

**UNIVERSIDAD DE SALAMANCA**

**DEPARTAMENTO DE ESTADÍSTICA**

---



**TESIS DOCTORAL**

**EVALUACIÓN DEL CAMBIO EN PACIENTES**

**ONCOLÓGICOS:**

**UN ENFOQUE ESTADÍSTICO**

**SARA MONTEIRO MORGADO DIAS NUNES**

**2009**

# EVALUACIÓN DEL CAMBIO EN PACIENTES ONCOLÓGICOS: UN ENFOQUE ESTADÍSTICO

---

Memoria que para optar al Grado de Doctor,  
por el Departamento de Estadística de la  
Universidad de Salamanca, presenta:

*Sara Monteiro Morgado Dias Nunes*

Salamanca

2009



*Universidad de Salamanca*  
*Departamento de Estadística*

---

**M<sup>a</sup> PURIFICACIÓN GALINDO VILLARDÓN**

**M<sup>a</sup> PURIFICACIÓN VICENTE GALINDO**

*Profesoras del Área de Estadística e Investigación Operativa  
de la Universidad de Salamanca*

CERTIFICAN: Que **Sara Monteiro Morgado Dias Nunes**, Licenciada en Enseñanza de Matemática e Ciências da Natureza, ha realizado en el Departamento de Estadística de la Universidad de Salamanca, bajo su dirección, el trabajo que para optar al Grado de Doctor, presenta con el título: *"Evaluación del Cambio en Pacientes Oncológicos: Un Enfoque Estadístico"*; y para que conste, firman el presente certificado en Salamanca, en 19 de Octubre de 2009.

Ma. Purificación Galindo Villardón

Ma. Purificación Vicente Galindo

**"I know who I was when I got up this  
morning, but I've changed several  
times since then."**

*Alice in Wonderland  
by Lewis Carroll*

***A mi Madre***

*por haberme proporcionado una vida maravillosa*

***A mis Abuelos***

*por haberla llenado de amor*

***A Rui***

*por darle sentido*

## AGRADECIMIENTOS

---

En primero lugar agradezco a la Dra Galindo, mi tutora y directora de este trabajo, por tantas cosas que me ha enseñado en estos años de Doctorado, por su excepcional profesionalismo y también por su entera disponibilidad y dedicación a este proyecto. Gracias por tantos consejos, por incentivar me siempre a mejorar y por todas las palabras de ánimo.

A Purificación Vicente, codirectora de este trabajo, por poner a mi disposición todos sus conocimientos sobre el tema de la Calidad de Vida.

A José, Maria José, Inma, Carmelo, Javi, Antonio, Santi, Mercedes y Rosita, profesores del Departamento de Estadística, por todo lo que de ellos he aprendido durante mi formación de Doctorado.

Al Instituto Politécnico de Castelo Branco, donde trabajo, por haber criado, a lo largo de estos años de formación, todas las condiciones para que pudiera conciliar las tareas del profesorado con la investigación, un arte nada fácil.

A las Instituciones Hospitalarias que aceptaran en colaborar en este estudio, abriendo sus puertas para la recogida de datos. Gracias al Dr Sanches Pires, director del Hospital Amato Lusitano (HAL) en Castelo Branco; a la Dra Rita Resende, responsable del Hospital de Día de Oncología de la misma institución; al Dr Eduardo Pereira del Servicio de Gastroenterología del HAL; al Dr. António Lagoa, director del Servicio de Senología del Hospital de Faro; a la Dra Teresa Figueiredo, radióloga en el mismo hospital; a la Dra Joana Veloso, psicóloga en el Hospital de Faro; a la Dra Estela Landeiro, psicóloga en el Centro Hospitalar Cova da Beira (CHCB); al Dr José Alberto Moutinho, director del Servicio de

Ginecología/Obstetricia del CHCB y a todos los profesionales de Salud que, de alguna manera, han contribuido para la aplicación de los cuestionarios. Todos ellos han, de formas distintas, colaborado en este proyecto. Un agradecimiento muy especial a las enfermeras Manuela Cruz del Servicio de Senología del Hospital de Faro, a Ana Monteiro de la Unidad de Cirugía del HAL en Castelo Branco y Helena Santos del Hospital de Día de Oncología de la misma institución, que han persistido hasta el final en esta tarea a veces tan difícil de recoger los datos. Sin ellas esta tesis no existiría.

A mis compañeros del Doctorado, Guillermo, Juan, Roxana y Susana. De todos ellos he aprendido también. Un agradecimiento muy especial a Carmen por tantas y tantas horas compartidas. ¡Gracias, Amiga, por apoyarme en los momentos difíciles y por compartir los buenos!

Al Prof. João Pires, Prof. Carlos Maia, Prof. Ana Maria Vaz, Prof. Helena Ribeiro, Prof. Paula Sapeta del Instituto Politécnico de Castelo Branco que, siendo expertos en el tema de la Salud, me han facultado material sobre el tema de la Calidad de Vida e incentivado a seguir. Gracias por sus palabras de ánimo.

A mi madre, mi amiga y compañera de todas las horas, que siempre ha creído que este momento llegaría y que, de tantas maneras ha contribuido para ello. Ella sabe que lo que siento no cabe en papel.

A mis abuelos que me han visto empezar este proyecto pero ya no han podido disfrutar de la alegría de este momento. Su amor y dedicación siempre han sido una fundación fuerte en mi vida.

A Rui por compartir conmigo la tarea, a veces dura, de “estar en tesis”, incentivándome y apoyándome en cada momento. Gracias por formar parte de mi vida.

No puedo dejar de expresar mi profunda gratitud a cada uno de los pacientes que, luchando contra el cáncer y atravesando un momento tan difícil de sus vidas, han aceptado en colaborar en este estudio dejándonos participar en su experiencia. ¡Ellos son el motivo de este trabajo!

*A todos y cada uno... MIL GRACIAS!*



---

# ÍNDICE

---

# ÍNDICE

## PARTE I – MARCO TEÓRICO

<b>INTRODUCCIÓN</b> .....	<b>2</b>
<b>1. CALIDAD DE VIDA RELACIONADA CON LA SALUD</b> .....	<b>7</b>
1.1. EL CONCEPTO DE CVRS.....	7
1.2. ¿PORQUÉ MEDIR CALIDAD DE VIDA RELACIONADA CON LA SALUD?.....	16
1.3. ¿CÓMO MEDIR CVRS?.....	22
1.4. ¿QUIÉN DEBE MEDIR CVRS? .....	26
<b>2. INSTRUMENTOS PARA MEDIR CVRS</b> .....	<b>33</b>
2.1. MEDIDAS DE CVRS.....	33
2.2. CARACTERÍSTICAS DE LOS INSTRUMENTOS PARA MEDIR CVRS .....	35
2.2.1. <i>Fiabilidad</i> .....	36
2.2.2. <i>Validez</i> .....	38
2.2.3. <i>Viabilidad</i> .....	39
2.2.4. <i>Sensibilidad</i> .....	40
2.3. CONSTRUCCIÓN DE INSTRUMENTOS PARA MEDIR CVRS .....	44
2.4. INSTRUMENTOS, MEDIDAS Y ESCALAS UTILIZADOS PARA MEDIR CVRS.....	47
2.4.1. <i>Instrumentos Genéricos VS Instrumentos Específicos</i> .....	47
2.4.1.1. <i>Medical Outcomes Study 36 – Item Short Form (SF-36)</i> .....	52
2.4.1.2. <i>Rotterdam Symptom Check List (RSCL)</i> .....	55
2.4.2. <i>Ítems y Escalas de Respuesta</i> .....	61
2.5. EL PROBLEMA DE LOS DATOS FALTANTES.....	68
<b>3. EL FENÓMENO DE <i>RESPONSE SHIFT</i></b> .....	<b>75</b>
3.1. CAMBIO EN CALIDAD DE VIDA RELACIONADA CON LA SALUD.....	75
3.2. <i>RESPONSE SHIFT</i> .....	81
3.2.1. <i>¿Qué es Response Shift?</i> .....	81
3.2.2. <i>El Modelo Teórico del Response Shift</i> .....	86
3.2.3. <i>Estudios que evidencian Response Shift</i> .....	89
3.2.4. <i>El Fenómeno de la Paradoja de la Discapacidad</i> .....	92
3.2.5. <i>Response Shift y Comparación Social</i> .....	94
3.3. ¿CÓMO MEDIR <i>RESPONSE SHIFT</i> ?.....	98
3.3.1. <i>Métodos Individualizados</i> .....	100
3.3.2. <i>Utilidades y Preferencias</i> .....	102
3.3.3. <i>Métodos de Aproximación por Comparaciones Sucesivas</i> .....	104
3.4. EL MÍNIMO CAMBIO CLINICAMENTE IMPORTANTE Y <i>RESPONSE SHIFT</i> .....	104
<b>4. MÉTODOS ESTADÍSTICOS PARA ANALIZAR DATOS DE CVRS Y DETECTAR <i>RESPONSE SHIFT</i></b> .....	<b>111</b>
4.1. DATOS DE CVRS.....	112
4.2. MÉTODOS DE ANÁLISIS TRADICIONALES .....	115

4.2.1. <i>El Then-Test</i> .....	117
4.3. MÉTODOS MULTIVARIANTES PARA DETECTAR CAMBIO EN CVRS .....	120
4.3.1. <i>El Método de Ahmavaara</i> .....	120
4.3.2. <i>La Propuesta de Vicente-Galindo</i> .....	123
4.3.2.1. <i>El Método STATIS</i> .....	123
4.3.2.2. <i>El Procedimiento de Krzanowski</i> .....	128
4.3.2.3. <i>STATIS Dual y Metacomponentes en la Evaluación del Cambio</i> .....	132
4.4. EL COEFICIENTE RV PARA DETECTAR CAMBIO .....	136
4.5. EL MODELO DE ZWINDERMAN .....	139
4.5.1. <i>Una Contribución al Modelo de Zwinderman utilizando Técnicas Factoriales</i> .....	144
4.6. MODELOS DE ECUACIONES ESTRUCTURALES .....	151
4.6.1. <i>Modelos de Curvas de Crecimiento Latente</i> .....	169
4.6.1.1. <i>El Modelo de Zee</i> .....	176
4.6.2. <i>El Procedimiento de Oort</i> .....	183
4.7. EL MODELO DE MAYO ET AL.....	200

## PARTE II – APLICACIÓN A DATOS DE CALIDAD DE VIDA EN PACIENTES ONCOLÓGICOS

5. CVRS EN PACIENTES ONCOLÓGICOS.....	206
5.1. EL PROBLEMA DEL CÁNCER.....	206
5.2. CVRS EN PACIENTES ONCOLÓGICOS .....	214
5.2.1. <i>Los Tratamientos y Síntomas Asociados</i> .....	217
5.2.2. <i>La Adaptación al Cáncer</i> .....	222
5.2.3. <i>La Comunicación Médico-Paciente</i> .....	225
5.2.4. <i>El Apoyo Psicosocial</i> .....	227
5.2.5. <i>Un Relato en Primera Persona</i> .....	228
6. INSTRUMENTOS, MÉTODO Y MUESTRA.....	233
6.1. INSTRUMENTOS UTILIZADOS.....	233
6.1.1. <i>Medical Outcomes Study 36 – Item Short Form (SF-36)</i> .....	234
6.1.2. <i>Rotterdam Symptom Check List (RSCL)</i> .....	237
6.2. OBJETIVO Y MÉTODO .....	240
6.3. DESCRIPCIÓN DE LA MUESTRA .....	243
6.3.1. <i>Variables Sociodemográficas</i> .....	244
6.3.2. <i>Variables Clínicas</i> .....	244
7. RESULTADOS OBTENIDOS.....	248
7.1. CALIDAD DE VIDA DE LOS PACIENTES EVALUADA CON UN INSTRUMENTO GENÉRICO.....	248
7.1.1. <i>Estructuras Latentes del SF-36 y Consistencia Interna</i> .....	248
7.1.2. <i>Puntuaciones Obtenidas en las Subescalas del SF-36</i> .....	251
7.1.3. <i>Comparación de las puntuaciones obtenidas en el Then-Test y Post-Test con el Pre-Test para el SF-36</i> .....	255
7.2. CALIDAD DE VIDA DE LOS PACIENTES EVALUADA CON UN INSTRUMENTO ESPECÍFICO .....	262
7.2.1. <i>Estructuras Latentes del RSCL y Consistencia Interna</i> .....	262

7.2.2. Puntuaciones obtenidas en la subescalas del RSCL .....	263
7.2.3. Comparación de las puntuaciones obtenidas en el Then-Test y Post-Test con el Pre-Test para el RSCL.....	266
7.3. APLICACIÓN DEL MODELO DE ZWINDERMAN A LOS DATOS OBTENIDOS MEDIANTE EL SF-36 .....	271
7.4. ANÁLISIS DE LA ESTRUCTURA FACTORIAL DEL INSTRUMENTO GENÉRICO PARA LOS DIFERENTES TIEMPOS .....	273
7.5. ANÁLISIS DEL SF-36 MEDIANTE EL STATIS DUAL .....	276
7.6. ANÁLISIS DEL SF-36 MEDIANTE EL METABILOT.....	279
7.7. ANÁLISIS DEL SF-36 MEDIANTE UN MODELO DE CURVAS DE CRECIMIENTO LATENTE .....	282
7.8. ANÁLISIS DEL SF-36 MEDIANTE EL MODELO DE ECUACIONES ESTRUCTURALES ..	286
<b>CONCLUSIONES .....</b>	<b>314</b>
<b>REFERENCIAS .....</b>	<b>318</b>
<b>ANEXO I – SF-36 .....</b>	<b>343</b>
<b>ANEXO II – RSCL .....</b>	<b>346</b>
<b>ANEXO III – EL MÉTODO STATIS.....</b>	<b>349</b>
<b>ANEXO IV – EL MODELO DE ZWINDERMAN.....</b>	<b>356</b>
<b>ANEXO V – CUESTIONARIO SOCIODEMOGRÁFICO.....</b>	<b>366</b>
<b>ANEXO VI – CONSENTIMIENTO INFORMADO .....</b>	<b>370</b>

---

# INTRODUCCIÓN

---

## INTRODUCCIÓN

El discurso efusivo del presidente de los EEUU Lyndon Johnson, en los años 60, afirmando que la concretización de los objetivos de una Sociedad no puede medirse por el balance de las entidades bancarias sino por la calidad de vida proporcionada a las personas, ha supuesto un avance importante en el concepto de Calidad de Vida. Hasta entonces los indicadores de la calidad de vida de las poblaciones estaban basados exclusivamente en indicadores sociales como el Producto Interno Bruto, la Tasa de Desempleo o el coste de las habitaciones. Desde entonces ha surgido una nueva conciencia de que, aunque tales indicadores sean importantes, no reflejaban necesariamente el nivel de calidad de vida individual de las personas.

En una sociedad moderna, es absolutamente imprescindible que se piense en la calidad de vida de los individuos más allá de la calidad de vida global de la población, medida a través de indicadores sociales. Progresivamente, la Calidad de Vida ha dejado de constituir un tema de exclusivo interés en el dominio político y económico y ha pasado a centrarse más en el individuo y sus valores personales, extendiéndose hoy a todos los dominios de la Ciencia.

En el dominio de la Salud, el tema de la Calidad de Vida ha cobrado especial importancia. El número de trabajos publicados en este ámbito ha registrado, en los últimos años, un crecimiento exponencial. La preocupación por la calidad de vida de los pacientes ha recibido especial atención en el dominio de la Oncología: más que extender los años de vida de los pacientes, importa conferir calidad a esos años.

El impacto producido por la enfermedad oncológica es brutal y tiene repercusiones en todas las áreas de la vida de cualquiera que se ve sentado delante

del médico y escucha la palabra “cáncer”. El impacto producido por el diagnóstico, el estigma todavía asociado al cáncer, la cirugía que elimina permanente una parte del cuerpo, los efectos de los tratamientos, etc. pueden constituir una experiencia aterradora y traumática. Investigaciones recientes, ponen de manifiesto que tales eventos pueden alterar la forma en la que cada individuo percibe y evalúa su calidad de vida. En la literatura, este fenómeno es conocido como *response shift* y se define como un cambio en la autoevaluación de un individuo resultante de una recalibración de la escala de medida o debido a cambios en sus prioridades o incluso en el propio concepto de Calidad de Vida.

Desde el punto de vista estadístico el problema del análisis de datos de Calidad de Vida no es trivial. La aplicación de las herramientas estadísticas convencionales a esto tipo de datos puede volverse problemática si tuviéremos en cuenta que una escala de medida ha sido recalibrada. En la literatura son escasas las propuestas metodológicas que tratan de aportar algún tipo de solución al respecto.

Por todo esto, hemos considerado de extrema relevancia el desarrollo de la investigación **“Evaluación del Cambio en Pacientes Oncológicos: Un Enfoque Estadístico”**. El objetivo del presente trabajo está centrado en las técnicas estadísticas aplicadas al análisis de datos de Calidad de Vida. En particular se pretende entender cómo percibe el paciente oncológico su calidad de vida y cómo evoluciona esta percepción a lo largo de todo el complejo proceso de diagnóstico y tratamiento de la enfermedad. Con este propósito, hemos desarrollado un estudio con una muestra de pacientes oncológicos pertenecientes a cinco hospitales portugueses. El proceso de recogida de estos datos ha sido extenso y arduo en lo que respecta a la obtención de los permisos necesarios y a los muchos días pasados en el hospital, sin embargo, extremadamente gratificante y compensador por el contacto con cada paciente. La lucha de estos pacientes contra la

enfermedad, su fuerza, su empeño, sus ganas de vivir, su sonrisa, han sido también nuestra fuerza y aliento para seguir.

La **primera parte** de esta tesis contiene cuatro capítulos que consisten en el desarrollo de los contenidos teóricos subyacentes a la problemática en estudio. La **segunda parte** contiene tres capítulos dedicados a la aplicación de los conceptos presentados a la evaluación de Calidad de Vida en pacientes oncológicos.

En particular, en el **primer capítulo** se aborda el concepto de Calidad de Vida Relacionada con Salud, la importancia de medirla, cómo y quién debe medirla. El **segundo capítulo** está dedicado a los instrumentos para medir Calidad de Vida Relacionada con Salud y sus características. En el **tercer capítulo** se aborda el fenómeno de *response shift* en Calidad de Vida y cómo detectarlo. En el **cuarto capítulo** se presentan y discuten algunos métodos estadísticos para analizar datos de Calidad de Vida y, en particular, detectar *response shift*. El **quinto capítulo** aborda la problemática específica de la Calidad de Vida en pacientes oncológicos. En el **sexto capítulo** se presentan los instrumentos, la metodología seguida y la muestra en estudio. Finalmente, el **séptimo capítulo** está dedicado a la presentación de los resultados obtenidos mediante la aplicación de varias herramientas estadísticas.

La forma en la que hemos presentado el contenido de esta tesis obedece a su carácter híbrido: Métodos Estadísticos en Calidad de Vida Relacionada con la Salud. Pretendemos que sea útil a los especialistas en Calidad de Vida y, por tanto, hemos limitado el desarrollo de los métodos estadísticos al mínimo. Pero los estadísticos interesados en todos los detalles técnicos de los métodos, los encontrarán en los correspondientes anexos.



**PARTE I**  
**MARCO TEÓRICO**

## **CAPÍTULO I**

---

# **CALIDAD DE VIDA RELACIONADA CON LA SALUD**

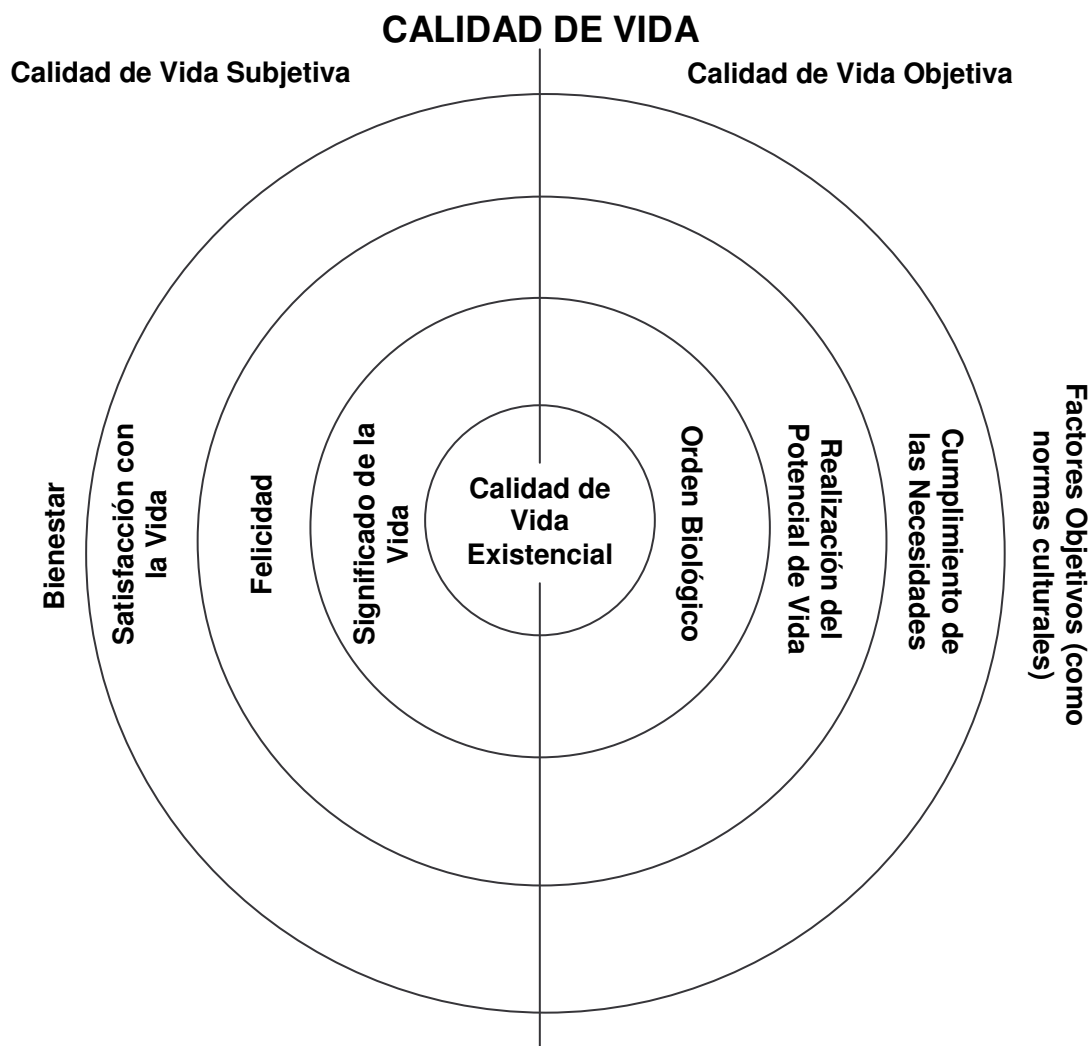
---

## 1. CALIDAD DE VIDA RELACIONADA CON LA SALUD

### 1.1. EL CONCEPTO DE CVRS

No existe una definición única y cierta de Calidad de Vida, sino un sin número de definiciones y aproximaciones posibles. Fayers & Machin (2000) mencionan una de las primeras referencias a lo que se puede entender por Calidad de Vida, debida a Aristóteles (384-322 A.C.) que escribe: *“Tanto la persona más modesta como la más refinada entiende ‘vida buena’ o ‘estar bien’ como lo mismo que ‘estar feliz’.* Sin embargo, lo que es entendido como felicidad es discutible. Unos dicen una cosa y otros otra y la misma persona dice cosas diferentes en tiempos diferentes: cuando está enferma piensa que la salud es la felicidad, cuando es pobre felicidad es riqueza.” Muchos siglos después, otros autores siguen entendiendo “Calidad de Vida” como “una vida buena”, que es lo mismo que vivir la vida con una alta calidad, subrayando además que el concepto de “vida buena” está íntimamente relacionado con la realidad cultural en cuestión (Ventegodt *et al*, 2003; Camfield & Skevington, 2007). Según la Teoría Integrante de la Calidad de Vida de Ventegodt *et al* (2003), esta es la combinación de 3 importantes aspectos (Fig.1):

- 1) La **Calidad de Vida Subjetiva** es el sentimiento que cada individuo posee sobre si tiene una buena vida o no;
- 2) La **Calidad de Vida Existencial** tiene que ver con la percepción que cada uno posee sobre si tiene una vida buena a un nivel más profundo. Se trata de vivir en armonía con su propia naturaleza, por ejemplo de acuerdo con determinados ideales espirituales o religiosos;
- 3) La **Calidad de Vida Objetiva** se relaciona con la forma en que la Calidad de Vida de uno es percibida por las otras personas. Esa percepción está influida por la cultura en la cual cada individuo vive.



**Fig.1.** La Teoría Integradora de la Calidad de Vida Ventegodt *et al* (2003): el individuo puede compararse con una manzana verde con manchas rojas (las dimensiones subjetiva y objetiva de la calidad de vida, respectivamente, en la superficie de la existencia de un individuo) con un núcleo oculto (interior profundo de la humanidad). Por otra parte, la imagen de la humanidad es comparable una cebolla, con un número de capas entre la superficie y el núcleo: entre la esencia de la vida y la superficie está el bienestar, la satisfacción, la armonía y el significado profundo de la vida.

Hasta el siglo XX, la Calidad de Vida ha sido muy poco mencionada. En 1948, la Organización Mundial de Salud (OMS) define Salud como algo próximo al concepto de Calidad de Vida, centrándose en tres dimensiones: física, mental y social. La OMS pasa entonces a entender Salud no solamente como la ausencia de enfermedad sino también como un estadio de completo bienestar físico, mental y social (OMS, 1948). En el cotidiano, el concepto de Calidad de Vida se ha generalizado después de la 2ª Guerra Mundial cuando en la década de los 60 se

asiste a un cambio de los valores sociales que enfatizaba las necesidades sociales y psicológicas más que los intereses materiales (Frisch, 2000).

Citando a Vicente-Galindo (2007), en 1974 se incluye el término Calidad de Vida en la revista norte-americana *Social Indicators Research* y en 1979 en la publicación *Sociological Abstracts*. En el contexto de la Salud, este concepto se convirtió desde 1977 en categoría de búsqueda en el *Index Medicus* y en palabra clave en la base *MEDLINE*. Estos hechos contribuyeron a su difusión teórica y metodológica y ya en la década de los 80, se produjo el despegue definitivo de la investigación en torno a este término. Paradójicamente, transcurridos casi 30 años, aún no existe consenso ni en la definición del constructo ni en la forma más adecuada de medida.

Aunque, actualmente, el concepto de Calidad de Vida se utiliza en contextos muy variados, en general se refiere a nociones como riqueza, ocio, autonomía, libertad, o sea, todo lo que proporcione un cotidiano agradable. La OMS (1993), define Calidad de Vida como *“la percepción que el individuo tiene de su lugar en la vida, en el contexto de la cultura y del sistema de valores en los cuales vive, en relación con sus objetivos, sus deseos, sus normas y sus inquietudes. Es un concepto muy amplio que puede ser influenciado de manera compleja por la salud física del individuo, por su estado psicológico y por su nivel de independencia, sus relaciones sociales y sus relaciones con elementos esenciales de su medio”*. Se destaca la idea de que la Calidad de Vida está influenciada por el entorno físico y social así como por las reacciones del individuo a ese entorno.

En la misma línea, Levine & Croog (1984) y Camfield & Skevington (2007) definen Calidad de Vida en base a la interacción del individuo con el medio en sus diversas vertientes social, física, emocional e intelectual y a la percepción subjetiva que engloba el sentido general de satisfacción del individuo y de su propia salud. Thoma *et al* (2008) definen Calidad de Vida como la adecuación de las circunstancias materiales y sentimientos de cada individuo a sus circunstancias.

Según Calman (1984) la percepción de la Calidad de Vida está relacionada con el estilo de vida actual, con las experiencias pasadas, con los objetivos y ambiciones. Precisamente por eso, algunos autores entienden que la Calidad de Vida solamente puede ser medida y descrita en términos individuales.

Wu (2009) relaciona Calidad de Vida con la discrepancia percibida entre el estado de vida actual y el deseado. Cuanto mayor es la discrepancia, menor es la Calidad de Vida. Otro aspecto importante enfocado por Wu (*op.cit.*) en la definición de Calidad de Vida es la percepción que el individuo tiene de las diferentes dimensiones de su vida. Según este autor se puede mejorar la Calidad de Vida enfocando los dominios en que el individuo se encuentra mejor y quitando importancia a los otros. Independientemente del contexto en cuestión, está claro que Calidad de Vida asume diferentes significados para diferentes personas y para la misma persona en distintos contextos (Rapkin & Schwartz, 2004).

En la Sociedad Occidental el concepto de Calidad de Vida en el contexto de la Salud pasa a tener gran importancia cuando la esperanza de vida aumenta y el papel de la medicina ya no es sólo proporcionar muchos años de vida a las personas, sino, sobre todo aportar calidad a los años vividos (Badia & Carné, 1998). El término **Calidad de Vida Relacionada con la Salud (CVRS)** ha sido propuesto por Kaplan & Bush (1982), con el objetivo de enfocar la importancia de los efectos de la salud en la percepción individual del bienestar, distinguiendo tales aspectos de otros como la satisfacción con el trabajo o con factores ambientales. Para Testa & Simonson (1996), el término CVRS se refiere a los dominios físico, psicológico y social de la salud, vistos como áreas distintas, e influenciados por las experiencias personales, convicciones, expectativas y percepciones.

Interesa también pensar sobre el concepto de Salud. Brody & Sobel (1979) definen Salud como la capacidad de un sistema para contestar de modo ajustado a

los desafíos presentados por el ambiente. La Salud es pues, entendida, como un proceso positivo y no solamente como la ausencia de síntomas de enfermedad. Por otra parte, la Enfermedad se relaciona con la incapacidad de contestar adecuadamente a los desafíos ambientales, que trae como consecuencia el desequilibrio del sistema. Ribeiro (2005) debate la cuestión de cómo evaluar la Salud subrayando que, si la forma de diagnosticar, tratar y evaluar enfermedades están definidas en los protocolos, la forma de evaluar Salud no está definida. Si las variables indicadoras de Enfermedad están relativamente claras, las que son susceptibles de indicar Salud, no lo están todavía. La Salud tiende a ser evaluada a partir de la percepción del propio sujeto. Aunque esta forma de evaluación no reúna todavía consenso, distintas investigaciones evidencian que la autoevaluación del estado de salud es un predictor más fiable de la supervivencia, que los registros médicos o la descripción de las enfermedades (Idler & Benyamini, 1997).

En un paciente, la Calidad de Vida es un concepto relativo que se refiere a su nivel de satisfacción en función de sus posibilidades actuales, comparadas con aquellas que cree que son posibles o ideales (Cella & Cherin, 1988; Wu, 2009). El concepto de CVRS es necesariamente diferente del concepto de Calidad de Vida global. Postulart & Adang (2000) definen CVRS como el efecto funcional de una enfermedad, incluyendo la terapia sobre un paciente y la forma cómo es percibida por el paciente. El término "estado de salud" es utilizado para describir la salud de un individuo en un instante particular, mientras CVRS está relacionada con la preferencia asignada a un estado de salud. En el ámbito de los ensayos clínicos, Fayers & Machin (2000), destacan que, CVRS se refiere a la evaluación de los aspectos afectados por la enfermedad o su tratamiento.

Muldoon *et al* (1998) subrayan que, aunque no existe una definición exacta, que reúna el consenso de todos, la investigación en el área de la CVRS busca, sobre todo, dos tipos de información: el estado funcional del individuo y la evaluación del propio paciente sobre la forma en la cual su estado de salud afecta a su calidad

de vida. Según este autor, se puede entonces definir operacionalmente CVRS como el nivel funcional del individuo y el estado de salud o bienestar que este individuo percibe. Si por una parte evaluar la capacidad de alguien para desempeñar determinadas tareas o actividades es algo extremadamente objetivo, por otra parte, pedirle que evalúe los efectos de su estado de salud en su bienestar personal, es algo manifiestamente subjetivo. Por ejemplo, la pregunta “¿Puede llevar dos bolsas de compras a lo largo de 20 metros?” busca información sobre un comportamiento, luego tiene una respuesta objetiva. Por otra parte, la pregunta “¿Su salud interfiere con su disfrute de la vida?” implica una evaluación subjetiva. Esta evaluación subjetiva del estado de salud es extremadamente importante dado que, en parte, la calidad de vida es independiente del estado de salud. En cierta medida, la CVRS puede entenderse como un reflejo de la forma como los pacientes perciben su estado de salud y reaccionan frente a otros aspectos no médicos de sus vidas.

Según Pimentel (2004), la CVRS es la parte de la Calidad de Vida general del individuo constituida por los componentes relacionados con la Salud (enfermedad y terapia), debiendo incluir como mínimo los siguientes aspectos:

- síntomas producidos por la enfermedad o tratamiento;
- funcionalidad física;
- aspectos psicológicos;
- aspectos sociales, familiares, laborales o económicos.

Todos estos aspectos están interrelacionados y se influyen mutuamente. En la misma línea, Wilson & Clearly (1995) proponen un modelo conceptual que establece relaciones entre varios aspectos relacionados con la salud, identificando algunos factores que influyen en la evaluación global de la CVRS:

- factores biológicos y psicológicos que tienen que ver con el verdadero estado del individuo;



- síntomas que se definen como la percepción que el paciente tiene sobre su estado físico, emocional o cognitivo;
- el estado funcional que incluye las dimensiones física, psicológica y social;
- la percepción del estado de salud general, que incluye la evaluación del paciente sobre su estado anterior y actual de salud, su perspectiva futura y preocupaciones con el estado de salud.

Tras el análisis de un conjunto de artículos sobre CVRS, Badia & Carné (1998) destacan tres ideas comunes:

- la importancia del estado funcional del paciente (tanto físico, como social o mental) en el contexto de un concepto como el de Calidad de Vida, que se considera multidimensional;
- el hecho de que la medición del estado de salud se considere un fenómeno innegablemente subjetivo;
- la necesidad de obtener un valor numérico que represente la preferencia del paciente por un determinado estado de salud.

Aunque cada paciente posea su escala de valores, es posible encontrar elementos comunes y construir un concepto operacional de CVRS. Según Calman (1984), el impacto de la enfermedad puede afectar al individuo de diversas formas, dependiendo mucho de la situación en que este se encuentre. Algunos individuos consideran que poseen una elevada Calidad de Vida aunque tengan problemas graves, lo que se debe a las bajas expectativas. Burgess *et al* (2000) también consideran que la CVRS no puede ser considerada como un constructo unitario y destacan dos factores importantes en la obtención de una medida de CVRS:

- la componente física que está relacionada con las limitaciones funcionales;

- el impacto de la salud en el estado psicológico que debe ser entendida como una medida de la CVRS psicológica.

Por otra parte, hay que tener en cuenta que los aspectos psicológicos y físicos son influenciados por innumerables factores.

Como hemos dicho, durante los últimos años, varias definiciones de CVRS han sido propuestas y no existe todavía una definición formal que sea unánimemente aceptada. Sin embargo, existe un acuerdo general en cuanto al hecho de que se trata de un concepto multidimensional, subjetivo, que enfoca el impacto de una enfermedad y su tratamiento, así como el bienestar del individuo y envuelve tres dimensiones importantes: física, psicológica y social, pudiendo todavía incluirse una amplia gama de aspectos como por ejemplo el funcionamiento cognitivo o la espiritualidad (Boini *et al*, 2004; Camfield & Skevington, 2007; Thoma *et al*, 2008). Naturalmente, en un individuo enfermo, la percepción del estado general de salud es más importante que todas las demás vivencias. Resultados de un estudio sobre factores predictivos de la satisfacción con el estado de salud han confirmado esta idea, revelando que la percepción individual del estado de salud es el factor más importante ya que explicaba un 53% de la variabilidad en las respuestas sobre felicidad, el 68% de las respuestas sobre satisfacción global con la vida y el 63% de la variabilidad de las respuestas relativas a la satisfacción con la Calidad de Vida global (Michalos *et al*, 2000).

### ***La CVRS Física y Psicológica***

La relación entre CVRS física y psicológica es compleja. Es bien sabido que la salud física influencia el bienestar psicológico y que, a su vez, el estado psicológico ejerce una fuerte influencia sobre la percepción de bienestar físico. Burgess *et al* (2000) destacan la influencia de la componente psicológica en la CVRS y refieren que, algunas características individuales de la personalidad como la Afectividad Negativa o el Neuroticismo influyen ampliamente la forma como

las personas se sienten. También los mecanismos de *coping*, entendidos como el conjunto de estrategias que el individuo utiliza para adaptarse a circunstancias adversas o estresantes (Suls *et al*, 1996), influyen la componente psicológica de la CVRS. Tales estrategias pueden modelar el impacto de agentes estresores y actúan como un parachoques protector o amplificador de las consecuencias negativas. En este proceso, también asumen particular importancia otros rasgos de la personalidad de cada individuo como la Extroversión, que influyen la forma como se manifiesta el Neuroticismo.

Wrosch & Scheier (2003) argumentan que la personalidad afecta la CVRS. La personalidad influye la forma como cada persona reacciona delante de las situaciones críticas de la vida. Estos autores atribuyen al optimismo y al ajuste de objetivos un papel importante en la promoción de la CVRS. En la misma línea, estudios recientes sobre los niveles de fatiga en pacientes oncológicos, atribuyen al optimismo un papel importante como facilitador del bienestar subjetivo y de la salud a través de las estrategias de *coping* (Westerman *et al*, 2007a; Ridder *et al*, 2008; Santana *et al*, 2009). En el fondo, el paciente intenta distanciarse del estereotipo del paciente con cáncer. Wrosch & Scheier (*op.cit.*) concluyen que, las diferencias individuales en la capacidad de cada persona para adaptar los objetivos inalcanzables está directamente relacionada con una alta calidad de vida.

Schwartz *et al* (2007) entienden CVRS como el estado de salud percibido por un individuo, grupo o población que refleja el grado según el cual cada persona es capaz de participar física, emocional y socialmente, con o sin asistencia. La CVRS abarca una amplia cantidad de aspectos de la experiencia de vida del individuo, pasando por el estado físico, emocional, adaptabilidad y experiencia existencial. Estos autores distinguen los conceptos de "salud" y "discapacidad" ya que el entorno donde vive una persona con discapacidad puede contribuir para muchas de las limitaciones relacionadas con la discapacidad que experimentan. Una persona en una silla de ruedas se encuentra limitada en su capacidad para caminar 100 metros, sin embargo, puede moverse utilizando su silla de ruedas. Por eso

puede percibir su limitación a nivel del funcionamiento físico, no reflejando necesariamente una salud o CVRS pobre.

En realidad, no existe una definición única de CVRS, sin embargo, podemos describirla, de forma funcional, como la percepción de los pacientes sobre sus capacidades en cuatro importantes dimensiones: bienestar físico y actividades cotidianas, bienestar psicológico, relaciones sociales y síntomas. En la última década se ha asistido a una intensa actividad de investigación en el dominio de la CVRS que ha originado diferentes concepciones y modelos de CVRS. No obstante de las muchas indefiniciones e incertidumbres que todavía persisten, apuntamos como consensos entre los diferentes autores algunos aspectos importantes:

- La CVRS se define como un concepto multidimensional, subjetivo e individual que traduce el bienestar subjetivo del paciente en las vertientes física, psicológica y social.
- Existen instrumentos de medida de CVRS que posibilitan su evaluación de forma estandarizada.
- La CVRS debe ser considerada como el objetivo principal en las enfermedades crónicas e incurables.

## **1.2. ¿PORQUÉ MEDIR CALIDAD DE VIDA RELACIONADA CON LA SALUD?**

En los países desarrollados, los principales problemas de salud ya no son las enfermedades infecciosas, sino las enfermedades crónicas que a veces persisten durante muchos años y requieren terapias durante periodos largos. Los avances científicos han hecho en muchos casos posible obtener la cura para muchas enfermedades y, en otros, extender la vida mucho más allá de lo que se esperaría por la historia natural de la enfermedad. Merecen destacarse las enfermedades oncológicas. Para muchos pacientes, el cáncer ha dejado de ser una enfermedad

con un desenlace rápido y se ha vuelto una enfermedad crónica que perdura meses o años, con tratamientos complejos y muchas veces tóxicos.

Desde que la OMS (1948) ha definido Salud no solamente como la ausencia de enfermedad sino como un estadio de bienestar físico, mental y social, el tema de la Calidad de Vida ha pasado a asumir un grado de importancia continuamente creciente, tanto en la práctica como en la investigación relacionada con los cuidados de salud. Una nueva medicación puede prolongar la vida de un paciente, sin embargo, existe una necesidad creciente de comprobar que también mejora su CVRS. Más que sumar años a la vida, el desafío que se coloca en la actualidad es sumar vida a los años (Boini *et al*, 2004).

Efficace *et al* (2007) subrayan que, desde un punto de vista psicológico y antropológico interesa entender la percepción que el paciente tiene de su enfermedad y no solamente la enfermedad que realmente tiene. En general, los médicos tienden a centrarse en el proceso de la enfermedad, intentando resolverlo y dando poca atención a la percepción que el paciente posee de la enfermedad. También Pimentel (2004) destaca la necesidad de medir CVRS, especialmente en pacientes oncológicos, para obtener la percepción del impacto de la enfermedad y de los tratamientos en el día a día del paciente. Los eventos a que un paciente está expuesto debido a su enfermedad (señales, síntomas, efectos colaterales, experiencias traumáticas, beneficios de la terapia) forman un tejido complejo que va siendo integrado por el paciente. La importancia que este atribuye a cada uno de estos eventos influencia los dominios físico y psicológico de la CVRS, repercutiéndose tales influencias en los dominios social y económico. Un mejor conocimiento y evaluación de la CVRS puede proporcionar datos para mejorar procedimientos terapéuticos a utilizar en otros pacientes. Al optar entre dos tipos de terapia con resultados en términos de cura o supervivencia semejantes, se opta por el que proporciona mejor CVRS (Thoma *et al*, 2008).

La obtención de medidas de CVRS posibilita la evaluación del impacto de una enfermedad crónica, siendo un factor importante en el momento de tomar decisiones clínicas (Guyatt *et al*, 1993; Postulart & Adang, 2000; Goodwin *et al*, 2003; Efficace *et al*, 2004, 2007; Pucciarelli *et al*, 2008; Thoma *et al*, 2008; Westerman *et al*, 2008). Más que alargar el tiempo de vida de los pacientes, la preocupación actual se centra en incrementar la CVRS durante ese tiempo (Froberg & Kane, 1989d). Este hecho, asume una importancia particular en pacientes con enfermedades crónicas o discapacitados, sin esperanza de cura y en condiciones que tienen impacto en su bienestar físico, psicológico y social (Addington-Hall & Kalra, 2001). La problemática de la CVRS es cada vez más vista como un objetivo en sí misma, tan importante o más que los cuidados físicos. La evaluación de la CVRS permite analizar la percepción que el paciente posee sobre los efectos somáticos y psicosociales inducidos por la terapia.

Testa & Simonson (1996) destacan el crecimiento exponencial registrado en la utilización de las técnicas de evaluación de la calidad de vida en la investigación clínica señalando que en 1973 únicamente cinco artículos mencionaban “Calidad de Vida” como palabra-clave en la base de datos *Medline*, mientras en los cinco años subsecuentes se han registrado 195, 273, 490 y 1252, respectivamente. Efficace *et al* (2007) han identificado 159 ensayos clínicos randomizados con 58.635 pacientes entre 1990 y 2004. Los autores destacan que, de los estudios clasificados como probablemente robustos, el 64,3% han sido publicados después del año 2000, registrándose una mejoría significativa en lo que respecta a la calidad de los ensayos clínicos en el ámbito de CVRS desde 1990.

En la secuencia de una búsqueda relativa a ensayos clínicos en el ámbito del cáncer de mama, Goodwin *et al* (2003) afirman que la medición de la CVRS en pacientes de cáncer en general constituye el mayor foco de investigación a lo largo del último cuarto de siglo. Este interés en la CVRS queda bien reflejado a través del desarrollo y validación de diversos instrumentos tanto en Europa como Norte

América y por su inclusión en los ensayos clínicos con diversos grupos de pacientes. Sin duda, una de las razones que ha motivado el rápido desarrollo de medidas de CVRS, ha sido el reconocimiento creciente de la importancia de entender el impacto de intervenciones relacionadas con cuidados de salud en la vida de los pacientes y no únicamente en sus cuerpos. Con este objetivo, la *European Organization Research for Treatment of Cancer* (EORTC) desarrolló, en 1986, un cuestionario de CVRS para pacientes oncológicos. En esa misma época, en Francia, la Sociedad Francesa de Oncología, inicia la evaluación sistemática de la CVRS. En EEUU, en 1990, el *Nacional Cancer Institute* (NCI), constituye un grupo de trabajo para estudiar la implementación de la evaluación de la CVRS en los ensayos clínicos de Oncología. En Inglaterra, el *Medical Research Council* ha reconocido la necesidad de integrar la evaluación de la CVRS en los ensayos clínicos. Actualmente, en EEUU existen dos organismos direccionados para el estudio de la CVRS en Oncología: la *American Association for Clinical Oncology* (ASCO) e el *Nacional Cancer Institute* (NCI). Se trata de un dominio relativamente reciente en la práctica clínica, inicialmente desarrollado en el contexto de Oncología pero que es, en la actualidad, objeto de investigación en muchas otras áreas del conocimiento. En realidad, la evaluación de la CVRS es hoy un componente integral de la investigación centrada en el paciente.

Diversos autores afirman que la simple evaluación de la CVRS puede tener un efecto terapéutico, no solamente porque permite mejorar la satisfacción de los pacientes con los cuidados prestados sino también porque, desde el punto de vista psicológico, todos los procedimientos que refuercen la “capacidad de escuchar”, tienen un efecto positivo (Zittoun, 1986). Además de eso, la utilización regular de instrumentos específicamente concebidos para evaluar CVRS en la práctica clínica permite a los pacientes describir mejor los síntomas y los efectos adversos del tratamiento, compensando eventuales insuficiencias en la valoración de estos síntomas. Por otra parte, la utilización de información estandarizada sobre CVRS,

facilita la comunicación entre médicos y enfermeros, pudiendo ser el primer paso para implementar la utilización de medidas de CVRS en el día a día ya que el contenido y calidad de estas informaciones puede influenciar decisiones relativas a la terapia.

Aunque el interés de diversos organismos, de los médicos en general y de los oncólogos en particular por la CVRS haya aumentado significativamente en los últimos tiempos, todavía no se acepta definitivamente como componente importante en la práctica clínica. En 1998, un estudio ha revelado que el 50% de un grupo de 154 oncólogos afirmaba utilizar la evaluación de la CVRS en la práctica clínica y el 80% reconocía su importancia (Morris, 1998). Posteriormente, en 2001, una encuesta a 271 oncólogos ha permitido constatar que el 84% afirmaba tener pocos conocimientos sobre CVRS, aunque el 82% creía tratarse de un tema útil en la prestación de cuidados al paciente y afirmaba hacer planes para incorporarla en su práctica clínica (Bezjak *et al*, 2001). En Portugal se ha aplicado una encuesta a 396 médicos, de los cuales apenas 82 contestaron. De estos, el 95% consideraba que el tema de la CVRS era esencial para proporcionar buenos cuidados y el 93% estaban de acuerdo en que era útil para tomar decisiones a nivel individual. Sin embargo, apenas el 40% refería utilizar el tema de la CVRS en su práctica clínica, el 73% consideraba más importante la toxicidad y los efectos colaterales que la CVRS y el 35% confiaba en su experiencia personal para evaluar la CVRS (Pimentel, 2002). Hay todavía mucho por hacer con respecto al tema de la CVRS ya que su evaluación debería formar parte de la práctica clínica diaria de los profesionales de salud.

Badia & Carné (1998) también destacan la importancia de evaluar la CVRS en contextos de ensayos clínicos, considerando que tal es fundamental cuando las medidas habituales de resultado en los pacientes son limitadas, como es el caso de enfermedades como la artritis reumatoide, donde la naturaleza de la enfermedad hace difícil la utilización de medidas que reflejen el resultado del tratamiento.



Estos autores consideran particularmente importante la evaluación de la CVRS en tres situaciones clínicas concretas:

- cuando la mejoría de la CVRS es el principal resultado del tratamiento aplicado, como son, por ejemplo, los cuidados paliativos del cáncer;
- cuando dos tratamientos son de una eficacia y una seguridad equivalentes, y el tratamiento nuevo puede aportar ventajas adicionales en la CVRS;
- cuando un tratamiento ofrece ventajas menores en cuanto a la supervivencia del paciente, a cambio de una evidente mejoría en su CVRS.

En suma, muchas de las decisiones tomadas por los profesionales de la salud con respecto al tratamiento de estos pacientes, están basadas en la percepción que los profesionales de salud poseen de la CVRS de sus pacientes, la cual no siempre coincide con la percibida por el propio paciente. En este contexto, es muy importante pedir a los pacientes que evalúen su propia CVRS, recurriendo para eso a instrumentos concebidos para tal efecto (Addington-Hall & Kalra, 2001). Desafortunadamente, un papel formal de la evaluación de la CVRS en la práctica clínica, en paralelo con la consulta médica y los exámenes tradicionales, está todavía lejos de ser algo que forma parte de la rutina de los profesionales de salud. Walters & Brazier (2003) subrayan que, en general, los médicos están muy familiarizados con medidas de la presión arterial o de la capacidad torácica, lo que permite que hagan una interpretación significativa de los resultados. En contraste, el significado de un cambio en la puntuación asociada a un instrumento de CVRS es menos intuitivo, no sólo porque las unidades de la escala no son familiares sino también porque los profesionales de la salud raramente utilizan este tipo de medidas en la práctica clínica diaria.

### 1.3. ¿CÓMO MEDIR CVRS?

El interés de encontrar indicadores que engloben el Hombre en su totalidad (física, psicológica, cultural y social) así como la comprensión de la gran variedad de factores que integran la Salud y la CVRS es una constante en la literatura (Figueiredo *et al*, 2006). Como hemos visto, las definiciones propuestas de CVRS enfatizan, de forma general, la naturaleza subjetiva de su evaluación, enfocando la evaluación individual de cada sujeto. En realidad, el bien estar o el estado de salud de un individuo no puede ser directamente medido. Únicamente es posible hacer inferencias a partir de indicadores mensurables de síntomas o percepciones relacionadas. En este sentido, se verifica un consenso sobre la inexistencia de una medida patrón para CVRS. Sin embargo, a pesar de la dificultad en cuantificar el concepto de CVRS y transformar las percepciones del paciente en un número, la cantidad de instrumentos disponibles para medir CVRS ha crecido rápidamente (Boini *et al*, 2004). Actualmente es posible evaluar CVRS de muchas formas diferentes utilizando medidas genéricas o específicas.

Considerando que CVRS contempla los dominios físico, psicológico y social de la salud, los cuales son influenciadas por las experiencias personales, convicciones, expectativas y percepciones, Testa & Simonson (1996) consideran que cada uno de estos dominios puede ser medido en dos dimensiones: evaluando objetivamente el estado de salud y evaluando percepciones de salud más subjetivas. En realidad, dos personas con el mismo estado de salud pueden tener niveles de calidad de vida muy diferentes, una vez que las percepciones, expectativas sobre la salud y la capacidad de adaptación afectan la percepción que una persona tiene sobre su estado de salud, su satisfacción con la vida y luego la CVRS experimentada (Sinclair & Blackburn, 2008; Wu, 2009).

Según Testa & Simonson (*op.cit.*), si por una parte, cada dominio de la Salud posee muchas componentes que deben ser medidas, por otra parte, debido a la multidimensionalidad subyacente al concepto de CVRS, se podría considerar un

número casi infinito de estados de salud. Precisamente por eso, traducir estos dominios y componentes de la salud en un valor cuantitativo, es una tarea compleja. Pedir por ejemplo “Por favor, cuantifique su Calidad de Vida o su Estado Global de Salud en una escala de 1 a 10” puede constituir una evaluación global útil, sin embargo, los conceptos de “Calidad de Vida” y “Estado Global de Salud” están ambigüamente definidos por lo que la cantidad a ser medida es muy vaga y no puede ser interpretada con exactitud. Por otra parte, considerando únicamente los datos clínicos referentes a los síntomas que indican un estado de salud objetivo, se omiten factores relevantes como la tolerancia al malestar. Las diferencias entre los instrumentos que miden CVRS se relacionan sobretodo con el grado según el cual enfatizan las dimensiones objetivas frente a las subjetivas, la extensión según la cual los diferentes dominios son cubiertos, el formato de las cuestiones, y no con diferencias en la definición básica de CVRS. El creciente interés por evaluar CVRS ha estimulado, en las últimas décadas, la proliferación de instrumentos destinados a ese fin.

Según Froberg & Kane (1989a), cualquier abordaje que intenta obtener preferencias sobre el estado de salud, debe contemplar tres pasos:

- definición de un conjunto de estados de salud de interés;
- identificación de un grupo de expertos que se pronuncien sobre las características de cada estado de salud;
- la determinación de una escala de valores para cada estado de salud.

Froberg & Kane (*op.cit.*) agrupan las estrategias de formulación de preguntas a los pacientes en dos grandes categorías: el abordaje holístico y el abordaje descompuesto. El **abordaje holístico** implica la atribución de una escala de valores a cada posible estado de salud, teniendo en cuenta que un estado de salud representa una combinación de muchos atributos. Todo el trabajo pionero en la medición de las preferencias relativas a estados de salud se ha basado en el abordaje holístico, según el cual los pacientes evalúan cada estado de salud (con

múltiplos atributos) de interés para el investigador, no siendo analizados efectos separados de cada atributo. Una de las limitaciones de los diseños holísticos es que no proporcionan información sobre la forma en la cual los diferentes atributos son sopesados y combinados de forma que produzcan los valores asociados a cada estado de salud.

Posteriormente, se ha registrado una tendencia en los estudios de preferencias relacionadas con el estado de salud, en sustituir los abordajes holísticos por los métodos de descomposición. El **abordaje por descomposición** permite al investigador obtener valores para todos los estados de salud sin necesidad de atribuir valores a cada uno. Esto simplifica la tarea de evaluación una vez que se expresa el valor total de un estado de salud como una función descompuesta de los atributos en causa. Así, al contrario de los diseños holísticos, los métodos de descomposición, reducen en gran medida la subjetividad inherente a la atribución de valores a los estados de salud completos, permitiendo que los atributos asociados a los estados de salud estén separados y sus efectos individuales analizados, de forma que proporcionen información sobre la forma como se combinan los atributos para llegar a una evaluación completa. Dentro de estos modelos, existen los modelos de descomposición explícitos, según los cuales se pide al encuestado que evalúe cada nivel de un atributo particular asumiendo que todos los demás son constantes. Comparando los diseños holísticos con los de descomposición, la principal ventaja de los últimos es que reducen significativamente la subjetividad.

Además del tipo de método utilizado, existen varios aspectos inherentes al proceso de medida que pueden afectar la evaluación que el paciente hace de su estado de salud. Estudios demuestran que determinadas características como el conocimiento médico o experiencias con enfermedades anteriores, pueden influenciar el juicio emitido por los pacientes. También la forma como los estados de salud están definidos y presentados o alteraciones sutiles en la formulación de

las frases, pueden producir alteraciones en el orden de las preferencias. Por otra parte, algunas inconsistencias pueden relacionarse con la propia capacidad de juicio humano, cuando la misma alternativa puede ser vista bajo diferentes puntos de referencia. Para superar estas dificultades, el investigador debe determinar el punto de referencia al menos de tres formas (Froberg & Kane, 1989c, d):

- proporcionar referencias concretas como por ejemplo “salud perfecta” y “muerte”. Estudios demostraron que los valores atribuidos a una escala cuyos extremos eran “salud perfecta” y “muerte” eran sistemáticamente más elevados que los obtenidos cuando se sustituía el extremo “muerte” por otros estados. De la misma forma, los valores obtenidos eran más bajos cuando se sustituía el extremo “salud perfecta” por otros conceptos;
- atribuir nombres a las enfermedades y a los tratamientos en vez de dejarlos por identificar. En un estudio de decisión clínica, la radioterapia ha sido elegida en un 42% por los individuos no estando explícitamente identificada (únicamente estaba referida como un tratamiento con resultados y consecuencias específicas) y elegida en un 26% de los casos cuando sí estaba identificada, o sea, el rótulo atribuido a un determinado fenómeno, tiene un efecto significativo en las preferencias;
- elegir una forma particular de describir las consecuencias de un tratamiento. Estudios demostraron que variaciones en la forma en la que se describen las consecuencias de un tratamiento, afectan las preferencias. Por ejemplo, formular los resultados de un tratamiento en términos de probabilidad de morir, produce resultados distintos a los obtenidos cuando se formula lo mismo en términos de probabilidad de sobrevivir.

Las formas más frecuentemente utilizadas para evaluar CVRS son las entrevistas, los diarios y los cuestionarios. Es cierto que las entrevistas y los diarios

posibilitan más información, sin embargo, el cuestionario es actualmente la técnica más ampliamente utilizada en la medición de la CVRS. El cuestionario es un método práctico, reproducible, pudiendo siempre ser complementado con una entrevista para clarificar algún aspecto.

Teniendo en cuenta que muchas de las componentes de la CVRS no pueden ser observadas directamente, es habitual recurrir a los principios clásicos de la Teoría de la Respuesta al Ítem (TRI). La TRI considera que existe un verdadero valor de CVRS que no puede ser medido directamente sino de forma indirecta, a través de una serie de cuestiones (ítems) que miden el verdadero constructo. Estas cuestiones son colocadas al paciente y la respuesta es convertida en una puntuación numérica. Si los ítems han sido debidamente elegidos, la medida resultante difiere del verdadero valor de CVRS únicamente en error aleatorio (Testa & Simonson, 1996).

#### **1.4. ¿QUIÉN DEBE MEDIR CVRS?**

Es común preguntar quién debe medir CVRS. Según varios autores, es el propio paciente quien debe evaluar su propia CVRS (Efficace *et al*, 2004, 2007; Pimentel, 2004, Schwartz *et al*, 2007; Shimozuma *et al*, 2007, Schwartz *et al*, 1997). La autoevaluación hecha por el paciente, no siempre coincide con la evaluación hecha por los familiares o por los profesionales de la salud. Los observadores son frecuentemente malos evaluadores tanto de los síntomas como de la CVRS en general y por eso la correspondencia entre las respuestas obtenidas en los cuestionarios rellenos por pacientes y observadores, es baja (Fayers & Machin, 2000; Rapkin & Schwartz, 2004; Shimozuma *et al*, *op.cit.*; Westerman *et al*, 2008; Brunelli *et al*, 1998). Es curioso notar que, incluso médicos con experiencia, tienden a subestimar o sobrestimar muchos aspectos relacionados con el desempeño físico y psicológico de los pacientes.

Otro aspecto importante en la explicación de diferencias entre la evaluación del paciente y la de los demás tiene que ver con el hecho de que la percepción de CVRS del paciente puede modificarse como resultado de una adaptación a la enfermedad (Postulari & Adang, 2000; Wiel *et al*, 2007; Sinclair & Blackburn, 2008). Esto explica porque los médicos hacen frecuentemente subestimaciones al nivel de los síntomas y desempeño de los pacientes, sin embargo, tienden a referir que los pacientes poseen peor CVRS que la percibida por los propios pacientes. En un estudio, la aplicación de un cuestionario sobre CVRS ha revelado la existencia de síntomas (fatiga, depresión, dolor, alteraciones del sueño) en un 44% de los pacientes cuando todos habían sido considerados como asintomáticos por los médicos (Glimelius *et al*, 1994). Tener en cuenta únicamente la evaluación hecha por los profesionales de la salud, es ignorar información valiosa de cómo el paciente percibe su salud y su calidad de vida, especialmente los aspectos del dominio emocional y social. Por otra parte, algunos pacientes que aparentemente están gravemente enfermos pueden realizar la mayor parte de sus actividades diarias. La CVRS no puede ser inferida por las apariencias, necesitando una evaluación rigurosa.

Ribeiro (2005) menciona un estudio desarrollado por Mosey & Shapiro (1982) como uno de los más contundentes en la demostración de la validez de las medidas de salud basadas en la percepción personal. Este estudio se basa en una muestra de 3128 sujetos a lo largo de siete años y ha permitido verificar que la autoevaluación del estado de salud constituía un predictor de supervivencia más eficaz que una evaluación objetiva basada en los datos facilitados por el médico personal, relativos a la gravedad de los problemas de salud. El estudio evaluaba la mortalidad en dos momentos: en los primeros tres años y en los cuatro años siguientes. El riesgo de fallecimiento encontrado para quienes autoevaluaban su salud como pobre en oposición a los que autoevaluaban su salud como excelente era de 2,92 mientras para la evaluación objetiva era de 2,55 en los primeros tres

años, y respectivamente de 2,77 versus 1,56 para los cuatro años siguientes. Estos autores explican que probablemente la autoevaluación refleja una percepción de cambios fisiológicos y biológicos sutiles que no serían detectados por métodos objetivos.

También Kaplan & Camacho (1983) constatan la existencia de una asociación significativa entre la percepción que los individuos tienen de su salud y la mortalidad en una investigación con 6928 individuos. Estos autores concluyen que, aunque el nivel de salud percibida esté asociado a muchas otras variables, una percepción baja de salud, actúa de forma intensa e independiente de esas variables. Estos autores sugieren que muchas limitaciones funcionales, están asociadas a percepciones de salud menos favorables. Por ejemplo, una percepción débil de salud puede ser la expresión de estados psicosociales adversos tales como aislamiento social, acontecimientos de vida negativos, depresión y estrés profesional. En esta línea Benyamini *et al* (2000) constatan que el humor positivo, el funcionamiento físico, el humor negativo, enfermedad grave, soporte social, edad, son las variables con mayor correlación (por el orden indicado) con la percepción del Estado de Salud.

Con el objetivo de demostrar el efecto de adaptación a la enfermedad de un grupo de pacientes sometido a un trasplante de riñón, Postulard & Adang (2000) pidieron a un grupo de observadores externos que estimaran la CVRS de estos pacientes. La idea subyacente al estudio era que, al no experimentar un estado de salud de alguna forma imperfecto, se asume que las medidas resultantes de las evaluaciones de los observadores, excluyen el efecto de adaptación. Los resultados demostraron que las evaluaciones de los observadores diferían significativamente de las evaluaciones de los pacientes.

Otro estudio desarrollado por Rothwell *et al* (1997) con el objetivo de comparar las evaluaciones de los médicos con las evaluaciones de pacientes con esclerosis múltiple en las varias dimensiones de la CVRS, permitió concluir que médicos y pacientes no concuerdan en relación a cual es el dominio de CVRS con



mayor importancia. Por otra parte, la evaluación de los pacientes relativa a su discapacidad física, tenía alta correlación con la evaluación de los médicos, sin embargo los médicos tienden a infravalorar las demás dimensiones como vitalidad, salud general o salud mental. Estos autores concluyen que pacientes con esclerosis múltiple, y probablemente otros con enfermedades crónicas, están menos preocupados con su discapacidad física que sus médicos. Sin embargo, los pacientes consideran importantes los efectos del tratamiento que se reflejan en otras dimensiones de su CVRS y que los médicos no valoran, o sea, la discapacidad física no siempre es el factor más determinante en la CVRS global. Estos resultados están en conformidad con los encontrados por Pinson *et al* (2009).

También Ubel *et al* (2001, 2003) subrayan las discrepancias existentes entre la CVRS percibida por el paciente y la CVRS estimada por los demás. Entre los resultados de varios estudios, estos autores destacan, por ejemplo, que en una escala donde 0 representa muerte y 1 salud perfecta, el público en general estima que la CVRS en un contexto de diálisis es un 0,39, mientras los pacientes de diálisis perciben su CVRS en un 0,56. También personas sin colostomías, estiman la CVRS de personas colostomizadas en un 0,8, mientras pacientes con colostomías evalúan su CVRS en un 0,92. Por otra parte, estudios demuestran que personas que han quedado parapléjicas o cuadripléjicas en accidentes, no se sienten significativamente menos felices que personas que han ganado recientemente la lotería. Estas discrepancias sugieren una de dos cosas: que las personas en general no entienden el valor de la vida de personas con discapacidades, o que estas personas sobrevalúan su CVRS. Ubel *et al* (*op.cit.*) identifican tres potenciales fuentes de las discrepancias encontradas entre las evaluaciones de CVRS de los pacientes y del público en general:

- Imposibilidad de los pacientes y el público compararen los mismos estados de salud – las descripciones que el público en general posee sobre un determinado estado de salud o enfermedad que no ha experimentado son vagas e incompletas. Además de eso, un individuo

que evalúe la CVRS de una persona con paraplejía, tenderá a centrarse en las actividades afectadas por la paraplejía como caminar o practicar los deportes favoritos y no ver televisión o convivir con la familia. En general, las personas sobreestiman el impacto emocional de estos eventos porque se centran en los dominios influenciados por la discapacidad, mientras el paciente considera todos los aspectos de su enfermedad.

