

**PROPIETATS DELS INSTRUMENTS DE
MESURA DE QUALITAT DE VIDA O
ALTRES RESULTATS PERCEBUTS PELS
PACIENTS (PROs)**

Aida Ribera

Unitat d'Epidemiologia

Servei de Cardiologia. Àrea del cor

OBJECTIUS

1. Conèixer els criteris per:

Escollir l'instrument més adequat

Avaluar les propietats de mesura

2. Discussió d'exemples pràctics

- 1. Quants esteu interessats en utilitzar PROs per a la mesura de resultats en els vostres estudis?**
- 2. Quants esteu interessats en l'estudi metodològic de PROs? És a dir, voleu adaptar, validar, etc, algun instrument?**
- 3. En quins camps?**

A QUÈ ENS REFERIM QUAN PARLEM DE PROs?

Resultats percebuts pels pacients

PRO: 'Patient reported outcomes'

Qualitat de vida relacionada amb la salut:

Impacte de la malaltia i del seu tractament tal com és percebut pel malalt i relatiu a la seva salut general i als àmbits del funcionament físic, emocional i social

= Salut percebuda

Altres PROs: inventoris de símptomes, estat funcional, satisfacció, cribatge, utilitats, ...

CONCEPTES

Qüestionaris genèrics

Salut general.

Permeten comparacions amb població general i amb grups de pacients amb problemes de salut diferents.

Qüestionaris específics

Donen informació sobre aspectes que són característics dels pacients que pateixen un problema de salut concret.

Son més sensibles al canvi clínic experimentat pels pacients.

PER QUÈ PARLEM DE PROPIETATS PSICOMÈTRIQUES?

Dues tradicions de mesura

TRADICIÓ CLINIMÈTRICA	TRADICIÓ PSICOMÈTRICA
Pensament categòric	Pensament multidimensional i continu
Diagnòstic, tractament	Grau de severitat dins un continu
La validesa s'estableix per criteri i experiència clínica	Es desenvolupen tècniques sofisticades per al desenvolupament i avaluació d'instruments de mesura

INTERÈS DE LA MESURA DE RESULTATS PERCEBUTS PELS PACIENTS PER QUÈ MESURAR-LOS?

1. Rellevància en un context de canvi de paradigma en la medicina: augment de l'esperança de vida i de la supervivència en malalts crònics

2. Aporten informació que no aporten les mesures clíniques:

Visió més àmplia

Des de la perspectiva del malalt

Inclouen judicis de valor, preferències, satisfacció. Aspectes subjectius rellevants i susceptibles de mesura

3. La mesura d'aquests resultats por ser robusta: vàlida i fiable

INTERÈS DE LA MESURA DE RESULTATS PERCEBUTS PELS PACIENTS. PER QUÈ MESURAR-LOS?

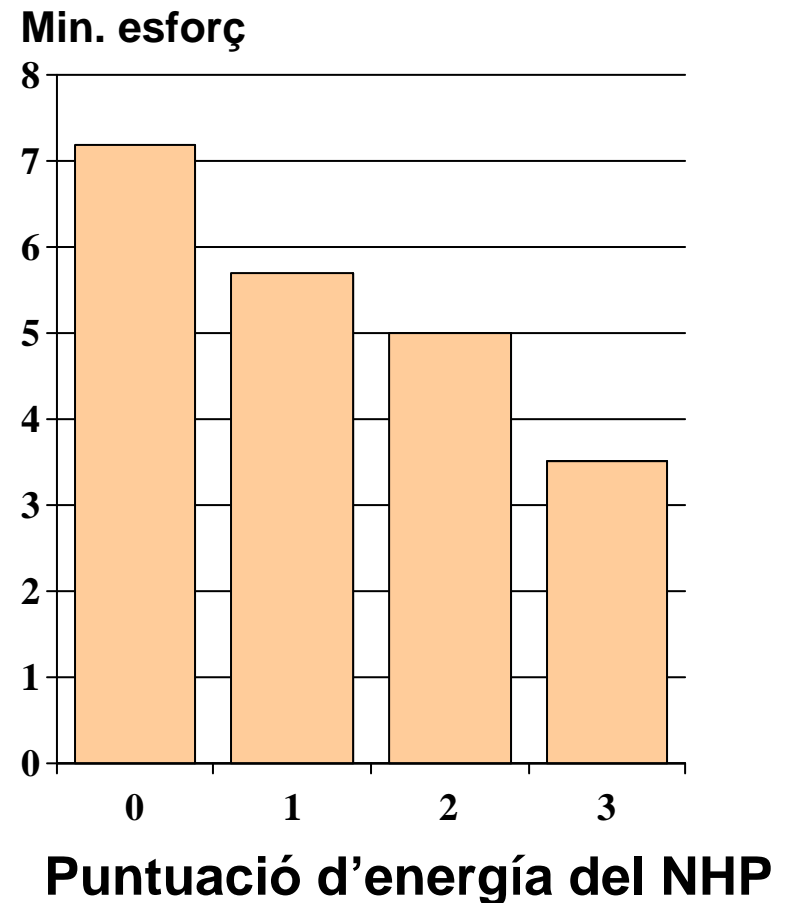
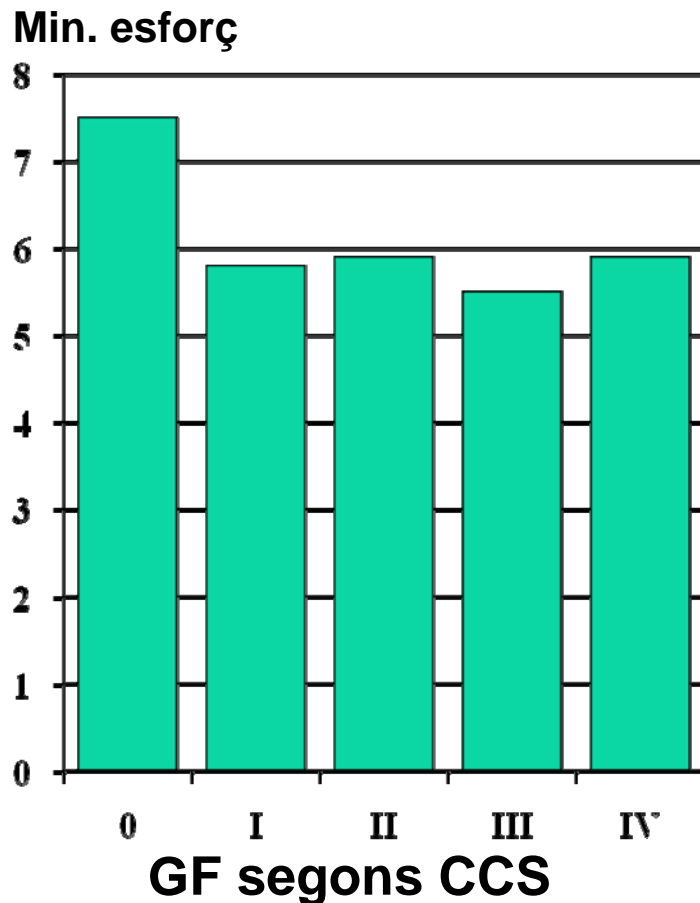
Recerca clínica: estudi dels determinants de la salut o de l'impacte de la malaltia (qüestionaris específics o genèrics), com a variable secundària en assajos clínics aleatoritzats (qüestionaris específics), estudis de cost-efectivitat o cost-utilitat (càlcul de preferències).

