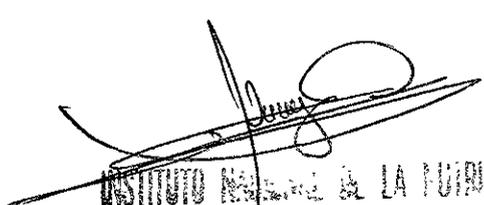


1221
34

TESIS DE ESPECIALIDAD MEDICA.

MEDICINA INTERNA

UTILIZACIÓN DE TRES INSTRUMENTOS DE CALIDAD DE VID EN PACIENTES HOSPITALIZADOS EN EL INSTITUTO NACIONAL DE CIENCIAS MÉDICAS Y NUTRICIÓN "SALVADOR ZUBIRÁN"


INSTITUTO NACIONAL DE LA NUTRICION
SALVADOR ZUBIRAN
SUS DIRECCION DE ENSEÑANZA
México, D. F.
Alumno: RUY LÓPEZ RIDAURA


* SET 4 2000 *
Programa de Maestría y Doctorado en Ciencias Médicas, Odontológicas y de la Salud

Asesor. (a): DRA. FLORENCIA VARGAS VORACKOVA

Profesor del curso: DR. LUIS F. USCANGA DOMÍNGUEZ

2000

282634



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

ÍNDICE

Resumen	i
Introducción	1
Justificación	7
Objetivos.....	8
Materiales y métodos	9
Resultados	24
Discusión	50
Conclusiones	56
Bibliografía	57

Utilización de 3 instrumentos de calidad de vida en pacientes hospitalizados en el Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INCMNSZ).

Antecedentes. La medición de calidad de vida de pacientes hospitalizados ha sido poco explorada en nuestro medio. La medición de calidad de vida en pacientes con enfermedades crónico-degenerativas es indispensable para evaluar la atención médica que se les puede ofrecer.

Objetivo. Utilizar y validar tres instrumentos de calidad de vida en un grupo de pacientes con alguna de cuatro enfermedades crónicas (SIDA, cirrosis hepática, lupus eritematoso generalizado, insuficiencia renal crónica terminal) atendidos en el INCMNSZ.

Pacientes. Se estudiaron 75 pacientes con algunas de las siguientes enfermedades: SIDA, cirrosis hepática, lupus eritematoso generalizado, insuficiencia renal crónica terminal, que fueron hospitalizados por cualquier causa en el INCMNSZ de agosto 1999 a mayo 2000.

Observaciones. Se midió la calidad de vida al ingreso con dos instrumentos de utilidades en salud (Time trade-off y Termómetro) y un instrumento de perfil de salud (SF-36).

Resultados. El termómetro, Time trade-off y las dos subescalas del SF36 (salud física y salud mental) presentaron buenos índices de correlación intraclase en la prueba repetida de reproducibilidad: 0.77, 0.98, 0.92 y 0.89 respectivamente. A diferencia del TTO, el termómetro y la escala de salud mental del SF-36, discriminaron adecuadamente a los pacientes según la causa de la hospitalización (urgente, electiva y diagnóstica). El termómetro correlacionó de forma significativa con la mayoría de los parámetros clínicos específicos por enfermedad. El SF-36 se asoció de la forma esperada con las características globales y no con las específicas por enfermedad. El TTO no se asoció con ningún supuesto teórico esperado. El termómetro correlacionó mejor con la escala de salud física que con la escala de salud mental del SF-36. No hubo diferencias significativas entre las cuatro enfermedades.

Conclusiones. El SF-36 y el Termómetro son válidos y reproducibles en nuestro medio y ofrecen una visión complementaria de la calidad de vida. El método del Time trade-off resultó menos aplicable en nuestra población. La evaluación de las preferencias del paciente por medio del Termómetro complementa la descripción del perfil de salud que hace el SF-36.

Utilización de 3 instrumentos de calidad de vida en pacientes hospitalizados en el Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán

Introducción.

El avance tecnológico de la medicina ha llegado al punto de poder modificar el curso natural de muchas enfermedades. Desafortunadamente este avance tecnológico ha ocasionado que en las últimas décadas, los costos en salud se hayan incrementado de forma constante y alarmante. Lo anterior ha provocado que las decisiones de incorporar acciones determinadas de salud, contemplen criterios adicionales a los de eficacia, efectividad y seguridad (1,2). Estos criterios, principalmente socioeconómicos, han dado lugar a lo que se conoce como eficiencia u optimización de los recursos, donde se busca derivar el beneficio total máximo de los recursos disponibles (3). Por otro lado, la limitación de recursos tiene un carácter regional, por lo que definir parámetros internacionales de la distribución de estos en los diferentes tipos de atención es inconcebible. Este hecho ha provocado una preocupación cada vez mayor sobre la calidad de la atención, cómo evaluarla y cómo mantenerla al menor costo posible, para así tener programas nacionales eficientes.

Desde el punto de vista asistencial, existen tres razones principales por las que se ofrece cualquier tipo de atención médica al paciente. Estas intervenciones terapéuticas pueden: 1) incrementar longevidad, 2) prevenir morbilidad en el futuro, y por último 3) hacer que el paciente se "sienta" mejor. Los primeros dos puntos son relativamente fáciles de medir y por lo tanto de cuantificar sus logros y, en caso de requerirlo, relacionarlo con los costos asociados a la intervención específica; siendo este un tipo de análisis ampliamente definido y utilizado. El último punto, el que traduce calidad de vida, tiene un importante grado de subjetividad, por lo que su medición se hace más compleja pero necesaria para evaluar completamente el beneficio al paciente y definir el empleo óptimo de los recursos para la atención médica, en función de la modificación o mantenimiento de una adecuada calidad de vida (4).

Por otro lado, en las últimas décadas, se ha visto un cambio importante en el perfil de morbilidad de los países menos industrializados. Este cambio radica en un aumento en la incidencia y prevalencia de enfermedades crónico-degenerativas, que han rebasado en algunos casos al de las enfermedades agudas. Estas enfermedades, por lo general no curables, deterioran principalmente la calidad de vida y con esto la productividad, lo que representa un importante costo en salud pública. El desenlace de las acciones medicas hacia este tipo de pacientes, no puede ser evaluado exclusivamente con medidas "duras" de mortalidad, morbilidad o sobrevida,

sino que requiere de una evaluación integral que incluya varios aspectos de la vida diaria del paciente y cómo éstos son modificados por su estado de salud; evaluando así no sólo la cantidad de vida sino también la calidad. Por esto, el cambio en la calidad de vida se reconoce actualmente como una de las principales demandas de atención médica; principalmente en pacientes que cursan con enfermedades crónico-degenerativas (5). Desde 1984, se definió la calidad de la atención médica, como aquella atención que consistentemente contribuyera a la mejoría y mantenimiento de la calidad y duración de la vida. De esta forma se contempla a la calidad de vida, como parte de los resultados fundamentales de una adecuada calidad de atención (6, 7).

Todo lo anterior, ha generado una búsqueda exhaustiva de tanto un concepto concreto de calidad de vida, como de métodos para medirla. La calidad de vida es un concepto que apareció en la década de los 70's e inicialmente se utilizó para describir la función física y emocional de un individuo (8). Desde entonces hasta la actualidad, ha existido una gran controversia entre los expertos y aún no se tiene una definición operacional uniforme. Por lo anterior se ha intentado definir una serie de dominios o características en la vida diaria de los individuos que funcionen como indicadores de calidad de vida. Actualmente los indicadores o dominios de calidad de vida se han definido desde algo tan simple como la posibilidad de regresar al trabajo, hasta una serie de cuestionarios complejos sobre actividades sociales y problemas psicológicos que modifican la calidad de vida. De esta forma diferentes autores han propuesto grupos de dominios necesarios de incluir y aunque aún en esto no existe un acuerdo concreto, por lo general se incluyen aspectos de la función física de los individuos (habilidades físicas, limitaciones, movilidad, etc.), estados mentales o emocionales (ansiedad, depresión, función cognitiva, etc.), función social (relaciones interpersonales, aspectos familiares, etc.), función laboral, y aspectos de la enfermedad como son síntomas y percepción general de salud (9).

Independientemente de lo complejo que resulta definir calidad de vida y del poco acuerdo entre los expertos en el tema, se debe entender que la calidad de vida representa la percepción del paciente por preferir ciertas características de su vida, que le permitan lograr objetivos y metas que le sean útiles para su función física, social y emocional. (9, 10). De esta forma la calidad de vida representa un conjunto de dominios, dados por aspectos de bienestar físico, mental y emocional más aspectos no relacionados con la salud como el trabajo, la familia, los amigos y otras circunstancias. Para facilitar su comprensión, se ha acuñado el término de "calidad de vida relacionada a la salud", como aquellos aspectos de la calidad de vida que son modificables por el estado de salud, pudiendo así definir una serie de dominios relacionado con los síntomas y estado de salud del paciente. De esta forma la calidad de vida relacionada a la salud está diferenciada de

todos aquellos aspectos que están más relacionados con aspectos culturales, socioeconómicos y morales de los individuos y que no están asociados con el estado de salud.

La medición de la calidad de vida relacionada a la salud, se realiza de varias formas. Según Guyatt (11), en su artículo de taxonomía de los instrumentos de estados de salud, plantea que los instrumentos pueden ser discriminativos o evaluadores. Los primeros son útiles para diferenciar pacientes con mejor calidad de vida contra aquellos con peor; en este sentido se requiere que sujetos estables demuestren poca variabilidad en evaluaciones repetidas. Es decir, que la variabilidad entre sujetos sea mayor que la variabilidad intra-sujeto, que se traduce en reproducibilidad. Los segundos evalúan principalmente los cambios en calidad de vida, que un mismo paciente presenta en determinado tiempo o después de alguna acción terapéutica específica. En estos instrumentos es importante que, además de reproducibilidad en los estados basales, la variación en los pacientes que cambiaron su estado de salud sea mayor que en aquellos que no modificaron su estado de salud, lo que se traduce como sensibilidad al cambio.

Los instrumentos que miden calidad de vida también pueden ser genéricos o específicos dependiendo de su aplicabilidad a una sola enfermedad o a varias. Por lo general los instrumentos específicos son buenos instrumentos evaluadores, y por lo tanto tienen buena sensibilidad al cambio. Se han utilizado en varios ensayos clínicos, para evaluar la calidad de vida como desenlace, demostrando así sus virtudes. Este tipo de instrumentos tiene la desventaja de que, por lo general, sólo contemplan un aspecto específico de la calidad de vida, sin reflejar en su totalidad la calidad de vida y, por definición, no son comparables con otros instrumentos específicos de otras enfermedades. Ejemplos de este tipo de instrumentos son innumerables y están contruidos principalmente con él o los dominios de calidad de vida que teóricamente están mas afectados por la enfermedad que evalúan, resultando en ocasiones, varios instrumentos para cada enfermedad, según el dominio de calidad de vida que se quiera evaluar.

Los instrumentos genéricos por lo general son buenos instrumentos discriminativos y contemplan varios aspectos de salud o dominios, sin embargo, la mayoría tiene poca sensibilidad al cambio en condiciones específicas, es decir son poco evaluadores.

Los instrumentos genéricos a su vez se pueden dividir, según Guyatt (11), en perfiles de salud o índices y utilidades en salud. Los primeros, están compuestos por cuestionarios que contienen varias preguntas y cada una de ellas esta dirigida a calificar un dominio específico. De esta forma evalúan de forma diferencial cada dominio y se obtiene una calificación por cada uno de ellos sin poder lograr una calificación global. Este tipo de instrumento se ha utilizado en varios estudios, demostrando su validez y reproducibilidad en varias situaciones y con diferentes

enfermedades, comprobando así su poder de discriminación. Debido a la calificación independiente por dominio, tienen la ventaja de detectar el impacto de la enfermedad sobre cada uno de los aspectos de la calidad de vida, sin embargo, es difícil cuantificar la importancia que tiene cada uno de estos dominios en la calidad de vida general, desde el punto de vista del paciente. Ejemplos de este tipo de instrumento son el Perfil del Impacto de la Enfermedad (Sickness Impact Profile) y el SF36, que han sido ampliamente utilizados en la práctica clínica. (12).

El otro tipo de instrumentos son aquellos que miden utilidades en salud. Estos instrumentos son de los más utilizados derivándose de la teoría económica y del análisis de toma de decisiones (13). Según esta teoría las utilidades son una medida de preferencia que en el caso de la salud, implica preferencia por diferentes estados de salud. Por el tipo de medición, este tipo de instrumentos evalúa de forma global una serie de dominios que, por lo general, se le presentan al paciente de forma implícita y éste, en su respuesta, contempla el valor que le da a cada dominio. De esta forma refleja la preferencia del paciente por los diferentes aspectos de su estado de salud y cómo modifican éstos su calidad de vida. El paciente integra todos los efectos positivos y negativos tanto de su enfermedad como del tratamiento y logra darle un solo valor entre muerte = 0 (es decir, el peor estado de salud imaginable), y el mejor estado de salud probable = 1. Lo anterior le proporciona una ventaja sobre los instrumentos específicos y los de perfiles de salud, en los cuales, como ya se comentó, se miden por separado los dominios, y no se conoce cuál es el valor diferencial que el paciente le da a cada dominio. Se ha aplicado a varias enfermedades, teniendo buenos índices de reproducibilidad y validez tanto de contenido como de constructo. Tiene la ventaja que en una sola puntuación el paciente pondera los efectos positivos y negativos de la intervención en salud.

Por ser una medida única y relacionada con la preferencia del paciente, la medición de utilidades, permite evaluar diferentes intervenciones en la misma enfermedad, así como comparar intervenciones entre enfermedades, permitiendo evaluar la eficiencia del programa de salud, con la construcción de árboles de decisiones. Por último, otra ventaja que tiene el medir utilidades, es que cuando se relaciona a sobrevida esperada, se puede calcular *QALY's* (años de sobrevida ajustados a calidad), que es una de las principales unidades de calidad vida utilizadas en los análisis de costo-utilidad, con lo que se han definido algunas políticas de atención médica (9, 14, 15).

Para medir utilidades existen dos formas generales. La primera consiste en clasificar al paciente en categorías, con base en sus respuestas sobre su estado de salud, logrando un índice

único, ponderado por el paciente. Un ejemplo de este es el EuroQOL (European Quality of Life Instrument). Este instrumento es relativamente nuevo y aún está en la etapa de validación y medición de su reproducibilidad. La otra forma consiste en preguntar directamente al paciente qué valor le da a su estado de salud y a la calidad de vida que se desprende de éste. Lo anterior se ha realizado con 4 diferentes métodos: Ruleta de probabilidades (Standard gambling), trueque (Time Trade-Off), disposición a pagar, y termómetro. Los diferentes métodos, han sido evaluados en diferentes contextos, demostrando buenos índices de reproducibilidad, además de validez. El problema principal que se ha tenido con la medición de utilidades es que la calificación final depende totalmente del método escogido, por lo que comparar de forma absoluta los índices obtenidos con diferentes instrumentos resulta imposible (16).

Los estudios que evalúan calidad de vida cada vez son más frecuentes. En general se han realizado estudios que determinan cuál es la calidad de vida relacionada a diferentes enfermedades, logrando tener un valor cuantificable de cuál es la repercusión que tienen estas enfermedades (principalmente crónico-degenerativas) en la calidad de vida de los pacientes. Estos estudios han tenido una variación importante, y aunque se tiene una idea general de cual es la calificación en calidad de vida de algunas enfermedades, es difícil aún tener una adecuada comparación entre diferentes poblaciones de estudios independientes.

