).  
Y, en sus días malos ¿Cuánto sería lo más que puede... (Actividad mental de la pregunta 1)? \_\_\_\_\_ (B2).

C) Emocional -- Mis siguientes preguntas tienen que ver con las cosas emocionalmente difíciles o preocupantes que pueden haberle pasado recientemente. Le voy a preguntar si alguna de las cosas que han estado pasando en su vida ha sido particularmente preocupante. Esto podría incluir su salud, preocupaciones familiares o económicas, su trabajo, o cualquier otra preocupación que haya experimentado.

1. ¿Hay alguna cosa que está pasando en su vida que haya sido particularmente preocupante o emocionalmente difícil? \_\_\_\_\_ (C1).

2. (Si contesta si) ¿Qué tipo de problema es? \_\_\_\_\_ (C2).

3. ¿Cómo lo (la) está afectando esta situación? \_\_\_\_\_ (C3).





3. (Si contesta si) ¿Es esta aún la actividad mentalmente más fatigante que usted podría hacer? \_\_\_\_\_(B3).

4. (Si contesta no) ¿Cual es la actividad mentalmente más fatigante que usted podría hacer ahora? \_\_\_\_\_(B4).

C) Emocional.

1. Estoy interesado en como está usted afrontando ahora las preocupaciones de su vida. ¿Diría usted que su habilidad para afrontar las preocupaciones, desde la última vez que platicamos, es mucho mejor, ligeramente mejor, igual, ligeramente peor, o mucho peor? \_\_\_\_\_(C1).

2. ¿Ha experimentado alguna cosa desde la última vez que platicamos, que usted consideraría preocupante? \_\_\_\_\_(C2).

3. (Si contesta si) ¿Cómo lo está afrontando? \_\_\_\_\_(C3).

## Instrumento No. 2.

El instrumento No. 2 es el diseñado por nuestro grupo en base a escalas visuales análogas; consta también de una parte basal o inicial (ver figura IV-2), en relación a la cual se juzga el cambio en la segunda parte (subsecuente o de seguimiento).

El desarrollo de la forma final del instrumento incluyó los siguientes pasos:

1.- Elección de los rubros a interrogar (6). Los investigadores involucrados en el estudio (tres médicos y dos trabajadoras sociales) se reunieron, aportando cada uno de ellos los rubros que consideraran importantes para juzgar calidad de vida, en el sentido amplio del término y tomando como base la definición de salud de la OMS (39). En reunión posterior se acordó incluir en el instrumento las siguientes áreas:

- Física.
- Mental.
- Actividades diarias.
- Emocional.
- Convivencia.
- Dolor.

2.- Como segundo paso, se decidió utilizar escalas visuales análogas de 10 cm, horizontales y con divisiones cada centímetro (36).

3.- A continuación se dió un formato homogéneo a las preguntas, fraseándolas más como indicaciones directas que como interrogaciones; y designando homogéneamente el extremo derecho de la escala basal como el de ausencia de compromiso por la enfermedad. Las escalas de seguimiento se constituyeron de la manera mas simétrica posible (29), correspondiendo el centro de la línea a "sin cambio".

4.- Se realizaron aproximadamente 20 entrevistas piloto por cada uno de los entrevistadores participantes en el estudio. En base a esta experiencia se decidió hacer algunas modificaciones y formular lineamientos para el uso del instrumento (apéndice II-3).



3.- Señale cuánta ayuda necesita para sus actividades diarias ahora que ha estado enfermo.

Me tienen  
que ayudar \_\_\_\_\_ Ninguna  
en todo

4.- Señale cómo siente su estado de ánimo ahora que ha estado enfermo.

Peor que \_\_\_\_\_ Mejor que  
nunca \_\_\_\_\_ nunca

5.- Señale que tanto convive con los demás (familiares, amigos, compañeros de trabajo) ahora que ha estado enfermo.

No puedo \_\_\_\_\_ Más que  
convivir \_\_\_\_\_ nunca  
con nadie

6.- En caso de que tenga usted dolor, señale qué tanto lo afecta.

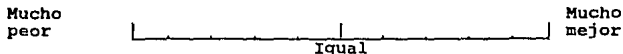
No lo \_\_\_\_\_ No me da  
soporto \_\_\_\_\_ problema

7.- En general, señale cómo siente que es su vida actualmente.

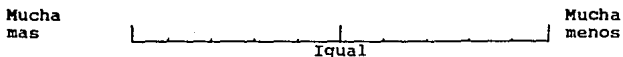
Peor que \_\_\_\_\_ Mejor que  
nunca \_\_\_\_\_ nunca



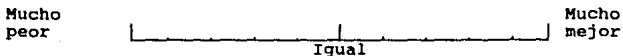
2.- Su capacidad para pensar, en relación a la última vez que nos vimos, está...



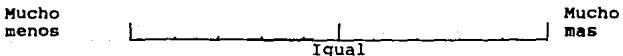
3.- La ayuda que necesita para sus actividades diarias, en relación a la última vez que nos vimos, es...



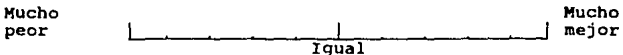
4.- Su estado de ánimo, en relación a la última vez que nos vimos está...



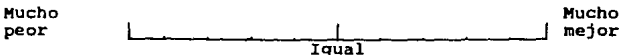
5.- En relación a la última vez que nos vimos, usted puede convivir con los demás...



6.- En relación a su dolor (si es que tiene dolor), ahora se siente usted...



7.- En general, usted siente que su vida, en relación a la última vez que nos vimos, es...



5.- Durante la aplicación piloto fué evidente la necesidad de un ejemplo inicial que permitiera familiarizar al enfermo con el uso de la escala visual análoga. El ejemplo debía ser lo más similar posible a las preguntas integrantes del índice, y después de varios ensayos se decidió incluir la pregunta en relación con la lluvia que aparece en la forma final del instrumento.

6.- Si bien inicialmente se había pensado en una escala de 100 puntos (es decir, en milímetros), pronto fué aparente que diferencias de unos pocos milímetros raramente señalaban diferencias reales, y que los pacientes tendían a señalar los puntos marcados cada centímetro en la escala. Se decidió entonces utilizar la medición en cm, es decir de 10 puntos.

Consideramos que es particularmente importante la inclusión y debida ilustración del ejemplo inicial, si se quiere garantizar un uso confiable de las escalas visuales que constituyen las respuestas del instrumento.

En la medición inicial (basal) se confeccionó un índice global con la suma de las 5 primeras preguntas: como cada una va de cero (en el extremo izquierdo) hasta 10 ( en el extremo derecho), la suma va de cero a 50.

Para la medición de cambio (entrevista subsecuente) el paciente tuvo a la vista sus respuestas previas, a fin de aumentar la reproducibilidad del instrumento (18); en este caso a cada pregunta se asigna un valor que va de -5 (máximo empeoramiento, extremo izquierdo) a +5 (máxima mejoría, extremo derecho). El índice agregado o global es la suma algebraica de las cinco primeras respuestas, y tiene un rango posible de -25 a +25.



### Parámetros de Comparación.

A fin de realizar un juicio sobre la validez y sensibilidad al cambio de los instrumentos se registraron simultáneamente las siguientes variables:

Cambio en Karnofski: Medido como la diferencia entre dos aplicaciones de esta escala por un mismo observador médico (22,24); las formas utilizadas se consignan en el apéndice II-4, y el rango posible es de -9 (máximo empeoramiento) a +9 (máxima mejoría).

Opinión del paciente: Pregunta directa en forma de escala visual análoga, que corresponde a la pregunta 7 del instrumento 2. Ya que se tienen dos versiones de esta pregunta ( dos entrevistadores), se calculó un promedio.

Opinión del médico tratante y del familiar: Pregunta directa (apéndice II-5) graduada en 5 puntos (mucho peor, ligeramente peor, igual, ligeramente mejor, mucho mejor).

Asimismo con cada instrumento se incluyó una hoja (apéndice II-6) donde se consignó el tiempo necesario para completar la entrevista y la facilidad o dificultad relacionadas con ella, según el entrevistado y el entrevistador.

#### IV.- MATERIAL Y METODOS.

##### D) El Análisis.

Además de describir los resultados en tablas, figuras y proporciones se consideró necesario estimar matemáticamente las diferentes propiedades de los instrumentos; es decir : reproducibilidad, consistencia interna, validez, sensibilidad al cambio y factibilidad. Para la estimación numérica de estas propiedades se consideraron significativos aquellos valores con probabilidad de azar menor al 0.05, utilizándose siempre pruebas estadísticas bilaterales (de dos colas).

##### Reproducibilidad.

Se estimó en base a la concordancia inter-observador, para dos aplicaciones del instrumento por diferentes entrevistadores ( un médico y una trabajadora social), espaciadas 24 hs entre sí. Para el instrumento 1 se calcularon el porcentaje de acuerdos y la prueba de Kappa ponderada ( $K_w$ , ver formulario en apéndice III). Para el instrumento 2 se calculó además de el coeficiente de correlación de Pearson ( $r$ ) y de la prueba "t" pareada, un índice de concordancia más estricto (27): el coeficiente de correlación intraclase ( $R_i$ , ver formulario), obtenido a partir de un análisis de varianza para mediciones repetidas.

Kramer y Feinstein (27) recomiendan interpretar de manera similar los valores de  $K_w$  y  $R_i$ , de forma que podemos considerar:

VALOR	CONCORDANCIA
0.81 a 1.0	Casi perfecta.
0.61 a 0.80	Substantial.
0.41 a 0.60	Moderada.
0.21 a 0.40	Pobre.
0.01 a 0.20	Escasa.
Menor de 0	Mala.

##### Consistencia Interna.

La consistencia interna del instrumento No. 1 se estimó calculando la congruencia entre las preguntas por medio del coeficiente  $F_i$  (27) (ver formulario en el apéndice). Para el instrumento No. 2 se estimó como la correlación entre cada pregunta con la calificación total por medio de coeficientes de correlación de Spearman ( $r_s$ ).

**Validez.**

El aspecto más importante para juzgar la validez de los instrumentos en estudio se explora por medio de la validación de constructo, ya que como se comentó previamente, carecemos de un patrón de oro para juzgar calidad de vida.