Salut pública: enquestes de salut (qüestionaris genèrics), avaluació de necessitats per a la planificació sanitària. Exemple ESCA (enquesta de salut de Catalunya).

El seu ús en **la pràctica clínica** habitual és més limitat i alhora controvertit: desconfiança per part de clínics, falta de criteris d'interpretació, estàndards elevats de validesa i fiabilitat per a l'ús en el malalt individual.

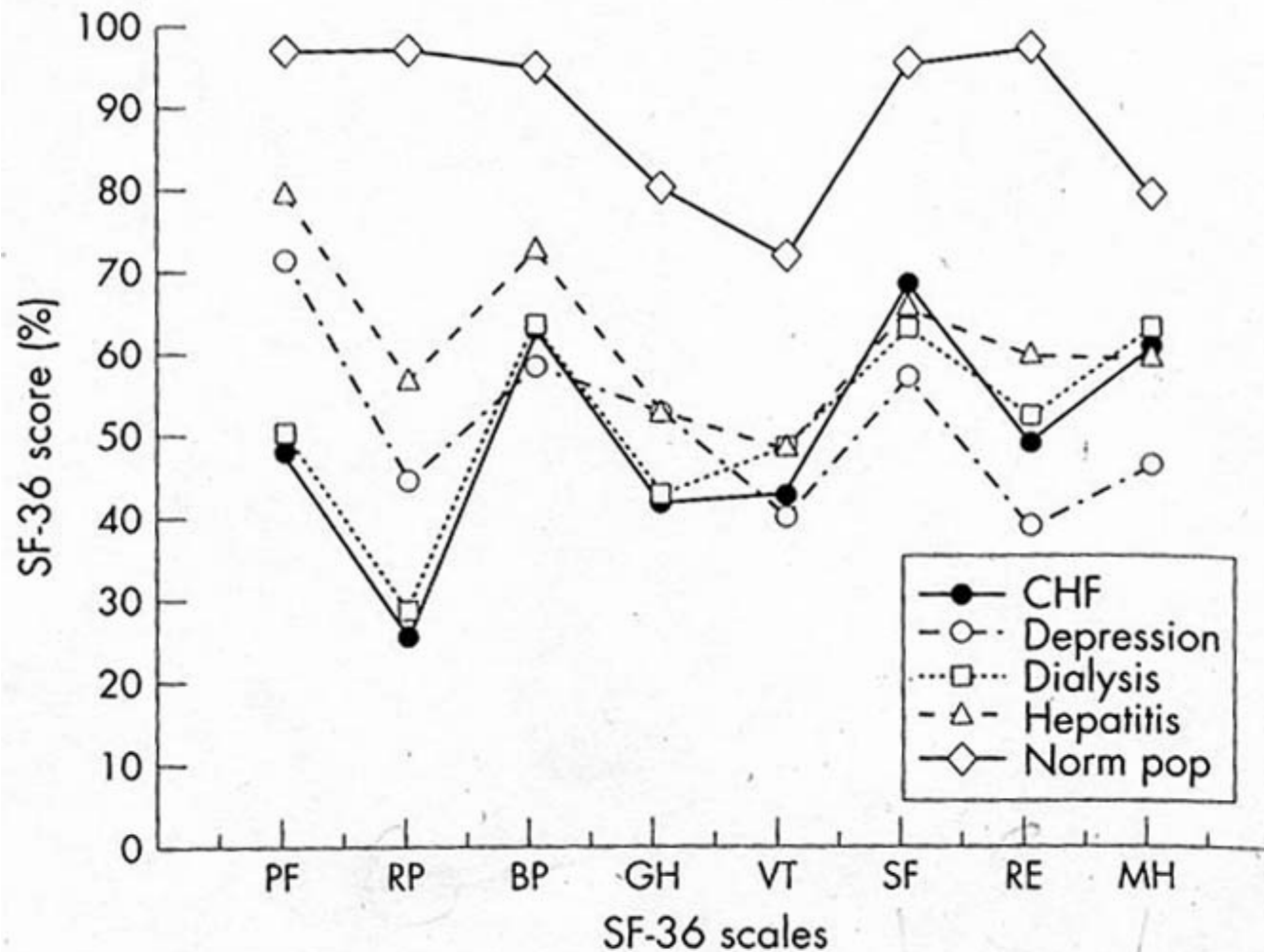
CORRELACIÓ AMB MESURES CLÍNIQUES

Relació entre capacitat d'esforç, grau funcional clínic i salut percebuda (energia NHP)



IMPACTE DE LA MALALTIA

Comparació de QV genèrica entre diferents patologies



VALOR PRONÒSTIC

Predictors basals de mortalitat als 3 mesos després d'un ingrés per insuficiència cardíaca

	OR	(IC 95%)
PCS inicial	0,96	(0,93 – 0,99)*
MCS inicial	0,99	(0,97 – 1,01)
Edat	1,09	(1,04 – 1,13)*
Dones	0,64	(0,32 – 1,27)
IC prèvia	2,28	(0,99 – 5,24)
Diabetis	1,51	(0,81 – 2,81)
Comorbiditat	0,80	(0,53 – 1,20)
IC biventricular	1,31	(0,69 – 2,47)
NYHA IV	1,79	(0,96 – 3,32)

QUÈ FER QUAN NECESSEM UNA MESURA DE PRO?