Por otro lado existen muchos estudios que utilizan a la calidad de vida, como desenlace principal de diferentes maniobras específicas de atención médica. Con estos estudios se han podido evaluar, por medio de análisis de costo-utilidad, maniobras terapéuticas en pacientes con enfermedades crónicas, con poca posibilidad de curación. Los más comunes son aquellos que evalúan maniobras paliativas en pacientes con cáncer terminal, o la conveniencia de diferentes tratamientos agresivos en enfermedad con poca posibilidad de curación completa, como pueden ser enfermedades reumáticas y SIDA. Por último también se han realizado unos pocos estudios para evaluar la atención médica más general, como puede ser la hospitalización, utilizando como desenlace la calidad de vida.

El Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán (INCMNSZ), es un hospital público, de tercer nivel, de concentración y universitario, que atiende una población numerosa de enfermos. En 1996 se hospitalizaron 8764 pacientes y los diagnósticos principales fueron los siguientes (17):

TABLA 1. Causas principales de hospitalización en el INCMNSZ en 1996

Diagnóstico	%
Tumores malignos excepto linfomas y leucemias	13.4
Enfermedades del aparato circulatorio	7.2
Cirrosis y otras enfermedades crónicas del hígado.	4.5
Enfermedades del sistema osteomuscular y tejido conectivo	4.3
Colelitiasis y colecistitis	3.7
Enfermedades del aparato urinario	3.1
Diabetes Mellitus	2.8
Neumonía	2.1
Hernia de la cavidad abdominal	1.9
Leucemia y linfomas	1.8

Fuente: Departamento de archivo clínico. INCMNSZ 1996.

Como puede apreciarse en la TABLA 1, el INCMNSZ hospitaliza predominantemente pacientes con enfermedades crónico-degenerativas, que por su evolución natural, causan un deterioro importante en la calidad de vida. El estudio de la calidad de vida en este tipo de pacientes ha sido poco explorado en el instituto y en general en todo el país. Los estudios que existen se han enfocado principalmente a evaluar calidad de vida con utilización de instrumentos específicos para ciertas enfermedades y probando acciones médicas específicas. Sin embargo cada vez existe una mayor necesidad de encontrar una forma útil de medir calidad de vida con instrumentos genéricos y así poder comparar acciones más generales de salud, como puede ser el impacto de la hospitalización en este tipo de enfermos, tema de por sí controvertido. La utilización de instrumentos genéricos permite no solo comparar entre enfermedades que se estudian en un mismo estudio, sino también compararlo contra otros grupos de pacientes de otros estudios y en condiciones distintas.

Por otro lado, en la actualidad existe una tendencia a intentar validar en la población que atendemos, instrumentos utilizados en otros escenarios, alejándose de la tendencia previa que era de generar instrumentos para cada situación en particular, lo que limitaba su extrapolación.

Por lo anterior en el presente estudio se utilizaron tres diferentes instrumentos de calidad de vida ampliamente utilizados en otras circunstancias, en cuatro enfermedades que se hospitalizan frecuentemente en nuestro instituto, valorando principalmente su validez en nuestra población.

Justificación

La calidad de vida en pacientes con enfermedades crónico-degenerativas en el ambiente hospitalario ha sido poco explorada en nuestro instituto. Por lo general cuando se ha intentado evaluar calidad de vida se ha utilizado instrumentos o abordajes muy particulares que no permiten su comparación entre enfermedades o con otras poblaciones. Por otro lado no existen estudios que evalúen el impacto en calidad de vida durante la hospitalización de enfermos con padecimientos crónico-degenerativos en el INCMNSZ. La evaluación de la atención médica por medio de modificaciones en la calidad de vida, es necesaria principalmente en este grupo de pacientes en quienes la calidad de vida se encuentra deteriorada, y el impacto que tiene la atención médica en otros desenlaces como morbi-mortalidad es muy poco o incluso nulo. Es por lo anterior que se ha dudado de la pertinencia de hospitalizar pacientes quienes por su enfermedad de base, las posibilidades terapéuticas tienen poco impacto en mortalidad. Un análisis del impacto en calidad de vida, proporcionará una visión mas completa del impacto de la hospitalización en este tipo de pacientes.

En un país como México, donde los recursos para la atención médica son limitados, es difícil establecer las políticas de intervención médica, de la misma forma que lo hacen en otros escenarios sociales, ya que no se pueden extrapolar datos de otros países y mucho menos de países con un amplio desarrollo que cuentan con un contexto social totalmente diferente al nuestro. Es por esto, que en cualquier hospital, y con mayor razón en un hospital público, de tercer nivel y con una cobertura nacional importante, es indispensable conocer el impacto de su atención para posteriormente relacionarlo con los costos que genera, para poder así en un futuro jerarquizar la atención y optimizar el uso de recursos. Sin embargo para poder iniciar una evaluación de este tipo, se debe explorar diferentes abordajes de medición de calidad de vida para encontrar cual o cuales instrumentos pueden ser útiles en este tipo de pacientes.

Objetivos

Objetivo general:

Validar tres instrumentos de calidad de vida (SF-36, Termómetro y TTO) en pacientes hospitalizados con 4 de las principales enfermedades crónicas atendidas en el INNSZ.

Objetivos específicos:

Objetivo específicos.

1. Elaborar y validar un instrumento para la medición de utilidades en salud con técnica de termómetro y trueque para cada una de las enfermedades seleccionadas, midiendo además su reproducibilidad.
2. Validar, medir reproducibilidad y del instrumento SF 36 en cada una de las enfermedades seleccionadas.
3. Comparar la calidad de vida al ingreso entre enfermedades

Material y métodos

Diseño

Se realizó un estudio de tipo observacional, transversal y descriptivo.

Muestra

De cada una de las cuatro enfermedades estudiadas, se seleccionaron pacientes consecutivos que cumplieron con los criterios de selección del estudio.

Criterios de inclusión:

- Diagnóstico de DPCA, CH, SIDA, o LEG acorde a las siguientes definiciones.

Diálisis Peritoneal Continua Ambulatoria (DPCA): Pacientes con insuficiencia renal crónica terminal de cualquier etiología, que al momento del ingreso y por lo menos desde un mes antes, utilizaran diálisis peritoneal continua ambulatoria como método sustitutivo de la función renal.

Cirrosis Hepática (CH): Pacientes, quienes independientemente de la etiología, tuvieran el diagnóstico establecido de cirrosis hepática, ya sea por biopsia registrada en el expediente clínico, o por datos clínicos y/o de laboratorio de insuficiencia hepática según la clasificación de Child-Pugh asociada a datos objetivos de hipertensión portal (varices esofágicas, ultrasonido hepático compatible, y/o esplenomegalia).

Síndrome de Inmuno-Deficiencia Adquirida (SIDA): Serología positiva para el virus de la inmunodeficiencia humana y presencia o historia de alguna de las enfermedades de definición del diagnóstico, según los criterios del Centro de Control de Enfermedades (CDC) de 1993 **(18)**.

Lupus Eritematoso Generalizado (LEG): Pacientes con el diagnóstico establecido por un reumatólogo, con presencia o historia de al menos 3 criterios de clasificación del Colegio Americano de Reumatología de 1982 para LEG **(19)**

- Cualquier edad y sexo.
- Ingreso por cualquier causa al área de hospitalización del INCMNSZ.
- Firma de consentimiento informado.

Criterios de exclusión:

- Pacientes quienes por sus condiciones físicas o mentales estuvieran incapacitados para contestar los instrumentos de calidad de vida en las primeras 24 horas de ingreso al hospital.
- Pacientes hospitalizados en el Servicio de Urgencias, quienes por razones propias del servicio, no pudieran completar la evaluación inicial durante las primeras 24 horas de su internamiento.
- Pacientes foráneos, quienes no pudieran acudir a su visita de seguimiento (30 días después del ingreso).

Mediciones e instrumentos:

Variables sociodemográficas:

Al momento del ingreso se registraron, de forma prolectiva y basándose en entrevista con el paciente y sus familiares, las variables sociodemográficas rutinarias: Edad, género, estado civil, escolaridad, años de educación formal, ocupación principal, y clasificación socioeconómica asignada por el Departamento de Trabajo Social del INCMNSZ.

Variables globales de la enfermedad:

Para valorar la estabilidad de la enfermedad principal en cada pacientes se registró, basándose en la respuesta del paciente y en la revisión posterior de los expedientes, el número de hospitalizaciones y de consultas al Servicio de Urgencias que tuvieron en el último año. Como indicador de gravedad de la enfermedad y de la intensidad terapéutica de forma global en las cuatro enfermedades se registró el número de medicamentos tomados al ingreso, de forma regular, sin tomar en cuenta la dosis. De la enfermedad también se registró el tiempo de evolución, y la presencia de comorbilidad definida como cualquier otra enfermedad crónica agregada.

Variables específicas por enfermedad:

En cada una de las cuatro enfermedades se registraron variables particulares de gravedad y de descripción que se consideraron importantes. Se utilizaron diferentes índices clínicos o de laboratorio de gravedad que han sido ampliamente utilizados en las enfermedades de interés, y que a continuación se describen:

Cirrosis Hepática: Se registró como descripción de la enfermedad, su etiología, la presencia o no de hipertensión portal hemorrágica, el número de sangrados variceales en el último año, la historia o no de cirugía derivativa de hipertensión portal y por último la valoración del estado nutricional por medio del Índice de Masa Corporal (IMC).

Por otro lado se utilizó la clasificación de Child-Pugh, utilizada ampliamente para definir gravedad de la insuficiencia hepática y que comprende evaluaciones laboratoriales y clínicas (bilirrubina total, albúmina sérica, tiempo de protrombina, grado de encefalopatía y grado de ascitis), y se califica en una escala de menor (5 puntos) a mayor (15 puntos) gravedad de la insuficiencia hepática para después categorizarla en tres estadios: A menor gravedad y C mayor gravedad.

Lupus Eritematoso Generalizado: Para la descripción basal de la enfermedad se utilizó la escala de daño crónico del Colegio Americano de Reumatología (SLICC/ACR) que ya ha sido validada en población mexicana (20) y utilizada en el INCMNSZ. Es una escala del daño acumulado desde el diagnóstico de LEG hasta la actualidad en 12 órganos o sistemas con un puntaje de 0 (ausencia de daño acumulado) a 48 (grado máximo de daño). También se registró y clasificó la intensidad terapéutica en 0 (prednisona <15 mg./día sin drogas citotóxicas a excepción de cloroquina) 1 (prednisona >20 mg../día sin drogas citotóxicas) o 3 (utilización de cualquier droga citotóxica a excepción de cloroquina independientemente de la dosis de prednisona). Esta calificación también ha sido utilizada y validada previamente (21).

Por último se clasificó a los pacientes con la escala de actividad lúpica validada en pacientes del INCMNSZ (Mex SLEdai) y que tiene una buena sensibilidad al cambio. Es una escala que comprende variables clínicas y de laboratorio, midiendo la actividad en varios órganos o sistemas con un puntaje de 0 (sin actividad) a 34 (máxima actividad posible) (22).

Síndrome de Inmunodeficiencia Adquirida (SIDA): Como descripción basal de la enfermedad se registró el Índice de Masa Corporal (IMC), el Karnofsky al ingreso, el tipo de evento de definición de SIDA, la última cuenta de CD4 total, su categorización en A, B o C según la OMS y el estadio clínico también definido por la OMS en 1, 2 o 3.

Para la gravedad de la enfermedad, se utilizó una escala de medición de los 9

principales síntomas relacionados con SIDA, es decir: alteraciones en memoria, convulsiones, debilidad de extremidades, fiebre por más de 4 días, escalofríos, diaforesis, pérdida de peso de >5 kilos, disnea y diarrea. Se interrogó la presencia y frecuencia de cada uno de estos síntomas en los últimos tres meses y se codificó cada uno en 0 (sin síntoma), 1 (ocasional, ≤ 1 vez por semana) o 2 (frecuentemente, más de una vez por semana). En el caso de pérdida de peso solo se calificó en 0 (sin pérdida de peso > de 5 kilos) o 1 (con pérdida de peso >5 kilos). Posteriormente se sumó la calificación de cada uno de síntomas y se obtuvo un calificación total con un intervalo de 0 (lo menos grave) a 17 (lo más grave). Esta escala ha sido utilizada para validar medidas de calidad de vida (23).

Por otro lado se utilizó y midió una escala de estadio pronóstico llamada CAPS (Clinical AIDS Prognostic Staging), que contempla indicadores clínicos y de laboratorio para determinar la presencia o ausencia de linfopenia, anemia, trombocitopenia, compromiso neurológico, hipoxemia, hipoalbuminemia y diarrea constante, para posteriormente clasificarla en 0 (ningún compromiso), 1 (un compromiso), 2 (dos compromisos) y 3 (tres o más compromisos). Esta escala ha sido validada previamente con una adecuada sensibilidad al cambio, además de que es muy fácil obtener los datos que contempla (24).

Diálisis Peritoneal Continua Ambulatoria (DPCA): Para la descripción de la enfermedad se registró la etiología de la insuficiencia renal, el IMC, y la historia de trasplante renal. Para valorar el nivel de gravedad y estabilidad se registró el número de eventos de peritonitis en el último año, el número de baños dialíticos requeridos y las cifras de hemoglobina en gr/dl.

Por último se utilizó un índice bioquímico (CAS: Chemistry Abnormality Score) de funcionamiento de diálisis, utilizado en estudios previos para validar instrumentos de calidad de vida (25). Mide el funcionamiento dialítico por medio de 13 estudios de laboratorio asignando un puntaje de 0 (al mejor funcionamiento) y de 23 (al peor funcionamiento). Estos estudios de laboratorio se toman de rutina a cualquier paciente con diálisis al ingreso hospitalario, y en el caso de la evaluación de seguimiento en la consulta externa se tomó la cifra de laboratorio de la fecha de la evaluación ± 2 semanas. En el caso de no tener la medición de laboratorio no se contabilizó este índice.

Causa de hospitalización:

Se registró la causa de la hospitalización describiéndola detalladamente para posteriormente, clasificarla de forma exhaustiva y mutuamente excluyente en: Tratamiento electivo, tratamiento urgente o procedimiento diagnóstico.

Tratamiento electivo: Aquellos pacientes que ingresaron con el plan de algún tipo de tratamiento médico o quirúrgico, de forma electiva, y que no tuvieran una indicación totalmente necesaria. No se tomaron en cuenta a aquellos pacientes estables que ingresaron para tratamientos necesarios, entendiéndose por esto a aquellos tratamientos dirigidos a una causa que pusiera en peligro la vida.

Tratamiento urgente: Aquellos pacientes que ingresaron para recibir un tratamiento necesario y/o urgente.

Procedimiento diagnóstico Pacientes estables o inestables que ingresaron para diagnóstico sin un plan terapéutico definido.

Así mismo, se registró la procedencia del paciente en términos de ingreso directo de su domicilio o ingreso por urgencias, terapia intensiva o consultorio de urgencias.