Con base en el análisis de las características propias de ambos instrumentos, y de los diversos parámetros de comparación, podemos realizar las siguientes predicciones o hipótesis:

1.- Debe existir, una correlación, al menos moderada, entre la valoración de Karnofski y ambos instrumentos. Sin embargo, considerando que la escala de Karnofski se refiere a capacidad física y función, se predice que habrá una mejor correlación entre Karnofski y el área física del instrumento No. 1; asimismo se espera una mejor correlación entre Karnofski y las preguntas 1 y 3 del instrumento No. 2, que hacen referencia a capacidad física y autonomía en actividades de la vida diaria, respectivamente.

2.- Ya que ambos instrumentos cuestionan directamente al enfermo, se espera una buena correlación con la opinión del paciente, la cual probablemente sea más fuerte que con la opinión del familiar y del médico tratante. Además, dado que la opinión del paciente está consignada en forma de escala visual análoga es posible que correlacione más con el instrumento No. 2 que con el No. 1.

3.- En relación a la opinión del médico; si bien se espera una moderada correlación con ambos instrumentos, se predice una mayor correlación con las preguntas del área física, y menor con las preguntas de convivencia y del área emocional.

4.- Por último, se espera una buena correlación con la opinión del familiar, si bien dicha correlación será probablemente más fuerte con las áreas de convivencia y actividades diarias del instrumento No. 2, y con el área emocional del instrumento No. 1.

La correlación entre cada uno de los instrumentos y las variables de contraste se analizará con tabulación de los valores obtenidos, y con el cálculo del coeficiente de correlación de rangos de Spearman ( $r_s$ ) y la probabilidad de azar asociada.

### **Sensibilidad al cambio.**

Esta propiedad será analizada de dos maneras:

1.- Tabulando el grado de cambio en ambos instrumentos cuando existen cambios discretos en los parámetros de comparación.

2.- Calculando un índice de sensibilidad al cambio como el cociente de la variación absoluta promedio en los pacientes que efectivamente cambiaron según cada parámetro de comparación, dividida por la variación absoluta promedio en los pacientes que no cambiaron según ese mismo parámetro de comparación. Para mayores detalles se puede consultar la fórmula (apéndice III-D) y ejemplos (apéndices IV-3 y IV-9) que se adjuntan. Los valores cercanos a 1.0 se interpretan como carencia de sensibilidad al cambio; los valores mayores a 1 significan que existe cierta sensibilidad al cambio, y mientras mayores sean reflejan una mayor sensibilidad.

### **Factibilidad.**

Se analizó en base al tiempo promedio requerido para las diferentes entrevistas, y la proporción de entrevistados y entrevistadores que consideró sencilla o muy sencilla su aplicación. La duración de la entrevista se presenta como promedio y desviación estandar.

V.- RESULTADOS.

A) Datos Generales.

Se captaron de manera inicial 40 pacientes que ingresaron al Instituto Nacional de la Nutrición Salvador Zubirán durante los meses de Septiembre y Octubre de 1988. De los cuales, 23 pacientes completaron satisfactoriamente el protocolo de estudio.

Las razones de exclusión de los pacientes se ilustran en la figura V-1.

De ella se desprende que los datos sobre valoración inicial se basan en un total de 39 casos, mientras que los datos de seguimiento y por tanto las mediciones de cambio se han calculado a partir de una muestra de 23 pacientes. Las características generales de ambos grupos se consignan en la tabla V-2.

Los diagnósticos principales de los pacientes estudiados fueron diversos, y se enumeran en la tabla V-3.

Figura V-1.

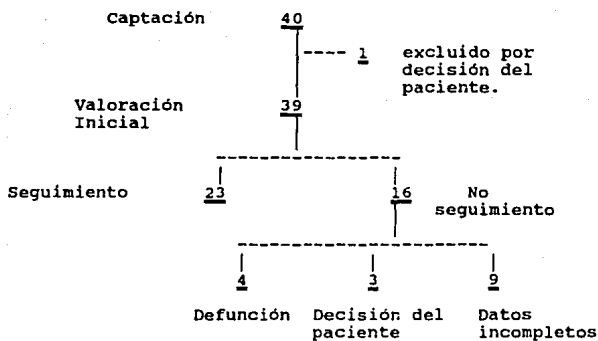


Tabla V-2. Características de los pacientes estudiados.

Característica	Grupo Inicial	Grupo de Seguimiento
Número	39	23
Edad*	55 (13-83)	52 (13-80)
Sexo (F/M)	21/18	13/10
Casos quirúrgicos	11 (28%)	5 (22%)
Casos médicos	28 (72%)	18 (78%)
Enf. agudas	12 (31%)	8 (35%)
Enf. crónicas	25 (64%)	14 (61%)
Sanos**	2 (5%)	1 (4%)
Karnofski < 80	18 (46%)	13 (57%)
Clasificación económica baja	29 (74%)	18 (78%)

\* Mediana (rango).

\*\* Un donador para trasplante renal y un caso de fiebre en estudio no confirmada.

**Tabla V-3. Diagnósticos individuales principales de los enfermos estudiados.**

Seguimiento Completo	Seguimiento Incompleto
<p>Carcinoma metastásico            CA pulmonar estadio III            Leucemia aguda            Lupus Eritematoso            Lupus Eritematoso            Cirrosis            Fibrosis pulmonar            Insuficiencia renal            Enf. injerto contra            huesped            Diabetes insípida            Insuficiencia cardiaca            Diabetes y neuritis            post-herpética            Absorción intestinal            deficiente.            Intolerancia a lacteos            Diabetes y osteomielitis            Pielonefritis aguda            Hepatitis            Ureterolitiasis            Trombosis venosa            Trastorno de la estática            pélvica.            Duodenitis            Gastritis            Sano (fiebre en estudio)</p>	<p>Carcinoma metastásico            Carcinoma metastásico            Carcinoma metastásico            Carcinoma gástrico            Carcinoma pancreático            Cirrosis            Cirrosis            Insuficiencia pancreática            Espondilitis tuberculosa            Crecimiento de silla            turca            Obesidad            Depresión            Infartos renales            Fistulas ileocutáneas            Nefrolitiasis            Colecistolitiasis            Donadora renal sana</p>



## V.- RESULTADOS.

### B) Instrumento No. 1 .

Las actividades más frecuentemente señaladas como las más fatigantes se consignan en la tabla V-4; al comparar las respuestas que se dieron a ambos entrevistadores se obtuvo una concordancia interobservador del 45% (14/31) para la actividad física, un 56.7% (17/30) para la actividad mental, y un 60% (18/30) para la esfera emocional (tabla V-5).

Cuando se contestó la misma actividad a ambos entrevistadores, y la cuantificación se realizó en las mismas unidades se obtuvo una concordancia de 60% para la cantidad de actividad física máxima, y del 50% para la cantidad máxima de actividad mental. Los tipos de unidades más frecuentemente utilizados fueron: unidades de tiempo (42.6% de las ocasiones), y el número de objetos o actividades concretas que se podían realizar (40.1% de las respuestas).

### Reproducibilidad.

Los datos de concordancia interobservador en la medición del cambio se resumen en la tabla V-6. La reproducibilidad fué menor para el área física que en las otras dos, con una Kappa ponderada de 0.39 . Las tablas de contingencia en que se basan tales cálculos se consignan en el apéndice IV. Para el índice global se calculó  $Kw=0.45$  . Cabe hacer notar que para el cambio global (suma de las 3 áreas) la medición permite un rango de 13 grados: desde -6 a +6 pasando por cero.

Al calcular las diferencias en la calificación global de un entrevistador y la del otro en cada paciente y promediarlas se obtuvo una media de 1.37 y una desviación estandar de 1.15 . Lo que sugiere que un cambio de hasta 1.37 unidades en este instrumento debe ser considerado con reserva, pues equivale a la variación inter-observador promedio .

### Consistencia Interna.

La consistencia interna del Instrumento No. 1 se puede estimar por la congruencia entre las respuestas a las preguntas uno y dos de cada área (figura IV-1): de manera que si el paciente dice estar igual o mejor (en la entrevista de seguimiento), se espera que diga que puede seguir haciendo la actividad que mencionó en la entrevista inicial; mientras que si asevera estar peor, debe contestar que no puede seguir realizando la actividad en cuestión.

**Tabla V-4. Actividades informadas como las más extenuantes durante la aplicación del Instrumento No. 1.**

Actividad física (%)	Actividad mental (%)	Emocional(%)
Labores domésticas 27.4	Actividades laborales 28.6	Salud propia 46.8
Subir escaleras 17.7	Leer 23.8	Problemas familiares 24.2
Caminar 14.5	Presupuesto familiar 20.6	Problemas económicos 9.7
Actividades laborales 12.9	Cálculos matemáticos 7.9	
Correr 8.1	Ver televisión 6.3	

**Tabla V-5. Actividades reportadas como las más extenuantes. Comparación de ambos entrevistadores.**

Area	Misma Respuesta	Diferente Respuesta	Desconocido
Actividad física	14	17	8
Actividad mental	17	13	9
Emocional	18	12	9

**Tabla V-6. Concordancia interobservador en la medición de cambio con el instrumento No. 1.**

	Area física	Area mental	Area emocional	Global
Acuerdo	7½	11	12	4
Desacuerdo un grado	9	6	7	6
Desacuerdo dos grados	1	-	-	2
Desacuerdo tres grados	-	1	-	4
Kw	0.394	0.57	0.584	0.45
n	17	17	19	16

½ .Número de pacientes en cada categoría de acuerdo o desacuerdo.

Como se observa en la Figura V-7, la consistencia interna fué significativamente mayor cuando aplicó el instrumento la T. Social. Y, a excepción del área mental interrogada por el médico, se aprecia una congruencia aceptable congruencia, lo que fortalece razonablemente la existencia de consistencia interna.

**Validez.**

En la tabla V-8 se consigna la comparación de los valores obtenidos en la calificación global del Instrumento No. 1 con los parámetros de contraste: cambio en Karnofski, opinión del paciente, opinión del médico tratante, y opinión del familiar. Para facilitar una comparación más directa, también se incluyeron en la tabla las calificaciones obtenidos con el instrumento No. 2; esta tabla constituye la parte medular de los resultados del estudio.

A partir de los datos de dicha tabla se puede calcular la correlación del instrumento con cada una de las variables de contraste; esto se muestra de manera gráfica en la figura V-9 y cuantitativamente en las tablas V-10 y V-11, destacando la ausencia de correlación estadísticamente significativa con el cambio en la escala de Karnofski, y una muy buena correlación con la opinión del familiar. En términos generales la correlación fué mejor cuando el instrumento lo aplicó la trabajadora social.

Cuando se analizó la correlación del área física del instrumento No. 1 con el cambio en Karnofski, se obtuvieron coeficientes de correlación de Spearman de 0.369 (p=0.1) cuando lo aplica el Médico, y de 0.504 (p=0.0381) cuando lo aplicó la trabajadora social.