1. Cerca
2. Adaptació cultural i lingüística de qüestionaris existents
3. Desenvolupament de nous qüestionaris

BiblioPRO

- BiblioPRO és una biblioteca virtual de qüestionaris de Qualitat de Vida Relacionada amb la Salut i altres resultats percebuts pels pacients (Patient Reported Outcomes, PRO) en espanyol **d'accés lliure i gratuït.**
- Pretén **promoure l'ús adequat dels instruments** de QVRS en espanyol en la investigació, pràctica clínica y gestió socio-sanitària, proporcionant informació exhaustiva online, avaluacions basades en l'evidència, així com realitzant formació especialitzada en aquest camp.
- Actualment la biblioteca compta amb informació sobre **més de 500 qüestionaris** i s'actualitza periòdicament.
- Es preveu que el juny de 2010 disposarà dels descarregables del 50% dels instruments.

QUÈ FER QUAN NECESSITEM UNA MESURA DE PRO?

1. Cerca
2. Adaptació cultural i lingüística de qüestionaris existents
3. Desenvolupament de nous qüestionaris

QUÈ FER QUAN NECESSITEM UNA MESURA DE PRO?

- 1. Cerca**
Triar en funció de les seves propietats i grau de validesa
- 2. Adaptació cultural i lingüística de qüestionaris existents**
Avaluar les propietats psicomètriques i validesa en el nostre context
- 3. Desenvolupament de nous qüestionaris**
Avaluar les propietats psicomètriques i validesa

CRITERIS D'AVALUACIÓ D'INSTRUMENTS DE MESURA. EMPRO

Vuit atributs per a l'avaluació d'instruments:

- 1. Model conceptual**
- 2. Fiabilitat**
- 3. Validesa**
- 4. Sensibilitat als canvis**
- 5. Interpretabilitat**
- 6. Facilitat d'administració**
- 7. Formes alternatives**
- 8. Adaptació transcultural**

Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 2002; 11: 193-205

Valderas JM et al. Development of EMPRO: A tool for the standardized assessment of patient reported outcomes measures. Value in Health

CRITERIS D'AVALUACIÓ D'INSTRUMENTS DE MESURA

- 1. Adaptació transcultural**
- 2. Fiabilitat**
- 3. Validesa**
- 4. Sensibilitat als canvis**
- 5. Interpretabilitat**

Exemples: MacNew QLMI. Específic postinfart
 SF-36. Genèric
 CHIP-AE. Específic pediatria
 Inventori de xerostomia. Específic de símptomes

El qüestionari de QVRS per al postinfart de miocardi McNew QLMI

Versió del qüestionari	Dominis	Puntuacions sumàries	Sistema de puntuació	Propietats de mesura estudiades
QLMI (Hillers et al, McMaster)	Síntomes (5) Restriccions(4) Confiança (4) Autoestima (7) Emocions (6) 26 ítems	Global Físic Global Emocional	Clinimètric	Fiabilitat Validesa discriminativa i avaluativa
McNew-QLMI (Linn et al, Newcastle)	Físic (14) Emocional (14) Social (12) 27 ítems	Puntuació Global	Psicomètric	Fiabilitat Validesa de constructe

Exemple. MacNew

6. ¿Cuántas veces, durante las 4 últimas semanas, se ha sentido agotado/a o bajo/a de energías?
 1. SIEMPRE.
 2. CASI SIEMPRE.
 3. A MENUDO.
 4. ALGUNAS VECES.
 5. POCAS VECES.
 6. CASI NUNCA.
 7. NUNCA.

ETAPES EN L'ADAPTACIÓ DE LA VERSIÓ ESPANYOLA DEL MACNEW QLMI

Traducció	Procés de traducció-retrotraducció en col.laboració amb els autors de la versió original. Estudi pilot en 10 pac per valorar comprensió i facilitat d'administració
Equivalència lingüística i cultural	Anàlisi factorial confirmatòria en una mostra de 143 pac en el marc d'un estudi de prevenció secundària
Fiabilitat	Estudi de reproducibilitat i consistència interna en una mostra de 50 pac en situació estable amb reevaluació als 15 dies
Validesa i sensibilitat als canvis	Fiabilitat, validesa i sensibilitat als canvis en una mostra de 97 pac avaluats després de l'IAM i seguits durant un any

Brotos C et al. Med Clin 2000; 115: 768-771
Ribera A et al. Qual Life Res 2006; 15: 357-365

ADAPTACIÓ TRANSCULTURAL

- Traducció: traducció-retrotraducció, participació d'experts i pacients
- Avaluació de l'equivalència lingüística i conceptual: anàlisi factorial confirmatòria, participació d'experts i pacients

PROCÉS DE TRADUCCIÓ-RETROTRADUCCIÓ.

Ex. Inventori de Xerostomia

- Versió original
 - Traducció
 - Discussió experts
 - 1^a versió consensuada
 - Retrotraducció
 - Discussió experts, participació pacients
 - 2a versió consensuada
- I sip liquids to aid in swallowing food
- Bebo líquidos para ayudarme a tragar los alimentos
- Tomo sorbos para facilitar la deglución de los alimentos
- Tomo sorbos de líquido para ayudarme a tragar los alimentos
- I take sips of liquid to help me swallow food

AVALUACIÓ DE L'EQUIVALÈNCA LINGÜÍSTICA I CONCEPTUAL

Anàlisi factorial confirmatòria

	F1	F2	F3
1. Frustrado	0.749	0.089	0.160
2. Inútil	0.402	0.208	0.631
3. Confiado	0.608	0.408	0.023
4. Desanimado	0.689	0.245	0.358
5. Relajado	0.729	0.251	0.248
6. Agotado	0.554	0.503	0.251
7. Satisfecho	0.695	0.206	0.137
8. Agitado	0.743	0.160	0.321
9. Falto de aire	0.487	0.500	0.200
10. Lloroso	0.486	0.097	0.501
11. Dependiente	0.348	0.355	0.577
12. Actividades sociales	0.309	0.157	0.610
13. Confianza de otros	0.317	0.297	0.481
14. Dolor en el pecho	0.260	0.619	0.231
15. Inseguro	0.371	0.311	0.678
16. Dolor piernas	0.435	0.457	0.132
17. Limitado deporte	0.209	0.748	0.213
18. Asustado	0.469	0.285	0.563
19. Mareado	0.590	0.217	0.257
20. Limitado	0.224	0.750	0.259
21. Inseguro ejercicio	0.281	0.663	0.224
22. Sobreprotegido	-0.116	0.380	0.534
23. Carga	0.300	0.215	0.647
24. Excluído	0.127	0.643	0.436
25. Incapaz relación	0.049	0.249	0.601
26. Limitació física	0.297	0.630	0.431
27. Relación sexual	0.092	0.693	0.024

Tabla 3. Pesos de los 27 ítems del cuestionario en los tres factores extraídos mediante análisis de componentes principales con rotación *varimax*

ANÀLISI FACTORIAL CONFIRMATÒRIA

1. Frustrado	0.749
2. Inútil	0.402
3. Confiado	0.608
4. Desanimado	0.689
5. Relajado	0.729
6. Agotado	0.554
7. Satisfecho	0.695
8. Agitado	0.743
9. Falto de aire	0.487
10. Lloroso	0.486
16. Dolor piernas	0.435
18. Asustado	0.469
19. Mareado	0.590

Ítems relacionats amb l'estat d'ànim

~ dimensió emocional?