Instrumentos de calidad de vida

Se evaluó la calidad de vida con los siguientes instrumentos:

1. Perfil de salud con la escala SF-36 en sus ocho dominios y subescalas estandarizadas,
2. Utilidades en salud por la técnica del termómetro y de trueque

En presencia de una gran diversidad de instrumentos, en cualquier estudio que intente evaluar calidad de vida, es importante justificar la aplicación de alguno de ellos, además de evaluar su reproducibilidad, validez y en el caso que sea necesario sensibilidad al cambio, ya que como se ha comentado la definición de calidad de vida y su medición aún no es lo suficientemente sólida para asumir validez y reproducibilidad en cualquier contexto (10).

El principal factor de decisión para escoger algún instrumento radica en sus propiedades y funciones que se requieren, para cumplir con los objetivos del estudio. Además de lo anterior, es recomendable la utilización de instrumentos, que ya hayan sido validados y medida su reproducibilidad, y preferentemente en el tipo de población que se intenta estudiar. En los últimos años, se ha dado la tendencia de utilizar varios instrumentos, ya que para muchos estudios son complementarios y por otro lado su correlación puede indicar cierto grado de validez

de la medición. De esta forma la utilización de instrumentos genéricos junto con específicos, es una estrategia recomendada, ya que unos tienen mejor función evaluadora y otros tienen mejor función discriminativa y, en conjunto, pueden dar las dos funciones en un mismo estudio.

También la utilización de perfiles de salud y utilidades en salud en un mismo estudio, permite correlacionar el índice único proporcionado por el instrumento de utilidades en salud, con la calificación de cada uno de los dominios de calidad de vida definidos en el perfil de salud, y así poder definir cuál de los dominios se refleja con mayor precisión en el índice de utilidades.

En este estudio se decidió la utilización de varios instrumentos de calidad de vida, y ya que el objetivo principal fue aplicarlos en varias enfermedades, fue necesario incluir en la selección algún instrumento genérico como lo es el SF-36. Sin embargo, con la utilización del termómetro para medir utilidades, aunque se tiene una estructura común entre las enfermedades, los escenarios clínicos que se utilizan, abordan síntomas específicos de cada enfermedad, por lo que se comparten algunas características con los instrumentos específicos, principalmente la de asociación con gravedad de la enfermedad y la sensibilidad al cambio. Es por esto que en este estudio se decidió la utilización de un instrumento genérico del tipo de perfil de salud (SF-36), y dos instrumentos de utilidades en salud específicamente, el termómetro y el trueque o Time Trade Off (TTO).

SF-36

El Measures of Outcomes Study Short Form health survey (SF-36) se creó a partir de un estudio que realizó la corporación Rand, en el que se pretendió evaluar el impacto de un sistema de seguros alternativo en el estado de salud de la población (26). De su versión original, de 245 preguntas, se derivaron versiones más cortas que fueran lo suficientemente sensibles y discriminadoras, de las que surgió en 1992 la forma reducida de 36 preguntas (SF 36) para medir la calidad de vida. Esta versión, evalúa ocho dominios y una escala de transición del estado de salud que no se considera para su calificación. Los ocho dominios son: Función Física (FF), Rol Físico (RF), Dolor corporal (DC), Salud General (SG), Vitalidad (VT), Función Social (FS), Rol Emocional (RE) y Salud Mental (SM) (Tabla 2)

TABLA 2. Dominios del SF-36

Dominio.	No. de preguntas	Aspectos evaluados
Función Física	10	Limitaciones para realizar las actividades físicas
Rol Físico	4	Problemas con el trabajo u otras actividades por problemas físicos
Dolor Corporal	2	Dolor y sus limitaciones
Salud General	5	Salud general y sensación de empeoramiento
Vitalidad	4	Energía, cansancio, agotamiento, entusiasmo
Función Social	2	Interferencia en actividades sociales normales por problemas físicos y emocionales
Rol Emocional	3	Problemas con el trabajo u otras actividades por problemas emocionales
Salud Mental	5	Nerviosismo, depresión, felicidad, calma
Transición en salud	1	Comparación de salud actual con la de un año antes
TOTAL	36	

Es un instrumento de auto administración que toma aproximadamente 10 - 15 minutos en su llenado, logrando así un adecuado porcentaje de respuesta. Para cada pregunta, existe una forma de codificar y re-codificar cada respuesta en una escala del 0 (peor estado de salud) al 100 (mejor estado de salud), y posteriormente se obtiene un promedio para cada dominio obteniendo así una escala cuantitativa continua para cada dominio. Para este estudio se siguió con el algoritmo de calificación recomendado por los autores del instrumento y que se explica detalladamente en el manual del usuario (27). Esta forma de calificación permite la generación de dos subescalas mediante una ponderación diferenciada de cada dominio, resumiendo así a los ocho dominios mismas que son la Subescala de Salud Física (SSF) y la Subescala de Salud Mental (SSM). Estas dos subescalas se generaron con análisis factorial y su validez y reproducibilidad han sido ampliamente demostradas (28).

El SF-36 ha sido utilizado en numerosos estudios, demostrando su validez y reproducibilidad en diferentes enfermedades y situaciones. La reproducibilidad se ha medido mediante consistencia interna en un estudio de 3,445 pacientes con diferentes enfermedades crónicas, derivados del estudio original de la corporación Rand, mostrando índices alfa de Cronbach mayores de 0.8 en todos los dominios (29). También se ha probado reproducibilidad con prueba repetida, excediendo un coeficiente de reproducibilidad mayor a 80%. (30). En

cuanto a la validez, ésta ya ha sido de mostrada en varias enfermedades principalmente con validez de constructo, utilizando diferentes escalas clínicas que valoran gravedad, y por lo tanto teóricamente diferentes grados de calidad de vida. Este instrumento ya ha sido validado en las cuatro enfermedades de interés para nuestro estudio: SIDA (23), hepatopatía crónica (31), lupus (21) e IRCT (32). En población mexicana, ha sido validado en la población general (33), pero los estudios en enfermedades específicas son pocos, por lo que consideramos necesario probar su validez dentro de este estudio.

Existen dos variantes principales de este instrumento, la de cambios crónicos que evalúa los cambios en el último mes, y la de cambios agudos, que los evalúa en la última semana. Para este estudio, por su corto tiempo de seguimiento utilizamos la segunda. Esta variante, a pesar de ser un instrumento genérico, ha demostrado su sensibilidad al cambio en varios estudios, lo que ofrece una gran ventaja para responder la pregunta de investigación de este estudio. (26)

Otra ventaja que tiene este instrumento, es que en un estudio previo realizado en el INCMNSZ, se demostró su validez y reproducibilidad, en un grupo de población mexicana. En este estudio se estudiaron 24 pacientes con diferentes diagnósticos reumáticos, a los que se les realizó reemplazo total de cadera o rodilla. Se compararon instrumentos específicos como el AIMS (Escala de medición de impacto en artritis), genérico como el SF-36 y el HAQ (cuestionario de evaluación en salud), que es un instrumento que evalúa capacidad física diaria. Se evaluaron los instrumentos antes de la cirugía, una semana y un mes después, midiendo reproducibilidad, validez y sensibilidad al cambio. Se demostró con prueba repetida, un índice de reproducibilidad mayor a 0.9 en todos los instrumentos. La sensibilidad al cambio del SF-36 medida como porcentaje de cambio fue de 12%, comparable con la del instrumento específico. (34)

Instrumentos de Utilidades en Salud

El otro tipo de instrumentos que se utilizaron en este estudio es el de utilidades en salud, ya que la utilización con instrumentos genéricos ofrece las ventajas antes mencionadas, y consideramos, que es de las medidas, que en un solo índice compara tanto los efectos positivos como negativos de la intervención, además de las preferencias de cada paciente, lo que puede facilitar su análisis y comprensión. Este tipo de instrumentos igualmente ha sido ampliamente utilizado, y validado en las enfermedades de interés en nuestro estudio.

La medición de utilidades, se realizó con dos instrumentos. La razón de utilizar dos instrumentos, a pesar de que su comparación absoluta es imposible, fue que cada una de ellos tiene ventajas importantes.

Termómetro

Consiste en poner un termómetro o escala visual análoga vertical, que va del 0 (peor estado de salud imaginable incluso la muerte) al 100 (mejor estado de salud imaginable). Se le dan al paciente 4 tarjetas que contienen el nombre y una descripción de enfermos supuestos, con el mismo tipo de patología que padece el paciente pero con 4 distintos escenarios cuya calidad de vida va de muy bien (cerca de 1) a muy mal (cerca de 0). El enfermo acomoda dichas tarjetas en el termómetro y finalmente se le pide que se sitúe a sí mismo en un punto. Se obtuvieron 3 medidas con este instrumento:

1. El punto donde se situó el paciente en la escala (**Paciente**).
2. La medida del paciente menos la medida del escenario hipotético más grave.(**Ajuste 1**), y
3. La medida del paciente ajustada a la escala comprendida entre el escenario hipotético más grave y el menos grave (**Ajuste 2**).

La calificación más utilizada es la primera, sin embargo en este estudio decidimos probar los dos ajustes mencionados, ya que si la concepción de muerte del paciente y su valor no son claros, la diferencia contra su propia referencia puede dar un valor más real. El resultado de esta calificación, traducido a una escala de 0 a 1, es la medición de utilidades en salud.
(12)

Este método difiere de una escala visual análoga, en que al paciente se le proporcionan estados de salud diferentes (escenarios) como marcadores, y referencia para el mismo paciente. Además, estos marcadores pueden servir para determinar las utilidades, y también es una forma de medir tanto reproducibilidad como sensibilidad al cambio. Para esto se valora la variabilidad de la calificación de los diferentes escenarios, que debe ser menor a la variabilidad de la calificación del paciente si es que en realidad hubo un cambio en su calidad de vida.

Este método, aunque es un instrumento genérico, su estructura le da ciertas ventajas sobre el resto de los instrumentos genéricos, principalmente en la sensibilidad al cambio. Como se puede ver en los escenarios para cada enfermedad, están compuestos de 10 a 11 puntos o características, que definen al paciente imaginario. En las cuatro enfermedades, los primeros

5 puntos, son exactamente iguales, e intentan cubrir los dominios de función física, función social, estado emocional o mental, función ocupacional, el resto de los puntos son diferentes para cada enfermedad, ya que intentar cubrir el dominio de sensación somática, y para esto cada enfermedad, tiene principales síntomas. Por esto consideramos que este tipo de instrumentos tiene la ventaja de funcionar como genérico, ya que comparten la misma estructura, y por lo tanto son comparables, además de funcionar como específicos en su función evaluadora, ya que se hace mención a los síntomas principales de cada enfermedad. Por último también se incluyen aspectos negativos tanto somáticos como económicos de la enfermedad, además de evaluación general del estado de salud.

Este tipo de instrumentos también ha sido validado y medida su reproducibilidad en pacientes del INCMNSZ. En 22 pacientes con AR. se obtuvo un índice Kappa de reproducibilidad de 0.71 lográndose discriminar a los pacientes con diferentes grados de gravedad de la enfermedad. (35).

Time Trade- Off (TTO)

El otro método de utilidades que utilizamos fue el de trueque o Time Trade-Off (TTO). En este instrumento se utiliza la respuesta del paciente a diferentes situaciones de riesgo, para determinar el valor que el paciente le da a su calidad de vida. Este tipo de métodos, son los que originalmente surgieron de la teoría económica y de la de toma de decisiones. Se decidió utilizar esta método, ya que ha sido ampliamente utilizado en varios ensayos clínicos, demostrando su validez, sensibilidad al cambio y reproducibilidad. Esto se ha realizado en las enfermedades que evaluamos.

A pesar de su gran difusión y utilización en todo el mundo, aún no ha sido probado en población mexicana, por lo que será importante medir sus cualidades clinimétricas, ya que puede ser un método factible, rápido y útil. Se decidió por la modalidad de trueque, ya que consideramos, que el manejo de probabilidad que se utiliza en el método de ruleta de probabilidades es complicado, y pensamos que no es entendido por la población general.

La modalidad de trueque, es una técnica implícita que no incluye la valoración de riesgo. En este método se le plantean al paciente dos alternativas. La alternativa 1 es que el paciente mantenga su propio estado de salud por el resto de su vida (tiempo t), mientras que la otra alternativa es menor tiempo (tiempo x) en mejor estado de salud. El tiempo x varía hasta que el paciente es indiferente a las dos alternativas; en este punto, la medida de utilidad es x/t .

Este método ha sido utilizado en varios estudios y ensayos clínicos controlados, como variable final de interés. Se ha encontrado altos índices de reproducibilidad con prueba

repetida (mayor de 0.8) tanto en población general, como en diferentes enfermedades. En México no se encontraron informes de este método, sin embargo, en el mundo ha sido ampliamente utilizado, y validado en muchas enfermedades y contextos diferentes. (36, 37, 38).

Otras mediciones

Se utilizó una escala visual análoga (EVA) de 10 cm. con la pregunta: ¿Cómo se siente hoy? en donde el 0 correspondía a "MUY BIEN" y el 10 a "MUY MAL". Con esto se intentó evaluar el estado de salud del paciente.

Evaluación de confiabilidad (reproducibilidad):

La reproducibilidad de los instrumentos que se utilizaron para medir calidad de vida actual, se evaluó por medio de prueba repetida (test-retest). Esta fue realizada en una sub-muestra de todos los pacientes con representación de cada una de las enfermedades. Debido a que la posibilidad de cambio en las primeras 24 horas de la hospitalización era alta, sólo se realizó en aquellos pacientes que ingresaron para procedimientos electivos y que se encontraban sin una alteración aguda concurrente, y en aquellos pacientes que ingresaron para procedimientos diagnósticos y que a juicio del médico tratante y del investigador principal se encontraban estables. A los pacientes se aplicaron nuevamente los tres instrumentos principales en un lapso de tiempo no menor de 12 horas y no mayor de 24 horas posterior a la aplicación de los instrumentos al ingreso de la hospitalización. La reproducibilidad se ponderó mediante coeficientes de correlación intraclase. No se consideró necesario realizar la medición de reproducibilidad previo al trabajo, ya que los dos instrumentos principales que se utilizaron (utilidades con técnica de termómetro y el SF-36), ya han sido utilizados en el INCMNSZ, obteniendo buenos índices de reproducibilidad (34, 35). Para la técnica de trueque, se consideró que el realizar la medición antes del estudio, resultaría poco factible, y ya que no fue el único instrumento utilizado, tampoco se consideró realizarla previa al estudio

Consistencia Interna:

Se evaluó la consistencia interna en el SF-36 por medio de la determinación de la α de Cronbach de los cuestionarios al ingreso, para cada uno de los dominios estudiados. Este análisis, considerado por otros autores como prueba de reproducibilidad, compara la varianza de cada una

de las preguntas de un dominio con la varianza de la calificación promedio.

Validación:

A este respecto, se evaluó la validez de constructo y de contenido solamente en el caso del termómetro.

Validez de contenido.

Sólo se evaluó en el del termómetro. En el caso del SF-36, esto ya se ha realizado en muchas ocasiones, por lo que no consideramos necesaria una nueva validación. En cuanto a la técnica de trueque, ya que el contenido no se expone de forma explícita, no se puede realizar este tipo de evaluación.

Se realizó un cuestionario de preguntas cerradas y abiertas a por lo menos 5 expertos en cada una de las enfermedades seleccionadas. Los expertos fueron médicos internistas especialistas de reumatología para el caso de LEG, nefrología para el caso de DPCA, gastroenterología para el caso de CH e infectología para el caso de SIDA. En este cuestionario se solicitó a los expertos que juzgaran si los dominios en los escenarios clínicos de los termómetros eran adecuados y si existía algún otro dominio o síntoma que agregar. Posterior a esto, y con las respuestas, nos juntamos el investigador principal y los tutores, para modificar los escenarios con base en las respuestas del cuestionario.

