

Primera validación en español del método INTERMED: Un sistema de temprana detección de problemas biopsicosociales y de consumo de servicios en pacientes médico-quirúrgicos.

E. Lobo*, M. Bellido*, R. Campos*, P. Saz**, F. Huyse***, P. de Jonge***, A. Lobo*

Resumen

Objetivos: *En el contexto de la alta prevalencia de pacientes médicos “complejos” y de la “complejidad de cuidados” en plantas médicas, INTERMED ha sido diseñado para la detección temprana de problemas biopsicosociales que puedan influenciar la pronta remisión y una planificación y tratamiento tempranos. El objetivo principal de este estudio es validar la versión española del método.*

Método:

Muestra: *Pacientes consecutivos, tanto ambulatorios como hospitalizados, incluidos en la lista de espera de Trasplante Hepático del Hospital Clínico Universitario de Zaragoza.*

Procedimiento: *a) Traducción y retrotraducción al castellano de INTERMED por investigadores expertos. Además, los investigadores examinaron el instrumento para asegurar que cumple estándares de validez “aparente” y de “contenido”.*

b) Estandarización de las enfermeras en el método (entrevista e interpretación de resultados). Dos enfermeras administraron la entrevista a cada paciente, de acuerdo al procedimiento estándar (revisión de la historia clínica, entrevista al paciente y al médico a cargo del mismo). Los roles de entrevistadora y observadora se alternaron, y se puntuó INTERMED ciegamente.

c) Para documentar fiabilidad inter-examinadores, se calcularon los siguientes coeficientes: Acuerdo Total (OA), Índice de Acuerdo de Ausencia de Síntomas (IAa), Índice de Acuerdo de Presencia de Síntomas (Iap), Acuerdo Específico (AE), Índice de Total Desacuerdo (TDI), Índice de Grave Desacuerdo (SDI), Kappa y Weighted Kappa (WKappa).

* Servicio de Psicología y Psiquiatría de Enlace. Hospital Clínico Universitario, Zaragoza.

** Departamento de Medicina y Psiquiatría. Universidad de Zaragoza.

*** Liaison Psychiatry Service. Free University Hospital, Amsterdam, Holanda.

Correspondencia: Dr. D. Antonio Lobo
Hospital Clínico Universitario Lozano Bleza
Avda. San Juan Bosco, 15
50009 Zaragoza

Resultados: Se entrevistó a 19 pacientes. Su edad media fue de 53,9 años, siendo la mayoría de clase social media y con estudios primarios. Los expertos investigadores concluyeron que la versión española de INTERMED cumple estándares mínimos de validez aparente y de contenido. El método es también aceptable para los pacientes, los médicos y las enfermeras investigadoras. La administración y su puntuación es factible, y su duración es de aproximadamente 15 minutos. Los coeficientes de acuerdo fueron en general satisfactorios ($\kappa > 0,6$), con puntuaciones óptimas para algunos ítems. Sin embargo, se documentaron valores κ bajos en varios ítems, en particular aquellos sobre pronóstico de las áreas psicológica, social y de atención sanitaria.

Conclusiones: La versión española de INTERMED es "factible" y cumple estándares de validez aparente y de contenido. Los coeficientes de fiabilidad inter-examinadores son en general satisfactorios. Estudios futuros deberían intentar documentar que es posible mejorar los coeficientes de fiabilidad en algunos ítems mediante la revisión del procedimiento de entrenamiento.

Palabras clave: Pacientes complejos. INTERMED. Detección temprana. Biopsicosocial.

Summary

First validation of the Spanish version of INTERMED: a method for the early detection of bio-psycho-social problems and in the use of services by medical patients.

Objective: In the context of the high prevalence of "complex" medical patients and "complexity of care" in medical wards, the INTERMED model has been designed for the early detection of bio-psycho-social problems that might influence an early referral and early treatment and discharge planning. The main objective of this paper is to validate the Spanish version of this model.

Methods:

Sample: Consecutive patients both, hospitalised and out-patients included in the waiting list for liver transplant in the University Hospital (Clínico) of Zaragoza, Spain.

Procedure: a) Translation and back-translation of INTERMED was previously completed by researchers with experience in this method. Furthermore, the researchers examined the instrument to assure it fulfils standards of face validity and content validity.

b) Trained nurses were then standardized in the use of the systematic interview and scoring system. Then, two nurses administered the interview to each patient, according to standard procedures (review of medical chart, interview of patients and physicians in charge). The nurses acted alternatively in the position of interviewer and observer and scored INTERMED blind to each other.

c) To document inter-rater reliability, the following coefficients were calculated: Overall Agreement (OA), Cohen's IA (a) and IA (p), Specific Agreement (SA), Kendell's Total Disagreement Index (TDI), Severe Disagreement Index (SDI) and both Cohen's Kappa and Weighted Kappa (W.Kappa).