ANÀLISI FACTORIAL CONFIRMATÒRIA

3. Confiado	0.408
6. Agotado	0.503
9. Falto de aire	0.500
14. Dolor en el pecho	0.619
16. Dolor piernas	0.457
17. Limitado deporte	0.748
20. Limitado	0.750
21. Inseguro ejercicio	0.663
24. Excluído	0.643
26. Limitació física	0.630
27. Relación sexual	0.693

Ítems relacionats amb el dolor i limitacions físiques

~ dimensió física?

ANÀLISI FACTORIAL CONFIRMATÒRIA

2. Inútil	0.631
10. Lloroso	0.501
11. Dependiente	0.577
12. Actividades sociales	0.610
13. Confianza de otros	0.481
15. Inseguro	0.678
18. Asustado	0.563
22. Sobreprotegido	0.534
23. Carga	0.647
24. Excluído	0.436
25. Incapaz relación	0.601
26. Limitació física	0.431

Ítems relacionats amb
incapacitats o percepcions
en relació amb les activitats
socials o amb la família

~ dimensió social?

ANÀLISI FACTORIAL CONFIRMATÒRIA

- **Bona equivalència amb la versió original**
- **Es detecten diferències subtils:**

Es van detectar alguns termes que no tenien equivalència lingüística exacta: per exemple “lloroso” ~ “tearful” o “mareado” ~ “dizzy”

Així com activitats amb diferent connotació cultural: “Seguro de cuanto ejercicio debería hacer”, en la versió anglesa té també molt de pes en la dimensió “social”

FIABILITAT

En quin grau la mesura està lliure d'error aleatori?

Reproductibilitat	Consistència interna
Estabilitat de la mesura en un estudi longitudinal en pacients estables	En quin grau els diferents ítems d'una escala són mesures del mateix concepte
$CCI = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2}$	$\alpha = \frac{k\bar{r}}{1 + (k - 1)\bar{r}}$
0,7 / 0,9	0,7 / 0,9

s: subjectes; e:error;

K: N° d'ítems

FIABILITAT

consistència interna y reproducibilitat del Macnew QLMi

	Consistència interna	Reproductibilitat
	Alfa de Cronbach (n=97)	Correlació intraclasse (n=50)
Físic (14)	0,88	0,82
Emocional (14)	0,92	0,80
Social (12)	0,87	0,74
Puntuació global (27)	0,92	0,83

FIABILITAT

En quin grau la mesura està lliure d'error aleatori?

La fiabilitat augmenta si augmentem el nombre d'ítems

Una alfa de Chronbach massa alta pot indicar redundància. Més de 0.7, però no més de 0.9?

VALIDESA

En quin grau l'instrument mesura allò que pretén mesurar?

- Per què hem de validar? A diferència d'altres variables (temperatura, nivell de creatinina, colesterol, etc) les mesures de PROs no mesuren una dimensió directament observable.
- “Tipus” de validesa (problema de terminologia): validesa de contingut, de criteri, de constructe, convergent/discriminant, ...

VALIDESA

En quin grau l'instrument mesura allò que pretén mesurar?

- Validesa de contingut ~ model conceptual
- Validesa de criteri: quan existeix *gold standard*
- Validesa de constructe

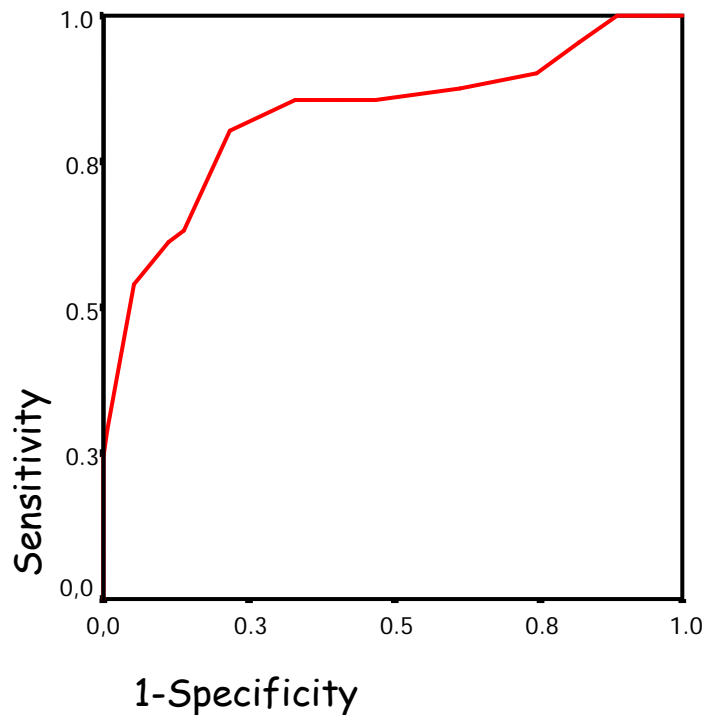
VALIDESA DE CONTINGUT

- S'aconsegueix en el procés de selecció d'ítems del qüestionari. Estem preguntant tot el que és rellevant per a la mesura d'interès?
- De vegades millorar la validesa de contingut pot significar reduir la fiabilitat (heterogeneïtat del contingut).

Exemple: ítems de 'Burning mouth Syndrome' en l'Inventori de Xerostomia

VALIDESA DE CRITERI

Ne. d'assignatures
suspeses (cap / 1 o més)
vs subdimensió rendiment
acadèmic del CHIP-AE



Àrea sota la corva ROC: 0,83 (IC 95% = 0,75-0,92)

VALIDESA DE CONSTRUCTE

- Constructe: “Atribut que no és directament mesurable però que es relaciona amb altres atributs o comportaments observables a partir dels quals podem construir hipòtesis (p.ex. ansietat, intel.ligència)”
- ~ síndrome = constel.lació de símptomes
- La validació de constructe no s’acaba amb un estudi: procés de formulació i contrast de noves hipòtesis

VALIDESA DE CONSTRUCTE

- Formulació d'hipòtesis: Basant-nos en el constructe teòric que pretenem mesurar, en quin grau les persones que puntuïn alt (o baix) amb el nostre instrument tindran un (o altre) diagnòstic, puntuaran alt (o baix) en aquesta altra mesura, etc?
- P.ex. si volem mesurar la qualitat de vida específica en pacients amb EPOC:
La mesura es correlacionarà positivament amb:
 - altres mesures de qualitat de vida (genèrica)
 - Mesura clínica de dísipnea en les activitats de la vida diària
 - resultat de proves funcionals respiratòries.
 - ...