Validez de constructo.

Para la validez de constructo se utilizaron tanto supuestos hipotéticos generales para toda la población como específicos para cada una de las enfermedades. Estos constructos o supuestos teóricos, han sido utilizados en otros estudios con el mismo fin, y lo que se busca es que la medición en calidad de vida se relacione con estos supuestos, aunque por definición no pueden tener una correlación perfecta, ya que corresponden a mediciones de características diferentes. Todas las pruebas de validación de constructo se realizaron exclusivamente con los instrumentos aplicados al ingreso.

Los supuestos para la evaluación global son los siguientes:

- a) La edad tiene una relación inversa con la calidad de vida, principalmente cuando se valoran dominios de función física. (30)
- b) Las mujeres por lo general tiene una calidad de vida menor, principalmente en dominios de salud mental. (30)

- c) Los pacientes con nivel socioeconómico menor tienen una calidad de vida peor.
- d) Los pacientes que han tenido hospitalización o consultas en urgencias en el año previo a su ingreso al estudio, tienen una enfermedad por lo general menos estable y posiblemente más grave, y por lo tanto se espera que tengan una calidad de vida peor
- e) A mayor co-morbilidad peor calidad de vida.
- f) La causa de la hospitalización está relacionada con la calidad de vida de los pacientes, en el sentido de que los pacientes que ingrese por problemas urgentes tendrán peor calidad de vida.
- g) El lugar de procedencia (Urgencias, UTI o directo de su domicilio) está relacionado con la gravedad de los pacientes y por lo tanto con la calidad de vida.

Los primeros tres supuestos se han utilizado principalmente en población general, por lo que en este estudio con padecimientos crónicos, es probable que estas características tengan menor influencia sobre la calidad de vida.

Para los constructos hipotéticos de cada enfermedad se utilizaron medidas de gravedad definidas previamente, que son escalas o variables que se han relacionado en estudios previos con calidad de vida. Además en todas las enfermedades se probó el supuesto teórico de que a mayor tiempo de evolución, peor calidad de vida.

Para cada una de las enfermedades se utilizaron los siguientes supuestos teóricos:

CH

- a) El grado de insuficiencia hepática medida con la clasificación de Child- Pugh, tiene una relación inversa con la calidad de vida.
- b) La historia de hipertensión portal hemorrágica se asocia con mayor gravedad de la cirrosis hepática y por lo tanto a peor calidad de vida.
- c) El grado de desnutrición medido con IMC esta asociado a peor calidad de vida.
- d) La cantidad de sangrados en el último año, tiene una relación inversa con la calidad de vida

LEG

- a) El grado de actividad lúpica (Mex-SLEdai) tiene una relación inversa con la calidad de vida.

- b) El índice de daño acumulado (SLICC/ACR) tiene una relación inversa con la calidad de vida.
- c) La dosis de prednisona al momento del ingreso, probablemente se asocia a mayor gravedad, y por lo tanto tiene una relación inversa con la calidad de vida.
- d) A mayor necesidad terapéutica (clasificación de medicamentos) peor calidad de vida.

DPCA

- a) A mejor funcionamiento dialítico (índice de funcionamiento), mejor calidad de vida.
- b) A mejor estado nutricional mejor (IMC), mejor calidad de vida.
- c) A mejores cifras de Hb al ingreso, mejor calidad de vida
- d) El haber tenido eventos de peritonitis por el catéter peritoneal, se asocia a peor calidad de vida.

SIDA

- a) A mayor número de CD4 mejor calidad de vida.
- b) A mayor Karnofsky mejor calidad de vida, preferentemente la función física.
- c) A mayor puntaje en la escala de síntomas, peor calidad de vida.
- d) A mejor estado nutricional, mejor calidad de vida.

Estrategia analítica.

Resumen de datos

Para las características de los pacientes, tanto demográficas como clínicas se utilizaron frecuencias y porcentajes, para resumir las variables nominales y ordinales, o promedio y desviaciones estándar o medianas y mínimos y máximos, para las variables numéricas dependiendo de su distribución de frecuencias.

Para las medidas de calidad de vida se utilizaron medianas y valores mínimos y máximos, ya que presentaron distribuciones asimétricas en la mayoría de los instrumentos de calidad de vida. En el caso de la medición en las dos subescalas agrupadoras del SF-36, la distribución permitió la descripción en términos de medias y desviación estándar.(DE).

Análisis de datos:**Instrumentos de medición.**

Como se comentó en la sección de reproducibilidad, este análisis se realizó por medio de coeficientes de correlación intra-clase entre los resultados de prueba repetida y los de ingreso. La prueba de hipótesis analizó su diferencia de cero.

Se calculó la consistencia interna por medio del α Cronbach para cada uno de los dominios del SF-36.

Para la validación se utilizó la Ji-cuadrada o la prueba de *Kruskal-Wallis* para variables categóricas, y el índice de correlación no paramétrica (*rho de Spearman*), para variables cuantitativas. También se compararon los resultados entre instrumentos con el mismo tipo de correlación.

Consideraciones Éticas:

Se siguieron los principios que para la investigación en humanos postula la declaración de Helsinki y la Guía de la Buena Práctica Clínica de los Estatutos de la Organización Europea para la Evaluación de los Productos Médicos.

A cada paciente se le explicó de manera detallada en qué consistía la investigación, los riesgos y beneficios de la misma obteniéndose consentimiento informado en todos ellos.

Se hizo hincapié en que cada paciente tendría la total libertad para participar, sin que su negativa o retiro afectara su atención habitual.

Se aseguró que los resultados de la investigación serían usados en forma confidencial y únicamente para fines de este estudio.

Los pacientes no recibieron ningún tipo de remuneración económica por su participación.

Resultados

Pacientes

En un período de reclutamiento de 10 meses aproximadamente (agosto de 1999 a mayo del 2000), se incluyeron 75 pacientes que se hospitalizaron por motivos varios en INCMNSZ y que entre sus diagnósticos principales tenían alguna o algunas de las cuatro enfermedades en estudio: Cirrosis hepática (26), diálisis peritoneal continua ambulatoria (16), lupus eritematoso generalizado (19) y síndrome de inmunodeficiencia adquirida (14).

Se incluyeron pacientes consecutivos que ingresaban a los sectores de hospitalización, ya sea directamente del domicilio o provenientes de Urgencias con menos de 24 horas de internamiento, que cumplieron con los criterios de inclusión y exclusión, y que en la solicitud de internamiento se especificara la presencia de alguno de los diagnósticos de interés (CH, DPCA, LEG, SIDA). No se incluyeron pacientes cuyo motivo de ingreso al hospital fue un evento agudo grave que condicionó un período de vigilancia en Urgencias o Terapia Intensiva mayor de 24 horas, situación que es muy frecuente en estos cuatro padecimientos.

En dos pacientes con DPCA, la insuficiencia renal era secundaria a LEG, por lo que en los análisis independientes por enfermedad principal, estos pacientes se analizaron en los dos grupos, resultando en un total de 20 pacientes con LEG y 17 con DPCA. Para el análisis global, se tomó en cuenta un paciente en el grupo de DPCA, ya que la causa de internamiento estaba relacionada con este diagnóstico (peritonitis por catéter de Tenckhoff) y el otro paciente se incluyó en el grupo de LEG, ya que la causa de la hospitalización era artritis séptica recidivante, que estaba más en relación con LEG y su tratamiento inmunosupresor.

Características Sociodemográficas de la muestra

Las características generales y sociodemográficas de cada grupo diagnóstico y del grupo total se presentan en la Tabla 3. Es notoria la distribución diferencial de varias características sociodemográficas en los grupos diagnósticos de interés. Después de realizar el análisis estadístico pertinente, se encontró que en todas las variables presentadas en la tabla a excepción de los años de educación formal y la clasificación socioeconómica, existió por lo menos un grupo diagnóstico que fue diferente al resto a un nivel de significancia de $p < 0.05$. En el caso de la edad, es evidente que el grupo de pacientes con cirrosis es el más viejo, con una diferencia de promedios de casi 20 años con los otros tres grupos, lo que se refleja también el porcentaje de solteros en este mismo grupo. El género fue otra de las variables donde hubo mucha diferencia

principalmente por lo opuesto entre SIDA y LEG, que era totalmente esperable, por la epidemiología de estas enfermedades. En las otras variables, principalmente los grupos de SIDA y LEG presentaron mayor escolaridad y mayor porcentaje de profesionistas que el resto.

TABLA 3. Características sociodemográficas por diagnóstico y por el grupo total

CARACTERÍSTICA	CH (n=26)	DPCA (n=17)	LEG (n=20)	SIDA (n=14)	Total (n=75)
Edad (años):	55.2 ± 13.5	35.6 ± 12.8	33.2 ± 9.1	36.2 ± 8.3	41.9 ± 14.9
Sexo: hombres	50 (13)	65 (11)	10 (2)	93 (13)	52 (39)
Estado civil					
Soltero	7.7 (2)	23.5 (4)	45 (9)	57.1 (8)	30.7 (23)
Casado	80.8 (21)	70.6 (12)	45 (9)	28.6 (4)	58.7 (44)
Viudo	7.7 (2)	5.9 (1)	0 (0)	0 (0)	4 (3)
Divorciado	3.8 (1)	0 (0)	10 (2)	14.3 (2)	6.7 (5)
Escolaridad					
Secundaria o menos	50 (13)	64.7 (11)	35 (7)	14.4 (2)	41.3 (31)
Técnica o preparatoria	15.4 (4)	23.5 (4)	20 (4)	42.8 (6)	24 (18)
Licenciatura o más	34.6 (9)	11.8 (2)	45 (9)	42.8 (6)	34.7 (26)
Años de educación	10.4 ± 6.5	10.1 ± 4.5	13.0 ± 4.9	14.0 ± 3.5	11.7 ± 5.4
Ocupación					
Desempleado	11.5 (3)	23.5 (4)	20 (4)	0 (0)	13.3 (10)
Hogar o estudiante	42.3 (11)	23.5 (4)	40 (8)	7.2 (1)	30.6 (23)
Empleado	23.1 (6)	35.3 (6)	15 (3)	64.4 (9)	32.1 (24)
Profesional	23.1 (6)	17.7 (3)	25 (5)	28.4 (4)	24 (18)
Clasificación socioeconómica					
1 y 2	23 (6)	35.3 (6)	50 (10)	28.6 (4)	33.3 (25)
3 y 4	46.2 (12)	58.8 (10)	40 (8)	57.1 (8)	49.4 (37)
5 y 6	30.8 (8)	5.9 (1)	10 (2)	14.3 (2)	17.3 (13)

Las variables cuantitativas se expresan en términos de $X \pm DE$; las variables cualitativas se expresan como % (n)

Características clínicas por grupo diagnóstico.

En cuanto a las características clínicas basales de los pacientes, la distribución por grupo diagnóstico y grupo total se presentan en la tabla 4.

TABLA 4. Características clínicas por diagnóstico y por grupo total

CARACTERÍSTICA	CH (n=26)	DPCA (n=17)	LEG (n=20)	SIDA (n=14)	TOTAL (n=75)
comorbilidad	35 (9)	100 (17)	68 (13)	28 (4)	55 (41)
Causa de hospitalización					
Electiva	23 (6)	50 (8)	10 (2)	0 (0)	21 (16)
Urgente	31 (8)	38 (6)	58 (11)	64 (9)	45 (34)
Diagnóstica	46 (12)	12 (2)	32 (6)	36 (5)	34 (25)
Procedencia de Urgencias	34(9)	19 (3)	47 (9)	50 (7)	37 (28)
Sin consultas en el último año a Urgencias	35 (9)	25 (4)	32 (6)	43 (9)	33 (25)
Sin Hospitalización en el último año	46 (12)	31 (5)	63 (12)	71 (10)	52 (39)
Número de medicamentos	2.5 (0-6)	5.5 (1-9)	4 (1-10)	4.5 (0-7)	4 (0-10)
Tiempo de evolución ,meses	17 (1-99)	30 (4-178)	30 (2 -251)	34 (2-74)	23 (1-251)

*Las variables cuantitativas se expresan en términos de Mediana (min-max).
las variables cualitativas se expresan como % (n).*

La comorbilidad fue claramente diferente ($p < 0.0001$) entre diagnósticos, ocasionada por la presencia de hipertensión arterial en todo el grupo de pacientes con DPCA , y en lupus también un número importante presentaba actividad renal y por lo tanto cierto grado de hipertensión arterial.

En cuanto a la causa de la hospitalización, también se encontraron diferencias significativas ($p < 0.01$) entre enfermedades, ocasionadas principalmente por un porcentaje alto de pacientes con LEG y SIDA que ingresaron por problemas urgentes, reflejando la inestabilidad de estas enfermedades, en comparación con los otros grupos diagnósticos, en donde los pacientes que ingresaron al servicio de Urgencias, muchos de ellos requirieron más de 24 horas de vigilancia en este servicio, por lo que no se incluyeron en este estudio. El porcentaje tan alto de tratamientos electivos en pacientes con DPCA se debió a que 7 de los 17 pacientes ingresaron para trasplante renal.

Las diferencias significativas ($p < 0.05$) en el número de medicamentos por día, se debió a la poca cantidad de medicamentos administrados en el grupo de cirrosis.

En el resto de las características presentadas en la Tabla 4 las diferencias entre grupos diagnósticos no fueron estadísticamente significativas ($p > 0.05$).

Aplicación de los instrumentos de calidad de vida

Los instrumentos de calidad de vida, fueron aplicados bajo la vigilancia del investigador principal, por lo que no se obtuvieron respuestas faltantes en los cuestionarios. El tiempo de llenado promedio de los tres instrumentos fue el siguiente: SF-36 11.6 ± 3.5 min, termómetro 6.2 ± 2 min y TTO 4.1 ± 1.2 min. El 45 % de los pacientes se sintieron mejor evaluados con el termómetro, el 28 % con el TTO, y el 27% con el SF-36%.

Reproducibilidad de los instrumentos de calidad de vida

En el estudio de prueba repetida (test-retest) se evaluó un total de 24 pacientes; 7 con CH, 8 con DPCA, 6 con LEG y 3 con SIDA. Todos ellos reportaron estar en las mismas condiciones durante la prueba repetida que en el ingreso.

Los resultados de los coeficientes de correlación intra-clase correspondientes al termómetro y al método de TTO de utilidades en salud, se presentan en la Tabla 5. En la medición del termómetro se incluyen la medición cruda del estado actual del paciente (paciente), la medida del paciente ajustada al escenario hipotético más grave (ajuste 1), la medida del paciente ajustada a la escala comprendida entre el escenario hipotético más grave y el menos grave (ajuste2) y la de los cuatro escenarios hipotéticos que conformaron en instrumento.

Todos los índices de correlación intra-clase fueron superiores a 0.50 oscilando en un intervalo de 0.52 a 0.98. Solo el TTO superó el 0.90 mínimo requerido para comparaciones clínicas individuales. En el caso de los escenarios hipotéticos, la mayoría de los coeficientes estuvo por debajo de 0.70, sin embargo éstos no se utilizan directamente para la medición de calidad de vida. Todos los índices de correlación fueron significativamente diferentes a un nivel de $p < 0.0001$.

TABLA 5. Correlación intra-clase de la prueba repetida del termómetro y del TTO.