Results: Nineteen patients were interviewed. Their mean age was 53.9 years and most were of middle class and had primary educational background. The experienced researchers concluded that the Spanish version of INTERMED fulfils standards of face and content validity. The method is also acceptable to patients, physicians and research nurses. It is feasible, the administration and scoring lasting 15 minutes approximately. Agreement coefficients were in general satisfactory ($\kappa > 0,6$) and optimal in some items. However, low kappas were documented in several items, in particular the prognostic items in psychological, social and use of services areas.

Conclusions: The Spanish version of INTERMED is feasible and fulfils standards of face and content validity. Inter-rater reliability coefficients are in general satisfactory. Future studies should try to document that reliability coefficients in some items improve by refining the training procedure.

Key words: Complex patients. INTERMED. Early detection. Bio-psycho-social.

INTRODUCCIÓN

Se ha documentado una alta prevalencia de problemas psico-sociales en pacientes médico-quirúrgicos, que incrementan: la “complejidad” de su manejo hospitalario, las estancias medias, y el consumo de servicios, tanto intra como extra hospitalarios(10, 14, 15, 19, 33). Así mismo, se ha hipotetizado que una temprana detección de los problemas podría mejorar la organización de la asistencia intra y extra hospitalaria; y que las consultas tempranas con los Servicios o agencias especializados podría evitar largas estancias y disminuir los costes médicos (7, 9, 26).

Nuestro grupo de trabajo lleva diez años estudiando el tema desde una perspectiva psico-social y epidemiológica, dentro del Grupo Europeo ECLW financiado por los programas BIOMED de la Unión Europea. Entre los datos que ha documentado(15, 23, 27) cabe destacar aquí que la remisión de pacientes médico quirúrgicos hospitalizados a las Unidades de Psicopatología y Psiquiatría de Enlace (UPPE), por problemática psico-social asociada, es en hospitales europeos del 1,4% (4% en el Hospital Clínico Universitario de Zaragoza). Esa cifra es muy inferior a la prevalencia de morbilidad psíquica en esos enfermos (30%, y 10% la de nivel más grave/complicado). Otro dato es que las UPPE están saturadas y no podrían atender mayor número de enfermos, y que los enfermos remitidos son en general muy “complejos” por la co-morbilidad somática y psíquica; y por la frecuencia de problemas sociales y de alto consumo de servicios. Sus estancias medias en hospitales europeos son 2,5 veces más altas que la media de estancia hospitalaria, en todos los países de la U.E. (14, 15).

El Grupo Europeo ECLW, en el que está integrado nuestro Grupo Investigador, ha diseñado y validado varios instrumentos. En primer lugar, el ARSI (Admisión Risk Screening Instrument) para temprana detección de los pacientes “complejos” (16). Otro instrumento es el COMPRI (Complexity Prediction Instrument), surgido del análisis factorial y reducción de los ítems del ARSI a un total de 13 ítems, de modo que se pueda pasar por enfermeras en cinco minutos(12), como se explica en otro artículo(13).

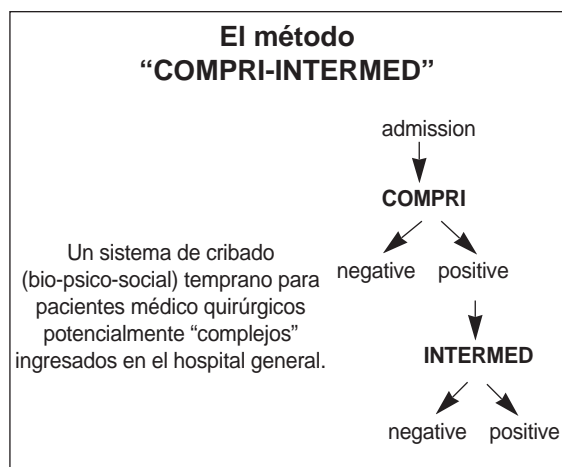


Figura 1

El método COMPRI-INTERMED consta de un primer cribado con el COMPRI seguido de la administración del INTERMED a los COMPRI + (potencialmente “complejos”) (8, 18, 21).