- Imposibilidad de utilizar la misma escala de mensuración – público y pacientes no interpretan las escalas de CVRS de la misma forma, o sea, lo que para una persona es un “10”, para otras es un “11”. Por otra parte, cambios en el estado de salud, conducen a cambios en los estándares internos subyacentes a la evaluación de la CVRS. El significado de los números “1” y “10” en una escala de CVRS puede cambiar, para la misma persona, antes y después de una enfermedad. Este fenómeno se llama *response shift* y será abordado adelante con detalle.
- Diferencias reales de opinión sobre la severidad de varias enfermedades o discapacidades – el estado de salud actual de una persona, condiciona su evaluación sobre el estado de salud de otra persona. Por ejemplo, para una persona con salud normal, la diferencia entre hemiplejía y hemiplejía con afasia puede parecer pequeña ya que ambos estados son extremadamente severos. Sin embargo, para un paciente que viva con ambas, la diferencia parecerá enorme porque hablar o no, supone una gran diferencia para estos pacientes.

Asume pues gran importancia la evaluación del paciente sobre su propia CVRS. Esta importancia ha sido confirmada por la *European Medicines Agency* (EMA) al publicar un documento regulador de la utilización de medidas de CVRS en la evaluación de productos médicos (EMA, 2005). En este documento se

enfatisa la evaluación de la CVRS por autoinforme y se define *Patient-Reported Outcomes* (PRO) como cualquier evaluación hecha por el propio paciente, basada en su percepción de la enfermedad o tratamiento. También las recientes orientaciones de la *US Food and Drug Administration* (FDA) enfatizan la importancia de las PRO para acceder al impacto de intervenciones sobre la salud (FDA, 2006), atribuyendo a la evaluación del propio paciente una importancia extrema en CVRS. Las PRO deben comprender medidas simples y multidimensionales de síntomas, estado de salud, adherencia a un tratamiento, satisfacción con el tratamiento, etc.

En suma, las medidas basadas en el autopercepción son cada vez más aceptadas y diversas investigaciones han evidenciado su validez y eficacia. Siendo la CVRS un fenómeno subjetivo, lo más importante al medirla es la percepción del individuo en estudio y por eso los instrumentos de autoinforme son actualmente los más utilizados para recoger datos de CVRS.

## CAPÍTULO II

---

---

# INSTRUMENTOS PARA MEDIR CVRS

---

---

## 2. INSTRUMENTOS PARA MEDIR CVRS

### 2.1. MEDIDAS DE CVRS

A partir de la década de 60 empiezan a surgir diversos indicadores de salud, desde los que cubrían un dominio pequeño y limitado de elementos relativos al funcionamiento físico hasta los que incluían una amplia variedad de aspectos subyacentes al funcionamiento físico, social y mental. Aunque muchas de estas medidas hayan sido desarrolladas con recurso a técnicas psicométricas, la mayor parte ha sido concebida en una perspectiva de evaluar la ausencia de enfermedad y no la presencia de salud. O sea, muchos de estos instrumentos, constituyen medidas de evaluación de la enfermedad, siendo útiles para medir la salud en un hospital (la presencia o ausencia de enfermedad) sin embargo no la Salud Pública (la salud de personas que no están enfermas). En realidad los indicadores negativos de salud siguen estando más fácilmente disponibles y por eso muchos de los indicadores utilizados actualmente reflejan una gran carga de enfermedad. Lentamente se ha invertido esa tendencia y, durante los últimos 20 años, se han desarrollado instrumentos sofisticados para medir CVRS (Thoma *et al*, 2008). Ribeiro (2005) destaca tres aspectos importantes en el cambio de la evaluación de la percepción del estado de salud que se han observado desde la década de 60:

- el cambio de medidas negativas por medidas positivas;
- el recurso a otras variables además de las fisiológicas;
- basar la evaluación en la percepción del sujeto evaluado.

Cómo ya hemos mencionado, es muy frecuente utilizar cuestionarios para obtener medidas de CVRS. Estos cuestionarios son constituidos por un determinado número de ítems o cuestiones, que se agrupan en varios dominios o

dimensiones, referentes a una determinada área de la CVRS que se pretende medir (dolor, movilidad, depresión, ansiedad, etc.). En general, los ítems son descriptivos y ordenan los diferentes niveles de cada dimensión (ej: "Tiene usted dolor?" 1. "No, en absoluto", 2. "Sí, un poco", 3. "Sí, mucho"). En algunos instrumentos los ítems tienen pesos diferentes, siendo el peso atribuido en función de la importancia de cada ítem, aunque lo más frecuente es que los ítems posean pesos iguales, asumiendo que poseen importancias iguales.

Guyatt *et al* (1993) definen dos grandes grupos de objetivos subyacentes a los instrumentos para medir CVRS:

- Diferenciar entre las personas que tienen mejor CVRS y las que no tienen. Para ello se construyen los **instrumentos discriminantes** con el objetivo de medir diferencias entre pacientes en un determinado instante;
- Medir cuánto ha cambiado la CVRS. Para ello se construyen los **instrumentos evaluativos** con el objetivo de medir diferencias entre las respuestas de un paciente durante un periodo de tiempo.

Las propiedades de los instrumentos para estos dos tipos de propósitos son diferentes, como veremos enseguida.

Guyatt *et al* (*op.cit.*) definen todavía dos tipos de medidas generales de CVRS:

- evaluación del estado de salud;
- evaluación de las preferencias de los pacientes.

La **evaluación del estado de salud** se encuadra en la perspectiva psicométrica y, en general, las medidas subyacentes a esta evaluación, contienen aspectos relativos al bienestar percibido por los pacientes que resultan de una autoevaluación. La puntuación obtenida es resultante de un conjunto de cuestiones y refleja la CVRS relativa del paciente cuando comparado con otros pacientes, permitiendo también comparar puntuaciones obtenidas en diferentes

instantes por el mismo paciente. Estas medidas han sido inicialmente concebidas para comparar grupos de pacientes recibiendo diferentes tratamientos o para identificar cambios a lo largo del tiempo dentro de grupos de pacientes y han sido ampliamente utilizadas en ensayos clínicos con el objetivo de comparar diferentes terapéuticas. Las medidas relativas al estado de salud enfocan sobre todo el impacto percibido producido por la enfermedad o por la terapia, la frecuencia y severidad de los síntomas o el estado general de salud.

La **evaluación de las preferencias** de los pacientes ha resultado de la perspectiva econométrica, está basada en el concepto de utilidad y pretende expresar una decisión individual tomada bajo incertidumbre. Estas medidas pueden ser utilizadas para evaluar la relación entre calidad y cantidad de vida. En general, los valores de estas escalas oscilan entre 0 y 1, donde el 0 está asociado a la muerte y el 1 a la salud perfecta.

## **2.2. CARACTERÍSTICAS DE LOS INSTRUMENTOS PARA MEDIR CVRS**

El desarrollo de un instrumento para medir CVRS es un proceso que requiere esfuerzo y habilidad técnica. Muchas veces es un proceso que exige mucho tiempo ya que implica múltiples estudios, muchas aplicaciones y refinamientos. Existen numerosos instrumentos para establecer las propiedades psicométricas de un instrumento, siendo ampliamente utilizadas la Teoría Psicométrica Tradicional y la Teoría de la Respuesta al Ítem. Al medir constructos psicológicos y, en particular aspectos relacionados con estados de salud, es importante que los instrumentos utilizados posean algunas características importantes (Froberg & Kane, 1989b; Guyatt *et al*, 1987; Guyatt *et al*, 1993; Testa & Simonson, 1996) que veremos enseguida.

### 2.2.1. Fiabilidad

Se dice que una medida es fiable si es relativamente libre de errores de medida. La fiabilidad está relacionada con la capacidad que un determinado procedimiento de medida tiene para producir resultados consistentes en estudios repetidos independientes, o sea, se supone que el instrumento produce los mismos resultados en mediciones repetidas cuando es administrado en tiempos diferentes a sujetos estables bajo las mismas condiciones. Con respecto a los instrumentos para medir estados de salud, la fiabilidad se entiende de tres formas:

- la **fiabilidad intra** se refiere a la consistencia interna e implica homogeneidad de contenido, o sea, los ítems del instrumento deben medir aspectos distintos del mismo concepto y se mide habitualmente utilizando el *alpha de Cronbach*:

$$\alpha_{Cronbach} = \frac{m}{m-1} \left( 1 - \frac{\sum var(x_i)}{var(S)} \right)$$

siendo  $m$  el número de ítems de una determinada dimensión,  $var(x_i)$  la varianza del  $i$ -ésimo ítem de la escala, calculado a partir de la muestra y  $S = \sum x_i$ . Si los ítems no están correlacionados la  $var(S)$  será igual a la suma de las varianzas individuales, implicando  $\alpha_{Cronbach} = 0$ . Por otra parte, si todos los ítems son idénticos y presentan correlación perfecta, las varianzas de los ítems serán iguales y la suma de la varianza de todos los ítems originaría  $\alpha_{Cronbach} = 1$ . En general se acepta el valor  $\alpha_{Cronbach} = 0,7$  como razonable en escalas psicométricas, aunque sea recomendable que este valor sobrepase el  $0,8$  o el  $0,9$ , valores que se consideran como buenos y excelentes, respectivamente (Cronbach, 1951; Bland & Altman, 1997; Allen & Yen, 2002)



- la **fiabilidad test-retest** se refiere a la estabilidad de los valores de la escala al aplicarla en tiempos distintos separados por un periodo de tiempo corto, a pacientes estables (que no hayan sufrido cambios en su estado de salud), y se mide habitualmente utilizando el Coeficiente de Correlación de Pearson o el Coeficiente de Correlación Intraclase (CCI<sub>Intra</sub>) que se define como:

$$CCI_{Intra} = \frac{var(i)}{var(t)} = \frac{var(i)}{var(i) + var(e)}$$

donde  $var(i)$  denota la varianza entre los individuos (atribuida a características propias de los sujetos) y  $var(t)$  la varianza total que incluye  $var(i)$  y  $var(e)$ , la varianza debida a errores aleatorios de medida. Para variables continuas, el CCI<sub>Intra</sub> es equivalente al estadístico *kappa* y expresa la concordancia entre los valores obtenidos en repeticiones del test con los mismos individuos bajo las mismas condiciones.

- la **fiabilidad inter** dice respecto a la consistencia entre la respuesta de diferentes personas delante de la misma escala. Guyatt *et al* (1987) definen la fiabilidad de una observación simple como:

$$Fiabilidad = \frac{var(diferencias\ entre\ sujetos)}{var(total)}$$

siendo la varianza total, la suma de la varianza atribuida a las verdaderas diferencias entre sujetos, y la varianza debida a errores aleatorios de medida, y asumiendo que los errores son independientes.

La fiabilidad es una característica importante de los instrumentos discriminantes. Un instrumento fiable evidencia generalmente que pacientes

estables obtienen más o menos los mismos resultados en administraciones repetidas.

### 2.2.2. Validez

En un contexto de estudios de preferencias relacionadas con la salud, un método presenta validez si permite medir con precisión la característica para la cual ha sido concebido, o sea, si de hecho mide lo que se desea medir. Badia & Carné (1998) resumen los varios tipos de validez necesarios en los instrumentos para medir CVRS en tres:

- la **validez de criterio** se refiere a la capacidad de un instrumento para producir resultados patrón en términos de dimensiones que deben medirse y las relaciones esperadas entre ellas y otras variables. Cuando existe un patrón (habitualmente llamado *gold standard*) se evalúa la validez de criterio de un instrumento frente a ese patrón;

- la **validez de contenido** se refiere a la capacidad de un instrumento de medición para cubrir todas las dimensiones importantes del constructo, o sea, todos los aspectos relevantes del dominio que se está intentando medir deben estar debidamente representados. Un instrumento destinado a medir CVRS debe asegurar que están contempladas todas las componentes subjetivas y objetivas de este constructo (síntomas, condición, aspectos sociales, etc.) que son importantes para los elementos de una determinada población y son susceptibles de ser afectados, positiva o negativamente por intervenciones;

- **validez de concepto** (o de **constructo**) se refiere a la capacidad que un instrumento de medida tiene de relacionarse con otros de manera consistente y con las hipótesis derivadas de los conceptos subyacentes al constructo que se está

midiendo. Se espera que el instrumento para medir CVRS confirme las relaciones subyacentes al modelo en causa, evidenciando las dimensiones latentes esperadas.

En contextos de CVRS nuestro objetivo es medir una variable abstracta o constructo y no una variable observable. Además, cuándo se trata de medir CVRS, no disponemos de un *gold standard*. Esto hace que no siempre sea fácil establecer la validez de un instrumento de CVRS. En la definición de una variable abstracta y determinación exacta de lo que mide, es habitual el recurso a métodos de validación del constructo. En realidad, existen diversas aproximaciones para la validación del constructo, sin embargo, en estudios de Salud, las más comunes son el examen de los resultados obtenidos a través de métodos diferentes y el examen de las relaciones previstas entre las preferencias y todas las otras variables que se espera sean empíricamente soportadas. El objetivo es demostrar que la medida se comporta de forma consistente con lo que esperamos y que se correlaciona con aspectos observables, a su vez relacionados con CVRS.

Cuantas más veces un instrumento es utilizado y cuanto más fueren las situaciones en que se comporta como es esperado, más confianza tenemos en su validez. Guyatt *et al* (1993) refieren incluso que quizás nunca se debiera concluir que un cuestionario ha sido validado pero sí que aplicaciones en un vasto número de contextos y estudios, sugieren fuerte evidencia de validez.

### **2.2.3. Viabilidad**

Los métodos utilizados deben ser aplicables, viables y aceptables en términos de tiempo y costes. Por otra parte, en estudios clínicos, es importante que los pacientes sean cuidadosamente informados sobre el estudio y debidamente aclaradas todas las dudas.

#### 2.2.4. Sensibilidad

Aunque en contextos de estudios clínicos, una de las principales preocupaciones sea demostrar la fiabilidad y la validez del instrumento, la capacidad de un instrumento para detectar cambio en las respuestas de pacientes a lo largo del tiempo no depende únicamente de las características anteriormente referidas sino también de su habilidad para detectar diferencias mínimas clínicamente importantes. Esta característica se llama habitualmente sensibilidad en el contexto médico y epidemiológico (Guyatt *et al*, 1987).

Aunque muchos autores utilicen los términos *sensitivity* (sensibilidad) y *responsiveness* indiscriminadamente, otros hay, como Fayers & Machin (2000), que diferencian sensibilidad de *responsiveness*, relacionando el primero con la habilidad para detectar diferencias entre pacientes o grupos de pacientes y el segundo con la detección de cambios cuando el paciente mejora o empeora. Lo que sí es cierto es que se espera que el instrumento produzca diferentes resultados en mediciones repetidas si las cosas cambian, o sea, si el verdadero constructo cambia. Así, tests concebidos para medir CVRS, deben poseer tres características:

- diferenciar los sujetos en las dimensiones de interés;
- predecir los resultados de tests que hayan sido construidos para el mismo efecto;
- medir cambios en los mismos sujetos a lo largo del tiempo.

Si por una parte tests destinados a medir características como la inteligencia son, casi exclusivamente, medidas discriminantes, instrumentos cuyo objetivo es medir estados internos como el dolor o constructos como la CVRS, son medidas de carácter evaluativo. En realidad, las características requeridas para un instrumento utilizado en un estudio clínico, dependen del diseño del estudio. Para diseños con grupos independientes, el instrumento debe discriminar entre sujetos y dentro de los sujetos a lo largo del tiempo. Sin embargo, un test que haya sido diseñado con propósitos evaluativos, tiene como objetivo sobre todo medir cambios en los

sujetos a lo largo del tiempo y no tanto diferencias entre los sujetos en una dimensión considerada. Así, en general, lo que realmente es importante en un instrumento evaluativo es la capacidad que tiene para detectar efectos de tratamientos clínicamente importantes, aunque tales efectos sean pequeños. Por lo tanto, si la fiabilidad es un factor determinante en la utilidad de instrumentos concebidos para fines discriminantes, en un instrumento concebido para fines evaluativos, la habilidad para detectar diferencias mínimas clínicamente importantes es fundamental para que el instrumento sea útil.

La sensibilidad de un test está directamente relacionada con la magnitud del cambio en las puntuaciones de los sujetos. Badia & Carné (1998) sintetizan estas ideas, subrayando que, cuanto mayor sea la diferencia en la puntuación del cuestionario en los pacientes en los que existe un cambio real de salud, mayor será la sensibilidad del mismo. Por el contrario, cuanto mayor sea la diferencia en la puntuación obtenida en los sujetos estables, menor será la sensibilidad de la prueba. Estos autores enfatizan la necesidad de realizar una valoración de la relevancia o significación clínica de las diferencias observadas, diferenciando claramente significación estadística de significación clínica. En realidad, en la práctica clínica, el juicio clínico de lo que supone una diferencia clínicamente significativa se basa en la experiencia clínica. Sin embargo, este no es un criterio objetivo. Abordaremos más adelante el concepto de “mínimo cambio clínicamente importante”.

No existe una medida que sea universalmente aceptada para cuantificar la sensibilidad de un instrumento. Shun *et al* (2007) refieren que existen al menos 25 definiciones distintas y 31 medidas distintas de sensibilidad. En un ensayo clínico, a medida que la variabilidad inicial aumenta, se necesitan efectos de tratamientos más significativos para demostrar la eficacia del instrumento, sin embargo, la sensibilidad es inversamente proporcional a la variabilidad de los cambios individuales obtenidos en las puntuaciones de los sujetos a lo largo del tiempo (y

no a la variabilidad dentro de las puntuaciones de los sujetos). En términos ideales, el investigador obtendría múltiples medidas iniciales en una muestra simple de sujetos y, después de una intervención o tratamiento, obtendría las mismas medidas múltiples ahora relativas al periodo de tiempo posterior. A partir de las medidas obtenidas antes y después de la intervención, se estima la variabilidad en las puntuaciones del test en sujetos para saber si la variabilidad ha sido afectada por la intervención y así estimar las diferencias clínicamente importantes. Sin embargo, en la práctica clínica, no siempre es posible implementar este procedimiento. Incluso en el desarrollo de nuevos instrumentos (o en la validación de instrumentos ya existentes), muchos investigadores utilizan muestras separadas para examinar la variabilidad en sujetos estables y para analizar cómo las puntuaciones obtenidas a partir de ese instrumento cambian con la institución de una determinada terapia. En estas circunstancias el efecto de la terapia, en la variabilidad de los sujetos, no puede ser analizado.

Nótese que la diferencia mínima clínicamente importante y la variabilidad entre sujetos con respecto al cambio en la puntuación de cada sujeto son dos parámetros importantes que, en un contexto de ensayos clínicos, determinan el número de individuos necesario en la muestra. Si la variabilidad en sujetos estables es pequeña relativamente al cambio en la puntuación de los sujetos, lo que indica la existencia de diferencias clínicamente importantes, podrá recurrirse a una muestra pequeña en el ensayo clínico. La medida más utilizada en la cuantificación de cambios clínicos significativos, consiste en el cociente entre la magnitud de un efecto (por ejemplo, la diferencia de la puntuación media pre y pos tratamiento) y una medida de variabilidad (por ejemplo, la desviación típica del cambio o error estándar de sus puntuaciones pre tratamiento).

Guyatt *et al* (1987) proponen como medida de sensibilidad el índice siguiente:

$$\text{Sensibilidad} = \frac{\Delta}{\sqrt{2 \times \text{Error Cuadrático Medio}}}$$

donde  $\Delta$  denota las diferencias clínicamente importantes. Cuantificando la sensibilidad de un test a través de este estadístico, podremos elegir el tamaño muestral necesario en cualquier ensayo clínico en el que dispongamos de las puntuaciones antes y después de la intervención o tratamiento y en el que el cambio a lo largo del tiempo sea el foco.

Uno de los problemas con el índice de sensibilidad descrito es que, para instrumentos nuevos (y a veces para instrumentos ya validados), no conocemos el cambio en la puntuación que constituye una diferencia clínicamente importante. Faltando esta información para obtener una estimación inicial de la sensibilidad de un instrumento, puede compararse la desviación estándar dentro de las puntuaciones de cada sujeto con el cambio observado en la puntuación después de un tratamiento de eficacia conocida. Si el tratamiento en estudio no cambia el estado del paciente (o produce un cambio que no es clínicamente importante), la sensibilidad será subestimada. Por otra parte, si la intervención o tratamiento produce una mejoría mayor que la diferencia clínicamente importante mínima, se obtendrá una sobrestimación de la sensibilidad.

La interpretación de la significación clínica anteriormente referida, está basada en la distribución estadística de las puntuaciones. Badia & Carné (1998) sugieren también que tal interpretación debe basarse en cambios clínicos externos. En este caso las puntuaciones obtenidas en un cuestionario de CVRS se comparan con otras medidas clínicas o externas para conseguir una valoración de la significación clínica. Habitualmente se compara la puntuación obtenida en el cuestionario con una pregunta de evaluación general de la salud que contesta el médico, el paciente o ambos. La magnitud del efecto se puede conocer preguntando al paciente directamente por el cambio en su salud que percibe como importante.

### 2.3. CONSTRUCCIÓN DE INSTRUMENTOS PARA MEDIR CVRS

En la construcción de instrumentos destinados a evaluar CVRS, Fayers & Machin (2000) subrayan la importancia de formular claramente las cuestiones de investigación, especificar los objetivos subyacentes a la medición de CVRS, definir el significado de CVRS, identificar la población en estudio y las dimensiones de CVRS a evaluar. Tal como hemos mencionado antes, por ejemplo, los instrumentos serán necesariamente diferentes si se destinan a comparar grupos de tratamiento en estudios clínicos (instrumentos discriminantes) o a la evaluación individual de los pacientes (instrumentos evaluativos). Por otra parte, diferentes definiciones de CVRS atribuyen diferentes grados de importancia a los síntomas y a aspectos del dominio psicológico y espiritual.

La construcción de instrumentos para medir CVRS es un proceso complejo al cual pueden surgir asociados diversos problemas. Según Guyatt *et al* (1987), cuando estamos interesados en medir el estado funcional o la CVRS en pacientes con enfermedades crónicas, uno de los problemas que surge con la construcción del cuestionario tiene que ver con el hecho de que las actividades de los pacientes y la importancia de cada actividad varía ampliamente entre individuos. Una sugerencia para contornear esta cuestión es pedir al paciente que elija actividades con las cuales tiene dificultad y analizar el efecto de una intervención o tratamiento en esa actividad.

Asimismo, el cuestionario es el instrumento más comúnmente utilizado en la obtención de medidas de CVRS por autoinforme. Al elaborar un cuestionario el palabreo debe ser adecuado. Las palabras utilizadas deben ser examinadas de forma que el contenido de cada cuestión sea relevante para la población de interés. Aunque expertos puedan hacer esta evaluación, es siempre conveniente verificar la validez de las cuestiones con los propios pacientes, ya que estos pueden tener una perspectiva distinta.



En general, las escalas de CVRS se refieren a una evaluación basada en un periodo de tiempo específico (los últimos 7 días o las últimas 4 semanas como en el SF-36). Nótese que este período de tiempo debe ser suficientemente corto para que posibilite la detección de diferencias entre los tratamientos y suficientemente largo como para permitir minimizar pequeñas fluctuaciones (ruido) que no traducen un cambio real.

En el caso de que la aplicación del cuestionario a una determinada población, implica su traducción, debe utilizarse lenguaje nativo en cada paso y proceder a una validación transcultural. Es frecuente el recurso a técnicas como el test cognitivo que consiste en pedir a los sujetos que describan verbalmente lo que están pensando y como construyen sus respuestas. Los métodos basados en la Teoría de la Respuesta al Ítem y en los Modelos de Rasch han facilitado el examen del funcionamiento diferencial de los ítems en función de la cultura y del lenguaje. El *Scientific Advisory Committee* (2002), recomienda que, en la adaptación transcultural de cualquier técnica de evaluación, sean tomados en consideración tres aspectos importantes:

- equivalencia lingüística;
- equivalencia conceptual;
- equivalencia psicométrica.

Para garantizar la equivalencia en las traducciones y adaptaciones transculturales de instrumentos de medida en el área de la salud, Guilhemmin *et al* (1993) sugieren que se empiece por producir varias traducciones, por medio de traductores cualificados. Se procede después a una retrotraducción (*back-translation*) de cada una de las traducciones iniciales. Las versiones traducidas deben entonces ser comparadas por un equipo multidisciplinar por si fuese necesario modificar instrucciones o el formato, modificar o rechazar ítems no apropiados, asegurándose de que la traducción está completamente comprensible y chequeando la equivalencia transcultural de la versión traducida. Guilhemmin *et*

*al* (1993) subrayan la importancia del pre-test, que permitirá verificar la equivalencia mediante la utilización de técnicas apropiadas, así como considerar posibles adaptaciones de las puntuaciones al nuevo contexto cultural.

Según Pimentel (2004), los cuestionarios deben ser específicos, adaptados a cada tipo de patología y diseñados de forma a que sus calidades psicométricas sean consistentes, debiendo obedecer a algunas características, que sintetiza en los siguientes puntos:

- Los ítems iniciales deben ser seleccionados o creados por los equipos donde participen profesionales de salud con experiencia en la situación clínica para la cual se pretende construir el cuestionario.
- Normalización, que asegura la uniformidad de las cuestiones colocadas.
- Diseño y validación en base a poblaciones apropiadas (por ejemplo, si se pretende utilizar un cuestionario en pacientes oncológicos, este debe ser desarrollado en esta población).
- Utilizar un lenguaje accesible a todos los individuos, independiente de su escolaridad.
- Ser fácil de contestar y cuantificar.
- Los ítems deben ser percibidos como relevantes por los pacientes;
- Incluir ítems y respuestas alternativas, capaces de reflejar aspectos positivos de la CVRS.
- Incluir un número suficiente de ítems totales por dimensión de CVRS, capaz de proporcionar puntuaciones que permitan reconocer modificaciones en la CVRS Global o de alguna de sus componentes. Este aspecto incluye sensibilidad a los cambios producidos por la enfermedad o tratamientos; capacidad para distinguir entre la componente física, psicológica y social; poder de discriminación entre los efectos de la enfermedad y de los tratamientos;

- Viabilidad - el cuestionario debe ser razonablemente corto, de forma que pueda ser efectuado por los pacientes con algún tipo de limitación, no debiendo sobrepasar los 20 minutos.
- Respetar las susceptibilidades de los pacientes.
- Bajo coste económico en la ejecución.
- Capacidad de generar una base de datos para futuras tomas de decisión.

## 2.4. INSTRUMENTOS, MEDIDAS Y ESCALAS UTILIZADOS PARA MEDIR CVRS

En las últimas décadas, varias generaciones de cuestionarios han surgido con el propósito de medir CVRS. Los primeros cuestionarios que intentaban evaluar la CVRS en pacientes sometidos a quimioterapia datan de las décadas de 40-50 y únicamente consideraban la componente funcional, la cual, como es bien sabido, representa una parte de la CVRS. Posteriormente, diversos instrumentos han sido aplicados en varios ámbitos para medir CVRS, los cuales contemplaban no solamente los síntomas físicos y psicológicos, sino también el impacto de la enfermedad, la percepción de estrés y la satisfacción con la vida. La mayoría de los cuestionarios aplicados tienen como objetivo la obtención de una única medida que traduzca la CVRS.

En este apartado veremos algunos instrumentos genéricos y específicos para evaluar CVRS. Dedicaremos especial atención a dos de estos instrumentos: el *Medical Outcomes Study 36 – Item Short Form (SF-36)* y el *Rotterdam Symptom Check List (RSCL)*, dado que han sido los instrumentos utilizados en nuestro estudio para la recogida de datos.

### 2.4.1. Instrumentos Genéricos VS Instrumentos Específicos

Muchos de los cuestionarios creados para medir CVRS son **instrumentos genéricos** que evalúan el estado general de salud y pueden ser aplicados

independientemente del diagnóstico o del estado general del paciente e incluso a individuos sanos. Estos instrumentos no están relacionados con ningún tipo de enfermedad en particular y por lo tanto son aplicables tanto a la población en general con diferencias sociales y demográficas, como a grupos de pacientes con diferentes tipos y grados de enfermedades, sometidos a diferentes tratamientos e intervenciones, dado que incluyen un amplio espectro de dimensiones de la CVRS compatibles con diferentes pacientes, enfermedades y poblaciones (Badia & Carné, 1998).

Fayers & Machin, (2000) subrayan que, aunque sean comúnmente descritos como escalas de CVRS, muchos de los instrumentos genéricos son medidas del estado de salud ya que enfocan los síntomas físicos, basándose en el supuesto de que un bajo grado de salud implica un bajo nivel de CVRS. Según Patrick & Deyo (1989), los instrumentos genéricos para evaluar el estado de salud contemplan conceptos de CVRS y cubren esencialmente cinco aspectos distintos: duración de la vida, empeoramiento, estado funcional, percepción y oportunidades sociales. Estos instrumentos proporcionan definiciones operacionales de diversos conceptos que pueden ser sumariados en un índice numérico simple. Describimos sucintamente algunos de estos instrumentos genéricos:

- El *Sickness Impact Profile (SIP)* (Bergner *et al*, 1981) es una medida del estado general de salud que describe actividades cotidianas con el objetivo de evaluar el nivel de desempeño del paciente. El SIP permite acceder a disfunciones en 12 categorías distintas (física, psicosocial, trabajo, alimentación, sueño y descanso, tareas domésticas, ocio, etc.) a las cuales se asocia una puntuación;
- El *Nottingham Health Profile (NHP)* (Hunt *et al*, 1981) mide el estrés emocional, social y físico. Es similar al SIP, sin embargo evalúa directamente sentimientos y emociones y no cambios en el comportamiento, siendo más corto y fácil de rellenar. Los ítems se

distribuyen en 6 dimensiones que cubren el sueño, dolor, reacciones emocionales, aislamiento social, movilidad física y nivel de energía;

- El *Medical Outcomes Study 36 – Item Short Form (SF-36)* (Ware *et al.*, 1993) evalúa conceptos genéricos de salud y no es específico para ninguna patología o edad, enfatizando la *performance* física, social y emocional. Dedicaremos especial atención al SF-36 ya que ha sido uno de los instrumentos utilizados en el estudio desarrollado;
- El *EuroQol (EQ-5D)* (Brooks, 1996) mide también el estado general de salud, reuniendo información sobre aspectos mentales, físicos y sociales. Este instrumento tiene la ventaja de que es bastante sencillo, pudiendo ser rellenado en 2 minutos y se recomienda su aplicación conjuntamente con otros instrumentos específicos.

Los instrumentos genéricos tienen la ventaja de posibilitar la comparación de puntuaciones obtenidas por pacientes con diferentes patologías o incluso individuos sanos. Sin embargo, tales instrumentos fallan en la evaluación de problemas asociados a patologías o tratamientos específicos como por ejemplo en el dominio oncológico. Se han desarrollado pues **instrumentos específicos**. Los instrumentos específicos han sido concebidos para acceder a grupos con diagnósticos específicos o poblaciones de pacientes, siempre con el objetivo de medir sensibilidad o cambios clínicamente importantes (Patrick & Deyo, 1989). Este tipo de escalas se caracteriza por incluir dimensiones de la CVRS de un determinado ámbito que se pretende estudiar, con más detalle que los genéricos y más ítems relativos a funciones específicas. En general, incluyen preguntas sobre el impacto de los síntomas en dimensiones de la CVRS y, con respecto a los instrumentos genéricos, tienen la ventaja de presentar una alta sensibilidad a los cambios ante el problema específico de salud que se está evaluando (Badia & Carné, 1998). En los últimos años se han desarrollado numerosos instrumentos

específicos, en varios dominios, de entre los cuales destacamos los siguientes en el ámbito de la Oncología:

- El *Breast Cancer Chemotherapy Questionnaire (BCQ)* (Levine *et al*, 1988) concebido para mujeres con cáncer de mama en quimioterapia;
- El *Cancer Rehabilitation Evaluation Systems (CARES)* (Heinrich *et al*, 1984) es otro instrumento específico para evaluar los problemas asociados al cáncer;
- El *European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC-QLQ)* (Aaronson *et al*, 1993) es específico para pacientes de cáncer. Este cuestionario está constituido por 30 ítems y es uno de los que más se utiliza en Europa por ser corto y de fácil aplicación;
- El *Functional Assessment of Cancer Therapy - General (FACT-G)* (Cella *et al*, 1993) posee 27 ítems distribuidos en las subescalas de bienestar físico, bienestar social y familiar, bienestar general, bienestar emocional y bienestar funcional;
- El *Functional Living Index Scale (FLIC)* (Schipper *et al*, 1984) se centra en la intensidad e impacto de los síntomas en las actividades diarias;
- El *Linear Analogue Self-Assessment* (Selby *et al*, 1984) que evalúa el estado de salud general y los problemas específicos asociados al cáncer;
- El *Multidimensional Quality of Life Scale* (Padilla *et al*, 1983) es otro instrumento específicamente diseñado para pacientes de cáncer;
- El *Quality of Life Index (QL-index)* (Spitzer *et al*, 1981) es contestado por el médico o enfermero y pretende proporcionar una medida basada en la observación;
- El *Rotterdam Symptom Check List (RSCL)* (De Haes *et al*, 1990) ha sido ampliamente utilizado en ensayos clínicos para medir CVRS en pacientes oncológicos. Dedicaremos especial atención a este instrumento

ya que ha sido uno de los instrumentos utilizados en el estudio desarrollado.

Como hemos mencionado, en oposición a los instrumentos genéricos, los instrumentos específicos son, en general, clínicamente sensibles a pequeños cambios. Sin embargo, no posibilitan comparaciones entre diferentes poblaciones. La opción por un instrumento genérico o específico depende sobre todo del objetivo de la investigación.

Según Patrick & Deyo (1989) el recurso a un instrumento genérico y a un instrumento específico en la misma investigación puede ser ventajosa ya que el instrumento específico es más sensible a la eficacia terapéutica, mientras el instrumento genérico posibilita la evaluación de efectos clínicos adversos que pueden ser directamente comparados. Otras soluciones intermedias consisten en utilizar un instrumento genérico modificado para una población específica, utilizar un instrumento genérico con un suplemento específico que contenga los aspectos específicos que no están contemplados en el instrumento genérico o utilizar simplemente baterías de instrumentos específicos. Esta última solución es particularmente utilizada en ensayos clínicos e investigaciones epidemiológicas donde se aplican escalas, subescalas o ítems individuales de instrumentos ampliamente utilizados de forma a testar efectos específicos. Sin embargo, la utilización de baterías de instrumentos específicos tiene la desventaja de no proporcionar un índice del estado del paciente o una indicación de la importancia relativa de cada dimensión del estado de salud.

En nuestro estudio, hemos optado por utilizar un instrumento genérico (el SF-36) y un instrumento específico (el RSCL) reuniendo las ventajas de utilización de ambos tipos de instrumentos. El SF-36 ha posibilitado la obtención de resultados que futuramente pueden ser comparados con otras poblaciones, mientras el RSCL, siendo más sensible y específicamente concebido para pacientes

de cáncer, permite captar mejor los efectos de la enfermedad y los tratamientos. Abordamos enseguida estos instrumentos con más pormenor.

#### **2.4.1.1. Medical Outcomes Study 36 – Item Short Form (SF-36)**

El *Medical Outcomes Study 36 – Item Short Form* (SF-36) es un instrumento genérico breve de evaluación del Estado de Salud y ha sido desarrollado por Ware *et al* (1993). Actualmente es el instrumento de CVRS más utilizado en el Reino Unido, ha sido traducido en 45 idiomas distintos y es considerado uno de cuestionarios más válidos y fiables (McClimans, 2006). Contiene 36 ítems de los cuales 35 se agrupan en 8 escalas; 4 de estas 8 escalas proporcionan un resumen de estados mentales, mientras las restantes 4 constituyen medidas físicas.

Siendo una medida genérica del estado de salud, el SF-36 no incide específicamente sobre un grupo etario, enfermedad o tratamiento. Este instrumento permite evaluar el impacto de intervenciones médicas en el estado de salud en lo que respecta a los 8 dominios en estudio desde el punto de vista del individuo y evaluar conceptos de salud que representan los valores humanos básicos relevantes para el bienestar funcional y mental. Según los autores, el SF-36 se ha revelado útil en la comparación entre poblaciones generales y específicas, en la evaluación de los efectos de diversas enfermedades, en la diferenciación de los beneficios para la salud que son consecuencia de diversos tratamientos, en el rastreo de pacientes individuales y en las decisiones políticas y económicas para la Salud.

#### ***Historia y Desarrollo del SF-36***

El origen del SF-36 se relaciona con el *The Rand Health Insurance Study* (HIS), desarrollado por Brook *et al* (1979) en la década de 70, del cual resultó el *General Health Rating Index* (GHRI). El HIS ha proporcionado técnicas para evaluar el estado de salud de las personas, en particular en lo que respecta a la Salud Física, Salud Mental, Salud Social y Percepción General de Salud. Al HIS ha seguido el



*Medical Outcomes Study* (MOS) que ha originado el *Functioning and Well-Being Profile* (FWBP) con 149 ítems. Todos estos estudios se han basado en la definición formal de Salud de la OMS, como concepto multidimensional. El SF-36 ha tenido origen en los 149 ítems del FWBP y se debe a Ware *et al* (1993).

Según Ware *et al* (*op.cit.*) la primera tentativa para construir un cuestionario reducido destinado a una evaluación del Estado de Salud, fue hecha en 1984 y incluía 18 ítems que medían el funcionamiento físico, las limitaciones funcionales resultantes de una salud débil, la salud mental general y la percepción general de salud. En 1986 han sido añadidos dos ítems a este instrumento para medir el funcionamiento social y el dolor corporal, originando una versión que es conocida como el MOS SF-20. En realidad, todos los ítems elegidos habían ya sido utilizados anteriormente en múltiples cuestionarios que pretendían evaluar otros constructos y que han sido seleccionados para el MOS SF-20 por ser considerados buenos para evaluar Salud. Ware *et al* (*op.cit.*) han perfeccionado este cuestionario, presentando la forma reducida de 36 ítems, conocida por SF-36.

Posteriormente el SF-36 ha sido traducido y adaptado a diversos países y se han desarrollado normas para la utilización internacional de este instrumento, siendo en la actualidad uno de los más utilizados para evaluar la Percepción del Estado de Salud. Ha tenido gran importancia el Proyecto IQOLA (*Internacional Quality of Life Assessement*) que tenía como objetivo la validación y adaptación del SF-36 a diversos países y cuyos resultados han sido publicados en 1998 incluyendo 10 países: Alemania, Dinamarca, España, Estados Unidos de América, Francia, Holanda, Italia, Noruega, Reino Unido y Suecia (Keller *et al*, 1998).

En este estudio se ha utilizado la versión portuguesa traducida y validada de Ferreira (2000a, 2000b) y Ferreira & Santana (2003) (Anexo I).

### ***Contenido, Utilización y Análisis del SF-36***

Los 36 ítems del SF-36 se distribuyen por 8 dimensiones: Funcionamiento Físico, Desempeño Físico, Dolor Corporal, Percepción General de Salud, Vitalidad,

Funcionamiento Social, Desempeño Emocional y Salud Mental y más un ítem de Transición de Salud. A su vez, estas 8 dimensiones se asocian en 2 grandes componentes: Componente Física y Componente Mental.

La subescala de **Funcionamiento Físico** está constituida por 10 ítems expresados en una escala de Likert de 3 puntos que va desde “*muy limitado*” a “*nada limitado*”. Una puntuación mínima en esta dimensión significa que, por razones de salud, el individuo se encuentra muy limitado en la realización de todas las actividades físicas, incluyendo ducharse o vestirse, mientras una puntuación máxima traduce que el individuo realiza todo tipo de actividad física, incluyendo las más exigentes, sin limitaciones.

La subescala de **Desempeño Físico** está constituida por 4 ítems expresos en una escala de Likert de 5 puntos que va desde “*siempre*” a “*nunca*”. Una puntuación baja en esta dimensión significa que, por razones de salud, el individuo tiene problemas con el trabajo u otras actividades diarias. A su vez, una puntuación elevada, traduce que el individuo no presenta problemas en la realización de su trabajo o de otras actividades diarias.

La subescala de **Dolor Corporal** es constituida por 2 ítems expresos en una escala de Likert de 5 puntos que va desde “*absolutamente nada*” a “*inmenso*”. Una puntuación mínima se interpreta como dolor muy intenso y extremadamente limitativo, mientras una puntuación máxima significa sin dolor o limitación debida al dolor.

La subescala de **Salud General** es constituida por 5 ítems expresos en una escala de Likert de 5 puntos que va desde “*absolutamente verdad*” a “*absolutamente falso*” y permite evaluar el estado general de salud. Así, una puntuación baja traduce que el individuo percibe un bajo nivel de salud y que probablemente su salud va a empeorar. A su vez, una puntuación alta significa que el individuo evalúa su salud como excelente.

La subescala de **Vitalidad** es constituida por 4 ítems expresos en una escala Likert de 5 puntos que va desde “*siempre*” a “*nunca*”. Puntuaciones bajas en esta

escala evidencian sentimientos de cansancio o extenuación durante la mayor parte del tiempo, mientras una puntuación elevada significa que el individuo se siente animado y con energía.

La subescala de **Funcionamiento Social** contiene 2 ítems expresados en una escala Likert de 5 puntos que se encuentran formulados en sentido inverso. Puntuaciones bajas en esta subescala refleja que los problemas emocionales y físicos interfieren de forma extrema y frecuente en las actividades sociales normales. Por otra parte, puntuaciones elevadas significan que el individuo realiza sus actividades sociales normalmente, sin que problemas físicos o emocionales interfieran.

La subescala de **Desempeño Emocional** contiene 3 ítems expresados en una escala de Likert de 5 puntos que va desde “*siempre*” a “*nunca*”. Puntuaciones bajas en esta subescala reflejan dificultades con el trabajo u otras actividades diarias como resultado de problemas emocionales. A su vez, puntuaciones elevadas evidencian que el individuo no tiene dificultades con el trabajo u otras actividades diarias debido a problemas emocionales.

La subescala de **Salud Mental** está constituida por 5 ítems expresados en una escala de Likert de 5 puntos que va desde “*siempre*” a “*nunca*”, para los cuales puntuaciones bajas evidencian que el individuo se siente nervioso y deprimido, mientras puntuaciones elevadas indican la sensación de paz, felicidad y calma.

Por último, una puntuación baja en el ítem relativo a la **Transición de Salud**, que se encuentra expresado en una Escala de Likert de 5 puntos, indica que el individuo cree que su estado de salud general es mejor ahora que hace un año, mientras una puntuación elevada indica precisamente el contrario.

#### 2.4.1.2. Rotterdam Symptom Check List (RSCL)

El *Rotterdam Symptom Check List* (RSCL) es un instrumento específico de CVRS y ha sido propuesto por De Haes *et al* (1990) para la evaluación de síntomas físicos y psíquicos en pacientes de cáncer. El RSCL ha sido originalmente validado

en un estudio holandés y posteriormente utilizado en diversas investigaciones holandesas y británicas. Este instrumento ha sido desarrollado con el objetivo de evaluar objetivamente la CVRS en pacientes con neoplasias bajo tratamiento oncológico, permitiendo monitorizar niveles de ansiedad, depresión y reflejando la presencia de enfermedad psicológica. Se trata de una escala simple, constituida por 30 ítems, que deberá ser rellenada por el propio paciente, pudiendo ser utilizada tanto en ambulatorio como en régimen de internamiento. Incluye 8 ítems más, referentes a actividades diarias, con el objetivo de caracterizar el estado funcional del paciente y otro ítem que evalúa la percepción general del paciente relativamente a su nivel de CVRS.

### *Historia y Desarrollo del RSCL*

El RSCL ha sido construido en base al análisis de datos resultantes de la aplicación de diferentes instrumentos:

- el *Hopkins Symptom Checklist*, el cual ha sido utilizado en una población de 352 pacientes psiquiátricos, 147 pacientes con artritis reumatoide y 308 individuos normales para control (Luteijn *et al*, 1979);
- una escala de síntomas utilizada en el control de síntomas de 150 pacientes de cáncer de mama (Linssen *et al*, 1979);
- una versión alemana del *Symptom Distress Scale* aplicada a un grupo de 49 pacientes de cáncer hospitalizados (Leendertse *et al*, 1979).

La selección de los ítems de estas escalas ha sido hecha por un grupo de expertos en Oncología y se ha basado en las saturaciones de los ítems en los factores. Se han excluido los ítems excesivamente asimétricos, resultando un total de 34 ítems que contemplan síntomas físicos y psicológicos. A los pacientes se ha pedido que indicaran el grado de preocupación relativamente a estos síntomas durante los últimos tres días en una escala de Likert de cuatro puntos con las categorías “no en absoluto”, “un poco”, “bastante” y “mucho”. Más tarde se ha añadido un ítem que pretendía evaluar el nivel de CVRS global. La investigación

de la estructura y estabilidad del RSCL y su validación se ha basado en tres estudios:

- Se ha estudiado una muestra de 86 mujeres con cáncer de mama que se desplazaban al hospital para la quimioterapia o seguimiento;
- La segunda investigación se ha basado en el estudio de una muestra de 56 pacientes que participaban en un ensayo clínico con vista a la comparación de dos tipos distintos de quimioterapia para tratamiento de cáncer de ovario en estado avanzado. Las pacientes han completado el cuestionario varias veces a lo largo del tratamiento;
- En un tercero estudio se ha comparado la CVRS en un grupo heterogéneo de 216 pacientes con cáncer que habían sido operados en los últimos tres meses o que estaban haciendo quimioterapia y un grupo constituido por 192 pacientes de cáncer sin síntomas hace 3 años o más con un grupo de control constituido por una muestra aleatoria de la población holandesa de 201 individuos.

Después de estos trabajos que han posibilitado la validación del RSCL, se han realizado diversos estudios con el objetivo de mejorar su fiabilidad y validez. Además de la versión holandesa (De Haes *et al*, 1990), se han desarrollado también las versiones inglesa (Watson *et al*, 1992), italiana (Paci, 1992) y portuguesa (Santos *et al*, 1994) que utilizamos en este estudio (Anexo II).

La primera versión del RSCL presentaba una estructura tetra factorial, siendo identificables las siguientes dimensiones:

- dimensión **psicológica**, donde todos los ítems con saturaciones elevadas describían angustia psicológica, siendo este el factor que explicaba mayor cantidad de varianza;

- la dimensión siguiente se relacionaba con el **dolor** e incluía síntomas como dolor de cabeza, en los músculos y en la espalda. Síntomas como dificultad en respirar, escalofríos o mareos estaban correlacionados con este factor, sin embargo, su contenido estaba menos relacionado con la experiencia de dolor;
- la tercera dimensión se refería a **perturbaciones gastrointestinales** como vómitos, náuseas o pérdida de apetito. El ítem que mide el nivel de irritabilidad presentaba también saturación elevada en este factor;
- la cuarta dimensión se relacionaba con la **fatiga** y pérdida de energía.

El patrón emergente de este análisis depende también de los aspectos inherentes a las especificidades de la propia enfermedad y su tratamiento. El contenido de la dimensión psicológica es, sin duda, más claro que el de las demás dimensiones.

El RSCL ha sido posteriormente mejorado. Los varios análisis de su estructura han puesto de manifiesto que las dimensiones psicológica y física son las esenciales para describir la experiencia de síntomas en pacientes de cáncer. El RSCL presenta pues una estructura de dos factores: la dimensión física e la dimensión psicológica.

La **dimensión psicológica** se ha revelado un elemento estable en la estructura del RSCL. Diversos estudios han confirmado la elevada fiabilidad y consistencia de la escala psicológica. Los ítems inicialmente incluidos en esta escala han cambiado. Por ejemplo, en los primeros estudios el ítem relativo a la irritabilidad, pertenencia al factor gastrointestinal. También la falta de energía y el dolor de cabeza surgían asociados a la dimensión física en la población homogénea de pacientes de cáncer y claramente a la dimensión psicológica en la población heterogénea. También el ítem relativo a la dificultad en dormir surgía de forma menos consistente en la subescala psicológica. En realidad, estos factores pueden ser considerados psicosomáticos. Síntomas físicos que habitualmente

acompañan la morbilidad psicológica asumen diferentes significados para los pacientes de cáncer. Algunos de estos síntomas están relacionados con la enfermedad o con los efectos del tratamiento del cáncer y como tal no deben incluirse en instrumentos concebidos para medir la ansiedad o la depresión. Por eso se ha decidido incluir únicamente ítems de naturaleza psicológica en la subescala psicológica del RSCL que así contiene ítems relacionados con irritabilidad, preocupación, depresión, nervios, sentimientos de desesperanza en el futuro, tensión, ansiedad y problemas de concentración. Los estudios revelan que la angustia psicológica experimentada por los pacientes es en alto grado independiente de los síntomas físicos.

El patrón subyacente a la **dimensión física** se ha revelado menos estable. En los primeros estudios los autores han encontrado tres factores: dolor, fatiga y perturbaciones gastrointestinales. Sin embargo, esta distinción no ha quedado clara en los estudios posteriores lo que se ha explicado debido a la homogeneidad de las poblaciones estudiadas. En los primeros estudios, muchos de los participantes recibían quimioterapia, la cual puede originar distintos patrones en la experiencia de los síntomas y luego una estructura dimensional específica. En la población heterogénea del tercer estudio la toxicidad del tratamiento y la experiencia resultante de la enfermedad ha sido más diversificada y consecuentemente la relación entre los síntomas en las que se basa el análisis resultó más débil. El bienestar físico, náuseas, trastornos alimentarios, pérdida de pelo, deterioro físico y social, fatiga, entre otras, son también dimensiones asociadas a la dimensión física en pacientes de cáncer referidas en la literatura (Schipper *et al*, 1984; Selby *et al*, 1984, Aaronson *et al*, 1987; Padilla *et al*, 1983). En realidad las dimensiones que pueden considerarse son tan diversas que no es posible formular un modelo general para la dimensión física en pacientes de cáncer. El patrón subyacente a la dimensión física en pacientes de cáncer depende de la población en estudio y de los instrumentos utilizados.

El RSCL se ha revelado una herramienta útil en los estudios clínicos de cáncer en la medida que cubre los dominios relevantes en la experiencia de estos pacientes, siendo fácil de administrar.

### *Contenido, Utilización y Análisis del RSCL*

El RSCL tarda unos 8 minutos en ser rellenado, se destina a evaluar la CVRS en pacientes de cáncer y cubre cuatro dominios: síntomas físicos, síntomas psicológicos, nivel de actividad y CVRS global. Estas son las cuatro escalas principales. Los ítems de la dimensión física están intercalados con los de la dimensión psicológica de forma que se evite que las respuestas sean influenciadas unas por otras.

La escala relativa a los **Síntomas Físicos** contiene 23 ítems. Algunos síntomas como dolor de cabeza y fatiga pueden ser experimentados por las personas en general y no solamente por los pacientes de cáncer. Otros síntomas como los trastornos gastrointestinales o los síntomas inherentes a la quimioterapia, están más específicamente relacionados con el cáncer o con su tratamiento.

La escala relativa a la **Angustia Psicológica** consta de 7 ítems contemplando diferentes síntomas que pueden ser experimentados tanto por pacientes de cáncer como por otras poblaciones. Ninguno de estos ítems se refiere a síntomas psicológicos que pueden ser considerados psicósomáticos (como dolor de cabeza o problemas de sueño). El ítem relativo al sentimiento de soledad ha sido inicialmente excluido por tener una distribución bastante asimétrica. También el ítem relativo a la concentración ha sido excluido de esta escala por haberse considerado que tenía más relación con la subescala de síntomas físicos, en particular con la fatiga.

La escala relativa al **Nivel de Actividad** consiste en 8 ítems que contemplan el estado funcional, cubriendo tanto la movilidad como las actividades sociales. Esta escala no está específicamente relacionada con el cáncer.

Por último, la **CVRS Global** es medida por un único ítem.



Las categorías de respuesta para la escala de Síntomas Físicos y Angustia Psicológica están medidas en una escala tipo Likert con 4 opciones de respuesta desde “*no en absoluto*” hasta “*mucho*”. Para la escala de Nivel de Actividad se utiliza igualmente una escala de Likert de 4 puntos que varía entre “*incapaz*” de efectuar una determinada actividad y efectuarla “*sin ayuda*”. La CVRS Global es evaluada en una escala de Likert de 7 puntos entre “*excelente*” y “*muy mala*”.

#### **2.4.2. Ítems y Escalas de Respuesta**

Algunos instrumentos tratan de evaluar un único concepto como el estado funcional, otros pretenden acceder a varias dimensiones de CVRS y por lo tanto incluyen ítems relacionados con diversos conceptos. Aunque no exista un consenso sobre las componentes a ser evaluadas, la mayoría de los investigadores entiende la CVRS como un constructo multidimensional y considera que los cuestionarios deben contener un determinado conjunto de dimensiones (Fayers & Machin, 2000). Sin embargo, existiendo tantas dimensiones, es imposible intentar acceder a todos los conceptos implicados a la vez en un único instrumento. La mayor parte de los instrumentos intenta incluir ítems relativos a las dimensiones física, emocional y social. Por ejemplo, si se considera la dimensión emocional como parte de la CVRS, deberán incluirse ítems que evalúen la ansiedad, la tensión, la irritabilidad, la depresión, etc.

Además de los ítems que miden dimensiones específicas, gran parte de los instrumentos, contiene una cuestión que pretende evaluar el nivel global de CVRS. Los ítems a través de los cuales se pretende evaluar el nivel de CVRS global implican que el sujeto considere todos los aspectos del fenómeno que para sí son importantes, ignorando los aspectos que no son relevantes para su situación y atribuyendo pesos diferenciados a otros aspectos según la importancia que tienen en su CVRS. Por este motivo, algunos autores consideran que las cuestiones globales son difíciles de interpretar y contestar, defendiendo la utilización de

múltiples cuestiones sobre CVRS a través de las cuales se obtiene una puntuación global.

Fayers *et al* (1997) consideran que los ítems contenidos en los cuestionarios de CVRS pueden agruparse en dos grupos distintos:

- Los **Ítems Causales** evalúan síntomas de una determinada enfermedad ya que la ocurrencia de estos síntomas producen cambios en la CVRS. Sin duda, los síntomas o efectos colaterales relativos a una determinada enfermedad pueden afectar la CVRS del paciente la cual, consecuentemente, será más pobre. Parece entonces razonable asumir que los síntomas causan deterioro de la CVRS.
- Los **Ítems Indicadores** reflejan el nivel de CVRS. Por ejemplo, la ansiedad y la depresión reflejan la CVRS del paciente. Si un sujeto posee una pobre CVRS como resultado de una enfermedad, puede sentirse deprimido o ansioso y por lo tanto estos ítems sirven como indicación de que el paciente tiene una CVRS baja.

Los Ítems Causales y los Indicadores tienen diferentes relaciones con la CVRS. Según Fayers *et al* (*op.cit.*), muchos de los ítems contenidos en los instrumentos para medir CVRS son causales mientras en la mayor parte de los instrumentos psicométricos están presentes ítems indicadores. Las escalas psicométricas tradicionales intentan cuantificar conceptos como la inteligencia o estados como la ansiedad o la depresión. Por otra parte, las escalas clínicas, inciden, en general, sobre los síntomas de una enfermedad y otros índices clínicos. Aunque exista una distinción clara entre las escalas psicométricas tradicionales, escalas clínicas y las escalas que miden CVRS, la mayor parte de las escalas de CVRS reúne características de las escalas psicométricas y de las escalas clínicas ya que contienen tanto ítems causales como ítems indicadores.

Siendo la CVRS un constructo latente, no es directamente observable o mensurable. La CVRS es medida a través de ítems causales que afectan la CVRS

(síntomas como vómitos o dolor). Sin embargo, el impacto producido por los síntomas en la CVRS varía de paciente a paciente. Además de eso, la percepción del propio paciente sobre la severidad de los síntomas, puede ser influenciada por diferentes factores. Por ejemplo, un paciente que sufre de dolor puede estar más sensible a problemas adicionales como vómitos. Otro tipo de indicadores de la CVRS puede ser la ansiedad, la depresión y el insomnio, los cuales pueden ser vistos como manifestaciones del nivel de CVRS. Sin embargo, tales relaciones deben ser analizadas con cuidado, no solamente porque la distinción entre causa e indicador no siempre es clara, sino también porque existen correlaciones entre los ítems. En realidad, muchos ítems de los cuestionarios de CVRS poseen interrelaciones complejas.

En muchos instrumentos para medir CVRS se asume que el deterioro en el estado de salud general implica una reducción del nivel de CVRS, la cual se manifiesta por un aumento de síntomas como ansiedad, tensión e irritabilidad. Sin embargo, factores externos no relacionados con la enfermedad o su tratamiento pueden influenciar la CVRS. Por ejemplo, si un paciente teme que la terapia no funcione puede empezar a sentirse deprimido y ansioso y tales ítems indicadores pasan a ser causales, cambiando el estado de CVRS percibida.

Fayers *et al* (1997) subrayan que la naturaleza causal de algunos ítems en escalas de CVRS puede originar factores inesperados y algunas veces inestables, difiriendo de estudio a estudio, de acuerdo con la enfermedad o terapia. Estando el Análisis Factorial basado en el análisis de la matriz de correlaciones entre los ítems, no permite distinguir con eficacia los dos tipos de ítems, mientras un Modelo de Ecuaciones Estructurales permitirá distinguir los ítems causales en situaciones restringidas. Teniendo en cuenta que temporalidad y causalidad están íntimamente relacionadas, quizás la forma más eficaz para establecer una relación causal, está en base al hecho de que una variable precede a otra en el tiempo, aunque esta relación no siempre sea fácil de establecer, como sucede

frecuentemente en el contexto de la CVRS (por ejemplo, no es trivial determinar si la ansiedad precede un cambio en la CVRS o viceversa).

Fayers *et al* (1997) sugieren que, en CVRS siempre es posible investigar la asimetría inherente a la relación entre los ítems causales y el constructo latente. Los cuestionarios de CVRS frecuentemente envuelven muchos ítems que, en principio, son causales y se espera que, un deterioro en cualquier de estos ítems, produzca un descenso en la CVRS. Sin embargo, no es posible afirmar que cualquier disminución en la CVRS sea debida a un ítem específico. En términos epidemiológicos, se puede decir que los ítems causales constituyen condiciones suficientes para alterar la CVRS. Sin embargo, los ítems individuales no constituyen condiciones necesarias al cambio de la CVRS. La clasificación de los ítems en causales e indicadores, no puede hacerse en base al sentido común. El método propuesto por estos autores se basa en el cruce de los ítems del cuestionario con una medida de CVRS global, en el cálculo del estadístico ji-cuadrado y en el análisis del grado de asimetría de esta relación, donde resultan algunos aspectos importantes:

- Se espera que los **ítems indicadores** posean alta correlación con el coeficiente global de CVRS. Un ítem indicador perfecto, tendría correlación 1 con la CVRS global, mientras un ítem causal presentaría correlación 1 con la CVRS global únicamente si fuera el único factor que influenciara o cambiara la CVRS del paciente. En realidad, existen muchos factores que influyen en la CVRS, tanto los relacionados con la enfermedad como con el tratamiento, y también factores extraños independientes. Así, en la práctica, se espera que pacientes para los cuales este ausente un determinado ítem causal, presenten una gran variabilidad de niveles de CVRS, mientras pacientes con valores elevados del ítem, presenten una pobre CVRS.

- Los ítems causales tienden frecuentemente a presentar valores medios bajos. Esto pasa porque muchas veces representan síntomas que no ocurren con frecuencia.
- Teniendo en cuenta que los ítems indicadores miden el mismo constructo, se espera que presenten correlaciones entre sí. Este es en realidad el fundamento de muchas de las técnicas utilizadas en el desarrollo y validación de escalas. Por otra parte, los ítems causales no tienen necesariamente que estar correlacionados entre sí. Por ejemplo, no existe razón para esperar que pacientes que sufren de diarrea sufran también de falta de aire o vómitos. Alguna correlación detectada entre estos ítems es consecuencia del tratamiento y no del nivel de CVRS.
- Un ítem causal representa únicamente uno de factores que pueden perjudicar la CVRS. Consecuentemente, un ítem causal aislado no constituye una medida fiable de CVRS. De la misma forma, un paciente puede considerarse que posee una pobre CVRS aunque no presente efectos adversos con respecto a cualquier ítem específico. Sin embargo, los ítems causales permiten identificar pacientes con baja CVRS ya que, la presencia de un síntoma implica el deterioro de la CVRS. Los ítems causales no son, en general, eficientes en la medición de una alta CVRS, mientras un ítem indicador es eficiente en la medición tanto de una alta como baja CVRS. Por lo tanto, ambos ítems son eficientes en la detección de una baja CVRS aunque solamente los ítems indicadores aseguren la detección de una alta CVRS.
- Cuando una escala se basa, sobretodo, en ítems causales, es importante asegurar que la mayor parte de los aspectos relativos a síntomas, toxicidad y efectos colaterales están cubiertos. Este aspecto es importante ya que, una respuesta alta con respecto a alguno de estos ítems potencialmente causales, es suficiente para producir una reducción en la CVRS. Por este motivo, muchos cuestionarios contienen

un módulo específico para una enfermedad o tratamiento conteniendo una alta proporción de ítems causales, como es el RSCL utilizado en el estudio que hemos desarrollado.

En suma, Fayers *et al* (1997) demuestran que los cuestionarios de CVRS contienen no solamente ítems indicadores (como se asume en la Teoría Psicométrica Tradicional) sino también ítems causales, los cuales se comportan de forma distinta. La distinción entre ítems causales e indicadores es de fundamental importancia en el diseño y validación de nuevos instrumentos. Los ítems indicadores originan escalas homogéneas con alta fiabilidad, mientras los ítems causales deben ser tratados con más cuidado.

En esta línea Vicente-Galindo (2003) ha propuesto un procedimiento para distinguir variables indicadoras de variables causales, que está basado en el análisis de tablas de contingencia con información de Calidad de Vida.

En cuanto al formato de las respuestas, los cuestionarios de CVRS pueden diferir. El formato más ampliamente utilizado es la **Escala de Likert** que contiene un número limitado de respuestas ordenadas con una descripción asociada a cada nivel. En este caso, la puntuación final, es la suma (o el valor medio) de las puntuaciones atribuidas a los diversos ítems. Existen también variaciones de estas escalas en que únicamente a los extremos se asocia una descripción. Estudios demuestran que los individuos pueden discriminar como máximo entre 7 y 10 categorías, mientras la sensibilidad al cambio decrece cuando se utilizan menos de 5 niveles (Streiner & Norman, 1995). Aunque las escalas más utilizadas sean las Escalas de Likert, también es común recurrir a **Escalas Dicotómicas** o a **Escalas Visuales Analógicas**. Las Escalas Visuales Analógicas, consisten en una línea, en general con 10 cm, con descripciones asociadas a cada extremo de la línea (usualmente “mejor CVRS posible” y “peor CVRS posible”). Efectivamente se pensaba que una medida continua sería más ventajosa en la discriminación que una Escala de Likert. Sin embargo, la experiencia no confirmó este hecho,

revelando que las Escalas Visuales Analógicas presentan algunas limitaciones porque requieren una coordinación ojo mano, de la que no siempre el paciente dispone y, además de eso no pueden ser aplicadas por teléfono o entrevista.

Guyatt *et al* (1993) hacen también la distinción entre las medidas en forma de Perfil y las medidas en forma de Índice. Los **Perfiles** consisten en escalas que reflejan dimensiones múltiples de la CVRS como el bienestar físico, emocional, funcional y social de los pacientes. Los **Índices** pueden consistir en una cuestión simple para acceder a la evaluación de la CVRS del sujeto o bien combinar respuestas múltiples de forma a proporcionar un índice simple de CVRS. La obtención de un índice simple tiene la ventaja de proporcionar información que sirve directamente para tomar decisiones, cuestión que asume particular importancia en contextos de ensayos clínicos. Por otra parte, es imposible construir un índice que agregue las múltiples dimensiones de CVRS y funcione en todos los contextos. Por lo tanto, un índice simple de CVRS no es el indicado para captar cambios en los dominios individuales. A su vez, los perfiles permiten evidenciar la naturaleza multidimensional de la CVRS. Ejemplos de medidas de perfil son el SF-36 (instrumento genérico) y RSCL (instrumento específico) que hemos utilizado en el estudio desarrollado.

La elección de una medida para CVRS depende, antes que nada, del propósito de estudio. Las medidas genéricas pueden revelarse particularmente útiles en estudios cuyo propósito sea documentar la extensión de alguna limitación en una población general o en un grupo de pacientes. Por otra parte, las medidas específicas se revelan clínicamente sensibles cuando son aplicadas a grupos específicos de pacientes. Algunas medidas específicas pueden ser utilizadas en conjunto con el fin de obtener una comprensión del impacto de diferentes intervenciones en la CVRS (Guyatt *et al*, *op.cit.*).