**Sensibilidad al cambio.**

Tomando cada uno de los parámetros de comparación como si se tratara de la regla de oro, se construyeron tabulaciones para la variación de la calificación obtenida con el instrumento No. 1 --tanto cuando lo aplica el médico, como cuando lo aplica la T. social--. Los resultados se anotan en las tablas V-12 a V-15.

La interpretación de dichas tablas es simple, por ejemplo:

Tabla V-7. Consistencia interna del Instrumento No. 1.

## A) Cuando lo aplica el médico.

Puede seguir haciendo  
la actividad.

		SI	NO	
Area Física	Igual o mejor	13	5	76% de congruencia fi= 0.52
	Peor	0	3	
Area Mental	Igual o mejor	14	5	68.2% de congruencia fi= 0.13
	Peor	2	1	

## B) Cuando lo aplica la trabajadora social.

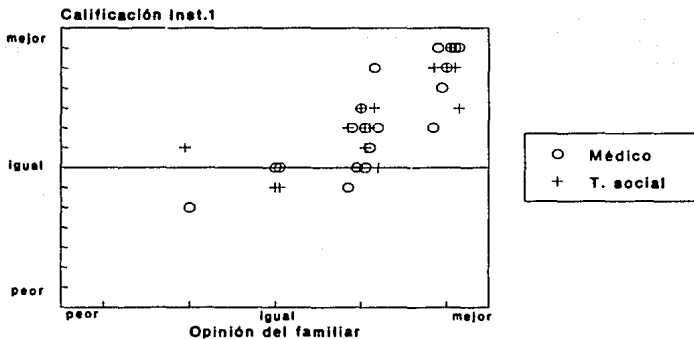
Puede seguir haciendo  
la actividad.

		SI	NO	
Area Física	Igual o mejor	11	2	86.7% de congruencia fi= 0.65
	Peor	0	2	
Area Mental	Igual o mejor	13	1	93.7% de congruencia fi= 0.79
	Peor	0	2	

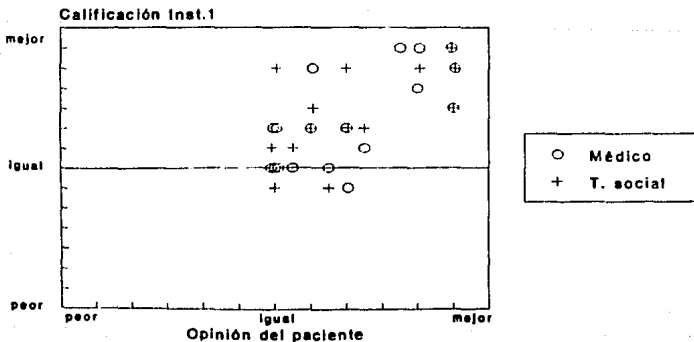
**Tabla V-8. Resultados de las mediciones de cambio de ambos Instrumentos y los parámetros de contraste.**

Caso	Inst. No. 1		Cambio en Karnofski	Opinión			Inst. No. 2	
	Med.	T.S.		Pac.	Med	Fam.	Med.	T.S.
1	3	3	0	5	2	1	18	9
2	5	5	2	5	0	2	21	21
3	2	2	1	1	0	1	5	12
4	0	-1	1	0	-1	0	-1	-3
5	6	6	2	5	2	2	25	25
6	4	-	3	4	2	2	17	20
7	-2	-	-5	-	-2	-1	-	-8
8	0	1	2	0.5	1	1	12	0
9	0	-	-1	-	1	1	-	-4
10	0	-1	-	1.5	-2	0	-7	-5
11	2	5	3	2	1	0	9	8
12	1	2	2	2.5	0	1	-1	5
13	6	5	0	4	2	2	10	23
14	2	-	-2	-	1	1	-	11
15	6	-	1	3.5	2	2	19	20
16	6	3	1	-	0	2	-	20
17	-	1	-1	-	-1	-1	0	-
18	5	3	0	1	1	1	8	-4
19	-1	2	-1	2	1	1	4	-1
20	2	5	1	0	1	2	15	10
21	2	0	0	0	2	1	5	0
22	0	-	0	0	1	-	0	-3
23	0	1	2	0	0	-	1	4
Mediana	2	2	1	1.75	1	1	9	6.5
Rango	-2	-1	-5	0	-2	-1	-7	-8
	6	6	3	5	2	2	25	25

**Figura V-9a.**  
Inst.1 / Opinión familiar.



**Figura V-9b.**  
Inst.1 / Opinión paciente.



**Tabla V-10.** Comportamiento de los parámetros de comparación cuando el Instrumento No. 1 es aplicado por el médico.

CALIFICACION DEL INST. 1	No. casos	Cambio en Karnofski	Pac.	Opinión Méd.	Fam.
< 0	2	-3	2.0	-0.5	0.0
0 o 1	7	1.0	0.75	0.0	0.6
2 o 3	6	0.5	1.6	1.17	1.0
4 o 5	3	1.67	3.3	1.0	1.67
6	4	1.0	4.17	1.5	2.0
Desconocido	1	--	--	--	--
$r_s$		0.311	0.521½	0.557½	0.751*

½ p < 0.05

\* p < 0.001

**Tabla V-11.** Comportamiento de los parámetros de comparación cuando el instrumento No. 1 es aplicado por la T. social.

CALIFICACION DEL INST. 1	No. casos	Cambio en Karnofski	Pac.	Opinión Méd.	Fam.
< 0	2	1.0	0.75	-1.5	0.0
0 o 1	4	0.75	0.17	0.5	0.33
2 o 3	6	0.5	2.3	0.67	1.67
4 o 5	4	1.5	2.75	1.0	1.5
6	1	2.0	5.0	2.0	2.0
Desconocido	6	--	--	--	--
$r_s$		0.289	0.655½	0.555½	0.758*

½ p < 0.05

\* p < 0.001



Si la regla de oro es la escala de Karnofski, el instrumento No. 1 aplicado por el médico carece de sensibilidad al cambio (tabla V-12); pues no importa que el Karnofski mejore una o dos unidades, el instrumento sigue marcando valores muy similares a cuando no hay cambio (3.2 y 2.4 comparados con 3.2).

Por otro lado, si la regla de oro es la opinión del familiar entonces el instrumento No. 1 aplicado por el médico es extraordinariamente sensible (tabla V-14), ya que el cambio de una o dos unidades en la opinión del familiar se refleja en amplios cambios en las calificaciones del instrumento (1.56 y 5.0 comparado con cero).

El segundo enfoque para estimación de sensibilidad al cambio es el cálculo del índice propuesto por Guyat (ver material y métodos). Los resultados se muestran al pie de las tablas mencionadas. En general el Instrumento No. 1 es muy sensible a los cambios en la opinión del familiar y el paciente, y poco sensible a los cambios en Karnofski y en la opinión del médico tratante.

#### **Factibilidad.**

La entrevista inicial fué completada por el médico en 4.53 +/- 1.43 minutos (media +/- desviación estándar), mientras que la trabajadora social necesitó de 7.1 +/- 3.28 minutos. El paciente consideró que las preguntas eran sencillas o muy sencillas el 91 % de las ocasiones; mientras que el entrevistador consideró sencilla o muy sencilla la aplicación del instrumento el 85.1 % de las veces.

Los tiempos promedio para la entrevista de seguimiento fueron de 3.68 +/- 1.0 y 4.25 +/- 1.0 (médico y trabajadora social respectivamente). La opinión del paciente fué de sencillo o muy sencillo el 97.4 % de las veces, obteniéndose el mismo valor para el entrevistador.

**Tabla V-12.** Sensibilidad al cambio del instrumento No. 1 según el cambio en el Karnofski.

Cambio en Karnofski	No. de casos	Inst. 1 Médico*	Inst. 1 T. social*
< -2	2	0.0	--
-1	3	-0.5	1.5
0	5	3.2	2.75
1	5	3.2	2.25
2	5	2.4	3.0
3	2	3.0	5.0
Desconocido	1	---	---
Indice de sensibilidad		0.76	1.03

\* Calificación promedio de el No. de casos anotado en cada renglón.

**Tabla V-13.** Sensibilidad al cambio según opinión del paciente.

Opinión paciente	No. de casos	Inst. 1 Médico*	Inst. 1 T. social*
< 0	0	--	--
0.0 - 0.9	6	0.67	1.2
1.0 - 2.9	6	1.5	2.16
3.0 - 5.0	6	5.0	4.75
Desconocido	5	---	---
Indice de sensibilidad		4.36	1.98

\* Calificación promedio de el No. de casos anotado en cada renglón.

**Tabla V-14. Sensibilidad al cambio según opinión del familiar.**

Opinión familiar	No. de casos	Inst. 1 Médico*	Inst. 1 T. social*
-1	2	-2.0	0.0
0	2	0.0	1.0
1	9	1.56	1.86
2	7	5.0	4.8
Desconocido	3	--	--
Indice de sensibilidad		infinito <sup>‡</sup>	2.92

\* Calificación promedio de el No. de casos anotado en cada renglón.

‡ División por cero.

**Tabla V-15. Sensibilidad al cambio según opinión del médico tratante.**

Opinión M. tratante	No. de casos	Inst. 1 Médico*	Inst. 1 T. social*
-2	2	-1.0	-1.0
-1	2	0.0	0.0
0	5	2.8	2.6
1	8	1.25	3.2
2	6	4.5	3.5
Indice de sensibilidad		0.86	0.99

\* Calificación promedio de el No. de casos anotado en cada renglón.

## V.- RESULTADOS.

### c) Instrumento No. 2.

En la aplicación basal del instrumento No. 2 la calificación promedio cuando lo aplicó el médico fué de 25.3, y para la T.social fué de 23.2 ; recordemos que cero corresponde a pésima calidad de vida, mientras 50 corresponde a máximo bienestar.

La gama de calificaciones obtenidas se muestra en la figura V-16.

En cuanto a la medición del cambio, es decir, la calificación obtenida durante la entrevista de seguimiento, los valores obtenidos se han anotado en la tabla V-8, y se muestran en forma gráfica en la figura V-17.

### Reproducibilidad.

El instrumento mostró una concordancia substancial en la medición basal, con un  $R_i=0.7$  para la calificación global ( $r=0.75$ ,  $t=0.53$ ,  $p>0.5$ ). Sin embargo, al analizar la concordancia de cada una de las preguntas integrantes de el instrumento, esta varió desde un  $R_i=0.14$  para la pregunta 2, hasta un  $R_i=0.455$  para la pregunta 5. Los datos completos a este respecto se muestran en el apéndice IV-6.