- Entrevista semi-estructurada, 15-20 min
- Enfermeras
 - Definiciones operativizadas
 - Cuantificación fiable*, visualización en colores
 - Antecedentes/Estado actual/Pronóstico
 - Ejes: Biológico, Psicopatológico, Social y Cuidados de Salud
 - “Validez”: inter-Servicios, trans-cultural**
 - Pretende facilitar:
 - Una temprana planificación del tratamiento interdisciplinar (y post-alta)
 - La comunicación interdisciplinar
 - Una temprana remisión a PE
 - Descripción del “case-mix”
 - Estudios de intervención
 - Potencial utilidad como instrumento de gestión

*La fiabilidad trata de demostrarse en este estudio en nuestro medio

**Esta validez en nuestro medio es también objetivo del estudio.

Figura 2
INTERMED

Por último, miembros del ECLW han diseñado la entrevista INTERMED que, como indica la Fig. 2 y se explica en otro artículo (13, 18, 21), puede administrar en quince minutos una enfermera para confirmar de modo fiable la existencia de “complejidad” (32). Esta entrevista pretende facilitar una temprana planificación de la asisten-

cia. La presentación en símbolos de acuerdo con la gravedad/“complejidad” [○ indica la puntuación 3 más grave; seguido del ❖ (2), □ (1) y ◆ (0)], en el impreso que figura en la primera página de la historia clínica, sirve para visualizar de inmediato la existencia de problemas en alguna de las áreas (Biológica, Psicológica, Social y Consumo de Servicios) que recomiendan una temprana organización de la asistencia y posiblemente de consultas con distintos servicios o agencias psico-sociales. La Fig. 3 muestra los resultados de la administración del INTERMED en un paciente “complejo” en nuestro medio hospitalario en Zaragoza.

	HISTORY	CURRENT STATE	PROGNOSIS
Biological	Chronicity m Complexity n	Severity v Complexity u	n
Psychological	Coping n Disfunctioning m	Compliance n Symptoms n	n
Social	Integration n Disfunctioning n	Living situation n Network n	n
Health care	Intensity n Experience n	Complexity u Referral n	v

INTER

		M
		E
		D

Figura 3

Los símbolos representan con el grado de “complejidad”. La puntuación más baja corresponde al símbolo ◆, las siguientes son el símbolo ■ (puntuación=1) y el ❖ (puntuación=2), hasta llegar al ○ (puntuación=3) que representa el grado más alto de “complejidad”.

El método COMPRI-INTERMED, en cuyo diseño y primeros estudios ha intervenido nuestro grupo investigador, no ha sido utilizado todavía en España. Una serie de trabajos en la bibliografía internacional avalan la utilidad del instrumento en varios trabajos realizados, principalmente en Holanda y en Suiza (8, 13, 18, 21). Consecuentemente, parece oportuno ahora proceder a una primera fase en la validación completa de INTERMED en nuestro medio, y aplicado por enfermeras estandarizadas. Se hipotetiza que el método es factible también en nuestro medio por

su fácil administración; que cumple estándares mínimos de validez y que tiene una adecuada fiabilidad inter-examinadores.

MATERIAL Y MÉTODOS

1. Muestra: Se incluyeron en el estudio todos los pacientes consecutivos, tanto ingresados como ambulatorios, vistos en el Hospital Clínico Universitario de Zaragoza que habían sido incorporados a la lista de espera de Trasplante Hepático en el período de Enero a Noviembre (ambos incluidos) de 2002.

2. Instrumentos

Entrevista INTERMED: Es un instrumento diseñado por F.J. Huyse y colabs.(13, 17) y desarrollado para mejorar la comunicación interdisciplinar entre especialistas (para-) médicos y proporcionar un método que describa complejidad de caso con fines clínicos, científicos y educativos. Este método se lleva a cabo mediante una entrevista con el paciente que sintetiza información de cuatro áreas: 1) biológica, 2) psicológica, 3) social y 4) consumo de servicios, valoradas en el contexto de tiempo (historia, estado actual y pronóstico). En cada uno de los apartados resultantes, dos variables pertinentes fueron elegidas en términos clínicos y/o científicos reconocidos como importantes para cada una de ellas, representando un indicador para necesidades pasadas, presentes o futuras (en dicho apartado). Las áreas, apartados y variables se resumen así:

BIOLÓGICA:

-Historia biológica, con dos variables; Cronicidad y Dilema diagnóstico.

-Estado biológico actual, con dos variables; Severidad de los síntomas y Complejidad diagnóstica.

-Pronóstico biológico.

PSICOLÓGICA:

-Antecedentes psicológicos, con dos variables: Afrontamiento y Disfunción psiquiátrica.

-Estado psicológico actual, con dos variables: Resistencia al tratamiento y Síntomas psiquiátricos.

-Pronóstico psicológico.

SOCIAL:

-Antecedentes sociales, con dos variables: Integración social y Disfunción social.