Guyatt *et al* (1992) subrayan que las escalas de CVRS deben ser validadas en términos de dos propósitos: comparaciones longitudinales o comparaciones transversales. En este contexto, Hyland (2003) sugiere que, en un estudio longitudinal, el cuestionario sea corto (máximo 40 ítems) y los ítems sean de respuesta múltiple (escalas de más o menos 7 puntos), mientras en los estudios transversales los cuestionarios sean más largos (entre 20 a 100 ítems) y de respuesta simple (de preferencia binarios). Por otra parte, en un estudio longitudinal los ítems describen problemas comunes a la mayor parte de los pacientes o a la población en estudio, debiendo ser relevantes para todos los pacientes. En un estudio transversal los ítems deben cubrir todos los aspectos de la CVRS ya que no necesitan ser relevantes para todos los pacientes. Estos autores subrayan que la selección de un instrumento de CVRS no es una ciencia exacta ya que es difícil predecir la eficacia de una escala de antemano.

## 2.5. EL PROBLEMA DE LOS DATOS FALTANTES

La ocurrencia de datos faltantes es también un problema inherente al análisis de datos de CVRS para el cual se necesitan técnicas adecuadas. En ensayos clínicos es muy frecuente que, por alguna razón, los pacientes no completen el cuestionario o que simplemente se queden incompletos determinados ítems. El hecho de que raramente todos los pacientes contesten a todos los ítems en todos los instantes es una de las mayores dificultades en la implementación de estudios de CVRS. En general, los datos faltantes se relacionan con tres aspectos distintos:

- algunos pacientes pueden morir durante el estudio;
- algunos pacientes pueden abandonar el estudio por algún motivo relacionado o no relacionado con el tratamiento;
- en general los pacientes entran en el estudio en diferentes instantes, pudiendo no haber participado en todas las evaluaciones (mecanismos de censura).



Los mecanismos causantes de datos faltantes deben ser tenidos en consideración durante el proceso de estimación de los efectos del tratamiento y del tiempo ya que ignorar tales mecanismos producirá resultados sesgados. Algunos instrumentos, tienen ya subyacente una estrategia para la atribución de puntuaciones en la presencia de respuestas faltantes en una pequeña proporción de cuestiones. La mayor parte de los métodos destinados a abordar este problema, asume que el patrón de datos faltantes es **aleatorio**, o sea, no sistemático. En oposición, un patrón **no aleatorio** (habitualmente notado por **MNAR**, *missing not at random*) o sistemático de datos faltantes significa que, por alguna razón, los casos incompletos difieren de los casos completos. Dentro de los datos faltantes aleatorios se suelen distinguir dos tipos de situaciones: los datos faltantes **aleatorios** y los datos faltantes **completamente aleatorios** (habitualmente designados por **MAR**, *missing at random*, y **MCAR**, *missing completely at random*, respectivamente). Si los datos faltantes ocurren de forma idéntica en todas las variables, o sea, si la ocurrencia de estos valores es independiente de las variables, los datos faltantes se dicen **MCAR**. Si por el contrario los valores faltantes se relacionan sistemáticamente con alguna(s) variable(s), estos se dicen **MAR**.

Vriens & Melton (2002) califican los métodos destinados a manejar datos faltantes en cuatro grupos distintos:

- 1) Los **procedimientos tradicionales** de imputación asumen que los datos faltantes son MCAR, siendo los más utilizados el *listwise deletion* y el *pairwise deletion*. El *listwise deletion* es el método que se utiliza con mayor frecuencia y consiste en trabajar únicamente con los casos para los cuales se dispone de información completa para todas las variables, excluyendo todos los demás del análisis. Aunque este método presente la ventaja de que todo el análisis será conducido con el mismo número de casos, también es verdad que el número de casos

disponibles para el análisis puede ser significativamente inferior a la muestra original. A su vez, en el *pairwise deletion* apenas se excluyen los casos para los cuales existen datos faltantes en las variables implicadas en un determinado análisis, lo que significa que el tamaño de la muestra puede cambiar de un análisis al otro.

- 2) Los **métodos de imputación simple** asumen también que los datos faltantes son MCAR y consisten en sustituir el valor faltante por una estimación. El método más común es imputar la respuesta faltante en base a la **media** de las demás respuestas en una subescala específica, siempre que al menos la mitad de las cuestiones haya sido contestada. Este método es simple, sin embargo tiende a distorsionar la distribución subyacente al conjunto de datos, reduciendo la variabilidad. Otro método de imputación simple es la imputación por **regresión**, que consiste en predecir el valor de una variable  $Y$  a partir de un grupo de variables  $X_1, X_2, \dots, X_n$ , eliminando las observaciones con datos incompletos y ajustando una ecuación de regresión para predecir  $Y$ . El *pattern matching* consiste en sustituir la observación faltante por otro caso con perfil semejante en las puntuaciones relativas a las demás variables. Otro método de imputación es el *hot-deck* que consiste en separar los registros completos de los incompletos y dentro de cada uno de estos dos grupos identificar casos con perfiles semejantes. El objetivo es llenar los registros vacíos (receptores) con información de campos con información completa (donantes), y los datos faltantes se reemplazan a partir de una selección aleatoria de valores observados, lo cual no introduce sesgos en la varianza del estimador.

- 3) En la **imputación basada en el modelo**, el método más conocido es el **Algoritmo EM** (*Expectation-Maximization*) desarrollado por Dempster *et al* (1977). El Algoritmo EM es una técnica general que permite estimar los datos faltantes haciendo inferencia sobre los parámetros. Esta técnica se centra en la relación existente entre los datos que faltan y los parámetros desconocidos de un determinado modelo de datos:
- si conociéramos los valores faltantes entonces sería fácil estimar los parámetros de modelo;
  - de forma similar, si conociéramos los parámetros del modelo de datos, entonces sería posible obtener estimadores no sesgados para los valores faltantes.

Esta interdependencia entre los datos faltantes y los parámetros del modelo, sugiere un método iterativo entre los pasos de “estimación” y “maximización” donde primero se predicen los valores que faltan en base a valores asumidos para los parámetros y después se utilizan esas predicciones para actualizar las estimaciones del parámetro. Se demuestra que cada iteración aumenta el valor de la función verosimilitud y por lo tanto la secuencia de estimadores de máxima verosimilitud converge para un máximo local. Nótese que la tasa de convergencia de los estimadores de máxima verosimilitud es determinada por la cantidad de información que falta en el conjunto de datos: si no hay información faltante, la convergencia será inmediata, mientras si faltan grandes cantidades de información sobre uno o más parámetros entonces la convergencia requerirá más iteraciones.

- 4) Los métodos de **imputación múltiple** en general asumen que los datos faltantes son MAR y no consisten en eliminar casos incompletos o imputar datos faltantes, sino en utilizar métodos de simulación de

Monte Carlo y sustituir los datos faltantes a partir de simulaciones. En cada simulación se analiza la matriz de datos completos a partir de métodos estadísticos convencionales y posteriormente se combinan los resultados para generar estimadores robustos, su error estándar e intervalos de confianza.

Cualquiera que sea la situación y la metodología utilizada, los datos reales son siempre preferibles a los obtenidos por mecanismos generadores de datos faltantes. Por otra parte, dado que el número de medidas repetidas de CVRS en ensayos clínicos puede ser particularmente elevado, especialmente en estudios de larga duración, estas aproximaciones pueden requerir un conjunto de parámetros bastante largo, por lo que la función de verosimilitud puede no contener información suficiente para la estimación de los parámetros.

Fayers & Machin (2000) subrayan la importancia de evaluar si los datos faltantes son o no aleatorios y que proporción de datos faltantes es aceptable en un estudio de CVRS. En algunas situaciones la existencia de datos faltantes puede ser informativa en el sentido de que nos revela algo sobre el paciente en cuestión. La ocurrencia de datos faltantes produce sesgo y, de esta forma, los resultados obtenidos no traducen la verdadera situación. En las situaciones en que la proporción de datos faltantes es grande, debe analizarse si las características de los pacientes con datos faltantes son diferentes de los restantes.

Otro problema analizado por Fayers & Machin (*op.cit.*), es la ocurrencia de ítems faltantes. Datos provenientes de la experiencia, revelan que, en general, la proporción de no respuestas a los ítems, varía entre 0,5% y 2,0%. El análisis del patrón permitirá percibir si la ocurrencia de estos datos es aleatoria o si, por lo contrario, pacientes que no contestan a determinados ítems tienden a no contestar a otros ítems. En términos formales, se asume que las no respuestas a un determinado ítem son aleatorias si la probabilidad de ocurrir un dato faltante es

independiente de la ocurrencia de datos faltantes en otros ítems. Al hilo de esta problemática, estos autores citan como ejemplo el RSCL para el cual la pregunta relativa a la disminución del interés sexual, suele registrar una proporción de no respuestas entre los 4% y los 14%. Una explicación factible es que los pacientes que experimentan este tipo de problemas se muestran más resistentes a contestar a esta cuestión, siendo por lo tanto más probable la ocurrencia de no respuestas cuando el problema se verifica. Por otra parte, para pacientes que no se encuentran sexualmente activos hace algún tiempo, el ítem no tiene sentido. Otro aspecto habitualmente constatado es que, en las personas mayores, la probabilidad de encontrar no respuestas es más elevada que en los más jóvenes. Se revela pues de extrema importancia el análisis cuidadoso de las no respuestas, pudiendo este revelar aspectos importantes de la CVRS de los pacientes.

## CAPÍTULO III

---

# EL FENÓMENO DE *RESPONSE SHIFT*

---

### 3. EL FENÓMENO DE *RESPONSE SHIFT*

#### 3.1. CAMBIO EN CALIDAD DE VIDA RELACIONADA CON LA SALUD

En la obtención de medidas de autoevaluación en estudios longitudinales, se asumen muchas veces dos supuestos:

- a los juicios emitidos por los individuos en estudio está subyacente un estándar interno;
- este estándar interno se mantiene constante a lo largo del tiempo.

También en gran parte de los estudios relativos a la medición de la CVRS percibida, se asume la existencia de una métrica común entre las evaluaciones y la estabilidad de los estándares internos lo que no siempre es verdadero. Recientemente, varios autores han abordado esta problemática subrayando que, cambios en el estado de salud, conducen a cambios en los estándares internos subyacentes a la evaluación de la CVRS (Visser *et al*, 2000; Middel *et al*, 2006; Osborne *et al*, 2006; Westerman *et al*, 2008; Andrykowski *et al*, 2009; Balain *et al*, 2009; Barclay-Goddard *et al*, 2009). Obviamente que, al registrarse un cambio en el constructo que se está evaluando, respuestas a las mismas preguntas dejan de ser comparables. Si por una parte, las Teorías Psicométricas tienden a considerar las diferencias individuales en el uso de las escalas como fuentes de error, en la aplicación de escalas de CVRS, este fenómeno asume gran importancia. En realidad, la CVRS no puede ser medida solamente en términos de “puntuaciones verdaderas” y “puntuaciones latentes” y las diferencias individuales en el proceso de evaluación de la CVRS no pueden ser vistas como fuentes de error, sino como un proceso intrínseco al propio proceso de evaluación de CVRS (Schwartz & Rapkin, 2004).

Antes de nada, conviene especificar la diferencia entre dos aspectos que no siempre coinciden:

- la percepción del estado general de salud;
- el estado de salud actual de un paciente.

En realidad, el individuo que está enfermo y percibe su enfermedad puede, después de un período de adaptación, retomar sus objetivos o adaptarlos, readquiriendo un bienestar general idéntico al del periodo anterior a la enfermedad. Otra persona, en la misma situación, puede no adaptarse y progresivamente volverse más insatisfecha con la vida y, consecuentemente, la CVRS percibida va empeorando (Pimentel, 2004; Sinclair & Blackburn, 2008). Por ejemplo, Kaarlola *et al* (2003) han desarrollado un estudio longitudinal para evaluar la CVRS en pacientes de Unidades de Cuidados Intensivos que se habían quedado con algún grado de invalidez. Este estudio ha revelado que, seis años después de los tratamientos, estos pacientes percibían su CVRS como más pobre al nivel del desempeño físico, dolor y estado de salud general, sin embargo, se sentían menos limitados al nivel emocional. Además de eso, seis años después, más de la mitad de los pacientes ha evaluado su CVRS global como mejor o igual que antes.

El análisis del cambio a través del tiempo es una cuestión de suma importancia en estudios de CVRS en los cuales se pretenden detectar modificaciones en las valoraciones de los individuos a lo largo del tiempo. En un abordaje inicial en el estudio del cambio, puede hacerse la distinción entre cambios cualitativos y cambios cuantitativos.

Un **cambio cualitativo** es más sencillo de evaluar ya que, en general, envuelve respuestas dicotómicas. A su vez, el **cambio cuantitativo** es más complejo de evaluar. Muchas veces la evaluación de los cambios cuantitativos se limita a una comparación de los valores medios de las respuestas de los



individuos en los distintos momentos en estudio. Sin embargo, si las dimensiones en cuestión no son estables, tal procedimiento no tiene sentido ya que, en realidad, son distintas las dimensiones que se están evaluando. Por lo tanto, antes de nada, es necesario hacerse la distinción entre **cuatro tipos de estabilidad**, conforme propone Taris (2000):

- **Invarianza Estructural** – se puede definir invarianza estructural como el grado de continuidad en la naturaleza de un determinado fenómeno en estudio (Mortimer *et al*, 1982). En la misma línea, Golembiewski *et al* (1976) afirman que dos (o más) estructuras factoriales son iguales cuando tienen el mismo patrón de los factores de carga, pero no es necesario que la magnitud de los factores de carga se mantenga. Invarianza estructural implica pues que el constructo en cuestión se caracterice por el mismo número de dimensiones a lo largo del tiempo y por el mismo patrón de relación entre sus componentes (los factores latentes están definidos por variables relativas al mismo aspecto del constructo con importancia similar en los grupos en cuestión, aunque la magnitud de los factores de carga no se mantenga), o sea, se considera que hay invarianza estructural cuando se mantiene constante la estructura factorial en cada momento o situación experimental.

Si se confirma la existencia de invarianza estructural, se presupone que el constructo en cuestión no ha cambiado y es posible evaluar si los resultados son estables y/o consistentes. Se evalúa pues la existencia de estabilidad normativa y/o cuantitativa:

- **Estabilidad Normativa** – se relaciona con la estabilidad observada en los rangos de orden de las preferencias de los individuos sobre un atributo de interés a lo largo del tiempo (Kagan, 1980). En particular,

autocorrelaciones elevadas entre las medidas en cuestión indican alto nivel de estabilidad normativa en cuanto que autocorrelaciones bajas indican cambios significativos en el orden de los pacientes a lo largo del tiempo, presuponiendo que el significado del atributo no ha cambiado.

- **Estabilidad Cuantitativa** – se refiere al nivel de estabilidad en la magnitud o cantidad media en la preferencia de un atributo dado a lo largo del tiempo (Baltes & Nesselroade, 1973). En general, se evalúa a través del cambio en los valores medios asociados a los grupos a lo largo del tiempo.
  
- **Consistencia Intraindividual** – se relaciona con el nivel de estabilidad en las preferencias de los individuos, o sea, se refiere a la consistencia en el orden que establece cada individuo (Emmerich, 1968). Se evalúa a través del coeficiente de correlación de los rangos de orden de los atributos en los diferentes momentos de tiempo, para cada individuo.

Es importante tener en cuenta que la evaluación de los niveles de estabilidad normativa, consistencia intraindividual y estabilidad cuantitativa sólo tiene sentido cuando el constructo en cuestión es estructuralmente invariante. Por lo tanto, antes que nada, el primer paso es probar la invarianza estructural.

Las personas cambian y sus valores también, lo que puede reflejarse en inversiones en los estándares internos de los individuos, en sus valores y prioridades o en la conceptualización de la CVRS percibida, las cuales son catalizadas por alteraciones en el estado de salud (Schwartz & Sprangers, 1999, 2002). Por ejemplo, es muy frecuente que las prioridades de pacientes con enfermedades crónicas cambien. Varios estudios indican que los valores existenciales, espirituales y sociales se vuelven más importantes. Cambios en la

autopercepción de CVRS resultan, en parte, del proceso de adaptación a la situación en que estos pacientes se ven (Addington-Hall & Kalra, 2001; Wiel *et al*, 2007; Sinclair & Blackburn, 2008). Los estándares internos, los valores y la conceptualización de calidad de vida puede cambiar a lo largo de la trayectoria de una enfermedad, siendo estos cambios inherentes al proceso de acomodación a la enfermedad (Sprangers & Schwartz, 1999b, Schwartz & Sprangers, 2002; Schwartz *et al*, 2007). Por lo tanto, medir CVRS en personas con alguna limitación, asumiendo que los estándares internos se mantienen constantes lo largo del tiempo, puede ser problemático teniendo en cuenta que frecuentemente ocurren cambios que afectan los índices psicométricos como la validez o la fiabilidad.

En realidad, el impacto producido por cambios en el estado de salud en la calidad de vida de un individuo ha ganado una importancia creciente en investigación clínica y a este proceso se le llama habitualmente *response shift* (Schwartz & Sprangers, 1999; Middel *et al*, 2006; Osborne *et al*, 2006; Sinclair & Blackburn, *op.cit.*; Westerman *et al*, 2008; Andrykowski *et al*, 2009; Balain *et al*, 2009; Barclay-Goddard *et al*, 2009). Sin embargo, esta problemática no es exclusiva del contexto de la Salud. En otros contextos, evaluaciones paradójicas conducen a consideraciones similares. Postulart & Adang (2000) mencionan estudios que revelan que un grupo de personas a quienes había tocado la lotería, no se revelaba significativamente más feliz que el grupo de control.

Aunque se pueden considerar diferentes concepciones y diferentes puntos de vista en el abordaje de la problemática del Cambio, tomaremos la concepción de Golembiewski *et al* (1976) y Schwartz & Sprangers (1999). Según la tipología de Golembiewski *et al* (*op.cit.*) se consideran tres tipos de cambio distintos: el **cambio Alpha**, el **cambio Beta** y el **cambio Gamma**. Por otra parte, Schwartz & Sprangers (1999) introducen el concepto de *response shift* en el contexto de la Salud, el cual es entendido como un cambio en el significado de una autoevaluación y comporta tres componentes distintas: **recalibración**, **repriorización** y **reconceptualización**.

### EL CAMBIO ALPHA

Es el verdadero cambio e implica una variación en el nivel de algún estado existencial, relacionado con un dominio conceptual constante, teniendo en cuenta que el instrumento de medición debe estar constantemente calibrado. Este tipo de cambio es el convencional y es justo el concepto opuesto al de estabilidad cuantitativa o estabilidad de nivel. Se evalúa observando si se ha registrado cambio en la media.

### EL CAMBIO BETA

Presupone una variación en los intervalos de medida de las dimensiones de una realidad estable y se debe a cambios en la escala de medida del individuo y/o en el punto de referencia, tratándose pues de una **recalibración** en la escala de medida del encuestado. Este cambio ocurre con frecuencia en contextos de CVRS y envuelve diferentes estimaciones o percepciones de la realidad por parte de los pacientes como ocurre por ejemplo en procesos de adaptación. Implica una variación en el nivel de algún estado existencial, complicado por el hecho de que algunos intervalos del continuo de medida asociado con el dominio conceptual constante, han sido recalibrados.

### EL CAMBIO GAMMA

Supone una redefinición del significado del constructo. Puede implicar una **reconceptualización** del significado del constructo o una **repriorización** de sus dimensiones. En la perspectiva de Golembiewsky *et al* (1976) el cambio Gamma es precisamente el opuesto de invarianza estructural.

En ausencia de invarianza estructural, puede haberse producido un cambio Beta o un cambio Gamma. Por lo tanto, hay que valorar si el constructo en estudio es estable y han cambiado las unidades de medida o bien si el propio dominio ha

cambiado y no tiene sentido considerar las mismas unidades de medida. Así que, para clasificar un cambio hay que saber si:

- el constructo en cuestión es constante a lo largo del tiempo;
- la calibración del instrumento de medida es constante en las diferentes evaluaciones.

Antes de intentar detectar un cambio Alpha hay que descartar la existencia de los cambios Gamma y Beta. De igual modo, antes de evaluar la posible ocurrencia de un cambio Beta, hay que estar seguros de que no se ha producido un cambio Gamma.

## 3.2. *RESPONSE SHIFT*

### 3.2.1. ¿Qué es *Response Shift*?

El concepto de *response shift* se refiere a alteraciones en los estándares internos de medida, en los valores y en la propia definición de calidad de vida para el sujeto. Este fenómeno se verifica por ejemplo en personas que ven sus capacidades seriamente afectadas por problemas de salud y, sin embargo, muchas veces consideran poseer un nivel alto de calidad de vida, como han demostrado numerosos estudios (Reynolds & Prior, 2003; Nieuwkerk *et al*, 2007; Schwartz *et al*, 2007; Mayo *et al*, 2008; Pucciarelli *et al*, 2008; Ridder *et al*, 2008; Sinclair & Blackburn, 2008; Westerman *et al*, 2008; Wiles, 2008; Pinson *et al*, 2009).

El fenómeno de *response shift* se relaciona con la separación aparente entre el funcionamiento físico y el bien estar psicológico y es definido por Sprangers & Schwartz (1999a) como un cambio en el significado de la autoevaluación sobre un determinado constructo *target*, como resultado de:

- un cambio en los estándares internos de medida, o sea, una **recalibración** de la escala (en el sentido psicológico);

- un cambio en los valores, o sea, se altera la importancia relativa de los dominios de las componentes que constituyen el constructo *target* (**repriorización**);
- una redefinición del constructo *target*, o sea, una **reconceptualización**.

El concepto de *response shift* ha tenido origen en el contexto educativo (Howard *et al*, 1979a, 1979b, 1979c) y organizacional (Golembiewski *et al*, 1976). Mientras Howard *et al* (*op.cit.*) definen *response shift* en términos de cambio en los estándares internos de medida, Golembiewski *et al* (*op.cit.*) introducen la componente de reconceptualización. A la recalibración de la tipología de Schwartz & Sprangers (1999) está subyacente el cambio Beta de Golembiewski *et al* (*op.cit.*) y a los fenómenos de repriorización y reconceptualización está inherente el cambio Gamma. Aunque la investigación del fenómeno de *response shift* haya tenido su origen en la esfera de la Educación, posteriormente ha recibido una importancia creciente en el ámbito de la Psicología de la Salud.

Korfage *et al* (2007) definen *response shift* como una adaptación a algún cambio al nivel de la salud y presentan el ejemplo de una mujer a la que se pide que evalúe su calidad de vida. Al evaluar su calidad de vida ella piensa en su trabajo que le ocupa 50 horas semanales, en su pareja, en los partidos de voleibol con las amigas y evalúa su calidad de vida como “muy buena”. Desafortunadamente, esta mujer se enferma. Durante algunos meses no puede trabajar. El apoyo de su pareja se vuelve muy importante. Gradualmente ella se va recuperando y empieza a trabajar 20 horas a la semana. Sin embargo ya no puede jugar al voleibol. Esta mujer evalúa nuevamente su calidad de vida como “muy buena”. Esta vez los principales aspectos que ha considerado han sido su pareja, su familia y su trabajo. Aunque su evaluación no ha cambiado, ocurrieran tres formas de *response shift*: el trabajo ha pasado de 50 horas semanales a 20 (recalibración), su pareja ha pasado a ser más importante (repriorización) y su

concepto de Calidad de Vida ha cambiado: desaparece el deporte y cobran mayor importancia la familia (reconceptualización).

Jansen *et al* (2000) abordan también las tres componentes del fenómeno *response shift*. La **recalibración de escala** es la primera componente del *response shift*. Según estos autores, los pacientes poseen una percepción interna de su nivel funcional con respecto a una dimensión dada. Este estándar de medida conduce a una determinada comprensión de los puntos que forman la escala de respuesta. El significado de expresiones como “la mayor parte de las veces” o “algunas veces” está basado en su percepción y experiencia particular. Nuevas experiencias a lo largo del tiempo pueden disminuir o aumentar su consciencia de las posibilidades existentes, lo que puede originar que la percepción de la escala de respuesta se extienda o se condense para encajar la nueva experiencia. Los sujetos pueden utilizar diferentes estándares internos de medida como resultado de un evento específico (como una terapia, un curso o un entrenamiento). Por ejemplo, pacientes con cáncer relatan con frecuencia que ahora conocen el significado de “estar exhausto”, aunque antes del tratamiento pensarán que “estaban exhaustos”. Ahora consideran que antes del tratamiento “se sentían simplemente cansados”.

La segunda componente del *response shift* es un cambio en los valores de respuesta (**repriorización**) a lo largo del tiempo. La CVRS (así como muchos otros constructos) es constituida por diferentes dominios como la salud física, psicológica o social. Jansen *et al* (*op.cit.*) subraya que la importancia de los diferentes dominios puede cambiar como resultado de cambios en el estado de salud. Es común que, personas con diagnóstico de cáncer, pasen a atribuir menos importancia a la salud física al percibir que el soporte que encuentran en la familia y los amigos es extremadamente importante. En realidad el concepto de CVRS no ha cambiado ya que el constructo sigue estando constituido por los mismos dominios. Sin embargo, ha cambiado la importancia atribuida a los diferentes dominios.

La tercera componente del *response shift* se relaciona con una redefinición del constructo, o sea, una **reconceptualización**. Ocurre reconceptualización siempre que los dominios que constituyen el constructo *target* cambian a lo largo del tiempo. Jansen *et al* (2000) refieren que, el paciente a quien ha sido diagnosticado cáncer, puede descubrir un nuevo apoyo, por ejemplo, en la religión, añadiendo el dominio de la salud espiritual a su percepción de CVRS, al mismo nivel de la salud física, psicológica y social.

Frecuentemente, resultados de ensayos clínicos indican que pacientes sometidos a cirugía o quimioterapia adyuvante registran alteraciones en los estándares internos. En comparación con las medidas de cambio convencionales, el cambio basado en el *response shift* puede ampliar la mejoría total en todos los ramos de tratamiento pero disminuir las diferencias entre los efectos del tratamiento en los diferentes ramos de los estudios. Por otra parte, según Sprangers & Schwartz, (1999b) es posible encontrar un patrón común en los resultados de varios ensayos clínicos, revelando que la ocurrencia de *response shift* puede darse en dos sentidos distintos: pacientes que han experimentado una mejoría de su estado de salud después de un tratamiento tienden, retrospectivamente, a ampliar el malestar que sentían antes del tratamiento, mientras los pacientes que han empeorado, retrospectivamente, tienden a minimizar los niveles de malestar inicial. Este hecho levanta otra cuestión importante: la adaptación del paciente a la toxicidad del tratamiento o a la mejoría inducida por este, puede originar evaluaciones imprecisas de la Calidad de Vida a lo largo del tiempo. Un estudio con pacientes de cáncer de mama sometidas a radioterapia desarrollado por Jansen *et al* (*op.cit.*) ha evidenciado que el *response shift* puede conducir a estimaciones atenuadas de los efectos colaterales de la radioterapia si los pacientes se adaptan a la toxicidad del tratamiento. El *response shift* puede atenuar o aumentar los efectos de un tratamiento.



Aunque el fenómeno de *response shift* no sea nuevo desde una perspectiva clínica, es un fenómeno relativamente reciente desde el punto de vista metodológico y teórico. En realidad, el *response shift* tiene implicaciones potencialmente considerables ya que, en términos metodológicos, un cambio en los estándares internos, valores o conceptualizaciones de CVRS, puede originar evaluaciones completamente diferentes que no son comparables a lo largo del tiempo. Si las unidades de medida cambiaron, la comparación ha perdido el significado. Por ejemplo, en una situación de evaluación de tratamientos, el *response shift* puede atenuar las estimaciones de los efectos del tratamiento si los pacientes se adaptan a tales efectos (Sprangers & Schwartz, 1999b). Por lo tanto, es importante considerar el fenómeno de *response shift* en las evaluaciones de tratamientos, sobre todo cuando existe la tendencia para subestimar o sobrestimar sus efectos. Asimismo, los métodos utilizados para identificar *response shift* se revelan útiles no solamente para identificar efectos sesgados de tratamientos, sino también para analizar el impacto de la enfermedad a lo largo del tiempo. El estudio del *response shift* permite una mejor comprensión de cómo la calidad de vida es afectada por cambios en el estado de salud, así como por intervenciones médicas. Clínicamente, el análisis del *response shift* puede ayudar a comprender el fenómeno de respuesta completa a la enfermedad y tratamientos y el proceso subyacente de reajuste y reevaluación cuando la salud ya no forma parte del estado homeostático. Además de eso, el *response shift* puede ayudar a entender algunos problemas clínicos tales como desordenes somáticos e hipocondriasis, los cuales pueden interpretarse como resultado de una deficiente respuesta adaptativa normal (Sprangers & Schwartz, *op.cit.*).

### 3.2.2. El Modelo Teórico del *Response Shift*

De acuerdo con el modelo teórico de Sprangers & Schwartz (1999a), cambios en la autoevaluación de CVRS pueden resultar de la interacción entre:

- un **agente catalítico**, como un cambio en el estado de salud, resultante o no de un tratamiento;
- **antecedentes**, o sea, características particulares individuales como por ejemplo los rasgos de personalidad (optimismo, autoestima, sentido de control, etc.), variables sociodemográficas (sexo, educación, etc.), las expectativas o la identidad espiritual;
- **mecanismos**, o sea, procesos comportamentales, cognitivos o afectivos, inherentes al estado de salud, como por ejemplo, compararse a sí mismo con los otros, redefinir objetivos o adoptar rutinas diferentes;
- ***response shift***, o sea, cambio en el significado de la evaluación de su propia CVRS, resultante de cambios en los estándares internos, valores o conceptos;
- **calidad de vida percibida**, la cual se define como un constructo multidimensional que incorpora al menos tres amplios dominios: físico, psicológico y social. Estos dominios asumen grados de importancia distintos para diferentes individuos y para grupos de pacientes específicos, dependiendo de los dominios funcionales afectados por la enfermedad.

Como hemos mencionado, *response shift*, es el fenómeno de adaptación interna, según el cual, los pacientes cambian sus estándares internos, valores o concepto de CVRS, pasando a entenderla de forma diferente a la percepción que tendrían si no se hubieran adaptado a la situación. Estos cambios en los estándares internos, en los valores o en la conceptualización de CVRS, son catalizados por cambios en los estados de salud e influyen el proceso de acomodación a la

enfermedad y la CVRS percibida. El hecho de que el individuo perciba una baja CVRS, puede conducir al desarrollo de nuevos mecanismos de adaptación, tratándose por lo tanto de un proceso iterativo y dinámico. Según Ridder *et al* (2008), aunque muchos pacientes logren el ajuste psicológico, se estima que el 30% de los pacientes prolonga la fase de ajuste, muchas veces sin éxito.

No está todavía claro porque algunas personas presentan *response shift* y otras no. A pesar de que el fenómeno de *response shift* ha recibido poca atención hasta poco tiempo atrás, tiene importantes implicaciones al medir CVRS y asume una importancia extrema en las Ciencias Sociales y Médicas.

Según Rapkin & Schwartz (2004), el fenómeno de *response shift* conduce a cuatro componentes distintas en el proceso de evaluación:

- inducir un marco de referencia;
- recordar y seleccionar las principales experiencias;
- utilizar estándares de comparación para evaluar experiencias;
- utilizar un algoritmo subjetivo para priorizar y combinar evaluaciones en la escala de CVRS.

Estas componentes de evaluación pueden estar relacionadas con la cultura, personalidad, situación y pueden variar de persona a persona y a lo largo del tiempo. Por eso, cualquier puntuación de CVRS es ambigua si no tenemos en cuenta este proceso.

También Postulard & Adang (2000) entienden el fenómeno de *response shift* como un proceso de adaptación que se refiere a las estrategias individuales utilizadas para superar la enfermedad y restaurar el bienestar subjetivo. En la misma línea, Visser *et al* (2000) y Reavley *et al* (2009) entienden el fenómeno de *response shift* como consecuencia de un mecanismo de adaptación deseable que permite, por ejemplo, atenuar los efectos del tratamiento cuando los pacientes se adaptan a sus efectos o progresión de la enfermedad a lo largo del tiempo. Un

cambio en una o varias estrategias cambiará la percepción retrospectiva de CVRS, por ejemplo, después de un trasplante. Si un determinado tratamiento produce un cambio de actitud del paciente hacia su estado de salud, la interpretación de las medidas de salud debe ser cautelosa ya que, después de la intervención, la percepción del paciente ha cambiado debido a las estrategias de adaptación.

Por otra parte, estudios han demostrado que pacientes integrados en un grupo de personas con estados de salud mejores y peores, los lleva a cambiar sus estándares internos. Las intervenciones en grupo llevan a que estos pacientes reconsideren sus objetivos y prioridades, cambiando sus valores y a percibir que es posible tener una CVRS razonable incluso en una condición peor, reconceptualizando la CVRS y evidenciando que la intervención psicológica puede inducir *response shift* (Wiles, 2008; Reavley *et al*, 2009).

Albrecht & Devlieger (1999) conceptualizan el equilibrio entre el cuerpo, la mente y las componentes espirituales y las relaciones de los individuos con su entorno físico para demostrar que la salud y la percepción de CVRS resulta del balance de estos factores. En las personas con enfermedades que perciben una baja CVRS, el cuerpo tiende a ser el aspecto dominante, mientras las que, aún así, perciben un alto grado de CVRS, centran su atención en las relaciones con los demás y tienden a revalorizar la componente espiritual.

En un estudio llevado a cabo con pacientes de cáncer de pulmón, Westerman *et al* (2008) analizaron las estrategias de respuesta de estos pacientes al ser cuestionados sobre sus problemas y limitaciones. Los autores han observado que, en realidad, las respuestas de estos pacientes sugieren menos limitación que lo que de hecho se observó, habiendo concluido que pacientes terminales evalúan su funcionamiento en términos de lo que perciben ser normal bajo las circunstancias. Al estudiar longitudinalmente la percepción que un grupo de pacientes con la Enfermedad de Paget sobre su CVRS, Robertson *et al* (2009) han concluido que el cuadro de referencia cambia, cambiando también el significado

que atribuyen por ejemplo a los ítems de un cuestionario y luego la percepción que estos pacientes poseen de su CVRS. Por lo tanto, el cambio en el proceso de evaluación, debe ser tenido en cuenta a la hora de interpretar las respuestas de estos pacientes. Winter *et al* (2009) han encontrado resultados similares en un grupo de personas mayores.

Schwartz *et al* (2007) subrayan que, vivir con alguna discapacidad es un proceso de cambio constante y de ajuste constante que es difícil medir y categorizar. Existe una tensión dinámica entre la respuesta a la discapacidad por el individuo y la respuesta a la discapacidad en cuestión por los observadores externos.

### 3.2.3. Estudios que evidencian *Response Shift*

En los últimos años se han multiplicado los estudios que ponen de manifiesto la ocurrencia del fenómeno de *response shift*. Es muy común detectar discrepancias entre las evaluaciones de los propios pacientes. Por ejemplo, un estudio ha revelado que, pacientes esperando por un trasplante de riñón, al evaluar su CVRS, registraron un valor medio de 5,23 en una escala de 10 puntos. Este valor subió hasta 7 después del trasplante. Sin embargo, 5 meses, 12 meses y 18 meses después del trasplante, se ha pedido a estos pacientes que evaluaran su CVRS retrospectivamente, antes del trasplante y se obtuvieron valores medios de 3,27, 3,14 y 3,05, respectivamente. Antes del trasplante estos pacientes se han adaptado con éxito a su condición y por eso atribuyeron una puntuación más elevada a su CVRS que más tarde, al pedirse que la reevaluaran después del trasplante, cuando su estado de salud había mejorado (Addington-Hall & Kalra, 2001).

Otro estudio, desarrollado por Postulard & Adang (2000), con el objetivo de detectar la ocurrencia de *response shift* en un grupo de pacientes sometido a un

trasplante de riñón, ha revelado diferencias significativas entre la evaluación antes del trasplante y la evaluación retrospectiva después del trasplante. Estos autores consideran que el punto de referencia cambia y, desde este punto de vista, la situación antes de la intervención es entendida como más desfavorable. Medidas retrospectivas de CVRS reflejan tanto los efectos del tratamiento como los efectos de la adaptación.

Visser *et al* (2000) han llevado a efecto un estudio que pretendía identificar la ocurrencia de *response shift* en términos de cambio en los estándares internos de medida con respecto a la percepción de fatiga en un grupo de pacientes sometidos a radioterapia. Los resultados han revelado que, después del tratamiento, los pacientes han evaluado retrospectivamente la fatiga sentida antes como menos intensa que en la evaluación hecha en el *Pre-Test*. Por otra parte, no se han encontrado diferencias al comparar el *Pre-Test* y el *Post-Test*, evidenciando la importancia de identificar cambios en los estándares de medida, en la evaluación de la CVRS. Posteriormente, Michielsen *et al* (2007) han llevado a efecto un estudio similar y han encontrado resultados semejantes.

Rees *et al* (2005) ha desarrollado un estudio con pacientes de cáncer de próstata cuyo objetivo era el de evaluar la ocurrencia de *response shift* tanto en los pacientes como en sus parejas. A los pacientes y sus parejas se les ha pedido que completaran el *Prostate Cancer Patient & Partner Questionnaire* (PPP), después del diagnóstico y nuevamente 3 y 6 meses después, incluyendo una evaluación retrospectiva en las dos últimas aplicaciones. El estudio ha revelado que los pacientes con cáncer de próstata evalúan como más baja su CVRS cuando lo hacen retrospectivamente 3 y 6 meses después. Por otra parte, en algunos casos, las parejas han evaluado la CVRS más bajo que los pacientes. Rees *et al* (*op.cit.*) han observado *response shift* en los pacientes y en sus parejas. El *response shift* ha sido, en algunos casos, más evidente en las parejas que en los propios pacientes, sobre todo en lo que respecta a la ansiedad y a la dimensión social.

Al investigar la ocurrencia de *response shift* en pacientes con cáncer de próstata, Korfage *et al* (2007) han encontrado que estos pacientes evalúan retrospectivamente su CVRS como más alta de lo que habían hecho antes del diagnóstico. Por otra parte, retrospectivamente, evalúan como peor la CVRS en el principio del tratamiento. Los autores concluyen que el *response shift* ha sido inducido primero por el diagnóstico y después por el tratamiento. Además, el *response shift* inducido por el diagnóstico ha sido más significativo que el *response shift* inducido por el tratamiento.

Ring *et al* (2005) han evaluado el impacto de implantes dentarios en la calidad de vida de un grupo de pacientes. Para eso han aplicado un cuestionario antes del implante y 3 meses después, incluyendo a los 3 meses una evaluación retrospectiva. Los resultados no han indicado un impacto significativo del tratamiento. Sin embargo, la evaluación retrospectiva ha revelado que 3 meses después, los pacientes evalúan retrospectivamente su calidad de vida como más baja que lo que habían considerado en ese momento. Ring *et al* (*op.cit.*) concluyen que el impacto positivo de los implantes dentarios únicamente es visible cuando se tiene en cuenta el fenómeno de *response shift*.

En un estudio desarrollado en grupo de pacientes con cáncer metastizado, Sharpe *et al* (2005) concluyen que el fenómeno de *response shift* es común entre pacientes que reciben un diagnóstico de cáncer. Al pedirles que refiriesen las cinco áreas más importantes en su vida, Sharpe *et al* (2005) han observado que la mitad de los pacientes ha invertido sus prioridades 3 meses después del diagnóstico. Sin embargo, el fenómeno de *response shift* no siempre está asociado a una adaptación positiva (Osborne *et al*, 2006). En particular, Sharpe *et al* (*op.cit.*) han concluido que este fenómeno se ha revelado positivo para los pacientes que habían pasado a valorar más algún dominio que antes no valoraban tanto y que se ha revelado positivo para su CVRS. Sin embargo, el *response shift* ha sido negativo para los pacientes que han enfocado su atención en alguna área en la que la CVRS se había deteriorado. Al evaluar el impacto del dolor y fatiga en pacientes de cáncer en fase

avanzada, otros autores han encontrado resultados similares en lo que respecta a la dirección (positiva o negativa) del *response shift* (Chow *et al*, 2007; Echteld *et al*, 2007; Hagelin *et al*, 2009).

En un grupo de pacientes con cáncer de pulmón no operable, Broberger *et al* (2006) han investigado la ocurrencia de *response shift* como resultado de la adaptación al deterioro del estado de salud. Estas autoras han examinado cambios en los estándares internos de los pacientes al nivel de la fatiga, CVRS global y funcionamiento físico 3 y 6 meses después del diagnóstico. Los resultados obtenidos no han permitido concluir claramente la ocurrencia de cambios al nivel de los estándares internos de los pacientes (recalibración), concluyéndose que, debido a la naturaleza de los síntomas asociados a esta enfermedad, los pacientes podrían haberse adaptado a los síntomas antes del estudio.

En los últimos dos años numerosos estudios han surgido en que el objetivo principal era identificar la ocurrencia de *response shift* en grupos de pacientes con diagnósticos de cáncer colon-rectal (Pucciarelli *et al*, 2008), cáncer de mama (Andrykowski *et al*, 2009), cáncer de pulmón (Westerman *et al*, 2008; Santana *et al*, 2009), accidente cardiovascular (Mayo *et al*, 2008), artritis reumatoide (Sinclair & Blackburn, 2008), esclerosis múltiple (Wiles, 2008; Pinson *et al*, 2009), fractura de rodilla (Razmjou *et al*, 2009; Balain *et al*, 2009) y otras enfermedades crónicas (Ridder *et al*, 2008; Winter *et al*, 2009).

### **3.2.4. El Fenómeno de la Paradoja de la Discapacidad**

Recurriendo al concepto de *response shift*, Schwartz & Sprangers (1999) abordan un fenómeno denominado Paradoja de la Discapacidad ("*disability paradox*"). Este fenómeno ha sido descrito inicialmente por Levine (1987) al verificar que muchas personas con incapacidades serias y persistentes, consideran poseer una buena o incluso excelente CVRS cuando las circunstancias evidencian precisamente el contrario. Muldoon *et al* (1998) subrayan que la forma en la que



los pacientes evalúan su CVRS puede cambiar a lo largo del tiempo. Muchos pacientes de cáncer asocian a la enfermedad relatos que van desde una mayor capacidad de apreciar cada día a una mayor sensación de fuerza personal, evidenciando a veces una mayor satisfacción con su calidad de vida que grupos de comparación sanos. Podríamos concluir que el cáncer ha mejorado su CVRS. En realidad, esta paradoja refleja el proceso de *response shift* que ocurre frecuentemente en pacientes de cáncer y otras enfermedades crónicas como diabetes, insuficiencia renal, etc. Este bienestar subjetivo refleja adaptación psicológica.

Como hemos mencionado, cambios drásticos en el estado de salud producen muchas veces en el individuo cambios en los patrones de medida (recalibración), cambios en sus valores (repriorización) e incluso cambios en el propio concepto de CVRS (reconceptualización). El fenómeno de *response shift* explica los resultados inconsistentes obtenidos en varios estudios que evalúan CVRS en el dominio de la Psicooncología. Por ejemplo, varios estudios refieren disminuciones en el nivel de CVRS en pacientes de cáncer de mama (el grupo de pacientes de cáncer más ampliamente estudiado), mientras otros estudios no encuentran diferencias entre mujeres con cáncer de mama y mujeres sanas (Wiel *et al*, 2007). Andrykowski *et al* (2009) subrayan la misma idea al medir longitudinalmente los niveles de fatiga en pacientes de cáncer de mama. Estos autores han encontrado que, la ocurrencia de *response shift* sugiere que no ha ocurrido cambio en los niveles de fatiga de estas pacientes a lo largo del tiempo, enmascarando el cambio real. Santana *et al* (2009) han encontrado resultados similares en un grupo de pacientes con cáncer de pulmón.

Stiefel *et al* (2008) han evaluado el significado de la vida utilizando la escala SMiLE (*Schedule for Meaning in Life Evaluation*) en una muestra de pacientes con cáncer, comparando los resultados a los obtenidos en una muestra de estudiantes. Sorprendentemente, estos autores concluyeron que los pacientes con cáncer relataban los mismos niveles de satisfacción que los estudiantes. También Albrecht

& Devlieger (1999) han estudiado una muestra de 153 pacientes con diversas enfermedades, constatando que el 54,3% de los encuestados con problemas moderados o graves percibían una excelente o buena calidad de vida, confirmando la existencia de la Paradoja de la Discapacidad. Delante de tales resultados, estos autores preguntaban por qué tantas personas con discapacidades serias y persistentes consideraban experimentar una calidad de vida buena o excelente, cuándo para la mayor parte de los observadores externos, estas personas parecían vivir una existencia indeseable. Los autores han concluido que las relaciones entre cuerpo, mente y espíritu y el establecimiento de relaciones sociales son aspectos importantes en la explicación de estas contradicciones aparentes. La Paradoja de la Discapacidad contradice así la idea de que la percepción de una buena calidad de vida implica tener buena salud y de que una persona enferma no puede percibir una elevada calidad de vida.

En suma, el fenómeno de *response shift* puede explicar la aparente paradoja en las puntuaciones obtenidas cuando se pide a estos pacientes que evalúen su CVRS. Varias investigaciones han permitido concluir que pacientes con enfermedades graves, evaluaban positivamente su calidad de vida.

### **3.2.5. *Response Shift* y Comparación Social**

Según Fayers *et al* (2007), el proceso de respuesta a un cuestionario de CVRS es complejo: la información utilizada por el sujeto puede ser objetiva (como la presencia o ausencia de un determinado síntoma) o subjetiva. La información subjetiva deriva del conocimiento interno de los pacientes y de su experiencia pasada, produciendo un cuadro de referencias personales al evaluar CVRS. O sea, las puntuaciones atribuidas por los individuos dependen de sus experiencias de vida y del contacto con otras personas. Fayers & Sprangers (2002) subrayan que no está claro lo que los pacientes tienen en mente cuando contestan a una pregunta como “¿Cómo ha sido su calidad de vida global durante la semana pasada?”.

Como hemos mencionado anteriormente, los pacientes recurren a diferentes cuadros de referencia y en realidad la respuesta resulta de comparaciones implícitas con otras personas o con ellos mismos en otro instante. En general, las comparaciones que los pacientes tienen en mente incluyen personas sanas (como familia o amigos), otros pacientes (“En comparación con los pacientes que he visto en la clínica, estoy muy bien”) o con ellos mismos antes de la enfermedad. Fayers *et al* (2007) han realizado un estudio con el objetivo de identificar las comparaciones implícitas y cuadros de referencia que los pacientes tienen en mente y percibir cual es el impacto que tales comparaciones tienen en respuestas sobre bienestar o CVRS. Estos autores han concluido que los pacientes se referían a un año atrás, al periodo anterior a la enfermedad o se comparaban a otras personas (sanas o con la misma enfermedad). Las evaluaciones obtenidas estaban fuertemente asociadas al cuadro de referencia utilizado, siendo los que utilizaban como referencia personas con la misma enfermedad, los que percibían mejor CVRS y los que se referían a su estado antes de la enfermedad que percibían una CVRS más baja.

En parte, el fenómeno de *response shift* puede deberse a comparaciones sociales. Algunos pacientes evalúan su CVRS a través de comparaciones con otros pacientes y no con personas sanas. La comparación social es común en el contexto de enfermedades crónicas (Dibb & Yardley, 2006), pudiendo asumir un papel mediador entre eventos que ocurren y que son significativos para el sujeto y la ocurrencia de *response shift* (Gibbons, 1999; Ubel *et al*, 2003; Fayers *et al*, *op.cit.*). En general, los problemas de salud intensifican el proceso de comparación social, lo que conduce a cambios en la autoevaluación y consecuentemente a la ocurrencia de *response shift*. La sensación de dolor o placer resulta, en parte, de comparaciones que hacemos con otras personas o con experiencias que hemos vivido en otros instantes del tiempo. Ubel *et al* (2003) presentan el siguiente ejemplo: las primeras semanas conduciendo un coche nuevo son muy agradables porque tendemos a

comparar el coche nuevo con el coche antiguo. De forma análoga, la experiencia del paciente con una enfermedad, sobre todo si es repentina, puede ser muy angustiante cuando compara su estado de CVRS actual con el anterior. A lo largo del tiempo, esta comparación va perdiendo importancia y la angustia se va disipando. A lo largo del tiempo, la importancia de esta comparación va disminuyendo. De la misma forma, cuando los pacientes contactan con otros pacientes, algunos de ellos con síntomas más severos que ellos mismos, sus niveles de angustia disminuyen muchas veces debido a la comparación hecha. Personas exteriores ignorarán estos efectos, lo que llevará a una subestimación de la CVRS del paciente. Los eventos negativos, tienden a desencadenar la comparación social, alterando hábitos en este procedimiento como consecuencia de alteraciones en la comparación de dimensiones y niveles. El tipo de alteración en el proceso de comparación social que ocurre, puede influenciar el tipo de *response shift* resultante.

Gibbons (1999) ha comparado la percepción que un grupo de estudiantes con dificultades académicas tenía de sus propios problemas al observar individuos con problemas de salud serios, concluyendo que, observando los problemas de salud de otros, las dificultades académicas de un individuo asumen una importancia totalmente diferente. Según este autor, el fenómeno de *response shift* tiene, en general, un objetivo de adaptación, aunque esta función adaptativa pueda aparecer enmascarada por manifestaciones comportamentales. Por otra parte, no existen evaluaciones empíricas sobre la utilidad del *response shift* en situaciones controlables *versus* no controlables. Gibbons (*op.cit.*) refiere además que el proceso de comparación social puede volverse peligroso en escenarios en que el individuo se encuentra en situaciones de inevitable deterioro. Así pues, en situaciones en que el deterioro es inevitable, el proceso de comparación social puede motivar un *response shift* negativo. Gibbons (*op.cit.*) refiere la depresión y el optimismo como factores moderadores del papel de la comparación social en el

*response shift*. En particular, tener una perspectiva positiva puede constituir un predictor del *response shift* positivo, después de algún evento negativo. El optimismo actúa como una especie de parachoques contra el deterioro del autoconcepto o *response shift* negativo (Santana *et al*, 2009).

Schwartz & Sendor (1999) han investigado el efecto producido en un grupo de personas que prestaban ayuda por teléfono a pacientes con esclerosis múltiple, habiendo detectado también la ocurrencia de *response shift* en los que prestaban ayuda a estos pacientes. Curiosamente, estos autores han detectado beneficios más significativos al nivel psicosocial, de adaptabilidad, bienestar, autoconfianza y autoestima en los prestadores de ayuda que en los propios pacientes. En particular, estas personas han mejorado significativamente en dominios positivos como la autoeficacia, satisfacción global con la vida, crecimiento personal y objetivos de vida y también en dominios negativos como depresión, fatiga o limitaciones físicas.

Además del proceso de comparación social, Ubel *et al* (2003) relacionan el fenómeno de *response shift* con el cambio de expectativas. Supongamos que una persona de 30 años y otra de 95 evalúan su CVRS en 95 en una escala de 0-100. En realidad, sus respuestas no representan el mismo nivel de salud. Una puntuación de 95 en un anciano puede reflejar su expectativa de la mejor CVRS posible en su edad. De la misma forma, pacientes con enfermedades crónicas se adaptan físicamente y emocionalmente a sus estados de salud. Una parte de esta es física ya que cuando una persona pierde una determinada facultad, tiende a desarrollar habilidades alternativas. Otra parte de esta adaptación es psicológica y pasa por la reducción de las expectativas.

### 3.3. ¿CÓMO MEDIR *RESPONSE SHIFT*?

Cómo hemos mencionado, integrar el fenómeno de *response shift* en la investigación de la CVRS es fundamental para una mejor comprensión de cómo la CVRS es afectada por cambios en el estado de salud y para el desarrollo de medidas validas que permitan cuantificar cambios en CVRS. Evaluar el efecto de una intervención comparando simplemente los valores medios relativos a las puntuaciones antes y después de la intervención puede conducir a resultados erróneos. Hay que tener en cuenta que la intervención puede afectar los estándares internos de evaluación de los participantes en el estudio y haber ocurrido *response shift*. Según Visser *et al* (2000), un cambio en los estándares internos puede originar evaluaciones a lo largo del tiempo que son incomparables ya que las unidades de medida han cambiado y la comparación ha perdido su significado. Para medir explícitamente los efectos del *response shift* en la evaluación de tratamientos e intervenciones hay que integrar medidas que permitan acceder al cambio en los estándares internos. Desde una perspectiva teórica, integrar *response shift* en el estudio de la CVRS permitirá una mejor comprensión de cómo la CVRS es afectada por cambios en el estado de salud y cómo juega un papel importante en la adaptación del paciente a la enfermedad. Jansen *et al* (2000) subrayan también la importancia de desarrollar métodos que permitan acceder al *response shift*, posibilitando a los investigadores estudiar la eventual ocurrencia de adaptación, su dirección, magnitud y su impacto en la CVRS. Solo así podremos entender porque algunos pacientes se adaptan a su enfermedad, mientras otros no.

El hecho de que los pacientes tiendan a adaptarse a las circunstancias que cambian, como una enfermedad o tratamiento constituye un verdadero desafío en la medición y interpretación de medidas de autoevaluación de CVRS. Como hemos visto, resultados paradójicos de relatos de pacientes sobre su CVRS comparables a personas sanas, relatos de pacientes que evalúan mejor su CVRS que aquellos que les prestan cuidados, pueden explicarse por el fenómeno de la

adaptación. La adaptación puede cambiar el valor de referencia según el cual los pacientes evalúan su CVRS a lo largo del tiempo, lo que, a su vez, puede invalidar comparaciones longitudinales de CVRS en que se asume un valor de referencia constante (Nieuwkerk *et al*, 2007). Por otra parte, aunque muchos autores entiendan el fenómeno de *response shift* como un sesgo que debe ser eliminado antes de evaluar el verdadero cambio debido a un tratamiento, en el ámbito de la CVRS este fenómeno puede tener interés en si mismo ya que inducir *response shift* a un paciente con una enfermedad crónica puede ayudarlo a percibir su CVRS de forma más favorable (Schwartz *et al*, 2004; Ridder *et al*, 2008; Barclay-Goddard *et al*, 2009).

Las tres componentes del *response shift* están inevitablemente interconectadas y cualquier método para acceder a estos dominios debe respetar la interconexión entre estándares internos, valores y conceptualización. Este es un aspecto crucial para estudiar este fenómeno y para explicar algunas paradojas en la investigación de CVRS. Por otra parte, algunos investigadores en el dominio de los cambios Alpha, Beta y Gamma, expresan la interconexión entre las tres componentes del *response shift*, adoptando la siguiente jerarquía: primero es necesario descartar la ocurrencia de cambio Gamma antes de detectar cambio Beta; después es necesario descartar la existencia de cambio Beta, antes de poder detectar cambio Alpha. Esto pasa porque un cambio en los estándares internos perderá su significado si el constructo en análisis no es el mismo en los dos instantes en estudio, o sea, no es posible afirmar que ha ocurrido un cambio en los estándares internos sin afectar al significado del constructo. Si la escala de medida es expandida o contraída, entonces los extremos y los intervalos tienen diferentes significados.

Según Schwartz & Sprangers (1999), una cuestión muy importante y que debería ser mejor explorada tiene que ver con el hecho de que los pacientes no

siempre tienen conciencia del *response shift*. En realidad se trata de un fenómeno precognitivo y por tanto, cualquier método para evaluarlo debe centrarse en demostrar cambios en los estándares internos, valores y conceptualizaciones, en vez de pedir a los pacientes que se centren en estos aspectos directamente.

Evaluar *response shift* asume por eso una gran importancia en la investigación en CVRS. En realidad, se disponen de muy pocos métodos específicamente concebidos para detectar *response shift*. Por otra parte, los datos empíricos en el área de CVRS son escasos, por lo que existe una enorme necesidad de desarrollar y profundizar la investigación en esta área.

Integrando medidas del cambio en los estándares internos en la evaluación de la CVRS, los efectos del *response shift* pueden ser explícitamente medidos y tenidos en cuenta al comparar los efectos de diferentes tratamientos. En seguida se refieren algunos métodos habitualmente utilizados para evaluar CVRS y que pueden ser adaptados para medir diferentes aspectos de *response shift*.

### **3.3.1. Métodos Individualizados**

Los Métodos Individualizados se basan en la definición de CVRS por el propio individuo y requieren la identificación de los dominios relevantes en CVRS para cada paciente individualmente (Stiggelbout *et al*, 2008). Schwartz & Sprangers (1999) presentan como ejemplo de estos métodos, el *Cantrill's Ladder*, en el que se supone que la CVRS percibida se expresa en una escala del 0 al 10, donde 0 significa peor satisfacción posible con la vida y 10 la mejor posible. Se pide al paciente que describa los estados "peor posible" y "mejor posible". Otros ejemplos son el *Repertory Grid Technique* y el *Schedule for the Evaluation of Individual Quality of Life* (SEIQoL). El SEIQoL consiste en pedir al paciente que refiera los dominios más relevantes de su CVRS, evaluando enseguida su nivel de funcionamiento en cada uno de los dominios mencionados. Por fin estas evaluaciones son cuantificadas.



Browne *et al* (1997) han desarrollado una versión más simple de esta técnica, el SEIQoL-DW (*SEIQoL-Direct Weighting*) que consiste en pedir al paciente que rellene un diagrama de sectores en el cual el tamaño relativo de cada sector, representa el peso que el individuo atribuye a cada uno de los dominios de CVRS.

Aunque los métodos individualizados presenten algunas limitaciones (Stiggelbout *et al*, 2008), se identificaría *response shift*, aplicando estos métodos en un instante T1 y en un instante T2 y, además, en el instante T2, se podría pedir al paciente que haga la misma evaluación pero refiriéndose al instante T1. Esta es una forma de evaluar cambio en la descripción de los conceptos a lo largo del tiempo y en la relación de estos conceptos a lo largo del tiempo. Un cambio en la conceptualización puede indicar un cambio en el contenido de los dominios (conceptos) y refleja reconceptualización del constructo *target*. Por otra parte, un cambio en los valores puede sugerir un cambio en el orden de los dominios. Estos procedimientos permiten acceder a al menos dos de las tres componentes del *response shift*: cambio en los conceptos (reconceptualización) o cambio en los valores (repriorización).

El PGI (*Patient Generated Index*) es otro método individualizado que consiste en pedir al paciente que elija cinco áreas importantes en su CVRS y después que puntúe esas áreas reflejando su importancia. Cambios en el orden de los cinco dominios elegidos indican repriorización, cambios en el contenido de estos dominios refleja reconceptualización (Barclay-Goddard *et al*, 2009).

Las entrevistas permiten también obtener información sobre los objetivos de los pacientes y sobre sus autoevaluaciones acerca de de los mismos. Cambios en los objetivos entre una entrevista y otra, evidencian *response shift*.

La ventaja de estos métodos es que permiten que cada individuo defina las áreas de CVRS con importancia personal. Por otra parte, aunque posibiliten la identificación de *response shift*, no permiten una cuantificación de su efecto.

### 3.3.2. Utilidades y Preferencias

El concepto de utilidad que se utiliza actualmente en el área de la CVRS ha tenido origen en 1944 con la teoría económica de John von Neumann y Oscar Morgenstern (Torrance & Feeny, 1989). El principio subyacente a este paradigma es que los individuos tienen preferencias y toman decisiones mediante algún grado de incertidumbre. En el contexto de la CVRS, la utilidad refleja las preferencias por determinados estados de salud delante de la incertidumbre. Torrance *et al* (2001) diferencian preferencia y utilidad: preferencia es el concepto más amplio que comprende el concepto de utilidad. Utilidad es un tipo específico de preferencia medido bajo condiciones de incertidumbre.

Una de las medidas de utilidad más utilizadas son los años de vida ajustados por la calidad (QALYs – *Quality-Adjusted Life Years*). De hecho, la efectividad clínica puede expresarse en términos de mortalidad (años de vida conquistados, número de vidas salvadas), morbilidad (complicaciones o procedimientos evitados, días de trabajo perdidos, etc.), identificación de la enfermedad o terapia (número de casos diagnosticados, número de casos tratados apropiadamente) o combinando la morbilidad y mortalidad, o sea, a través de los años de vida ajustados por la calidad- QALYs (Campolina & Ciconelli, 2006).

La idea subyacente a estas medidas compuestas es asociar pesos a los diferentes estados de salud de forma que se refleje los niveles de salud asociados a estos estados. Una de las técnicas que ha surgido en el campo de la Psicología para asignar este tipo de datos ha sido la Escala Visual Analógica (VAS - *Visual Analog Scale*) que ha sido ampliamente utilizada en Medicina, sobre todo para evaluar el dolor. Como ya hemos visto anteriormente, la VAS es una línea con uno de los extremos coincidente con 0 y el otro con el 1. Habitualmente, el valor 0 equivale al peor estado de salud (muerte) posible y el valor 1 al mejor estado de salud posible (salud perfecta). El paciente marca en esta línea su estado de salud. Cada punto corresponde al valor de utilidad atribuido a cada uno de los estados indicados Sin

embargo, existen algunos inconvenientes de orden teórica y práctica asociados a la utilización de estas técnicas.

Krabbe *et al* (2007) han propuesto una metodología alternativa para generar valoraciones de estados de salud, basada en el orden de las diferencias entre los estados de salud combinados con un modelo de escalamiento asociado que transforma las órdenes individuales en valores agrupados en niveles intercalares. Esta estrategia consiste en presentar a los sujetos un conjunto simple de tareas o juicios y luego buscar la información necesaria para obtener las valoraciones del estado de salud. En particular, estos autores sugieren empezar por el cálculo de las diferencias en los valores de la VAS para cada par de estados de salud en base a los datos individuales. Después de obtener las órdenes de estas diferencias se utiliza el Escalamiento Multidimensional No Métrico para recalcular los valores de la escala para cada estado en base a las diferencias ordenadas transformando datos no métricos en datos métricos. Krabbe *et al* (*op.cit.*) han concluido que el escalamiento de las diferencias ordenadas entre los estados de salud, utilizando el Escalamiento Multidimensional produce valores cardinales prácticamente idénticos a la media de las valoraciones VAS. Los valores basados en las órdenes explican el 98% de la variabilidad de los valores VAS.

Otros ejemplos de estos métodos son el *Q-Twist* y el *Preference Mapping* cuyo principio es acceder a un listado de preferencias, valores del individuo u otros datos útiles. En suma, la idea subyacente a los Métodos de Preferencia es que el paciente se sitúe en una escala de CVRS. Aplicando estas técnicas en diferentes instantes, es posible detectar cambios en las preferencias y valores, permitiendo identificar *response shift* (Schwartz *et al*, 1997)

### 3.3.3. Métodos de Aproximación por Comparaciones Sucesivas

El principio en el que se basa la familia de los Métodos de Aproximación por Comparaciones Sucesivas, es la emisión de un conjunto de juicios sobre el orden de un conjunto de objetos a lo largo de un *continuum* psicológico o físico con el objetivo de percibir su importancia relativa. Un ejemplo es el *Pairwise Comparison*, que consiste en pedir a los pacientes que ordenen, según el grado de importancia, una serie de dominios por comparación sucesiva, a lo largo del tiempo. La alteración de la importancia relativa a lo largo del tiempo significa cambio en los valores y refleja repriorización. Si al paciente se permite también la selección de los dominios que están a ser comparados, también es posible identificar reconceptualización. El *Card Sort Approach* es un método semejante en el que se pide a los pacientes que ordenen una colección de cartones referentes a diferentes dominios de CVRS segundo su grado de importancia, posibilitando así la detección de cambios en los valores.

### 3.4. EL MÍNIMO CAMBIO CLINICAMENTE IMPORTANTE Y *RESPONSE SHIFT*

Como hemos mencionado, interpretar puntuaciones de CVRS cuando estas cambian a través del tiempo entre los mismos pacientes es todavía un desafío. Por una parte no existe un patrón (*gold standard*) para interpretar puntuaciones en estudios de CVRS y, por otro, en el análisis del cambio a través del tiempo no siempre una diferencia estadísticamente significativa tiene una gran relevancia clínica. Es por eso importante definir el concepto de “**mínimo cambio clínicamente importante**” (MCCI). Obviamente, la definición de lo que se considerará un cambio importante influenciará el juicio sobre el suceso de una intervención o tratamiento, el tamaño muestral necesario y el diseño del estudio (Shimozuma *et al*, 2007).

Badia & Carné (1998) definen MCCI como la diferencia en la puntuación más pequeña, en un dominio de interés, percibida por los pacientes como beneficiosa y que obligaría a un cambio en el tratamiento del sujeto en ausencia de efectos indeseables graves, o en presencia de un coste excesivo. De forma más general, otros autores definen MCCI como la diferencia más pequeña en las puntuaciones, que es percibida como importante, sea benéfica o nociva, y que llevará al clínico a considerar un cambio en el tratamiento del paciente (Jaeschke *et al*, 1989; Shun *et al*, 2007). Esto significa que estimar el MCCI implica también interpretar la percepción de los pacientes sobre lo que es un cambio importante. Varios autores coinciden sobre la importancia de conocer el MCCI en las puntuaciones de los instrumentos de CVRS en contextos de ensayos clínicos. Esto permitirá interpretar cambios clínicos detectados en los pacientes, permitirá determinar el tamaño de la muestra en un ensayo clínico y incluso ayudará a interpretar otros resultados de estudios en los que se incluyen medidas de CVRS. Además de eso, estimar el MCCI para un instrumento dado, es una forma de evaluar la sensibilidad al cambio (Badia & Carné, *op.cit.*; Shun *et al*, *op.cit.*; Wyrwich *et al*, 2002).