VALIDESA DE CONSTRUCTE (CONVERGENT I DISCRIMINANT)

Correlacions esperades (matriu multi-tret multi-mètode)

		Mètode X		Mètode Y	
		Tret A	Tret B	Tret A	Tret B
Mètode X	Tret A	<i>R</i>			
	Tret B	Baixa	<i>R</i>		
Mètode Y	Tret A	Alta	Molt baixa	<i>R</i>	
	Tret B	Molt baixa	Alta	Baixa	<i>R</i>

Convergent

Discriminant

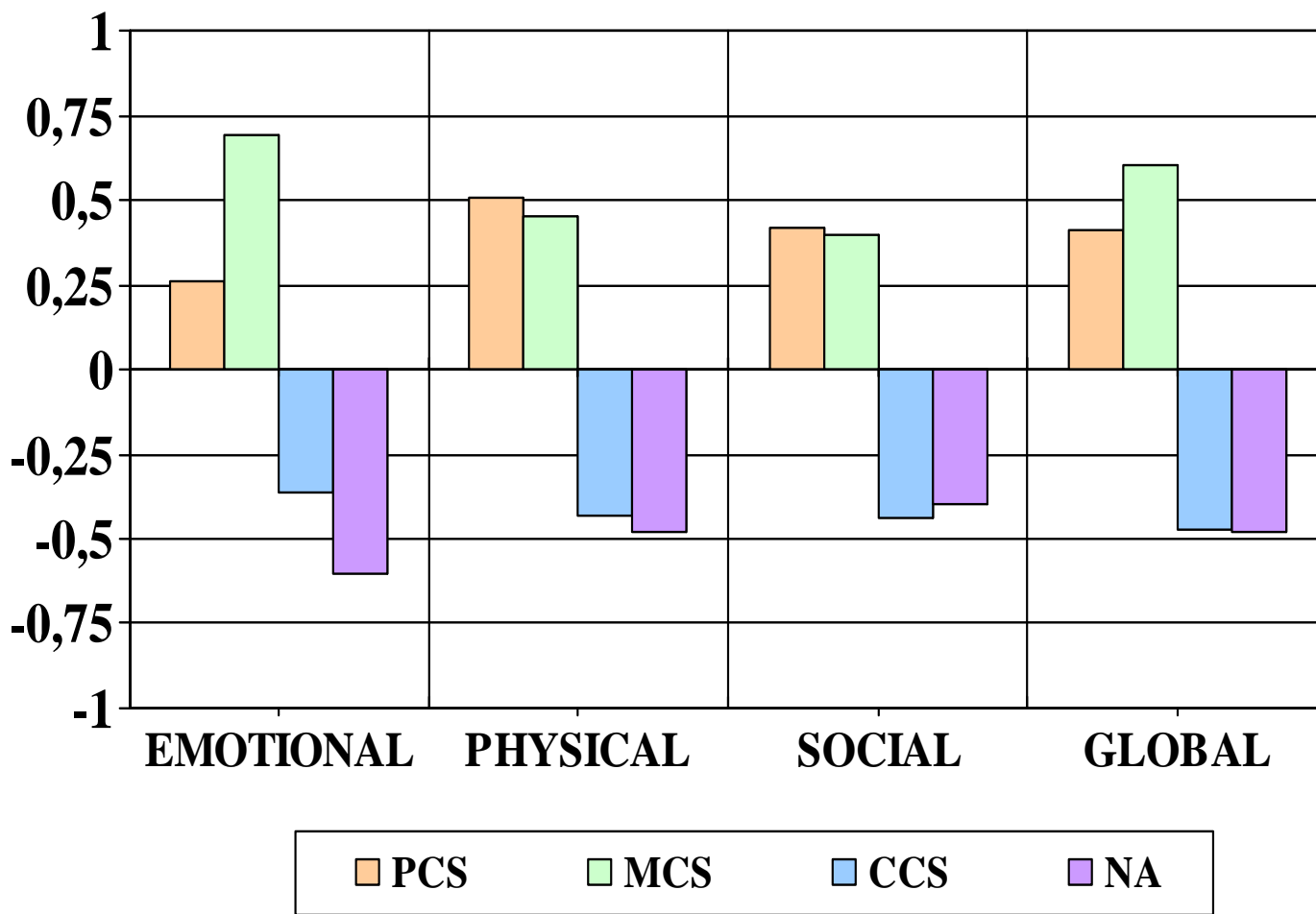
Matriu multi-tret multi-mètode

	MacNew QLMI							
	Emotional		Physical		Social		Global	
SF-36	Exp	Obs	Exp	Obs	Exp	Obs	Exp	Obs
Physical function	L	0.34	H	0.51	L	0.52✓	M	0.48
Role physical	L	0.35	H	0.42☑	L	0.41✓	M	0.41
Bodily Pain	L	0.32	M	0.41	L	0.29	M	0.41
General Health	M	0.56✓	M	0.43	M	0.43	H	0.51
Vitality	H	0.60	H	0.65	M	0.53✓	H	0.66
Social Function	M	0.52✓	L	0.46✓	H	0.45☑	M	0.53✓
Role emotional	H	0.56	L	0.44✓	M	0.41	M	0.51✓
Mental Health	H	0.79	L	0.54✓	M	0.48	M	0.66✓

Expected correlations; L: low (Spearman cor.: 0.2 - 0.35), M: medium (Spearman cor.: 0.35 - 0.5), H: high (Spearman cor.: over 0.5).

VALIDESA DE CONSTRUCTE

Correlació entre les tres dimensions del MacNew QLMI i altres indicadors de salut



SENSIBILITAT ALS CANVIS

Quina capacitat té l'instrument per detectar canvis?

- Magnitud del canvi. Quocient entre el canvi i la variança

Resposta mitjana estandarditzada (RME):

$$\frac{\bar{C}_{(P. inicial - P. final)}}{DE_c}$$

- Estudis longitudinals. Inductors de canvi (tractament, història natural). Comparació amb grups estables

SENSIBILITAT ALS CANVIS

Magnitud del canvi entre les dues setmanes i els dos mesos després del IAM

Dimensions del MacNew QLMI	RME
Físic	0,73
Emocional	0,27
Social	0,68
Puntuació global	0,56

Puntuacions sumàries y funció social del SF-36	RME
PCS	0,46
MCS	0,41
Funció social	0,50

PREGUNTA DE TRANSICIÓ

Com creu que es la seva salut actual comparada amb la de l'última entrevista?