-Estado social actual, con dos variables: Situación de vivienda y Redes sociales.

-Pronóstico social.

ATENCIÓN SANITARIA:

-Antecedentes de atención sanitaria, con dos variables: Intensidad de tratamientos y Experiencias en tratamientos.

-Situación actual de atención sanitaria, con dos variables: Organización de los cuidados y Adecuación del ingreso o remisión.

-Pronóstico de atención sanitaria.

Cada una de las variables de los diferentes apartados han de ser puntuadas de acuerdo a un sistema definido, en un rango que va del 0 (no vulnerabilidad o no necesidad) al 3 (alta vulnerabilidad o necesidad). INTERMED está diseñado para que lo realice un "lay interviewer" estandarizado (entrevistando de entre 5 a 10 pacientes), en aproximadamente quince minutos. Ha sido validado originalmente en Holanda por Huyse y colabs.(13), habiéndose documentado buenos coeficientes de fiabilidad (interexaminadores), así como aceptables coeficientes de validez.

En la versión española, la entrevista que se le realiza al paciente es la fundamental; se ve complementada por unas breves preguntas al médico a cargo del mismo, que ofrece información casi exclusivamente biológica sobre gravedad de la patología y pronóstico. Además se examina muy brevemente la historia clínica para obtener información sobre los antecedentes biológicos. El "lay interviewer" puede ser, como en este caso español, una enfermera cualificada y estandarizada en INTERMED.

3. Procedimiento: Previamente al estudio, nuestro grupo investigador realizó la traducción y retrotraducción al castellano del instrumento, así como el examen por expertos de la disciplina. Para el estudio de campo, las enfermeras del estudio, previamente estandarizadas procedieron del modo siguiente con cada paciente de la muestra:

a. Se revisó la historia clínica, y la información que se obtuvo se valoró dentro del área "Biológica".

b. Cada paciente fue entrevistado por dos enfermeras estandarizadas; una de ellas actuaba como entrevistadora y la otra como observadora. Ambos

roles, de entrevistadora y observadora, se alternaron de tal manera que cada una de ellas realizase un número similar de entrevistas en cada rol.

c. En cada caso, la misma enfermera que ejercía de entrevistadora hacía las preguntas al médico especialista a cargo del enfermo (el hepatólogo correspondiente).

d. En base a la información obtenida puntuaron, por separado, los parámetros de INTERMED con un procedimiento "ciego".

Para el análisis estadístico se utilizaron los siguientes coeficientes de acuerdo, con un método informatizado original de este grupo investigador: *Overall agreement (OA)*, *Índice de Acuerdo de Ausencia de Síntomas (IAa)*, *Índice de Acuerdo de Presencia de Síntomas (Iap)*, *Acuerdo Específico (AE)*, *Total disagreement index (TDI)*, *Severe disagreement index (SDI)*, *Kappa* y *weighted Kappa (WKappa)*.

RESULTADOS

La muestra que se obtuvo fue de 19 pacientes, 14 hombres (73%) y 5 mujeres (27%), y la edad media de la muestra global fue de 53,9 años. Se trata de pacientes predominantemente de clase social media y con estudios primarios.

Se calculó el acuerdo inter-examinadores para cada variable con los coeficientes de acuerdo citados y que figuran en las Tablas 1 a 4.

La Tabla 1 se refiere al cálculo de coeficientes de acuerdo en los parámetros biológicos. Se observa que el acuerdo es óptimo en algunos items, como *Cronicidad*, y más bajo en items como *Dilema diagnóstico* o *Severidad de los síntomas*. Como era de esperar, los coeficientes kappa son más bajos que los índices brutos de acuerdo como el OA, IA etc.

La Tabla 2 muestra el acuerdo en los parámetros psicológicos, donde encontramos el coeficiente de acuerdo más bajo, esto es, *Pronóstico psicológico* con un valor kappa de 0,063. Otro coeficiente también bajo de este área es el de *Resistencia al tratamiento* con kappa de 0,477; mientras que el resto de variables tienen valores kappa por encima de 0,8, muy por encima de lo que se estima como acuerdo inter-examinadores.