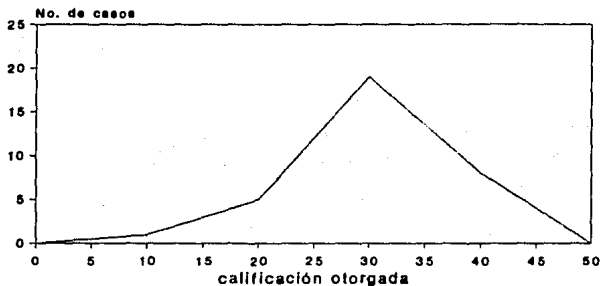
En cuanto a la reproducibilidad en la medición de cambio, que es la más importante, la concordancia global fué substancial ( $R_i=0.69$ ,  $r=0.78$ ,  $t=0.73$ ). Destacando la alta concordancia de las preguntas 1, 3 y 5 ( $R_i$  de 0.76, 0.74, y 0.67, respectivamente).

La variación promedio inter-observador fué de 4.9 (DE=4.14), por lo que los cambios de hasta 5 unidades en la calificación del instrumento deben ser tomados con reserva.

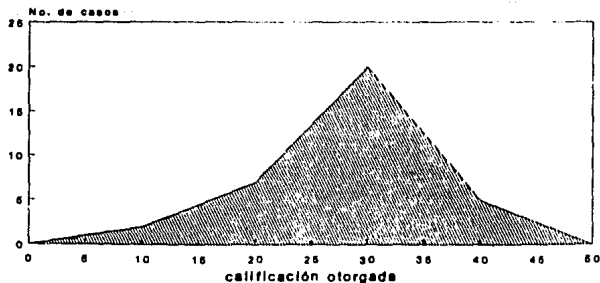
### Consistencia interna.

Al analizar la correlación entre cada pregunta y la calificación global en el índice de cambio, se demostró una muy buena congruencia; particularmente para las preguntas 1 y 2 cuando lo aplica el médico ( $r_s$  de 0.952 y 0.9 respectivamente), y para las preguntas 1,2 y 5 cuando lo aplica la T. social ( $r_s$  de 0.863, 0.8 y 0.916 respectivamente). Los interesados en mayor detalle, pueden consultar el apéndice IV-7.

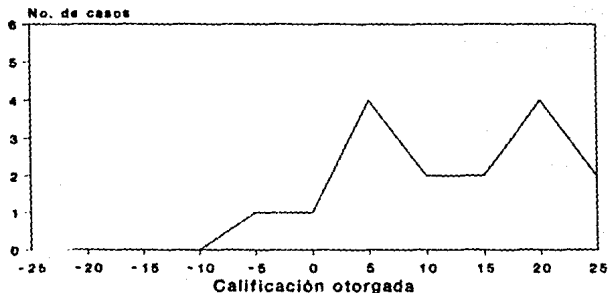
**Figura V-16a.**  
**Inst.2/Médico. Calificación basal**



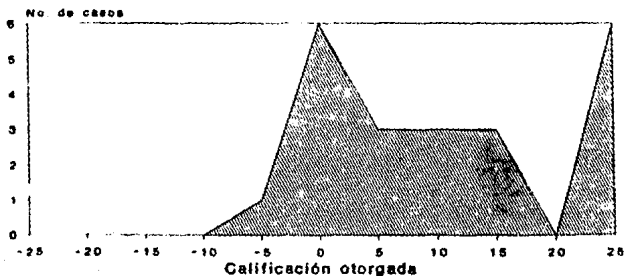
**Figura V-16b.**  
**Inst.2/T.social. Calificación basal.**



**Figura V-17a.**  
Inst.2/Médico. Calif. subsecuente.



**Figura V-17b.**  
Inst.2/T. social. Calif. subsecuente.



**Validez.**

La tabulación para validación de el instrumento No. 2 se anota en las tablas V-18 y V-19. Y se ilustra en la figura V-20.

Como se puede apreciar la medición de cambio en calidad de vida con el Instrumento No. 2 mostró una correlación estadísticamente significativa con todos los parámetros de comparación, excepto con la calificación de Karnofski cuando el instrumento es aplicado por el médico ( $r_s=0.337$ ,  $p=0.17$ ). La correlación más fuerte se demostró entre la calificación del instrumento aplicado por la trabajadora social, y la opinión del familiar y del paciente ( $r_s$  igual a 0.858 y 0.708, respectivamente).

Como se había predicho, la pregunta No. 3, que interroga sobre la ayuda requerida por el paciente para sus actividades diarias, correlaciona mejor con el Karnofski que el índice global; esto es válido cuando lo aplica el médico ( $r_s$  de 0.53 para la pregunta 3, contra un  $r_s$  global de 0.337), y cuando lo aplica la T. social ( $r_s$  de 0.502 contra 0.316). Sin embargo la predicción de que la pregunta 1 correlacionaría con mayor fuerza que la calificación global no pudo confirmarse ( $r_s=0.31$ ).

**Sensibilidad al cambio.**

La sensibilidad al cambio de el instrumento No. 2 se ilustra en las tablas V-21 a V-24, tomando como patrón de referencia cada uno de los parámetros de comparación.

Los resultados del índice de sensibilidad según Guyatt se anotan al pie de cada una de las tablas. Similarmente a lo observado con el instrumento 1, el instrumento 2 es poco sensible a los cambios en el Karnofski y la opinión del médico tratante, y bastante más sensible a la opinión del paciente y del familiar.

**Factibilidad.**

La entrevista inicial fué considerada sencilla o muy sencilla por el 97% de los entrevistados, y el 89.5% de los entrevistadores. Su aplicación fué bastante rápida: 3.56 minutos (DE=1.35) cuando lo hizo el médico, y 5.06 (DE=2.15) cuando lo hizo la T. social.

En cuanto a la entrevista de seguimiento, el tiempo utilizado por el médico fué 4.05 minutos (DE=1.68), y por la T. social 4.11 (DE=1.97) . Los entrevistados consideraron sencilla o muy sencilla la entrevista el 92.5% de las ocasiones, y los entrevistadores el 84.2% de las veces.



**Tabla V-18.** Análisis de validez de el Instrumento. No. 2 aplicado por el médico.

CALIFICACION DEL INST. 2	No. casos	Cambio Karnofski	Pac	Opinión Tratante	Fam
0 - 4	6	1.17	0.9	0.67	0.75
5 - 9	5	1.0	1.1	1.2	0.75
10 - 14	2	1.0	2.25	1.5	1.5
15 - 19	4	1.25	3.12	1.75	1.75
20 - 25	2	2.0	5.0	1.0	2.0
Desc.	4				
$r_s$		0.337	0.576½	0.703½	0.688*

½ p < 0.005

\* p < 0.001

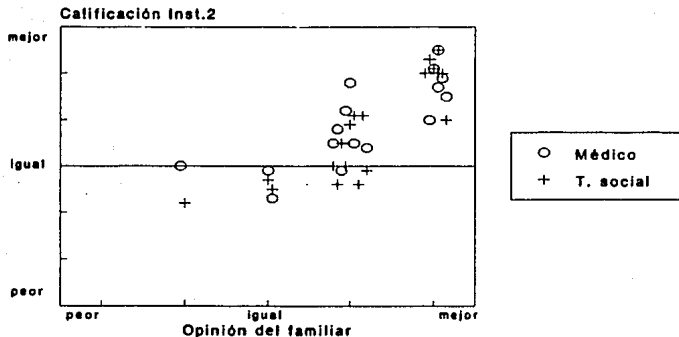
**Tabla V-19.** Análisis de validez del Instrumento. No. 2 aplicado por la T. social.

CALIFICACION DEL INST. 2	No. casos	Cambio Karnofski	Pac	Opinión Tratante	Fam
0 - 4	8	0.87	0.5	1.0	0.83
5 - 9	5	2.5	2.75	1.4	0.75
10 - 14	3	1.33	0.5	0.67	1.33
15 - 19	0	-	-	-	-
20 - 25	6	1.5	4.3	1.33	2.0
Desc.	1				
$r_s$		0.465½	0.708*	0.479½	0.858*

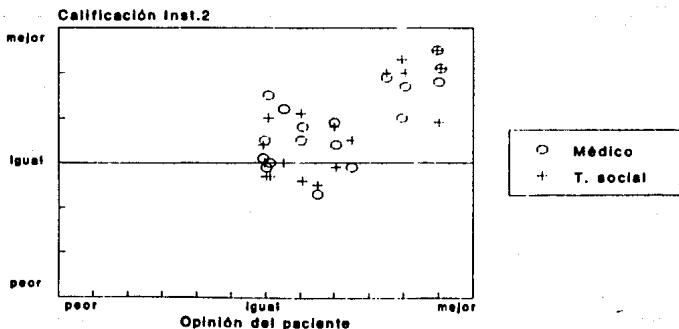
½ P < 0.05

\* P < 0.001

**Figura V-20a.**  
**Inst.2 / Opinión familiar.**



**Figura V-20b.**  
**Inst.2 / Opinión Paciente.**



**Tabla V-21.** Sensibilidad al cambio del instrumento No. 2 según el Karnofski.

Cambio en Karnofski	No. casos	Inst. 2 Médico*	Inst. 2 T. Social*
< -2	2	--	1.5
-1	3	2.0	-2.5
0	5	8.2	5.0
1	5	9.5	11.8
2	5	11.6	11.0
3	2	13.0	14.0
Desconocido	1	--	--
Indice de sensibilidad		1.22	1.38

\* Calificación promedio de el No. de casos anotado en cada renglón.

**Tabla V-22.** Sensibilidad al cambio del instrumento No. 2 según la opinión del paciente.

Opinión de Paciente.	No. casos	Inst. 2 Médico*	Inst. 2 T. Social*
< 0	0	--	--
0 - 0.9	6	5.33	1.33
1 - 2.9	6	3.0	2.5
3 - 5.0	6	18.3	19.6
Desconocido	5	--	--
Indice de Sensibilidad		2.05	3.83

\* Calificación promedio de el No. de casos anotado en cada renglón.

**Tabla V-23. Sensibilidad al cambio del instrumento No. 2 según la opinión del médico tratante.**

Opinión Tratante	No. casos	Inst. 2 Médico*	Inst. 2 T. Social*
-2	2	-7.0	-6.6
-1	2	-1.0	-3.0
0	5	6.5	12.4
1	8	8.0	2.1
2	6	15.7	16.2
Desconocido	1	--	--
Indice de Sensibilidad		1.43	0.73

\* Calificación promedio de el No. de casos anotado en cada renglón.