	Correlación intra-clase	IC 95%*
Termómetro		
Paciente	0.77	0.54-0.89
Ajuste 1	0.72	0.46-0.87
Ajuste 2	0.76	0.51-0.88
Escenario 1	0.66	0.36-0.84
Escenario 2	0.74	0.48-0.88
Escenario 3	0.60	0.26-0.80
Escenario 4	0.52	0.15-0.75
TTO	0.98	0.95-0.99

*IC 95%. Intervalo de confianza al 95%

En la Tabla 6 se presentan los coeficientes de correlación intra-clase de la prueba repetida de los 8 dominios del SF-36 y de las dos subescalas estandarizadas, además de los valores de las α de Cronbach, para cada uno de los dominios en la prueba de ingreso.

TABLA 6. Reproducibilidad y consistencia interna del SF-36 por dominios y subescalas.

	Correlación intra-clase	IC 95%	α de Cronbach
SF-36			
Función Física	0.91	0.81-0.96	0.88
Rol Físico	0.89	0.76-0.95	0.83
Dolor Corporal	0.93	0.84-0.97	0.95
Salud General	0.91	0.80-0.96	0.77
Vitalidad	0.85	0.69-0.93	0.83
Función Social	0.83	0.65-0.92	0.70
Rol Emocional	0.86	0.70-0.94	0.65
Salud Mental	0.88	0.75-0.95	0.87
Subescala Física.	0.92	0.82-0.96	NA
Subescala Mental	0.89	0.76-0.95	NA

IC 95%: intervalo de confianza al 95%.

NA: No aplicable.

Los índices de correlación intra-clase fueron buenos oscilando en un intervalo de 0.83 a 0.93, y en la mayoría de los intervalos de confianza no se incluye el valor de 0.70. Las α de Cronbach todas menos una superaron el 0.70 requerido para comparaciones de grupos. Sólo el dominio que evalúa dolor del SF-36, superó el 0.90 requerido para comparaciones clínicas.

Validez de contenido

Este tipo de validación sólo se realizó en el instrumento del termómetro por las razones mencionadas previamente.

Como parte de la validez de contenido se realizó un cuestionario a expertos de cada enfermedad, para que analizaran el contenido teórico de los escenarios hipotéticos del termómetro. Estos escenarios intentaban cubrir los siguientes dominios de la calidad de vida: Función física, función social, estado emocional, función laboral, sensación somática, problemas secundarios al tratamiento tanto físicos como económicos, y autopercepción general de salud.

En total se enviaron 27 cuestionarios (7 a reumatología, 7 a infectología, 7 a gastroenterología y 6 a nefrología. Se obtuvo un porcentaje de respuesta general de 81.5 % (22/27), el cual se distribuyó entre las especialidades de la siguiente forma: Reumatología 85.7% (6/7), Infectología 85.7% (6/7), gastroenterología 85.7% (6/7), y nefrología de 66.6% (4/6)

En general hubo una mayoría de expertos que estuvo de acuerdo con los escenarios hipotéticos inicialmente planteados. En el caso de gastroenterología para CH, y reumatología para LEG estuvieron de acuerdo arriba de 90% de los expertos y en el caso de los otros dos grupos de especialistas (infectología y nefrología) alrededor de 70 %.

Se recibió un número importante de comentarios por parte de los expertos, mismos que se incluyeron en la modificación posterior de los escenarios. Después de la discusión con el comité tutorial, se analizaron los comentarios agregados en los cuestionarios. Los comentarios fueron de varios tipos y para su análisis se dividieron en tres:

1. Comentarios relacionados con la adición de una serie de variables que habíamos considerado inicialmente en el protocolo como variables de confusión y que consideramos que era imposible reflejarlas en los escenarios hipotéticos. El origen de este tipo de comentarios seguramente fue secundario al desconocimiento que tenían los expertos de los objetivos y alcances del protocolo, además del funcionamiento del instrumento de medición. Sólo una minoría (principalmente reumatólogos) había trabajado con calidad de

vida y muchos menos con la técnica del termómetro para medir utilidades. Estos comentarios no se consideró necesario que formaran parte de los escenarios clínicos hipotéticos, ya que se contemplaron en la hoja de recolección del estudio y además no eran parte de un dominio en especial de calidad de vida.

Ejemplos de este tipo de comentarios son los siguientes.

- El tipo de tratamiento recibido. Se considera que el tipo de medicamentos y tratamiento modifica la calidad de vida desde varios puntos de vista, ya sea mejorando síntomas, mejorando sobrevida, ocasionando efectos secundarios, etc. Por lo anterior no se agregaron en los escenarios pero si se incluyo en la recolección de información, para considerarlos posteriormente como potenciales confusores.
 - Nivel educativo. También se consideró en la hoja de recolección y se manejó como confusor.
 - Causa de la IRCT, nivel socioeconómico, tiempo de evolución, etc. que también se consideraron en la hoja de recolección como variables de confusión.
 - Actividad de LEG con Mex-Sledai. También se consideró como variable independiente.
 - Comorbilidad. También se consideró.
2. Comentarios que sugerían la modificación de redacción o vocabulario de ciertos puntos ya contemplados en los escenarios clínicos. En otros puntos estos comentarios iban dirigidos a modificar escalas ordinales de las diferentes características para que la definición de cada paciente fuera más clara. Después de la discusión con el comité tutorial se decidió modificar algunas escalas y redacción, además de no contemplar algunas otras. Los cambios que se hicieron fueron los siguientes.
- Se cambiaron las palabras como ocasionalmente, frecuentemente, ligeramente, etc. ya que se consideró lo suficientemente ambiguo para su correcta interpretación. En cuanto a la frecuencia de presentación, esta se graduó como "nunca", "pocas veces" "muchas veces" y "siempre". La intensidad se graduó como "nada", "poco" y "mucho" repitiendo en dos de los escenarios, ya que el termómetro tenía 4 pacientes con estado de gravedad progresiva. También se cambió el término de "algunos" por el término de "pocos" y el término "varios" por el término "muchos".

- En general, y también en relación con las escalas, se decidió hacer los escenarios lo más homogéneos posible en todos los dominios y no mezclar intensidad con frecuencia.
 - La palabra "desempleado" en el dominio de función laboral se decidió eliminar, ya que existía la posibilidad de que los pacientes no tuvieran la idea de desempleo.
 - La palabra "avergüenza" se eliminó por imprecisa. Se cambió por "incómoda".
 - Se cambió la palabra "enfermarse" por "infectar" principalmente en los escenarios 1 y 2 de los pacientes con SIDA.
 - En pacientes con diálisis, se propuso y modificó que los 4 escenarios hipotéticos tuvieran diálisis, y no contemplar a priori que la diálisis deteriora la calidad de vida.
 - Se consideró numerar los medicamentos, sin embargo este cambio no se realizó por la dificultad de llegar a un consenso entre expertos de cuántos medicamentos se consideraban pocos y muchos. Esto se dejó a la interpretación de los pacientes.
 - Se recomendó mayor brevedad en cada característica con el fin de simplificación, para lo cual se hicieron los cambios pertinentes.
3. Comentarios que proponían agregar dominios o aspectos clínicos propios del tipo de enfermedad que no se habían considerado. Los aspectos y dominios sugeridos por especialidad y las modificaciones realizadas fueron los siguientes:
- **Reumatología:**
 - **Vida familiar y apoyo de la pareja:** Se consideró incluido en el dominio de función social, donde se dice que puede o no puede disfrutar a sus familiares y/o amigos.
 - **Relaciones sexuales.** Aunque sí se consideró como un dominio muy importante en la evaluación de calidad de vida, y que en efecto se encuentra muy alterado en pacientes con lupus, se eliminó ya que es un dominio que los pacientes frecuentemente no contestan o se sienten agredidos cuando se les pregunta.
 - **Conocimiento de su enfermedad.** Este aspecto también se eliminó por la dificultad de expresarlo en forma graduada en los cuatro escenarios hipotéticos.
 - **Imagen corporal.** Sí se decidió incluir, en forma de aspecto general.

-
- **Actividades recreativas.** Se consideró un dominio importante; sin embargo, para simplificar el instrumento se decidió generalizarlo dentro del aspecto social.
 - **Esfuerzo físico:** Se consideró incluido en el aspecto de función física.
 - **Efecto en la imagen corporal por esteroides.** Esta característica se incluyó en el escenario hipotético en el aspecto de imagen corporal.
 - **Alteraciones en piel.** Se decidió si tomarlo en cuenta.
- **Nefrología:**
 - **Molestias por la diálisis:** No se incluyó en los escenarios de forma explícita.
 - **Esparcimiento.** Aplica el mismo comentario hecho a actividades recreativas en pacientes con lupus.
 - **Vida sexual.** Aplica el mismo comentario hecho a relaciones sexuales en pacientes con lupus.
 - **Encefalopatía.** Sí se decidió incluir.
- **Gastroenterología:**
 - **Limitaciones en la calidad de la vida por la dieta.** Aunque sí es un dominio importante, se decidió no considerarlo para simplificar el termómetro.
 - **Actitud general ante el pronóstico de la enfermedad.** Aunque puede ser importante, se consideró que era muy complicado para graduarlo de forma ordinal.
 - **Prurito:** No se consideró como los síntomas principales en frecuencia.
 - **Fatiga.** Este síntoma si se consideró como principal en este tipo de pacientes por lo que se incluyó.
- **Infectología:**
 - **Ansiedad.** Se incluyó dentro de las características emocionales comunes en todas las enfermedades.

- **Diaforesis.** No se incluyó ya que se consideró un síntoma muy agudo para incluirlo en el escenario hipotético.
- **Ataque al estado general,** se consideró incluido en la autopercepción de salud comentada en las características comunes de todas las enfermedades.
- **Disnea y tos.** No se incluyó, ya que se consideraron síntomas intermitentes, que por lo general no se presentan de forma crónica, y difícilmente ayuda a diferenciar a los pacientes en los escenarios clínicos.
- **Rechazo por familiares y amigos, y personal de salud.** No se incluyó por simplificar los escenarios hipotéticos.
- **Fatiga.** Sí se decidió incluir de forma explícita.

De esta forma, tomando en cuenta los dominios y aspectos incluidos en los escenarios originales, además de los recomendados por el grupo de expertos, se decidió incluir en cada enfermedad de 2 a 3 síntomas principales, y manejarlos de forma independiente, con el fin de simplificar los escenarios clínicos.

En los escenarios finales, se incluyeron los siguientes dominios comunes para todas las enfermedades: Función física, función social, estado emocional, función laboral, percepción general de salud y molestias físicas por el tratamiento.

Se decidió eliminar los aspectos relacionados a problemas económicos, ya que se consideró que era muy confuso y complejo, además por el tipo de pacientes tan homogéneos en estado socioeconómico que vemos en el INCMNSZ, este dominio tendría un poder discriminativo muy bajo.

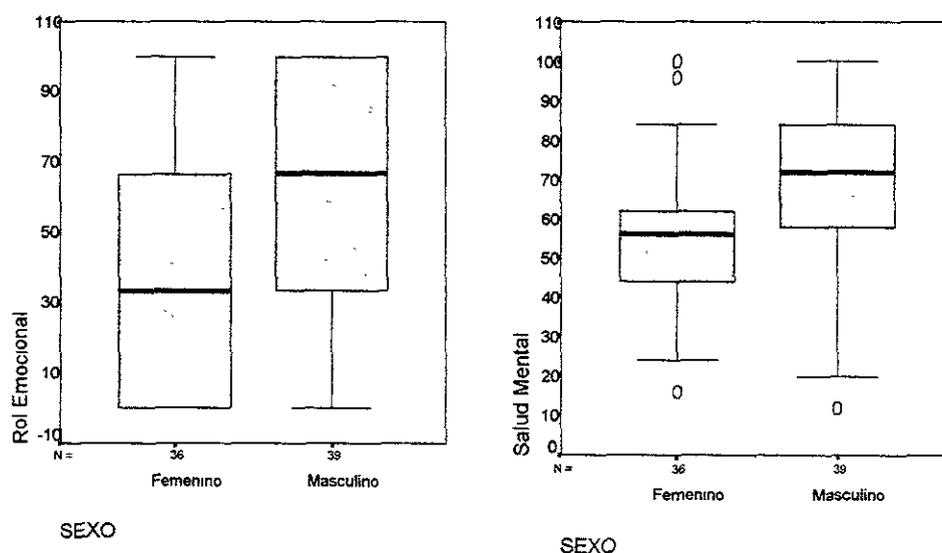
En cuanto al dominio de función somática se decidió incluir los siguientes síntomas principales para cada enfermedad: CH: encefalopatía, retención hídrica y fatiga. LEG: Dolores articulares, aspecto general y actividad renal. SIDA: Pérdida de peso, diarreas y aspecto general. DPCA: Encefalopatía, retención hídrica y fatiga.

Validez de constructo

Para la validez de constructo se utilizaron los supuestos teóricos mencionados en el apartado de material y métodos.

En cuanto a los supuestos generales, la edad no correlacionó con ninguno de los instrumentos de forma significativa. Comparando las calificaciones de calidad de vida entre ambos géneros se encontró asociación en el SF-36 en los dominios de rol emocional, salud mental, dolor y en la subescala de salud mental. Las medianas de calificación en estos dominios fueron más bajas en las mujeres específicamente: 33.3 vs. 66.67 ($p=0.004$) 56 vs 72 ($p=0.006$), 73 vs 84 ($p=0.040$) y 38.4 vs 49.9 ($p=0.015$) respectivamente. En la Figura 1 se grafican las diferencias más significativas

FIGURA 1. Distribución porcentilar por género de las calificaciones obtenidas en los dominios de Rol Emocional y Salud Mental del SF-36



Con el resto de los dominios e instrumentos no se encontraron diferencias significativas cuando se analizó género, y las demás variables sociodemográficas.

Continuando con los supuestos hipotéticos globales, cuando se comparó la calidad de vida con los tres instrumentos en pacientes con por lo menos una hospitalización en el último año ($n=36$) contra los que no habían tenido hospitalización en el último año ($n=39$), y por otro lado se comparó a aquellos con alguna consulta de urgencias ($n=50$) contra a aquellos sin consulta de urgencias en el último año ($n=25$), no se encontraron diferencias significativas.

En cuanto a la comorbilidad asociada, se comparó la calificación de calidad de vida en aquellos con alguna enfermedad importante agregada a la enfermedad principal (n=41) contra aquellos que sólo tenían la enfermedad principal (n=34). Se encontraron diferencias significativas en el dominio de función física, dolor, rol emocional y subescala de función física en el SF-36, y con la medida cruda sin los ajustes del termómetro. Las medianas de estos dominios e instrumentos para pacientes sin comorbilidad y con comorbilidad fueron: 65 vs. 50 ($p = .031$), 79 vs. 41 ($p = .026$), 66.6 vs. 33.3 ($p = .020$), 40.5 vs. 33.1 ($p = .024$), 80 vs. 65 ($p = .010$) respectivamente

Se probó la hipótesis de que la calidad de vida es mejor en los pacientes hospitalizados para tratamiento electivo que en aquellos a recibir tratamiento urgente o necesario. El grupo de pacientes internados para procedimientos diagnósticos fue muy heterogéneo, por lo que no se realizaron comparaciones especiales con este grupo. Los valores resumidos de los 8 dominios del SF-36, las tres mediciones del termómetro y la medición de TTO, se presentan en la Tabla 7 comparando entre las diferentes causas de hospitalización. Se agrega también el resultado de la escala visual análoga de auto evaluación del paciente de su estado de salud al ingreso

Comparando por lugar de procedencia, es decir aquellos pacientes que ingresaban desde Urgencias o Terapia Intensiva contra los que ingresaban directamente al sector de hospitalización, no se encontraron diferencias significativas en las calificaciones de calidad de vida

TABLA 7. Calidad de vida al ingreso por causa de hospitalización.