En la tabla 3, referente al área social, encon-

Tabla 1
Coefficiente de acuerdo inter-examinadores en la valoración de parámetros “biológicos” con el método INTERMED

	OA	IA(a)	IA(p)	AE	TDI	SDI	Kappa	Wkappa
Cronicidad	1.0000	0.0000	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000
Dilema diagnóstico	0.8421	0.8000	0.5714	0.5714	0.4286	0.4286	0.639	0.692
Severidad de los síntomas	0.8421	0.0000	1.0000	0.8421	0.1579	0.0000	0.587	0.596
Complejidad diagnóstica	0.8421	0.8235	0.4000	0.4000	0.6000	0.6000	0.500	0.607
Pronóstico biológico	0.8947	0.0000	1.0000	0.8947	0.1053	0.0000	0.801	0.810

Tabla 2
Coefficiente de acuerdo inter-examinadores en la valoración de parámetros “psicológicos” con el método INTERMED.

	OA	IA(a)	IA(p)	AE	TDI	SDI	Kappa	Wkappa
Afrontamiento	0.8947	0.9000	0.9000	0.8000	0.2000	0.1000	0.842	0.910
Disfunción psiquiátrica	0.8421	0.8750	0.9167	0.7500	0.2500	0.0833	0.762	0.828
Resistencia al tratamiento	0.8421	0.8235	0.4000	0.4000	0.6000	0.6000	0.477	0.477
Síntomas psiquiátricos	0.8421	0.7778	0.8333	0.7500	0.2500	0.1667	0.754	0.838
Pronóstico psicológico	0.4737	0.0000	0.8947	0.4737	0.5263	0.1053	0.064	0.063

Tabla 3
Coefficiente de acuerdo inter-examinadores en la valoración de parámetros “sociales” con el método INTERMED.

	OA	IA(a)	IA(p)	AE	TDI	SDI	Kappa	Wkappa
Integración social	0.8421	0.8000	0.5714	0.5714	0.4286	0.4286	0.648	0.698
Disfunción social	0.8947	0.8750	0.6000	0.6000	0.4000	0.4000	0.703	0.759
Situación de vivienda	0.7895	0.7647	0.3333	0.3333	0.6667	0.6667	0.433	0.475
Redes sociales	0.6316	0.6000	0.6923	0.4615	0.5385	0.3077	0.398	0.453
Pronóstico social	0.6842	0.6154	0.5455	0.4545	0.5455	0.4545	0.400	0.398

tramos que son tres las variables que puntúan con valores kappa inferiores a 0,5: *Situación de vivienda*, *Redes sociales* y *Pronóstico social*. Los primeros parámetros, por el contrario, son satisfactorios.

Finalmente, la Tabla 4 representa el área de atención sanitaria, donde todas las variables presentan valores kappa por encima de 0,5. Además,

la variable *Adecuación del ingreso* muestra acuerdo total (kappa=1).

Por tanto, los parámetros biológicos y de atención sanitaria tienen los coeficientes de acuerdo más altos que los parámetros psicológicos y sociales.

DISCUSIÓN

Los resultados que se presentan en este estu-

Tabla 4
Coefficiente de acuerdo inter-examinadores en la valoración de parámetros sobre el “consumo de servicios” con el método INTERMED

	OA	IA(a)	IA(p)	AE	TDI	SDI	Kappa	Wkappa
Instensidad del tratamiento	0.7895	1.0000	1.0000	0.7778	0.2222	0.0000	0.578	0.578
Experiencias en tratamientos	0.9474	0.9444	0.5000	0.5000	0.5000	0.5000	0.642	0.642
Organización de los cuidados	0.7895	0.0000	1.0000	0.7895	0.2105	0.0000	0.664	0.721
Adecuación de ingreso	1.0000	1.0000	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000
Pronóstico de cuidados sanitarios	0.6842	0.0000	1.0000	0.6842	0.3158	0.0000	0.470	0.520

dio, el primero de su género que se realiza en nuestro país, apoyan el criterio de que el INTERMED es un método válido para ser utilizado en pacientes en nuestro país y en idioma castellano, y para describir su situación clínica en una serie de ejes que incluyen distintas variables en las áreas generales Biológica, Psicológica, Social y de Atención sanitaria.

Instrumentos de este género deberían cumplir, por lo menos, estándares mínimos de validez, es decir, validez aparente y validez de contenido. Aunque se estima que dichos estándares son discutibles (22, 28) pensamos que pueden defenderse en este caso, por los siguientes motivos:

La validez aparente se refiere al juicio de que un nuevo instrumento tiene sentido para un investigador y, consecuentemente, incrementa la aceptabilidad de los clínicos, incluyendo aquí a enfermeras. En apoyo de este tema, investigadores europeos de varios países han considerado que INTERMED es factible y aceptable, y ahora lo corroboran autores españoles. Aunque estas características no han sido formalmente verificadas y aunque inicialmente surgió alguna crítica en relación a la complejidad del instrumento, enfermeras sin un entrenamiento especial y simplemente estandarizadas en el método pudieron aplicarlo en un promedio de quince minutos.