Varias técnicas han sido utilizadas para determinar la importancia clínica de cambios detectados en CVRS. Según Sloan *et al* (2003), filosóficamente, existen tres perspectivas para abordar el tema:

- una medida de la importancia clínica es el grado según el cual la CVRS de un paciente debe cambiar hasta que perciba que un cambio ha ocurrido;
- otra forma de medir el MCCI es en base a la opinión del médico sobre la magnitud del cambio en las puntuaciones de CVRS que justifiquen algún tipo de intervención clínica;

- finalmente, la perspectiva estadística/epidemiológica, según la cual el MCCI se basa en el análisis de la proporción de pacientes que experimentan cambios en las puntuaciones de CVRS.

Walters & Brazier (2003) sintetizan los estadísticos más utilizados como índices de sensibilidad que permiten cuantificar cambios clínicamente importantes:

1) El *Effect Size* relaciona el cambio a través del tiempo con la desviación estándar de las puntuaciones iniciales:

$$ES = \frac{\text{Cambio en las puntuaciones}}{\text{Desvío estándar en el instante 1}}$$

Este estadístico ignora la variación en el cambio.

2) *Estadístico t* para muestras apareadas es indicada para evaluaciones *Pre-Test/Post-Test* de eficacia conocida:

$$t = \frac{\text{Cambio en las puntuaciones}}{\text{Error estándar de la diferencia}}$$

Este estadístico ignora la información sobre la variación en las puntuaciones en los pacientes clínicamente estables (se consideran sujetos estables aquellos cuyo estado se mantiene constante a lo largo del tiempo).

3) *Respuesta Media Estandarizada* compara el cambio con la desviación estándar del cambio:

$$SRM = \frac{\text{Cambio en las puntuaciones}}{\text{Desvío estándar del cambio en las puntuaciones}}$$

4) *Coefficiente de Cambio (Responsiveness)* evalúa el cambio en CVRS según la variabilidad en los pacientes clínicamente estables:

$$\text{Cambio} = \frac{\text{Cambio en las puntuaciones}}{\text{Desviación estándar del cambio en las puntuaciones entre sujetos estables}}$$

Este estadístico ignora la información sobre la variación en las puntuaciones en los sujetos no estables clínicamente.

En la práctica, varios autores (Sloan *et al*, 2003; Walters & Brazier, 2003; Schünemann & Guyatt, 2005; Shimozuma *et al*, 2007; Shun *et al*, 2007) consideran dos aproximaciones complementarias para abordar este tema: una exclusivamente basada en el propio conjunto de datos disponible y otra en la teoría estadística. La primera perspectiva se basa en la relación entre una medida de CVRS y una medida independiente (ancla) para evaluar el significado de un cambio particular. Se utiliza por ejemplo la puntuación global de cambio en las puntuaciones de CVRS atribuidas por el propio paciente, relacionando la percepción del propio paciente sobre lo que es un cambio significativo en su estado de salud con el cambio correspondiente en la puntuación de CVRS. Esta aproximación es referida habitualmente como *anchor-based approach* una vez que requiere un estándar independiente o una ancla, interpretable y relacionada con el instrumento en cuestión. Por ejemplo, la importancia clínica de los cambios en CVRS está anclada a la autoevaluación del paciente sobre su CVRS.

La segunda perspectiva está basada en la distribución o métodos estadísticos y utiliza el “tamaño del efecto” (*effect size* - ES), que expresa la magnitud del efecto en términos de la desviación estándar de la distribución, cuantificando la superposición entre dos distribuciones y relacionando, en general,

grupos de tratamiento y control, según alguna medida de variabilidad. Cohen (1988) propone un índice  $d$  para calcular el ES:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{(S_1^2 + S_2^2)/2}}$$

donde  $\bar{X}_1$  y  $\bar{X}_2$  son las medias relativas a los instantes 1 y 2 y  $S_1^2$  y  $S_2^2$  las respectivas desviaciones estándar.

Una  $d$  entre 0,2 y 0,5 puede considerarse pequeña, un ES entre 0,5 y 0,8 es moderado y mayor que 0,8 se considera grande. Un ES pequeño indica una gran superposición, mientras un ES grande indica poca superposición. Un ES moderado indica que el cambio es clínicamente significativo para alguna dimensión, o sea, puede considerarse como criterio mínimo para el MCCI, 0,5 veces la desviación estándar. El ES de Cohen puede estar influenciado por el grado de homogeneidad o heterogeneidad en la muestra.

Razmjou *et al* (2009) utilizan el índice de Cohen para analizar la ocurrencia de *response shift* en un grupo de pacientes sometidos a artroplastia total de rodilla. En una primera fase, se identificaron los pacientes en los que había ocurrido *response shift* mediante la comparación de los valores del *Then-Test* y *Post-Test* (ver apartado 4.2.1). El ES obtenido se utilizó para examinar la magnitud de la recalibración en el grupo de pacientes en los que el *response shift* había ocurrido en sentido positivo y en los que había ocurrido en sentido negativo.

Schünemann & Guyatt (2005) argumentan que la terminología “mínimo cambio clínicamente importante” (MCCI) enfatiza la esfera clínica en vez de enfatizar la percepción del paciente. Según estos autores, cualquier instrumento para medir CVRS debe primeramente tener en cuenta la perspectiva del paciente, su experiencia en el día a día y no en la interpretación clínica. Por lo tanto,



Schünemann & Guyatt (2005) sugieren la eliminación de la “C” de esta terminología: “mínimo cambio importante” (MCI). En realidad es cada paciente que experimenta su propia CVRS y por eso sólo el paciente puede evaluar cuales los cambios realmente importantes. Schünemann & Guyatt (*op.cit.*) definen MCI como la diferencia mínima en la puntuación de una dimensión de interés que el paciente percibe como importante (sea benéfica o perjudicial) y que llevaría al paciente o al médico a cambiar algo.

Varios investigadores utilizan los valores del MCCI para interpretar mejor los cambios al nivel de puntuaciones obtenidas mediante instrumentos para medir CVRS y, en particular, saber si los cambios observados son clínicamente relevantes. Este análisis se complica si consideramos el fenómeno de *response shift*. Todas las aproximaciones anteriormente referidas para evaluar el MCCI, asumen que no ha ocurrido *response shift*. Un cambio en la CVRS puede ser subestimado si el fenómeno de *response shift* no es considerado, así como la importancia clínica de un cambio. Por lo tanto, en el cálculo del MCCI, es conveniente asegurarse de que las puntuaciones son precisas. Una posible área de investigación sería entender cómo el *response shift* y MCCI se relacionan y hasta qué punto el fenómeno de *response shift* en sí mismo es un cambio clínicamente importante.

## CAPÍTULO IV

---

# MÉTODOS ESTADÍSTICOS PARA ANALIZAR DATOS DE CVRS Y DETECTAR *RESPONSE SHIFT*

---

## 4. MÉTODOS ESTADÍSTICOS PARA ANALIZAR DATOS DE CVRS Y DETECTAR *RESPONSE SHIFT*

El método más utilizado en la literatura para detectar *response shift* es el *Then-Test* que consiste en pedir al paciente que haga una evaluación retrospectiva de su CVRS con el objetivo de detectar cambios. Posteriormente han surgido algunas aplicaciones de los métodos factoriales a datos de calidad de vida. Basándose estos métodos en el análisis de la estructura de varianza-covarianza de los ítems, alteraciones al nivel de esta estructura sugieren alteraciones al nivel de la autopercepción de los pacientes. En esta línea, métodos como el STATIS o el Procedimiento de Krzanowski pueden constituir una herramienta eficaz en la evaluación y detección de cambio. Vicente-Villardón *et al* (2004) proponen un modelo factorial aplicable a varias matrices de datos, procedentes de distintos tiempos, grupos o contextos. El modelo propuesto está basado en la existencia de factores comunes compartidos por todos los grupos y factores específicos para cada grupo. Mientras los factores comunes se obtienen de las direcciones compartidas de los subespacios latentes de cada grupo, los factores específicos se obtienen de los residuales tras la proyección del grupo en el subespacio común.

Zwinderman (1990) propone un modelo que permite algo que hasta entonces ningún otro método permitía: cuantificar el cambio ocurrido al nivel de la CVRS de pacientes en tratamiento, separando la componente debida al tratamiento y la debida al paso del tiempo. Sin embargo, este autor no considera la ocurrencia de cualquier cambio asociado al fenómeno de *response shift*.

Otro abordaje importante del problema de la detección del cambio está basado en los Modelos de Ecuaciones Estructurales. Dentro de esta familia de métodos, el abordaje más clásico se basa en las Curvas de Crecimiento Latente. Posteriormente, Oort (2005) propone un procedimiento basado en los Modelos de

Ecuaciones Estructurales que permite detectar *response shift*, aislando las diferentes componentes de recalibración, repriorización y reconceptualización.

Recientemente, Mayo *et al* (2008) han propuesto un método para identificar cambios al nivel individual. Mientras la mayor parte de los métodos sólo permiten identificar cambio cuando este es experimentado por una parte significativa de los pacientes en un grupo, el modelo propuesto por estas autoras permite identificar los sujetos que han experimentado *response shift*.

Una búsqueda en PubMed ha identificado más de 5000 trabajos que mencionaban el *Then-Test* en contextos de datos de calidad de vida. Por otra parte, las referencias al STATIS, al Procedimiento de Krzanowski o a aplicaciones del Modelo de Zwinderman son inexistentes en este contexto, mientras la modelización de datos mediante Modelos de Ecuaciones Estructurales es mencionada en 172 trabajos. Sin embargo, la aplicación de los Modelos de Ecuaciones Estructurales a la problemática específica de la detección del Cambio sólo es mencionada en 9 trabajos.

En este capítulo enfocamos la contribución que estos métodos pueden dar para la evaluación y tratamiento del Cambio en CVRS. Además de los métodos tradicionales, consideraremos algunos métodos ya conocidos que podrán constituir herramientas útiles en el análisis y detección del cambio y modelos especialmente concebidos o adaptados para identificar y analizar cambio. Muchas veces la mejor opción es combinar varios métodos, aprovechando los puntos fuertes de cada uno y atenuando sus puntos débiles. Hemos optado por limitar el desarrollo estadístico al mínimo y presentar los detalles técnicos en anexo.

#### **4.1. DATOS DE CVRS**

En el dominio de la Psicología y Ciencias de la Salud es común disponer de medidas repetidas relativas a secuencias de observaciones obtenidas en diferentes instantes o bajo diferentes tratamientos. Durante mucho tiempo, los estudios con medidas repetidas se restringían a los estudios experimentales. Sin embargo, los

estudios longitudinales son hoy ampliamente utilizados y constituyen foco de gran interés por parte de investigadores desde una perspectiva metodológica y estadística. Los datos de CVRS son, frecuentemente, datos longitudinales con características muy particulares. Así, al enfocar los métodos estadísticos para analizar datos de CVRS, importa también, considerar las técnicas estadísticas en las que habitualmente se recurre en el análisis de datos longitudinales.

Bono *et al* (2008) presentan un análisis de los principales modelos estadísticos utilizados en el tratamiento de datos longitudinales en el ámbito de la Psicología y Medicina. Estos autores han realizado una revisión bibliográfica de los artículos publicados durante el período 1985-2005 en las bases *PsycInfo* y *Medline*, concluyendo que la cantidad de estudios longitudinales ha aumentado significativamente. Se constató también que ha aumentado el uso de modelos multinivel. De acuerdo con los trabajos de Singer & Willett (2005, 2006) en base a una revisión de 10 revistas de la APA, en 1999 el 33% de los estudios publicados fueran longitudinales y en 2003 el 47%. Además de eso, una búsqueda en varias bases de datos, ha permitido constatar que este aumento es más notable en los dominios de la Medicina, Psicología y Física. Bono *et al* (*op.cit.*) refieren como principales factores justificativos del incremento de estudios de carácter longitudinal el desarrollo de técnicas de análisis avanzadas de modelización estadística y la amplia disponibilidad de programas informáticos.

En contextos de ensayos clínicos donde se pretende evaluar el impacto de terapias, por ejemplo en diagnósticos oncológicos, es común proceder a una evaluación longitudinal de la CVRS durante el periodo de tratamiento. En estos contextos cobra particular importancia el estudio del cambio. Muchas veces, a los pacientes que participan en estudios de CVRS se les pide que contesten a un cuestionario validado (un instrumento con varias escalas o diferentes instrumentos de medida que evalúan dominios generales o específicos de la CVRS), en varios instantes del periodo de tratamiento programado y después del

término del tratamiento. Los instantes de aplicación del cuestionario son en general determinados de forma que coincidan con el ciclo de tratamiento o con visitas programadas durante el período posterior. La evaluación de la CVRS termina en el instante previamente especificado en el protocolo o cuando el desarrollo de la enfermedad obliga a que el paciente sea sometido a otro tratamiento no establecido en el protocolo que confunde el efecto de la terapia original en la CVRS. Por todo esto, en la práctica, los datos de CVRS generan diversos problemas con cierto grado de complejidad, inherentes no solamente a la obtención del propio conjunto de datos como también a su análisis.

En primer lugar, al tratarse de datos longitudinales, el análisis de datos de CVRS es más complejo debido a la dependencia entre las medidas repetidas y, en segundo lugar, muchas veces el investigador no puede controlar las circunstancias bajo las cuales se obtienen medidas repetidas. Así, las evaluaciones pueden no coincidir con los intervalos de tiempo especificados en el protocolo, debido a retrasos en el tratamiento producidos, por ejemplo, por efectos colaterales del propio tratamiento o, simplemente, debido a problemas logísticos, los pacientes no han comparecido a las visitas programadas. Otro problema bastante común tiene que ver con la dificultad de evaluar la CVRS en pacientes que están siendo sometidos a tratamientos con grados de eficacia diferentes y que confunden la evaluación.

Cómo hemos referido en el capítulo anterior, gran parte de los instrumentos que tradicionalmente se utilizan en el análisis de CVRS, presuponen que el referencial del paciente es lo mismo en las diferentes valoraciones. Sin embargo, hemos visto que, en un contexto de CVRS, esta suposición no siempre se verifica. En realidad, las valoraciones del constructo conseguidas por autoinforme presentan una problemática muy particular ya que, en general, no es posible aislar, ni evaluar, el verdadero cambio en la valoración del constructo cuando el

individuo es sometido a una intervención experimental. Los cambios Beta y Gamma introducen un sesgo que necesariamente debe ser tenido en cuenta y aislado convenientemente, para poder medir los cambios reales. Asumen pues gran importancia métodos que permitan detectar *response shift*, detectar la ocurrencia de cambio y cuantificar el cambio ocurrido.

En los últimos años, han surgido metodologías alternativas para el análisis de datos longitudinales como los Modelos de Ecuaciones Estructurales, los Modelos de Series Temporales, los Modelos de Coeficientes Aleatorios, los Modelos de Curvas de Crecimiento. En el estudio desarrollado por Bono *et al*, (2008) antes citado se ha observado que estos modelos pasan a ser ampliamente utilizados a partir de la segunda mitad de la década de 90 en detrimento de los modelos tradicionales como el Análisis de la Varianza (ANOVA) con medidas repetidas cuya utilización disminuye lentamente en el ámbito de la Psicología y, de forma más contundente, en el campo de la Medicina.

No obstante del número creciente de estudios basados en el análisis de datos longitudinales, en la literatura revisada siguen siendo escasas las alusiones a los problemas subyacentes al trabajo con datos de CVRS y, en particular, en lo que respecta a la evaluación y tratamiento del cambio, y mucho más escasos los trabajos que tratan de aportar alguna solución al respecto. En los apartados siguientes consideramos algunas de las aportaciones en el sentido de medir cambios en CVRS.

## 4.2. MÉTODOS DE ANÁLISIS TRADICIONALES

Como hemos mencionado, los datos de CVRS están, en general, relacionados con medidas repetidas a lo largo de un intervalo de tiempo fijo. La mayor parte de las veces, el objetivo es captar cambios en las respuestas de los pacientes así como sacar conclusiones relativas al tipo de tratamiento más

ventajoso. Datos de CVRS pueden analizarse recurriendo a diversas metodologías. Quizás la forma más sencilla de tratar datos de CVRS sea utilizar estadísticos descriptivos cruzados, en los diferentes intervalos de tiempo, con el objetivo de describir la tendencia global de los diferentes grupos de tratamiento o sintetizar datos longitudinales en un estadístico resumen antes de efectuar comparaciones entre grupos de tratamiento. Sin embargo, estos métodos no tienen en cuenta variaciones intrínsecas a cada paciente y no captan cambios importantes en la CVRS a lo largo del tiempo, además de que producen incrementos del error Tipo I.

Con el objetivo de evaluar el nivel de CVRS en pacientes de cáncer sometidos a diferentes tratamientos, Fairclough (1997) propone un método basado en el uso de estadísticos síntesis, reduciendo considerablemente el número de contrastes. Este procedimiento permite controlar la tasa de error Tipo I, al revés de lo que pasa cuando el análisis de datos de CVRS se basa en tests univariantes a medidas individuales. Una forma de reducir la dimensionalidad de los datos es reducir un conjunto de  $n$  medidas individuales a un valor simple. Otra vía será a través de una media ponderada de las puntuaciones estandarizadas, donde las ponderaciones están basadas en la correlación entre las medidas.

Otra vía tradicionalmente utilizada en el análisis de este tipo de datos se basa en el estudio de las respuestas a los ítems con base en el Análisis Univariante de la Varianza (ANOVA) o el Análisis Multivariante de la Varianza (MANOVA) con medidas repetidas. Este tipo de análisis se basa en el cálculo de la suma de las puntuaciones atribuidas por los pacientes en cada instante en los ítems que miden cada variable latente, asumiendo que la suma de las puntuaciones es normalmente distribuida. Esto presenta algunos inconvenientes debido a la correlación producida entre medidas repetidas, lo que afecta la estimación de los errores estándar, ampliando los errores Tipo I. Además de eso, en el modelo ANOVA, se está asumiendo que las variables latentes son constructos unidimensionales y los ítems sobre los cuales se calcula la suma de las puntuaciones se refieren al mismo



constructo. Sin embargo, la CVRS es un constructo de naturaleza dinámica y multidimensional. Para resolver los problemas generados en el análisis de datos de medidas repetidas, por la dependencia entre las observaciones, también es habitual el recurso a las técnicas de Análisis de la Covarianza (ANCOVA), de Análisis Multivariante de la Covarianza (MANCOVA) y Regresión Múltiple.

Para detectar *response shift* no sólo en contextos de CVRS como también en otras situaciones, la técnica más ampliamente utilizada es el *Then-Test* que pasamos a desarrollar en seguida.

#### 4.2.1. El *Then-Test*

Aunque la utilización del *Then-Test* en la detección de *response shift* haya tenido su origen en el contexto educativo (Howard *et al*, 1979b), este método ha sido considerado válido para evaluar cambios en el contexto de la CVRS. En el contexto de la Salud el *Then-Test* consiste en pedir a los pacientes que, después de la intervención o tratamiento, procedan a un juicio retrospectivo de su estado de salud inicial. Nótese que el *Post-Test* y el *Then-Test* son obtenidas bajo los mismos estándares de medida porque son hechos al mismo tiempo. Por tanto, la comparación de las puntuaciones obtenidas en el *Then-Test* y en el *Post-Test* eliminan los efectos del *response shift* inducidos por el tratamiento. La comparación de las puntuaciones medias obtenidas en el *Pre-Test* y el *Then-Test* constituye una estimación de la magnitud y dirección del *response shift* (Sprangers & Schwartz, 1999b; Schwartz *et al*, 2004; Barclay-Goddard *et al*, 2009). El *Then-Test* puede contemplar el instrumento de medida completo o un conjunto de ítems seleccionados en los cuales se espera que ocurra *response shift*.

En general, la detección de cambios Beta se hace recurriendo al *Then-Test*, como forma de evaluar **recalibración**. Para entenderlo es conveniente aclarar los siguientes conceptos:

- *Pre-Test* – evaluación antes de un determinado tratamiento;
- *Post-Test* - evaluación después de un determinado tratamiento;
- *Then-Test* – evaluación hecha después del tratamiento pero referente al estado del individuo antes del tratamiento, o sea, consiste en obtener medidas relativas al *Pre-Test*, en el momento del *Post-Test*.

El procedimiento más común es pedirle al paciente que evalúe su estado en el momento actual (*Post-Test* convencional) y, algún tiempo después, que haga una nueva evaluación sobre su estado en el momento del *Pre-Test* (*Pre-Test* retrospectivo). Si se ha producido algún tipo de recalibración en la escala de medida del encuestado entre el *Pre-Test* y el *Post-Test*, los valores en cuestión no serán comparables. Sin embargo, al evaluar las dos experiencias en el momento del *Post-Test*, la recalibración afectará de forma igual ambas evaluaciones. De esta forma se confirma que los estándares de medida son los mismos para las dos evaluaciones y se evita que el **cambio Beta** sesgue el verdadero **cambio Alpha**.

Según Howard *et al* (1979b), diferencias estadísticamente significativas entre los valores medios del *Pre-Test* y el *Then-Test*, indican la posible ocurrencia de un **cambio Beta**, mientras la diferencia entre los valores subyacentes al *Then-Test* y al *Post-Test*, se acepta como un verdadero **cambio Alpha**. Si no se detecta recalibración, puede estimarse el cambio Alpha comparando el *Pre-Test* con el *Post-Test*. Estas comparaciones se hacen sin más que recurrir a un test estadístico para la comparación de valores medios.

Son varias las referencias al *Then-Test* en la literatura para identificar *response shift* en pacientes de cáncer de mama (Andrykowski *et al*, 2009), cáncer de próstata (Rees *et al*, 2005; Korfage *et al*, 2007), pacientes de cáncer en fase de

tratamiento (Visser *et al*, 2000; Michielsen *et al*, 2007), pacientes trasplantados (Addington-Hall & Kalra, 2001), pacientes con fractura de rodilla (Balain *et al*, 2009; Razmjou *et al*, 2009), pacientes con esclerosis múltiple (Pinson *et al*, 2009) o con implantes dentarios (Ring *et al*, 2005). En todos estos estudios se ha asumido que, al administrar el *Post-Test* y el *Then-Test* en instantes próximos, las respectivas puntuaciones han reflejado la misma perspectiva y, consecuentemente la misma escala de calibración, o sea, los mismos estándares internos de medida. Así, se asume que la comparación de ambas las puntuaciones ha proporcionado una estimación del efecto del tratamiento corregido por el *response shift*.

Una variante del *Then-Test* es el *Ideal Scale Approach* (Schwartz & Sprangers, 1999) en el que se pide a los pacientes que completen el mismo cuestionario refiriéndose a su estado actual (por ejemplo, cómo perciben su CVRS actualmente) y refiriéndose a lo que consideran ser el estado ideal (por ejemplo, cómo les gustaría que fuera su CVRS). Estos cuestionarios serían administrados varias veces. Cambios en los estándares internos serían detectados a través de cambios significativos en las puntuaciones obtenidas a lo largo del tiempo con respecto al estado ideal. Aunque el objetivo principal de este método sea detectar cambios en los estándares internos, puede también revelarse útil para detectar conceptualizaciones y cambios en los valores. Por ejemplo, cambios en la jerarquía ideal de los dominios de CVRS, indicarían cambios en los valores y, cambios en la definición de CVRS ideal podrían indicar reconceptualización.

Aunque en general se acepte el *Then-Test* como una herramienta válida para evaluar cambios en los estándares internos y luego para detectar *response shift*, varios autores hacen algunas críticas a esta metodología. Visser *et al* (2000) y Hill & Betz (2005) subrayan que no debe ignorarse la posibilidad de que los datos del *Then-Test* aparezcan contaminados por problemas de memoria o dificultades que los pacientes presenten en recordar su situación anterior. En situaciones crónicas,

la memoria puede volverse poco precisa y luego la validez del *Then-Test* puede ser cuestionable. También Sprangers & Schwartz, (1999b) refieren que el empleo de este método ha sido críticamente examinado, destacando la posibilidad de que los resultados del *Then-Test* resulten sesgados. Se está asumiendo que los estándares internos son los mismos en el *Then-Test* y en el *Post-Test*, lo que de hecho debe ser empíricamente testado.

### **4.3. MÉTODOS MULTIVARIANTES PARA DETECTAR CAMBIO EN CVRS**

La utilización del Análisis Multivariante en general y de las Técnicas Factoriales en particular es prácticamente nula en el contexto de la CVRS a pesar del claro carácter multidimensional y dinámico que presentan estos datos. Las aportaciones en la literatura se limitan al Método de Ahmavaara (1954), utilizado por Golembiewski *et al* (1976) para detectar posible cambio Gamma en dos matrices de datos que se corresponden con dos momentos de evaluación de la Calidad de Vida y la propuesta de Vicente-Galindo (2003) que incluye los métodos STATIS y las Metacomponentes de Krzanowski (1979), métodos multivariantes ya conocidos pero nunca usados en este contexto.

Dedicamos los apartados siguientes a la presentación del Método de Ahmavaara (*op.cit.*) y a la propuesta de Vicente-Galindo (*op.cit.*).

#### **4.3.1. El Método de Ahmavaara**

Aunque relativamente poco conocido y utilizado, el procedimiento más anticuo del cual hay conocimiento, destinado a la comparación de estructuras factoriales, ha sido propuesto por Ahmavaara (1954) y fue utilizado por Golembiewski *et al* (1976) para comparar evaluaciones de CVRS. El método de Ahmavaara (*op.cit.*) permite comparar dos soluciones factoriales obtenidas sobre

un mismo conjunto de datos o bien comparar dos estructuras factoriales generadas por el mismo procedimiento. La idea subyacente a este procedimiento consiste básicamente en aplicar una transformación a una de las matrices factoriales a comparar, haciendo con que la estructura del primero grupo sea tan similar cuanto posible a la del segundo.

Supongamos que disponemos de dos matrices  $X_1$  y  $X_2$  relativas a dos evaluaciones de la calidad de vida de un grupo de pacientes en tratamiento que pueden ser el *Pre-Test* y el *Post-Test*. Pretendemos comparar estas evaluaciones con el fin de evaluar la ocurrencia de cambio. El primer paso en la aplicación del método, pasa por desarrollar una matriz de transformación  $L$ . Las columnas de esta matriz  $L$  están definidas por los factores comunes de  $X_2$  los cuales surgen en  $L$  en el mismo orden que en  $X_2$ . Por otra parte, las líneas de  $L$  están definidas por los factores comunes de  $X_1$ , manteniendo el orden. Los elementos de la matriz  $L$  así definida, representan los cosenos de los ángulos entre los factores de los dos estudios en cuestión o, más concretamente, entre los vectores asociados a esos factores. Consecuentemente, si algún elemento de  $L$  es exactamente igual a uno, significa que los respectivos factores de del *Pre-Test* y *Post-Test* son idénticos.

Del producto elemento a elemento entre las matrices  $X_1$  y  $L$ , resulta una matriz  $C$ , la cual representa el mejor ajuste mínimo cuadrático de  $X_1$  sobre  $X_2$ . Después se calculan los Coeficientes de Correlación Intraclase (ver apartado 2.2.1.) entre las matrices  $C$  y  $X_2$ , los cuales son de extrema importancia en el análisis subsiguiente. La Correlación Intraclase indica el grado según el cual las dos estructuras son similares tanto en patrón como en magnitud. El cuadrado de estos coeficientes indica el porcentaje de varianza que las dos estructuras en comparación tienen en común.

El Método de Ahmavaara puede sintetizarse gráficamente como se esquematiza en la Fig.2.

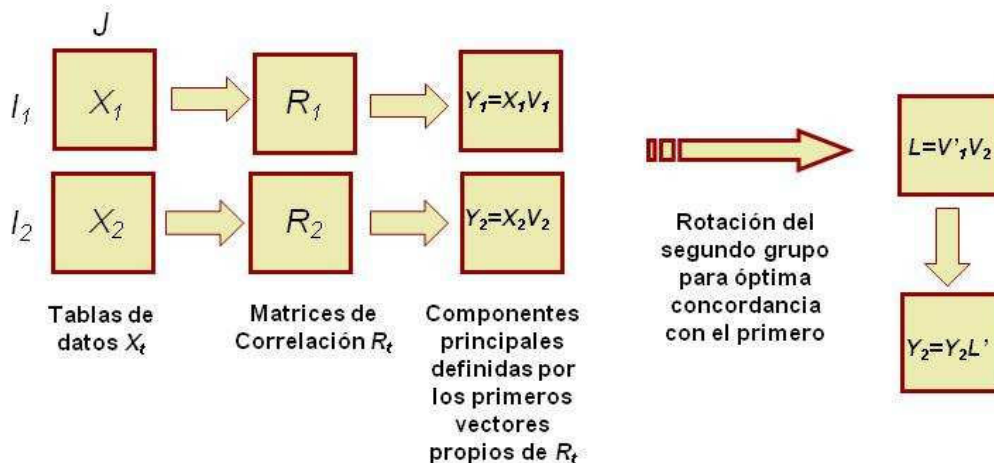


Fig.2. El Método de Ahmavaara

Aunque el Método de Ahmavaara se revele como un procedimiento útil para comparar dos soluciones factoriales, en un contexto de datos de CVRS, este procedimiento presenta algunas limitaciones, entre las cuales se destacan:

- posibilita únicamente la comparación de dos estructuras, lo que impide conseguir una visión general del problema cuando el objetivo es comparar estructuras en más de dos tiempos o situaciones;
- se supone que las matrices en estudio se refieren a los mismos individuos y variables (ítems), lo que en la práctica no siempre sucede;
- la aplicación del procedimiento implica el cálculo de un elevado número de correlaciones intraclase, lo cual produce incrementos significativos en el riesgo de errores Tipo I (declarar como significativas correlaciones que de hecho no lo son).

Posteriormente, Golembiewski *et al* (1976) proponen el **Método de Transformación**, concebido para detectar cambios en la estructura factorial de dos

evaluaciones ya que, alteraciones en la estructura factorial relativa a varios instantes en datos longitudinales pueden indicar *response shift*. Se propone una comparación factor a factor para identificar el aspecto del constructo que ha sido reconceptualizado. Sin embargo, no está claro cual es el nivel de disimilaridad que debe existir entre los factores (patrón y/o magnitud) para garantizar la ocurrencia de reconceptualización. En esta línea, Armenakis *et al* (1977) y Bedeian *et al* (1988) proponen el **Método de Congruencia** que tiene la ventaja de, además de examinar la similaridad de las estructuras factoriales, proporcionar un test de significancia estadística de los Coeficientes de Congruencia (Korth & Tucker, 1975).

#### 4.3.2. La Propuesta de Vicente-Galindo

Teniendo en cuenta la escasez de métodos propuestos en la literatura para analizar datos de CVRS y, en particular, para detectar cambio a lo largo del tiempo, Vicente-Galindo (2003) propuso un procedimiento que está basado en dos técnicas estadísticas ya existentes: el STATIS y la Comparación de Subespacios de Krzanowski. En este apartado presentamos una síntesis de estas dos técnicas así como la aportación de Vicente-Galindo (*op.cit.*).

##### 4.3.2.1. El Método STATIS

Los métodos STATIS (*Structuration des Tableaux A Trois Índices de la Statistique*) y STATIS Dual han sido propuestos por L'Hermier des Plantes (1976) y Escoufier (1976) y posteriormente desarrollados por Lavit (1988) y Lavit *et al* (1994), teniendo como principal objetivo la extracción de información relevante contenida en conjuntos de datos múltiples. Siendo, en general, los datos de CVRS datos de tres vías, en los cuales tenemos las respuestas de un conjunto de pacientes (individuos) a los ítems de un cuestionario (variables) en distintos tiempos, estas técnicas pueden revelarse útiles a la hora de detectar *response shift*.

Estos métodos posibilitan la exploración simultánea de  $T$  matrices  $X_t$  de dimensión  $I \times J$ ,  $t=1, \dots, T$ , de datos cuantitativos obtenidos en diferentes tiempos sobre los mismos individuos pudiendo eventualmente las variables ser diferentes (STATIS) o sobre las mismas variables pudiendo eventualmente los individuos ser diferentes (STATIS Dual). Nótese que, el método STATIS enfatiza las posiciones relativas de los individuos, mientras el STATIS Dual enfatiza las relaciones entre las variables.

STATIS y STATIS Dual se desarrollan en cuatro etapas esenciales:

- 1) Estudio de la **interestructura**, cuyo objetivo es proceder a una comparación global de la estructura de las  $T$  matrices de datos;
- 2) Búsqueda de una estructura común a las matrices en estudio - la **matriz compromiso**;
- 3) Estudio de la **intraestructura** que consiste en representar la estructura de cada matriz de datos en un espacio de baja dimensión;
- 4) Interpretación de las **trayectorias** individuales en el diagrama factorial.

Para detectar *response shift* consideramos la situación en que disponemos de las mismas variables (STATIS Dual) y únicamente necesitamos del primer paso que es el **Análisis de la Interestructura**. Una descripción detallada de los métodos STATIS y STATIS Dual se encuentra en anexo (Anexo III).

Supongamos que disponemos de  $T$  matrices  $X_t$  referentes a distintos momentos de evaluación de un conjunto de variables de calidad de vida en un grupo de pacientes en tratamiento. En el caso del STATIS Dual, la estructura de cada matriz  $X_t$  es captada a través de la configuración



$$C_t = X_t' D_t X_t$$

que es la matriz de varianzas-covarianzas entre variables, siendo  $D_t$  la métrica del espacio de las variables.

A fin de hacer posible la comparación de las  $T$  matrices de datos y averiguar la existencia de una eventual estructura común a las variables, se define el producto interno de Hilbert-Schmidt entre configuraciones, el cual se define como

$$\langle C_t | C_{t'} \rangle = \text{Tr}(C_t M C_{t'}' M)$$

donde  $M$  define la métrica en el espacio de los individuos.

En el análisis de la interestructura se recurre habitualmente al Coeficiente de Correlación Vectorial entre dos estudios  $t$  e  $t'$  propuesto por Escoufier (1973) e definido de la siguiente forma:

$$RV(t, t') = \left\langle \frac{C_t}{\|C_t\|_{HS}} \middle| \frac{C_{t'}}{\|C_{t'}\|_{HS}} \right\rangle_{HS} = \frac{\text{Tr}(C_t M C_{t'}' M)}{\sqrt{\text{Tr}(C_t M)^2} \sqrt{\text{Tr}(C_{t'} M)^2}}$$

Los coeficientes  $RV$  surgen organizados en una matriz  $RV$ , cuadrada de orden  $T$  y posibilitan una interpretación fácil de la interestructura en la medida en que son no negativos y varían entre 0 e 1. Si  $RV(t, t')=1$ , entonces la distancia entre los estudios  $t$  y  $t'$  es nula y las estructuras en cuestión son coincidentes. La distancia entre dos estudios normados  $t$  y  $t'$ , es dada por:

$$d_{HS} \left( \frac{C_t}{\|C_t\|_{HS}} \middle| \frac{C_{t'}}{\|C_{t'}\|_{HS}} \right) = \sqrt{2(1 - RV(t, t'))}$$

Con el objetivo de comparar globalmente la estructura de las  $T$  matrices de datos, es posible obtener una representación euclídea de las mismas sometiendo la matriz  $RV$  a un Análisis en Componentes Principales (ACP). En el espacio de las Componentes Principales, cada matriz viene representada por un punto y la distancia entre puntos en el plano refleja el grado de semejanza entre matrices, posibilitando el análisis de la interestructura. El procedimiento subyacente al análisis de la interestructura se sintetiza en la Fig. 3.

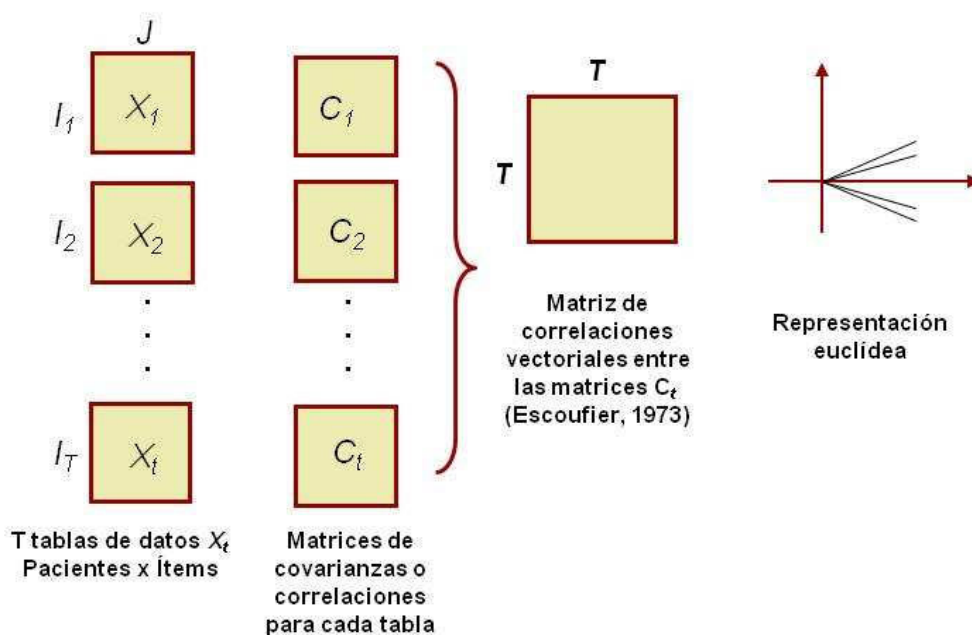


Fig. 3. Estudio de la interestructura.

Los coeficientes  $RV$  se interpretan como los cosenos de los ángulos entre los vectores que representan las matrices y aproximan la correlación vectorial entre las mismas, por lo que, si hay una estructura común, los ángulos son pequeños y la mayor parte de la variabilidad es explicada por el primer eje de la representación. Por eso, uniendo los puntos de la imagen euclídea con el origen del sistema de ejes, se obtiene una estimación gráfica de la correlación entre las estructuras mediante el ángulo que forman los vectores obtenidos. En esta imagen euclídea se pueden incluir matrices como elementos suplementarios, las cuales no

intervienen en la determinación de los ejes pero se desea observar su posicionamiento en la interestructura.

Supongamos que a un conjunto de pacientes en estudio se les ha aplicado un cuestionario de CVRS que mide Funcionamiento Físico, Salud Mental y Funcionamiento Social. Este cuestionario ha sido aplicado en el inicio del tratamiento (*Pre-Test* –  $T_1$ ) y en el final del tratamiento (*Post-Test* –  $T_2$ ). Además, después del tratamiento se ha aplicado un *Pre-Test* retrospectivo (*Then-Test* –  $T_3$ ). Imaginemos que, al someter las matrices de datos  $T_1$ ,  $T_2$  y  $T_3$  al estudio de la interestructura se había obtenido una de las representaciones euclídeas presentadas en la Fig. 4.

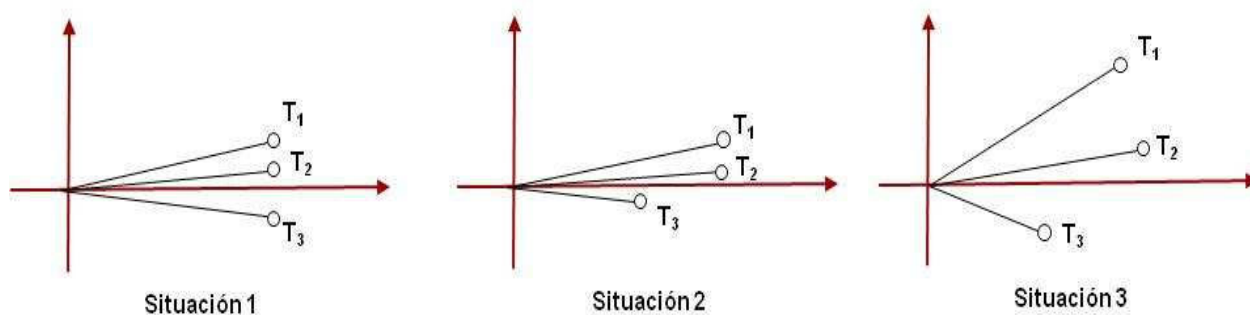


Fig. 4. Tres situaciones de posibles representaciones de la interestructura.

En la **situación 1**, los tres estudios en análisis están caracterizados por matrices  $C_t$  de normas similares y formando ángulos pequeños. En este caso, el Coeficiente de Correlación Vectorial presentará valores elevados y se concluye que existe una estructura de covariación estable y un elevado grado de congruencia entre las matrices  $T_1$ ,  $T_2$  y  $T_3$  en estudio. En este caso, se mantendrían las mismas dimensiones y su orden de importancia, o sea, si por ejemplo en el *Pre-Test* la dimensión Funcionamiento Físico fuese la que recoge más variabilidad, seguida de las dimensiones Salud Mental y Funcionamiento Social, eso se mantendría en las aplicaciones siguientes.

En la **situación 2** se obtendrían valores del Coeficiente de Correlación Vectorial igualmente elevados ya que los ángulos definidos entre los diferentes

estudios son pequeños indicando una estructura común de covariación. Sin embargo, las normas de las matrices en estudio no presentan igual magnitud, lo que se podría interpretar como un cambio de percepción de escala y llevaría a considerar la hipótesis de recalibración.

En la **situación 3** no solamente los ángulos formados por las tres matrices en estudio son grandes, sino también presentan normas distintas, lo que conllevaría a considerar la inexistencia de una estructura común a los diferentes estudios y luego a plantear la posibilidad de ocurrencia de cambio Gamma.

Nótese que si la primera dimensión de las distintas matrices de datos en estudio es coincidente, el STATIS declara que estas estructuras son similares. Si las matrices en estudio difieren en una o más dimensiones que no es la primera, el STATIS no va a detectarlo y declara que las estructuras son similares.

Si ocurre que el STATIS declara que las matrices en estudio no son similares, eso podrá indicar la ocurrencia de Cambio. Sin embargo, si declara que las estructuras son similares, podrá haber ocurrido cambio o no.

Si hubiéramos llegado a la conclusión que las matrices de datos en análisis poseen una estructura similar (como en la situación 1), se podría calcular una matriz compromiso, o sea, una estructura común que representara las matrices en estudio.

#### **4.3.2.2. El Procedimiento de Krzanowski**

La comparación de estructuras factoriales a través de los métodos clásicos se basa en la comparación de componente a componente. Sin embargo, puede ocurrir que dos estructuras sean similares salvo una transformación y que, al inspeccionar pares de componentes, los índices de congruencia sean bajos.

Tratándose de estructuras factoriales, es por tanto muchas veces necesario rotar las soluciones hasta conseguir resultados más interpretables, sin que eso tenga algún significado especial. Teniendo como base esta idea, Krzanowski (1979) propone la comparación de soluciones multidimensionales completas en vez de hacerse la comparación eje a eje, sugiriendo un procedimiento para la comparación de subespacios generados por componentes principales, fácilmente generalizable a cualquier técnica factorial y que se desarrolla en los siguientes pasos:

- 1) Se calculan las componentes principales de cada espacio;
- 2) Se compara cada componente de cada espacio con todas las componentes correspondientes a los espacios asociados a las otras matrices;
- 3) Se localiza la componente más similar;
- 4) Se calcula el coseno del ángulo formado por el par de componentes más similares;
- 5) Se repite el procedimiento para todas las componentes.

Supongamos que disponemos de dos matrices  $X_1$  e  $X_2$  que recogen información de  $p$  variables de CVRS sobre  $n$  pacientes en 2 tiempos. A estas matrices se les ha aplicado el mismo tipo de transformación: sean  $Y_1=X_1V_1$  y  $Y_2=X_2V_2$  las respectivas soluciones factoriales (obtenidas, por ejemplo, a través de un ACP) y  $V_1$  y  $V_2$  los vectores propios de las correspondientes matrices de correlaciones que definen el subespacio de las Componentes Principales.

La comparación de las soluciones factoriales en cuestión se basa en la búsqueda del ángulo mínimo entre los subespacios (en realidad se están comparando dos conjuntos de direcciones en el espacio). Si ese ángulo mínimo es nulo es porque existe algún par de dimensiones latentes coincidentes y si es muy

pequeño (aunque no nulo), significa que existe al menos un par de dimensiones muy similares. Por otra parte, cuanto mayor es el ángulo encontrado, menos parecidas serán las dimensiones.

Krzanowski (1979) demuestra que el ángulo mínimo entre un vector del espacio de las primeras Componentes Principales del primer conjunto y uno casi paralelo en el espacio de las Componentes Principales del segundo conjunto está dado por

$$\cos^{-1} \sqrt{\lambda_1}$$

donde  $\lambda_1$  es el valor propio más grande de la matriz  $S = V_1 V_2' V_2 V_1'$ .

Si  $\lambda_i$  es el  $i$ -ésimo valor propio mayor de  $S$  y  $a_i$  su vector propio asociado, y  $b_i = V_1' a_i$  ( $i=1, \dots, K$ ), entonces el ángulo entre el  $i$ -ésimo par,  $b_i$  y  $V_2' V_2 b_i$  está dado por  $\cos^{-1} \sqrt{\lambda_i}$ .

Una medida de similaridad entre los subespacios en análisis puede definirse en base a la suma de los valores propios ya que esta varía entre 0 (si no hay ningún factor común, o sea, tenemos espacios disjuntos y ortogonales) y  $K$  (si los espacios coinciden). Por lo tanto, un índice de similaridad entre dos subespacios puede definirse como

$$\text{Similaridad} = \frac{\sum_i \lambda_i}{K}$$

El procedimiento de Krzanowski puede sintetizarse en el esquema presentado en la Fig.5.

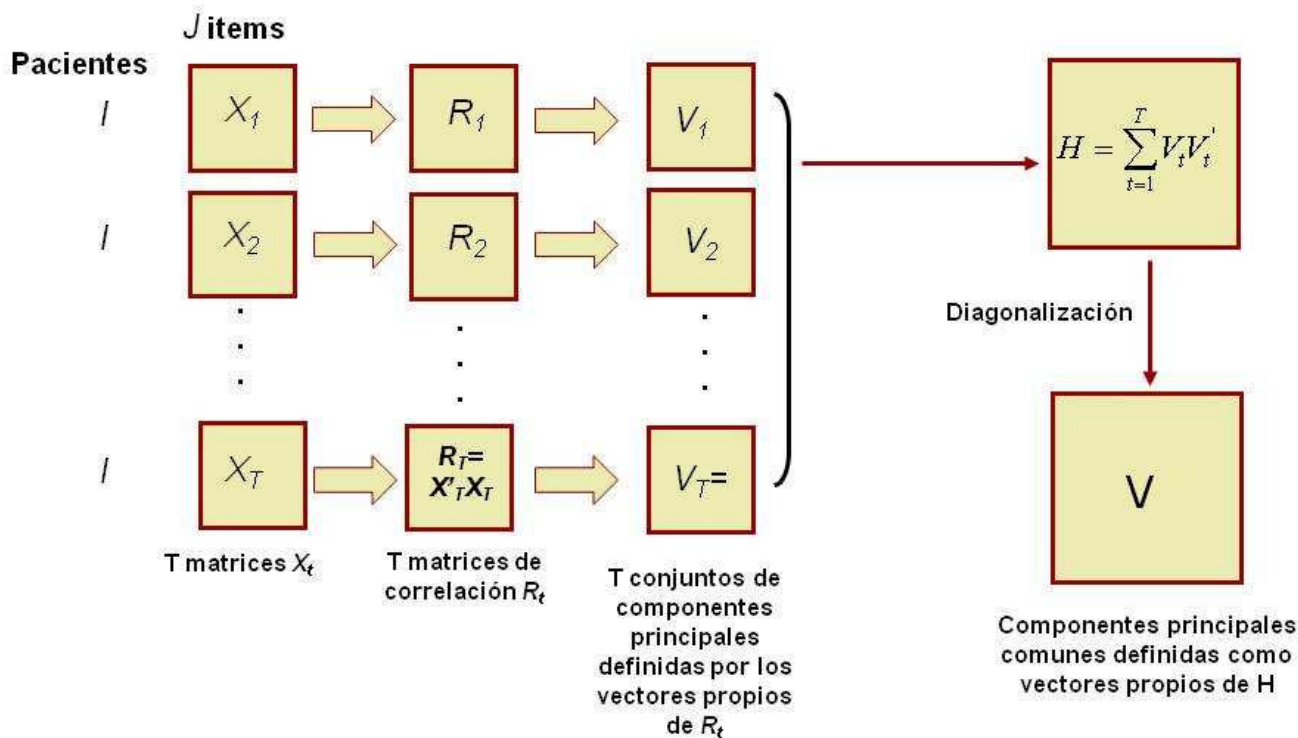


Fig.5. El Procedimiento de Krzanowski.

El procedimiento es fácilmente generalizable a más de dos matrices, constituyendo este aspecto una ventaja importante relativamente al Método de Ahmavaara, además de, en este caso, se inspeccionan soluciones factoriales multidimensionales completas y no pares de componentes.

El Procedimiento de Krzanowski compara cada una de las dimensiones latentes con las demás y, si detecta alguna dimensión similar, aunque no esté en el mismo orden, las declara similares, lo que constituye una ventaja relativamente al STATIS que únicamente trabaja con la primera dimensión. Por otra parte, al declarar similares dimensiones que han cambiado de orden (significando que los pacientes han pasado a valorar más una determinada dimensión de CVRS que antes), el Procedimiento de Krzanowski no permitiría detectar repriorización.

#### 4.3.2.3. STATIS Dual y Metacomponentes en la Evaluación del Cambio

Basándose en el STATIS Dual y la Comparación de Subespacios de Krzanowski, Vicente-Galindo (2003, 2007) propone un conjunto de procedimientos que pueden revelarse eficaces en la evaluación de cambios.

#### DETECCIÓN DE CAMBIOS GAMMA

La existencia de estructuras factoriales congruentes indica que no se ha producido ningún cambio Gamma (**reconceptualización**). Se considera que dos estructuras factoriales son congruentes cuando poseen los mismos factores latentes y los factores están configurados por los mismos patrones de carga (sin embargo, no es necesario que la magnitud de los factores de carga se mantenga).

Se puede probar que la ocurrencia de cambios Alpha y Beta puede afectar las correlaciones entre ítems y por tanto la estructura factorial interna. La ausencia de congruencia de las estructuras factoriales es condición necesaria pero no suficiente para detectar la presencia de un cambio Gamma, aunque una incongruencia en las estructuras factoriales es altamente sugerente de la ocurrencia de un cambio Gamma (Bartholomew, 2002).

El STATIS Dual y Procedimiento de Krzanowski pueden revelarse útiles a la hora de detectar cambios Gamma.

#### *El STATIS Dual en la detección de Cambio Gamma*

Dadas  $T$  matrices  $X_t$  de dimensión  $I \times J$ ,  $t=1, \dots, T$ , relativas a  $J$  variables de Calidad de Vida obtenidas sobre  $I$  pacientes en diferentes evaluaciones, se empiezan por calcular las configuraciones  $C_t$ , que son las matrices de varianza-covarianza relativas a los diferentes tiempos, las cuales se comparan recurriendo al coeficiente  $RV$ . En términos de *response shift*, interesa destacar que, si se obtiene



un coeficiente  $RV$  próximo a 1, significa que ambas estructuras son coincidentes y por lo tanto no se ha producido ninguna **reconceptualización**, o sea, no se ha producido un **cambio Gamma**. Cuanto más próximo a 1 sea  $RV$ , más similares son las estructuras en cuestión.

Además de eso, el STATIS Dual permite representar las matrices correspondientes a los diferentes tiempos en estudio como puntos en un espacio vectorial de baja dimensión. Como hemos dicho, en la imagen euclídea, la distancia entre puntos, se interpreta en términos de similitud entre las matrices representadas por tales puntos y, en este caso, en términos de similitud entre las estructuras de varianza-covarianza y, consecuentemente, en términos de congruencia entre las estructuras factoriales. Por lo tanto, si los ángulos formados por los vectores de la imagen euclídea son próximos a cero, las estructuras en cuestión son semejantes y por lo tanto no hay razones para suponer que se ha producido una **reconceptualización**. Consecuentemente, si se ha producido un **cambio Gamma**, la representación euclídea de las matrices debe exhibir algún ángulo grande o algún punto debe estar alejado del resto, aunque tal no sea condición suficiente.

En la imagen euclídea, si además de que los vectores forman entre sí ángulos pequeños, sus normas son iguales, se descarta la posibilidad de haberse producido una **recalibración**. En este caso se busca una representación compromiso, la cual proporciona una visión global de las dimensiones latentes del constructo a lo largo del tiempo. Por otra parte, si hay razones para suponer la ocurrencia de una **reconceptualización (cambio Gamma)** o de una **recalibración (cambio Beta)**, no es posible evaluar la ocurrencia de un eventual **cambio Alpha**. Proyectando los datos originales en la representación compromiso, es posible analizar el comportamiento de los ítems a lo largo del tiempo y esto puede ayudar a detectar los ítems implicados en la **reconceptualización** y en la **recalibración**.

Finalmente, la observación de las trayectorias puede proporcionar también alguna información útil. Una trayectoria con poca variación indica que el ítem es estable en el tiempo, mientras variaciones acentuadas en las trayectorias pueden estar asociadas a un **cambio Beta** o a un **cambio Gamma** si eso ha sido evidenciado en la interestructura. Trayectorias envolventes corresponden a ítems que siguen una evolución media a lo largo del tiempo, mientras trayectorias excéntricas pueden indicar cambios al largo del tiempo en el valor del ítem relativamente al valor medio.

#### *El Procedimiento de Krzanowski en la detección de Cambio Gamma*

En términos de *response shift*, puede afirmarse, siguiendo a Vicente-Galindo (2003) y Galindo & Vicente-Galindo (2004) que, si se concluye que las soluciones son congruentes, entonces las dimensiones y patrones del constructo son los mismos y eso será condición suficiente para poder afirmar que no se ha producido una **reconceptualización**. Por otra parte, si se concluye que las soluciones multidimensionales son similares, es posible encontrar una solución consenso, mientras que si las soluciones no son globalmente congruentes es posible conocer los ítems en los cuales se ha producido el cambio.

#### **DETECCIÓN DE CAMBIOS BETA**

Seguendo a Vicente-Galindo (*op.cit.*) y Galindo & Vicente-Galindo (*op.cit.*), una forma de detectar la eventual presencia de cambios Beta, es en base a la comparación de las respectivas estructuras factoriales. Una vez comprobado que las estructuras factoriales presentan el mismo patrón, es decir, el mismo número de ejes y los mismos ítems con altos factores de carga en cada eje, se puede detectar la presencia de un cambio Beta observando la magnitud de los factores de carga y/o las varianzas y covarianzas de las variables latentes y de las variables observables (Taris, 2000).

## DETECCIÓN DE CAMBIOS ALPHA

Descartada la ocurrencia de un cambio Beta, la comparación entre los resultados del *Pre-Test* y *Post-Test*, permiten identificar un cambio Alpha si hay diferencias estadísticamente significativas.

### En resumen:

- Antes que nada, hay que comparar las estructuras factoriales en cuestión intentando descartar la presencia de **cambio Gamma**;
- La comparación de las estructuras factoriales en cuestión puede ser una herramienta útil en la identificación de un **cambio Beta**. Si se concluye que las estructuras en estudio son congruentes en el sentido de que presentan el mismo número de ejes y los mismos ítems asociados a cada eje, se puede evaluar la ocurrencia de un eventual cambio Beta analizando la magnitud de los factores de carga así como las varianzas y covarianzas de las variables latentes y observables;
- La observación de la imagen euclídea de la interestructura que el STATIS Dual proporciona, permite también la detección de **cambio Beta**: una norma más pequeña o mayor que las restantes sugiere que ha ocurrido una recalibración;
- Sólo después de descartar la ocurrencia de los cambios Gamma y Beta, se evalúa la posible ocurrencia de **cambio Alpha**.

#### 4.4. EL COEFICIENTE $RV$ PARA DETECTAR CAMBIO

Como hemos visto, el coeficiente  $RV$  se utiliza en el método STATIS con el objetivo de cuantificar el grado de semejanza entre dos estructuras factoriales. Sin embargo, además de eso, este coeficiente puede constituir una herramienta útil en la detección del Cambio (Baccalá, 2004). En este apartado, veremos con más detalle como el coeficiente  $RV$  puede revelarse útil a la hora de detectar cambio.

Recuérdese que el coeficientes  $RV$  entre dos estructuras  $t$  y  $t'$  es dado por

$$RV(t,t') = \left\langle \frac{C_t N}{\sum_s (\lambda_s^t)^2}, \frac{C_{t'} N}{\sum_s (\lambda_s^{t'})^2} \right\rangle$$

siendo  $C_t$  la configuración obtenida a partir del producto escalar entre individuos,  $N$  la matriz de pesos de los individuos y  $\lambda_i$  los valores propios asociados a las estructuras en estudio.

El coeficiente  $RV$  está acotado entre 0 y 1.  $RV(t,t')=0$  significa que cada variable de la tabla  $t$  tiene covarianza nula variable de la tabla  $t'$ . Por otra parte,  $RV(t,t')=1$  significa que las nubes  $t$  y  $t'$  son homotéticas.

Valores de  $RV$  próximos de 1 indican estructuras similares en términos del comportamiento de los individuos (STATIS) o de estructura de covariación entre las variables (STATIS Dual), indicando que no ha ocurrido cambio en las estructuras en estudio.

Para comprobar la existencia de invarianza estructural entre dos grupos de variables  $t$  y  $t'$ , hay que averiguar si:

- tienen el mismo número de dimensiones (con absorción de porcentajes de inercia similares);

- los patrones de carga son similares (las dimensiones están formadas por las mismas variables contributivas).

Teniendo en cuenta lo que hemos dicho sobre la naturaleza del coeficiente  $RV$ , se puede afirmar que:

- Si dos grupos de variables  $t$  y  $t'$  presentan invarianza estructural,  $RV$  es igual a 1 sin depender del número de dimensiones (toma el valor máximo independientemente del número de dimensiones comunes);
- Si dos grupos de variables  $t$  y  $t'$  no presentan invarianza estructural (ha ocurrido un cambio estructural), los valores de  $RV$  dependen de la importancia de la dimensión donde se presenta el cambio:
  - o si el cambio ocurre en las principales direcciones de inercia,  $RV$  decrece;
  - o si el cambio se produce en las direcciones de inercia menos importantes,  $RV$  decrece a medida que crece el número y la importancia de las dimensiones no comunes, respecto a la primera dimensión de inercia.

La ocurrencia de un cambio en un constructo dado podrá implicar:

- cambio en el número de dimensiones (puede aparecer, por ejemplo, una nueva dimensión en uno de los grupos);
- que el número de dimensiones se mantenga pero las variables contributivas cambian de un grupo a otro.

En la presencia de un **cambio Gamma** (presupone cambio en las principales direcciones de inercia de los grupos en estudio),  $RV$  decrece a medida que aumentan las dimensiones comunes con respecto a la dirección de inercia principal en número e importancia.

La ocurrencia de un **cambio Beta** implica que alguna de las variables contributivas cambia su importancia de un grupo a otro, pudiendo así cambiar los valores de los factores de carga y/o la variabilidad de las dimensiones latentes. Si la estructura en cuestión es unidimensional y las variables modifican su importancia, *RV* decrece porque no considera esa dimensión como común.

Al ocurrir una **recalibración** en una dirección de menor absorción de inercia, *RV* no capta ese cambio dado que el hecho de que la primera dimensión sea coincidente es suficiente para mantener *RV* elevado.

Una **reprorización** implica un cambio en la importancia de las dimensiones de los grupos, aunque sean las mismas (por ejemplo, la primera componente pasa a ser la tercera). Si se produce una **reprorización**, *RV* decrece si se trata de la primera dimensión pero, en caso contrario, no la captará.

El Test de Permutaciones también puede revelarse una herramienta útil a la hora de evaluar una **reprorización** o **inconsistencia intraindividual** (Baccalá, 2004).

**En resumen:**

- *RV* no es útil en la identificación del tipo de cambio que se produce;
- Cuanto más próximo *RV* esté de 1, más evidencia existe de que no se han producido cambios en la dirección de máxima inercia;
- La utilización de *RV* resulta insuficiente a la hora de valorar la ocurrencia de cambios por lo que se debe proceder a un análisis más detallado y no solamente de los coeficientes de correlación entre matrices.

#### 4.5. EL MODELO DE ZWINDERMAN

Aunque los métodos abordados anteriormente puedan constituir una herramienta útil en la evaluación del cambio, ningún de ellos posibilita una cuantificación de la magnitud del cambio observado. En este sentido, Zwinderman (1990) propone un modelo para evaluar cambio en CVRS a lo largo del tiempo, en contextos de ensayos clínicos, para pacientes bajo la influencia de uno o más tratamientos. La metodología propuesta por este autor, posibilita la cuantificación del cambio observado. En concreto, se pretende evaluar el cambio debido al proceso temporal y el cambio debido a los tratamientos clínicos en base al modelo de regresión logística. En este modelo la CVRS es vista como una variable latente multidimensional y se mide a través de respuestas dicotómicas al mismo conjunto de ítems en diferentes instantes.

Este modelo se basa en los trabajos de Rash (1961)<sup>1</sup> y trata de explicar el Cambio en CVRS a partir del modelo de Regresión Logística Latente que puede incluir parámetros relacionados con el proceso temporal y con los efectos de los tratamientos clínicos, teniendo en cuenta la interacción entre parámetros.

Zwinderman considera un modelo de medida en el que  $X_v$  es la suma de las puntuaciones relativas al paciente  $v$  sobre todos los ítems relativos a una determinada variable latente (por ejemplo, la movilidad). Además, sea

$$X_v = \tau_v + \varepsilon_v$$

---

<sup>1</sup> El modelo de Rash es uno de los más usados en la Teoría de la Respuesta al Ítem (TRI). Este modelo presupone que la respuesta a un ítem depende de la competencia del sujeto ( $\theta$ ) y de la dificultad del ítem ( $b$ ) y que la curva característica de los ítems viene dada por la función logística:

$$P_i(\theta) = \frac{\exp[D(\theta - b_i)]}{1 + \exp[D(\theta - b_i)]}$$

donde  $P(\theta)$  es la probabilidad de acertar el ítem  $i$  al nivel  $\theta$ ,  $\theta$  los valores de la variable medida,  $b_i$  el índice de dificultad del ítem  $i$  y  $D$  una constante adicional cuyo objetivo es aproximar la curva logística a la curva normal acumulada (en general se hace  $D=1,7$ ). Asimismo, conocido el índice de dificultad de un ítem ( $b$ ) y la competencia de los sujetos ( $\theta$ ), es posible predecir la probabilidad ( $P(\theta)$ ) de acertar el ítem.

donde  $\tau_v$  es el valor verdadero de una determinada variable latente (por ejemplo, la verdadera movilidad) del paciente  $v$  y  $\varepsilon_v$  una componente de error.

Consideremos ahora las respuestas de  $N$  pacientes a  $k$  ítems en dos momentos en el tiempo. Sea  $X_{vi1}$  la respuesta del paciente  $v$  al ítem  $i$  en el instante  $T_1$  y sea  $X_{vi2}$  la respuesta del paciente  $v$  al ítem  $i$  en el instante  $T_2$ .

Se asume una variable latente  $\theta_i$  para cada ítem y  $\theta_{vi}$  representa la posición del paciente  $v$  en el *continuum* latente  $\theta_i$  en el instante  $T_1$ . Debido a que se trata de un proceso temporal que envuelve uno o más tratamientos,  $\theta_{vi}$  cambia según una cantidad  $\delta_v$ . Luego, la posición del paciente  $v$  en el *continuum* latente  $\theta_i$  en el instante  $T_2$  es dada por  $\theta_{vi2} = \theta_{vi1} + \delta_v$ .

El cambio  $\delta_v$  es modelizado como una función del tiempo y del tratamiento que el paciente  $v$  recibe:

$$\delta_v = q_v \mu + \tau$$

donde  $\tau$  es el cambio inherente al proceso temporal y  $\mu$  el cambio debido al tratamiento. El peso  $q_v$  toma el valor 1 si  $v$  recibe el tratamiento y 0 en caso contrario. Los efectos de otros tratamientos pueden también influir en posibles cambios. Todavía, tales efectos, no pueden ser separados de los efectos del proceso temporal.

Los estimadores de máxima verosimilitud condicional  $\hat{\mu}$  y  $\hat{\tau}$  se obtienen igualando a cero las derivadas parciales de la verosimilitud condicional de  $X_1$  y  $X_2$  dado  $X_1 + X_2$  con respecto a  $\mu$  y  $\tau$ :

$$\frac{\partial L(\mu, \tau)}{\partial \mu} = 0 \quad \text{y} \quad \frac{\partial L(\mu, \tau)}{\partial \tau} = 0.$$

Solamente se utilizan las respuestas para las cuales  $x_{vi1} \neq x_{vi2}$ .

Una descripción detallada del Modelo de Zwinderman se encuentra en anexo (Anexo IV).