1. Molt millor ara que abans
2. Bastant millor ara que abans
3. Més o menys igual
4. Bastant pitjor ara que abans
5. Molt pitjor ara que abans

SENSIBILITAT ALS CANVIS

Magnitud del canvi en les puntuacions segons la resposta a la pregunta de transició

	Dimensions del MacNew QLMI			
	Emocional	Físic	Social	Global
	“...millor o molt millor...”			
RME	-0.58*	-0.76*	-0.71*	-0.69*
	“...més o menys igual...”			
RME	0.09	-0.35*	-0.43*	-0.18

SENSIBILITAT ALS CANVIS

Magnitud del canvi en les puntuacions segons el canvi en la classe funcional

	Emocional	Físic	Social	Global
	Millora en el CCS			
RME	-0.57*	-0.77*	-0.74*	-0.61*
	Sense canvis en el CCS			
RME	0.1	-0.34*	-0.42*	-0.36*

INTERPRETABILITAT

Significat “clínic” de la mesura

- Quina és la diferència o el canvi observable en la mesura de salut percebuda que té rellevància des del punt de vista clínic o social?
- ‘Anchor based’ / ‘Distribution based’

Symposium on Quality of Life in Cancer Patients. Sèrie de 6 articles de revisió.
Mayo Clin Proc 2002; 77: 371-583

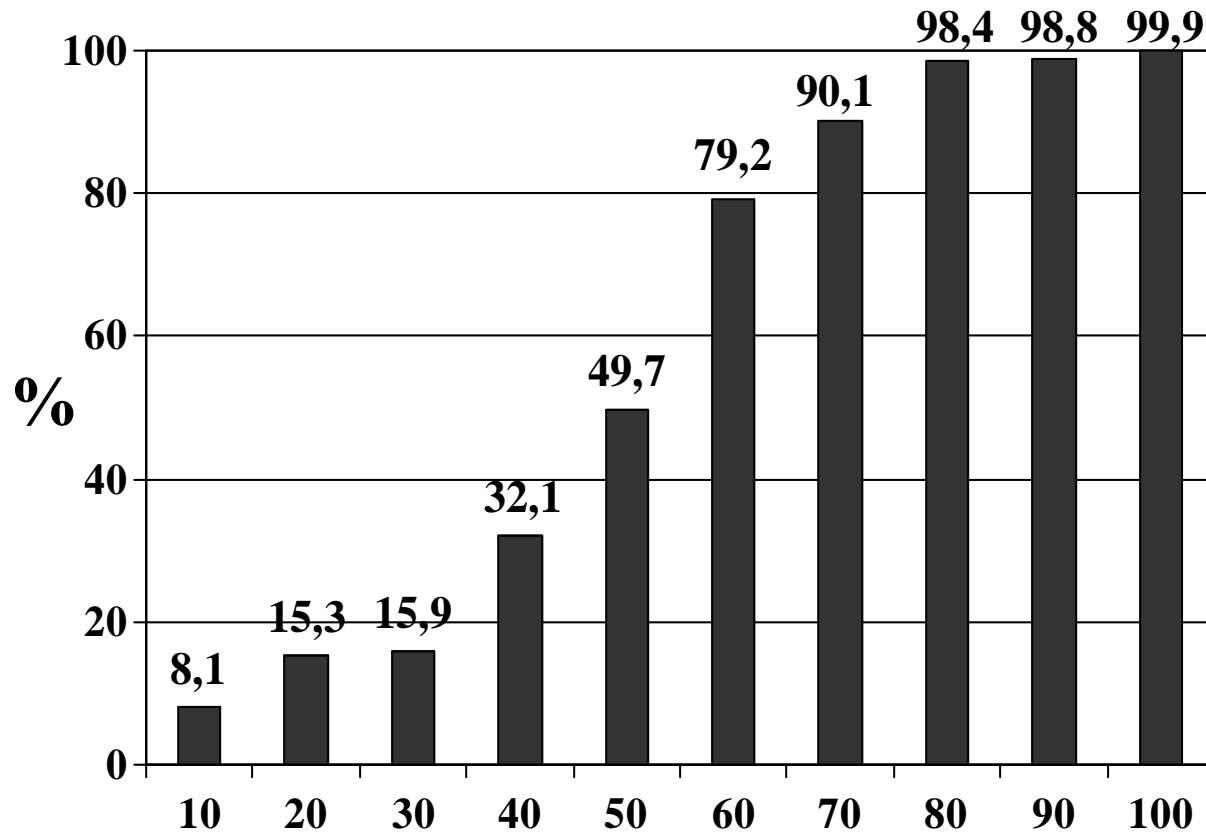
Guyatt GH, et al. Methods to explain the clinical significance of health status measures. Mayo Clin Proc 2002;77:371-383

INTERPRETABILITAT

‘ANCHOR BASED’

- ~ Validesa de constructe, ~ Sensibilitat als canvis
- Relació amb variables que tinguin un significat clínic o social: Una categoria o una altra té implicacions directes en la presa de decisions (prevenció, tractament...)
- **MID***: la diferència més petita en una puntuació que els pacients perceben com a beneficiosa (o perjudicial) i que comportaria algun canvi en el seu maneig.