La validez de contenido se refiere al examen sistemático de un nuevo instrumento por expertos en la disciplina, para asegurarse de que los

items cubren el tipo de información que sería necesaria para ser interpretada de modo correcto (28). Podía surgir la duda de si un registro de este tipo cubre por completo una descripción de la situación del paciente ingresado en un hospital general o ambulatorio; puesto que el instrumento ha sido examinado por una serie de expertos de distintos países europeos, incluyendo algún experto norteamericano, y ha seguido un cuidadoso desarrollo tras varios años, y ahora lo valoran expertos en nuestro país, se estima que la versión final cumple los criterios buscados.

El propósito principal del estudio ha sido documentar la fiabilidad inter-examinadores del instrumento en su versión española y en un entorno clínico concreto. Los resultados son en general satisfactorios, según criterios actuales (1, 30), como demuestran los índices de acuerdo en la Tabla 2, con valores kappa >0,6, incluso llegando al acuerdo total en dos de las variables (Cronicidad y Adecuación del ingreso). Estos buenos resultados prevalecen en la mayoría de secciones del instrumento.

Sin embargo, las siguientes variables fueron identificadas por tener un valor bajo (coeficiente Kappa <0,6): Severidad de los síntomas, Complejidad diagnóstica, Resistencia al tratamiento, Situación de vivienda, Redes sociales, Intensidad del tratamiento, Pronóstico psicológico, Pronóstico social y Pronóstico de cuidados sanitarios. Con respecto al índice de acuerdo en el

Pronóstico biológico, los autores originales en Holanda tienen la misma impresión: se trata de una variable muy difícil de predecir y por tanto no extraña que los valores globalmente considerados sean bajos. También son bajos el resto de indicadores pronósticos, y cualquier clínico experimentado conoce las dificultades de emitir un juicio en estas direcciones; por tanto, aunque en individuos concretos puede ser un valor de gran poder para la organización de servicios en relación con el juicio que se emite, globalmente considerados, estos ítems deben ser tomados con cautela hasta que en próximas ediciones y estudios de campo se intente mejorar la fiabilidad.

Con respecto a los valores conseguidos en el resto de ítems con bajos coeficientes kappa, también el investigador se encuentra ante la evaluación de aspectos que necesitan una fuerte inferencia estimativa. Es bien conocido en entrevistas psiquiátricas que la evaluación de signos y síntomas psicopatológicos en que hay que hacer este tipo de estimación resulta la más difícil en cuanto a conseguir una buena fiabilidad. Consecuentemente, también en estos ítems las futuras ediciones del instrumento y el entrenamiento y estandarización de los investigadores debe intentar afinar, para permitir que una mejor operativización consiga también mejores resultados. Entre tanto, habrá que utilizar con cautela los resultados que provienen de este tipo de variables.

En relación a posibles limitaciones del estudio de validación, debe considerarse el diseño y estadística utilizada. La mayoría de estudios que comunican coeficientes de acuerdo en entrevistas de tipo psiquiátrico o social, lo que será aplicable también a Enfermería, se refieren a coeficientes de fiabilidad (23). Autores como Tinsley y Weiss (34), han argumentado que la fiabilidad representa “el grado en que las puntuaciones de distintos jueces son proporcionales cuando se expresan como desviaciones del promedio” y han enfatizado la importancia de describir el modo en que se ha calculado dicho acuerdo. La medida de acuerdo es difícil, sobre todo cuando las variables de interés tienen más de dos posibles valores. Es difícil argumentar que los resultados de este estudio son exageradamente optimistas, pues no sólo se calcula el acuerdo global, que puede

ser inflado por “chance agreement” o “acuerdo por casualidad”. Por el contrario, existe un amplio consenso en relación con las ventajas de utilización del coeficiente kappa, más exigente (2-4, 20, 24, 29, 31). Nosotros, como en otros estudios recientes, hemos comunicado ambos tipos de coeficientes. Dewey (6), como otros autores, ha argumentado que el coeficiente kappa es la medida de elección cuando se comparan con nuevas mediciones tales como el coeficiente de Maxwell de error aleatorio. Sin embargo, las limitaciones del coeficiente kappa también han sido comentadas; su principal debilidad es que este coeficiente mide la frecuencia de acuerdo exacto y no el grado de acuerdo aproximado y, por tanto, podría ser arbitrario, sobre todo cuando se agrupan datos continuos en datos ordinales. En este tipo de situaciones podría ser importante la utilización de coeficientes intraclass de correlación (24), pero nosotros hemos utilizado además el índice kappa ponderado.