Instrumento de calidad de vida	Tratamiento Electivo n=16	Tratamiento Urgente n=34	Diagnóstico N=25	P*
TYO	0.95 (0.05-1)	0.70 (.05- 95)	0.65 (05-95)	0.033
EVA	1.50 (0-5)	6.25 (0-10)	4.00 (0-9)	<.0001
Termómetro				
Paciente	82.5 (50-100)	57.5 (10-97)	75 (21-98)	0.004
Ajuste 1	66.5 (40-85)	48 (0-91)	50 (15-95)	0.049
Ajuste 2	91.9 (50-100)	55.5 (0-100)	68.4 (18.6-100)	<0.0001
SF-36				
FF	52.5 (30-95)	50 (5-100)	65 (5-95)	NS
RF	25 (0-100)	0 (0-100)	25 (0-100)	NS
DC	73 (22-100)	41.5 (0-100)	62 (12-100)	NS
SG	45 (0-82)	28.5 (0-92)	45 (5-87)	0.067
VT	65(20-95)	35 (5-100)	45 (25-95)	0.004
FS	62.5 (12.5-100)	50 (0-100)	50 (12.5-100)	NS
RE	66.7 (0-100)	33.3 (0-100)	66.7 (0-100)	0.066
SM	72 (36-96)	52 (12-92)	64 (16-100)	0.035
SSF	37.7 (17.5-50.7)	31.2 (17.9-56.3)	39.7 (16.5-58.6)	NS
SSM	50.5 (32.3-58.8)	37.6 (16-59.4)	45.3 (28.1-63.2)	0.023

*Compara tratamiento electivo y de urgencias.

Los valores se expresan como mediana (mínimo -máximo)

NS = $p > 0.1$

Validaciones específicas por enfermedad.

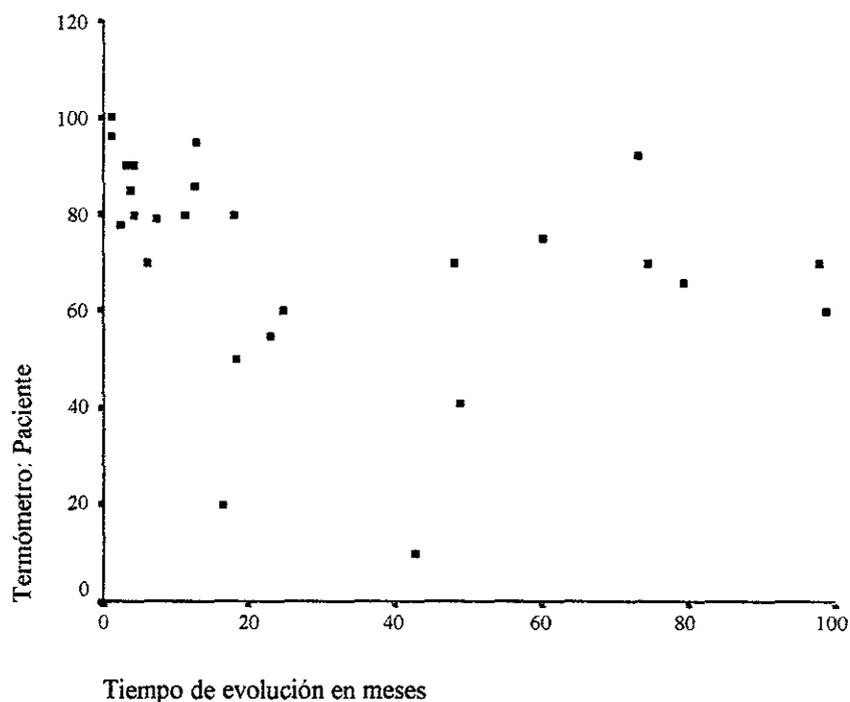
En el análisis de validación de constructo por enfermedad, se tomaron los parámetros clínicos mencionados en los métodos, y se evaluó la correlación de los instrumentos de calidad de vida.

Cirrosis Hepática

Tomando en cuenta los supuestos mencionados en material y método, el tiempo de evolución de la cirrosis, se asoció significativamente con la medida cruda del termómetro con una r de Spearman de -0.585 ($p = 0.001$). (Figura 2), y con el ajuste 2 del termómetro con una r Spearman -0.421 $p = .016$. También se asoció con un coeficiente r de .333 ($p=.048$) con el

dominio de dolor. Con el resto de dominios e instrumentos sus asociaciones fueron muy bajas y no significativas

FIGURA 2. Relación entre tiempo de evolución de la cirrosis hepática y calificación del termómetro



En cuanto a la gravedad de la insuficiencia hepática, calificada con la escala de Child-Pugh, el análisis se realizó tanto con el puntaje total, obteniendo índices de correlación no paramétricos, como con los 3 estadios reconocidos como A, B y C. En la Tabla 8 se presentan los resultados de los instrumentos de medición de calidad de vida, por estadio de Child Pugh. Se puede observar que sólo los parámetros del termómetro están asociados a las calificaciones de calidad de vida, correspondiendo peor calificación a los estadios más avanzados.

TABLA 8. Calidad de vida al ingreso por estadio de Child-Pugh

Instrumento de Calidad de Vida	Estadio de Child-Pugh			P*
	A (n=11)	B (n=9)	C (n=6)	
TTO	.95 (.05-.95)	.65 (.05-.95)	.50 (10-.85)	.080
Termómetro				
Paciente	80 (60-100)	70 (10-92)	60 (20-80)	.047
Ajuste 1	60 (15-95)	60 (10-91)	34 (10-65)	.065
Ajuste 2	83.3 (62.5-100)	66.6 (10-100)	55.6 (12.5-68.4)	.011
EVA	2. (0-5)	6 (1-9)	6 (1-8)	.007
SF-36				
FF	75 (30-95)	60 (10-90)	45 (15-100)	NS
RF	25 (0-100)	25 (0-100)	12 (0-75)	NS
DC	100 (22-100)	84 (32-100)	81 (0-100)	NS
SG	47 (15-87)	37 (0-82)	40 (5-42)	NS
VT	65 (35-95)	30 (5-80)	42.5 (15-80)	.060
FS	62.5 (37.5-87.5)	75(12.5-100)	56.25 (37.5- 75)	NS
RE	33.33 (0-100)	33.33 (0-100)	33.33 (0-100)	NS
SM	68 (36-100)	72 (24-100)	60 (16-96)	NS
SSF	45.2 (27.2-58.6)	39.2 (22.8-52.6)	31.1 (27.1-56.3)	NS
SSM	43.1 (28.9-57.3)	50.1 (20.7-63.2)	38.1 (28.1-62.8)	NS

* Prueba de Kruskal Wallis

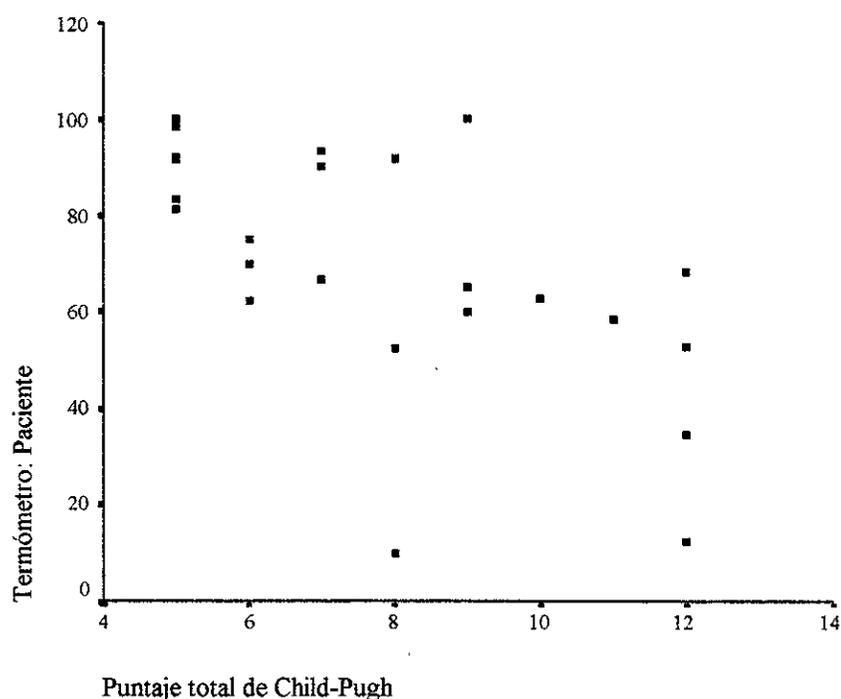
Los valores se expresan como mediana (mínimo-máximo)

El puntaje total de calidad de vida se correlacionó en forma estadísticamente significativa sólo con las medidas del termómetro dando coeficientes de correlación de Spearman entre -0.435 a -0.666. En la Figura 3 se ilustra la relación entre el puntaje total y la calificación de los pacientes con el método del termómetro. En el TTO y los dominios y subescalas del SF-36 no existieron correlaciones estadísticamente significativas.

Comparando a aquellos pacientes con historia de hipertensión portal hemorrágica (n=18) contra aquellos que no la tenían (n=8), no se encontraron diferencias en todos los instrumentos. En cuanto al estado nutricional, se encontraron correlaciones leves pero significativas con la subescala de salud mental, y los dominios de función social, rol emocional y salud mental (*rho de Spearman* de 0.48, 0.39, 0.35, y 0.35 respectivamente). Con el resto de los dominios e

instrumentos las correlaciones no fueron significativas. Por último, respecto al número de sangrados de tubo digestivo alto, tampoco se encontraron diferencias cuando se comparó el grupo de pacientes que habían tenido algún sangrado variceal en último año ($n = 16$) contra aquellos sin sangrado ($n = 10$).

FIGURA 3 Relación entre el puntaje total de Child-Pugh y calificación del termómetro



Lupus Eritematoso Generalizado

El tiempo de evolución de la enfermedad, definido como el momento en que se cumplieron tres criterios de la ARA para lupus, se asoció en forma estadísticamente significativa con los siguientes dominios del SF-36: Rol físico ($r = -0.612$), rol emocional ($r = -0.538$), y subescala de salud mental ($r = -0.481$), en forma marginal se asoció con la subescala de salud física y con los dominios de dolor, salud general y salud mental ($r = -0.342, -0.381, -0.324$ y -0.318 respectivamente). Con el resto de los dominios y de los instrumentos no se encontró relación.

Las correlaciones de los diferentes dominios e instrumentos de calidad de vida, con la escala de actividad lúpica (Mex Sledai), el índice de daño acumulado, y la dosis de PDN al ingreso se presentan en la Tabla 9. Las relaciones son negativas, ya que son escalas que tienen sentido invertido, a excepción de la EVA que tiene el mismo sentido de gravedad que el Mex Sledai.

TABLA 9. Correlación entre variables clínicas del lupus eritematoso generalizado y calidad de vida.

	Mex Sledai	Daño acumulado	Dosis de Prednisona
Mex Sledai	1.000		
Daño acumulado	.318	1.000	
Dosis Prednisona	.483	-	1.000
TTO	-	-	-
Termómetro			
Paciente	- .437	- .666	-
Ajuste 1	-	-.372	-
Ajuste 2	-.468	-.560	-
EVA	.436	.545	-
SF-36			
FF	-	-.401	-
RF	-	-	-
DC	-	-	-
SG	-	-	.336
VT	-	-.567	-
FS	-	-	-
RE	-	-	-
SM	-	-	-
SSF	-	-	-
SSM	-	-.337	.395
	-	-	-

Los valores corresponden a r de Spearman.
(-) correlación no significativa.

Se puede apreciar que el instrumento más asociado fue el termómetro, con un índice r por arriba de 0.5. Los dominios del SF-36 más asociados fueron los relacionados con la salud física más que emocional. También se realizó un análisis con la variable de actividad (Mex Sledai), categorizada en pacientes sin actividad (valor de 2 o menos) y con actividad (3 o más), sin encontrar diferencias significativas. (25)

Se realizó un análisis comparativo entre los pacientes que al momento del ingreso se encontraban sin drogas citotóxicas y dosis de Prednisona <15, y aquellos que tomaban dosis mayores de PDN y/o drogas citotóxicas sin tomar en cuenta la cloroquina. En este análisis no se encontraron diferencias significativas en las variables de medición de calidad de vida.

Diálisis Peritoneal Continua Ambulatoria

Se encontró correlación con r de Spearman de -0.512 ($p=.021$) entre el dominio de rol emocional del SF-36 y tiempo de evolución de la diálisis. Con el resto de los dominios e instrumentos no se encontró correlación significativa.

Los coeficientes de correlación (r de Spearman) entre los instrumentos de calidad de vida, y el funcionamiento dialítico, los valores de Hemoglobina y el IMC. se presentan en la Tabla 10.

TABLA 10. Correlación entre variables clínicas de la diálisis peritoneal continua ambulatoria y calidad de vida.

	Índice dialítico	Hemoglobina	IMC.
Índice dialítico	1.000		
Hemoglobina	-	1.000	
IMC.	-.501	-	1.000
TTO	-	-	-
Termómetro			
Paciente	-.474	-	-
Ajuste 1	-.485	-	-
Ajuste 2	-.416	-	-
EVA	.469)	-	-
SF-36			
FF	-	-	-
RF	-	-	-
DC	-.426	-	.454
SG	-.330	-	-
VT	-	-	-
FS	-	-	-
RE	-	-	-
SM	-	-	-
SSF	-	-	-
SSM	-	-	-

Los valores corresponden a r de Spearman.

(-) correlación no significativa.

TABLA 11. Correlación entre variables clínicas del síndrome de inmunodeficiencia adquirida y calidad de vida.

	Escala de Síntomas	Karnofsky	Cuenta de CD4	I.M.C.
Escala de Síntomas	1.000			
Karnofsky	-.785	1.000		
Cuenta de CD4	-.484	.567	1.000	
I.M.C.	-	.382	-	1.000
TTO	-	-	-	-
Termómetro				
Paciente	-.482	.407	-	-
Ajuste 1	-.466	.368	.374	-
Ajuste 2	-.485	.419	-	-
EVA	.458	-.463	-	-
SF-36				
FF	-.751	.848	.533	.416
RF	-	.384	.391	-
DC	-	-	-	-
SG	-.599	.495	-	-
VT	-.412	.514	-	-
FS	-.429	.657	.550	.467
RE	-	-	-	-
SM	-	.415	-	-
SSF	-.463	.534	.422	-
SSM	-	-	-	-

Los valores corresponden a r de Spearman.

(-) correlación no significativa

Como se aprecia en la tabla 11 en esta enfermedad a diferencia de las otras, se encontraron mejores correlaciones del SF-36 con las características clínicas específicas siendo esto más evidente en los dominios con componente físico que mental, con correlaciones tan altas como $r=0.848$ para la asociación de la escala de Karnofsky y el dominio de función física del SF-36. El termómetro también presentó adecuadas correlaciones con estos índices clínicos pero por lo general menores que las observadas con el SF-36.