Según Zwinderman, asumiendo la ausencia de interacción paciente-tiempo dentro de los grupos de tratamiento, los parámetros subyacentes al proceso temporal y a los efectos de los tratamientos, pueden ser estimados independientemente de los parámetros subyacentes a la CVRS latente en el principio del tratamiento, o sea, la verosimilitud condicional utilizada en la estimación es independiente de los valores individuales de  $\theta_i$ . Consecuentemente, mecanismos causadores de **datos faltantes** que dependen del valor de  $\theta_i$  pueden ignorarse cuando se estima  $\mu$  y  $\tau$  maximizando la verosimilitud condicional. Por lo tanto, solamente es necesario asumir que los abandonos, muertes o mecanismos de censura no dependen de las respuestas a los ítems y así es posible estimar  $\mu$  y  $\tau$  en base a la verosimilitud condicional, aunque algunos pacientes se mueran o abandonen el estudio. Los mecanismos causadores de datos faltantes se ignoran porque se parte del principio de que no existe interacción paciente-tiempo dentro de los grupos de tratamiento. Esto significa que pacientes que reciben el mismo tratamiento presentan el mismo patrón de cambio a lo largo del tiempo y que el patrón de cambio no depende de si la CVRS es buena o mala al principio del tratamiento. De hecho, esta suposición es equivalente a la suposición *hazard* proporcional del modelo de regresión de Cox para análisis de supervivencia. La razón *hazard* de nuestro modelo es proporcional, dentro de los estratos definidos por cada combinación ítem-paciente, luego

$$\text{logit}[p(X_{it})] - \text{logit}[p(X_{ij})] = \delta_{it} - \delta_{ij}.$$

La suposición *hazard* proporcional puede ser a veces restrictiva. Puede pasar que existan subgrupos de pacientes que reaccionan de forma diferente a un tratamiento dado. Por ejemplo, puede pasar que el efecto positivo de un tratamiento en la movilidad dependa de la edad del paciente. Esta hipótesis se puede testar incorporando la edad en el modelo. De forma similar, se puede testar

si los pacientes que han abandonado el estudio después del principio del tratamiento, presentan un patrón de cambio diferente de los pacientes que no han abandonado el estudio.

En este modelo, la mortalidad es vista únicamente como un mecanismo causador de datos faltantes. En la evaluación de tratamientos nuevos, a la muerte siempre se asocia la puntuación más baja de CVRS, la cual es conocida y constante. Partiendo del principio que existe una relación entre CVRS y mortalidad, un análisis basado en el Modelo de Zwinderman (1990) conducirá a resultados sesgados ya que no es posible comparar los efectos en la CVRS de dos tratamientos con diferentes tasas de mortalidad y abordar las muertes simplemente como datos faltantes. En este contexto, Zwinderman (*op.cit.*) sugiere combinar el análisis de las medidas de CVRS con el análisis de supervivencia.

El modelo propuesto por Zwinderman constituye una contribución significativa en la evaluación del cambio en contextos de CVRS. Se destaca el hecho de que permite la cuantificación del cambio ocurrido separando dos componentes importantes: el **cambio debido al tratamiento** y el **cambio debido al tiempo**. Esto constituye una importante **ventaja** sobre las demás técnicas a las que habitualmente se recurre en la detección y tratamiento del cambio y, desde nuestro punto de vista, este es el punto fuerte de este modelo.

Sin embargo, el modelo de Zwinderman, presenta algunas **limitaciones** que tienen que ver con los siguientes aspectos:

- Los datos en estudio tienen que ser datos apareados y dicotómicos;
- Solamente se utilizan los pares de respuestas que son diferentes en los instantes  $T_1$  y  $T_2$  en estudio, o sea, en que  $x_{u1} \neq x_{u2}$  y se ignoran por completo las situaciones en las cuales los pacientes no tenían síntomas o en que los tenían en ambas las evaluaciones;
- La evaluación completa del paciente no es tomada en cuenta sino únicamente los pares individuales de datos de una evaluación a otra;

- No es tenido en cuenta el nivel de CVRS en el inicio del tratamiento;
- Se asume que el cambio es igual para todas las variables latentes;
- Los efectos de otros tratamientos no pueden ser separados de los efectos debidos al tiempo;
- El hecho de asumir que la suma de las puntuaciones a lo largo del tiempo es un estadístico suficiente mínimo para los valores individuales iniciales  $\theta_{vi1}$ , exige la inexistencia de interacción paciente-tiempo dentro de los grupos de tratamiento y esto implica asumir que, dentro de un determinado grupo de tratamiento, el cambio a lo largo del tiempo es el mismo para todos los pacientes.
- Si el número de instantes en estudio es elevado, como ocurre con frecuencia en ensayos clínicos, el número de parámetros aumenta y el procedimiento de estimación en modelos latentes logísticos como el que propone Zwinderman, se vuelve bastante complejo, especialmente si existen datos faltantes.
- Es bien sabido que es prácticamente inevitable la existencia de **datos faltantes** en estudios longitudinales de CVRS. Zwinderman aborda el problema de los datos faltantes en contextos de CVRS asumiendo que los datos faltantes son aleatorios y, consecuentemente, el patrón de cambio medio puede ser estimado a partir de los datos observados. Este autor argumenta que los datos faltantes dependen únicamente de la historia observada del paciente y no de medidas futuras no observadas. Luego, los datos faltantes son ignorables y, comparaciones de CVRS entre distintos grupos de tratamiento, son relevantes únicamente para la historia pasada y evaluables en un dado instante de la comparación.

Teniendo en cuenta que, en el modelo propuesto por Zwinderman no es tenida en cuenta la eventual ocurrencia de fenómenos de alteración en la autopercepción de los sujetos, hemos tratado de presentar una posible solución de

forma a complementar la información obtenida mediante la aplicación de este modelo. Para ello, hemos desarrollado un pequeño estudio de simulación que presentamos en el apartado siguiente.

#### **4.5.1. Una Contribución al Modelo de Zwinderman utilizando Técnicas Factoriales**

Cómo hemos visto, el Modelo de Zwinderman (1990) tiene la ventaja de proporcionar estimaciones del cambio ocurrido en un conjunto de pacientes, aislando la componente debida al tratamiento y la componente asociada al paso del tiempo. Sin embargo, este autor no ha planteado la eventual ocurrencia de cambios en la autopercepción de los pacientes que ocurren a menudo en situaciones de evaluación de CVRS entre una situación de *Pre-Test* y *Post-Test*. Dada la importancia del fenómeno de *response shift* en el contexto oncológico, proponemos que las estimaciones del cambio proporcionadas por el Modelo de Zwinderman sean acompañadas de los resultados proporcionados por un método factorial de forma a obtener una visión de las relaciones entre las variables estudio. Enseguida veremos cómo las técnicas Biplot pueden revelarse útiles a la hora de expresar estas relaciones.

Los datos utilizados han sido reconstruidos a partir de los datos presentados por Zwinderman (*op.cit.*) los cuales han sido obtenidos sobre un grupo de 88 pacientes diagnosticadas de cáncer de mama. 44 de estas pacientes reciben un tratamiento complementario de APD (aminohydroxy propylidene-biphosphonate) el que se supone mejorar la calidad de vida de las pacientes, mientras las restantes 44 pacientes no reciben tratamiento. Los datos aquí considerados se refieren a las valoraciones de las 88 pacientes, en una escala de 0 (sin síntomas), 1 (síntomas moderados), 2 (síntomas severos) y 3 (síntomas muy

severos), sobre 6 ítems relativos a la movilidad, con los cuales se pretendía valorar si las pacientes tenían dificultades en los siguientes aspectos:

- caminar en casa (V1)
- caminar por la calle (V2)
- subir escaleras (V3)
- levantarse, inclinarse o doblarse (V4)
- lavarse o vestirse (V5)
- hacer las tareas del hogar o su trabajo (V6)

La encuesta fue aplicada un mes después de haber empezado el tratamiento y seis meses después. Las respuestas 0 y 1 han sido codificadas en 0 y las respuestas 2 y 3 codificadas en 1.

#### *El Método HJ-BIPLLOT*

Consideremos una matriz de datos  $X$  que contiene evaluaciones de CVRS de un conjunto de pacientes. Las columnas contienen la información relativa a las variables en estudio (V1 a V6) y las líneas contienen la información relativa a las pacientes en estudio. El Método HJ-Biplot, propuesto por Galindo (1986) consiste, en definir una representación gráfica multivariante mediante marcadores  $\mathbf{j}_1, \mathbf{j}_2, \dots, \mathbf{j}_n$  para las filas ( $n$  pacientes) y  $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_p$  para las columnas ( $p$  variables de movilidad) de  $X$ . Estos marcadores son elegidos de forma que ambos puedan superponerse en el mismo sistema de referencia con máxima calidad de representación. Generalmente las filas se representan como puntos y las columnas como vectores. El HJ-Biplot, se basa en la descomposición en valores singulares (DVS) de la matriz de datos. Cualquier matriz real  $X_{(n \times p)}$  de rango  $r$  ( $r \leq \min(n, p)$ ) se puede factorizar como el producto de tres matrices del tal forma:

$$X_{(n \times p)} = U_{(n \times r)} \Lambda_{(r \times r)} V'_{(r \times p)} \quad \text{con} \quad U'U = V'V = I_r$$

donde:

$U_{(n \times r)}$  es la matriz de vectores propios de  $XX'$ .

$V_{(p \times r)}$  es la matriz de vectores propios de  $X'X$ .

$\Lambda_{(r \times r)}$  es una matriz diagonal de  $\lambda_1, \lambda_2, \dots, \lambda_r$ , correspondientes a los  $r$  valores propios de  $XX'$  o  $X'X$ .

Los elementos de  $X_{(n \times p)}$  vienen dados por:

$$x_{ij} = \sum_{k=1}^r \sqrt{\lambda_k} u_{ik} v_{jk} \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, p$$

Así, partiendo de la DVS, la elección de marcadores en la dimensión  $q$  para las filas y columnas de la matriz  $X$  son:

$$J_{(q)} = U_{(q)} \Lambda_{(q)} \quad \text{y} \quad H_{(q)} = V_{(q)} \Lambda_{(q)}$$

La calidad de representación para las filas y para las columnas de la matriz de datos  $X$  es la misma y las filas y columnas vienen expresadas en coordenadas principales.

### *Aplicación*

Aplicando el Modelo de Zwinderman al conjunto de datos en estudio se han calculado las estimaciones de máxima verosimilitud de las componentes  $\mu$  (cambio debido al tratamiento) y  $\tau$  (cambio debido al tiempo) del cambio y se obtuvieron los siguientes resultados (el proceso de estimación se ha ejecutado mediante un script construido en MatLab para el efecto):

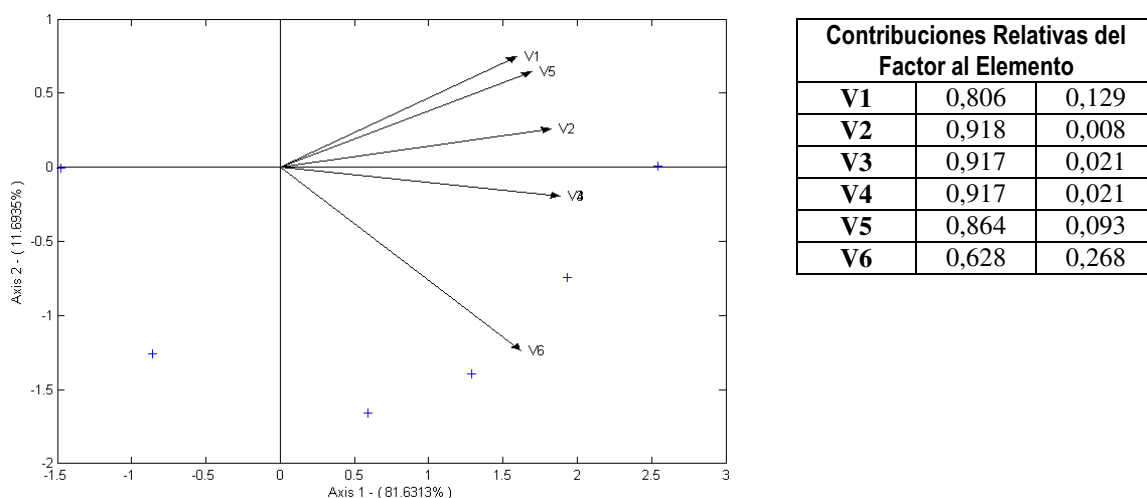
$$\hat{\mu} = -0,63 \quad (SE_{\hat{\mu}} = 0,15) \quad \text{y} \quad \hat{\tau} = -0,25 \quad (SE_{\hat{\tau}} = 0,11).$$

El valor de  $\hat{\tau} = -0,25$  indica una mejoría con el paso del tiempo de la movilidad tanto en el grupo que ha recibido el tratamiento como en el otro. Por otra parte, el valor de  $\hat{\mu}$  indica que la mejoría debida al tratamiento de APD ha

sido estimada en  $-0,63$ . Estos valores indican una mejoría global de la movilidad que, sin embargo, es mucho más significativa en el grupo que recibe el tratamiento de APD.

El conjunto de datos en estudio ha sido sometido a las Técnicas de Biplot. Se han estudiado separadamente las valoraciones de los dos grupos en estudio (con tratamiento y sin tratamiento) un mes después de haber empezado el tratamiento y seis meses después y se ha construido un HJ-Biplot para cada una de estas matrices de datos. Se obtuvieron los resultados que se presentan a continuación.

En una primera fase se han comparado las valoraciones de las pacientes que no han recibido el tratamiento de APD un mes después de haber empezado el tratamiento y seis meses después en los dos instantes. Las Fig.6 y Fig.7 contienen los resultados obtenidos para la primera evaluación y para la segunda evaluación, respectivamente.



**Fig.6.** HJ-Biplot, Inercia Absorbida y Contribuciones (primera evaluación, pacientes sin tratamiento).

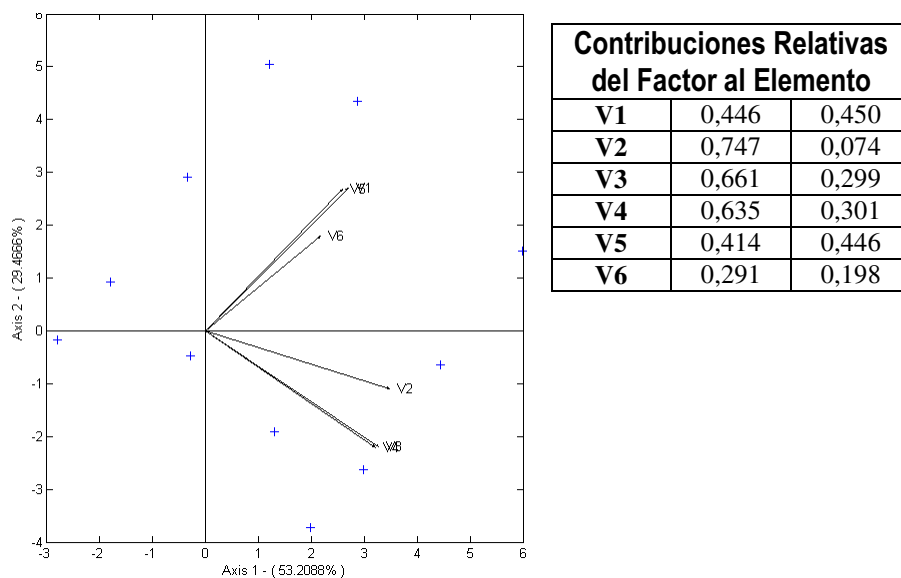


Fig.7. HJ-Biplot, Inercia Absorbida y Contribuciones (segunda evaluación, pacientes sin tratamiento).

Los resultados obtenidos evidencian la existencia de diferencias en cuanto a la proporción de variabilidad explicada por los dos primeros ejes entre las dos evaluaciones. En particular, se registra una disminución de la importancia del eje 1 y un aumento de la importancia de del eje 2, sugiriendo una modificación en la estructura factorial.

Se verifican alteraciones en el posicionamiento de las variables. La contribución del eje 1 a las variables en estudio disminuye entre las dos evaluaciones, mientras el eje 2 cobra importancia para todas las variables excepto la V2. Todas estas variables se relacionan con cuestiones inherentes al nivel de movilidad dentro de casa. Nótese que el grupo en cuestión no ha recibido el tratamiento lo que lleva a suponer que las diferencias registradas se deben al paso del tiempo o a alteraciones en el autopercepción de las pacientes.

En una segunda fase se han comparado las valoraciones de las pacientes que han recibido el tratamiento de APD en los dos instantes. Las Fig.8 y Fig.9



contienen los resultados obtenidos para la primera evaluación y para la segunda evaluación, respectivamente.

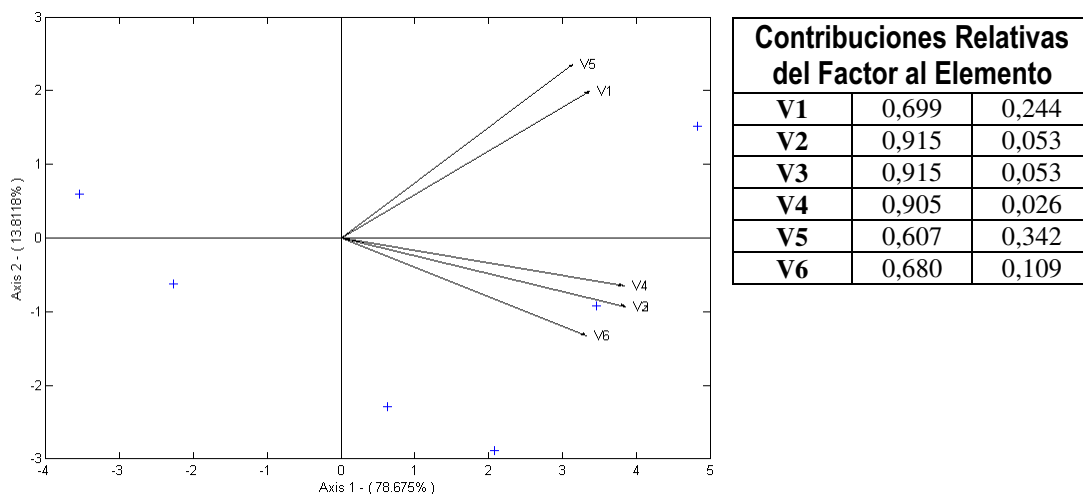


Fig.8. Biplot, Inercia Absorbida y Contribuciones (primera evaluación, pacientes con tratamiento).

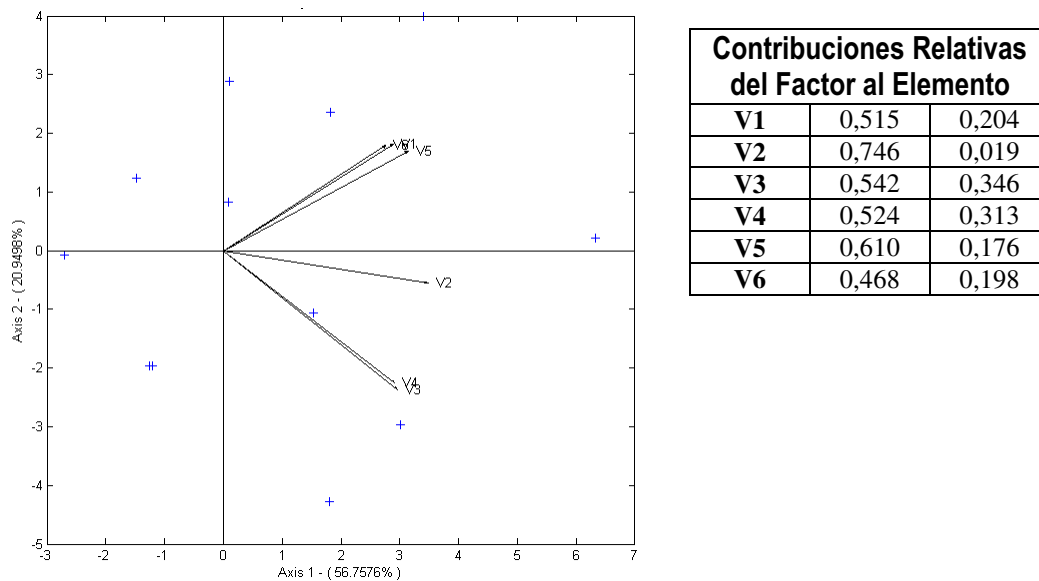


Fig.9. Biplot, Inercia Absorbida y Contribuciones (segunda evaluación, pacientes con tratamiento).

Se registran igualmente diferencias al nivel de la proporción de inercia absorbida en las estructuras factoriales relativas a los dos instantes con una disminución de la importancia del eje 1 y un aumento de la inercia absorbida por el eje 2. Entre las dos evaluaciones, se registra también una disminución de la

correlación entre las variables V3 y V4 y el eje 1 y una alteración significativa en la variable V6.

Estas alteraciones ocurren tanto en el grupo de las pacientes que reciben tratamiento como en el grupo de las pacientes que no reciben. El aumento de la cantidad de inercia absorbida por el eje 2 entre la primera y la segunda evaluación, ponen de manifiesto alteraciones al nivel de la estructura factorial.

Los resultados obtenidos mediante la aplicación de las Técnicas Biplot, llevan a considerar la ocurrencia de eventuales alteraciones al nivel de la autopercepción de las pacientes entre las dos evaluaciones, en ambos grupos. Tales alteraciones pueden deberse a la ocurrencia de cambios y en particular de fenómenos de recalibración, repriorización o reconceptualización, los cuales suceden a menudo en contextos de evaluación de CVRS.

El Modelo de Zwinderman no ha tenido en cuenta los mecanismos de adaptación por parte de los pacientes. En estas pacientes, la mejoría debida al tratamiento ha sido estimada en un  $\hat{\mu} = -0,63$ , mientras la mejoría debida al paso del tiempo se ha estimado en un  $\hat{\tau} = -0,25$ . Sin embargo, no se ha considerado la eventual ocurrencia de fenómenos de recalibración, repriorización o reconceptualización en los pacientes entre las diferentes evaluaciones los cuales, como es bien sabido, ocurren a menudo en contextos de evaluación de CVRS. Se necesitan de herramientas estadísticas que permitan identificar la ocurrencia de estos fenómenos, pudiendo las Técnicas Factoriales constituir una contribución importante en este sentido. En este caso, la utilización de las Técnicas Biplot ha permitido identificar una alteración al nivel de la estructura factorial entre la primera aplicación y la segunda, tanto para el grupo de las pacientes que habían sido sometidas al tratamiento como para el otro, sugiriendo la presencia de cambio.

## 4.6. MODELOS DE ECUACIONES ESTRUCTURALES

El término “Modelos de Ecuaciones Estructurales” (*Structural Equation Modeling* - SEM) no designa una técnica estadística específica sino una familia de procedimientos relacionados como el Análisis de la Estructura de Covarianza, Análisis de Variables Latentes, Análisis Factorial Confirmatorio, etc. Estos modelos se basan en el análisis de la covarianza y constituyen una eficaz herramienta en el estudio de relaciones causales entre variables, siempre que estas relaciones sean de tipo lineal. Sin embargo es importante tener en cuenta que estos métodos no prueban la causalidad, pudiendo ser utilizados más en una vertiente confirmatoria con el objetivo de confirmar relaciones preestablecidas a respecto de un conjunto de variables o bien en una vertiente más exploratoria en la cual el investigador pretende testar hipótesis sobre un conjunto de datos. Jöreskog (1993), sintetiza las aplicaciones de los Modelos de Ecuaciones Estructurales en 3 objetivos distintos:

- confirmar un modelo, cuando el investigador posee un modelo que se acepta o rechaza en base a la correspondencia con las relaciones evidenciadas en un determinado conjunto de datos y tiene un objetivo estrictamente confirmatorio;
- elegir modelos alternativos, cuando el investigador dispone de varios modelos posibles y pretende seleccionar uno;
- generar un modelo, que es quizás el objetivo más común y ocurre siempre que no se dispone de un modelo *a priori*. El modelo buscado debe tener dos propiedades importantes: tener sentido teórico y estadístico.

Siendo los Modelos de Ecuaciones Estructurales un conjunto de procedimientos estadísticos, no es posible identificar su origen con exactitud. En realidad parte de sus raíces datan del principio del siglo XX con los trabajos de

Charles Spearman en las técnicas de Análisis Factorial Exploratoria. Posteriormente, Sewall Wright desarrolla la base del *Path Analysis* en el dominio de la genética que ha sido extendida después al campo de la Sociología, Economía y Psicología. Desde entonces la utilización de los Modelos de Ecuaciones Estructurales ha crecido exponencialmente en muchas áreas distintas.

En el contexto de la **detección y tratamiento del cambio**, la idea subyacente a estas técnicas es utilizar el análisis de las estructuras de covarianza usando el Análisis Factorial Confirmatorio para identificar reconceptualización y recalibración a través del análisis de cambios en la estructura y patrón factorial a lo largo del tiempo. Schmitt (1982) y Schmitt *et al* (1984) proponen el Análisis de Estructuras de Covarianza utilizando el LISREL (Jöreskog & Sörbom, 1981). Además de la inspección de las estructuras factoriales Schmitt (1982) recomienda contrastar la homogeneidad de las estructuras de varianza-covarianza. La aproximación LISREL permite detectar reconceptualización. Posteriormente, Oort (2005) propone una metodología basada en los Modelos de Ecuaciones Estructurales para detectar las varias componentes del *response shift* y que abordaremos más adelante.

Aunque hemos optado por pasar los detalles algebraicos de los distintos procedimientos a un anexo, para facilitar la lectura, en este caso, dado que se trata de la parte central del estudio, detallaremos todo aquello que consideramos imprescindible para poder justificar todas las afirmaciones posteriores.

Kline (2005) sintetiza los 6 pasos básicos de la aplicación de los Modelos de Ecuaciones Estructurales:

- 1) **Especificar el modelo**, lo que significa que las hipótesis inicialmente formuladas por el investigador, son expresadas en la forma de un modelo de ecuaciones estructurales. Estas ecuaciones definen los

parámetros del modelo, y se basan en el establecimiento de relaciones entre variables observadas y latentes.

- 2) **Confirmar si el modelo se encuentra totalmente identificado**, lo que significa que, teóricamente, es posible derivar una única estimación para cada uno de los parámetros del modelo.
- 3) **Seleccionar las medidas de las variables** representadas en el modelo (operacionalizar el constructo), recoger y preparar los datos.
- 4) **Estimar el modelo**, lo que implica el recurso a un programa de ordenador. Esta etapa comprende los siguientes pasos:
  - Evaluar la bondad del ajuste, o sea, determinar hasta qué punto el modelo ajustado explica el conjunto de datos. Muchas veces el modelo tiene que ser reespecificado y, en ese caso, se pasa directamente al paso 5.
  - Interpretar las estimaciones de los parámetros.
  - Considerar modelos equivalentes, con diferentes configuraciones y relaciones, que expliquen los datos tan bien como el modelo elegido. Para un modelo dado de ecuaciones estructurales, existen muchas variaciones posibles, el investigador elige el que mejor se adecua a las relaciones establecidas y es estadísticamente equivalente.
- 5) **Reespecificar el modelo si es necesario** y evaluar la bondad de ajuste del modelo revisado sobre los mismos datos. Cualquier reespecificación debe ser conducida mediante las hipótesis de investigación.
- 6) **Encontrado un modelo satisfactorio hay que describir su análisis.**

Los Modelos de Ecuaciones Estructurales se basan en el análisis de relaciones entre variables. La explicación de la existencia de correlación entre dos variables  $X$  y  $Y$  puede incluir relaciones causales ( $X$  causa  $Y$ ) y relaciones no

causales, como asociaciones espurias entre  $X$  e  $Y$ . El objetivo de los Modelos de Ecuaciones Estructurales es determinar las componentes causales y no causales de las correlaciones observadas. Una relación de causalidad supone que el cambio en una variable produce un cambio en otra variable. Nótese que, la simple asociación entre dos variables no es suficiente ya que correlación no implica causalidad. Para poder inferir que  $X$  es la causa de  $Y$ , hay que verificar tres condiciones:

- 1) **Precedencia temporal**, o sea,  $X$  precede  $Y$ . En general, se acepta la prioridad temporal como una condición de causalidad, o sea, la causa debe preceder el efecto. El establecimiento de la prioridad causal es necesario para determinar si un indicador es causa o efecto de una variable latente.
- 2) La **dirección de causalidad** debe estar correctamente especificada, o sea,  $X$  causa  $Y$  y no  $Y$  causa  $X$ . En una relación de causalidad existe necesariamente una dirección subyacente.
- 3) La **asociación** entre  $X$  e  $Y$  no desaparece cuando variables externas (como causas comunes) se mantienen constantes. Aunque la asociación bivalente no sea condición necesaria y suficiente para una relación causal, cuando una causa y su efecto están aislados de otras influencias, entonces las dos variables podrán estar asociadas. Problemas como términos de perturbación heterocedásticos, autocorrelacionados o multicolinealidad pueden conducir a inferencias incorrectas a respecto a la asociación.

En los Modelos de Ecuaciones Estructurales las variables en estudio y las relaciones entre ellas se suelen representar a través del diagrama causal (*path diagram*) que constituye una representación gráfica de un sistema de ecuaciones donde se especifican tanto las relaciones entre variables dependientes e independientes como las correlaciones entre constructos y entre variables

indicadoras. Los elementos de un diagrama causal se pueden definir de la siguiente forma:

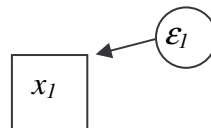
Las **variables observadas, manifiestas o indicadoras** se representan por medio de una casilla rectangular:



Las **variables no observadas, latentes o factores comunes** se representan por medio de un círculo o elipse:



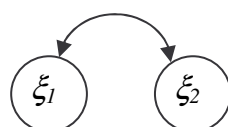
Los **términos de perturbación, error de medida o factor residual** se representan de la siguiente forma:



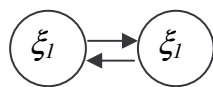
Una relación de **causalidad** se representa por medio de una flecha recta significando que  $\xi_I$  causa  $x_I$ :



Una relación de **asociación** sin interpretación causal entre dos variables se representa por medio de una flecha doble curvada:



Una relación de **reacción** o **causalidad recíproca**, entre dos variables se representa por medio de dos flechas rectas:



Cada flecha está afectada de un parámetro que indica la magnitud del efecto entre ambas variables, siendo este efecto cero cuando no existe flecha. Las variables a las que llega alguna flecha, se denominan por **variables endógenas** y aquellas a las que no llega ninguna flecha por **variables exógenas**. Las variables endógenas están siempre afectadas por un término de perturbación aleatorio.

Se asume que las variables exógenas y los errores están no correlacionados. A esta condición se llama pseudo-aislamiento y constituye una condición fuerte, que es muchas veces violada en los Modelos de Ecuaciones Estructurales.

Los Modelos de Ecuaciones Estructurales se ajustan a matrices de covarianzas, por lo que las variables se suelen considerar sólo centradas con respecto a su media y los parámetros no estandarizados. Las varianzas y covarianzas que se refieren únicamente a variables causa constituyen por sí mismas parámetros del modelo, siendo las reglas de descomposición para derivar las restantes varianzas y covarianzas, las siguientes:

- la covarianza entre dos variables es igual a la suma del efecto directo, los efectos indirectos, los efectos espurios y conjuntos;
- la varianza de una variable dependiente es igual a la varianza del término de perturbación más la varianza explicada por las otras variables del modelo.

Para determinar el número de parámetros en un Modelo de Ecuaciones Estructurales, se suele utilizar la regla que Kline (2005) enuncia de la siguiente forma:



*“El número total de varianzas y covarianzas (o sea, asociaciones no causales) de variables exógenas observadas o no observadas (o sea, perturbaciones) y efectos directos de otras variables observadas sobre las variables endógenas es igual al número de parámetros.”*

Las varianzas y covarianzas de las variables endógenas no son consideradas parámetros del modelo ya que, en estos modelos, las variables endógenas son representadas como efectos de otras variables medidas en el modelo, no siendo, por lo tanto, libres para variar o covariar.

Las dos reglas de descomposición atrás enunciadas, conducen a un sistema de ecuaciones estructurales que expresan cada elemento de la matriz de covarianzas en función de los parámetros del modelo y se puede representar en la forma  $\Sigma = \Sigma(\pi)$ , siendo  $\Sigma$  la matriz de varianza-covarianza de la población y  $\pi$  el vector de parámetros (efectos directos, varianzas y covarianzas de errores y variables exógenas). El modelo completo consiste en un sistema de ecuaciones estructurales y debe cumplir algunos supuestos estadísticos que Rios (2006) sintetiza como:

- Supuestos estadísticos sobre las fuentes de variación (factores exógenos, términos de perturbación y errores de medida) y en concreto sobre la forma de su función de distribución conjunta, que se requieren para la operatividad del modelo. Se suele asumir normalidad multivariante.
- Supuestos sobre el comportamiento de las variables no consideradas, cuyo efecto se recoge en los términos de error de medida o de perturbación. Se asume, o bien que no se han omitido dimensiones relevantes, o bien que si este no fuera el caso, que dichas dimensiones no están correlacionadas con cualquiera de las dimensiones incluidas u omitidas.

- Correlación entre determinados pares de términos de perturbación o términos de error de medida.

Cada parámetro del modelo, puede ser libre, fijo o restringido, dependiendo de las especificaciones del investigador. Un **parámetro libre** es estimado sin ninguna condición a partir del conjunto de datos. En oposición, un **parámetro fijo** es definido de forma a ser igual a una constante y todo el análisis subsiguiente se basa en el valor definido. Un **parámetro restringido** es estimado bajo alguna restricción, muchas veces relacionada con valores de otros parámetros restringidos. Se pueden estimar dos o más parámetros de forma a que sean iguales, proporcionales, mayores o menores que otro parámetro o incluso imponer relaciones no lineales.

El modelo contiene **variables aleatorias** (variables latentes, variables observadas y errores o perturbaciones), **parámetros estructurales** (que describen las relaciones entre las variables) y eventualmente **variables no aleatorias** (variables explicativas cuyos valores se mantienen iguales en muestreos aleatorios repetidos).

Existen dos grandes grupos de modelos: los recursivos y los no recursivos. Los **modelos recursivos** son más directos y tienen dos características básicas: los términos de perturbación son no correlacionados y todos los efectos causales son unidireccionales. En los **modelos no recursivos** los términos de perturbación pueden ser correlacionados. El hecho de que un modelo sea recursivo o no recursivo, tiene varias implicaciones. Las condiciones enunciadas subyacentes a los modelos recursivos simplifican el análisis estadístico, mientras los modelos no recursivos requieren, en general, métodos de análisis más sofisticados.

El modelo estructural debe incluir todas las ecuaciones estructurales que resumen las relaciones entre variables latentes, siendo estas ecuaciones lineales en

las variables y lineales en los parámetros. Siguiendo la notación de Rios (2006), el modelo estructural se representa como

$$\eta = B\eta + \Gamma\xi + \zeta$$

Se asume que

$$E(\eta) = 0$$

$$E(\xi) = 0$$

$$E(\zeta) = 0$$

$\zeta$  está no correlacionado con  $\xi$

$(I-B)$  es una matriz no singular

$\eta$  es un vector ( $Q \times 1$ ) de variables latentes endógenas

$\xi$  es un vector aleatorio de ( $R \times 1$ ) de variables latentes exógenas

$\zeta$  es un vector ( $Q \times 1$ ) de errores latentes en la ecuación, términos de perturbación

$B$  es una matriz ( $Q \times Q$ ) de coeficientes para las variables latentes endógenas

$\Gamma$  es una matriz ( $Q \times R$ ) de coeficientes para las variables latentes exógenas

$\Phi$  es una matriz ( $R \times R$ ) de varianza-covarianza de las variables latentes exógenas  $\xi$ ,  $E(\xi\xi')$

$\Psi$  es una matriz ( $Q \times Q$ ) de varianza-covarianza de los términos de perturbación  $\zeta$ ,  $E(\zeta\zeta')$

$$\eta = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \vdots \\ \eta_Q \end{bmatrix}_{Q \times 1} \quad \xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \vdots \\ \xi_R \end{bmatrix}_{R \times 1} \quad \zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \vdots \\ \zeta_Q \end{bmatrix}_{Q \times 1}$$

$$B = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \beta_{21} & 0 & \cdots & 0 & 0 \\ \beta_{31} & \beta_{32} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_{Q1} & \beta_{Q2} & \cdots & \beta_{Q(Q-1)} & 0 \end{bmatrix}_{Q \times Q} \quad \Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1R} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2R} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{Q1} & \gamma_{Q2} & \cdots & \gamma_{QR} \end{bmatrix}_{Q \times R}$$

Los coeficientes  $B$  son los parámetros estructurales que indican el cambio en el valor esperado de una variable  $\eta$  cuando las otras variables  $\eta$  se incrementan en una unidad, manteniendo constantes las variables  $\xi$ ; la diagonal principal de esta matriz es siempre cero, es decir, que se asume que una variable no es una causa instantánea e inmediata de sí misma; también el modelo asume que  $(I - B)$  es no singular, en consecuencia,  $(I - B)^{-1}$  existe. Los coeficientes de regresión  $\Gamma$  tienen una interpretación análoga. Se asume que los errores  $\zeta$  son homocedásticos y no autocorrelacionados. Las condiciones de homocedasticidad y no autocorrelación no significan que las perturbaciones de dos ecuaciones diferentes necesariamente estarán incorrelacionadas, ni que tendrán la misma varianza.

Este modelo estructural describe las relaciones entre las variables latentes. Sin embargo, para que el modelo exista, es necesario que existan medidas observables de estas variables: variables observadas, manifiestas o indicadoras. El modelo de medida incluye las ecuaciones estructurales que resumen las relaciones entre las variables latentes y las variables observadas. Siguiendo la notación de Rios (2006) las ecuaciones estructurales del modelo de medida se representan como

$$x = \Lambda_x \xi + \varepsilon$$

$$y = \Lambda_y \eta + \delta$$

donde la variable observada  $x$  es la variable indicadora de  $\xi$  e  $y$  indicadora de  $\eta$ . Se asume que

$$E(\eta) = 0$$

$$E(\xi) = 0$$

$$E(\zeta) = 0$$

$$E(\delta) = 0$$

$\varepsilon$  está no correlacionado con  $\eta$ ,  $\xi$  y  $\delta$

$\delta$  está no correlacionado con  $\xi$ ,  $\eta$  y  $\varepsilon$

$x$  es un vector ( $J \times 1$ ) de variables observadas indicadoras de  $\xi$

$y$  es un vector ( $I \times 1$ ) de variables observadas indicadoras de  $\eta$

$\varepsilon$  es un vector ( $J \times 1$ ) de errores de medida para  $x$

$\delta$  es un vector ( $I \times 1$ ) de errores de medida para  $y$

$\Lambda_x$  es una matriz ( $J \times R$ ) de coeficientes que relacionan  $x$  con  $\xi$

$\Lambda_y$  es una matriz ( $I \times Q$ ) de coeficientes que relacionan  $y$  con  $\eta$

$\Theta_\varepsilon$  es una matriz ( $J \times J$ ) de varianza-covarianza de  $\varepsilon$ ,  $E(\varepsilon\varepsilon')$

$\Theta_\delta$  es una matriz ( $I \times I$ ) de varianza-covarianza de  $\delta$ ,  $E(\delta\delta')$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_J \end{bmatrix}_{J \times 1} \quad \Lambda_x = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ \lambda_2 & 0 & \cdots & 0 \\ 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_J \end{bmatrix}_{J \times R} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_J \end{bmatrix}_{J \times 1}$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_I \end{bmatrix}_{I \times 1} \quad \Lambda_y = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ \lambda_2 & 0 & \cdots & 0 \\ 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_J \end{bmatrix}_{I \times Q} \quad \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \vdots \\ \delta_I \end{bmatrix}_{I \times 1}$$

Los coeficientes  $\Lambda$  son la magnitud del cambio esperado en las variables observadas por una unidad de cambio en la variable latente. Estos coeficientes estructurales son coeficientes de regresión para los efectos de las variables latentes sobre las variables observadas. Las columnas de las matrices  $\Lambda_x$  y  $\Lambda_y$  contienen algunos elementos iguales a cero y otros diferentes de cero, esto depende de la correspondencia de la variable indicadora con su respectiva variable latente indicada. En una columna que corresponde a una variable latente en particular, el

elemento que está en una fila que corresponde a una indicadora de la mencionada variable latente, será diferente de cero, en caso contrario será igual a cero.

Las variables aleatorias  $\delta$  y  $\varepsilon$  son los errores de medida para  $y$  y  $x$  respectivamente, son los disturbios que interrumpen la relación entre las variables latentes y las variables observadas. Se asume que los términos de error son homocedásticos y no autocorrelacionados a través de las observaciones. En el análisis factorial  $\delta$  y  $\varepsilon$  se denominan factores únicos, y cada uno de ellos está dividido en componente específica y no específica. Las matrices  $\Theta_\varepsilon$  y  $\Theta_\delta$  son varianzas y covarianzas de los términos de error de medida. Las diagonales principales contienen las varianzas de error asociadas con las indicadoras.

Sólo cuando es posible encontrar un único estimador para cada parámetro, el Modelo de Ecuaciones Estructurales se considera teóricamente **identificado**, siendo para eso necesario que las observaciones disponibles sean al menos tantas como los parámetros del modelo (el número de grados de libertad tiene que ser mayor que cero) y a todas las variables no observadas (latentes) debe ser asignada una escala. En la mayoría de las técnicas multivariantes, el número de grados de libertad se calcula en base a la diferencia entre el tamaño de la muestra y el número de parámetros estimados, mientras en un Modelo de Ecuaciones Estructurales el número de grados de libertad es dado por la diferencia entre el número de covarianzas y el número efectivo de parámetros a estimar (parámetros libres).

Los parámetros del modelo se suelen estimar a través del Método de **Máxima Verosimilitud**, aunque se puedan igualmente obtener estimaciones por Regresión Múltiple. Los parámetros estimados son interpretados como coeficientes de regresión (y los respectivos coeficientes de determinación en términos de variabilidad explicada). Para coeficientes estandarizados habitualmente se considera que valores absolutos inferiores a 0,1 traducen efectos

pequeños, valores en torno del 0,3 indican efectos típicos o medios y valores superiores a 0,5 traducen efectos largos (Kline, 2005).

El efecto total de una variable sobre otra es dado por la suma de todos los efectos directos e indirectos. El Test de Sobel permite calcular los efectos indirectos en modelos estructurales (Preacher & Hayes, 2008) de la siguiente forma:

- siendo  $a$  es el coeficiente no estandarizado del efecto  $X \rightarrow Y_1$  y  $SE_a$  el respectivo error estándar;
- siendo  $b$  es el coeficiente no estandarizado del efecto  $Y_1 \rightarrow Y_2$  y  $SE_b$  el respectivo error estándar;
- el efecto indirecto de  $X$  sobre  $Y_2$  a través de  $Y_1$  es dado por el producto  $ab$ ;
- el error estándar de  $ab$  es dado por  $SE_{ab} = \sqrt{b^2 SE_a^2 + a^2 SE_b^2}$ ;
- para muestras suficientemente largas, el cociente  $ab/SE_{ab}$  se interpreta como un test  $z$  del efecto indirecto no estandarizado.

El efecto total estandarizado de una variable sobre otra, se puede interpretar como la parte de su correlación observada debida a las relaciones causales supuestas. La suma de los efectos totales estandarizados y todas las restantes asociaciones no causales representadas en el modelo es igual a la correlación implícita en el modelo. Kline (2005) enuncia una regla antigua que permite calcular “a mano” las correlaciones en modelos recursivos, conocida por *tracing rule*:

*“La correlación implícita en un modelo es dada por la suma de todos los efectos causales y asociaciones no causales a partir de todos los trazados validos entre dos variables en un modelo recursivo. Un trazado «válido» significa que en una variable no entra y sale la misma flecha y que para cada trazado existe una única flecha.”*

La diferencia entre la correlación implícita en el modelo y la correlación observada es llamada correlación residual. En modelos posiblemente identificados (en los que puede existir una única solución para los parámetros que iguale la matriz de covarianza observada e implícita) todas las correlaciones residuales son cero. Sin embargo, en modelos posiblemente sobreidentificados (los que incluyen menos parámetros que las varianzas y covarianzas que recoge la matriz de entrada) no todas las correlaciones residuales tienen que ser cero. Una de las reglas comúnmente mencionadas en la literatura de los Modelos de Ecuaciones Estructurales establece que correlaciones residuales con valor absoluto superior a 0,1 indican que el modelo puede no explicar convenientemente la correlación observada.

Los Modelos de Ecuaciones Estructurales han sido pensados para muestras grandes (aunque algunos autores presenten algunas sugerencias para el análisis de muestras pequeñas). La dimensión de la muestra constituye un aspecto de gran importancia en los Modelos de Ecuaciones Estructurales (así como en la mayor parte de las técnicas estadísticas) ya que el error estándar asociado a los resultados obtenidos a partir de muestras largas es menor que en muestras de pequeña dimensión. Breckler (1990) considera que, para propósitos descriptivos una muestra de dimensión inferior a 100 es “pequeña”, entre 100 y 200 “media” y “larga” si la dimensión es mayor que 200. Por otra parte, es necesario tener en cuenta que, cuanto más parámetros tiene el modelo, mayor la dimensión requerida para la muestra. Se suele considerar como deseable el ratio de 20/1 entre las observaciones disponibles y parámetros libres, aunque en la práctica se acepte muchas veces el ratio 10/1.

Para evaluar la calidad de ajuste del modelo estimado se necesitan **medidas de bondad de ajuste**, las cuales miden la correspondencia entre la matriz de datos y la predicción del modelo propuesto. Existe un sin número de estadísticos que



permiten la cuantificación de la bondad de ajuste, de entre los cuales se destacan los siguientes:

- El **Ji-Cuadrado** ( $\chi^2$ ) mide la bondad de ajuste global del modelo. Si  $\chi^2=0$ , el modelo ajusta los datos perfectamente, mientras cuanto mayor es  $\chi^2$ , peor es la calidad de ajuste del modelo. El Test del Ji-Cuadrado contrasta la hipótesis nula de que el modelo es correcto, o sea, de que se ajusta perfectamente a los datos;
- El **NFI** (*Normed Fit Index*) es una medida de bondad de ajuste global que compara el valor del  $\chi^2$  del modelo con el de un modelo base que, en general, es el modelo de independencia, también llamado modelo nulo, y que asume covarianzas muestrales nulas entre las variables observadas.
- El **RMSEA** (*Root Mean Square Error of Approximation*) mide la parsimonia del modelo (dados dos modelos con igual poder explicativo, es mejor el que sea más simple). El RMSEA mide el error de la aproximación, o sea, la discrepancia entre la matriz de covarianza poblacional y la matriz de covarianza estimada en base al modelo. En general se considera que  $RMSEA \leq 0,05$  indica un buen ajuste, valores entre  $0,05$  y  $0,08$  sugieren un ajuste razonable y valores de  $RMSEA \geq 0,1$  un ajuste pobre;
- El **CFI** (*Comparative Fit Index*) evalúa el mejoramiento relativo del modelo en lo que respecta al ajuste, cuando el modelo en estudio es comparado con el modelo nulo. Valores de CFI superiores a  $0,9$  indican buena calidad de ajuste;

- El **SRMSR** (*Standardized Root Mean Square Residual*) es una medida del valor medio absoluto de la correlación de los residuos, o sea, mide la diferencia global entre las correlaciones observadas y previstas. Si el ajuste fuera perfecto, entonces  $SRMSR=0$ . Cuanto mayor el valor de SRMSR, más pobre es el ajuste. En general, valores de  $SRMSR \leq 0,1$  se consideran favorables;
- El **AIC** (*Akaike Information Criterion*) es un índice de ajuste predictivo, basado en la población y no en la muestra. Esta medida de la bondad del ajuste de un modelo estadístico estimado ha sido desarrollada por Akaike (1974). Se contrapone la complejidad de un modelo estimado a la calidad del ajuste a los datos, siendo el caso general:  $AIC=2k-2\ln L$ , donde  $k$  es el número de parámetros y  $L$  la función de máxima verosimilitud. Se busca el modelo que mejor explica los datos con el número mínimo de parámetros. El AIC es una función de la complejidad del modelo. Para la misma calidad de ajuste global, un modelo con menor AIC es preferible a uno con valor más elevado que será más complejo;
- El **GFI** (*Goodness of Fit Index*) es análogo al Coeficiente de Determinación ( $R^2$ ) en la Regresión Múltiple, varía entre 0 y 1 y cuantifica la proporción de varianza explicada.

Si se detectan fallos relativos al ajuste del modelo, es necesario buscar métodos para mejorarlo. Esto puede deberse a la complejidad del modelo, al gran número de restricciones definidas, a la inclusión o exclusión incorrecta de un parámetro. Al detectar tales problemas, hay que reespecificar el modelo añadiendo o eliminando parámetros del modelo original. Este proceso debe conducirse según una justificación teórica y no únicamente bajo criterios empíricos. El primer paso es distinguir las relaciones teóricas y empíricas (estimadas o no), ya que las

primeras no pueden ser modificadas, mientras las segundas pueden ser reespecificadas de forma que se mejore la calidad de ajuste del modelo. Rios (2006) reúne algunas indicaciones útiles para llevar a cabo una reespecificación del modelo:

- Introducir las modificaciones una a una, y volver a examinar los resultados antes de introducir la siguiente. A menudo ocurre que el hecho de introducir una modificación cambia substancialmente la estimación de las partes del modelo aún no modificadas, y aumenta o reduce la necesidad de introducir modificaciones adicionales;
- Introducir las modificaciones que mejoran la bondad del ajuste (añadiendo parámetros) antes de las que mejoran la parsimonia (eliminando parámetros);
- Evitar las modificaciones teóricamente no interpretables por muy significativos que sean los estadísticos de los modificadores de Lagrange. Tan desaconsejable es ignorar completamente la información contenida en los datos empíricos como basarse únicamente en ella y descartar un parámetro a favor de otro basándose en pequeñas diferencias en significación;
- Si el cambio en el valor de la ji-cuadrado es inferior al valor crítico se optará por el modelo más restrictivo, y si es superior al valor crítico por el menos restrictivo. Otros estadísticos adecuados para comparar modelos anidados son los índices de bondad de ajuste que tienen en cuenta la parsimonia del modelo: NNFI, CFI, AIC, CAIC y RMSEA.
- Si los modelos no son anidados pero incluyen las mismas variables, pueden emplearse también los índices NNFI, CFI, AIC, CAIC y RMSEA pero no el cambio en el valor ji-cuadrado.
- Si los modelos no incluyen las mismas variables sólo pueden emplearse los índices NNFI, CFI y RMSEA.

En la literatura son varias las situaciones en que es posible encontrar la modelación de datos de CVRS con recurso a las Ecuaciones Estructurales, verificándose recientemente un gran aumento de referencias al respecto.

Burgess *et al* (2000) han utilizado el Modelo de Ecuaciones Estructurales para modelizar un conjunto de datos de CVRS relativo a 279 pacientes con SIDA en varios estadios de la enfermedad. Los datos han sido obtenidos a través de un cuestionario que pretendía evaluar cuestiones relativas a CVRS, personalidad, estrategias de *coping* y soporte social. El análisis del modelo de ecuaciones estructurales construido, ha permitido concluir que el estado de salud está modestamente asociado al psicológico ( $\beta=0,44$ ), el neuroticismo está fuertemente asociado con el estado psicológico ( $\beta=-0,73$ ) aunque flacamente relacionado con el físico ( $\beta=-0,21$ ). Los autores concluyen que el neuroticismo tiene una fuerte influencia en la CVRS, la cual es independiente del estado de salud, o sea, el neuroticismo está más fuertemente asociado a los aspectos psicológicos de la CVRS que el estado de salud.

Yeh (2002) ha testado dos modelos de CVRS en 102 niños tailandeses con cáncer en base a las ecuaciones estructurales. Se han considerado como estímulos ambientales la gravedad de la enfermedad, la edad, el sexo, la comunicación con los otros, la comprensión de la enfermedad. Se consideró la gravedad de la enfermedad un constructo latente que consiste en la función física, en la función psicológica, en la función pares/escuela, síntomas asociados al tratamiento/enfermedad y funciones cognitivas. Se ha observado que el constructo gravedad de la enfermedad demuestra una fuerte asociación con el estado de la enfermedad, valores de laboratorio y con el número total de hospitalizaciones. Se concluyó también que la CVRS presentaba una alta validez de constructo con los cinco dominios testados y que los estímulos ambientales influyen las respuestas psicosociales.

Brossart *et al* (2002) han propuesto un método que combina el Modelo de Ecuaciones Estructurales con el Modelo de Curvas de Crecimiento. Sin embargo, la utilización de este modelo supone que se conoce *a priori* la forma del cambio ya que el objetivo es testar el ajuste del modelo construido de forma a parametrizar estos cambios, lo que construye una limitación.

Rios (2006) y Rios *et al* (2007) aplican los Modelos de Ecuaciones Estructurales a la detección de cambio, en un grupo de trabajadores asalariados, relativamente a sus niveles de satisfacción y autovaloraciones de las mismas cosas un año después de su situación laboral haber sido afectada por una alteración en el tipo de contrato. Los autores concluyen que ha ocurrido **recalibración**.

Dentro de la amplia familia de los Modelos de Ecuaciones Estructurales, los más utilizados en el análisis de datos longitudinales son las Curvas de Crecimiento Latente. Sin embargo estos modelos no contemplan el fenómeno de *response shift*. Posteriormente, Oort (2005) ha propuesto un procedimiento basado en los Modelos de Ecuaciones Estructurales específicamente destinado a detectar *response shift* y, en particular, a identificar recalibración, repriorización y reconceptualización. En los apartados siguientes presentamos los Modelos de Curvas de Crecimiento y el Procedimiento de Oort (*op.cit.*).

#### **4.6.1. Modelos de Curvas de Crecimiento Latente**

Los métodos basados en el Análisis de Curvas de Crecimiento (Potthoff & Roy, 1964) posibilitan el análisis de medidas repetidas. Actualmente, estos modelos son ampliamente utilizados para analizar la ocurrencia de cambios en contextos de datos de CVRS ya que permiten superar algunas de las

limitaciones de las técnicas tradicionales para trabajar con medidas repetidas, con la ventaja de posibilitar información adicional.

El Análisis de Curvas de Crecimiento permite investigar el patrón de cambio implícito en una variable respuesta a lo largo del tiempo. En particular, cuando asociado a los Modelos de Ecuaciones Estructurales, el Análisis de Curvas de Crecimiento Latente puede volverse una herramienta útil y versátil para estudiar cambio en datos longitudinales (McArdle & Epstein, 1987; Meredith & Tisak, 1990; Muthén, 1991; Willett & Sayer, 1994). El objetivo se centra en estimar los factores de crecimiento no observados que generan las medidas repetidas observadas. Las Curvas de Crecimiento Latente enfocan las trayectorias individuales de crecimiento de cada individuo, permitiendo modelizar la tendencia del cambio individual o de grupos de individuos. Esta técnica es bastante flexible y aplicable a una grande diversidad de situaciones, permitiendo al investigador explorar las variables que influyen la tendencia global y las trayectorias individuales.

En el contexto de la CVRS, las Curvas de Crecimiento proporcionan una estimación del cambio a lo largo del tiempo en la percepción individual de los pacientes sobre un conjunto de constructos en estudio.

Considerando cuatro evaluaciones de una variable observable  $Y$  sobre un conjunto de individuos, Curran & Hussong (2001, 2003) y Curran *et al* (2004) definen el Modelo de Curva de Crecimiento como:

$$y_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{it}$$

donde  $y_{it}$  es la medida obtenida sobre el individuo  $i$  en el instante  $t$ ,  $\lambda_t = 0,1,2,3$  para las  $T=4$  evaluaciones. El valor observado de  $Y$  para cada paciente  $i$  en el instante  $t$  se expresa como una combinación aditiva del intercepto de la trayectoria

del individuo  $i$  ( $\alpha_i$ ), su pendiente ( $\beta_i$ ) multiplicado por el coeficiente del tiempo ( $\lambda_t$ ) y un error aleatorio asociado al individuo  $i$  en el instante  $t$  ( $\varepsilon_{it}$ ). El **intercepto** representa el valor inicial asociado a cada individuo, que es el mismo a lo largo de los diferentes tiempos  $t$ . La **pendiente** se relaciona con la tasa de crecimiento de cada individuo  $i$ . El intercepto y la pendiente son modelizados como variables latentes, necesarias para describir el crecimiento. Las medidas repetidas observadas en cada instante  $t$  son utilizadas para estimar las trayectorias de crecimiento subyacente.

Los interceptos y pendientes individuales son las componentes de la función de crecimiento y pueden expresarse como:

$$\alpha_i = \mu_\alpha + \zeta_{\alpha i} \quad y \quad \beta_i = \mu_\beta + \zeta_{\beta i}$$

indicando que los interceptos y pendientes individuales pueden expresarse como una función aditiva de un intercepto medio global ( $\mu_\alpha$ ) y de una pendiente media global ( $\mu_\beta$ ) para el conjunto de individuos.

Del modelo presentado resulta que las medidas repetidas observadas de la variable  $Y$  proporcionan una trayectoria. Además, la varianza de las desviaciones proporciona una estimación directa del grado de variabilidad en los interceptos y pendientes en el conjunto de individuos.

El primer paso del análisis consiste en ajustar el modelo de curva latente presentado al conjunto de datos disponible. Este modelo proporciona una visión inicial. El proceso de estimación se basa en el Modelo de Ecuaciones Estructurales. En particular, las medidas repetidas son utilizadas como indicadores múltiples de los dos factores latentes. El primer factor representa el intercepto de la trayectoria

y el segundo factor la pendiente de la trayectoria. El modelo se define restringiendo las cuatro cargas factoriales del intercepto iguales a 1 y las cuatro cargas factoriales relativas a la pendiente igual a  $\lambda_t=t-1$  (en nuestro modelo,  $t=1,2,3,4$ ). La Fig. 10 contiene el Modelo de Curva de Crecimiento de trayectoria lineal para cuatro medidas repetidas  $y_1, y_2, y_3, y_4$ .

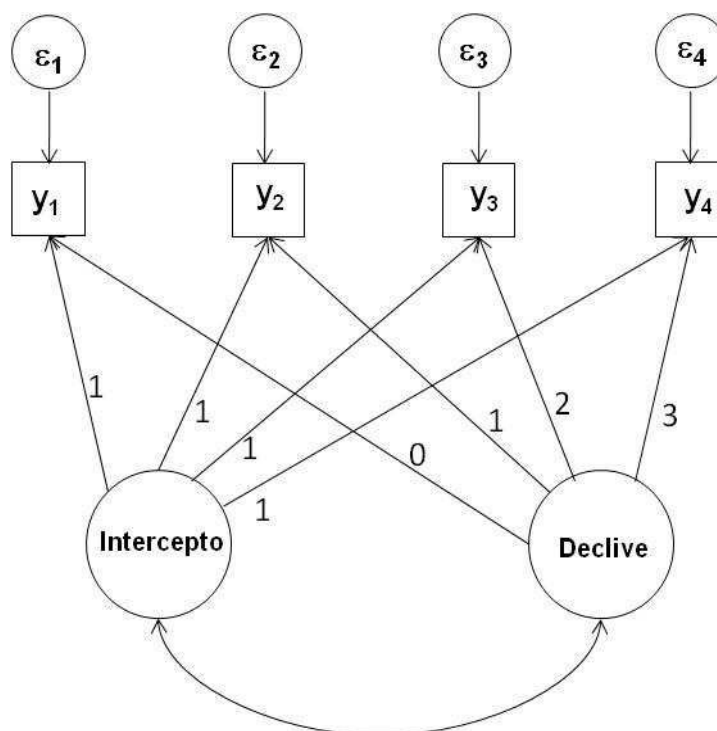


Fig. 10. Modelo de Curva de Crecimiento de trayectoria lineal.

Como se había determinado que el primer instante sería el cero, el intercepto refleja el valor implícito de la medida en el instante inicial. Para el caso de  $T=4$  medidas repetidas y una trayectoria lineal, el número de parámetros a estimar a partir de los datos es nueve. Los dos efectos fijos son las medias de cada factor latente ( $\mu_\alpha$  y  $\mu_\beta$ ) que son calculadas a partir de todos los individuos de la muestra. Los efectos aleatorios son representados por cuatro parámetros: la **varianza de cada factor latente** que refleja el grado de variabilidad individual en los interceptos y pendientes sobre todos los individuos de la muestra ( $\psi_\alpha$  y  $\psi_\beta$ ), la **covarianza entre los dos factores latentes** que se interpreta como el grado de asociación entre la variación individual de los interceptos y la variación individual



de las pendientes ( $\psi_{\alpha\beta}$ ) y la **varianza residual** para cada medida repetida que traduce la varianza no explicada ( $\sigma_t^2$ ). En conjunto, estos parámetros captan la trayectoria de la muestra en estudio, el grado de variabilidad sobre las trayectorias individuales en torno de los valores medios y la proporción de variabilidad no explicada por la trayectoria.

Tal como en los Modelos de Ecuaciones Estructurales, se evalúa la calidad del ajuste mediante el estadístico ji-Cuadrado ( $\chi^2$ ) y el RMSEA. Se observan los índices de modificación, la estructura de medias y covarianzas de los residuos.

En el modelo presentado se ha asumido que las medidas repetidas están linealmente relacionadas con el tiempo, o sea, el cambio de una unidad en el tiempo está asociado al cambio de una unidad ( $\beta$ ) en la variable observada, siendo la magnitud de esta relación constante a lo largo del tiempo. Sin embargo, en algunas situaciones, las medidas repetidas obtenidas pueden relacionarse con el tiempo de forma no lineal de forma que el cambio en  $y$  no es constante entre evaluaciones. Esta relación puede asumir muchas formas distintas como la cuadrática, la exponencial o funciones que son estimadas a partir de las características de la muestra en estudio. Por ejemplo, mientras el modelo lineal es definido únicamente por los factores de intercepto y pendiente, el modelo cuadrático incluye un tercer factor latente de forma a captar la curvatura presente en las trayectorias individuales. El modelo cuadrático para las trayectorias individuales vendría:

$$y_{it} = \alpha_i + \lambda_t \beta_{Li} + \lambda_t^2 \beta_{Qi} + \varepsilon_{it}$$

donde  $\alpha_i$  sigue siendo el intercepto de la trayectoria,  $\beta_{Li}$  es la componente lineal de la trayectoria y  $\beta_{Qi}$  la componente cuadrática de la trayectoria. Las cargas

factoriales tendrían que reflejar las relaciones entre el factor cuadrático latente y las medidas repetidas que son los cuadrados de los factores lineales (Fig.11).

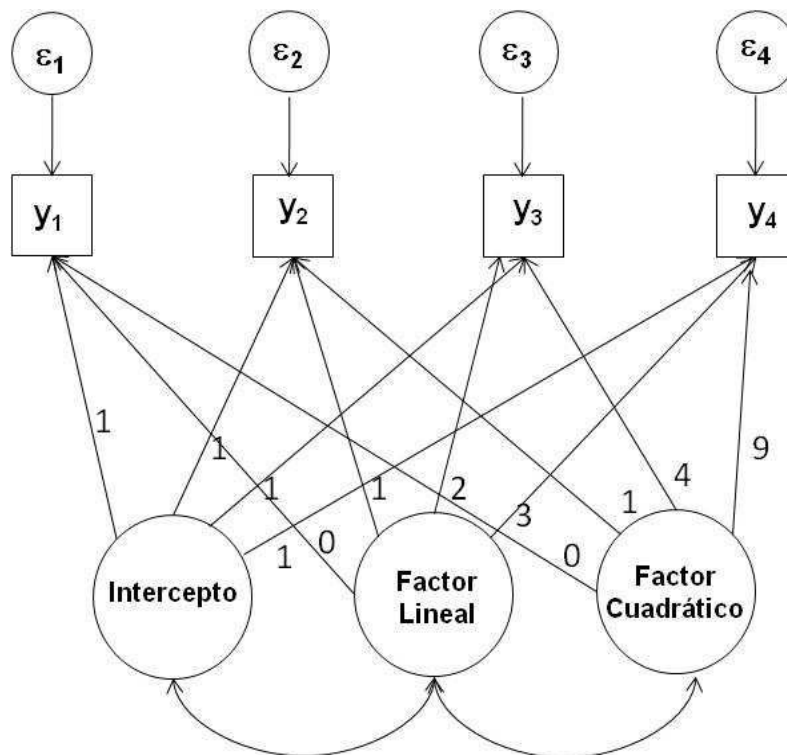


Fig.11. Modelo de Curva de Crecimiento de trayectoria cuadrática.

Mientras el modelo lineal implica un cambio constante en  $y$  entre evaluaciones consecutivas, el modelo cuadrático implica un cambio diferencial (por ejemplo, pueden registrarse grandes cambios iniciales en las trayectorias que disminuyen con el paso del tiempo o pueden registrarse pequeños cambios iniciales que después se acentúan).

Las funciones exponenciales también pueden ser utilizadas para modelizar curvas de crecimiento. En este caso, los incrementos o disminuciones son proporcionales a los incrementos o disminuciones anteriores. Un modelo exponencial sería:

$$y_{it} = \alpha_i + (1 - e^{-\lambda t})\beta_i + \epsilon_{it}$$

donde  $y_{it}$  es la medida definida anteriormente,  $\alpha_i$  el intercepto de la trayectoria,  $\beta_i$  la magnitud del cambio implícito en una observación, relativamente al nivel inicial y  $\gamma$  la tasa de crecimiento en  $y$  a lo largo del tiempo.