**Tabla V-24. Sensibilidad al cambio del instrumento No. 2 según la opinión del familiar.**

Opinión Familiar	No. casos	Lo aplica Médico	Lo aplica T. Social
-1	2	0.0	-8.0
0	2	1.0	0.0
1	9	7.3	3.1
2	7	17.8	19.9
Desconocido	3	--	--
Indice de sensibilidad		2.86	2.84

\* Calificación promedio de el No. de casos anotado en cada renglón.

## V.- RESULTADOS.

### d) Comparación de ambos instrumentos.

#### Reproducibilidad.

Es difícil comparar directamente los instrumentos, ya que utilizan escalas diferentes; sin embargo, en términos generales se pueden considerar como equivalentes los valores de concordancia calculados con Kappa ponderada ( $K_w$ ) y el coeficiente de correlación intraclase ( $R_i$ ).

Si aceptamos dicha equivalencia, entonces la reproducibilidad del Instrumento No. 2 en cuanto a la calificación global es mejor que la del Instrumento No. 1 ( $R_i=0.69$  versus  $K_w=0.45$ ). Sin embargo, la concordancia de las áreas mental y emocional consideradas por separado, es mejor para el Instrumento No. 1. Las relaciones comentadas se ilustran en la figura V-25.

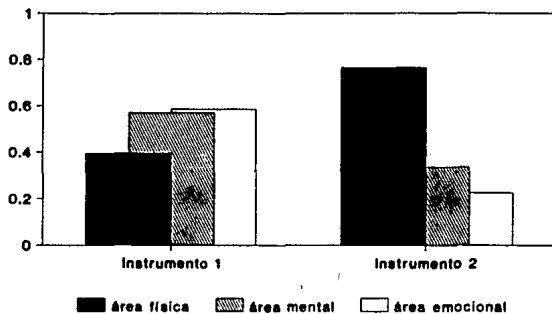
#### Consistencia interna.

Ambos instrumentos mostraron aceptable consistencia interna, aunque no es posible una comparación directa. Llama la atención el hecho de que la congruencia del instrumento 1 dependa fuertemente del entrevistador, siendo notoriamente mejor cuando lo aplica la T. social (Tabla V-7). Cuando se analizó la consistencia interna del instrumento 2, se vió en general una buena correlación entre cada una de las preguntas y la calificación global; al comparar los diferentes entrevistadores se observaron algunas variaciones en cada pregunta, destacando una mayor congruencia de la pregunta 5 (capacidad de convivencia) cuando la aplica la trabajadora social, en comparación a cuando lo aplica el médico ( $r_s$  0.916 VS 0.698).

#### Validez.

En cuanto a la validez de contenido, consideramos que el instrumento 2 cubre adecuadamente el espectro a considerar, ya que incluye las tres áreas de la salud, de acuerdo a la definición de la OMS (39); además de preguntas sobre actividades de la vida diaria y dolor. En lo que hace al instrumento 1, incluye directamente el área física y mental, e indirectamente el área social mediante el apartado emocional de dicho instrumento.

### Figura V-25. Reproducibilidad por áreas.\*



\*Valores de Kw y RI. Ver texto

Sin embargo, Jette (23) ha sugerido que un índice adecuado de estado de salud debe incluir como focos de atención tanto los síntomas, como los signos y el desempeño (funcionalidad). Bajo este criterio, el instrumento 1 es básicamente un índice de desempeño; Mientras que el instrumento dos incluye tanto parámetros de desempeño, como medidas de sensación, aunque no se encuentran deslindados unos de los otros. Ambos instrumentos excluyen observación directa de signos de enfermedad.

En la figura V-26 se ilustra la comparación de la validez de los instrumentos de acuerdo con el Karnofski; la correlación es pobre, y sólo alcanza significación estadística para el instrumento No. 2 aplicado por la trabajadora social. Cuando se consideró sólo el área física de los instrumentos (figura V-27) se observó una mejor correlación, destacando particularmente la pregunta 3 del Instrumento No. 2 (dependencia en actividades de la vida diaria).

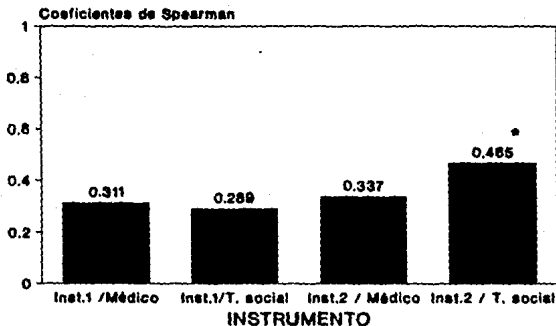
Cuando la validez se analiza en relación con la opinión del médico tratante (figura V-28) se encuentra una correlación estadísticamente significativa con ambos instrumentos, siendo mayor para el Instrumento No. 2, pero cuando lo aplica el médico y no la trabajadora social. Como se había predicho (ver material y métodos) la correlación mejora cuando se considera sólo el área física de los instrumentos, ya que los valores del coeficiente de Spearman (0.557, 0.555, 0.705, 0.479, figura V-28) cambian a 0.526, 0.624, 0.75, y 0.502 respectivamente.

Las mejores correlaciones de ambos instrumentos se obtuvieron al considerar la opinión del paciente (figura V-29) y del familiar (figura V-30), siendo homogéneamente superiores para el instrumento No. 2 aplicado por la T. social.

Curiosamente la correlación siempre es mejor con la opinión del familiar, y no con la del paciente como se había predicho. Para ahondar en este punto se analizó adicionalmente la correlación con las otras variables de contraste, encontrando que mientras las correlaciones de la opinión del familiar con el cambio en Karnofski y la opinión del médico tratante son bastante aceptables ( $r_g$  de 0.6 y 0.54 respectivamente); las correlaciones de la opinión del paciente con estos mismos parámetros no alcanzan significación estadística ( $r_g$  igual a 0.236 y 0.42, respectivamente).

# Figura V-26

## Correlación con Karnofski.

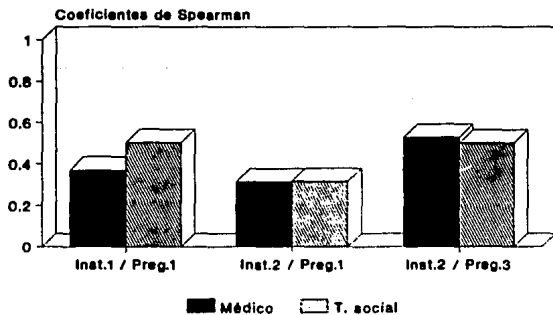


• p < 0.05



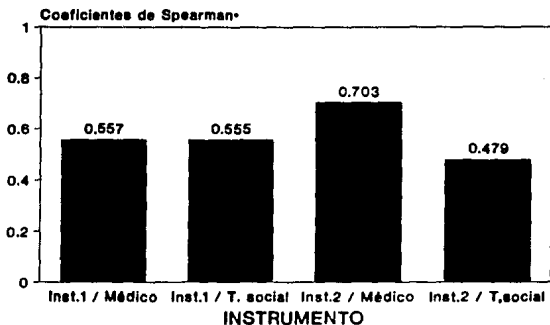
### Figura V-27

Correlación área física / Karnofski.



# Figura V-28

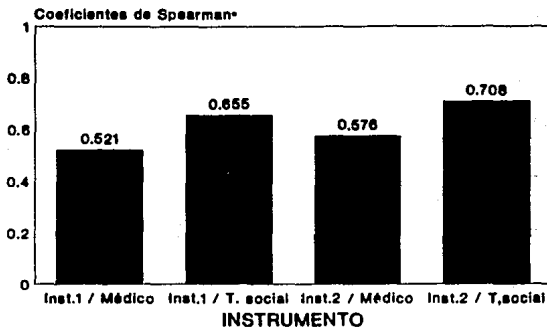
## Correlación con Tratante.



• en todos los casos  $p < 0.05$

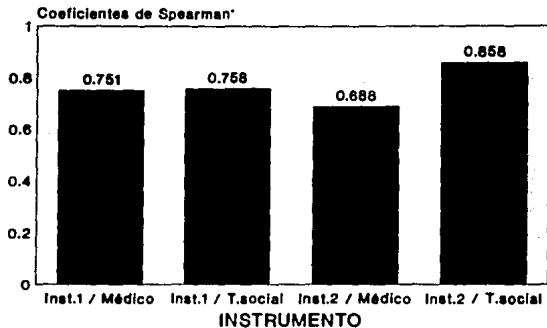
# Figura V-29

## Correlación con Paciente.



\* En todos los casos  $p < 0.05$

**Figura V-30**  
**Correlación con Familiar.**



\* En todos los casos  $p < 0.05$

No se confirmó la hipótesis de una mejor correlación de la opinión del familiar con el área emocional del Inst. 1 ( $r_s=0.49$  cuando lo aplicó el médico,  $r_s=0.31$  cuando lo aplicó la trabajadora social). Tampoco se confirmó la mejor asociación esperada con las áreas de convivencia ( $r_s$  de 0.694 y 0.765) y actividades diarias ( $r_s$  de 0.49 y 0.768) del instrumento No. 2.

#### **Sensibilidad al cambio.**

Al comparar la sensibilidad la cambio de ambos instrumentos según su índice de sensibilidad (Tabla V-31), se puede ver una baja sensibilidad a los cambios en Karnofski y en la opinión del médico tratante, lo cual era esperable, dada la modesta correlación de dichos parámetros de comparación con los instrumentos. Ahora bien, en relación con la opinión del paciente y a la opinión del familiar, la mejor sensibilidad al cambio correspondió al Instrumento No. 1 aplicado por el médico, seguido por el Instrumento No. 2 aplicado por la T. social (figura V-32).

#### **Factibilidad.**

En la figura V-33 se compara el tiempo necesario para la aplicación de los instrumentos, y en las figuras V-34 y V-35 el grado de dificultad asociado con su aplicación.