INTERPRETACIÓ BASADA EN EL CONTINGUT



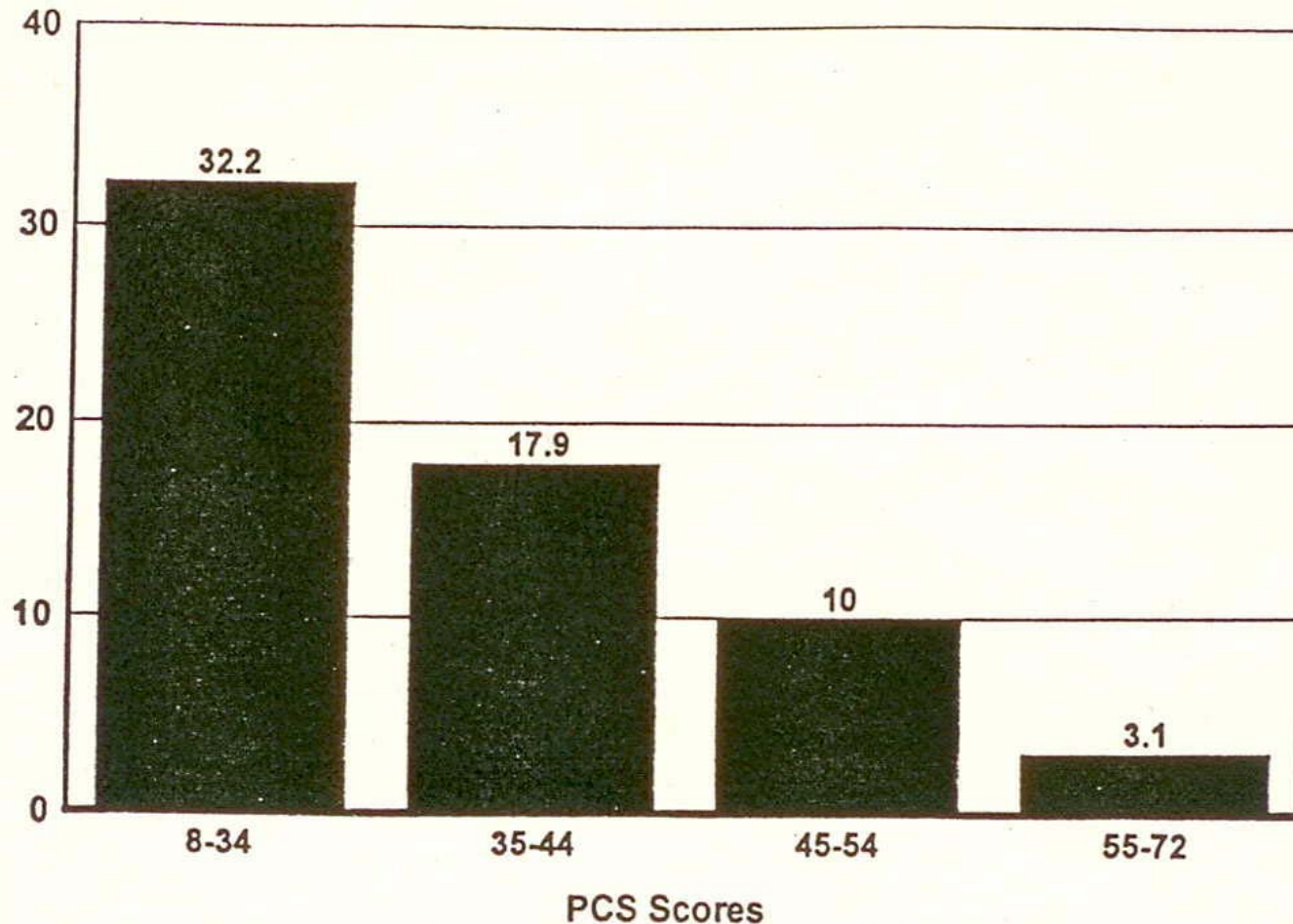
Physical Functioning Scale (MOS)

Percentage that can walk one block or more at each scale level

Ware JE, Keller SD. Interpreting general health measures, in: Quality of Life and Pharmacoeconomics in Clinical Trials, 1996

INTERPRETACIÓ BASADA EN CRITERIS EXTERNS

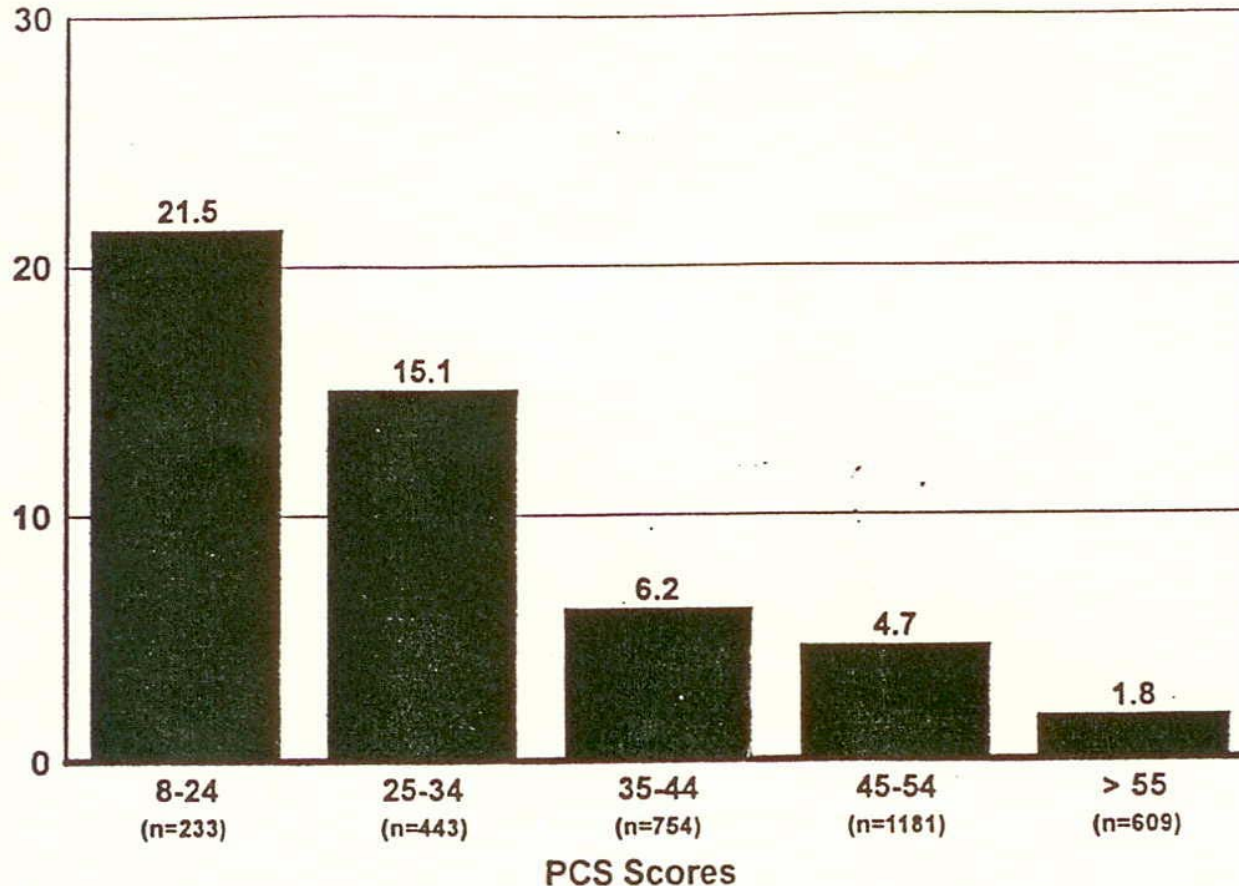
Criteri de significat social



Percentage of working patients who lost their job one year later, by levels of PCS scores