Puede ocurrir que las formas funcionales más utilizadas como la lineal, la cuadrática o la exponencial no conduzcan a un buen ajuste de los datos disponibles. Una aproximación alternativa es estimar la forma directamente a partir de los datos. Esto implica la estimación de un factor de intercepto y de un factor de cambio latente simple donde el conjunto de las cargas factoriales del factor latente pendiente son libremente estimados a partir de los datos en la vez de fijarlos. Por ejemplo, en la vez de fijar las cargas factoriales en 0, 1, 2 y 3, se fija la primera carga factorial del factor lineal en 0 y la segunda en 1 de forma a establecer una métrica y se estiman libremente la tercera y la cuarta cargas factoriales. Así, en este modelo, no se fijan las cargas factoriales a valores específicos asociados a una determinada forma funcional conocida, sino que uno o más factores son libremente estimados a partir de los datos. En este caso, en la vez de tener tres factores (por ejemplo, intercepto, lineal y cuadrático), se tienen dos factores: el intercepto y un factor “forma” que capta la no linealidad del cambio a través del tiempo. En estos casos, el estadístico ji-Cuadrado ( $\chi^2$ ) puede proporcionar una estimación del grado de mejoramiento al ajustar un modelo relativamente a otro.

Si se encuentra variabilidad individual significativa en los factores de intercepto y pendiente, el paso siguiente consiste en introducir variables explicativas de forma a predecir la variabilidad observada. Suponiendo que tenemos dos variables explicativas  $z_1$  y  $z_2$  (en un contexto de CVRS podría tratarse por ejemplo del sexo y tratamiento) el modelo vendría:

$$\alpha_i = \mu_\alpha + \gamma_1 z_{1i} + \gamma_2 z_{2i} + \zeta_{\alpha i} \quad y \quad \beta_i = \mu_\beta + \gamma_3 z_{1i} + \gamma_4 z_{2i} + \zeta_{\beta i}$$

Se trata de un modelo de crecimiento condicional ya que las diferencias individuales observadas en el punto inicial y el cambio a través del tiempo están siendo modelizados como función de las variables  $z$ . Nótese que estas variables varían entre los individuos sin embargo no varían a lo largo del tiempo ya que son medidas únicamente en el instante inicial. Por otra parte, los parámetros de regresión  $\gamma$  asociados a las variables explicativas son estimados de forma común a todos los individuos.

Aunque los modelos de curvas de crecimiento puedan revelarse útiles para detectar *response shift*, no permiten estimar directamente las componentes de recalibración, repriorización y reconceptualización. Estos métodos pueden constituir una herramienta eficaz como primer paso para evaluar la trayectoria de la muestra de pacientes en estudio.

Un caso particular de los Modelos de Curvas de Crecimiento Latente es el modelo propuesto por Zee (1998), específicamente pensado para modelizar datos de CVRS.

#### 4.6.1.1. El Modelo de Zee

Zee (1998) propone un modelo de curva de crecimiento bajo la ocurrencia de cambios en el estado de salud de los pacientes y bajo la existencia de datos faltantes para describir la CVRS. Esta curva es condicional a una variable dependiente del tiempo que representa los posibles estados de salud. Zee (*op.cit.*) aborda la problemática de la ocurrencia de datos faltantes y de medidas obtenidas a partir de visitas irregulares.

Sea  $Y_{ri}$  un vector  $q_{ri}$ -dimensional que contiene medidas de CVRS para el  $i$ -ésimo paciente del  $r$ -ésimo grupo de tratamiento. Entonces, los elementos de  $Y_{ri}$

son las  $q_{ri}$  medidas de CVRS, obtenidas en los instantes  $t_j$ ,  $j=1,\dots,q_{ri}$ . Se asume que  $Y_{ri}$  tiene una distribución Normal Multivariante con media

$$y_{rij} = f(\omega_r; t_j) = \beta_{r0} + \beta_{r1}t_j + \beta_{r2}t_j^2 + \dots + \beta_{rp}t_j^p$$

donde  $\omega_r$  es el vector de parámetros  $(\beta_{r0}, \beta_{r1}, \beta_{r2}, \dots, \beta_{rp})^T$

y matriz de varianza-covarianza simétrica compuesta<sup>2</sup>

$$\sigma_r^2[(1-\rho_r)I_i + \rho_r J_i]$$

donde  $I_i$  es la matriz identidad,  $J_i$  una matriz de “unos” de dimensión  $q_{ri}$  y  $\sigma_r^2$  y  $\rho_r$  son parámetros desconocidos. El instante  $t_j$  es el instante en que el  $j$ -ésimo elemento del vector de CVRS ha sido medido.

Los estimadores de máxima verosimilitud de los parámetros del modelo generan una curva de crecimiento polinomial que utiliza toda la información de los pacientes disponible, sin apagar del análisis registros de datos incompletos y luego no habiendo necesidad de recorrer a la imputación de datos para los campos incompletos. Por otra parte, datos resultantes de evaluaciones irregulares de la CVRS no constituyen obstáculo al proceso de modelización ya que, en base a la suposición de composición simétrica, no son necesarias ventanas para formatear los datos artificialmente en la forma rectangular.

Zee (1998) propone extender el modelo de curva de crecimiento presentado para controlar el efecto de los estados de salud experimentados por los pacientes.

---

<sup>2</sup> **Matriz Simétrica Compuesta** - matriz de varianzas-covarianzas en la cual las covarianzas residuales para cada individuo son no correlacionadas con las que dicen respecto a otros individuos.

Para tal, introduce una variable que varía con el tiempo y que representa el *status* (“estado de salud”), en el instante que se evaluaron los pacientes. Los “estados de salud” pueden por ejemplo ser definidos como “en tratamiento” *versus* “fuera de tratamiento” o “antes de la administración del tratamiento” y “después de la administración del tratamiento”. La introducción de los estados de salud, está basada en la suposición de que la CVRS de los pacientes es afectada por una condición específica. Este aspecto es particularmente importante cuando el objetivo del estudio es acceder a los impactos de corto y largo plazo de un tratamiento dado en oposición a un tratamiento estándar.

Sea  $\delta_{ris}$  un vector de variables indicadoras de dimensión  $q_{ris}$ , representando el *status* para el  $i$ -ésimo paciente en el  $r$ -ésimo grupo de tratamiento con el  $s$ -ésimo estado de salud, donde  $s=1, \dots, M$ , siendo  $M$  el número total de estados de salud. El elemento de  $\delta_{ris}$  en el instante  $t_{sj}$  es  $\delta_{risj}=1$  si las medidas fueran obtenidas durante un estado de salud específico  $s$  y  $\delta_{risj}=0$  si las medidas fueran obtenidas durante un estado de salud diferente. Dado que los datos de CVRS poseen variabilidad relativamente grande, incluso en el estado que precede a la administración del tratamiento, la bondad de ajuste del modelo puede ser mejorada si se incluye un estado de CVRS,  $x_{ri}$  con parámetro desconocido  $\alpha_r$ , cuando se estiman las curvas de crecimiento medias para los dos grupos. El modelo de curvas de crecimiento puede escribirse como

$$Y_{ri} \sim N \left( \alpha_r x_{ri} + \sum_{s=1}^M \delta_{ris} (\beta_{r0s} + \beta_{r1s} t_s + \beta_{r2s} t_s^2 + \dots + \beta_{rps} t_s^p), \sum_r \right)$$

donde

$$\delta_{risj} = \begin{cases} 0 & \text{si la medida no ha sido obtenida en el estado } s \\ 1 & \text{si la medida ha sido obtenida en el estado } s \end{cases}$$

$t_{sj}$  es un vector  $q_{ris}$ -dimensional que contiene los instantes en que han sido

medidos los datos de CVRS. El modelo de curva de crecimiento modificado para

un vector de observaciones  $Y_{ri}$ , tiene distribución Normal Multivariante con media

$$y_{rij} = f(\omega_{rs}; x_{ri}, \delta_{risj}, t_{sj}) = \alpha_r x_{ri} + \delta_{risj} (\beta_{r0s} + \beta_{r1s} t_{sj} + \beta_{r2s} t_{sj}^2 + \dots + \beta_{rps} t_{sj}^p)$$

donde  $\omega_r$  es el vector de parámetros  $(\alpha_r, \beta_{r0s}, \beta_{r1s}, \beta_{r2s}, \dots, \beta_{rps})^T$ , y matriz de varianza-covarianza intraclase  $\sum_r$  con parámetros  $\sigma_{rs}^2$  y  $\rho_{rs}$ ,  $s=1, \dots, M$ . La estructura de correlación intraclase puede ser definida dentro de cada grupo de tratamiento y dentro de cada estado de salud, con el objetivo de volver este modelo tan realista cuanto posible en la práctica. La matriz de varianza-covarianza tiene la siguiente estructura:

$$\sum_r = \left[ \begin{array}{cccc} \sigma_{r1}^2 \begin{pmatrix} 1 & \rho_{r1} & \dots & \rho_{r1} \\ \rho_{r1} & 1 & \dots & \rho_{r1} \\ \dots & \dots & \dots & \dots \\ \rho_{r1} & \rho_{r1} & \rho_{r1} & 1 \end{pmatrix} & \Theta & \dots & \Theta \\ \dots & \dots & \dots & \dots \\ \Theta & \Theta & \dots & \sigma_{rM}^2 \begin{pmatrix} 1 & \rho_{rM} & \dots & \rho_{rM} \\ \rho_{rM} & 1 & \dots & \rho_{rM} \\ \dots & \dots & \dots & \dots \\ \rho_{rM} & \rho_{rM} & \rho_{rM} & 1 \end{pmatrix} \end{array} \right]$$

donde  $\Theta$  son matrices de ceros y  $\sigma_{rs}^2$  y  $\rho_{rs}$ ,  $s=1, \dots, M$  son parámetros desconocidos.

La estructura simétrica de la matriz de varianzas-covarianzas permite acomodar datos faltantes y visitas irregulares, poseyendo a la vez flexibilidad suficiente para trabajar con diferentes estados de salud. La matriz de varianzas-covarianzas estimada permite no solamente efectuar inferencias entre tratamientos en los diversos instantes para cada estado de salud sino también proceder a la comparación global. Además, es preferible comparar los grupos de tratamiento utilizando la matriz de varianzas-covarianzas asintótica del modelo completo antes que considerar los diferentes estados de salud como grupos separados.

Trabajar con el modelo completo en vez de considerar subgrupos, mejora la eficiencia de los estimadores y reduce la ocurrencia de falsos positivos.

Las curvas de crecimiento globales asociadas a los tratamientos, pueden ser comparadas dado el valor medio del estado de CVRS. En un ensayo clínico randomizado, el objetivo principal es comparar la CVRS entre grupos de tratamiento. Se puede construir un contraste apropiado dentro del modelo para testar el efecto del tratamiento. El contraste entre grupos de tratamiento se puede realizar una vez que tengamos las estimaciones de los parámetros y las respectivas varianzas asintóticas.

Veamos cómo podría aplicarse el Modelo de Zee (1998) en la práctica: supongamos que tenemos dos grupos de pacientes de cáncer randomizados, uno de los cuales es sometido a la quimioterapia convencional y el otro a un nuevo tratamiento alternativo. Se pide a los pacientes que completen un cuestionario de CVRS de tres en tres semanas. En el análisis se consideran dos estados de salud “en tratamiento” y “fuera de tratamiento”. La puntuación para la CVRS global se exprime en una escala de 0 a 100 donde 0 traduce una CVRS muy mala y 100 excelente. De la aplicación del cuestionario resulta un conjunto de medidas repetidas de CVRS por paciente. Estos datos son sometidos a una regresión *loess*<sup>3</sup> con el objetivo de obtener los estimadores de regresión local y el gráfico de la función *loess*, el cual posibilita una visión general de los perfiles de CVRS. En particular, el gráfico de la función *loess* constituye una buena aproximación del perfil longitudinal global de la CVRS.

---

<sup>3</sup> **Regresión *Loess* o Regresión Local Ponderada** – es una generalización del método de las medias móviles en el que se obtiene un valor predicho para cada observación a través del ajuste de una regresión lineal local ponderada, en la cual los pesos decrecen a la medida que aumenta la distancia a la observación de interés. La conexión de los valores predichos produce una curva suave cuyo grado de suavización es definido por un parámetro que determina la amplitud del conjunto de datos a ser utilizado (o sea, la dimensión de la ventana). Tratase de un método no paramétrico para ajustar curvas mediante el ajuste de varias funciones de regresión en puntos vecinos y está indicado para casos de difícil decisión sobre la aplicación de una curva paramétrica o en presencia de valores discrepantes.



El proceso de ajuste del modelo empieza con el ajuste de un polinomio de grado 1 y, se ajusta un polinomio de grado 2 si el *AIC* es más bajo que el obtenido para el grado 1 o si la función polinomial de grado 2 es significativa y así sucesivamente. En este análisis son de gran importancia:

- las representaciones gráficas de las curvas de crecimiento para los diferentes estados de salud;
- las estimaciones asociadas a la varianza para los diversos tratamientos en los varios estados de salud;
- los contrastes entre los varios grupos de tratamiento para los varios estados de salud con el objetivo de comparar el nivel de CVRS global.

Como hemos visto, el Modelo de Zee (1998) tiene como objetivo evaluar cómo uno o más tratamientos afectan la CVRS de los pacientes, siendo esta una cuestión de investigación muy relevante y de extrema importancia. Zee (*op.cit.*) propone un modelo de curva de crecimiento que incorpora los “estados de salud” experimentados por los pacientes, con el objetivo de describir un patrón global de CVRS con datos faltantes. Al comparar tratamientos sin tener en cuenta los estados de salud, se obtienen resultados sesgados debido a la eventual diferencia de eficacia de los tratamientos. Por lo tanto, la relevancia del método propuesto, se justifica en base al hecho de que la CVRS es diferente en distintas fases del tratamiento y luego, en el proceso de estimación, se puede utilizar una variable dependiente del tiempo como partición de las medidas de CVRS repetidas. Este proceso puede ayudar a clarificar las interpretaciones relativas a la comparación de los tratamientos y el patrón de CVRS bajo un estado de salud específico.

Mientras en gran parte de las técnicas se ignoran los datos faltantes, Zee (*op.cit.*) argumenta que los distintos “estados de salud” experimentados por los pacientes tienen diferentes impactos en la CVRS, confundiendo la comparación porque los mecanismos generadores de datos faltantes son diferentes. Por ejemplo, el estado de salud durante el periodo de tratamiento puede ser bastante distinto

del estado de salud del paciente durante el periodo sin tratamiento, o sea, el tratamiento que los pacientes reciben puede afectar su calidad de vida y la ocurrencia de estados de salud. Así, desde este punto de vista, el argumento de Zwinderman (1990) sobre que los datos faltantes dependen apenas de las medidas observadas, sería válido únicamente cuando se compara CVRS en grupos de pacientes con el mismo estado de salud. En el modelo propuesto por Zee se asume que el mecanismo causante de datos faltantes es aleatorio, condicional a un dado estado de salud.

En suma, la idea subyacente al modelo propuesto por Zee (1998) es comparar la CVRS en grupos de pacientes sometidos a diferentes tratamientos dentro de cada estado de salud. El control de los estados de salud experimentados por los pacientes permite reducir eventuales sesgos en la comparación de tratamientos.

El Modelo de Zee (*op.cit.*) de las curvas de crecimiento condicionadas a los estados de salud, permite superar algunos problemas que subyacen generalmente al análisis de datos de CVRS. Se destacan como **puntos fuertes**:

- Los modelos de curvas de crecimiento posibilitan una idea general sobre el patrón evolutivo de la CVRS. En particular, las representaciones gráficas permiten una observación de los patrones de CVRS entre grupos de tratamiento y estados de salud;
- Es posible efectuar contrastes específicos con el objetivo de comparar la diferencia en el patrón global de CVRS entre tratamientos, a través de las estimaciones para la varianza-covarianza, evitando problemas de análisis en los subgrupos;
- La estructura de la matriz de varianzas-covarianzas posibilita que las estimaciones asociadas a la varianza y correlación, varíen entre los estados de salud y los tratamientos;
- Se asume que los datos faltantes son aleatorios dentro de cada estado de salud y se utiliza toda la información de los pacientes disponible (incluyendo los casos de visitas irregulares);

Desde nuestro punto de vista, algunas de las **limitaciones del Modelo de Zee (1998)**, tienen que ver con los siguientes aspectos:

- No posibilita la cuantificación (ni tampoco la evaluación) del cambio asociado a las variables latentes en estudio, sino únicamente el análisis de la evolución del patrón global de CVRS;
- Se supone que las medidas de CVRS obtenidas a partir de los pacientes en los diversos grupos poseen distribución Normal Multivariante;
- Suponer que la matriz de varianzas-covarianzas tiene estructura simétrica compuesta, implica asumir que las covarianzas residuales asociadas a los individuos, son no correlacionadas entre sí, siendo esta suposición un tanto irrealista.

El Modelo de Zee (1998) **ignora los mecanismos de adaptación** por parte de los pacientes. No se considera la eventual ocurrencia de fenómenos de **recalibración, repriorización o reconceptualización** en los pacientes entre las diferentes evaluaciones los cuales, como es bien sabido, ocurren a menudo en contextos de evaluación de CVRS.

#### 4.6.2. El Procedimiento de Oort

El primer abordaje en el sentido de identificar cambio a través de los Modelos de Ecuaciones Estructurales se debe a Schmitt (1982). Teniendo en cuenta que la definición operacional de cambio Gamma es un cambio en la estructura factorial, este autor consideró que este tipo de cambios pueden ser estudiados mediante el Análisis Factorial Confirmatorio y no sólo por medio de procedimientos rotacionales. Sin embargo, Schmitt (*op.cit.*) solamente ha considerado la estructura de covarianza, sin tener en cuenta cambios que se

manifiestan en la estructura de los valores medios, como la recalibración uniforme y los cambios Alpha.

Posteriormente Millsap & Hartog (1988) han aplicado el Modelo de Ecuaciones Estructurales tanto a la estructura de valores medios como covarianzas. Estos autores propusieron la utilización de estas técnicas en la evaluación del cambio a través del tiempo, planteando que el cambio Gamma puede ser detectado mediante la comparación de la estructura factorial de los dos tiempos (por ejemplo, la estructura factorial del *Pre-Test* y del *Post-Test*). En particular, se utiliza una técnica de rotación para rotar las matrices de patrones de los factores hasta conseguir máxima similitud, se calculan y evalúan índices de similitud, tales como el coeficiente de congruencia (Tucker, 1951). Si los valores de congruencia en la estructura factorial (por ejemplo, entre el *Pre-Test* y el *Post-Test*) son altos, entonces éstos indican ausencia de cambio Gamma. Por otra parte, para detectar cambio Beta, Millsap & Hartog (1988) proceden a una reparametrización similar a los Modelos de Regresión. La operacionalización de cambio Beta se basa en la regresión de las variables del segundo tiempo sobre el primero. Consecuentemente, una baja calidad de ajuste es atribuida a la no linealidad, que a su vez es atribuida a la ocurrencia de cambio Beta. Sin embargo, cambios Beta pueden ocurrir por razones distintas de la no linealidad, por lo que no linealidad no indica necesariamente la ocurrencia de cambios Beta. Las diferencias entre las estimaciones de los parámetros de regresión entre grupo experimental y grupo de control son igualmente tomadas como evidencia de cambio Beta, aunque en realidad tales diferencias puedan ocurrir sin haber ocurrido *response shift*.

Posteriormente, Oort (2005) aplica el Modelo de Ecuaciones Estructurales para investigar los diferentes tipos de cambio, basándose en operacionalizaciones y procedimientos de detección diferentes de los utilizados por Schmitt (1982) y Millsap & Hartog (1988). Oort (*op.cit.*) ha propuesto un procedimiento para detectar *response shift* y medir el verdadero cambio que es aplicable si uno o más constructos *target* (como CVRS) son medidos a través de múltiples ítems (como los

ítems o escalas de un cuestionario de CVRS). A este procedimiento están subyacentes los conceptos de:

- **reconceptualización**, entendida como un cambio en el significado del contenido de un ítem;
- **repriorización**, entendida como un cambio de la importancia relativa del ítem como indicador del constructo *target*;
- **recalibración**, entendida como un cambio en el significado de las opciones de respuesta de un ítem, o sea, en los estándares de medición internos de los individuos.

Además, este autor hace la distinción entre recalibración uniforme y no uniforme. Cuando existe recalibración, los individuos cambian su interpretación de las opciones de la escala de respuesta (valores, etiquetas o ambos). Si el cambio afecta todas las opciones de respuesta en la misma dirección y en la misma extensión, se trata de una **recalibración uniforme** y solo afecta las medias de las variables observadas. Por otra parte, la escala de medición puede estirarse o encogerse, completa o parcialmente y, en ese caso, la **recalibración** se dice **no uniforme** y también afecta a las varianzas de las variables observadas. El procedimiento propuesto por Oort (2005) se basa en el análisis de las medias y covarianzas. Por lo tanto, sólo detecta *response shift* y verdadero cambio si el fenómeno es experimentado por una parte substancial de los individuos.

Los métodos habitualmente propuestos para detectar *response shift*, envuelven aplicaciones adicionales del mismo cuestionario (como el *Then-Test*) o evaluaciones alternativas adicionales del mismo constructo. Este tipo de métodos para evaluar *response shift* tiene la ventaja de no necesitar medidas adicionales. El procedimiento propuesto por Oort (2005) para detectar *response shift*, utiliza datos de solamente dos tiempos de medida, posibilitando la detección de varios tipos de *response shift* y medición del verdadero cambio.

Supongamos que se está midiendo CVRS a través de varias variables observadas. El modelo para las puntuaciones observadas de un sujeto arbitrario  $i$ , es dado por

$$y_i = \tau + \Gamma \xi_i + \zeta_i$$

donde  $y_i$  es el vector de las puntuaciones observadas,  $\xi_i$  es el vector de las puntuaciones no observadas del factor común y  $\zeta_i$  es el vector de las puntuaciones no observadas del factor residual. La matriz  $\Gamma$  y el vector  $\tau$  contienen las cargas de los factores e interceptos.

Considerando los conceptos medidos a través del SF-36 (Funcionamiento Físico - FF, Desempeño Físico - DF, Dolor Corporal - DC, Salud General - SG, Vitalidad - VT, Funcionamiento Social - FS, Desempeño Emocional - DE y Salud Mental - SM), resultaría el modelo de medida de la Salud Física y Mental, presentado en la Fig.12.

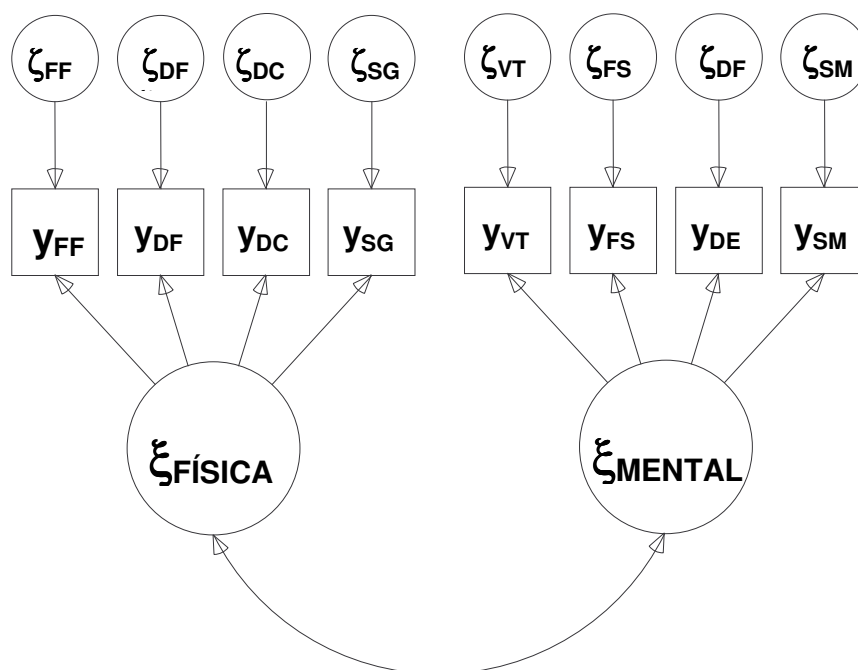


Fig.12. Modelo de Medida de la Salud Física y Mental (SF-36).

$y_{FF}$ ,  $y_{DF}$ ,  $y_{DC}$ ,  $y_{SG}$ ,  $y_{VT}$ ,  $y_{FS}$ ,  $y_{DE}$  y  $y_{SM}$  son las variables observadas, o sea, son operacionalizaciones de FF, DF, ..., SM. Las variables latentes son los factores comunes  $\xi_{FÍSICA}$  y  $\xi_{MENTAL}$ , que son operacionalizaciones de los conceptos de Salud Física y Mental. La variable latente  $\xi_{FÍSICA}$  es medida por  $y_{FF}$ ,  $y_{DF}$ ,  $y_{DC}$ ,  $y_{SG}$  y la variable latente  $\xi_{MENTAL}$  por  $y_{VT}$ ,  $y_{FS}$ ,  $y_{DE}$ ,  $y_{SM}$ . Las otras variables latentes son los factores residuales  $\zeta_{FF}$ ,  $\zeta_{DF}$ ,  $\zeta_{DC}$ ,  $\zeta_{SG}$ ,  $\zeta_{VT}$ ,  $\zeta_{FS}$ ,  $\zeta_{DE}$  y  $\zeta_{SM}$  que son no correlados con los factores comunes. Los factores residuales representan toda la varianza en FF, DF, ..., SM que no es explicada por los factores comunes. Por ejemplo, todo lo que FF tiene en común con DF, DC y SG es representado por la variable  $\xi_{FÍSICA}$  y todo lo demás de FF es representado por  $\zeta_{FF}$ . Las flechas representan las cargas factoriales ( $\Gamma$ ) que son los coeficientes para las regresiones de los factores sobre las variables observadas. Las cargas muestran cuán indicativa es cada variable observada en los factores comunes. Por ejemplo, si la carga de  $y_{DF}$  en  $\xi_{FÍSICA}$  es mayor que la carga de  $y_{DC}$  en  $\xi_{FÍSICA}$ , significa que DF es más indicativa de  $\xi_{FÍSICA}$  que DC. Los interceptos  $\tau$  (no representados en la figura) indican la "facilidad" con que se obtienen puntuaciones elevadas en la escala de CVRS. Por ejemplo, si DF y DC son medidos en la misma escala y  $y_{DF}$  tiene menor intercepto que  $y_{DC}$ , entonces es más "fácil" puntuar alto en  $y_{DF}$  que en  $y_{DC}$ , o sea, para un determinado nivel de Salud Física, los pacientes evidencian más DC que DF. En el contexto de la Teoría de la Respuesta al Ítem, las cargas e interceptos pueden interpretarse como parámetros de discriminación y dificultad.

En el Modelo de Ecuaciones Estructurales se modelizan valores medios y estructuras de covarianza en vez de las puntuaciones de los sujetos individuales. Los valores medios de las puntuaciones observadas son dadas por

$$\bar{y} = \mu = \tau + \Gamma\alpha$$

donde  $\alpha$  es el vector de medias de los factores comunes. Las covarianzas de las variables observadas son dadas por

$$COV(y, y') = \Sigma = \Gamma\Phi\Gamma' + \Omega$$

donde  $\Phi = COV(\xi, \xi')$  y  $\Omega = COV(\zeta, \zeta')$  son las matrices de varianza-covarianza de los factores comunes y de los factores residuales, respectivamente. Puede ser conveniente descomponer las matrices  $\Phi$  y  $\Omega$  en

$$\Phi = \sqrt{Diag(\Phi)\Phi^*} \sqrt{Diag(\Phi)} \quad \text{y} \quad \Omega = \sqrt{Diag(\Omega)\Omega^*} \sqrt{Diag(\Omega)}$$

donde  $\Phi^*$  y  $\Omega^*$  son las matrices de correlación y donde  $Diag(\Phi)$  y  $Diag(\Omega)$  son las matrices diagonales conteniendo las varianzas de los factores comunes y residuales, respectivamente.

El vector  $\alpha$  contiene los valores medios de "FÍSICA" y "MENTAL", la matriz  $\Phi^*$  contiene la correlación entre "FÍSICA" y "MENTAL" y  $Diag(\Phi)$  contiene las varianzas de "FÍSICA" y "MENTAL". La matriz  $\Omega$  contiene las varianzas de los factores residuales ( $\zeta_{FF}$ ,  $\zeta_{DF}$ , ...,  $\zeta_{SM}$ ). Sin embargo, algunas veces dos variables observadas tienen más en común que lo que puede ser explicado por la correlación entre sus factores comunes. Por ejemplo tanto DF como DE se refieren a limitaciones. Si la covarianza observada entre  $y_{DF}$  y  $y_{DE}$ , no es suficientemente explicada por la covarianza entre  $\xi_{FÍSICA}$  y  $\xi_{MENTAL}$  entonces, en la vez de complicar el modelo añadiendo otro factor común  $\xi_{LIMITACIONES}$ , se añade al modelo una correlación entre  $\zeta_{DF}$  y  $\zeta_{DE}$ .



Si tenemos dos momentos de medida, las matrices de parámetros son constituidas por matrices más pequeñas. El vector de los interceptos  $\tau$  y la matriz de las cargas  $\Gamma$  son dados por

$$\tau = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} \quad \text{y} \quad \Gamma = \begin{bmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{bmatrix}$$

donde los índices 1 y 2 dicen respecto al primer y segundo tiempos de medida, respectivamente. Así,  $\tau_1$  y  $\tau_2$  contienen los interceptos y  $\Gamma_1$  y  $\Gamma_2$  contienen las cargas del primer y segundo tiempos.

En este caso,  $\Gamma_1$  contiene las cargas de las variables observadas  $y_{FF,1}$ ,  $y_{DF,1}$ , ...,  $y_{SM,1}$  en los factores comunes  $\xi_{FÍSICA,1}$  y  $\xi_{MENTAL,1}$  y  $\Gamma_2$  contiene las cargas de las variables observadas  $y_{FF,2}$ ,  $y_{DF,2}$ , ...,  $y_{SM,2}$  en los factores comunes  $\xi_{FÍSICA,2}$  y  $\xi_{MENTAL,2}$ . Los ceros en  $\Gamma$  proporcionan matrices adecuadas en términos de dimensión a las operaciones definidas.

El vector de las medias de los factores comunes  $\alpha$  y la matriz de varianzas-covarianzas de los factores comunes  $\Phi$  son dados por

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \quad \text{y} \quad \Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}$$

donde  $\alpha_1$  y  $\alpha_2$  contienen los valores medios de los factores comunes relativos al primer y segundo tiempos;  $\Phi_{11}$  y  $\Phi_{22}$  contienen las varianzas-covarianzas de los factores comunes relativos al primer y segundo tiempos; y  $\Phi_{21}=\Phi_{12}$  contiene las covarianzas cruzadas de los factores comunes del primer y segundo tiempos. En el ejemplo,  $\Phi_{11}$  contiene las varianzas de  $\xi_{FÍSICA,1}$  y  $\xi_{MENTAL,1}$ , y la covarianza entre  $\xi_{FÍSICA,1}$  y  $\xi_{MENTAL,1}$ ,  $\Phi_{22}$  contiene las varianzas de  $\xi_{FÍSICA,2}$  y  $\xi_{MENTAL,2}$ , y la covarianza

entre  $\xi_{FÍSICA,2}$  y  $\xi_{MENTAL,2}$ , y  $\Phi_{21}$  y  $\Phi_{12}$  contienen las covarianzas entre  $\xi_{FÍSICA,2}$  y  $\xi_{FÍSICA,1}$ , entre  $\xi_{FÍSICA,2}$  y  $\xi_{MENTAL,1}$ , entre  $\xi_{MENTAL,2}$  y  $\xi_{FÍSICA,1}$  y entre  $\xi_{MENTAL,2}$  y  $\xi_{MENTAL,1}$ .

La matriz de varianzas-covarianzas de los factores residuales  $\Omega$  es dada por

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$$

donde  $\Omega_{11}$  y  $\Omega_{22}$  contienen las varianzas de los factores residuales del primer y segundo tiempos (y por veces una o más covarianzas); y  $\Omega_{21}=\Omega_{12}$  contienen las covarianzas cruzadas de los factores residuales del primer y segundo tiempos. Normalmente, los factores residuales de diferentes variables no son correlados, por eso, todas las matrices  $\Omega_{ij}$  son diagonales, dada una matriz  $\Omega$  con "estructura en banda".

En este caso, los factores residuales del primer tiempo  $\zeta_{FF,1}$ ,  $\zeta_{DF,1}$ , ...,  $\zeta_{SM,1}$  no son correlados unos con los otros; los factores residuales del segundo tiempo  $\zeta_{FF,2}$ ,  $\zeta_{DF,2}$ , ...,  $\zeta_{SM,2}$  no son correlados unos con los otros; sin embargo, los factores residuales de las mismas variables medidas en tiempos distintos, son correlados entre sí. La matriz  $\Omega_{21}$  es una matriz diagonal que contiene las covarianzas entre  $\zeta_{FF,1}$  y  $\zeta_{FF,2}$ , entre  $\zeta_{DF,1}$  y  $\zeta_{DF,2}$ , etc.

El modelo es extensible a más que dos tiempos de medida.

El objetivo de nuestro análisis es investigar *response shift* y otros tipos de cambio testando la invarianza entre tiempos de los elementos de las submatrices referidas. Oort (2005) asocia varios tipos de cambio a diferencias en los parámetros, como se sintetiza en la Tabla 1.

Matriz	Parámetro	Tipo de Cambio	Hipótesis de Invarianza
<b>Response Shift</b>			
Pat( $\Gamma$ )	Cargas de los Factores Comunes	Reconceptualización	Pat( $\Gamma_1$ )= Pat( $\Gamma_2$ )
$\Gamma$	Cargas de los Factores Comunes	Repriorización	$\Gamma_1 = \Gamma_2$
$\tau$	Interceptos	Recalibración (uniforme)	$\tau_1 = \tau_2$
Diag( $\Omega$ )	Varianzas de los Factores Residuales	Recalibración (no uniforme)	Diag( $\Omega_{11}$ )= Diag( $\Omega_{22}$ )
<b>Cambio Alpha</b>			
$\alpha$	Medias de los Factores Comunes	Cambio en las medias	$\alpha_1 = \alpha_2$
<b>Otros Tipos de Cambio</b>			
Diag( $\Phi$ )	Varianzas de los Factores Comunes	Verdadero cambio en las varianzas	Diag( $\Phi_{11}$ )= Diag( $\Phi_{22}$ )
$\Phi^*$	Correlaciones de los Factores Comunes	Reconceptualización o repriorización de nivel elevado	$\Phi_{11}^* = \Phi_{22}^*$
$\Omega^*$	Correlaciones de los Factores Residuales	Reconceptualización o repriorización de nivel bajo	$\Omega_{11}^* = \Omega_{22}^*$

**Tabla 1.** Tipos de Cambio e Hipótesis de Invarianza asociadas (Oort, 2005).

**Nota:** Pat( $\Gamma$ ) es el patrón de coeficientes cero y no-cero en la matriz  $\Gamma$ ; Diag() designa los elementos diagonales de una matriz.

### *Reconceptualización*

Las cargas de los factores comunes expresan la extensión según la cual las variables observadas son indicativas de los factores comunes. Los factores comunes son variables latentes, no observables. Por lo tanto, un concepto particular que es representado por un factor común, es definido por las variables observadas que son indicativas de ese factor. En el ejemplo, FF, DF, DC y SG cargan en el primer factor, por eso el nombre del primer factor es Salud Física. Variables observadas con cargas cero en un factor particular no pertenecen al dominio o concepto definido por ese factor.

Los conceptos que son representados por los factores comunes son definidos por el patrón de cero y no-cero en las cargas de la matriz  $\Gamma$ . Esta matriz  $\Gamma$  es constituida por dos submatrices  $\Gamma_1$  y  $\Gamma_2$ , que contienen las cargas de los factores comunes para el primer y segundo tiempos de medida. Si el patrón de la matriz  $\Gamma_1$  difiere del patrón de la matriz  $\Gamma_2$ , entonces uno o más conceptos son definidos por

conjuntos distintos de variables observadas y se puede concluir que ha ocurrido una **reconceptualización**.

Supongamos que la carga del factor común de  $y_{SG}$  en  $\xi_{\text{MENTAL}}$  es cero en el primer tiempo y diferente de cero en el segundo. Entonces  $y_{SG}$  no pertenece al conjunto de variables que definen el concepto de Salud Mental en el primer tiempo, mientras en el segundo ya pertenece. En este caso, el concepto de Salud Mental ha sido reconceptualizado. Además, se podría concluir también que el propio concepto de Salud General había sido reconceptualizado: en el primer tiempo apenas la Salud Física de los pacientes afecta las respuestas a los ítems de Salud General, mientras en el segundo tiempo, las respuestas son también afectadas por su Salud Mental. Cualquiera una de las interpretaciones pone de manifiesto que cambios en el patrón de las cargas de los factores comunes evidencian la ocurrencia de reconceptualización.

### *Repriorización*

Si la magnitud de la carga de una determinada variable observada cambia de un tiempo a otro, entonces la variable ha pasado a ser más o menos indicativa del concepto en cuestión. En este caso se asume que la variable ha pasado a ser más o menos importante en la medición del concepto.

Por ejemplo, si la carga del factor común  $y_{FS}$  en  $\xi_{\text{MENTAL}}$  es más bajo en el primer tiempo que en el segundo, entonces FS se ha vuelto más importante en la medición de la Salud Mental, o sea, el reconocimiento de FS como indicador de la Salud Mental ha cambiado. Se puede decir que los pacientes dan más valor a su FS en su evaluación de su Salud Mental. Esto se puede interpretar como una **repriorización** de la variable FS. Además, se puede decir que ha ocurrido una repriorización del factor Salud Mental ya que el efecto de la Salud Mental en las respuestas de los pacientes a los ítems de FS se ha vuelto más fuerte. En suma, un

cambio en los valores de las cargas de los factores comunes puede ser tomado como evidencia de la ocurrencia de *response shift*.

### *Recalibración*

Se ha definido recalibración como un cambio en los estándares internos de medida del individuo. Como resultado, el individuo cambia su interpretación de las opciones de respuesta de la escala (valores, etiquetas o ambos). Si este cambio afecta todas las opciones de respuesta en la misma dirección y magnitud, entonces se trata de una **recalibración uniforme** que apenas afecta los valores medios de las variables observadas. Caso contrario, la escala de medición se extiende o se reduce, total o parcialmente, afectando también la varianza de las variables observadas. Por lo tanto, la recalibración uniforme se manifiesta únicamente en la estructura de la media, mientras la **recalibración no uniforme** se manifiesta tanto en la estructura de la media como en la estructura de covarianza de las variables observadas (o sea,  $\mu$  y  $\Sigma$ ).

Por ejemplo, si observamos un cambio en la media de  $y_{DC}$  que no está justificado por el cambio en la media de  $\xi_{FÍSICA}$ , entonces podremos concluir que los individuos han cambiado su evaluación de lo que es específico en la medición de DC, así como la evaluación de DC. Quizás nuevas experiencias después de la primera evaluación hayan cambiado su idea del nivel de dolor que consideraban "severa", "moderada", etc. Si todos los puntos de la escala de respuesta están asociados a un nivel de dolor más alto que antes, entonces el cambio se dice uniforme. Por otra parte, si apenas algunas opciones de respuesta están asociadas a un nivel más elevado de dolor que antes, entonces el cambio es no uniforme.

La recalibración uniforme se refleja por un cambio en los valores medios de las variables observadas ( $\mu$ ) que no puede ser atribuido a un cambio en las medias de los factores comunes ( $\alpha$ ), mientras la recalibración no uniforme por un cambio en las varianzas de las variables observadas ( $\text{Diag}(\Sigma)$ ) que no puede ser atribuido a

un cambio en la varianza de los factores comunes ( $\text{Diag}(\Phi)$ ). Por lo tanto, diferencias a lo largo de los tiempos entre los interceptos  $\tau$  (o sea, dentro de  $\tau$ , entre  $\tau_1$  y  $\tau_2$ ) y entre las varianzas residuales (o sea, dentro de  $\text{Diag}(\Omega)$ , entre  $\text{Diag}(\Omega_1)$  y  $\text{Diag}(\Omega_2)$ ) son consideradas como evidencia de recalibración uniforme e recalibración no uniforme, respectivamente.

Oort (2005) explica porque los interceptos y varianzas residuales son indicativos de recalibración mirando los factores residuales  $\zeta$  que figuran en las puntuaciones observadas del modelo. Estos factores residuales son una composición de los factores específicos  $\eta$  y los errores  $\varepsilon$ . Los factores específicos  $\eta$  representan todo lo que es específico en una variable observada, mientras los factores de error  $\varepsilon$  representan fluctuaciones aleatorias, debidas a errores accidentales, como errores de lectura por parte del individuo o errores de digitación de los datos. Consecuentemente, cambios en los interceptos reflejan cambios en el grupo de medias de los factores específicos, mientras cambios en las varianzas residuales reflejan principalmente cambios en las varianzas de los factores específicos.

Por ejemplo,  $y_{FF}$  está entre los indicadores de  $\xi_{\text{FÍSICA}}$ . La variable  $y_{FF}$  tiene una componente de "Salud Física", representada por  $\xi_{\text{FÍSICA}}$ , una componente específica de "funcionamiento" representada por  $\eta_{FF}$ , y una componente de error representada por  $\varepsilon_{FF}$ . Si la media de  $\xi_{\text{FÍSICA}}$  cambia de un tiempo a otro, entonces, consecuentemente, la media de los indicadores también cambia. Sin embargo, si el cambio en la media de  $y_{FF}$  no puede ser totalmente explicado por el cambio en  $\xi_{\text{FÍSICA}}$ , entonces se puede únicamente explicar por un cambio en la media de  $\eta_{FF}$ . Quizás los pacientes hayan aprendido a hacer frente a su enfermedad y por eso su funcionamiento ha mejorado más (o se ha deteriorado menos) que lo que sería de esperar teniendo en cuenta su salud física. Pensando en el ejemplo del dolor,

supongamos que un grupo de pacientes de cáncer, además del tratamiento oncológico, recibe un tratamiento especial para el dolor. La mejoría en la escala de DC puede ser más significativa que la mejoría en los demás indicadores de Salud Física. En el modelo presentado, esto reflejaría un cambio en los interceptos y podría explicarse por un cambio en las medias de los factores específicos.

### *Cambio Alpha*

Un **cambio Alpha** implica un cambio en el nivel del individuo en el constructo *target*. Si el constructo es CVRS, entonces un cambio Alpha positivo se refiere a una mejor salud o a una mejoría de la CVRS. Los factores comunes  $\xi$  pueden ser vistos como operacionalizaciones de los constructos *target*. Por lo tanto, cambios en las medias  $\alpha$  de los factores comunes pueden ser utilizados como operacionalizaciones de cambios verdaderos. Los cambios Alpha pueden ser investigados testando invarianza en el vector  $\alpha$  a lo largo de los tiempos.

### *Otros Tipos de Cambio*

Otros tipos de cambio que no han sido mencionados son los cambios en las varianzas de los factores comunes  $\text{Diag}(\Phi)$ , en las correlaciones de los factores comunes  $\Phi^*$  y en las correlaciones de los factores residuales  $\Omega^*$ .

Un **cambio en las varianzas de los factores comunes** ( $\text{Diag}(\Phi)$ ), significa que el grupo de sujetos se ha vuelto más homogéneo o menos homogéneo con respecto a los factores comunes. Por ejemplo, si la varianza de  $\xi_{\text{FÍSICA}}$  es menor en el segundo tiempo que en el primer, entonces los pacientes se han vuelto más similares relativamente a la Salud Física. Paralelamente a la utilización de "cambios verdaderos en las medias" para el cambio en  $\alpha$ , se puede utilizar la designación "cambios verdaderos en las varianzas" para el cambio en  $\text{Diag}(\Phi)$ .

Un **cambio en las correlaciones entre los factores comunes**, cambio en  $\Phi^*$ , significa que las interrelaciones entre los factores comunes cambiaron. Tal cambio puede interpretarse como una repriorización o reconceptualización en la medición

de uno o más factores de orden elevada. En la Fig.12, los factores comunes  $\xi_{FÍSICA}$  y  $\xi_{MENTAL}$ , pueden ser vistos como indicadores de un factor de orden más elevada representando CVRS. Si las correlaciones entre los factores comunes cambian, entonces cambia la composición del constructo CVRS, indicando repriorización en la evaluación de la CVRS. Si tuviéramos más de dos factores comunes, incluso el número de factores de orden elevada puede cambiar, indicando reconceptualización. Oort (2005) refiere estos tipos de *response shift* como repriorización de "nivel elevado" y reconceptualización.

Finalmente, se existen **correlaciones entre los factores residuales** de las diferentes variables, entonces se puede testar la invarianza a lo largo de los diferentes tiempos. El rechazo de tal hipótesis de invarianza puede interpretarse como una repriorización de los factores residuales implicados, o, si la correlación desaparece, como una reconceptualización. Como los factores residuales se pueden interpretar como factores de orden más baja, Oort (2005) refiere estos tipos de *response shift* como repriorización de "nivel bajo" y reconceptualización.

El procedimiento descrito por Oort (2005) y Oort *et al* (2005) para detectar *response shift* y medir cambios Alpha, que se sintetiza en la Tabla 2.

Paso	Modelo Asociado	Identificación de las Restricciones	Restricciones de Invarianza
1. Establecer un modelo de medida	1. Modelo de Medida	$\alpha_1 = \alpha_2 = 0$ $\text{Diag}(\Phi_{11}) = \text{Diag}(\Phi_{22}) = I$	No
2. Test Global de <i>Response Shift</i>	2. Modelo sin <i>Response Shift</i>	$\alpha_1 = 0$ $\text{Diag}(\Phi_{11}) = I$	$\tau_1 = \tau_2$ ; $\text{Diag}(\Omega_{11}) = \text{Diag}(\Omega_{22})$
3. Detección de <i>Response Shift</i>	3. Modelo de <i>Response Shift</i>	$\alpha_1 = 0$ $\text{Diag}(\Phi_{11}) = I$	Como en el Modelo 2, excepto las restricciones insostenibles dentro de $\tau, \Gamma$ y $\text{Diag}(\Omega)$
4. Evaluación de Cambio Alpha	4. Modelo Final	$\alpha_1 = 0$ $\text{Diag}(\Phi_{11}) = I$	Como en el Modelo 3, más las restricciones sostenibles dentro de $\alpha, \text{Diag}(\Phi), \Phi^*$ y $\Omega^*$

**Tabla 2.** Procedimiento para detectar *Response Shift* y Cambio Alpha (Oort, 2005).

**Nota:** I es matriz identidad adecuada. La identificación puede hacerse de varias formas, sin embargo es conveniente fijar las medias de los factores del primer tiempo en cero y las varianzas de los factores del primer tiempo en uno.



### *Evaluación de Response Shift y Cambios Alpha*

Además de evaluar significancia estadística del *response shift* y cambios Alpha, los tests  $\chi^2$  para las diferencias son informativos sobre la magnitud de los efectos del *response shift* y verdadero cambio en el cambio observado. Se considera  $\mu_2 - \mu_1$  como el modelo para el cambio observado y se asume  $\alpha_1 = 0$  (como en los Modelos 3 y 4). El cambio observado puede ser descompuesto en tres componentes:

$$\mu_2 - \mu_1 = (\tau_2 - \tau_1) + (\Gamma_2 - \Gamma_1)\alpha_2 + \Gamma_1\alpha_2$$

donde la primera componente  $(\tau_2 - \tau_1)$  constituye la contribución de la recalibración, la segunda componente  $(\Gamma_2 - \Gamma_1)\alpha_2$  constituye la contribución de la reconceptualización y repriorización, y la tercera componente  $\Gamma_1\alpha_2$  constituye la contribución del verdadero cambio al cambio observado.

Las estimaciones de los parámetros del Modelo 4 (o del Modelo 3) pueden ser utilizadas para estimar las componentes del cambio observado. El cociente del cambio observado por la desviación estándar estimada, proporciona índices del "tamaño del efecto"  $d$ . Para una variable observada  $j$ , la desviación estándar es dado por  $\sqrt{\hat{\sigma}_{j1,j1} + \hat{\sigma}_{j2,j2} - 2\hat{\sigma}_{j2,j1}}$ , donde las varianzas  $\hat{\sigma}_{j1,j1}$ ,  $\hat{\sigma}_{j2,j2}$  y la covarianza  $\hat{\sigma}_{j2,j1}$  son elementos de la matriz de covarianzas  $\hat{\Sigma}$  implícita en el Modelo 4 (o Modelo 3). Según Oort (2005), valores de "tamaño del efecto"  $d=0.2$ ,  $0.5$  y  $0.8$  son considerados "pequeño", "medio" y "grande", respectivamente.

### *Bondad de Ajuste*

En el primer paso del procedimiento presentado, se testa si el Modelo 1 ajusta los datos de forma satisfactoria. El estadístico  $\chi^2$  para la bondad de ajuste global, es una prueba de ajuste "exacto". Sin embargo, en los Modelos de Ecuaciones Estructurales no se espera encontrar un ajuste exacto y por eso se utiliza una prueba de ajuste "aproximado". El RMSEA, ya referido anteriormente,

permite probar la hipótesis nula de ajuste aproximado (ejemplo:  $RMSEA < 0,05$ ). En este caso, un error Tipo I sería rechazar erróneamente el Modelo 1, mientras un error Tipo 2 sería no rechazar el Modelo 1 si este no es adecuado. Siendo la potencia de un test la probabilidad de rechazar correctamente la hipótesis nula (cuando la hipótesis alternativa es verdadera), la potencia del estadístico RMSEA es la probabilidad de rechazar un Modelo 1 que no ajusta bien los datos (ejemplo:  $RMSEA > 0,08$ ).

Oort (2005) ha estudiado el tamaño muestral mínimo necesario para aplicar el procedimiento descrito al modelo presentado, con dos factores comunes que miden cuatro indicadores cada uno en dos tiempos (totalizando 16 variables observadas y 4 factores comunes). Ha concluido que para una potencia suficiente (80% de probabilidad de rechazar correctamente a un nivel de significación de 5%) para el test de ajuste ( $RMSEA < 0,05$ ) vs. no ajuste ( $RMSEA = 0,08$ ), se necesita un tamaño muestral de 141 para el modelo menos restrictivo (con 90 grados de libertad). Para alcanzar una potencia de 90%, es suficiente un tamaño muestral de 178.

En los pasos 2, 3 y 4 del procedimiento de detección, cambios en los parámetros individuales del modelo son probados con el estadístico  $\chi^2$  para las diferencias. Con este estadístico se prueba la hipótesis nula de ausencia de cambio en uno (o más) parámetros particulares. La potencia de este test se define como la probabilidad de rechazar correctamente la hipótesis de ausencia de cambio en un parámetro particular, dado que el parámetro ha cambiado. En el ejemplo presentado una muestra de tamaño 170 es suficiente para detectar diferencias "moderadas" entre dos cargas factoriales correspondientes a diferentes tiempos con una potencia de 80%. Para alcanzar una potencia de 90% se necesitaría una muestra de 228.

Oort (2005) considera que para detectar repriorización, se necesitaría una muestra de 200 individuos ya que se buscan diferencias entre dos cargas factoriales diferentes de cero. Por otra parte, la reconceptualización y recalibración tienen mayor impacto en las medias y (co)varianzas observadas y por eso pueden ser detectadas con tamaños muestrales más bajos.

### *Cambios Individuales*

El procedimiento presentado ha sido concebido para detectar *response shift* y cambio Alpha en un grupo de individuos y no a nivel individual ya que los Modelos de Ecuaciones Estructurales se centran en las medias y covarianzas, que son estadísticos de grupo. Sin embargo, si muchos sujetos cambian según el mismo patrón, evidenciando el mismo nivel de cambio individual, entonces ese cambio será evidente en los estadísticos de grupo. Por ejemplo, una reconceptualización o repriorización, reflejada a través de un efecto fuerte (o flaco) de la variable latente (ejemplo: Salud Mental) en la variable observada (ejemplo: FF), aparece en la estructura de covarianza (a través de las cargas factoriales) únicamente si un número considerable de individuos experimentan ese tipo de *response shift* en la misma dirección. Si los cambios individuales son en direcciones distintas, entonces pueden aparecer como cambios en las varianzas de los factores residuales, indicando (al nivel del grupo) recalibración no uniforme. Por otra parte, puede haber otras razones que no la recalibración que pueda producir cambios en los interceptos y varianzas residuales.

#### 4.7. EL MODELO DE MAYO *ET AL*

Los modelos referidos anteriormente apenas permiten detectar el cambio experimentado por la mayor parte de los individuos que constituyen un determinado grupo, o sea, al nivel del grupo, y no cambios al nivel individual.

Mayo *et al* (2008) proponen un método para identificar los sujetos que han experimentado *response shift* en un grupo de pacientes sometidos a un estudio longitudinal. El método consiste en comparar longitudinalmente la CVRS resultante de la evaluación del paciente (CVRS observada) y la prevista (en base a los síntomas y funcionamiento del paciente) de forma a identificar los sujetos que potencialmente han experimentado *response shift*. Grandes fluctuaciones en las diferencias entre la CVRS observada y prevista pueden ser utilizadas para identificar la presencia de *response shift* y clasificar las personas en grupos según la dirección y instante del *response shift*. El modelo está basado en el análisis de los residuos resultantes de un modelo de efectos aleatorios que son centrados y utilizados para crear trayectorias. En particular, se utilizan los residuos generados a partir de un modelo de curvas de crecimiento que mide la CVRS observada y prevista. El análisis de los residuos al nivel individual, proporciona información sobre hasta qué punto la percepción de cada paciente (CVRS observada) se aleja de la CVRS prevista.

Estas autoras han llevado a efecto un estudio con supervivientes de AVC, evaluando *response shift* a través del *EuroQoL EQ-5D Visual Analogue Scale* (EQ-5D VAS). La aplicación del *Then-Test* 6 semanas y 24 semanas después del AVC, ha permitido observar que, en media, la reevaluación retrospectiva de su estado de salud era peor que la evaluación original. Esta tendencia para atribuir puntuaciones más bajas sugiere una recalibración y se ha registrado independientemente de que los sujetos hayan mejorado o empeorado.

El conjunto de datos recogido ha sido utilizado para crear un modelo de forma a predecir la CVRS observada (autoevaluación del paciente mediante el EQ-5D VAS) en función de variables sociodemográficas, físicas, mentales y cognitivas (la severidad del AVC, edad, sexo, comorbilidad, etc). Se utilizó un modelo longitudinal con un intercepto aleatorio. Inicialmente se han incluido todos los predictores. Se han testado efectos principales e interacciones con el tiempo. Una interacción significativa podría sugerir una repriorización en el sentido que la asociación entre el predictor y el resultado observado cambió a lo largo del tiempo. Se ha construido un modelo final únicamente con los predictores significativos del EQ-5D VAS.

Para estimar la proporción de variabilidad en el EQ-5D VAS que es explicada por el modelo predictivo se calculó el estadístico  $R^2$  que es dado por el cuadrado de la correlación entre la CVRS observada y la prevista. La comparación de este modelo frente a un modelo que incluya el factor tiempo, permitió obtener una estimación de la proporción de variabilidad potencialmente explicada por el fenómeno de *response shift*.

Los residuos obtenidos a partir del modelo traducen las diferencias entre la salud percibida (observada) y la salud prevista y son utilizados para identificar los pacientes que experimentaron *response shift*. Fluctuaciones largas en los residuos (diferencias entre CVRS observada y prevista) sugieren la ocurrencia de *response shift*. Los residuos referentes a cada paciente son centrados en el respectivo valor medio (residuo del paciente  $i$  en el instante  $j$  menos el valor medio residual del paciente  $i$ ) de forma a remover el impacto de la distancia entre las puntuaciones observadas y previstas. Un paciente cuyos residuos sean consistentes a lo largo del tiempo (no necesariamente cerca de los previstos) produce residuos centrados cercanos de cero y llevaría a considerar que no había ocurrido *response shift*. Por ejemplo, en el caso de un paciente que puntúe de forma consistente a lo largo del tiempo, sin embargo más alto que lo previsto, podrá tener una visión optimista, no evidenciando la ocurrencia de *response shift* ya que los residuos centrados son

consistentemente pequeños. De forma idéntica, personas que puntúan más bajo que lo previsto de forma consistente a lo largo del tiempo, podrán tener una visión pesimista de su CVRS, no habiendo sin embargo evidencia de que haya ocurrido *response shift*. Por otra parte, se considera que ha ocurrido *response shift* cuando los residuos centrados de un paciente presentan un patrón de fluctuación a lo largo del tiempo.

Mayo *et al* (2008) se basan en los trabajos sobre modelación de trayectorias de Jones *et al* (2001) y Nagin & Tremblay (2001), una forma de análisis de clases latentes para datos continuos, para identificar los distintos grupos de trayectorias de los residuos centrados. Se parte del principio que la población es constituida por una mezcla de grupos distintos que pueden definirse en base a sus trayectorias, bajo un determinado nivel de incertidumbre. Las autoras han creado un modelo cuya variable central son los residuos centrados a lo largo del tiempo. Se comparan modelos que incluyen diferentes grupos de trayectorias utilizando los Criterios de Información de Akaike (AIC) y Bayesianos (BIC). Se estima para cada modelo la probabilidad del paciente pertenecer a un grupo dado, se evalúa el ajuste, y se asignan los pacientes a distintos grupos. Considerando otros estudios, se espera encontrar un grupo substancial en los que no ha ocurrido *response shift*, donde se observan residuos centrados consistentes que gráficamente forman una recta que intercepta el origen; y pequeños grupos con desviaciones a este patrón, donde ocurrió *response shift*.

El modelo propuesto constituye una contribución importante para detectar *response shift* al nivel individual. Las autoras recomiendan disponer de un mínimo de cuatro instantes de evaluación. Un estudio profundo del grupo donde se ha detectado *response shift* podría posibilitar la identificación de predictores de *response shift*. Desde el punto de vista clínico sería importante acompañar estos pacientes. Sin embargo, una de las limitaciones del método propuesto es que no

permite identificar el tipo de *response shift* ocurrido (**recalibración**, **reconceptualización** o **repriorización**), sino únicamente subgrupos de la población en los cuales han ocurrido diferentes formas de *response shift*.

**PARTE II**

**APLICACIÓN A DATOS DE CALIDAD  
DE VIDA EN PACIENTES  
ONCOLÓGICOS**



## **CAPÍTULO V**

---

# **CALIDAD DE VIDA EN PACIENTES ONCOLÓGICOS**

---

## 5. CVRS EN PACIENTES ONCOLÓGICOS

Dado el carácter híbrido de este trabajo, para motivar y poner en contexto a los estadísticos estudiosos del tema de la CVRS, dedicamos este capítulo al problema del cáncer y de la calidad de vida en los pacientes oncológicos. Se trata de una introducción breve y muy básica que también nos ha ayudado a nosotros que no somos profesionales de la Salud a ubicar el problema en estudio. Al final incluimos un relato en primera persona de una joven que ha tenido cáncer de mama y ha aceptado en compartir con nosotros su experiencia.

### 5.1. EL PROBLEMA DEL CÁNCER

Desde la época de Hipócrates ya las poblaciones reaccionaban con miedo, angustia y desesperación al diagnóstico de cáncer. Las descripciones relativas a esta enfermedad y formas de tratarla son muy antiguas. La palabra “cáncer” encierra una pesada connotación, simbolizando frecuentemente dolor, sufrimiento y muerte. El progreso de la investigación asociada al desarrollo de técnicas terapéuticas, ha permitido extender la supervivencia de los pacientes oncológicos, despertando una mayor atención para los aspectos relacionados con la CVRS (Figueiredo *et al*, 2006).

“Cáncer” o “tumor maligno” es el término médico más utilizado para denominar un conjunto amplio de enfermedades caracterizadas por un crecimiento anormal y descontrolado de las células y que, en la mayor parte de las veces, forman una masa a que se llama “tumor”. La diferencia entre los tumores benignos y tumores malignos se debe al hecho de que estos últimos pueden invadir y destruir los tejidos y órganos vecinos, extendiéndose a otras partes del cuerpo. A este proceso de diseminación del tumor se da el nombre de “metástasis” (Carvalho, 2007).

El tratamiento del cáncer, depende del tamaño, tipo y localización del tumor, se ha habido diseminación a otras partes del cuerpo (metástasis) y también de la edad y estado de salud del individuo. El tratamiento puede actuar en una área específica (terapéutica local) o en todo el cuerpo (terapéutica sistémica). La cirugía y la radioterapia son **terapéuticas locales** que visan remover o destruir las células del tumor existentes en una parte específica del cuerpo, mientras la quimioterapia, la terapéutica hormonal y la terapéutica biológica (inmunoterapia) son **terapéuticas sistémicas**. La elección del tratamiento se hace caso a caso en base al análisis de los riesgos y beneficios que cada uno implica. Carvalho (2007), presenta una breve síntesis de las terapéuticas actualmente utilizadas:

- 1) La **cirugía** es, para muchos tipos de cáncer, el tratamiento inicial y consiste en la remoción (total o parcial) del tumor. La cirugía permite confirmar el diagnóstico, determinar hasta donde el cáncer ha avanzado y aliviar algunos síntomas que el paciente presente (por ejemplo, obstrucción intestinal).
- 2) La **quimioterapia** consiste en la administración de fármacos, por ingestión o directamente en la corriente sanguínea, que destruyen las células cancerígenas, interfiriendo con los procesos de crecimiento y división de las mismas. Estos fármacos evitan también que el cáncer se disemine para otras partes del cuerpo, sin embargo pueden afectar no solamente las células cancerígenas como también las células sanas. Así pues, al interferir con las células sanas del organismo, la quimioterapia puede llevar al apareamiento de síntomas indeseables como el cansancio, náuseas, vómitos y diarrea. Dependiendo del fármaco utilizado, algunas personas pueden sentir latencia en los brazos y piernas, presentar caída del cabello (alopecia), úlceras en la boca, pérdida de apetito, anemia (por la disminución de los glóbulos rojos

en la sangre), etc. Las infecciones son también más frecuentes ya que se registra también una disminución de los glóbulos blancos. Afortunadamente, en la actualidad, por una parte existen fármacos que producen menos efectos adversos, por otra existen medicamentos eficaces en el alivio de estos efectos. Dependiendo del estadio del cáncer, la quimioterapia puede tener objetivos distintos:

- La **quimioterapia adyuvante** es administrada después de un tratamiento con intuito curativo (por ejemplo, quimioterapia después de la cirugía o radioterapia) y su objetivo es evitar una posible diseminación de micro metástasis mediante la destrucción de cualesquiera células tumorales todavía existentes, eliminando la probabilidad de enfermedad residual;
  - La **quimioterapia neo-adyuvante** es administrada antes de la cirugía o radioterapia y tiene como objetivo disminuir el tamaño del tumor de forma a posibilitar la cirugía, reduciendo el número de células tumorales existentes;
  - La **quimioterapia paliativa** es administrada cuando el tumor se encuentra en una fase avanzada, con metástasis en otros locales que no únicamente el tumor de origen y su objetivo es tratar y aliviar la sintomatología del cáncer, no presentando intuito curativo.
- 3) La **radioterapia** consiste en utilizar radiación para destruir las células cancerígenas. Es considerado un tratamiento local ya que es generalmente aplicado únicamente en la región del cuerpo en la que el tumor se encuentra. La radiación puede ser aplicada a partir del exterior del cuerpo (radioterapia externa), siendo entonces orientada únicamente para el local donde se encuentra el tumor o a través de pequeños implantes que contienen el material radioactivo (radioterapia

interna o braquiterapia) aplicados en la zona del tumor. Tal como la quimioterapia, la radioterapia también puede afectar las células sanas del organismo lo que, consecuentemente, podrá conducir al apareamiento de efectos adversos como el cansancio, hinchazón, irritación de la piel, pérdida de cabello y otros síntomas, dependiendo del local donde es aplicada. Estos efectos desaparecen con el fin del tratamiento. Sin embargo, la radioterapia externa puede tener efectos adversos a largo plazo conforme el local donde se aplica la radiación: infertilidad (se aplicada sobre o próximo de los órganos genitales), problemas cardíacos o pulmonares (se aplicada en el pecho), problemas gastrointestinales (si aplicada en la barriga), problemas de tiroides y neurológicos (si aplicada en el cuello y cabeza), pudiendo también provocar osteoporosis. Estudios demuestran que, personas que recibieron radiación en el pecho presentan mayor riesgo de desarrollar cáncer de mama y pulmón.