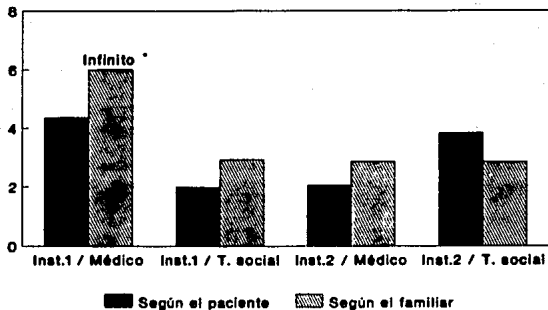
**Tabla V-31. Índices de sensibilidad al cambio de los diferentes instrumentos y entrevistadores.**

CRITERIO	INST. No. 1		INST. No. 2	
	Médico	T. social	Médico	T. social
Karnofski	0.76	1.03	1.22	1.38
Tratante	0.86	0.99	1.43	0.73
Paciente	4.36	1.98	2.05	3.83
Familiar	infinito*	2.92	2.86	2.84

\* División entre cero.

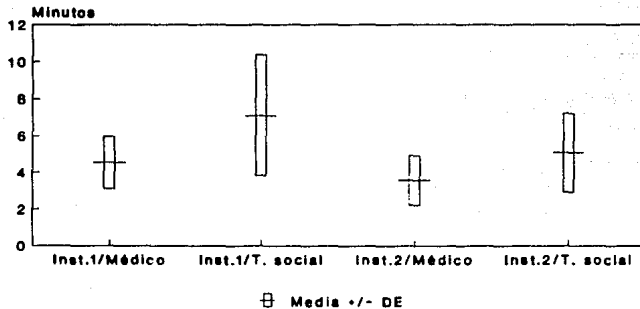
### Figura V-32

#### Indices de Sensibilidad al Cambio.



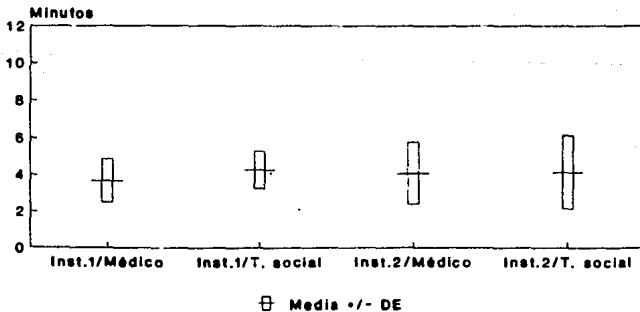
\* infinito (división por cero).

**Figura V-33a**  
Entrevista inicial.\*



\* Tiempo requerido para aplicación.

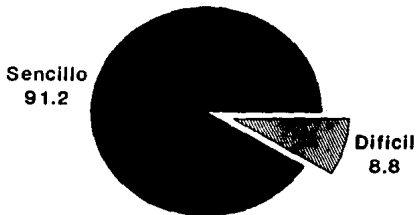
**Figura V-33b**  
Entrevista subsecuente.\*



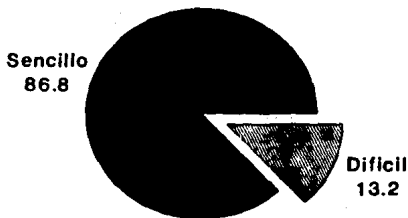
\* Tiempo requerido para aplicación.



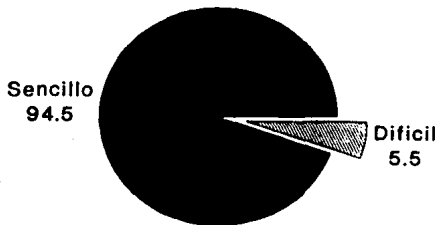
**Figura V-34a. Instrumento 1.**  
**Opinión del entrevistador (%),**



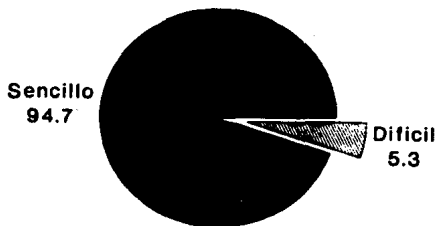
**Figura V-34b. Instrumento 2.**  
**Opinión del entrevistador (%).**



**Figura V-35a. Instrumento 1.**  
**Opinión del paciente (%).**



**Figura V-35b. Instrumento 2.**  
**Opinión del paciente (%).**



## VI.- Discusión.

En la comparación de los instrumentos estudiados resaltan más las similitudes que las diferencias, y se puede decir que en términos generales ambos instrumentos se desempeñan de manera equivalente, por lo que cualquiera de ellos podría ser utilizado en ensayos controlados para dar una estimación de los cambios en la calidad de vida de los individuos, quizá con alguna ventaja para el Instrumento No. 2 por su mejor reproducibilidad.

Analizando individualmente las diferencias encontradas: el instrumento No. 2 es superior en su reproducibilidad global al instrumento No.1; la baja reproducibilidad de las preguntas 2 y 4 de este instrumento analizadas por separado sugiere que la reformulación de dichas preguntas podría elevar aún más la reproducibilidad global. En cuanto al instrumento No. 1, no obstante que la reproducibilidad de la calificación global es moderada ( $Kw=0.45$ ), las discrepancias observadas en la actividad máxima señalada a uno y otro entrevistador arrojan algunas dudas al respecto.

La consistencia interna de ambos instrumentos parece aceptable, si bien no es posible comparar en este respecto uno y otro instrumentos.

En términos generales la correlación de los instrumentos con las variables de contraste (validez de constructo) fué equivalente, quizá con una ligera superioridad de el instrumento 2 en relación al Karnofski, por lo que se puede afirmar que ambos instrumentos tienen la misma validez. Además pudieron confirmarse las hipótesis de una mejor correlación de la opinión del médico tratante con el área física de los instrumentos; mientras que las hipótesis avanzadas en relación a una mejor correlación de la opinión del familiar con las áreas emocionales y de convivencia de los instrumentos no fueron apoyadas por los resultados.

La sensibilidad al cambio fué superior para el instrumento No. 1, pero curiosamente ello sólo se manifestó cuando lo aplicó el médico.

En cuanto a la factibilidad, no hubo diferencia importante entre los instrumentos. Ambos se aplican rápidamente (alrededor de 5 minutos por entrevista) y son bastante sencillos, tanto para el entrevistador como para el entrevistado.

Mención aparte merece la correlación observada entre el Karnofski y los instrumentos estudiados. En contra de lo esperado, dicha correlación fué muy débil, y sólo alcanzó significancia estadística para el inst. No. 2 aplicado por la trabajadora social.

Considerando que la correlación mejoró cuando se consideraron solamente las preguntas sobre capacidad física y actividades diarias, esto puede interpretarse diciendo que la escala de Karnofski mide básicamente desempeño físico en las actividades diarias, y que el dominio del concepto de calidad de vida va más allá, influyendo otros factores importantes en la opinión del paciente y del familiar. Cabe mencionar además que la concordancia inter-observador reportada para la escala de Karnofski es menor que la de los instrumentos probados en este estudio (22).

Es de interés analizar el efecto que el entrevistador tiene sobre el desempeño de los instrumentos, destacando particularmente la incongruencia de las respuestas dadas al médico durante la aplicación del Inst. 1; aunque también las calificaciones del Inst. 2 fueron más optimistas cuando lo aplicó el médico, esta tendencia no fué tan clara. En cuanto a la validez, cabe decir que se ve ligeramente incrementada cuando entrevista la trabajadora social. Ambas observaciones pueden reflejar cierto grado de compromiso del paciente ante el médico, que lo induce a ser menos verídico en sus respuestas.

La sensibilidad al cambio tiene un curioso efecto diferencial según el entrevistador y el instrumento, ya que mientras la sensibilidad del inst. 1 es mejor cuando lo aplica el médico, la del inst. 2 es mejor cuando la trabajadora social es la entrevistadora; es difícil saber si la falta de congruencia del instrumento No. 1 en la entrevista con el médico afectó el cálculo de la sensibilidad al cambio y generó parte de este efecto diferencial.

Como se recordará, el instrumento No. 1 es la versión en español de el índice propuesto por MacKenzie (30) por lo que es interesante comparar el desempeño del índice original con nuestra versión. La concordancia inter-observador en la entrevista inicial en nuestro estudio fué considerablemente menor que la reportada; tanto en el área física (45% vs 76%), como en la mental (57% vs 80%), y en la emocional (60% vs 80%). No tenemos una explicación satisfactoria para estas diferencias; podría tratarse de problemas con la traducción, diferencias en las instrucciones dadas a los entrevistadores, diferencias culturales, etc. En contraste, la consistencia interna del instrumento fué bien preservada, pero como ya se comentó, sólo cuando lo aplica la trabajadora social; así, los valores de congruencia calculados en nuestra muestra son

incluso mejores que los reportados previamente (86.7% vs 86% para el área física, 93.7% vs 85% para el área mental). Un aspecto importante que no fué específicamente abordado en las publicaciones originales (30,31) es la reproducibilidad en la medición de cambio. Es difícil comparar nuestras estimaciones de validez con la experiencia previa, pues en nuestro estudio no se utilizó el Perfil de Impacto de la Enfermedad (SIP, referencias 1 y 31) como parámetro de comparación.

Un hallazgo inesperado del estudio es la demostración de una mejor correlación de los instrumentos con la opinión del familiar que con la opinión del paciente. Una explicación plausible es que cuando el paciente es interrogado de manera directa por un representante del sistema de salud se siente comprometido a dar juicios optimistas; argumento a favor de esta hipótesis es el hecho de que ningún paciente aceptó haber empeorado, a pesar de que los demás índices señalaban la presencia de tal deterioro; asimismo, un análisis ulterior mostró que no existía la correlación esperable entre la opinión del paciente y los cambios en Karnofski y opinión del médico tratante, y aún más, la reproducibilidad de la pregunta directa al paciente fué inferior a la de las calificaciones de los instrumentos ( $R_i=0.35$ ).

Sin embargo, es claro que la investigación realizada tiene algunas limitaciones: Una de ellas es que la pregunta sobre la opinión del paciente fué incluida como parte integral del Inst. 2, por lo que la mejor correlación de este instrumento con la opinión del paciente puede reflejar tan solo esta incorporación. Además, el bajo número de casos en algunas de las subcategorías para estimación de sensibilidad al cambio hacen que los valores obtenidos deban tomarse con reserva.

## VII.-CONCLUSIONES.

Con los resultados obtenidos en el presente estudios y algunas experiencias ganadas durante su realización se pueden formular varias conclusiones interesantes.

En primer lugar, queda claro que es posible obtener mediciones de cambio en calidad de vida con adecuada reproducibilidad de manera rápida y sencilla. Asimismo parece claro que a fin de aumentar la reproducibilidad y otras propiedades deseables en los índices evaluativos de calidad de vida es conveniente estudiar su desempeño en condiciones de un ensayo clínico, lo que permitirá conocer las áreas débiles y hacer las modificaciones pertinentes.