Pueden surgir también dificultades en relación con kappa en situaciones de asimetría marginal (4,5). Estos coeficientes se consideran demasiado restrictivos cuando la distribución de las categorías es muy asimétrica; incluso altos porcentajes de acuerdo pueden dar como resultado bajos o negativos coeficientes kappa (11). Además, para los ítems en que las características a observar son muy claras, el índice kappa restrictivo no sería una medida adecuada de acuerdo (25). Por tanto, nosotros hemos utilizado además de este coeficiente, los más “crudos”, como el acuerdo global (Overall agreement OA), o los coeficientes IA, y, consecuentemente, estimamos que el conjunto de todos los coeficientes calculados da mayor solidez a las conclusiones.

Naturalmente, la forma más importante de validez o “validez de constructo” (28) es difícil de documentar. Una posibilidad hubiese sido el intento de hacer una validación concurrente con instrumentos ya acreditados. Sin embargo, dada la novedad del procedimiento, no existe ningún instrumento similar con el que podría ser comparado. Entre tanto, es una práctica común apoyarse en estándares mínimos de validez, tales como los de fiabilidad que hemos utilizado en este estudio. Además, está en marcha un procedimiento de validación en muestras más amplias y

distintos entornos, y teniendo en cuenta variables clínicas y epidemiológicas que serán oportunamente comunicadas.

BIBLIOGRAFÍA

1. **Bartko J, Carpenter W.:** On the methods and theory of reliability. *Journal of Nervous and Mental Diseases* 1976; 163: 307-317.
2. **Bech P, Malt UF, Dencker SJ, Ahlfors UG, Elgen K, Lewander T, Lundell A, Simpson GM, Lingjaerde O.:** Scales for assessment of diagnoses and severity of mental disorders. *Acta Psychiatr Scand* 1993; 87: Suppl.372.
3. **Berry KJ, Mielke PW.:** A generalization of Cohen's kappa agreement measure to interval measurement and multiple raters. *Educ Psychol Measur* 1988; 48: 921-933.
4. **Cicchetti DV, Feinstein AR.:** High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990; 43: 551-558.
5. **Collis GM.:** Kappa, measures of marginal symmetry and intraclass correlations. *Educ Psychol Measur* 1985; 45: 55-62.
6. **Dewey ME.:** Coefficients of agreement. *Br J Psychiatry* 1983; 143: 487-489.
7. **De Jonge P, Huyse FJ, Herzog T, Lobo A, Slaets JPJ, Lyons JS, Opmeer BC, Stein B, Arolt V, Balogh N, Cardoso G, Fink P, Rigatelli M.:** Risk Factors for Complex Care Needs in General Medical Inpatients: Results From a European Study. *Psychosomatics* 2001; 42: 213-221.
8. **De Jonge P, Huyse FJ, Slaets JPJ, Lyons JS, Lobo A, Herzog T, Opmeer BC, Stein B, Fink P, Gans ROB, Van Schijndel RS, Cardoso G, Rigatelli M, Creed F, House A, Balogh N.:** Complexity of care as a new concept for describing health service needs: a confirmatory factor analytic approach.
9. **De Jonge P, Huyse FJ, Ruinemans, GMF, et al.:** The timing of psychiatric consultation: the impact of social vulnerability and level of psychiatric dysfunction. *Psychosomatics* 2000; 41: 505-511.
10. **De Jonge P.:** Detection of complex patients in the general hospital: from psychiatric comorbidity to care complexity. Amsterdam, Thela Thesis 1999.
11. **Feinstein AR, Cicchetti DV.:** High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990; 43: 543-549.
12. **Huyse J, De Jonge P, Slaets JPJ, Herzog T, Lobo A, Lyons JS, Opmeer BC, Stein B, Arolt V, Balogh N, Cardoso G, Fink P, Rigatelli M.:** COMPRI-An Instrument to Detect Patients With Complex Care Needs. Results From a European Study. *Psychosomatics* 2001; 42: 222-228.
13. **Huyse FJ, De Jonge P, Slaets JPJ, Stiefel F, Söllner W, Latour CHM.:** Un nuevo "rol" para la Psiquiatría de Enlace: desde los Servicios ad-hoc a una asistencia integrada". *C. Med. Psicosom.* En prensa.
14. **Huyse FJ, Herzog T, Lobo A, et al.:** Detection and treatment of mental disorders in general health care. *European Psychiatry* 1997; 12: 70-78.
15. **Huyse FJ, Herzog T, Malt UF, Lobo A & ECLW.:** The European Consultation Liaison Workgroup (ECLW) Collaborative Study. I: General Outline. *Gen. Hosp. Psychiatry* 1996; 18(1): 44-55.
16. **Huyse FJ, Herzog T, Malt UF, et al.:** A screening instrument for the detection of psychosocial risk factors in patients admitted to general hospital wards, in Biomedical and Health Research, The Biomed 1 Programme, edited by Baert A. Amsterdam, IOS Press 1995.
17. **Huyse FJ, Lyons JS, Stiefel FC, Slaets JPJ, Lobo A, Guex P, De Jonge P.:** Intermed: A Method To Assess Health Service Needs: Development And First Results On Its Reliability. *Gen Hosp Psychiatry* 1999; 21: 39-48.
18. **Huyse FJ, Lyons JS, Stiefel FC, et al.:** Operationalizing the biopsychosocial model: the INTERMED. *Psychosomatics* 2001; 42: 1-9.
19. **Katon WJ, Von Korff M, Lin E, et al.:** High utilizers of medical care, DSM-III-R diagnoses and treatment needs. *Gen. Hosp. Psychiatry* 1990; 12: 355-362.
20. **Kendell RE.:** Psychiatric diagnosis: A study of how they are made. *Br J Psychiatry* 1973; 122: 437-445.
21. **Koch N, Stiefel F, De Jonge P, Fransen J, Chamot AM, Gerster JC, Huyse F, So AKL.:** Identification of Case Complexity and Increased Health Care Utilization in Patients With Rheumatoid Arthritis. *Arthritis Care and Research. American College of Rheumatology.* Published by Wiley-Liss, Inc. 2001; 45: 216-221.
22. **Lobo A, Campos R, Pérez-Echeverría MJ, Izuzquiza J, García-Campayo J, Saz P, Marcos G.:** A new interview for the multiaxial assessment of psychiatric morbidity in medical settings. *Psychol Med* 1993; 23: 505-510.
23. **Lobo A, Huyse FJ, Herzog T, Malt UM, Opmeer B, and the ECLW.:** The European Consultation Liaison Workgroup (ECLW) Collaborative Study. II: Patient Registration Form (PRF) instrument, training and reliability. *J. Psychosom. Res.* 1996; 40: 143-156.
24. **Maclure M, Willet WC.:** Misinterpretation and misuse of the kappa statistic. *Am J Epidemiol* 1987; 126: 161-169.