- 4) Varios tipos de cáncer del sistema reproductor, como algunos cánceres de mama y próstata, se desarrollan en la presencia de hormonas producidas por el propio organismo. La **terapéutica hormonal** u **hormonoterapia**, consiste en la alteración de la cantidad de estas hormonas en el organismo de forma a disminuir la probabilidad de crecimiento y propagación de estos tumores.
- 5) La **terapéutica biológica** o **inmunoterapia** consiste en la utilización de sustancias que estimulan los mecanismos de defensa del propio organismo (sistema inmunitario) para combatir el cáncer. En la actualidad, se está desarrollando una amplia investigación en esta área, siendo ejemplo el desarrollo de vacunas y anticuerpos monoclonales (sustancias producidas artificialmente en laboratorio que se conectan a

las células cancerígenas, atacándolas y impidiendo su crecimiento), trayendo grandes esperanzas en el futuro de la lucha contra el cáncer. Los efectos secundarios generalmente relacionados con estos tratamientos incluyen síntomas del tipo gripal (fiebre, náuseas, constipación, dolor muscular, etc.).

El cáncer es una de las principales enfermedades y causas de muerte a nivel europeo. Según los datos de la OCDE (2003), después de las enfermedades del sistema circulatorio, el cáncer sigue siendo la mayor causa de muerte en los países de la OCDE. Son diagnosticados en todo el mundo anualmente 14 millones de nuevos casos de cáncer y, 7 millones se mueren con esta enfermedad. El cáncer constituye un grave problema de salud pública, con elevados costes económicos y sociales ya que, a nivel individual, todas las dimensiones de la vivencia del paciente son afectadas.

Entre mujeres, el cáncer de mama es la forma de cáncer más frecuente en todos los países de la Unión Europea, siguiéndose el cáncer colon-rectal y del útero. El cáncer de mama representa en varios países de la OCDE al menos el 30% de los casos de cáncer en las mujeres y es responsable por 15 a 20% de las muertes por cáncer (OCDE, *op.cit.*). Entre los hombres europeos, el cáncer de próstata es el más frecuente, siguiéndose el cáncer del pulmón y colon-rectal.

En los últimos años, el **cáncer de mama** ha recibido especial atención. En realidad, el cáncer de mama presenta una problemática particular porque afecta la identidad de estas mujeres de forma peculiar y distinta de todos los otros tipos de cáncer. Montazeri (2008) subraya que las mujeres juegan un importante papel como compañeras, esposas y madres dentro de cualquier familia. Por lo tanto, cuando una mujer desarrolla cáncer de mama todos los miembros de la familia son seriamente afectados y el cáncer de mama se vuelve una enfermedad de toda la familia. El cáncer de mama es inusual antes de los 30 años y su incidencia

aumenta con la edad. Las mujeres entre los 40 y los 60 años son las más afectadas por este tipo de cáncer, cuándo están precisamente en el pico de su actividad (Shimozuma *et al*, 2007). Aunque, el diagnóstico temprano contribuya para un mejor pronóstico, el 30% de las pacientes experimenta una recidiva (Taguchi *et al*, 2008).

El número de mujeres con cáncer de mama está aumentando. Cada año más de 1,1 millones de mujeres en todo el mundo son diagnosticadas con cáncer de mama. Cifras de 2003 evidencian que 410.000 mujeres se mueren anualmente en el mundo a causa de esta enfermedad (Stewart & Kleihues, 2003), mientras según cifras de 2009 este número ha aumentado para 519.000 (WHO, 2009).

Según el Comité Consultivo para la Prevención del Cáncer (1999), en los países con estadísticas nacionales de la población sobre el cáncer, como los países nórdicos, la incidencia del cáncer de mama ha aumentado durante las últimas cuatro décadas. El comienzo de un programa de cribado se asocia a un incremento temporal en la incidencia de cáncer de mama. Datos posteriores (OCDE, 2003) han revelado que la incidencia de cáncer de mama ha aumentado en todos los países de la OCDE con excepción de Italia. Este fenómeno se explica sobre todo por los progresos médicos realizados al nivel del diagnóstico y por el número creciente de mujeres que se han sometido al cribado por mamografía, permitiendo un aumento significativo de detección de nuevos casos. En Portugal, actualmente con una población femenina de 5,5 millones, surgen todos los años 4500 nuevos casos de cáncer de mama y 1500 muertes por esta enfermedad, o sea, 12 casos por día y se mueren diariamente 4 mujeres víctimas de esta enfermedad (Liga Portuguesa Contra o Cancro, 2009).

Afortunadamente, las tasas de supervivencia mejoraron significativamente y, en los países de la OCDE, varían entre 74% en Inglaterra, 84% en los Estados Unidos y 85% en Japón (OCDE, 2003). La detección y tratamiento precoz son factores fundamentales. Nótese que, en el Reino Unido, la tasa de supervivencia es más baja debido al diagnóstico tardío.

Montazeri (2008) desarrolló un estudio que ha consistido en una exhaustiva revisión bibliográfica de las publicaciones en inglés entre 1974 y 2007 que incluyen en el título los términos “calidad de vida” y “cáncer de mama” en el título. Sintetizando los principales resultados esta pesquisa se destaca que:

- el EORTC QLQ-C30 (*European Organization for Research and Treatment of Cancer Core Cancer Quality of Life Questionnaire*) y el FACIT-G (*Functional Assessment Chronic Illness Therapy General*) con los respectivos módulos específicos para el cáncer de mama (EORTC QLQ-BR23 y FACIT-B) son los instrumentos más utilizados para medir CVRS en contextos de este tipo de cáncer;
- diferentes tipos de cirugía, conducen a resultados relativamente similares en términos de evaluación de la CVRS, aunque pacientes mastectomizadas en general reporten una imagen corporal y funcionamiento sexual más bajos que pacientes sometidas a cirugía conservativa;
- en lo que respecta a las terapias sistémicas, todos los estudios indican que pacientes de cáncer de mama recibiendo quimioterapia experimentan varios efectos colaterales que afectan negativamente su CVRS;
- las terapias hormonales adyuvantes tienen igualmente efectos colaterales negativos en la CVRS, aunque en general aparezcan asociadas a una mejoría de la supervivencia;
- en pacientes con cáncer de mama metastizado, además de los factores clínicos conocidos, la CVRS constituye un pronóstico y predictor del tiempo de vida;
- la ansiedad y depresión son síntomas comunes entre las pacientes de cáncer de mama, incluso varios años después del diagnóstico y tratamiento de la enfermedad;



- los factores psicológicos son también mencionados como predictores de la CVRS subsecuente e incluso de la supervivencia global;
- tratamientos, apoyo psicosocial y ejercicio físico contribuyen de forma positiva para mejorar la CVRS;
- fatiga, morbilidad del brazo y síntomas postmenopáusicos están entre los síntomas más comunes descritos por las pacientes de cáncer de mama;
- el funcionamiento sexual afecta negativamente estas pacientes, especialmente las más jóvenes.

El **cáncer colon-rectal** es la segunda causa de muerte debida al cáncer y la cuarta enfermedad maligna más frecuente en muchos países desarrollados. Este tipo de cáncer afecta hombres y mujeres de forma igual. Todos los años surge más de 1 millón de nuevos casos y ocurren en todo el mundo 639.000 muertes (WHO, 2009). Las tasas de supervivencia han crecido a lo largo de las últimas décadas debido a los diagnósticos tempranos, a la mejoría de los tests de diagnóstico, a la introducción de terapia adyuvante y a los avances en los tratamientos de metástasis. Aproximadamente el 80% de los pacientes supera ahora el primer año después del diagnóstico y aproximadamente el 62% supervive más de 5 años (Arndt *et al*, 2004; Pucciarelli *et al*, 2008). La resección abdominoperineal seguida de la colostomía permanente conduce no solamente a problemas físicos sino también a problemas psicológicos y sociales. Por eso es fundamental la evaluación de la CVRS de estos pacientes (Velenik *et al*, 2008).

Un estudio de la OCDE (1995) ha revelado que, en Portugal, los tumores malignos representaron para Portugal 1318 años perdidos de vida potencial en los hombres y 918 en las mujeres por 100 000 habitantes, siendo la segunda causa de muerte. Por la amplitud y gravedad del problema en términos de salud pública, en 2001, el gobierno portugués decidió adoptar medidas al nivel de la educación y

promoción de la salud, del diagnóstico y tratamiento y de la rehabilitación y cuidados paliativos elaborando el Plano Oncológico Nacional por Resolución del Consejo de Ministros (Resolução do Conselho de Ministros nº129/2001). Con este documento se pretendió definir una política de coordinación, intersección y complementariedad de los servicios de salud, abarcando todas las fases que acompañan la enfermedad, desde la prevención y rastreo al diagnóstico y tratamiento, terminando en la rehabilitación y cuidados paliativos. Este documento instituye que el rastreo al cáncer de mama se basa en la mamografía a mujeres de los 50 a los 69 años con intervalos de dos años, aunque se puedan también desarrollar programas de cribado en el grupo de los 40 a los 49 años. Se refiere también que, cuando dejan de ser necesarios los tratamientos específicos, dirigidos a la enfermedad, las necesidades de los pacientes y su familia continúan a exigir un apoyo humanizado y eficaz. A pesar de los avances en el rastreo y tratamiento de estos pacientes, según las estimaciones, cerca de la mitad de los pacientes oncológicos fallecerán de esta enfermedad.

## 5.2. CVRS EN PACIENTES ONCOLÓGICOS

Varios pacientes relatan que, estar sentado delante de un médico y escuchar la palabra “cáncer” es una experiencia que despierta miedo e incertidumbre. Personas que han vivido esta situación refieren que, después de haber escuchado el diagnóstico, se han vuelto incapaces de asimilar la información que les ha sido transmitida enseguida. La enfermedad oncológica presenta todavía un carácter altamente estigmatizante en la sociedad, sentida particularmente por el paciente y por la familia, con niveles de mortalidad significativos que varían en función de la patología en causa y el estadio en que se encuentra. Este hecho, asociado muchas veces a la incertidumbre del diagnóstico, pronóstico y al sufrimiento físico y psicológico, levantan importantes cuestiones al nivel de la CVRS de los individuos (Figueiredo *et al*, 2006; Sá, 2003).

El suceso de la terapia oncológica se describe muchas veces en términos de supervivencia y complicaciones. Al utilizar únicamente estos parámetros, se está ignorando toda la complejidad que caracteriza la enfermedad oncológica. El paciente posee una percepción mucho más global de todos los eventos relacionados con su enfermedad ya que esos asumen un papel central en su vivencia. El choque producido por el diagnóstico, el dolor, el estrés resultante de las terapias, las restricciones a su desempeño físico e intelectual, las limitaciones en las actividades diarias, la estigmatización social, la vivencia de situaciones que colocan en riesgo la vida o que disminuyen la esperanza de vida, son parámetros que deben ser considerados. Es por eso muy importante que, además de la evaluación clásica, se evalúe la CVRS, la cual constituye un objetivo fundamental a conseguir siempre que se establece una estrategia terapéutica para un paciente oncológico (Shimozuma *et al*, 2007; Pimentel, 2004).

El cáncer difiere en muchos aspectos de otras enfermedades crónicas. Una encuesta realizada en EEUU ha puesto de manifiesto que el problema del cáncer ocupa el primer lugar en la mente de los estadounidenses en lo que respecta a la preocupación con la salud y que casi la mitad cree que es imposible evitar el cáncer (Gotay, 2004). Al cáncer están siempre asociados varios síntomas que hacen que sea diferente de otras enfermedades.

Las personas que viven con cáncer, enfrentan una enorme cantidad de incertidumbres y miedos motivados por las alteraciones en la rutina diaria, en la familia y posibles alteraciones físicas relacionadas con la enfermedad y los efectos secundarios de los tratamientos. Es frecuente que una persona con cáncer sienta rabia, tristeza, miedo, frustración, ansiedad, alteraciones en la autoimagen y sensación de pérdida de control. Además de eso, los tratamientos pueden también producir alteraciones del sueño y del bienestar, lo que, a su vez, puede contribuir para una disminución de la disposición emocional. Personas que han pasado por esta experiencia relatan con frecuencia que gran parte de la ansiedad se prende

con el miedo de la muerte, miedo de la recidiva, miedo de perder el control sobre su vida, miedo del futuro, miedo de quedar desfiguradas debido al cáncer y ser abandonadas por eso. Las alteraciones que ocurren en la apariencia física pueden afectar la imagen que el paciente tiene de sí mismo y hacer con que se sienta inseguro en relación a su cuerpo. Aunque no todas las alteraciones sean visibles, la persona con cáncer puede pensar que los otros lo ven de forma diferente. También la pérdida de la resistencia física, problemas de infertilidad, incontinencia, impotencia, pérdida de una mama o un testículo, pueden llevar la persona a sentirse menos femenina o masculina (Barroso, 2007; Schroevers *et al*, 2004). Así pues, no es difícil comprender que la ansiedad y la depresión sean síntomas comunes en las personas con cáncer que se extiendan a sus familiares. Todo esto tiene importantes consecuencias al nivel de la CVRS de estos pacientes. Sin embargo, este aspecto no siempre es valorizado y la atención casi siempre se centra en la cura del cáncer.

Carlson & Bultz (2003) sintetizan las dimensiones que importa considerar en la evaluación de la CVRS en pacientes oncológicos:

- 1) **Dimensión Física y Biológica** – se destacan todos los aspectos que se relacionan con la enfermedad y sus implicaciones, en particular síntomas, tratamientos y malestar;
- 2) **Dimensión Psicológica** – el impacto de la enfermedad a nivel emocional que se traduce a veces en ansiedad, depresión y miedo de la experiencia con el cáncer;
- 3) **Dimensión Social** – se incluyen las relaciones a nivel familiar, laboral, social y sus interrelaciones con su cultura, costumbres y tradiciones;
- 4) **Dimensión Espiritual** – incluye el bienestar espiritual y significado de la enfermedad, esperanza, trascendencia, incertidumbre, religiosidad y fuerza interior;

- 5) **Estado Funcional** – se refiere a la capacidad de ejecutar actividades normalmente efectuadas por la mayoría de las personas, como el autocuidado, movilización y desempeño de su papel social;
- 6) **Evaluación Global** – incluye una evaluación global, que debe limitarse a las áreas relacionadas con la salud.

### 5.2.1. Los Tratamientos y Síntomas Asociados

Como es bien sabido, las terapias utilizadas en el tratamiento del cáncer poseen efectos asociados. La cirugía remueve permanentemente una parte del cuerpo, mientras la radioterapia puede producir náuseas durante el período de tratamiento y fatiga durante mucho tiempo (Andrykowski *et al*, 2009). Por otra parte, la quimioterapia está asociada a la disminución de las defensas del organismo, vómitos, caída del pelo y una disminución significativa de la CVRS. Todo esto afecta fuertemente el bienestar físico, psicológico y social del individuo con cáncer (Schroevers *et al*, 2004).

Según Figueiredo *et al* (2006), los pacientes oncológicos pasan por dos fases distintas a lo largo de la enfermedad y que, de alguna forma, influyen en su CVRS: la fase de **diagnóstico** y la fase de **tratamiento**. En la fase que antecede el **diagnóstico**, los pacientes ya experimentaron, muchas veces, alteraciones físicas (por ejemplo al nivel del tracto intestinal). El dolor constituye uno de los síntomas más importantes en la evaluación de la CVRS de cualquier paciente oncológico ya que tiende a presentarse como uno de los síntomas más frecuentes, encontrándose en todas las fases de la enfermedad, incluso en la fase de diagnóstico. Su frecuencia depende del tipo de tumor y del estadio de la enfermedad. Este síntoma afecta el 50% de los pacientes oncológicos, sin embargo, su frecuencia

aumenta cuando nos referimos a la fase final de la vida, cuando afecta del 75 al 85% de los pacientes.

En la fase de **tratamiento**, el paciente es frecuentemente confrontado con tratamientos más o menos invasores que pueden afectar su imagen corporal, sexualidad y, consecuentemente, el bienestar físico y psicológico, tratándose de una fase bastante problemática. Couvreur (2001) subraya la importancia de saber si el tumor es local o si existen metástasis, ya que estas diferentes situaciones exigen diferentes tratamientos y tienen, consecuentemente, diferentes repercusiones al nivel de la CVRS. Las consecuencias de una cirugía, de la quimioterapia o radioterapia producen ansiedad y tienen implicaciones al nivel del autoconcepto. En general, el paciente manifiesta mayores dificultades en el primer año después de la cirugía debido a la aceptación de su nueva condición. Estos pacientes demuestran algunas situaciones de ansiedad, depresión y pesimismo debido a la pérdida de una parte de su cuerpo, originando crisis emocionales que se caracterizan por tristeza y rabia. La actividad social y las relaciones interpersonales a veces se deterioran.

En el transcurso de los tratamientos, el paciente oncológico experimenta señales y síntomas que influyen su concepción de imagen corporal: dolor, pérdida de peso, anorexia, fatiga, náuseas, vómitos y malestar. Ocurren alteraciones en la apariencia física, en las funciones corporales, en el control de las actividades diarias y, consecuentemente, en la CVRS. En general, aunque la fatiga y las perturbaciones del sueño, las náuseas, los vómitos y la pérdida de apetito tiendan a aumentar con la evolución de la enfermedad, el dolor parece disminuir debido al efecto de la medicación. En este sentido, los tratamientos tradicionales como cirugía, quimioterapia y radioterapia, utilizados en la enfermedad oncológica, debido a su agresividad a nivel físico producida por los efectos colaterales, contribuyen para una representación todavía más crónica de la enfermedad (Figueiredo *et al*, 2006).

Couvreur (2001) considera que las terapias agresivas son validas, a pesar de su toxicidad, en las situaciones en las que el tratamiento es de carácter curativo. Asimismo, a pesar de la magnitud de consecuencias asociadas al cáncer y sus tratamientos, la CVRS es percibida por gran parte de los pacientes oncológicos de forma positiva. Esta autora cita un estudio realizado en Gran Bretaña que permitió constatar que estos pacientes sufren una disminución significativa en las capacidades de ejercicio de las actividades diarias. Se ha observado una disminución en un 87% de la capacidad para coger objetos pesados, en un 60% para hacer limpiezas, en un 56% para cuidar de los hijos y en un 56% para subir escaleras.

Un número elevado de investigaciones se centra en la evaluación de los problemas asociados al cáncer y su tratamiento. Ganz *et al* (2004) relata que, en pacientes de cáncer de mama que han finalizado los tratamientos primarios, se observa una disminución del Funcionamiento Físico, sobre todo en las mujeres mastectomizadas o recibiendo quimioterapia. Los síntomas más frecuentemente relatados han sido “dolor”, la “dificultad en concentrarse” y “disminución de energía”. Estos autores consideran que deberían ser emprendidas intervenciones clínicas de forma a favorecer el Funcionamiento Físico y Emocional en el final de los tratamientos primarios asociados al cáncer de mama. En otro estudio con mujeres con cáncer de mama, Caffo *et al* (2003) han observado que el dolor afectaba 39,7% de las mujeres dos años después de la cirugía, con mayor incidencia en las que habían sido sometidas a cirugía conservativa. Además, se ha constatado que las mujeres afectadas por el dolor poseen puntuaciones de CVRS significativamente más bajas en las dimensiones de bienestar físico, autonomía física, social y bienestar psicológico, que las demás. En realidad el dolor después de la cirugía afecta cerca de un tercio de las pacientes, independiente del tipo de tratamiento y tiene un efecto negativo en la CVRS de estas pacientes. El dolor puede representar la memoria continua de los tratamientos y de la enfermedad y

puede conducir al miedo de empeorar o de una recidiva. Por eso, pacientes de cáncer de mama que sienten dolor después que hayan terminado los tratamientos, tienden a sufrir angustia psicofisiológica.

Considerado como el síntoma más temido, el dolor es visto como una amenaza al modo de vida y a la propia existencia, constituyendo una causa profunda de perturbación de la CVRS. Frecuentemente el dolor no afecta sólo a un órgano pudiendo extenderse a otras partes del cuerpo o incluso a todo el cuerpo, produciéndose perturbaciones en las actividades, disminución de la concentración, alteraciones psíquicas, estrés al nivel familiar y empobrecimiento de las relaciones sociales. La interpretación del dolor se hace de diferentes formas, sufriendo a la influencia de la personalidad, cultura y medio social.

Dependiendo del diagnóstico oncológico, algunos pacientes experimentan problemas específicos. Por ejemplo, en las pacientes de cáncer de mama es frecuente el espesamiento y dolor en el brazo y problemas al nivel sexual. En un estudio sobre predictores de la CVRS en pacientes con cáncer de mama, Engel *et al* (2003) han encontrado que los problemas producidos en el brazo por la cirugía, la comunicación médico-paciente, la comorbilidad, la edad, el tipo de cirugía y, en menos grado, el estado civil, el nivel educacional y la situación laboral, tienen una asociación significativa con la CVRS. En particular, las pacientes sin problemas en el brazo, las pacientes con las cuales el médico ha establecido una comunicación clara y ha explicado que era lo que les estaba exactamente pasando, las pacientes sin comorbilidad, las pacientes que han sido sometidas a una cirugía conservativa (en oposición a una mastectomía) revelaron niveles de CVRS superiores en todas las dimensiones evaluadas. Considerando la variable edad, pacientes con menos de 50 años perciben una CVRS superior a las demás en lo que respecta al estado funcional, físico y global. Relativamente a las dimensiones emocional y social, son las pacientes con más de 70 años que perciben una CVRS superior. Estos resultados están en conformidad con los encontrados por Janz *et al* (2005). En otro



estudio sobre los efectos a largo plazo del cáncer de mama, Mols *et al* (2005) han encontrado que la condición médica, la cantidad de apoyo social y el actual nivel de ingresos eran fuertes predictores positivos de la calidad de vida. Por otra parte, la quimioterapia adyuvante ha surgido como un factor predictivo negativo.

Varios estudios demuestran que la fatiga es uno de los problemas que más afecta los pacientes oncológicos (Carlson *et al*, 2004; Aranda *et al*, 2005; Michielsen *et al*, 2007; Andrykowski *et al*, 2009; Hagelin *et al*, 2009; Santana *et al*, 2009). En un estudio de seguimiento de personas con cáncer, Carlson *et al* (*op.cit.*) ha encontrado que la mitad de ellos han identificado la fatiga como un problema. En un grupo de mujeres con cáncer de mama, Aranda *et al* (*op.cit.*) han concluido que la fatiga es el problema más relatado por estas mujeres. Además, entre un cuarto y un tercio de las mujeres manifestaron dificultades con su estado físico, el papel y el funcionamiento social, y un poco más de una cuarta parte de las mujeres evaluó como pobre su estado de salud global. Aranda *et al* (*op.cit.*) evaluaron la CVRS en estas mujeres un año y tres años después del diagnóstico, comparando además los resultados obtenidos con valores de referencia de la población general. Un año después del diagnóstico, los déficits en la CVRS eran todavía evidentes en el dominio emocional, cognitivo, funcionamiento social, síntomas de insomnio, fatiga, disnea y dificultades financieras. Las diferencias entre las pacientes con cáncer de mama y las mujeres de la población general han sido más evidentes en las mujeres más jóvenes. Tres años después del diagnóstico, las diferencias observadas han sido sobre todo a nivel funcional. Por otra parte, la recurrencia de cáncer, tuvo un efecto prejudicial sobre la CVRS durante el periodo de seguimiento.

### 5.2.2. La Adaptación al Cáncer

Barroso (2007) considera que existe una idea errada entre cáncer y depresión y que no todas las personas con cáncer se quedan deprimidas. Sentir tristeza, frustración o incluso rabia es natural en estas situaciones. La depresión es un disturbio biológico que produce cansancio, que a su vez aumenta la fatiga producida por los tratamientos. Esto tiene un impacto en la CVRS del paciente, en su capacidad para tomar decisiones y deteriora la energía necesaria para continuar el tratamiento. La persona con depresión, puede sentir tristeza profunda, irritabilidad, falta de esperanza, alteraciones del sueño y apetito, dificultades de concentración, desinterés por las actividades que anteriormente apreciaba, desarrollar pensamientos negativos como el suicidio y presentar alteraciones de comportamiento como aislamiento, crisis de lloro y desinterés sexual. Algunas personas con formas moderadas de depresión, desconociendo que están deprimidas, se culpan porque “no están luchando lo suficiente” y se vuelven reacios a pedir ayuda. Asimismo, no permanecen con este estado de ánimo por más de algunas semanas y responden bien al apoyo de la familia y amigos. Cuando la depresión es más severa, se prolonga por más tiempo y típicamente no responde al aliento de la familia y amigos que se pueden sentir frustrados con el aislamiento y sentir dificultad en comprenderlo. Sin embargo, algunas personas con cáncer también relatan alteraciones positivas en la forma como pasan a encarar la vida y el futuro, experimentando sentimientos de admiración por la capacidad de resistencia de su organismo, apreciación renovada de la vida con alteración de las prioridades, menos preocupación con cuestiones de peso o belleza corporal, paz o gratitud.

Según Santos *et al* (1994), la patología oncológica genera reacciones de adaptación mental y alteraciones al nivel psicológico las cuales es fundamental detectar. Mediante su conocimiento estas pueden ser evitables o tratables, teniendo como objetivo la terapéutica global de la persona con cáncer y no

solamente el cáncer de que el individuo padece. Sá (2003) define el proceso de adaptación mental al cáncer como la respuesta cognitiva, de comportamiento y emocional del individuo delante el diagnóstico de cáncer, determinada por la forma como el paciente percibe las implicaciones del cáncer, cómo reacciona, lo que piensa y lo que hace para reducir la amenaza transmitida por la enfermedad. El abordaje de los problemas físicos, psicológicos, sociales y familiares del paciente oncológico ha constituido, en los últimos años, foco de interés debido a la repercusión en las múltiples dimensiones de CVRS.

Varios estudios confirman el fenómeno de adaptación psicológica a la enfermedad experimentado por muchos pacientes de cáncer. Un estudio con pacientes de cáncer de mama, desarrollado por Ganz *et al* (2003), ha revelado que, un año después de la cirugía, las puntuaciones relativas a la salud física y mental presentan una disminución significativa, independiente de la edad. Sin embargo, estos autores han observado que las puntuaciones asociadas a los ítems relativos a la CVRS psicosocial específicos de cáncer, mejoraron a lo largo del tiempo, reflejando adaptación y ajuste a las especificidades del cáncer (como por ejemplo sentirse incómoda con las alteraciones del cuerpo, con miedo de una recidiva o nerviosa). Por otra parte, mujeres mayores que han empeorado al nivel físico, mental y emocional después de la cirugía perciben como más pobre su salud y su ajuste psicosocial un año después de la cirugía. La edad aisladamente no es un predictor significativo de la salud o del ajuste psicosocial.

Cómo hemos mencionado, en contextos de CVRS, el fenómeno de *response shift* ocurre frecuentemente y se define como un cambio en los estándares internos o en los valores, o una reconceptualización de la calidad de vida producido por un cambio en el estado de salud. En el contexto del tratamiento del cáncer, el *response shift* es un mecanismo normal psicológico de adaptación. A la medida que los pacientes se adaptan a los efectos colaterales del tratamiento, ocurren cambios en los estándares internos haciendo que la percepción de la severidad de los efectos colaterales del tratamiento sea atenuada. Donovan *et al* (2004) han encontrado que,

en pacientes de cáncer recibiendo tratamientos múltiples (como por ejemplo quimioterapia seguida de radioterapia) la ocurrencia de *response shift* durante el primer tratamiento, influencia la percepción de fatiga durante el segundo tratamiento. En particular, cuando comparadas con mujeres que no habían experimentado la fatiga resultante de la quimioterapia, las que habían hecho quimioterapia, perciben la radioterapia como menos inductora de fatiga. Jong *et al* (2004) mencionan resultados semejantes y afirman que este efecto es todavía más pronunciado en las mujeres mastectomizadas que en las demás.

Al evaluar *response shift* en la percepción de fatiga en pacientes de cáncer de mama en radioterapia a través del *Then-Test*, varios autores han obtenido puntuaciones significativamente más bajas en el *Then-Test* que en el *Pre-Test*. La experiencia de la quimioterapia adyuvante ha llevado a las pacientes a evaluar retrospectivamente más bajo la percepción que habían tenido de los niveles de fatiga experimentados en el inicio del tratamiento, indicando que estas pacientes se habían adaptado a un nivel más elevado de fatiga (Sprangers *et al*, 1999; Jansen *et al*, 2000; Visser *et al*, 2000; Westerman *et al*, 2007a; Westerman *et al*, 2007b; Andrykowski *et al*, 2009; Hagelin *et al*, 2009).

En un estudio cuyo objetivo era evaluar la CVRS en pacientes de cáncer haciendo quimioterapia, Bernhard *et al* (2004) han observado que la mejoría de la CVRS percibida durante los primeros 6 meses de tratamiento ha sido más pronunciada que el deterioro transitorio producido por los tratamientos. Estos resultados sugieren que la adaptación psicológica de los pacientes es más importante en la CVRS percibida que los efectos de la quimioterapia.

En un grupo de mujeres con cáncer de mama, De Bock *et al* (2004) no han encontrado diferencias significativas entre la CVRS percibida por pacientes mastectomizadas y las que habían sido sometidas a cirugía conservativa, en lo que respecta a la dimensión de funcionamiento psicológico. Estos autores han encontrado también niveles significativamente más elevados de ansiedad entre las

mujeres con diagnósticos recientes de cáncer que entre las que habían recibido el diagnóstico hace más de 12 meses.

En suma, la enfermedad oncológica es considerada como un agente de estrés, lo que implica la adopción de determinadas estrategias que tienen como objetivo mejorar la CVRS percibida por el paciente. El paciente se presenta como un agente activo en este proceso, desarrollando estrategias de adaptación. La percepción social relativamente al cáncer, los efectos directos de esta enfermedad y sus tratamientos, justifican la importancia y utilidad de recurrir a estrategias que permitan evaluar los efectos específicos del cáncer en la CVRS de los pacientes (Gotay, 2004).

### **5.2.3. La Comunicación Médico-Paciente**

Varios autores subrayan la importancia de la comunicación médico-paciente refiriendo la existencia de una relación positiva entre una comunicación clara y abierta y la CVRS (Engel *et al*, 2003; Kerr *et al*, 2003; Ganz *et al*, 2003; Gestel *et al*, 2007). Estudios demuestran que personas bien informadas sobre su enfermedad y opciones de tratamiento, tienden a lidiar mejor con la situación y sienten menos los efectos adversos que aquellas que apenas siguen las órdenes del médico (Barroso, 2007).

La información contribuye para que las expectativas del paciente sean realistas. Estos autores sugieren que los médicos deberían comunicarse mejor con sus pacientes y hacer un esfuerzo en el sentido de atenuar su ansiedad y angustia. Por otra parte, Velikova *et al* (2004) ha concluido, a partir de un estudio desarrollado con 286 pacientes de cáncer y 28 oncólogos, que la evaluación de la CVRS del paciente tiene un impacto positivo en la comunicación paciente-médico resultando un beneficio para los pacientes y contribuyendo positivamente para su bienestar emocional. Algunos defienden que la habilidad de comunicar es un

talento natural. Sin embargo, según estos autores, determinadas habilidades como ser un oyente activo, abierto, demostrar empatía, sintetizar la información y llevar el paciente a entenderla, son habilidades que pueden ser aprendidas y desarrolladas.

Otro estudio con pacientes de cáncer de mama, desarrollado por Kerr *et al* (2003), permitió concluir que, 5 años después del diagnóstico, casi la mitad (45%) de estas mujeres consideraban insuficiente la comunicación que habían recibido y más de la mitad (59%) quería hablar más con el médico. Las pacientes con menos de 50 años han considerado la ayuda social y psicológica como más importante. Aunque las pacientes de este grupo etario se hayan mostrado más informadas de estos servicios y hayan tenido más contacto con grupos de apoyo, se mostraban más insatisfechas con la información recibida. Las pacientes más jóvenes tienen más necesidad de información.

Kerr *et al* (*op.cit.*) subrayan, sin embargo que, el hecho de que el paciente considere incomprendible o incompleta la información transmitida, no significa necesariamente que el médico haya fallado en la transmisión de esta información. El choque producido por un diagnóstico de cáncer puede afectar el procesamiento de la información, haciendo que el paciente no registre la información dada. Las personas aceptan la enfermedad de diferentes formas y eso hace que el momento apropiado para transmitir la información no siempre sea el mismo. Sin embargo, los resultados de investigaciones confirman la falta de información por parte de los médicos. Por otra parte, en general, los médicos admiten que sus habilidades de comunicación son insuficientes. Kerr *et al* (*op.cit.*) han encontrado que pacientes sometidos a una pobre comunicación tienden a percibir peor CVRS. Según estos autores, no es la comunicación pobre que produce baja CVRS sino que los pacientes más angustiados tienden a presentar más problemas de comunicación. Por una parte, la ansiedad inhibe el procesamiento de la información, por otra, el vacío de información produce angustia.

#### 5.2.4. El Apoyo Psicosocial

Varios estudios efectuados en distintos países, subrayan que más de la cuarta parte de las personas con enfermedad oncológica sufren de problemas psicosociales durante un largo periodo o incluso permanentemente (Yates *et al*, 2005; Carlson & Bultz, 2003; Marchioro *et al*, 1996; Bordelau *et al*, 2003; Kerr *et al*, 2003; Northouse *et al*, 2007). Como hemos mencionado, la aparición de una enfermedad crónica significa el inicio de un proceso de adaptación a cambios físicos, psicológicos, sociales y ambientales. Es por eso importante entender cómo las personas enfrentan este proceso de adaptación. Bishop (2005) y Wiles (2008) subrayan que los cambios funcionales, psicológicos y sociales asociados a la aparición de una enfermedad crónica son experimentados en todas las dimensiones de la vida del individuo. Por otra parte, existe una gran variación interindividual en lo que respecta a la respuesta a estos cambios, que parece no estar relacionada con el tipo de enfermedad.

La intervención psicosocial puede volverse una ayuda importante para los pacientes y sus familias durante este periodo difícil (Wiles, *op.cit.*; Pinson *et al*, 2009; Reavley *et al*, 2009). Carlson & Bultz (2003) estiman que aproximadamente un 20% de los pacientes de cáncer necesitan de apoyo psiquiátrico debido a la depresión y ansiedad en algún momento del curso de la enfermedad, un 15% necesitan de un psicólogo para tratar la angustia y un 25% de los servicios de asistencia social para lidiar con cuestiones financieras.

Marchioro *et al* (*op.cit.*) han llevado a efecto un estudio con el objetivo de evaluar los efectos de la intervención psicosocial en un grupo de mujeres con cáncer de mama. Este estudio ha permitido encontrar una mejoría significativa al nivel de la CVRS percibida.

Posteriormente, Bordelau *et al* (*op.cit.*) han desarrollado un estudio similar con pacientes con cáncer de mama metastizado. La CVRS ha sido evaluada en base al *European Organization for Research and Treatment of Cancer Quality of Life*

*Questionnaire C30* (EORTC QLQ-C30). En particular, el grupo experimental ha sido sometido a un trabajo terapéutico en el sentido de restablecer la conexión con el sentimiento de sí mismo y facilitar la expresión verbal de la enfermedad en base a la terapia de expresión apoyada (*supportive expressive therapy*). Aunque otros estudios hayan permitido detectar los efectos positivos en la CVRS en pacientes oncológicos sometidos a intervención psicosocial, Bordelau *et al* (2003) no han encontrado diferencias estadísticamente significativas en la CVRS percibida por los dos grupos. Al revés, se ha detectado un deterioro significativo a lo largo del tiempo tanto al nivel global, como al nivel físico, social, funcionamiento cognitivo y escalas de síntomas (disnea, pérdida de apetito y fatiga). Los autores concluyen que los beneficios de la terapia psicosocial no son significativos en pacientes con cáncer de mama metastizado.

Northouse *et al* (2007) han desarrollado un estudio sobre el estado psicosocial y calidad de vida de pacientes con cáncer de próstata y sus parejas, habiendo concluido que se necesitan programas de apoyo psicosocial de forma a ayudar estos pacientes y sus parejas a lidiar con las consecuencias de la enfermedad

### **5.2.5. Un Relato en Primera Persona**

Exactamente por el significado y profundidad de la experiencia que es escuchar un diagnóstico de cáncer, cruzar el largo proceso de tratamiento, luchar contra esta enfermedad y los efectos de su tratamiento, hemos entrevistado una joven a quien, con 35 años ha sido diagnosticado cáncer de mama y que ha aceptado en colaborar en este estudio. El relato que transcribimos a continuación evidencia bien la violencia de recibir el diagnóstico, los sentimientos de miedo e incertidumbre y la adaptación.

Desde aquí le damos las gracias por dejarnos participar en su experiencia.



### Porquê justo a mim...

Na realidade é o que me pergunto todos os dias, porquê que eu tenho que passar por esta etapa tão difícil, porquê que o meu percurso terreno não é linear, ou melhor, menos sinuoso, porquê que eu tenho que ter um cancro de mama?...

Receber um diagnóstico de cancro aos 35 anos não me deu uma perspectiva futura muito promissora e, sem qualquer dó ou piedade castrou-me de qualquer sonho ou fantasia que naturalmente alimentei ao longo da minha vida! No momento parece que o meu corpo e mente explodiram literalmente, eu senti-me desfeita em todo o meu ser, como pessoa e como mulher! Em segundos ou minutos que me pareceram intermináveis deixei de pensar, senti-me envolta numa camisa-de-forças física e emocional!



A primeira reacção é sem dúvida uma sensação de pânico e medo da morte, foi incontável a sequência de pensamentos negativos que vieram à minha mente, isto tudo porque, todos nós de certa forma ainda crescemos numa era em que o cancro significava a morte, o que não deixa de infelizmente continuar a ser uma realidade.

Após a consciencialização da doença e do seu estado evolutivo, através de inúmeras pesquisas na Internet, leituras de entrevistas e reportagens de medicina e muitas questões colocadas aos técnicos de saúde e a outras pessoas com experiências semelhantes, fui me acalmando um pouco, mas nunca deixando de pensar, pensar, pensar...até cansar a alma!

A revolta é outro sentimento que me assolou o coração, senti-me muito triste e revoltada com tudo...comigo própria, com os amigos, com familiares, com o mundo e com Deus! Simplesmente para mim, todos tinham culpa por eu estar doente, até os meus pais eu culpava por me terem concebido, parece loucura de facto! Esta fase é muito longa e dolorosa e talvez em mim perdurou mais tempo porque para além de lutar pela vida tive também que o fazer sem o apoio do meu companheiro que simplesmente não teve a capacidade de se manter ao meu lado e preferiu me “abandonar” neste percurso.

A isto só posso qualificar como crueldade desumana de alguém que pensa que estas doenças só aparecem aos outros...e que nunca irão precisar de uma palavra reconfortante e de força, assim como, um simples abraço de apoio. Para ser sincera, precisava muito que ele me dissesse que me continuava a amar e que eu como mulher em nada tinha desvalorizado para ele...isso seria muito importante para mim, acreditem!!!

Qualquer cancro é devastador sem dúvida, mas para mim como mulher, o facto de ser na mama, mexeu com o meu ego, com a minha auto-estima, com a minha feminilidade, com a maternidade e sexualidade. Deu-me a sensação de que não voltaria a ser uma mulher bonita e desejada e que a partir daquele momento tinha que esconder algo que até ali eu tinha orgulho em discretamente evidenciar num pequeno decote ou vestido, o que faz parte do ser mulher!

As cirurgias, a recuperação do após-cirurgia, os tratamentos são devastadores quer física quer emocionalmente, não consigo descrever, é muito profundo e íntimo... Para dizer a verdade, acho que nunca consegui ver a quimioterapia como aliada à minha cura, mesmo com a ajuda da minha psicóloga que tão sabiamente me tentou mostrar isso. A quimioterapia é de facto um bicho papão que veio fazer sentir mal, com dores, enjoos, vômitos e fazer-me ficar sem cabelo. Esta é então de certeza a parte que mais me custou a encarar pessoalmente e socialmente...felizmente consegui sempre ultrapassar os preconceitos e a vergonha, arranjando mecanismos estéticos de disfarce, também porque fui sempre muito bem acompanhada pela unidade da psicologia que sempre me ajudou a elevar a minha auto-estima e formas de aceitar este processo todo.

Finalmente, algum tempo depois de esta luta se tornar uma rotina do meu dia a dia, veio a aceitação quando senti a necessidade de me agarrar à fé e voltar-me para Deus, comecei a frequentar a igreja e receber o apoio das irmãs do convento do Carmelo no Patacão que são uma força espiritual fantástica e incansáveis nas suas orações. É nesta fase que se dá a minha reconciliação com Deus e descubro um amor verdadeiro e incondicional...em oração falei com Deus, expus as minhas necessidades, os meus medos e todas as esperanças que me alimentam a vontade de viver.

Na oração apresentei a minha própria história nas suas diferentes etapas, embora sem receber uma resposta imediata e visível aos meus olhos físicos! Para mim, na oração aprendi a desenvolver, não só uma estratégia ou um remédio para eliminar o sofrimento mas especialmente um recurso para enfrentá-lo, embora actualmente se acredite que a fé pode curar, o que eu acredito plenamente, pois por experiência própria sinto que tem sido a minha maior arma de força, por vezes até fico surpreendida comigo própria e pergunto-me donde vem tamanha coragem, o que me dá um certo orgulho de mim mesma...

A oração é um caminho para me ajudar a amadurecer e expandir as minhas perspectivas pessoais para assim me aproximar de Deus e finalmente reconquistar a paz interior que é tão importante para o meu equilíbrio em todos os níveis.

Hoje, presentemente, ainda continuo na luta, após quimioterapias e radioterapias e após avaliação médica foi-me recomendado iniciar mais seis ciclos de quimioterapia a título preventivo por inúmeros factores (tipo de cancro, metástase num gânglio, idade, intolerância aos tratamentos anteriores, hereditariedade...) apesar do sofrimento e de ansiosamente querer voltar à minha vida normal tive que me consciencializar que seria a atitude mais segura para garantir um futuro mais tranquilo.

Aquí está a tal força que vos falo!...

Quando estou nos meus momentos de reflexão e interiorização, admito que este cancro veio por alguma razão especial e para me mostrar algo de novo.

Pergunto-me o que será?

Mas no fundo sei que as respostas estão dentro de mim e naturalmente vão brotando, em forma de emoções, gestos e atitudes do quotidiano para comigo, pessoas e mundo. É surpreendente a mutação de padrões que transportava do passado e que os tinha como os mais correctos mas que agora naturalmente deixam até de fazer sentido.

Este percurso é de facto uma evolução constante onde se aprende e reaprende, aquí faz todo o sentido dizer que a vida é uma escola e neste momento sinto-me na universidade a fazer um mestrado...Lol

Aprendi a ver, pensar, ouvir e sentir tudo de forma diferente! É muito engraçado aperceber-me que um pequeno passeio na praia emociona-me de tal forma que chego a chorar de alegria por me sentir uma felizarda por naquele momento poder ver a força das ondas, ouvir as gaivotas, sentir a brisa do mar e claro absorver a beleza dum pôr-do-sol...

Aprendi que não tinha que tolerar na minha vida mais situações desagradáveis, por muito pequenas que sejam, consigo convictamente dizer que não ao que realmente não me apetece fazer ou ouvir, de certa forma tornei-me mais intolerante a coisas fúteis e menos boas, simplesmente faço o que me dá mesmo prazer, acompanhada ou sozinha...

Uma boa experiência de vida e que me trouxe um grande ensinamento e que me provou de que sou realmente uma grande mulher, foi no espaço de recuperação entre as quimioterapias e as radioterapias, decidi fazer uma viagem de 10 dias sozinha fora do país a um destino que há muito ambicionava ir. Foi de facto a melhor terapia individual e íntima que alguma vez fiz, voltei renovada como mulher, com muitas “gavetinhas” arrumadas e decisões importantes tomadas...

Aprendi que afinal uma mulher com cancro de mama, continua a ser “mulher” no seu todo e mais valorizada e admirada por aqueles que nos rodeiam. Com isto quero dizer, e não podia deixar de partilhar, que afinal havia alguém muito especial destinado para me aceitar nesta condição e me amar verdadeiramente tal como sou. Agora já começo a perceber porquê que me tinha que cair o meu longo cabelo cheio de lindos caracóis...é que o meu príncipe adora cabelos curtos e foi precisamente isso que despertou a sua atenção! Hoje vivo uma linda história de amor...

Só neste percurso que dura há já 1 ano, pude conhecer e sentir essências de pessoas com atitudes totalmente distintas para comigo. Apesar de em parte ser doloroso, por outro lado é gratificante ter a oportunidade de partilhar e usufruir o amor e amizade de outras pessoas que eu própria não esperava! É tudo uma questão de ver o lado bonito e positivo das coisas...

Podia escrever horas intermináveis a testemunhar as minhas experiências, pois todos os dias vão surgindo momentos dignos de serem registados, o que vou fazendo no meu diário, para mais tarde relê-los e constatar que há sempre uma evolução positiva e que tudo se resolve. Aconselho vivamente a que todos o façam, mesmo sem cancros!!! É uma excelente terapia que eu aprendi...

Termino com a necessidade de partilhar com todas as pessoas, um pouco daquilo que me tornei como pessoa e a forma de ser e estar perante a minha passagem por cá.

A morte é de facto algo que luto com todas as minhas forças para adiar o mais possível e que tenho grande pavor, mas que tenho vindo a aceitar como condição que me será imposta por algo superior e que eu só tenho que respeitar...eu assinei esse acordo! Se Deus me permitir, aproveitarei todo o tempo do mundo para dar continuidade ao que me propus, como ser, esposa, amiga e especialmente como mãe, sem deixar de ser um “espírito livre”...respeitando sempre o espaço individual e íntimo de cada um, os desejos e os sonhos de cada um, a essência e os segredos de cada um. Assim sinto que estou a amar com toda a pureza e verdade!

Diariamente exercito, a verdade, a cumplicidade, a espiritualidade e especialmente o respeito para com todos, e abomino o orgulho, o rancor, a raiva de forma a não ser atingida por maus fluidos que me possam alterar a minha paz tão preciosa!

Com tudo isto, deu-se um pleno despertar interior da minha capacidade de recriar os meus dias e de crescer num poder consciente, que é o “amor incondicional por mim”. Deu-se um pleno despertar que apela a todos os meus cantos e recantos da alma como mulher, feita de amor e força, sensibilidade e firmeza, convicção e propósito.

Sinto-me uma vencedora e de alguma forma inteligente, pois tal como diz Fernando Pessoa, “...pedras no meu caminho? Apanharei e guardarei todas para construir um castelo...” Acreditem, que no meu caminho tenho apanhado autênticos pedregulhos, tal como este cancro, mas o lado positivo disto é que vislumbro que o meu castelo será uma fortaleza com alicerces bem fortes...

Assim, faço questão de continuar a perguntar-me **porquê justo a mim...**e sei que todos os dias terei respostas cada vez mais positivas!!!

## **CAPÍTULO VI**

---

# **INSTRUMENTOS, MÉTODO Y MUESTRA**

---

## 6. INSTRUMENTOS, MÉTODO Y MUESTRA

Con el objetivo de evaluar CVRS en un grupo de pacientes oncológicos y, en particular, identificar la ocurrencia de *response shift* entre la fase de diagnóstico y tratamiento de la enfermedad, se ha desarrollado un estudio. En este apartado se presentan algunas consideraciones sobre los instrumentos utilizados (en particular sobre la forma cómo están organizados los ítems y su puntuación), así como la metodología seguida y una breve descripción de la muestra en estudio.

### 6.1. INSTRUMENTOS UTILIZADOS

Para la obtención de los datos en el estudio desarrollado se han utilizado dos instrumentos de medida - el *Medical Outcomes Study 36 – Item Short Form* (SF-36) y el *Rotterdam Symptom Check List* (RSCL) anteriormente descritos (apartados 2.4.1.1 y 2.4.1.2, respectivamente). Como hemos mencionado, al trabajar con estos dos instrumentos, hemos intentado reunir las ventajas inherentes a la utilización de un instrumento genérico y de un instrumento específico. Siendo un instrumento genérico, el SF-36 ha posibilitado la obtención de resultados que pueden ser comparados con los obtenidos en pacientes con otros diagnósticos o incluso con personas sanas, mientras el RSCL, está específicamente concebido para diagnósticos de cáncer, incidiendo sobre los efectos de la enfermedad y su tratamiento.

Además de los instrumentos para evaluar CVRS referidos, se ha aplicado también un cuestionario con el cual se pretendía recoger alguna información sociodemográfica relativa a los pacientes en estudio (Anexo V).

### 6.1.1. Medical Outcomes Study 36 – Item Short Form (SF-36)

El *Medical Outcomes Study 36 – Item Short Form (SF-36)* es una medida genérica para evaluar CVRS (Ware *et al*, 1993) y está compuesto por 36 ítems que presentamos en la Tabla 3.

1. En general, usted diría que su salud es (excelente; muy buena; buena; regular; mala)
2. ¿Cómo diría que es su salud, <b>comparada con la de hace un año?</b> (mucho mejor; algo mejor; más o menos igual; algo peor; mucho peor)
<b>Su salud actual</b> , ¿le limita para hacer esas actividades o cosas? (sí, me limita mucho; sí me limita un poco; no, no me limita nada)
3a. <b>Esfuerzos intensos</b> , tales como correr, levantar objetos pesados, o participar en deportes agotadores
3b. <b>Esfuerzos moderados</b> , como mover una mesa, pasar la aspiradora, o caminar más de una hora
3c. Coger o llevar la bolsa de la compra
3d. Subir <b>varios</b> pisos por la escalera
3e. Subir <b>un solo</b> piso por la escalera
3f. Agacharse o arrodillarse
3g. Caminar <b>un kilómetro o más</b>
3h. Caminar <b>varias manzanas</b> (varios centenares de metros)
3i. Caminar <b>una sola manzana</b> (unos 100 metros)
3j. Bañarse o vestirse por si mismo
Durante <b>las 4 últimas semanas</b> , ¿ha tenido alguno de los siguientes problemas en su trabajo o en sus actividades cotidianas, <b>a causa de su salud física?</b> (siempre; la mayor parte del tiempo; algún tiempo; poco tiempo; nunca)
4a. ¿Tuvo que reducir el tiempo dedicado al trabajo o a sus actividades cotidianas?
4b. ¿Hizo menos de lo que hubiera querido hacer?
4c. ¿Tuvo que dejar de algunas tareas en su trabajo o en sus actividades cotidianas?
4d. ¿Tuvo dificultad para hacer su trabajo o sus actividades cotidianas (por ejemplo, le costó más de lo normal)?
Durante <b>las últimas 4 semanas</b> , ¿ha tenido alguno de los siguientes problemas en su trabajo o en sus actividades cotidianas, <b>a causa de algún problema emocional</b> (como estar triste, deprimido, o nervioso)? (siempre; la mayor parte del tiempo; algún tiempo; poco tiempo; nunca)
5a. ¿Tuvo que <b>reducir el tiempo</b> dedicado al trabajo o a sus actividades cotidianas, <b>por algún problema emocional?</b>
5b. ¿Hizo <b>menos</b> de lo que hubiera querido hacer, <b>por algún problema emocional?</b>
5c. ¿No hizo su trabajo o sus actividades cotidianas tan <b>cuidadosamente</b> como de costumbre, <b>por algún problema emocional?</b>
6. Durante <b>las 4 últimas semanas</b> , ¿hasta qué punto su salud física o los problemas emocionales han dificultado sus actividades sociales habituales con la familia, los amigos, los vecinos u otras personas? (nada, un poco; regular; bastante: mucho)
7. ¿Tuvo dolor en alguna parte del cuerpo durante <b>las 4 últimas semanas?</b> (no, ninguno; sí, muy poco; sí, un poco; sí, moderado; sí, mucho; sí, muchísimo)
8. Durante <b>las 4 últimas semanas</b> , ¿hasta qué punto el dolor le ha dificultado su trabajo

habitual (incluido el trabajo fuera de casa y las tareas domésticas)? (nada, un poco; regular; bastante: mucho)
Las preguntas que siguen se refieren a cómo se ha sentido y cómo le han ido las cosas durante <b>las 4 últimas semanas</b> . En cada pregunta responda lo que le parezca más a cómo se ha sentido usted. Durante <b>las últimas 4 semanas</b> , ¿cuánto tiempo... (siempre; casi siempre; muchas veces; algunas veces; sólo alguna vez; nunca)
9a. ¿Se sintió lleno de vitalidad?
9b. ¿Estuvo muy nervioso?
9c. ¿Se sintió tan bajo de moral que nadie podía animarle?
9d. ¿Se sintió calmado y tranquilo?
9e. ¿Tuvo mucha energía?
9f. ¿Se sintió desanimado y triste?
9g. ¿Se sintió agotado?
9h. ¿Se sintió feliz?
9i. ¿Se sintió cansado?
10. Durante <b>las 4 últimas semanas</b> , ¿con qué frecuencia la salud física o los problemas emocionales le han dificultado sus actividades sociales (como visitar a los amigos o familiares)? (siempre; casi siempre; muchas veces; algunas veces: sólo alguna vez; nunca)
Por favor, diga si le parece cierta o falsa <b>cada una</b> de las siguientes frases: (totalmente cierta; cierta; no lo sé; falsa; totalmente falsa)
11a. Creo que me pongo enfermo más fácilmente que otras personas
11b. Estoy tan sano como cualquiera
11c. Creo que mi salud va a mejorar
11d. Mi salud es excelente

Tabla 3. Ítems del SF-36.

Estos ítems se agrupan en 8 subescalas. En la Tabla 4 se presentan los ítems asociados a cada una de las subescalas y las respectivas puntuaciones mínimas y máximas resultantes de la suma de las puntuaciones obtenidas en los ítems de cada escala.

Subescala	Ítems	Mínimo	Máximo
<b>Funcionamiento Físico</b>	3a, 3b, 3c, 3d, 3e, 3f, 3g, 3h, 3i, 3j	10	30
<b>Desempeño Físico</b>	4a, 4b, 4c, 4d	4	20
<b>Dolor Corporal</b>	7, 8	2	10
<b>Salud General</b>	1, 11a, 11b, 11c, 11d	5	25
<b>Vitalidad</b>	9a, 9e, 9g, 9i	4	20
<b>Funcionamiento Social</b>	6, 10	2	10
<b>Desempeño Emocional</b>	5a, 5b, 5c	3	15
<b>Salud Mental</b>	9b, 9c, 9d, 9f, 9h	5	25

Tabla 4. Ítems, escalas y puntuaciones del SF-36.

Todos los ítems del SF-36 son puntuados de forma a que a la puntuación más elevada le corresponda una mejor percepción del estado de Salud. Por ejemplo, una puntuación elevada en la subescala de Funcionamiento Físico indica un buen funcionamiento, mientras una puntuación elevada en la subescala de Dolor Corporal, indica ausencia de dolor. Por lo tanto, existen 10 ítems que necesitan ser invertidos: ítem 1, 6, 7, 8, 9a, 9d, 9e, 9h, 11b e 11d.

Los autores sugieren la estandarización de las puntuaciones de forma que las subescalas sean interpretables, expresando las puntuaciones obtenidas en una escala de 0 a 100. Las puntuaciones estandarizadas se obtienen de la siguiente forma:

$$\frac{\text{puntuación obtenida en la escala} - \text{puntuación mínima en la escala}}{\text{puntuación máxima en la escala} - \text{puntuación mínima en la escala}} \times 100$$

Por ejemplo, se un individuo obtiene 24 en la subescala de Salud Mental y teniendo en cuenta que la puntuación mínima en esta escala es 5 y la máxima es 25, la puntuación estandarizada es  $[(24-5)/20] \times 100 = 95$ .

McHorney (1993) y Keller *et al* (1998) confirman que las 8 dimensiones del SF-36 se agrupan en las dos componentes referidas anteriormente: la Componente Física y la Componente Mental. De esta forma, el Funcionamiento Físico, el Desempeño Físico, el Dolor Corporal y la Percepción General del Estado de Salud son dimensiones con correlación más fuerte con la **Componente Física**, mientras la Vitalidad, la Salud Mental, el Funcionamiento Social y el Desempeño Emocional, integran la **Componente Mental**. Estos autores confirmaron también que las dimensiones Vitalidad y Percepción General del Estado de Salud presentaban correlación elevada con ambas las componentes. La dimensión Funcionamiento Social exhibía una correlación elevada con la Componente Mental y también una correlación moderada con la Componente Física.



### 6.1.2. Rotterdam Symptom Check List (RSCL)

El *Rotterdam Symptom Check List* (RSCL) es una escala específica de CVRS que ha sido desarrollada por De Haes *et al* (1990) para evaluar síntomas físicos y psíquicos en pacientes con cáncer. Este instrumento está constituido por los 30 ítems que presentamos en la Tabla 5.

Durante la semana pasada, ¿ha tenido Ud...? (nada; un poco; bastante; mucho)
s1. Falta de apetito
s2. Irritabilidad
s3. Cansancio
s4. Preocupación
s5. Dolor muscular
s6. Ánimo deprimido
s7. Falta de fuerzas (flojedad)
s8. Dolor lumbar (dolor de riñones)
s9. Nerviosismo
s10. Náusea
s11. Esperanza por el futuro
s12. Dificultad para dormir
s13. Dolores de cabeza
s14. Vómitos
s15. Mareos
s16. Disminución del interés sexual
s17. Estado de tensión (estrés)
s18. Molestias abdominales (estómago)
s19. Ansiedad
s20. Estreñimiento
s21. Diarrea
s22. Acidez/ ardor de estómago
s23. Escalofríos (tiritona)
s24. Hormigueos en manos y piernas
s25. Dificultad de concentración
s26. Dolor de boca o dolor al tragar
s27. Pérdida de pelo
s28. Escozor o dolor de ojos
s29. Dificultad para respirar
s30. Boca seca
<b>A continuación se relacionan un número de actividades. No queremos saber si realmente las hace, sólo si Ud. Sería capaz de hacerlas actualmente. ¿Podría, por favor, marcar con una X la respuesta que se adecue mejor a su estado de la semana pasada? (incapaz; solamente con ayuda; sin ayuda pero con dificultad; sin ayuda).</b>
act1. Cuidado Personal (lavarse, etc.)
act2. Caminar por la casa
act3. Tareas domésticas ligeras (pelar patatas, etc.)
act4. Subir escaleras
act5. Tareas domésticas pesadas(hacer la cama, etc.)

<b>act6.</b> Paseos fuera de casa (salir a la calle)
<b>act7.</b> Ir de compras
<b>act8.</b> Ir a trabajar
<b>cv1.</b> Considerando todas estas cosas, ¿cómo describiría su calidad de vida durante la semana pasada? (excelente; buena; moderadamente buena; ni buena ni mala; moderadamente mala; mala; muy mala)

Tabla 5. Ítems del RSCL.

Además de los ítems que integran el RSCL, hemos añadido a este cuestionario una última cuestión (cv2) que consistió en replicar el último ítem, pidiendo al profesional de salud que lo contestara. El objetivo ha sido comparar hasta qué punto la opinión del profesional difiere de la del paciente cuando se trata de evaluar el nivel de CVRS de este último.

En la Tabla 6 se presentan los ítems asociados a cada una de las escalas y las respectivas puntuaciones mínimas y máximas resultantes de la suma de las puntuaciones obtenidas en los ítems de cada escala.

Subescala	Ítems	Mínimo	Máximo
<b>Síntomas Físicos</b>	s1, s3, s5, s7, s8, s10, s12, s13, s14, s15, s16, s18, s20, s21, s22, s23, s24, s25, s26, s27, s28, s29, s30	23	92
<b>Angustia Psicológica</b>	s2, s4, s6, s9, s11, s17, s19	7	28
<b>Nivel de Actividad</b>	act1, act2, act3, act4, act5, act6, act7, act8,	8	32
<b>CVRS Global</b>	cv1	1	7

Tabla 6. Ítems, escalas y puntuaciones del RSCL.

Para las subescalas de Síntomas Físicos y Angustia Psicológica los ítems son puntuados de 1 a 4 y, cuanto más elevada es la puntuación obtenida, menor es el nivel de CVRS reflejado, con excepción del ítem s11 (“esperanza en el futuro”) que deberá ser previamente invertido.

Los ítems de la subescala relativa al Nivel de Actividad son igualmente puntuados de 1 a 4. Sin embargo, en este caso, cuanto más elevada es la puntuación obtenida, mayor es el nivel de funcionalidad e por tanto de CVRS. El

Nivel de CVRS global es puntuado en una escala de 1 a 7 y, de forma que cuanto más elevada es la puntuación obtenida, más bajo es el nivel de CVRS reflejado. Para una interpretación más fácil se ha optado por invertir los ítems de la subescala de Nivel de Actividad. Así, para todas las subescalas del RSCL, puntuaciones más elevadas reflejan menor nivel de CVRS percibida.

De Haes *et al* (1996) sugieren la estandarización de las puntuaciones en una escala de 0 a 100, de forma que los resultados sean más fácilmente interpretables y comparables. En esta escala, 0 significa que no existe deterioro de la CVRS, mientras 100 indica deterioro máximo. Esta estandarización puede hacerse de la siguiente forma:

$$\frac{\text{puntuación obtenida en la escala} - \text{puntuación mínima en la escala}}{\text{puntuación máxima en la escala} - \text{puntuación mínima en la escala}} \times 100$$

Por ejemplo, se un individuo obtiene 14 en la escala de angustia psicológica y teniendo en cuenta que la puntuación mínima en esta escala es 7 y la máxima es 28, la puntuación estandarizada es  $[(14-7)/21]*100=33$ .

En el caso de que existan datos faltantes, De Haes *et al* (*op.cit.*) sugieren la sustitución del valor por la media del individuo en la respectiva subescala. Los autores consideran que este procedimiento puede ser seguido siempre que el respondiente haya contestado al menos el 50% de los ítems de la subescala en cuestión, en caso contrario el individuo será considerado como un caso faltante para la subescala en cuestión.

El RSCL puede ser utilizado con propósitos descriptivos. En esos casos es frecuente dicotomizar los resultados y calcular el porcentaje de pacientes que contesta “*no en absoluto*” y el porcentaje que presenta algún tipo de síntoma. En las situaciones en que se pretende efectuar comparaciones entre distintos grupos es común recurrir al cálculo de la medias y desviación estándar para cada subescala y para cada ítem. Sin embargo, teniendo en cuenta que muchas veces la

distribución obtenida no siempre es Normal, los tests no paramétricos son, en esos casos, más apropiados.

El RSCL se ha revelado útil en la detección de casos de morbilidad psicológica o psiquiátrica. En este caso, se necesita un punto de corte para la escala de angustia psicológica que indique la presencia o ausencia de enfermedad. Diferentes estudios han conducido a distintos puntos de corte. Sin embargo, De Haes *et al* (1996) sugieren un punto de corte de 15 como el más indicado en poblaciones de pacientes estables, mientras en pacientes con enfermedad progresiva es conveniente utilizar el 16. Para la escala de síntomas físicos no es habitual utilizar un punto de corte.

En los últimos años, el RSCL ha sido utilizado en un amplio número de estudios oncológicos y se ha revelado apropiado en pacientes con cáncer tanto en pacientes sometidos a cirugía, como quimioterapia o radioterapia, así como en diferentes tipos de cáncer (mama, ovario, próstata, testículo, pulmón, colon-rectal, gástrico, vejiga, renal, cabeza y cuello) y en grupos heterogéneos de pacientes. El RSCL ha sido también utilizado tanto en estadios primarios de la enfermedad como en estados más avanzados y también en poblaciones normales como grupo de comparación.

## 6.2. OBJETIVO Y MÉTODO

El principal **propósito** del presente estudio, ha sido descubrir cambios en la autopercepción de pacientes con cáncer relativamente a su CVRS a lo largo del proceso de tratamiento de la enfermedad. En particular, se pretende saber si pacientes de cáncer cambian sus estándares internos en los cuales basan la percepción que tienen de su calidad de vida, cambian sus valores, o reconceptualizan el constructo. Si ocurre así, pretendemos analizar de qué manera

se altera la interpretación de CVRS de estos pacientes, describiendo el cambio e identificando diferencias entre individuos y grupos en términos de la naturaleza, dirección y magnitud del *response shift*.

La **obtención de los datos** ha sido un proceso largo y arduo. En este proceso han estado implicados los directores de los servicios de Cirugía General, Obstetricia, Oncología y/o Psicología de cinco instituciones hospitalarias portuguesas:

- Hospital Amato Lusitano, en Castelo Branco;
- Hospital de Faro EPE, en Faro;
- Centro Hospitalar da Cova da Beira EPE, en Covilhã;
- Hospital de Reynaldo dos Santos, en Vila Franca de Xira;
- Hospital Garcia de Orta, en Almada.

La primera etapa ha consistido en la formulación de la solicitud para recoger los datos al Consejo de Administración y Comisión de Ética de cada una de las instituciones referidas. También se había pedido autorización al Instituto de Portugués de Oncología Francisco Gentil (Lisboa). Sin embargo esta petición ha sido rechazada. Después de obtener los respectivos permisos, se ha iniciado el proceso de aplicación de los cuestionarios. Los datos utilizados en este trabajo han sido recogidos entre septiembre de 2007 y abril de 2009.

Además de los directores de servicio, han tenido una importancia fundamental en este proceso las enfermeras de los servicios donde se ha desarrollado el estudio. Afortunadamente, en cada uno de los hospitales hemos conseguido que al menos una enfermera se implicara en este estudio. Estas profesionales han sido incansables y una pieza clave en el seguimiento de cada paciente y aplicación de los cuestionarios. Se ha constatado que el paciente se siente más confortable rellenando el cuestionario con la enfermera o psicóloga que

con cualquier persona extraña al hospital y por eso, siempre que ha sido posible, la aplicación de los cuestionarios ha sido llevada a cabo por el profesional de salud.