Las condiciones de la entrevista, incluyendo la persona que la realiza y la imagen que esa persona tenga ante los ojos del paciente, pueden inducir desviaciones en las mediciones obtenidas; de donde se desprende que como parte integral de todo instrumento que pretenda medir calidad de vida debe describirse en detalle la forma y las personas que realizan la entrevista. Aunque nuestros datos no son concluyentes al respecto, parece recomendable que la persona encargada de dichas entrevistas sea lo más independiente posible de las encargadas de la atención del enfermo.

Nuestros datos sugieren que las preguntas directas al paciente tales como ¿Ha habido mejoría?, ¿Ha cambiado su calidad de vida por el tratamiento?, etc. son alteradas con mayor facilidad por las condiciones de la entrevista, y son bastante menos confiables que las preguntas indirectas.

Los índices que se refieren exclusivamente a conductas observables en el enfermo y a su grado de funcionalidad o independencia, tales como el Karnofski, reflejan sólo una parte del dominio conceptual de "calidad de vida", por lo que su uso como mediciones únicas de calidad de vida en ensayos controlados puede ser inadecuado.

Los índices de cambio en calidad de vida, al igual que todo proceso de medición tienen una variación propia del método; por lo que es deseable que todo nuevo instrumento propuesto se acompañe de estimaciones de esta variación, a fin de saber que valores deben considerarse significativos y cuales no. Una buena estimación de ella sería la variabilidad inter-observador determinada empíricamente en estudios adecuados.

**ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA**

-79-

Las estimaciones de sensibilidad al cambio son fuertemente dependientes de la eficacia de la(s) maniobra aplicada por lo que su comparación directa en diferentes estudios, o la asignación de valores mínimos aceptables, no parece adecuada. Además cabe comentar que el cálculo de sensibilidad al cambio según Guyatt, si bien es útil como índice general, no considera la posibilidad de que un instrumento sea mejor para detectar mejoría que deterioro o viceversa.

## REFERENCIAS.

- 1.- Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: Development and Final Revision of a Health Status Measure. *Medical Care* 1981; 19: 787-805.
- 2.- Bergner M, Kaplan RM, Ware JE Jr. Evaluating Health Measures. Commentary: Measuring Overall Health: an Evaluation of Three Important Approaches. *J Chron Dis* 1987; 40(suppl 1): 23S-26S.
- 3.- Bombardier C, Tugwell P. Methodological Considerations in Functional Assessment. *J Rheumatol* 1987; 14(suppl 15): 6-10.
- 4.- Bombardier C, Tugwell P. A Methodological Framework to Develop and Select Indices for Clinical Trials: Statistical and Judgmental Approaches. *J Rheumatol* 1982; 9: 753-757.
- 5.- De Haes JGJM, Van Knippenberg FCE. Quality of life Instruments for Cancer Patients: "Babel's Tower Revisited". *J Clin Epidemiol* 1989; 42: 1239-1241.
- 6.- Del Greco L, Walop W. Questionnaire Development: 1.Formulation. *CMAJ* 1987; 136: 583-585.
- 7.- Del Greco L, Walop W, McCarthy RH. Questionnaire Development: 2.Validity and Reliability. *CMAJ* 1987; 136: 699-700.
- 8.- Del Greco L, Walop W, Eastridge L. Questionnaire Development: 3.Translation. *CMAJ* 1987; 136: 817-818.
- 9.- Del Greco L, Walop W. Questionnaire Development: 5.The Pretest.
- 10.- Deyo RA. Pitfalls in Measuring the Health Status of Mexican-Americans: Comparative Validity of the English and Spanish SIP. *Am J Public Health* 1984; 6: 569-573.
- 11.- Deyo RA, Centor RM. Assessing the Responsiveness of Functional Scales to Clinical Change: An Analogy to Diagnostic Test Performance. *J Chron Dis* 1986; 39: 897-906.
- 12.- Deyo RA, Inui TS, Leininger JD, Coverman SS. Measuring Functional Outcomes in Chronic Disease: A Comparison of Traditional Scales and a Self-Administered Health Status Questionnaire in Patients with Rheumatoid Arthritis. *Medical Care* 1983; 21: 180-192.
- 13.- Editorial. Assessment of Disability. *Lancet* 1986; I: 591-2.



14.- Fayers PM, Jones DR. Measuring and Analysing Quality of Life in Cancer Clinical Trials: a Review. *Statistics in Medicine* 1983; 2: 429-446.

15.- Feinstein AR. *Clinimetrics*. Yale University Press, New Haven, CT. 1987.

16.- Feinstein AR. An Additional Basic Science for Clinical Medicine: IV. The Development of Clinimetrics. *Ann Intern Med* 1983; 99: 843-848.

17.- Feinstein AR, Josephy BR, Wells CK. Scientific and Clinical Problems in Indexes of Functional Disability. *Ann Intern Med* 1986; 105: 413-420.

18.- Guyat GH, Berman LB, Townsend M, Taylor DW. Should Study Subjects See Their Previous Responses? *J Chron Dis* 1985; 38: 1003-1007.

19.- Guyatt GH, Bombardier C, Tugwell PK. Measuring Disease-Specific Quality of Life in Clinical Trials. *CMAJ* 1986; 134: 889-895.

20.- Guyat GH, Veldhuyzen Van Zanten SJO, Feeny DH, Patrick DL. Measuring Quality of Life in Clinical Trials: A Taxonomy and Review. *CMAJ* 1989; 140: 1441-1448.

21.- Guyatt GH, Walter S, Norman G. Measuring Change Over Time: Assessing the Usefulness of Evaluative Instruments. *J Chron Dis* 1987; 40: 171-178.

22.- Hutchinson TA, Boyd NF, Feinstein AR. Scientific Problems in Clinical Scales, as Demonstrated in the Karnofsky Index of Performance Status. *J Chron Dis* 1979; 32: 661-666.

23.- Jette AM. Health Status Indicators: Their Utility in Chronic-Disease Evaluation Research. *J Chron Dis* 1980; 33: 567-579.

24.- Karnofski DA, Abelmann WH, Craver LF, Burchenal JH. The use of Nitrogen Mustards in the Palliative Treatment of Carcinoma. *Cancer* 1948; 1: 634-656.

25.- Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of Illness en the Aged. The Index of ADL: a Standardized Measure of Biological and Psychosocial Function. *JAMA* 1963; 185: 914-919.

26.- Kirshner B, Guyatt GH. A Methodological Framework for Assessing Health Indices. *J Chron Dis* 1985; 38: 27-36.

- 27.- Kramer MS, Feinstein AR. Clinical Biostatistics. LIV. The Biostatistics of Concordance. Clin Pharmacol Ther 1981; 29: 111-123.
- 28.- Liang MH. The Historical and Conceptual Framework for Functional Assessment in Rheumatic Disease. J Rheumatol 1987; 14(suppl 15): 2-5.
- 29.- MacKenzie CR, Charlson ME. Standards for the Use of Ordinal Scales in Clinical Trials. Br Med J 1986; 292: 40-43.
- 30.- MacKenzie CR, Charlson ME, DiGioia D, Kelley K. A Patient-Specific Measure of Change in Maximal Function. Arch Intern Med 1986; 146: 1325-1329.
- 31.- MacKenzie CR, Charlson ME, DiGioia D, Kelley K. Can the Sickness Impact Profile Measure Change? An Example of Scale Assessment. J Chron Dis 1986; 39: 429-438.
- 32.- Mahler DA, Weinberg DH, Wells CK, Feinstein AR. The Measurement of Dyspnea. Contents, Interobserver Agreement, and Physiological Correlates of Two New Clinical indexes. Chest 1984; 85: 751.
- 33.- Read JL, Quinn RJ, Hoefler MA. Measuring Overall Health: an Evaluation of Three Important Approaches. J Chron Dis 1987; 40(suppl 1): 7S-21S.
- 34.- Sokolow J, Silson JE, Taylor EJ, Anderson MA, Rusk HA. Functional Approach to Disability Evaluation. JAMA 1958; 167: 1575-1584.
- 35.- Spitzer WO. State of Science 1986: Quality of Life and Functional Status as Target Variables for Research. J Chron Dis 1987; 40: 465-471.
- 36.- Sriwatanakul K, Kelvie W, Lasagna L, Calimlim JF, Weis DF, Mehta G. Studies with Different Types of Visual Analog Scales for Measurement of Pain. Clin Pharmacol Ther 1983; 34: 234-239.
- 37.- Torrance GW. Utility Approach to Measuring Health-Related Quality of Life. J Chron Dis 1987; 40: 593-603.
- 38.- Walop W, Del Greco L, Eastridge L. Questionnaire Development: 4. Preparation for Analysis. CMAJ 1987; 136: 927-928.
- 39.- World Health Organization: The Constitution of the World Health Organization. WHO Chron 1947; 1: 29.

**APENDICE I**

**VERSION ORIGINAL DEL INSTRUMENTO 1.**

## The Functional Status Index.

### Baseline Assessment.

We will ask you a few questions about how you have been doing before your admission (or visit) to the hospital. The questions have to do with your ability to do physically demanding things and your ability to concentrate. We also would like to know how you have been coping with stress recently and will ask about your job.

**Physical activity.**- My first set of questions is about things that you do that are physically strenuous. These may be things that you do around the house, activities related to work or to just getting around. Specifically, I am going to ask you what physical activity that you do requires the most physical effort.

1. Of the things that you usually do, what physical activity requires the most physical exertion or is the most physically strenuous?

2. On your good days, what would be the most [physical activity from question 1] that you could do? On your bad days, what would be the most [physical activity from question 1] that you could do?

**Concentration.**- My next questions have to do with things that you do that require mental effort or concentration. Again, this may be activities related to work, to running the home, or to any other interest. I am going to ask you what activity that you do requires the most concentration or mental effort.

1. Of the things that you usually do, what mental activity requires the most concentration or mental effort?

2. On your good days, what would be the most [mental activity from question 1] that you could do? On your bad days, what would be the most [mental activity from question 1] that you could do?

**Emotional.-** My next questions have to do with stressful or emotionally difficult things that may have happened recently to you. I am going to ask you whether anything that has been going on in your life has been particularly stressful. This could include your health, family or financial worries, your job, or any other stresses you have experienced.

1. Is there anything going on in your life that has been particularly stressful or emotionally difficult?

2. [If yes] What kind of problem is it?