-
25. **Maxwell AE.:** Coefficients of Agreement between observers and their interpretation. *Br J Psychiatry* 1977; 130: 79-83.
26. **Naylor MD, Brooten D, Campbell R, et al.:** Comprehensive discharge planning and home follow-up of hospitalised elders. *JAMA* 1999; 281: 613-620.
27. **Opmeer BC.:** Patterns of C-L service delivery in Europe. Amsterdam, Vrije Universiteit Press 1999.
28. **Regier DA, Burke JD.:** Quantitative and experimental methods in psychiatry. En: Kaplan, H.I.; Sadock, B.J. ed. *Comprehensive Textbook of Psychiatry* (4° Ed.). Baltimore: Williams and Wilkins 1989; 308-326.
29. **Shrout PE, Spitzer RL, Fleiss JL.:** Quantification of agreement in psychiatric diagnosis revisited. *Arch Gen Psychiatry* 1978; 44: 172-177.
30. **Spitzer RL, Fleiss JL.:** A re-analysis of the reliability of psychiatric diagnosis. *British Journal of Psychiatry* 1974; 125: 341-347.
31. **Spitzer RL, Williams JBW.:** Classification of mental disorders. En: Kaplan, H.I.; Sadock, B.J. ed. *Comprehensive Textbook of Psychiatry* (5° Ed.). Baltimore: Williams and Wilkins 1989; 591-613.
32. **Stiefel FC, De Jonge P, Huyse F, Slaets JPI, Guex P, Lyons JS, Vannotti M, Fritsch C, Moeri R, Leyvraz P, So A, Spagnoli J.:** INTERMED-An Assessment and Classification System for Case Complexity. Results in Patients With Low Back Pain. *SPINE* 1999 vol 24, 4: 378-389.
33. **Strain JJ, Lyons JS, Hammer JS, et al.:** Cost offset from a psychiatric consultation-liaison intervention with elderly hip fracture patients. *Am. J. Psychiatry* 1991; 148: 1044-1048.
34. **Tinsley HEA, Weiss DJ.:** Interrater reliability and agreement of subjective judgements. *J Counsel Psychol* 1975; 22: 358-375.
35. **Wittchen HU, Robins LN, Cottler LB, Sarotius N, Burke JD, Regier D.:** Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). *Br J Psychiatry* 1991; 159: 645-653.