En los pacientes con SIDA no se realizó análisis con variables categóricas, ya que el tamaño de la muestra tan reducido no lo permitió.

Correlaciones entre instrumentos

Por último como parte de la validez se realizó un análisis de correlación entre los instrumentos por medio de correlaciones no paramétricas, (*r de Spearman*) en los cuestionarios al momento del ingreso. En la tabla 12 se presenta la matriz de correlación entre instrumentos de calidad de vida además de la EVA referente al estado de salud.

TABLA 12. Matriz de correlación no paramétrica (Spearman) entre instrumentos.

	TTO	EVA	SF-36							Termómetro						
			SSF	SSM	FF	RF	DC	SG	VT	FS	RE	SM	Pac.	Aju 1	Aju 2	
TTO	1															
EVA	-.585	1														
SSF	400	-.509	1													
SSM	194	-.387	-	1												
FF	282	-.315	.679	-	1											
RF	.355	-.404	.693	.415	.335	1										
DC	.227	-.332	.680	.330	.224	.556	1									
SG	.420	-.553	.664	.364	.382	.435	.341	1								
VT	.422	-.612	.635	.594	.509	.596	.484	.556	1							
FS	223	-.416	.375	.693	-	.439	.398	.426	.562	1						
RE	-	-	-	.699	-	.437	.322	-	.282	.240	1					
SM	.251	-.445	.268	.869	.314	.406	.365	.407	.595	.577	.503	1				
Pac	.515	-.683	.518	.396	.383	.371	.384	.504	.549	.513	.170	.420	1			
Aju 1	.361	-.533	.379	.347	.225	.304	.276	.420	.409	.430	-	.354	.747	1		
Aju 2	.486	-.686	.526	.425	.328	.394	.400	.510	.559	.545	-	.435	.939	.811	1	

Negritas = Correlaciones superiores a 0.5. **Recuadros** = correlaciones entre instrumentos principales.

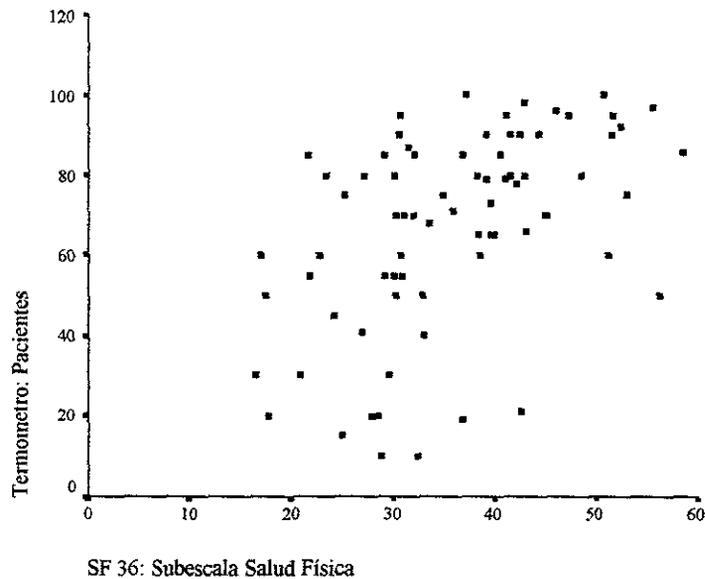
Pac. = Paciente no ajustada. Aju 1 = Ajuste 1. Aju 2 = Ajuste 2.

Todos los valores son *r* de Spearman. (-) = correlaciones no significativas $p > 0.05$.

Como se puede apreciar en la tabla 12, la gran mayoría de las correlaciones entre instrumentos y dominios fueron significativas y en el sentido esperado. El termómetro correlacionó mejor con la subescala de salud física del SF -36, que con la subescala de salud mental, teniendo correlaciones superiores a 0.5 y superiores también a las correlaciones presentadas anteriormente con los índices clínicos. El TTO aunque correlacionó adecuadamente con el termómetro, su correlación con el SF 36 mas bien fue baja, presentando coeficientes inferiores a 0.5. y en muchos de los dominios incluso por debajo de 0.3. En la figura 5 se presenta

de forma gráfica la correlación entre la medida del termómetro con la subescala de salud mental del SF-36 apoyando aun más lo comentado previamente.

FIGURA 5 Correlación entre la calificación del termómetro y la subescala de salud física del SF-36



Calidad de vida.

Las calificaciones basales del SF-36 presentaron distribuciones asimétricas, tanto en el grupo general como por diagnósticos, en todos los dominios no así en las dos subescalas agrupadores. En el caso de las calificaciones obtenidas por los instrumentos de utilidades en salud, las distribuciones a pesar de ser más amplias, fueron por lo general más simétricas que en el SF-36.

TABLA 13. Calificaciones basales del SF-36 por diagnóstico y grupo total

	CH	DPCA	LEG	SIDA	TOTAL
FF	65 (10-100)	43 (5-70)	60 (20-80)	78 (35-95)	60 (5-100)
RF	25 (0-100)	0 (0-100)	25 (0-100)	0 (0-100)	0 (0-100)
DC	84 (0-100)	62 (12-100)	41 (0-100)	41 (0-100)	62 (0-100)
SG	40 (0-87)	38 (0-82)	40 (5-82)	26 (0-92)	40 (0-92)
VT	45 (5-95)	43 (10-85)	50 (15-100)	48 (10-85)	45 (5-100)
FS	63 (12-100)	56 (0-100)	50 (0-100)	38 (0-100)	50 (0-100)
RE	33 (0-100)	67 (0-100)	33 (0-100)	67 (0-100)	33 (0-100)
SM	66 (16-100)	70 (24-88)	56 (16-100)	62 (12-92)	60 (12-100)
SSF	40 (23-59)	33 (17-43)	37 (17-52)	36 (22-56)	36 (17-59)
SSM	42 (21-63)	48 (25-58)	39 (17-61)	42 (16-56)	44 (16-63)

Los resultados se presentan como Mediana (min-max)

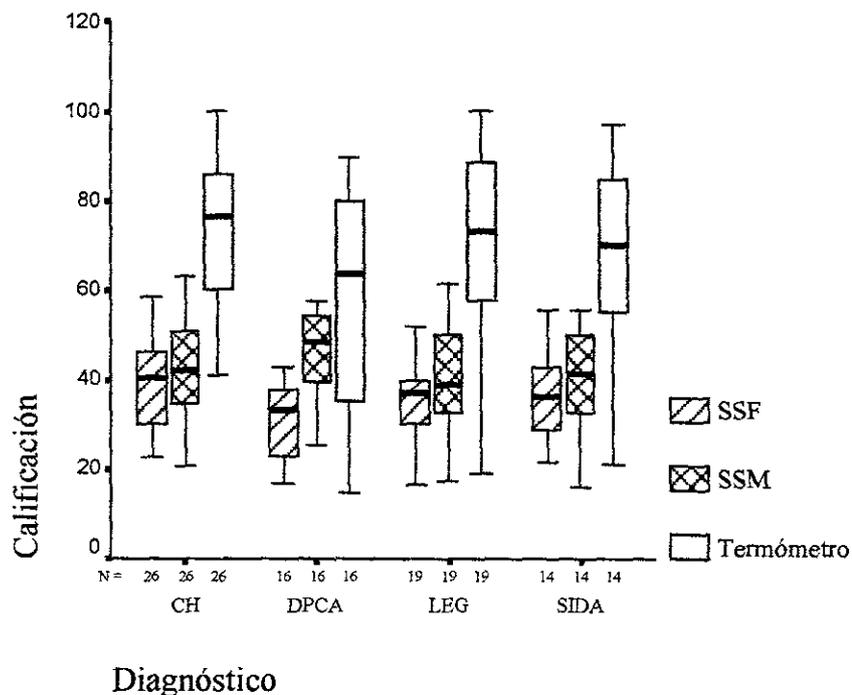
Las medianas de las calificaciones por dominios del SF-36 y diagnóstico se muestran en la tabla 13 y para los instrumentos de utilidades en salud en la tabla 14. Como se puede apreciar en estas tablas los grupos diagnósticos son comparables entre si y solo se encontró diferencia significativa en los dominios de función física y dolor corporal, ocasionado por la baja calificación relativa en el grupo de DPCA en el dominio de función física, y la calificación mas alta en grupo de CH para el dominio de dolor corporal. Otras diferencias que se observan, pero que no alcanzaron a ser significativas fueron la calificación relativamente baja en los dominios de salud general y función social en el grupo de SIDA, la calificación mas baja del termómetro en DPCA y la calificación mas baja con el TTO igualmente en el grupo de SIDA.

TABLA 14. Calificaciones basales del termómetro y TTO , por diagnóstico y grupo total

N	Diagnóstico				TOTAL
	CH	DPCA	LEG	SIDA	
PACIENTE	77 (10-100)	64 (15-90)	73 (19-100)	70 (10-97)	71 (10-100)
AJUSTE 1	53 (10-95)	58 (5-80)	50 (0-80)	53 (0-90)	55 (0-95)
AJUSTE 2	73 (10-100)	64 (8-100)	70 (0-113)	66 (0-102)	69 (0-113)
TTO	.78 (.05-.95)	.70 (.05-1)	.85 (.05-.95)	.20 (.05-.95)	.75 (.05-1)

Los resultados se presentan como Mediana (min-max)

FIGURA 7. Calificación de calidad de vida, por instrumentos y diagnósticos



Por otro lado también se puede apreciar que en todos los grupos diagnósticos el dominio más afectado es el de rol por problemas físicos, lo que puede estar reflejando el alto número de pacientes que ingresaron por problemas urgentes en la población estudiada. Además en todos diagnósticos la subescala de salud física es menor que la de salud mental compatible también con el tipo de enfermedades. En la figura 7 se puede apreciar la relación que guardan estas dos subescalas entre sí y con la calificación del termómetro que en todos los grupos tendió a calificar más alto a los pacientes aunque no es del todo comparable por tratarse de instrumentos distintos

En cuanto a los determinantes o factores que se asociaron de forma significativa con la calificación de calidad de vida al ingreso, ya se han comentado la gran mayoría de ellos en la sección de validación, tanto en el grupo general como por diagnóstico. En resumen las variables asociadas de forma univariada con los instrumentos de calidad de vida se presentan a continuación (del SF-36 solo se presentan las dos subescalas agrupadoras y del termómetro solo la medición del paciente y no los ajustes):

Subescala de salud mental:

- *Genero:* Las mujeres presentaron calificación más baja que los hombres.
- *Causa de la hospitalización:* Los que ingresaron para tratamiento electivo presentaron mejor calificación que los que ingresaron para tratamiento de urgencias.
- *Tiempo de evolución:* A mayor tiempo de evolución de la enfermedad peor calidad de vida en esta subescala, principalmente en los grupos de SIDA y LEG.

Subescala de salud física:

- *Comorbilidad:* La presencia de comorbilidad condicionó peor calificación en esta subescala.

Termómetro:

- *Comorbilidad:* La comorbilidad condicionó peor calificación en este instrumento.
- *Causa de la hospitalización:* Los que ingresaron para tratamiento electivo presentaron mejor calificación que los que ingresaron para tratamiento de urgencias.
- *Tiempo de evolución de la enfermedad:* A mayor tiempo de evolución de la enfermedad peor calidad de vida en este instrumento.

- *Índices clínicos de gravedad por enfermedad:* A mayor gravedad peor calidad de vida.

TTO:

- *Causa de la Hospitalización:* Los que ingresaron por tratamiento electivo presentaron mejor calificación que los que ingresaron en tratamiento de urgencias

Discusión.

Para empezar, es indispensable intentar contestar la pregunta de que si lo que se está midiendo con los instrumentos utilizados es en realidad calidad de vida. Debido a la gran complejidad en la definición de calidad de vida aún en la visión del concepto de calidad de vida relacionada al estado de salud, difícilmente llegaremos a contestar esta pregunta. Sin embargo, en un sentido más práctico, es importante que antes de cualquier conclusión, valoremos las propiedades clinimétricas de los instrumentos utilizados, principalmente su validez, reproducibilidad y sensibilidad al cambio, lo que ayudará a conocer mas a fondo lo que en realidad estamos midiendo. A continuación analizaremos las propiedades clinimétricas de los instrumentos que utilizamos en el presente estudio.

Reproducibilidad y consistencia interna de los instrumentos.

La reproducibilidad fue buena en todos los instrumentos si tomamos en cuenta que, tanto para el coeficiente de correlación intra-clase en la prueba repetida, como para el coeficiente α de Cronbach en la consistencia interna, se superaron en casi todos los instrumentos y escalas los valores recomendados superiores a 0.7 para poder hacer comparaciones entre poblaciones (39) Para las comparaciones clínicas individuales, la recomendación es que estos coeficientes sean mayores a 0.90, y esto solamente se logró con prueba repetida en el método de TTO y en tres dominios del SF 36 (función física, dolor corporal y salud general) y en consistencia interna solamente con el dominio de dolor corporal del SF-36.

Este hallazgo concuerda con lo señalado en la literatura principalmente en lo referente a la consistencia interna del SF 36 donde la mayoría de los estudios han encontrado estos coeficientes por arriba de 0.7, y por lo general con el mismo orden entre los dominios encontrados en el presente estudio. (27,30,33,40) Se ha referido que los peores dominios en cuanto a consistencia interna han sido función social y salud general, sin embargo en nuestro estudio el dominio menos reproducible fue el de rol emocional, con un valor inferior a 0.7. Este hallazgo es similar a lo encontrado por Zúñiga u col (33) en el estudio de población mexicana mas importante de validación de este instrumento,. Esta discrepancia se debe probablemente, a que este dominio requiere un mayor grado de interpretación por parte del paciente y esta influenciado por patrones culturales por lo que se requeriría de una mejor adaptación de la versión en español.

Los índices de reproducibilidad con prueba repetida, estuvieron en su mayoría superiores a los encontrados en la consistencia interna. Aunque éstas dos propiedades están estrechamente relacionadas, no miden lo mismo. Por un lado reproducibilidad esta midiendo la capacidad de un

instrumento de medir lo mismo en igualdad de condiciones y la consistencia interna se refiere a la correlación que guardan entre sí las preguntas de una misma escala. Por lo tanto, es necesario una adecuada consistencia interna para tener buena reproducibilidad y no necesariamente lo contrario. (41). En los estudios donde se han realizado las dos evaluaciones, se ha visto que la consistencia interna sobreestima la reproducibilidad con prueba repetida (30,42). La explicación de que en nuestro estudio las pruebas repetidas superaron a las de consistencia interna puede ser porque el tiempo entre las pruebas repetidas fue mas corto (24 horas) que lo realizado en otros estudios. Es posible que por esta razón tengamos en nuestros resultados un sesgo de memoria Será necesario evaluar esta propiedad en pacientes estables antes de afirmar que la reproducibilidad de este instrumento es buena.

La reproducibilidad en el termómetro y en el TTO resultó mas alta que lo encontrado en otros estudios aún en los realizados con población mexicana (33 38). Esto puede estar en relación con el sesgo de memoria ya comentado, principalmente en el método del TTO que comprende solamente una respuesta fácilmente recordable.

Validez de contenido:

La validación de contenido del termómetro fue buena, en cuanto a que la mayoría de los expertos estuvieron de acuerdo con los dominios evaluados, y que además se modificaron los escenarios a partir de sus sugerencias. Los comentarios y las razones de su modificación ya fueron ampliamente descritos en la sección de resultados.