3. How are you coping with this situation?

**Transition Assessment.**

**Physical.-** The last time we talked, you said that [physical activity from baseline or previous transition] was the most physically strenuous thing you could do.

1. In terms of physical activity, are you much better, slightly better, the same, slightly worse, or much worse?

2. Can you still [physical activity from the baseline or previous transition]?

3. [If yes] Is this still the most physically strenuous activity that you can do?

4. [If no] What is the most physically strenuous activity that you can do now?

**Concentration.-** The last time we talked, you said that [mental activity from the baseline or previous transition] was the most mentally taxing thing you could do.

1. In terms of mental activity and your ability to concentrate, are you much better, slightly better, the same, slightly worse, or much worse?

2. Can you still [mental activity from the baseline or previous transition]?

3. [If yes] Is this still the most mentally taxing activity that you can do?

4. [If no] What is the most mentally taxing activity that you can do now?

**Emotional.**

1. I am interested in how you are now coping with the stresses in your life. Would you describe your ability to cope with stress since we last talked as much better, slightly better, the same, slightly worse, or much worse?

2. Have you experienced anything since we last talked that you consider stressful?

3. [If yes] How are you coping with it?

**APENDICE II**

**FORMAS UTILIZADAS EN LA INVESTIGACION.**















**APENDICE III.    FORMULARIO.**

## A) Fórmula para cálculo de Kappa ponderada.

$$Kw = 1 - \frac{Qo'}{Qe'} \quad Qo' = \frac{Ewfo}{N} \quad Qc' = \frac{Ewfe}{N}$$

Donde:

- Kw = Kappa ponderada.  
 Qo' = Proporción ponderada de desacuerdos observados.  
 Qe' = Proporción ponderada de desacuerdos esperados.  
 E = Sumatoria.  
 w = Peso del desacuerdo.  
 fo = Frecuencia observada del desacuerdo.  
 fe = Frecuencia esperada del desacuerdo.  
 N = Número de individuos.

## B) Cálculo de el coeficiente de correlación intraclass.

$$Ri = \frac{msS - mse}{msS + mse + 2(msM)} \quad F = \frac{msS}{mse}$$

Donde:

- Ri = Coeficiente de Correlación Intraclase.  
 msS = Cuadrado medio entre sujetos.  
 msM = Cuadrado medio entre observadores.  
 mse = Cuadrado medio del error.

C) Cálculo del coeficiente de correlación de rangos de Spearman.

$$r_s = 1 - \frac{6 (Ed^2)}{N^3 - N}$$

Donde:

$r_s$  = Coeficiente de correlación de rangos.

$d^2$  = Cuadrado de la diferencia de rangos (rango asignado por criterio 1 - rango asignado por criterio 2).

$N$  = Número de individuos estudiados.

$E$  = Sumatoria.

D) Cálculo de sensibilidad al cambio.

$$Sc = \frac{E |c| / Nc}{E|nc| / Nnc}$$

Donde:

$Sc$  = Sensibilidad al cambio.

$E |c|$  = Sumatoria del valor absoluto de cambio en los sujetos que sí cambiaron.

$E|nc|$  = Sumatoria del valor absoluto de cambio en los sujetos que en realidad no cambiaron.

$Nc$  = Número de sujetos que sí cambiaron.

$Nnc$  = Número de sujetos que no cambiaron.



E) Cálculo de  $f_i$ .

$$f_i = \frac{ad - bc}{c_1 c_2 r_1 r_2}$$

Donde las letras corresponden a los valores de una tabla de contingencia:

		CRITERIO I		
		+	-	
CRITERIO II	+	a	b	$r_1$
	-	c	d	$r_2$
		$c_1$	$c_2$	

**APENDICE IV. TABLAS DE DATOS.**

Apéndice IV-1. Tablas de contingencia para análisis de reproducibilidad del Instrumento 1.

Area Física.

		T. Social				
		-2	-1	0	+1	+2
	-2	0	0	0	0	0
	-1	0	1	0	1	0
Médico	0	0	1	1	2	0
	+1	0	0	1	2	2
	+2	0	0	0	3	3

Area Mental.

		T. Social				
		-2	-1	0	+1	+2
	-2	1	0	0	0	0
	-1	0	0	1	0	1
Médico	0	0	1	5	1	0
	+1	0	0	1	3	0
	+2	0	0	0	2	2

Area Emocional.

		T. Social				
		-2	-1	0	+1	+2
	-2	0	0	0	0	0
	-1	0	0	0	0	0
Médico	0	0	0	5	1	0
	+1	0	0	1	3	3
	+2	0	0	0	2	4

Global.

## Trabajadora Social

Médico

	-1	0	+1	+2	+3	+4	+5	+6
:								
-1	0	0	0	1	0	0	0	0
0	2	0	2	0	0	0	0	0
+1	0	0	0	1	0	0	0	0
+2	0	1	0	1	0	0	2	0
+3	0	0	0	0	1	0	0	0
+4	0	0	0	0	0	0	0	0
+5	0	0	0	0	1	0	1	0
+6	0	0	0	0	1	0	1	1

**Apéndice IV-2. Cálculo de sensibilidad al cambio del Instrumento No. 1.**

**A) Cuando lo aplica el Médico.**

CRITERIO	Cambio	n	Calificación Promedio	Cociente
Karnofski	Si	17	2.44	0.76
	No	5	3.2	
Opinión del Paciente	Si	12	2.92	4.36
	No	6	0.67	
Opinión del Tratante	Si	18	2.41	0.86
	No	5	2.8	
Opinión del Familiar	Si	18	3.18	Infinito
	No	2	0.0	

**B) Cuando lo aplica la Trabajadora Social.**

CRITERIO	Cambio	n	Calificación Promedio	Cociente
Karnofski	Si	17	2.83	1.03
	No	5	2.75	
Opinión del Paciente	Si	12	3.17	1.98
	No	6	1.6	
Opinión del Tratante	Si	18	2.58	0.99
	No	5	2.6	
Opinión del Familiar	Si	18	2.92	2.92
	No	2	1.0	

**Apéndice IV-3. Distribución de Frecuencias. Calificaciones basales y subsecuentes del Instrumento No. 2.**

<b>Calificación basal</b>	<b>Según el Médico (No. de casos)</b>	<b>Según la T. social (No. de casos)</b>
41 - 50	0	0
31 - 40	8	5
21 - 30	19	20
11 - 20	5	7
0 - 10	1	2
Desconocido	6	5

<b>Calificación subsecuente</b>	<b>Según el Médico (No. de casos)</b>	<b>Según la T. social (No. de casos)</b>
20 a 25	2	6
15 a 19	4	0
10 a 14	2	3
5 a 9	4	3
0 a 4	4	3
-5 a -1	2	6
-10 a -6	1	1
Desconocido	4	1

**Apéndice IV-4. Concordancia interobservador del Instrumento No. 2.**

**Entrevista Inicial.**

PREGUNTA	$R_{ij}$	$r^+$	$t^*$
1	0.3 F=0.47	0.61 p=0.0003	-1.6 p=0.119
2	0.14 F=1.45	0.18 p=0.34	0.613 p=0.54
3	0.31 F=8.16	0.79 P=0.0001	2.67 p=0.0125
4	0.21 F=2.65	0.27 P=0.155	-0.6 P=0.55
5	0.455 F=2.65	0.45 p=0.0147	0.0 p=1.0
GLOBAL	0.7 F=7.09	0.75 p<0.0001	0.53 p=0.6

**Entrevista de Seguimiento.**

PREGUNTA	$Ri\frac{1}{2}$	$r^+$	$t^*$
1	0.763 F=7.9	0.77 p=0.0001	-2.7 p=0.79
2	0.337 F=2.3	0.40 p=0.1	0.556 p=0.58
3	0.741 F=6.85	0.75 P=0.0002	-0.14 p=0.89
4	0.225 F=3.16	0.54 P=0.02	1.64 P=0.12
5	0.67 F=5.135	0.675 p=0.0015	-0.14 p=0.89
GLOBAL	0.69 F=7.93	0.78 p<0.0001	0.73 p=0.47

½ Coeficiente de correlación intraclase y estadístico F asociado.

+ Coeficiente de correlación de Pearson; y probabilidad de que  $r=0$ .

\* Valor del estadístico "t" y probabilidad de que la media de ambos observadores sea la misma.



**Apéndice IV-5. Consistencia Interna de el Instrumento No. 2.  
Coeficientes de Correlación de Spearman entre cada pregunta  
y la calificación global.**

PREGUNTA	Médico		T. Social	
	$r_s$	p	$r_s$	p
1	0.952	<0.0001	0.863	<0.0001
2	0.9	<0.0001	0.8	<0.0001
3	0.693	0.0006	0.752	<0.0001
4	0.681	0.0009	0.791	<0.0001
5	0.698	0.0005	0.916	<0.0001

**Apéndice IV-6. Correlación de la calificación global del instrumento No. 2 con las variables de contraste.**

**Cuando lo Aplica el Médico.**

<b>CRITERIO</b>	<b>Spearman</b>	<b>Valor de p</b>
Karnofski	0.337	0.1737
Opinión del Paciente	0.576	0.0087
Opinión del Tratante	0.703	0.0005
Opinión del Familiar	0.688	0.0023

**Cuando lo Aplica la Trabajadora Social.**

<b>CRITERIO</b>	<b>Spearman</b>	<b>Valor de p</b>
Karnofski	0.465	0.0325
Opinión del Paciente	0.708	0.0001
Opinión del Tratante	0.479	0.023
Opinión del Familiar	0.858	<0.0001

**Apéndice IV-7. Cálculo de Sensibilidad al Cambio del Instrumento. No. 2.**

**A) Cuando lo aplica el Médico.**

CRITERIO	Cambio	n	Calificación Promedio	Cociente
Karnofski	Si	17	10.0	1.22
	No	5	8.2	
Opinión del Paciente	Si	12	11.64	2.05
	No	6	5.67	
Opinión del Tratante	Si	18	10.0	1.43
	No	5	7.0	
Opinión del Familiar	Si	18	11.43	2.86
	No	2	4.0	

**B) Cuando lo aplica la Trabajadora Social.**

CRITERIO	Cambio	n	Calificación Promedio	Cociente
Karnofski	Si	17	10.75	1.38
	No	5	7.8	
Opinión del Paciente	Si	12	12.75	3.83
	No	6	3.33	
Opinión del Tratante	Si	18	9.06	0.73
	No	5	12.4	
Opinión del Familiar	Si	18	11.35	2.84
	No	2	4.0	