Ware JE, Keller SD. Interpreting general health measures, in: Quality of Life and Pharmacoeconomics in Clinical Trials, 1996

INTERPRETACIÓ BASADA EN CRITERIS EXTERNS: Criteri de significat clínic/pronòstic



Five-year mortality rates for MOS patients, by levels of PCS scores

Ware JE, Keller SD. Interpreting general health measures, in: Quality of Life and Pharmacoeconomics in Clinical Trials, 1996

INTERPRETABILITAT

EL 'MID' I LES PREGUNTES DE TRANSICIÓ

- | | |
|-------------------------------------|------------|
| 1. Almost the same | ~0 |
| 2. A little worse (better) | } MID~0.5 |
| 3. Somewhat worse (better) | |
| 4. Moderately worse (better) | } ~0.8-0.9 |
| 5. A good deal worse (better) | |
| 6. A great deal worse (better) | } ~1.5 |
| 7. A very great deal worse (better) | |

INTERPRETABILITAT

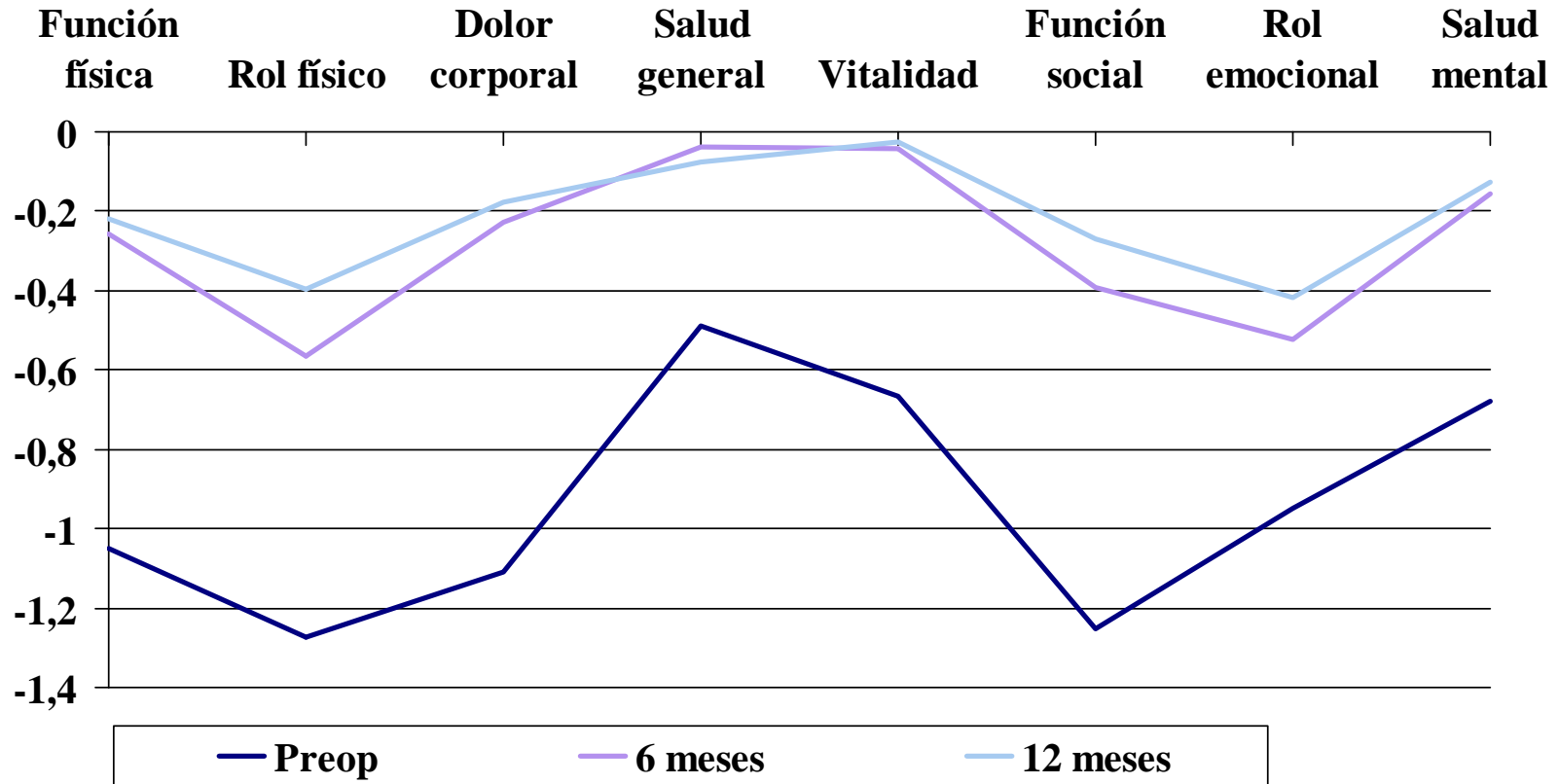
EL 'MID' I LES PREGUNTES DE TRANSICIÓ

- Validesa de la pregunta de transició?
- Problema de la proporcionalitat de les escales
- Útil en assajos clínics (proporció de casos amb canvi \geq MID)
- És el mateix per les diferències entre grups?

INTERPRETABILITAT 'DISTRIBUTION BASED'

- Població de referència
- Tamany de l'efecte. Criteris de Cohen:
Canvi petit: 0.2 DE
Canvi moderat: 0.5 DE
Canvi gran: 0.8 DE
- S'aplica a les diferències intra i interindividuals
- Relació amb el MID: 0,2 ~ 0,5 !

INTERPRETACIÓN BASADA EN NORMES



Evolución de las puntuaciones medias de las dimensiones de salud percibida (SF36) relativas a la población española de referencia ajustadas por edad y sexo, en pacientes operados de by-pass coronario

INTERÈS DE L'AVALUACIÓ DE PROPIETATS DE MESURA DE PROs EN L'ÀMBIT CLÍNIC

- Els estudis de **validesa, sensibilitat als canvis i interpretabilitat clínica** es poden fer en el context d'estudis clínics
- L'expertesa del clínic afavoreix la formulació d'hipòtesis coherents
- L'ús en el dia a dia de la pràctica clínica augmentaria de forma 'natural' la interpretabilitat de les escales.