Faltaron de evaluar algunas características que siendo importantes, no se incluyeron debido a la complejidad del instrumento o la complejidad de la expresión de estos atributos en el contexto de los escenarios hipotéticos. Algunos de estos aspectos fueron vida sexual, actividades recreativas, y preocupación por el pronóstico de su enfermedad. En relación con esta última característica, es probable que en pacientes con enfermedades crónicas, como es el caso de este estudio; esta sea la única característica de su enfermedad que repercute en su calidad, ya que muchos pacientes están conscientes de que su padecimiento conduce a un deterioro progresivo aun cuando en el momento de la evaluación no tenga limitaciones físicas o emocionales importantes.

Una limitante importante en la elaboración y validación del termómetro, es no haber incluido la opinión de los pacientes ya que el hacerlo hubiera implicado un estudio mas complejo y prolongado que excedía los propósitos de esta investigación . Es posible que si se incluyeran

estas opiniones , que reflejan las preferencias de los pacientes por ciertas características,, tendríamos una mejor validación en términos de calidad de vida.

Validez de constructo:

Los instrumentos de medición de calidad de vida presentaron asociaciones muy variables y en ocasiones poco consistentes con los parámetros sociodemográficos y clínicos tanto generales como específicos por enfermedad. En general el método del termómetro correlacionó de forma consistente con las variables clínicas tanto generales como específicas, y el SF-36 correlacionó mejor con las variables generales dada la característica de instrumento genérico.

En cuanto a los constructos hipotéticos generales, la edad no correlacionó con el SF-36 como se esperaba, ya que en la mayoría de los estudios se ha visto una asociación que entre mayor edad, mas afectados se encuentran los dominios relacionados con la función física. (30,43) Sin embargo este supuesto teórico tampoco se encontró en el estudio con población mexicana (33), lo que sugiere nuevamente que existen problemas en la traducción y definición del contenido. Una explicación alternativa a la falta de correlación con edad es que, la muestra estudiada está conformada por enfermedades crónicas que afectan pacientes de edades muy diferentes (CH vs SIDA o LEG) por lo que la relación entre estas variables podría perderse en el contexto de toda la muestra.

Por otro lado el SF-36 si se asoció con otras variables generales que al tener un componente más clínico reduce los sesgo mencionado previamente. Así se observó relación de comorbilidad con dominios de función física y de causa de la hospitalización con dominios de función emocional, lo que apoya en cierta medida su validez.

.Las correlaciones del SF-36 con los constructos específicos por enfermedad fueron muy bajas en los pocos casos que las hubo. Principalmente se encontró asociación con los índices que señalan, de alguna forma, daño crónico acumulado. Así, existió asociación con el índice de daño acumulado de LEG y con el Karnofsky y la escala de síntomas en SIDA, y no se encontró asociación alguna en las otras dos enfermedades (CH y DPCA), probablemente porque estas últimas además de no contar con buenas escalas clínicas que evalúen la repercusión crónica de la enfermedad, suelen cursar con un deterioro es más paulatino,. Los resultados encontrados en la literatura en este aspecto son contradictorios. Por ejemplo en LEG, que es la enfermedad mas evaluada en los aspectos clinimétricos y de calidad de vida, el SF-36 en algunos estudios tiene

una adecuada asociación con la actividad lúpica ($r=-0.38$ con FF) y el daño crónico ($r= 0.30$ con FF) (21) mientras que en otros dichas asociaciones no existen (44, 45).

A diferencia del SF-36, el termómetro correlacionó de la forma esperada con la mayoría de las variables clínicas específicas por enfermedad, y además con la presencia de comorbilidad y la causa de la hospitalización dentro de los constructos generales. Este tipo de correlaciones sugiere que este instrumento comparte características de instrumento genérico y específico, ya que aunque se asoció adecuadamente con las variables específicas por enfermedad presentó también la capacidad de discriminar según la causa de la hospitalización a pacientes con diferentes enfermedades. Por lo anterior considero que su utilización es válida para comparar entre diferentes enfermedades a pesar de que los escenarios hipotéticos tengan un componente específico para cada una de ellas. La consistencia de esta observación no ha sido probada ya que este instrumento se ha utilizado casi exclusivamente en análisis de costo utilidad, en enfermedades o intervenciones específicas.

En cuanto al método de TTO, sus correlaciones fueron aun más variables y prácticamente ninguna fue significativa. En el caso de la causa de hospitalización se encontraron diferencias significativas, sin embargo fue el único instrumento que en promedio calificó más bajo al procedimiento diagnóstico que a la causa urgente de hospitalización, situación que no tiene una explicación clínica. Desde el punto de vista teórico, este método, en contraste con el termómetro, mide mejor las preferencias del paciente Sin embargo el hecho de que esté basado en supuestos de duración de vida y de mejoría absoluta, hace compleja su comprensión, lo que explicaría la discordancia en prácticamente todos los supuestos hipotéticos.

Por otro lado y basándonos también en la inesperadamente alta reproducibilidad que presentó el método de TTO, es probable que el valor que se decide intercambiar de años de sobrevida por mejoría, sea un concepto definido por el paciente independientemente de su estado de salud. Este hallazgo no es del todo consistente en la literatura, ya que este método es de los más utilizados en los análisis de costo utilidad y su comportamiento clinimétrico ha sido bueno, comportándose en ocasiones mejor que el termómetro (13) No se encontró en la literatura el antecedente de la utilización de este instrumento en población mexicana, y ya que implica el concepto de intercambio, es probable que su interpretación se modifique ampliamente con diferentes patrones culturales. En este sentido será interesante realizar exploraciones más profundas a este tipo de instrumentos con el fin de que puedan ser utilizados en nuestra población.

Conclusiones

1. La reproducibilidad y consistencia interna de los instrumentos utilizados fueron buenas y permiten la comparación de grupos aunque no de individuos.
2. El instrumento del termómetro mostró una adecuada validez de contenido después de las modificaciones sugeridas por los expertos. Esta valoración quedó incompleta ya que no se contempló la opinión de los pacientes con las enfermedades estudiadas.
3. Los instrumentos de medición de calidad de vida presentaron asociaciones muy variables con los parámetros clínicos y sociodemográficos pero en general demostraron en varios aspectos su asociación con los constructos hipotéticos
4. . En general el método de termómetro correlacionó adecuadamente con las variables clínicas tanto generales como por enfermedad El SF-36 correlacionó mejor con las variables generales pero las correlaciones con los parámetros específicos por enfermedad fueron pobres y poco consistentes, situación esperable dadas la característica genéricas de este instrumento
5. En general los ajustes del termómetro no presentaron ventaja alguna en comparación con la medida cruda del paciente.
6. El TTO fue el menos válido probablemente por dificultades en su comprensión.
7. Debido a que el SF-36 es un método descriptivo no permitió interpretar las preferencias del paciente las cuales fueron mas fácilmente captadas por el termómetro
8. Por lo anterior concluimos que ambos instrumentos fueron útiles y complementarios ya que miden diferentes aspectos relacionados a la calidad de vida..

Bibliografía

1. Rutstein D, Berenberg W, Chalmers T, et al. Measuring the quality of medical care. *N Eng J Med*. 1976; 294:582-588.
2. Detsky AS, Naglie GA. Clinician's guide to cost-effectiveness analysis. *Ann Intern Med*. 1990;113: 147-154.
3. Vargas VF. ¿Cómo se evalúa el costo-beneficio de un tratamiento? En Ramiro R, Saita-Kamino O. Temas de medicina interna: Epidemiología clínica. México:Interamericana-McGraw-Hill. Vol. II, No2. 1994. págs:369-374.
4. Guyatt GH, Naylor D, Juniper E, et al. Users' guides to the medical literature. How to use articles about health-related quality of life. *JAMA* 1997; 277:1232-1237.
5. Leplège A, Hunt S. The problem of quality of life in medicine. *JAMA* 1997; 278:47-50.
6. Council of Medical Service. Quality of care. *JAMA*. 1986;256: 1032-1034.
7. Steffen G. Quality medical care. A definition. *JAMA* 1988; 260:56-61.
8. Cardiel HM. ¿Cómo se evalúa la calidad de vida? En Ramiro R, Saita-Kamino O. Temas de medicina interna: Epidemiología clínica. México:Interamericana-McGraw-Hill. Vol. II, No2. 1994. págs:359-368..
9. McSweeney J, Creer T. Health-related quality of life assessment in medical care. *DM* 1995;41:3-71.
10. Gill T, Feinstein A. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994;272:619-626.
11. Guyatt GH. A taxonomy of health status instruments. *J Rheumatol* 1995; 22: 1188-1190.
12. Guyatt GH, Feeney DH, Oatrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993; 118: 622-629.
13. Bakker C, Rutten M, van Doorslaer E, Bennett K, van der Lindern S. Health related utility measurement in rheumatology: An introduction. En: Bakker C. Patient-oriented outcome assessment in rheumatic diseases. Universitaire pers Maastricht 1995. págs:32-41.
14. Drummond M, Richardson S, Brien B, et al. Users' guide to the medical literature. XIII: How to use an article on economic analysis clinical practice. *JAMA* 1997; 277: 1552-1557.
15. Udvarhelyi S, Colditz GA, Rai A, Epstein A. Cost-effectiveness and cost-benefit analyses in the medical literature. Are the methods being used correctly? *Ann Intern Med* 1992; 116: 238-244.
16. Bellamy N, Boers M, Felson D, Fries J, et al. Health Status instruments: Utilities. *J Rheumatol* 1995; 22: 1203-1207.
17. Departamento de archivo clínico y bioestadística: Informe anual. Instituto Nacional de la Nutrición Salvador Zubirán, 1996.
18. Curran JW. Epidemiology of HIV infection and AIDS. En: Bennet JC, Plum F: Cecil textbook of medicine. 20th ed. U.S.A. Saunders.1996. págs: 1846-1851.
19. Tan EM, Cohen AS, Fries JF, et al. The 1982 revised criteria for the classification of SLE. *Arthritis Rheum* 1982;25:1271-1277.

20. Zonana-Nacach A, Camargo-Coronel A, Yañes P, et al. Measurement of damage in 210 Mexican patients with systemic lupus erythematosus: relationship with disease duration *Lupus* 1998 7:2 119-123
21. Stoll T, Gordon C, Seifert B, et al. Consistency and Validity of Patients Administered Assessment of Quality of Life by The MOS SF-36; Its association with Disease Activity and Damage in Patients with Systemic Lupus Erythematosus. *J. Rheumatol* 1997; 24: 1608-14.
22. Guzmán J, Cardiel MH, Arce-Salinas A, Sánchez-Guerrero J, Alarcón Segovia D. Measurement of disease activity in systemic lupus erythematosus. Prospective validation of 3 clinical indices. *J Rheumatol* 1992,19:1551-1558.
23. Wachtel T, Piette J, Mor V, et. al. Quality of life in patients with human immunodeficiency virus infection: Measurement by the Medical Outcome Study Instrument. *Ann Intern Med* 1993; 116: 129-137
24. Justice AC, Aiken LH, Smith LH, Turner JB. The role of functional status in predicting inpatient mortality with AIDS: A comparison with current predictors. *J Clin Epidemiol* 1996 49;2:193-201.
25. Wolcott DL, Nissenson AR. Quality of life in chronic dialysis patients: a critical comparison of continuous ambulatory peritoneal dialysis (CAPD) and hemodialysis. *Am J Kidney Dis* 1988,11:402-412.
26. McDowell Y, Newell C. Measuring health. A guide to rating scales and questionnaires. 2nd Ed. Oxford University Press 1996: págs: 446- 454.
27. Ware JE, Kosinski M, Keller SK. SF-36 physical and mental health summary scales: a user's manual. Boston Ma: The Health Institute 1994
28. Ware Jr JE, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA. *J Clin Epidemiol* 1998,51:945-952
29. McHorney CA, Ware JE, Lu JFR, et al. The MOS 36-item Short form Health Survey (SF 36): III. Test of data quality scaling assumption and reliability across diverse patient groups. *Medical Care* 1994; 32:40-66.
30. Brazier JE, Harper R, Jones NMB, et al. Validating the SF36 health survey questionnaire: New outcome measure for primary care. *Br Med J* 1992; 305: 160-164.
31. Carithers RL, Sugano D, Bayliss M. Health assessment for chronic HCV infection. Results of quality of life. *Dig Dis Sci* 1996. 41;12:75s-80s.
32. Kahan IH, Garrat AM, Kumar A, Cody DJ, et. al. Patients' perception of health on renal replacement therapy: Evaluation using a new instrument. *Nephrol Dial Tansplant* 1995. 10:5 684-689.
33. Zúñiga MA, Carillo-Jiménez GT, Fos PJ, Gandek B, Medina-Moreno MR. Evaluación del estado de salud con la Encuesta SF-36: resultados preliminares en México. *Salud Pub Mex* 1999,41:110-117
34. Díaz-Borjón E, Marcial-Barba LD, Cardiel MH., Hernández BE., Ruiz-Mercado R., Alcocer-Varela J. Changes in quality of life in rheumatic patients with joint replacement. Comparison of generic vs. specific instruments. *Arth Rheumat* 1995;38: 74
35. Hernández-Cruz B, Cardiel MH. Intra-observer reliability of commonly used outcome measures in rheumatoid arthritis. *Clin Exp Rheumatol* 1998,16:459-462

36. Hamel MB, Phillips RS, Davis RB, et al. Outcomes and cost-effectiveness of initiating dialysis and continuing aggressive care in seriously ill hospitalized adults. *Ann Intern Med* 1997;127:195-202.
37. Tsevat J, Solzan JG, Kuntz KM, et al. Health value of patients with HIV. Relationship to mental health and physical functioning. *Med Care* 1996;34:44-57.
38. Bakker C, Rutten M, Hidding A, van Doorslaer E, Bennett K, van der Lindern S. Patient utilities in ankylosing spondylitis and the association with other outcome measures. En: Bakker C. Patient-oriented outcome assessment in rheumatic diseases. Universitaire pers Maastricht 1995. págs:58-70.
39. Feinstein AR. *Clinimetrics*. Yale University Press. New Haven. 1987..págs: 167-190.
40. Alonso J, Prieto L, Ferrer M, Vilagut GL, Broquetas JM, Roca J, Batlle JS, Antó JM. Testing the measurement properties of the Spanish version of the SF-36 survey among male patients with chronic obstructive pulmonary disease. *J Clin Epidemiol* 1998;51:1087-1094
41. Streiner DL, Norman GR, *Health Measurement Scales : A practical guide to their development and use*. Oxford Medical Publications 2nd ed. 1995, págs: 104-126.
42. Stadnyk K, Calder J, Rockwood K. Testing the measurement properties of the Short Form-36 Health Suirvey in a frail elderly population. *J Clin Epidemiol* 1998;51:827-835.
43. Ware Jr. JE, Gandek B. Overview of the SF-36 Health Survey and the Intenational Quality of Life Assessment (IQOLA) project. *J Clin Epidemiol* 1998;51:903-902.
44. Hanly JG. Disease activity, cumulative damage and quality of life in systematic lupus erythematosus: results of a cross sectionl study. *Lupus* 1997;6:243-247.
45. Gladman DD, Urowitz MB Ong A, Gough J, MacKinnon A. A comparison of five health status instruments in patients with systemic lupus erythematosis (SLE). *Lupus* 1996;5:190-195.