Como **criterios de inclusión**, se ha establecido que participarían en nuestro estudio todos los pacientes que pasaran por los servicios mencionados de las cinco instituciones hospitalarias implicadas desde la fecha de obtención del permiso hasta abril de 2009. La muestra ha incluido pacientes con cáncer local, loco-regional o en proceso de metástasis.

Como **criterios de exclusión** se ha determinado que no participarían en este estudio pacientes cuya capacidad de entender las preguntas de los cuestionarios estuviera afectada así como pacientes con esperanza de vida inferior a tres meses. También se han excluido los pacientes que habían recibido tratamiento por cáncer anteriormente, con el objetivo de proceder a una evaluación desde el comienzo de la trayectoria del paciente.

Los pacientes han rellenado el SF-36 y el RSCL en fases distintas del proceso, de acuerdo con el *timing* siguiente:

- 1) El **Pre-Test (1ª Aplicación)** ha sido hecho antes de la intervención quirúrgica o en el inicio de la quimioterapia (para los pacientes que no han sido operados). En esta fase se ha aplicado el cuestionario sociodemográfico, el SF-36 y el RSCL;
- 2) La **2ª Aplicación** se ha hecho después de la intervención quirúrgica (día de la alta) o en el medio de los ciclos de quimioterapia (para los pacientes que no han sido operados). En esta fase se ha aplicado el SF-36 y el RSCL.
- 3) El **Post-Test (3ª Aplicación)** y el **Then-Test (4ª Aplicación)** se han completado en una visita posterior o por teléfono entre 3 y 6 meses después o 1 mes después del final de la quimioterapia (para los pacientes que no han sido operados).

Recuérdese que, mientras el *Post-Test* consiste en una aplicación convencional, el *Then-Test* (Howard *et al*, 1979b) consiste en evaluación retrospectiva hacia el momento antes de la intervención quirúrgica o el momento antes del tratamiento. A los paciente se les ha pedido que hicieran una nueva evaluación sobre cómo estaban en el momento del *Pre-Test*. Como hemos referido anteriormente, el *Then-Test* se puede utilizar para acceder al fenómeno de *response shift* ya que la comparación entre el *Pre-Test* y el *Then-Test* (evaluaciones que se reportan al mismo instante) reflejan la dirección y magnitud del cambio en los estándares internos (Sprangers & Schwartz, 1999b; Schwartz *et al*, 2004).

Los pacientes han sido informados sobre el objetivo del estudio en el cual eran libres de participar y de abandonar en el momento que ellos quisieran, sin dar ninguna explicación y sin que sus cuidados médicos y sus derechos legales salieran afectados. Ha sido garantizada la confidencialidad de la información solicitada y rellenado el **Consentimiento Informado** (Anexo VI).

Nuestro estudio no está basado en un diseño clínico ya que el principal objetivo no está enfocado en comparar ratios de respuesta entre diferentes tratamientos sino detectar *response shift* y medir cambio a través del tiempo. El **diseño** ha sido **descriptivo, longitudinal y observacional**.

### 6.3. DESCRIPCIÓN DE LA MUESTRA

Durante el estudio, 150 pacientes han completado la primera aplicación (*Pre-Test*) del SF-36 y del RSCL. Algunos pacientes han fallecido o abandonado el estudio, siendo la muestra que completó las aplicaciones finales (*Post-Test* y *Then-Test*) constituida por 132 pacientes. De los pacientes que iniciaran el estudio, la mayor parte ha sido seguido en el Hospital Amato Lusitano – Castelo Branco (45,33%) y en el Hospital de Faro EPE (36,67%).

### 6.3.1. Variables Sociodemográficas

La Tabla 7 contiene la distribución porcentual relativa a las variables sociodemográficas más relevantes para la muestra en estudio. La mayor parte de los pacientes que ha completado el estudio son mujeres (63,64%). En el momento de la primera aplicación la edad media de los pacientes era de 62,23 años (con una desviación estándar de 12,64 años).

La mayor parte de los pacientes estaba casado (70,23%), tenía hijos (93,94%), estaba jubilado (61,42%) y tenía estudios primarios (57,36%).

<b>Sexo</b>	<b>n</b>	<b>%</b>
Varón	48	36,36
Mujer	84	63,64
<b>Edad</b>	<b>n</b>	<b>%</b>
[30;40[	5	3,79
[40;50[	20	15,15
[50;60[	31	23,48
[60;70[	36	27,27
[70;80[	31	23,48
>=80	9	6,82
<b>Estado Civil</b>	<b>n</b>	<b>%</b>
Soltero	12	9,16
Casado	92	70,23
Divorciado	5	3,82
Viudo	22	16,79
<b>¿Tiene Hijos?</b>	<b>n</b>	<b>%</b>
Sí	124	93,94
No	8	6,06
<b>Actividad Actual</b>	<b>n</b>	<b>%</b>
Tiempo Completo	33	25,98
Media Jornada	6	4,72
Jubilado	78	61,42
Parado	10	7,87
<b>Nivel de Estudios</b>	<b>n</b>	<b>%</b>
Sin Estudios	14	10,85
Primarios	74	57,36
Secundarios	33	25,58
Universitarios	8	6,20

Tabla 7. Información sociodemográfica de la muestra en estudio.

### 6.3.2. Variables Clínicas

Se ha registrado la información clínica elemental relativa a los pacientes en estudio, cuya síntesis se presenta en Tabla 8. Los diagnósticos de enfermedad



oncológica más frecuentes han sido neoplasias de la mama (28,79%), colon o recto (28,79%), útero u ovario (13,64%) y gástrico (13,64%), habiendo casos de otros tipos de neoplasias como vesical o renal, pulmonar, próstata o piel. La mayor parte de los tumores diagnosticado era local (55,00%) y en la mayor parte de los casos el peso era estable (50,38%). La terapéutica más común ha sido la cirugía combinada con la quimioterapia (40,15%) o solamente la intervención quirúrgica (31,06%).

<b>Neoplasia</b>	<b>n</b>	<b>%</b>
Mama	38	28,79
Colon/Recto	38	28,79
Útero/Ovario	18	13,64
Gástrico	18	13,64
Otras	20	15,14
<b>Estadio</b>	<b>n</b>	<b>%</b>
Local	66	55,00
Loco - regional	34	28,33
Metástasis	20	16,67
<b>Peso</b>	<b>n</b>	<b>%</b>
Estable	66	50,38
Ha perdido menos de 5kg	33	25,19
Ha perdido entre 5kg y 10 kg	22	16,79
Ha perdido más de 10kg	10	7,63
<b>Terapéutica</b>	<b>n</b>	<b>%</b>
Cirugía	41	31,06
Cirugía+Quimioterapia	53	40,15
Cirugía+Radioterapia	14	10,61
Cirugía+Radioterapia+Quimioterapia	6	4,55
Quimioterapia	16	12,12
Radioterapia+Quimioterapia	2	1,52

**Tabla 8.** Información clínica de la muestra en estudio.

El 13,39% de los pacientes había realizado algún tratamiento en el pasado, el 13,18% tenía ya diagnósticos psiquiátricos anteriores, el 53,28% afirma que ya había contactado con alguien con cáncer, apenas el 8,00% tiene acompañamiento psicológico durante el proceso de enfermedad.

La muestra en estudio es integrada por pacientes con diferentes diagnósticos y sometidos a distintos tratamientos. Sin embargo, teniendo en cuenta que nuestro objetivo se centra en evaluar cambios al nivel de la autopercepción de estos pacientes a lo largo del proceso de tratamiento y dado el

tremendo impacto producido por la enfermedad y su tratamiento, creemos que la heterogeneidad de la muestra en estudio no compromete la fiabilidad de nuestro estudio. En este sentido también han trabajado Echteld *et al* (2007), Oort *et al* (2005), Visser *et al* (2005) y Winter *et al* (2009), entre otros.

## **CAPÍTULO VII**

---

# **RESULTADOS OBTENIDOS**

---

## 7. RESULTADOS OBTENIDOS

En este apartado presentamos los resultados obtenidos mediante la aplicación de los instrumentos genérico (SF-36) y específico (RSCL) utilizados en este estudio. Empezamos por describir las estructuras latentes y consistencia interna de los cuestionarios, las puntuaciones observadas en cada una de sus subescalas y los resultados obtenidos al comparar las puntuaciones referentes al *Then-Test* y *Post-Test* con las del *Pre-Test*. A la continuación presentamos los resultados de la aplicación de algunas de las técnicas descritas en el Capítulo IV a los datos obtenidos mediante el SF-36. La opción por trabajar solamente con el instrumento genérico ha sido justificada por la heterogeneidad de diagnósticos en la muestra en estudio.

### 7.1. CALIDAD DE VIDA DE LOS PACIENTES EVALUADA CON UN INSTRUMENTO GENÉRICO

#### 7.1.1. Estructuras Latentes del SF-36 y Consistencia Interna

Hemos empezado por confirmar la estructura factorial del SF-36 para la primera aplicación. Se ha obtenido el valor de 0,877 para el KMO y un valor de  $p < 0,001$  para el Test de Bartlett, indicando que el modelo de Análisis Factorial se adecua a la muestra en estudio. Para la extracción de los ejes se ha utilizado el Análisis de Componentes Principales. Para la determinación del número de factores a retener se ha recurrido al *scree plot* de Cattell. Además, precisamente 8 ejes tenían asociados valores propios mayores que uno. Recurriendo a la Regla del Codo, se han retenido 8 ejes, los cuales explican el 74,21 % de la varianza total.

La Tabla 9 contiene la matriz factorial rotada para la primera aplicación.

Dimensión	Ítems	Componente							
		1	2	3	4	5	6	7	8
Funcionamiento Físico	i3a					,623			
	i3b	,409		,407		,688			
	i3c			,495		,583			
	i3d			,654		,458			
	i3e			,786					
	i3f			,629		,422			
	i3g			,737					
	i3h			,865					
	i3i			,900					
	i3j			,718					
Desempeño Físico	i4a	,775		,305					
	i4b	,834							
	i4c	,820		,302					
	i4d	,779							
Dolor Corporal	i7						,919		
	i8						,767		
Salud General	i1				,494	,405			
	i11a								,721
	i11b				,750				
	i11c				,570			,355	
Vitalidad	i9a	,471	,480		,447				
	i9e	,357	,600		,417				
	i9g		,708	,306					
	i9i	,355	,648						
Funcionamiento Social	i6							,731	
	i10	,375	,433					,637	
Desempeño Emocional	i5a	,778	,337						,339
	i5b	,780							,340
	i5c	,603	,317						,313
Salud Mental	i9b		,731						
	i9c		,795						
	i9d		,780						
	i9f		,823						
	i9h		,553		,308				

Tabla 9. Matriz factorial tras rotación Varimax para el *Pre-Test*

Los valores de los factores de carga menores de 0,300 no se escriben para facilitar la interpretación.

Se observa que los ítems relativos a la dimensión Funcionamiento Físico distribuyen sus saturaciones por los ejes 3 y 5; los ítems asociados a la dimensión Desempeño Físico cargan el eje 1; los asociados al Dolor Corporal el eje 6; los ítems de Salud General concentran sus saturaciones en el eje 4, exceptuando el ítem 11a (“Creo que me pongo enfermo más fácilmente que otras personas”) que carga el eje 8; para los ítems de la dimensión Vitalidad las saturaciones más elevadas se

registran en el eje 2; los de Funcionamiento Social cargan el eje 7; Desempeño Emocional se asocia al eje 1; y Salud Mental al eje 2. Hemos encontrado que los ítems relativos a las dimensiones Desempeño Físico y Desempeño Emocional cargan el mismo eje. La aplicación del Análisis Factorial ha permitido confirmar las estructuras latentes teóricas del SF-36, no obstante algunos ítems presentan cargas en factores que no son los esperados.

Como medida de consistencia interna se ha utilizado el *alpha de Cronbach* el cual, siendo un coeficiente de fiabilidad, se utiliza como criterio para evaluar hasta qué punto el cuestionario está compuesto por ítems suficientemente homogéneos como para justificar que su suma constituya una medida del constructo subyacente. Los valores encontrados para el *alpha de Cronbach* para el SF-36 y respectivas subescalas, se presentan en la Tabla 10.

Escala	Alpha de Cronbach	N (ítems)
SF-36	0,948	36
Funcionamiento Físico	0,926	10
Desempeño Físico	0,958	4
Dolor Corporal	0,854	2
Salud General	0,738	5
Vitalidad	0,846	4
Funcionamiento Social	0,790	2
Desempeño Emocional	0,900	3
Salud Mental	0,878	5

Tabla 10. *Alpha de Cronbach* para el SF-36 y respectivas subescalas.

Los valores encontrados sobrepasan para todos los casos el valor 0,70 indicando un grado satisfactorio de consistencia interna (Bland & Altman, 1997), siendo compatibles con los encontrados en otros estudios (Anderson *et al*, 1996; Sullivan *et al*, 1995; Hays & Morales, 2001).

### 7.1.2. Puntuaciones Obtenidas en las Subescalas del SF-36

Las puntuaciones obtenidas para cada subescala del SF-36 han sido previamente estandarizadas y se han calculado los estadísticos descriptivos básicos para las subescalas del SF-36 para todas las aplicaciones (Tabla 11). La Fig.13 contiene las puntuaciones medias relativas a las subescalas del SF-36 para las diferentes aplicaciones.

Subescala	Pre-Test			2ª Aplicación			Post-Test			Then-Test		
	N	Media	Error Estándar Media	N	Media	Error Estándar Media	N	Media	Error Estándar Media	N	Media	Error Estándar Media
Funcionamiento Físico	145	68,38	2,28	103	60,19	2,66	124	66,73	2,42	124	73,79	2,61
Desempeño Físico	147	54,89	2,87	104	52,28	3,02	129	48,06	2,89	130	62,88	3,14
Dolor Corporal	145	65,17	2,49	102	62,25	2,98	128	73,05	2,55	129	70,25	3,02
Salud General	142	53,24	1,52	93	52,26	1,54	131	56,60	1,81	130	55,04	1,90
Vitalidad	139	51,89	2,15	100	53,75	2,68	128	56,49	2,24	128	59,28	2,58
Funcionamiento Social	130	70,29	2,65	86	70,20	2,88	115	77,391	2,47	118	76,59	2,75
Desempeño Emocional	145	61,15	2,52	101	57,51	2,92	131	62,85	2,70	131	66,35	2,93
Salud Mental	142	61,69	2,01	99	60,51	2,54	128	65,12	2,20	129	62,56	2,29

Tabla 11. Descriptiva básica relativa a las subescalas del SF-36 para el Pre-Test, 2ª Aplicación, Post-Test y Then-Test.

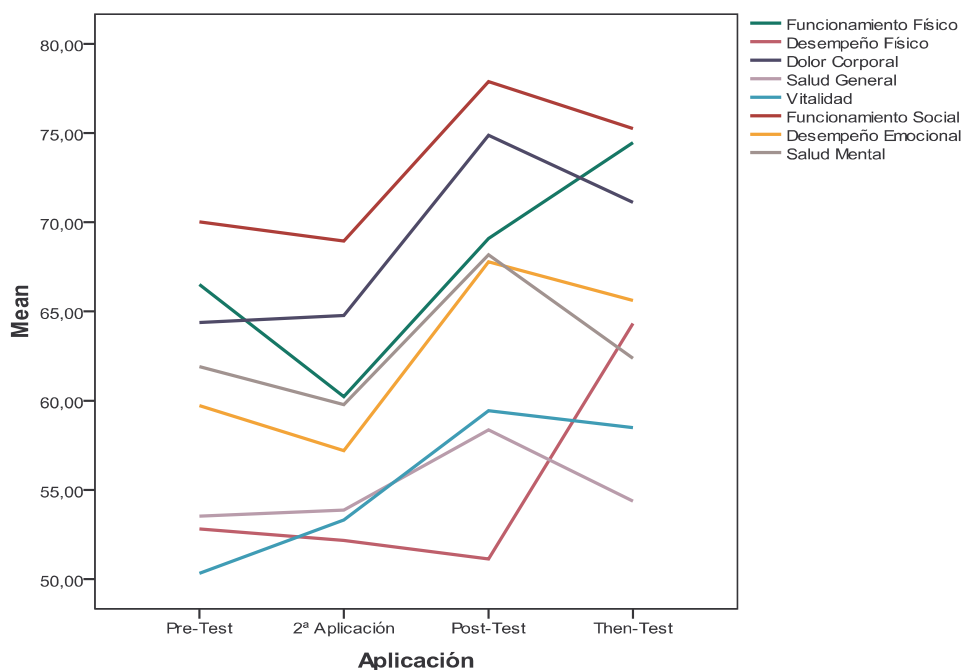


Fig.13. Puntuaciones medias relativas a las subescalas del SF-36 para las diferentes aplicaciones.

En la Tabla 11 el valor del N no es constante ya que, al no contestar a un determinado ítem, la puntuación del paciente en cuestión se ha eliminado de la dimensión respectiva.

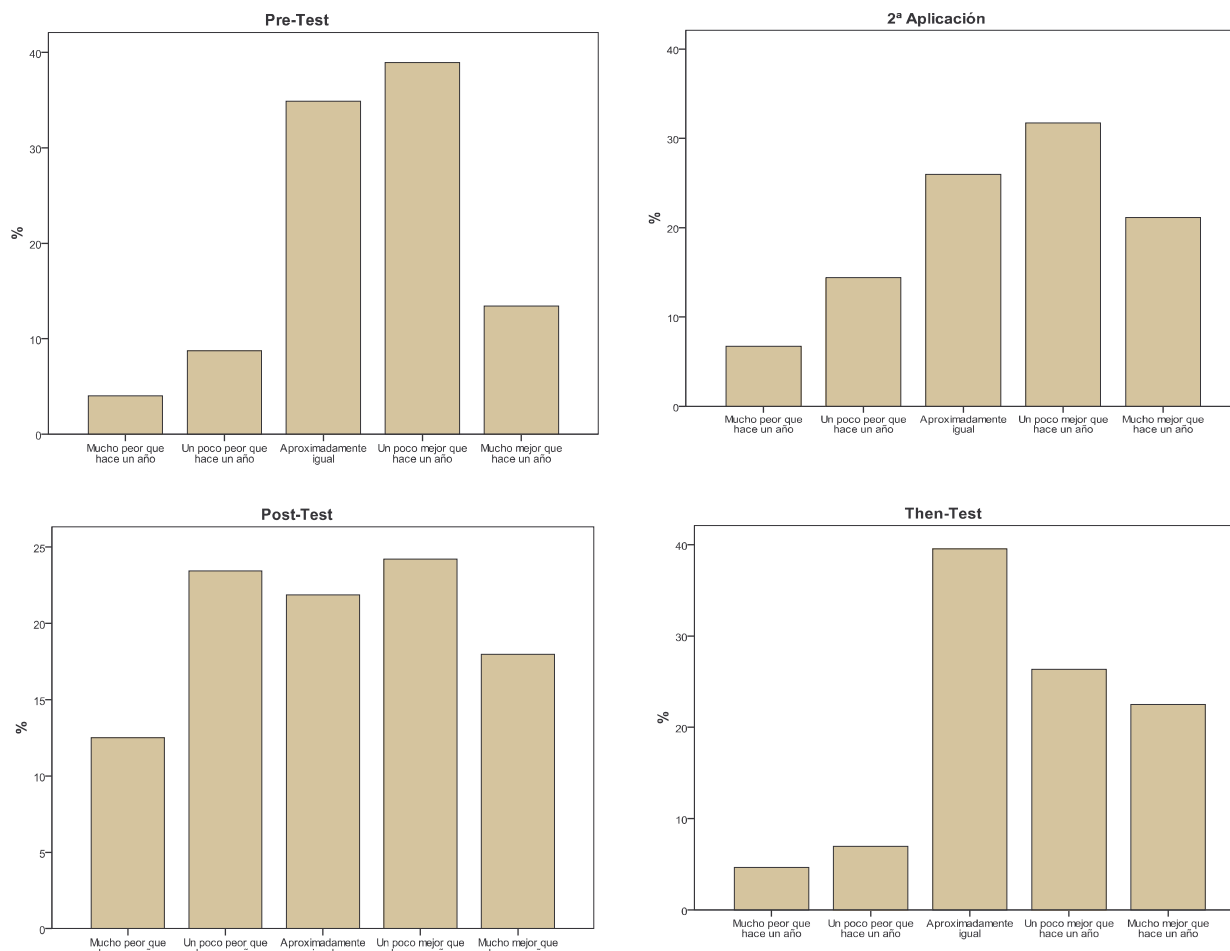
El Test ANOVA identifica diferencias estadísticamente significativas entre las cuatro aplicaciones para las subescalas de Funcionamiento Físico ( $p=0,003$ ), Desempeño Físico ( $p=0,005$ ) y Dolor Corporal ( $p=0,032$ ). Además, del análisis de la Tabla 11 y la Fig.13 se destaca que:

- Exceptuando las dimensiones Dolor Corporal y Vitalidad, las puntuaciones medias obtenidas en la 2ª Aplicación son más bajas que en el *Pre-Test*, indicando menor CVRS percibida. Nótese que en el momento de la 2ª Aplicación la mayor parte de los pacientes estaban saliendo del postquirúrgico, estando todavía recuperándose de la cirugía. Por otra parte, los que han iniciado el estudio con el inicio de la quimioterapia estaban, en el momento de la 2ª Aplicación justo a la mitad de los ciclos de quimioterapia. Esto puede explicar la disminución del nivel de calidad de vida percibida por una gran parte de los pacientes.
- Exceptuando la subescala de Desempeño Físico, las puntuaciones medias son más elevadas en el *Post-Test* que para cualquier una de las aplicaciones anteriores. Estos resultados pueden estar relacionados con el hecho de que, en el momento del *Post-Test*, la mayor parte de los pacientes habían ya finalizado los tratamientos y por eso tenían una percepción mejor de su CVRS. Nótese que los tiempos más próximos a la cirugía o al inicio de la quimioterapia son marcados por la incertidumbre, la angustia y el miedo del desconocido. En el momento del *Post-Test*, aunque se sientan más limitados al nivel del desempeño físico, los pacientes perciben como más positiva su calidad de vida.



- Los resultados obtenidos con el *Then-Test* son altamente reveladores sobre la dirección del *response shift*. Lo primero que destacamos es que, para todas las dimensiones las puntuaciones medias obtenidas con el *Then-Test* son superiores a las obtenidas en el *Pre-Test*. Reportándose el *Then-Test* precisamente al instante del *Pre-Test*, esto indica que, haciendo una evaluación retrospectiva, los pacientes consideran que se encontraban mejor que lo que habían dicho al principio. En segundo lugar, comparando los resultados obtenidos en el *Post-Test* y *Then-Test*, destacamos que, en lo que se refiere a las dimensiones de Funcionamiento Físico y Desempeño Físico, los pacientes consideran que su CVRS inicial (*Then-Test*) era superior a la que experimentan ahora (*Post-Test*). Sin embargo, para las restantes dimensiones las puntuaciones medias obtenidas en el *Then-Test* son inferiores a las obtenidas con el *Post-Test*, evidenciando que creen que ahora están mejor que antes.

El ítem 2 del SF-36 no pertenece a ninguna de las ocho dimensiones del SF-36 y por eso se ha procedido a un análisis individual. La Fig.14 contiene las distribuciones de frecuencia observadas para este ítem según las diferentes aplicaciones. Observando los porcentajes de respuesta a este ítem, (“Comparando con lo que pasaba hace un año atrás, ¿cómo evalúa su estado de salud general ahora?”) se ha constatado que en el *Pre-Test*, solamente el 12,8% de los pacientes considera que se encuentra “mucho peor que hace un año” o “un poco peor que hace un año”; en el momento de la 2ª Aplicación este porcentaje se eleva para un 21,2%; en el *Post-Test* para un 35,9%. El *Then-Test* revela que, evaluando retrospectivamente, el 11,6% considera que al empezar el estudio estaba peor que un año antes.



**Fig.14.** Distribución de la frecuencia de respuesta para el ítem 2 del SF-36 (“Comparando con lo que pasaba hace un año atrás, ¿cómo evalúa su estado de salud general ahora?”) para las diferentes aplicaciones.

En un estudio desarrollado por Walters & Brazier (2003) a partir de los datos de siete estudios longitudinales en nueve grupos de pacientes, mediante el SF-36, estos autores han examinado la relación entre el ítem 2 y la puntuación global de cambio. Al calcular el cambio global entre la primera y la segunda evaluación para cada paciente, se ha constatado que esta medida se encontraba altamente correlacionada con el estado presente y no correlacionada con el estado inicial. Los autores justifican los resultados encontrados destacando que la evaluación del cambio es psicológicamente difícil: los pacientes pueden cuantificar sus estados de salud presente e inicial y enseguida hacer una sustracción mental. Sin embargo, los pacientes pueden ser incapaces de recordar su estado inicial, basando esta evaluación en su estado actual. Según Walters & Brazier (C),

cualquier medida de cambio que refleje la diferencia no sesgada entre el estado final y inicial, debería evidenciar una correlación positiva con el estado final y una correlación negativa de igual magnitud con el estado inicial. De hecho, en estudios desarrollados, estos autores han encontrado correlaciones fuertes entre las medidas globales de cambio en CVRS y el estado presente y correlaciones mucho más bajas con el estado inicial. En nuestro estudio no se han confirmado estos resultados. En realidad, la correlación más elevada se ha registrado entre el ítem 2 en el *Pre-Test* y la subescala de Salud Mental en el *Post-Test* ( $r=0,191$ ,  $p=0,032$ ).

### 7.1.3. Comparación de las puntuaciones obtenidas en el *Then-Test* y *Post-Test* con el *Pre-Test* para el SF-36

La forma más comúnmente utilizada para detectar la presencia de cambio en contextos de CVRS está basada en la utilización del *Then-Test* (Howard *et al*, 1979b). Se asume que las medidas resultantes del *Post-Test* y *Then-Test* son obtenidas bajo los mismos estándares de medida, ya que son obtenidas al mismo tiempo. Por lo tanto, la comparación de las puntuaciones medias obtenidas en el *Pre-Test* y el *Then-Test* constituye una estimación de la magnitud y dirección del *response shift*. En esta línea, hemos comparado las puntuaciones obtenidas en el *Pre-Test* con las del *Post-Test* y *Then-Test*, de la siguiente forma:

- en primer lugar, se ha procedido a la comparación de las puntuaciones obtenidas en el *Pre-Test* y en el *Then-Test* para identificar **cambios Beta**;
- una vez descartada la ocurrencia de cambios Beta o Gamma, se han comparado las puntuaciones relativas al *Pre-Test* y *Post-Test* para identificar **cambios Alpha**, ya que la ocurrencia de cambios Beta o Gamma puede enmascarar los cambios Alpha.

La Tabla 12 contiene la información anterior para los 132 individuos que han completado las tres aplicaciones.

Subescala	N	Media	Desviación Estándar
Funcionamiento Físico (Pre-Test)	127	69,76	27,21
Funcionamiento Físico (Post-Test)	124	66,73	26,96
Funcionamiento Físico (Then-Test)	124	73,79	29,02
Desempeño Físico (Pre-Test)	129	57,95	34,24
Desempeño Físico (Post-Test)	129	48,06	32,85
Desempeño Físico (Then-Test)	130	62,88	35,84
Dolor Corporal (Pre-Test)	127	65,55	29,91
Dolor Corporal (Post-Test)	128	73,05	28,86
Dolor Corporal (Then-Test)	129	70,25	34,28
Salud General (Pre-Test)	124	52,82	18,26
Salud General (Post-Test)	131	56,60	20,69
Salud General (Then-Test)	130	55,04	21,64
Vitalidad (Pre-Test)	122	53,74	25,46
Vitalidad (Post-Test)	128	56,49	25,31
Vitalidad (Then-Test)	128	59,28	29,22
Funcionamiento Social (Pre-Test)	114	71,71	30,21
Funcionamiento Social (Post-Test)	115	77,39	26,54
Funcionamiento Social (Then-Test)	118	76,59	29,89
Desempeño Emocional (Pre-Test)	127	63,39	29,62
Desempeño Emocional (Post-Test)	131	62,85	30,94
Desempeño Emocional (Then-Test)	131	66,35	33,50
Salud Mental (Pre-Test)	125	62,64	24,23
Salud Mental (Post-Test)	128	65,12	24,91
Salud Mental (Then-Test)	129	62,56	26,05

Tabla 12. Descriptiva básica relativa a las subescalas del SF-36 para el *Pre-Test*, *Post-Test* y *Then-Test*.

### *Comparación Then-Test – Pre-Test*

Si no hubiera ocurrido ningún tipo de alteración al nivel del auto percepción de los pacientes, se esperaría que, los resultados obtenidos en el *Pre-Test* y en el *Then-Test*, fueran los mismos, ya que ambos se refieren al inicio del tratamiento. Sin embargo, al comparar los resultados obtenidos en el *Pre-Test* con los obtenidos en el *Then-Test* se observa que, exceptuando la subescala de Salud Mental, las puntuaciones obtenidas en el *Then-Test* son superiores a las obtenidas con el *Pre-Test* para todas las subescalas, sugiriendo que, mirando hacia atrás, los pacientes perciben ahora como más positiva su calidad de vida al principio que lo

que habían dicho en ese instante. En particular, se han identificado diferencias estadísticamente significativas para las dimensiones Funcionamiento Físico ( $p=0,040$ ) y Vitalidad ( $p=0,024$ ) (Tabla 13). Estos resultados indican la ocurrencia de **cambios Beta o Gamma**, traduciendo una recalibración en la escala de medida, una repriorización o incluso una reconceptualización del concepto de CVRS en lo que respecta a estas dimensiones ya que se obtienen valoraciones significativamente diferentes entre el *Pre-Test* y el *Then-Test*. Los pacientes perciben ahora de forma distinta su Funcionamiento Físico y Vitalidad y tienden a considerar que se encontraban mejor antes que lo que efectivamente habían referido en ese momento.

Dimensiones del SF-36	p-valor
Funcionamiento Físico	,040
Desempeño Físico	,068
Dolor Corporal	,082
Salud General	,438
Vitalidad	,024
Funcionamiento Social	,079
Desempeño Emocional	,254
Salud Mental	,832

Tabla 13. p-valores asociados al Test t (muestras pareadas) para las diferencias entre *Pre-Test* y el *Then-Test* en el SF-36.

### *Comparación Post-Test – Pre-Test*

En una fase siguiente, se compararan las puntuaciones obtenidas en *Pre-Test* y *Post-Test*. Nótese que, utilizando el *Then-Test*, se había detectado la presencia de **cambio Beta** para las dimensiones Funcionamiento Físico y Vitalidad. Por lo tanto, para estas dimensiones, hay que tener en cuenta que verdaderos cambios (**cambio Alpha**) podrán aparecer enmascarados por un cambio Beta.

Al comparar las puntuaciones obtenidas en el *Pre-Test* y *Post-Test* (Tabla 12), se encuentran valores medios más bajos para el *Post-Test* en lo que respecta a las dimensiones Funcionamiento Físico, Desempeño Físico y Desempeño Emocional sugiriendo que, para estas dimensiones, los pacientes tienden a percibir un nivel más bajo de calidad de vida en el momento de la última aplicación que en la

percepción inicial. Se han identificado diferencias estadísticamente significativas para las dimensiones Desempeño Físico, Dolor Corporal y Funcionamiento Social (Tabla 14). Nótese que la CVRS percibida es inferior para el Desempeño Físico, sin embargo superior para Dolor Corporal y Funcionamiento Social. Teniendo en cuenta que para estas dimensiones se había ya descartado la presencia de cambios Beta o Gamma, estos resultados indican la presencia de un **cambio Alpha**. Por lo tanto, se registra una variación real en las dimensiones Desempeño Físico, Dolor Corporal y Funcionamiento Social, indicando una mejoría de la CVRS percibida al nivel de las dos últimas y un deterioro al nivel de la primera.

Dimensiones del SF-36	p-valor
Funcionamiento Físico	,178
Desempeño Físico	,006
Dolor Corporal	,012
Salud General	,086
Vitalidad	,177
Funcionamiento Social	,038
Desempeño Emocional	,880
Salud Mental	,105

**Tabla 14.** p-valores asociados al Test t (muestras pareadas) para las diferencias entre el *Pre-Test* y el *Post-Test* en el SF-36.

Walters & Brazier (2003) concluyen que, en la dimensión de Salud General (escala de 0-100), se puede considerar como cambio mínimo clínica y socialmente relevante, una diferencia de 5 puntos. Angst *et al* (2001a) han encontrado que, en pacientes con osteoartritis, el mínimo cambio clínicamente importante (MCCI) varía entre 3,3 y 5,3 puntos en la dimensión de Funcionamiento Físico y 7,2 y 7,8 puntos en la dimensión de Dolor Corporal. Por otra parte, Hays & Morales (2001) han concluido el MCCI para el SF-36 se sitúa típicamente entre 3 y 5 puntos. En nuestro estudio, comparando el *Post-Test* y el *Pre-Test* hemos encontrado una diferencia de -3,03 para el Funcionamiento Físico, -9,88 para el Desempeño Físico, 7,5 para el Dolor Corporal, 3,78, para la Salud General y 5,68 para el Funcionamiento Social, siendo las demás inferiores a 3. Recuérdese que se había detectado la eventual presencia de cambio Beta o Gamma para la subescala de

Funcionamiento Físico, por lo que el cambio real vendría enmascarado por cambios Beta o Gamma.

*Comparación Then-Test – Pre-Test y Post-Test – Pre-Test según los diferentes grupos de diagnóstico*

Teniendo en cuenta que el tipo de cáncer diagnosticado determina el tipo de tratamiento utilizado, puede tratarse de un factor importante en la evolución de la autopercepción que el paciente posee de su calidad de vida. De esta forma, se ha procedido a una comparación de los resultados obtenidos en el *Pre-Test*, *Post-Test* y *Then-Test*, dentro de los **diferentes grupos de diagnóstico**.

Relativamente a las pacientes con **cáncer de mama**, se han encontrado puntuaciones más elevadas en el *Then-Test* que en para el *Pre-Test* para las dimensiones Funcionamiento Físico ( $p=0,041$ ) y Vitalidad ( $p=0,025$ ), indicando la presencia de cambio Beta, tal como había ocurrido en la muestra completa. Por otra parte, al comparar el *Post-Test* con el *Pre-Test*, se han encontrado resultados más elevados en el *Pre-Test* para la dimensión Desempeño Físico ( $p=0,016$ ), indicando cambio Alpha. Otros estudios basados en el SF-36 confirman que la percepción de la CVRS cambia con el tiempo en diferentes grupos de pacientes. En un estudio con pacientes de cáncer de mama, Wiel *et al* (2007) han registrado mejorías significativas al nivel del Funcionamiento Social, Desempeño Físico y Desempeño Emocional, al final de 12 meses. El estudio refiere que la Vitalidad y el Dolor Corporal siguen una U invertida, o sea, se registran mejorías durante los primeros 6 meses de seguimiento y un deterioro a lo largo de los 6 meses siguientes. No se detectaron alteraciones en lo que respecta al Funcionamiento Físico, Salud Mental y Salud General. Los autores explican el deterioro de la Vitalidad y aumento del Dolor Corporal durante la segunda mitad del seguimiento con base en los efectos colaterales de la quimioterapia y/o desarrollo de linfedema.

En nuestro estudio, hemos encontrado que los pacientes con **cáncer colon-rectal** puntúan más alto en el *Then-Test* que en el *Pre-Test* para la dimensión Salud General ( $p=0,020$ ), Vitalidad ( $p=0,011$ ) y Desempeño Emocional ( $p=0,016$ ) indicando cambio Beta. Mientras, al comparar el *Post-Test* con el *Pre-Test*, únicamente se han encontrado puntuaciones más elevadas en el *Pre-Test* para la dimensión Funcionamiento Físico ( $p=0,031$ ), indicando cambio Alpha. Gervaz *et al* (2008) han descrito mejorías significativas al nivel de la CVRS General, Funcionamiento Físico y Desempeño Físico en un grupo de pacientes de cáncer colon-rectal que habían sido sometidos a resección abdominoperineal. Los autores atribuyen los resultados obtenidos a la ocurrencia de *response shift* en el sentido de que la experiencia del cáncer produce una apreciación mejor de cada día de vida. En otro estudio, Pucciarelli *et al* (2008) han encontrado resultados semejantes.

En el grupo de pacientes con **cáncer de útero o ovario**, se han encontrado puntuaciones más elevadas en el *Post-Test*, comparativamente con el *Pre-Test*, para las dimensiones Dolor Corporal ( $p=0,008$ ), Vitalidad ( $p=0,002$ ), Funcionamiento Social ( $p=0,013$ ) y Salud Mental ( $p=0,004$ ). Mientras en el grupo de pacientes con **cáncer gástrico**, se han encontrado puntuaciones más elevadas en el *Post-Test*, comparativamente con el *Pre-Test*, para las dimensiones Desempeño Físico ( $p=0,038$ ) y Desempeño Emocional ( $p=0,021$ ), Funcionamiento Social ( $p=0,013$ ), indicando cambio Alpha.

También Mangione *et al* (1997) han evaluado la sensibilidad del SF-36 a cambios clínicos en tres grupos de pacientes sometidos a diferentes tipos de cirugías: implante del fémur, cirugía torácica para tratamiento de células pulmonares cancerígenas o reparación de un aneurisma aórtico abdominal. Se ha aplicado el SF-36 1 mes, 6 meses y 12 meses después de la cirugía y se ha concluido que las limitaciones físicas eran más significativas un mes después de la cirugía, para los tres grupos. Sin embargo, 6 meses después de la cirugía, muchos



pacientes presentaban mejorías significativas al nivel de varias dimensiones de su salud, lo que se mantenía 12 meses después. Se ha concluido que, en el contexto del estudio, el SF-36 es un instrumento sensible a los cambios en CVRS. En el grupo de pacientes que han sido sometidos a implante del fémur, el cambio ha sido mayor en las dimensiones que evaluaban aspectos físicos. Sin embargo, para los otros dos grupos, se han observado cambios significativos al nivel de las 8 dimensiones del SF-36.

**Sintetizando los resultados obtenidos, se concluye que:**

- **el SF-36 es un instrumento válido para detectar cambio en contextos de evaluación de la CVRS en pacientes con cáncer;**
- **el *Then-Test* permitió detectar la ocurrencia de cambio Beta en lo que respecta a las dimensiones Funcionamiento Físico y Vitalidad;**
- **se ha identificado cambio real en la percepción que los pacientes tienen de su calidad de vida para las dimensiones Desempeño Físico, Dolor Corporal y Funcionamiento Social.**

## 7.2. CALIDAD DE VIDA DE LOS PACIENTES EVALUADA CON UN INSTRUMENTO ESPECÍFICO

Aunque los instrumentos genéricos tengan la ventaja de posibilitar la comparación de la percepción que pacientes con diferentes patologías poseen de su CVRS, tienden a ser poco sensibles a síntomas específicos del dominio oncológico. Exactamente por eso, hemos optado por incluir también los resultados de la evaluación de la CVRS de estos pacientes mediante un instrumento específicamente concebido para diagnósticos oncológicos: el RSCL. Este instrumento incluye ítems sobre el impacto de los síntomas asociados al cáncer y su tratamiento.

### 7.2.1. Estructuras Latentes del RSCL y Consistencia Interna

Tras la aplicación del Análisis Factorial a los datos referentes a la primera aplicación, se ha confirmado la estructura latente teórica del RSCL. Se ha obtenido el valor de 0,877 para el KMO y un valor de  $p < 0,001$  para el Test de Bartlett, indicando que el modelo de Análisis Factorial se adecua a la muestra en estudio. Para la extracción de los ejes se ha utilizado el Análisis de Componentes Principales y se ha determinado el número de factores a retener recurriendo al *scree plot* de Cattell. La Regla del Codo, indica la existencia de 3 ejes, los cuales explican el 44,29 % de la varianza total.

Se ha utilizado el *alpha de Cronbach* como medida de consistencia interna. Los valores encontrados para el RSCL y respectivas subescalas, se presentan en la Tabla 15.

Escala	Alpha de Cronbach	N (ítems)
RSCL	0,908	38
Síntomas Físicos	0,846	23
Angustia Psicológica	0,833	7
Nivel de Actividad	0,922	8

Tabla 15. Alpha de Cronbach para el RSCL y respectivas subescalas.

Los valores encontrados sobrepasan para todos los casos el valor 0,80 indicando un grado bueno de consistencia interna (Bland & Altman, 1997) siendo compatibles con los encontrados en otros estudios (Martins, 2009).

### **7.2.2. Puntuaciones obtenidas en la subescalas del RSCL**

Los estadísticos descriptivos básicos relativos a las puntuaciones estandarizadas obtenidas para las subescalas del RSCL en las distintas aplicaciones se encuentran en la Tabla 16. La Fig.15 contiene las puntuaciones medias relativas a las subescalas del RSCL para las diferentes aplicaciones.

Además de la evaluación de la CVRS Global por el paciente, hemos pedido a los profesionales de salud que evaluaran ellos la CVRS del paciente en la misma escala que estos habían hecho y en el mismo instante. Varios autores consideran que la evaluación que el paciente hace de su propia calidad de vida no siempre coincide con la evaluación hecha por el médico o enfermero (Fayers & Machin, 2000; Rapkin & Schwartz, 2004; Shimozuma *et al*, 2007; Westerman *et al*, 2008). Muchas veces el profesional de salud tiende a subestimar el impacto psicológico de los síntomas (como dolor, náusea, vómitos) porque, siendo estos ya esperados, son muchas veces olvidados. Por otra parte, en lo que respecta a las limitaciones inherentes al desempeño funcional, es frecuente que el profesional de salud considere que el estado del paciente es peor que lo que se constata en la evaluación hecha por el propio paciente (Brunelli *et al*, 1998). Por otra parte, hemos visto que, la percepción que el paciente posee de su calidad de vida, tiende a modificarse a lo largo del proceso de enfermedad y tratamiento, lo que puede explicar también eventuales diferencias observadas entre las evaluaciones de pacientes y profesionales de salud. Por todo esto, hemos considerado pertinente incluir en este estudio la evaluación del profesional de salud.

Subescala	Pre-Test			2ª Aplicación			Post-Test			Then-Test		
	N	Media	Error Estándar Media	N	Media	Error Estándar Media	N	Media	Error Estándar Media	N	Media	Error Estándar Media
Síntomas Físicos	124	28,60	1,41	89	32,78	1,70	117	31,59	1,59	117	27,63	1,80
Angustia Psicológica	139	43,24	2,04	101	39,89	2,33	128	36,38	2,31	129	41,82	2,45
Nivel de Actividad	130	18,59	2,25	96	21,88	2,42	121	14,29	1,89	119	13,13	2,22
CVRS Global	150	38,67	1,85	106	41,51	2,10	132	33,33	1,79	132	37,75	2,19
CVRS Global (Professional Salud)	118	34,75	1,92	80	37,92	1,97	81	33,95	2,06	58	35,92	3,22

Tabla 16. Descriptiva básica relativa a las subescalas del RSCL para el Pre-Test, 2ª Aplicación, Post-Test y Then-Test.

En la Tabla 16 el valor del N no es constante ya que, al no contestar a un determinado ítem, la puntuación del paciente en cuestión es eliminada de la dimensión respectiva.

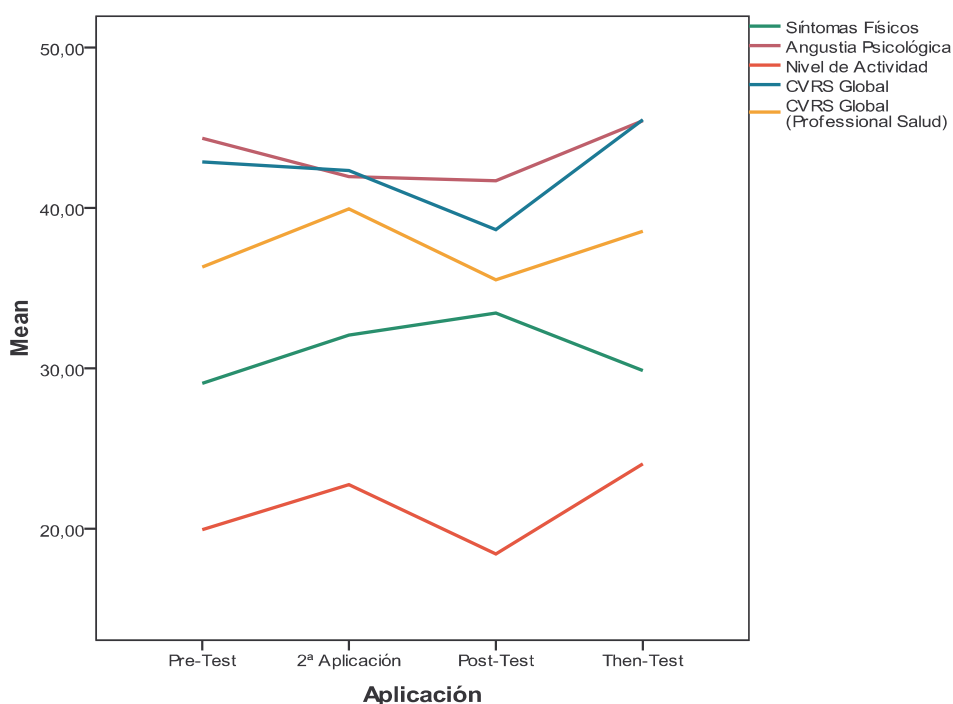


Fig.15. Puntuaciones medias relativas a las subescalas del RSCL y evaluación del profesional de salud para las diferentes aplicaciones.

El Test ANOVA identifica diferencias estadísticamente significativas entre las cuatro aplicaciones para las subescalas Nivel de Actividad ( $p=0,026$ ) y CVRS Global ( $p=0,044$ ). Además, el análisis de la Tabla 16 y la Fig.15 evidencia que:

- Exceptuando la subescala de Angustia Psicológica, las puntuaciones medias obtenidas en la 2ª Aplicación son más elevadas que en el *Pre-Test*, indicando la presencia de más síntomas y por tanto menor CVRS. Estos resultados están en perfecta sintonía con los encontrados para el SF-36. Como hemos mencionado, en el momento de la 2ª Aplicación, una gran parte de los pacientes, estaba todavía recuperándose de la cirugía en la mitad de los ciclos de quimioterapia. Por otra parte, es esperable que la Angustia Psicológica sea más evidente en el momento que antecede la cirugía o antes del inicio de la quimioterapia, ya que estos momentos son marcados por la incertidumbre y el miedo a lo desconocido.
  
- Las puntuaciones medias son más bajas en el *Post-Test* para la subescala de Nivel de Actividad y en lo que respecta a la evaluación de CVRS Global tanto por parte del paciente como por parte del profesional de salud, indicando por eso un nivel más alto de CVRS. Tal como habíamos concluido para el SF-36, en el momento del *Post-Test*, los pacientes perciben como más positiva su CVRS, aunque los síntomas físicos sean más evidentes.
  
- Al revés de lo que habíamos encontrado con el SF-36, para el RSCL, no se encuentran grandes diferencias entre las puntuaciones medias asociadas al *Then-Test* y a las referentes a al *Pre-Test*. Por otra parte, comparando las puntuaciones obtenidas en el *Post-Test* y *Then-Test*, los resultados están en conformidad con lo que habíamos encontrado para el SF-36: en lo que se refiere a la dimensión Síntomas Físicos los pacientes consideran que su CVRS inicial era superior a la de ahora. Además, para las restantes dimensiones las puntuaciones medias

obtenidas en el *Then-Test* son superiores a las obtenidas con el *Post-Test*, evidenciando que creen que ahora están mejor que antes.

- Los **profesionales de salud** tienden a evaluar como más positiva la CVRS del paciente que el propio paciente. El Test t para muestras apareadas ha permitido confirmar que las diferencias entre la evaluación del paciente y la del profesional de salud son estadísticamente significativas para todas las aplicaciones ( $p < 0,001$ ). Estos resultados concuerdan con los referidos en la literatura (Fayers & Machin, 2000; Rapkin & Schwartz, 2004; Shimozuma *et al*, 2007; Westerman *et al*, 2008; Brunelli *et al*, 1998) y nos llevan a concluir que los profesionales de salud tienden a subestimar el impacto psicológico de los síntomas asociados a la enfermedad oncológica.

### 7.2.3. Comparación de las puntuaciones obtenidas en el *Then-Test* y *Post-Test* con el *Pre-Test* para el RSCL

Hemos comparado las puntuaciones obtenidas con el RSCL en el *Pre-Test*, *Post-Test* y *Then-Test* con el objetivo de detectar la ocurrencia de cambio. Tal cómo habíamos procedido en el análisis de los resultados obtenidos mediante la aplicación del instrumento genérico, en primer lugar, se ha efectuado una comparación de las puntuaciones obtenidas en el *Pre-Test* y en el *Then-Test* para identificar **cambios Beta**. Una vez descartada la ocurrencia de cambios Beta o Gamma, se han comparado las puntuaciones relativas al *Pre-Test* y *Post-Test* para identificar **cambios Alpha**, ya que la ocurrencia de cambios Beta o Gamma puede enmascarar los cambios Alpha.

La Tabla 17 contiene la información anterior para los 132 individuos que han completado las tres aplicaciones.

Subescala	N	Media	Desviación Estándar
Síntomas Físicos (Pre-Test)	108	28,09	16,30
Síntomas Físicos (Post-Test)	117	31,59	17,22
Síntomas Físicos (Then-Test)	117	27,64	19,52
Angustia Psicológica (Pre-Test)	121	43,76	24,27
Angustia Psicológica (Post-Test)	128	36,38	26,16
Angustia Psicológica (Then-Test)	129	41,82	27,85
Nivel de Actividad (Pre-Test)	112	16,74	25,06
Nivel de Actividad (Post-Test)	121	14,29	20,81
Nivel de Actividad (Then-Test)	119	13,13	24,29
CVRS Global (Pre-Test)	132	37,37	22,66
CVRS Global (Post-Test)	132	33,33	20,59
CVRS Global (Then-Test)	132	37,75	25,21
CVRS Global - Professional Salud (Pre-Test)	105	33,49	20,35
CVRS Global - Professional Salud (Post-Test)	81	33,95	18,53
CVRS Global - Professional Salud (Then-Test)	58	35,92	24,54

**Tabla 17.** Descriptiva básica relativa a las subescalas del RSCL para el *Pre-Test*, *Post-Test* y *Then-Test*.

### *Comparación Then-Test – Pre-Test*

Al comparar los resultados obtenidos en el *Pre-Test* y en el *Then-Test* (Tabla 17), se observan valores próximos. La mayor diferencia se registra en la subescala Nivel de Actividad indicando que, en una evaluación retrospectiva, los pacientes perciben que tenían menos síntomas al principio que lo que habían referido en ese instante y por lo tanto un mayor nivel de CVRS. No se han identificado diferencias estadísticamente significativas entre el *Then-Test* y el *Pre-Test* para las dimensiones del RSCL (Tabla 18). Estos resultados indican que no se han detectan **cambios Beta o Gamma**.

Dimensiones del RSCL	p-valor
Síntomas Físicos	,352
Angustia Psicológica	,557
Nivel de Actividad	,063
CVRS Global	,865
CVRS Global - Professional Salud	,545

**Tabla 18.** p-valores asociados al Test t (muestras apareadas) para las diferencias entre el *Pre-Test* y el *Then-Test* en el RSCL.

Descartada la presencia de cambios Beta e Gamma, se identificaron diferencias estadísticamente significativas entre las puntuaciones obtenidas en el *Pre-Test* y *Post-Test* únicamente para la subescala de Angustia Psicológica (Tabla 19). Este resultado confirma la ocurrencia de un **cambio Alpha** en el sentido de una disminución efectiva de los síntomas asociados a esta dimensión.

Dimensiones del RSCL	p-valor
Síntomas Físicos	,275
Angustia Psicológica	,005
Nivel de Actividad	,356
CVRS Global	,056
CVRS Global - Professional Salud	,303

Tabla 19. p-valores asociados al Test t (muestras pareadas) para las diferencias entre el *Pre-Test* y el *Post-Test* en el RSCL.

#### *Comparación Then-Test – Pre-Test y Post-Test – Pre-Test según los diferentes grupos*

Siendo el RSCL un instrumento específicamente destinado a pacientes oncológicos, hemos valorado si la percepción que los pacientes poseen sobre los síntomas físicos, la angustia psicológica y su calidad de vida en general, evoluciona de diferentes formas según el sexo, la edad, los distintos grupos de diagnóstico y entre el grupo de pacientes que estaba haciendo algún tratamiento en el momento del *Then-Test*. Para ello, hemos procedido a una comparación de los resultados obtenidos en el *Pre-Test*, *Post-Test* y *Then-Test*, en función de estas variables.

Los resultados obtenidos han evidenciado diferencias estadísticamente significativas entre el *Pre-Test* y el *Then-Test* para:

- las **mujeres** ( $p=0,034$ ) en lo que respecta a la dimensión Nivel de Actividad, siendo evidente que, al final, estas pacientes evalúan retrospectivamente como más positiva su CVRS de lo que lo habían hecho inicialmente. De hecho parece normal que, debido a las tareas del



hogar, las mujeres se sientan más limitadas que los hombres a nivel de su actividad física, justificando el cambio en esta dimensión;

- los **pacientes con menos de 50 años** ( $p=0,006$ ) en lo que respecta al nivel de CVRS Global, también consideran que se encontraban mejor, al evaluar retrospectivamente.

Por otra parte, las diferencias entre el *Pre-Test* y el *Post-Test* son más evidentes en la subescala de **Angustia Psicológica**. Así, hemos encontrado niveles más bajos de angustia psicológica en el momento de la última aplicación para los siguientes grupos:

- pacientes de cáncer de mama ( $p=0,002$ );
- pacientes de cáncer de colon-rectal ( $p=0,011$ );
- pacientes que no hicieran tratamiento ( $p=0,001$ );
- pacientes con más de 50 años ( $p=0,003$ ).

Hemos encontrado también diferencias entre el *Pre-Test* y el *Post-Test* en la subescala de **Síntomas Físicos** para el grupo de pacientes que había terminado los tratamientos en el momento de la última aplicación ( $p=0,001$ ), siendo en este caso menos los síntomas percibidos.

Los resultados encontrados están en concordancia con los descritos en la literatura. Avis *et al* (2005) llevaron a cabo un estudio con el objetivo de evaluar la CVRS en mujeres jóvenes (con menos de 50 años) entre 4 y 42 meses después de haber recibido el diagnóstico de cáncer de mama, intentando identificar los factores asociados con el deterioro de la calidad de vida. Más de un 70% de las mujeres han relatado dolor, molestias en general y se han manifestado insatisfechas con su apariencia. Sofocos, problemas al nivel sexual y dificultad en controlar la vejiga aumentaron con la edad. La CVRS global percibida ha sido significativamente inferior que la CVRS en una muestra de mujeres jóvenes sanas.

Por otra parte, los días de trabajo perdidos y la disminución en la actividad habitual inmediatamente después del diagnóstico, los problemas al nivel sexual, al nivel de la imagen corporal después del diagnóstico y las estrategias de afrontamiento, se relacionan con casi todos los dominios de la CVRS. Durante los tratamientos, los problemas al nivel sexual y la falta de preparación para el impacto producido por el cáncer de mama, se relacionan con algunos dominios. Estos problemas pueden persistir durante varios años después del diagnóstico (Avis *et al*, 2005; Andrykowski *et al*, 2009).

En la misma línea, Arndt *et al* (2005) han concluido que déficits al nivel emocional, cognitivo, y el funcionamiento social persisten durante años en las mujeres con cáncer de mama y afectan predominantemente a las pacientes más jóvenes. Estos resultados están en conformidad con los presentados por Andrykowski *et al* (*op.cit.*) y Taguchi *et al* (2008), que además subrayan la necesidad de desarrollar sistemas de apoyo que tengan en cuenta las características individuales psicológicas, espirituales y socioeconómicas de estas mujeres.

En pacientes de cáncer colon-rectal, un año después del diagnóstico, Arndt *et al* (2004), han encontrado déficits al nivel del funcionamiento emocional y social (además de limitaciones específicas como fatiga, insomnio, diarrea, entre otros), más evidentes en los pacientes más jóvenes. En la misma línea, Sajid *et al* (2007) destacan los problemas experimentados por pacientes de cáncer colon-rectal sometidos a cirugía y subrayan la importancia (tanto para el paciente como para el médico) de evaluar CVRS en estos pacientes y desarrollar una comunicación más abierta con estos pacientes sobre las consecuencias de la cirugía en su CVRS.

### 7.3. APLICACIÓN DEL MODELO DE ZWINDERMAN A LOS DATOS OBTENIDOS MEDIANTE EL SF-36

Se ha aplicado el Modelo de Zwinderman (1990) a los datos obtenidos a través del SF-36 en el *Pre-Test* y en el *Post-Test*. Se han considerado los 132 pacientes que han completado el estudio, de los cuales 41 habían sido sometidos a cirugía sin tratamiento posterior.

Para aplicar este modelo hemos convertido los datos en binarios, codificando en cero las opciones de respuesta que no reflejaban síntomas y en uno las que reflejaban.

Cómo hemos visto, el Modelo de Zwinderman posibilita la cuantificación del cambio observado en un conjunto de pacientes, estimando el cambio debido al tratamiento y el cambio debido al paso del tiempo. Las estimaciones se han obtenido mediante un script construido en MatLab para el efecto.

Aplicando el modelo al conjunto de datos en estudio se obtienen los siguientes valores para las estimaciones de máxima verosimilitud de las componentes  $\mu$  y  $\tau$  del cambio:

$$\hat{\mu} = 0,13 \text{ (SE}_{\hat{\mu}} = 0,10) \quad \text{y} \quad \hat{\tau} = 0,03 \text{ (SE}_{\hat{\tau}} = 0,08).$$

El valor de  $\hat{\tau} = 0,03$  indica un empeoro con el paso del tiempo de la calidad de vida tanto en el grupo que ha recibido tratamientos posteriores como en el grupo que únicamente ha sido sometido a cirugía. Por otra parte, el valor de  $\hat{\mu}$  indica que el empeoro asociado a los tratamientos ha sido estimado en un 0,13. Estos valores indican un empeoro global de la CVRS que, sin embargo, es más significativo en el grupo que ha recibido tratamientos. Tal como esperábamos, los pacientes recibiendo radio o quimioterapia, poseen una percepción peor de su nivel de calidad de vida.

Además de las estimaciones del cambio asociado a la CVRS global, hemos calculado las estimaciones asociadas a cada una de las 8 dimensiones del SF-36 (Tabla 20).

Dimensiones del SF-36	$\hat{\mu}$	$SE_{\hat{\mu}}$	$\hat{\tau}$	$SE_{\hat{\tau}}$
Funcionamiento Físico	0,01	0,21	0,26	0,16
Desempeño Físico	0,15	0,30	0,54	0,22
Dolor Corporal	0,75	0,45	-1,10	0,37
Salud General	0,10	0,27	-0,26	0,21
Vitalidad	0,29	0,30	-0,05	0,22
Funcionamiento Social	0,53	0,43	-0,41	0,32
Desempeño Emocional	0,20	0,34	0,14	0,26
Salud Mental	-0,15	0,07	0,28	0,21

**Tabla 20.** Estimaciones de máxima verosimilitud de las componentes del cambio asociadas al tratamiento ( $\hat{\mu}$ ) y al tiempo ( $\hat{\tau}$ ) obtenidas mediante el Modelo de Zwinderman para las dimensiones del SF-36.

Observando los valores estimados para el **cambio debido al tratamiento**, se constata un empeoro de la calidad de vida para todas las dimensiones con excepción de la dimensión Salud Mental. Para el cambio debido al tratamiento, el deterioro más acentuado se observa en la dimensión Dolor Corporal, siguiéndose la dimensión Funcionamiento Social.

Las estimaciones del **cambio debido al paso del tiempo** reflejan una mejoría de la calidad de vida para las dimensiones Dolor Corporal, Salud General, Vitalidad y Funcionamiento Social, y un deterioro de la calidad de vida asociada a las dimensiones Funcionamiento Físico, Desempeño Físico, Desempeño Emocional y Salud Mental. La mejoría más acentuada se registró en la dimensión Dolor Corporal y el empeoro más significativo en la dimensión Desempeño Físico.

Aunque permite estimar el cambio aislando las componentes tiempo y tratamiento, el Modelo de Zwinderman no tiene en cuenta mecanismos de adaptación por parte de los pacientes. Algunos de estos cambios pueden estar sesgados por el fenómeno de *response shift*.

#### 7.4. ANÁLISIS DE LA ESTRUCTURA FACTORIAL DEL INSTRUMENTO GENÉRICO PARA LOS DIFERENTES TIEMPOS

Se ha realizado un Análisis Factorial a partir de las respuestas de los pacientes al SF-36 con el objetivo de valorar cómo ha evolucionado la estructura factorial subyacente a lo largo del proceso de tratamiento. Para ello, hemos considerado las puntuaciones asociadas al *Pre-Test*, *Post-Test* y *Then-Test* para la muestra de los 132 pacientes que completaron el estudio.

Para la extracción de los ejes se ha utilizado el Análisis de Componentes Principales. Se han retenido los 8 ejes correspondientes a las 8 dimensiones del SF-36 y se ha aplicado una rotación Varimax a la matriz de las componentes.

El KMO obtenido para las tres aplicaciones en análisis varió entre el 0,872 (*Pre-Test*) y el 0,891 (*Post-Test*) y para el Test de Bartlett se obtuvieron valores de  $p < 0,001$ , indicando una adecuación correcta de los datos al modelo de Análisis Factorial.

La Tabla 21, contiene los valores propios y varianza explicada las tres aplicaciones en estudio.

Componentes	<i>Pre-Test</i>			<i>Post-Test</i>			<i>Then-Test</i>		
	Valores Propios	% Varianza Explicada	Acumul.	Valores Propios	% Varianza Explicada	Acumul.	Valores Propios	% Varianza Explicada	Acumul.
1	15,00	42,85	42,85	15,99	45,69	45,69	17,35	49,58	49,58
2	3,95	11,28	54,13	2,81	8,02	53,71	3,31	9,45	59,03
3	2,15	6,14	60,27	2,23	6,36	60,08	1,84	5,27	64,29
4	1,45	4,15	64,42	1,50	4,30	64,37	1,58	4,52	68,81
5	1,33	3,80	68,23	1,41	4,02	68,39	1,34	3,82	72,64
6	1,11	3,18	71,41	1,22	3,49	71,88	1,13	3,23	75,86
7	1,03	2,95	74,36	1,11	3,16	75,04	0,96	2,75	78,61
8	0,90	2,57	76,93	0,97	2,78	77,82	0,89	2,55	81,16

Tabla 21. Valores Propios y Varianza Explicada asociados a las 8 componentes retenidas para el *Pre-Test*, *Post-Test* y *Then-Test*.

Se observa que, para las tres aplicaciones, la cantidad de inercia absorbida por las 8 componentes retenidas varía entre los 76,93% (para el *Pre-Test*) y los

81,16% (para el *Then-Test*), no existiendo por lo tanto diferencias acentuadas al nivel de la proporción de inercia explicada a lo largo de las tres aplicaciones.

La Tabla 22, contiene las matrices de componentes rotadas para el *Pre-Test*, *Post-Test* y *Then-Test*.

Dim	Ítems	Pre-Test								Post-Test								Then-Test								
		1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	
FF	i3a	,443	,325	,327					-,390		,616							-,370	,419	,459						,469
	i3b	,624		,570						,632	,523								,591	,572						
	i3c	,641	,370	,384						,641	,482								,653	,448						
	i3d	,767								,755									,732	,334				,396		
	i3e	,817								,805									,811							
	i3f	,752								,664									,643	,426						
	i3g	,790								,713									,694	,352				,306		
	i3h	,865								,795									,826							
	i3i	,837								,815								,367	,885							
	i3j	,597							,651	,526								,650	,719							
DF	i4a	,370		,726	,310						,840								,415	,810						
	i4b			,782	,324						,833								,336	,843						
	i4c	,368		,739	,310						,796								,313	,856						
	i4d	,393		,742							,808								,306	,840						
DC	i7					,894										,863									,882	
	i8		,307	,350		,735										,819									,875	
SG	i1	,302		,341	,506					,332				,412		,357								,568		
	i11a						,819					,343	,361			,564									,819	
	i11b				,747					,300						,579									,762	
	i11c				,422		,618									,761										,777
i11d				,722					,308		,493				,418									,404	,526	
VT	i9a		,446	,324	,615					,326				,684						,323				,631		
	i9e		,575		,515					,338		,313		,615						,492				,416		
	i9g		,724								,324	,374	,396		,334						,348			,602	,378	
	i9i		,641								,482			,407	,370					,365	,347			,614	,390	
FS	i6		,427	,331		,650				,594		,353		,365					,432		,376	,572				
	i10		,551	,418		,527				,394		,381	,374	,342					,321		,596	,354				,306
DE	i5a		,336	,779				,316				,857								,333		,789				
	i5b			,789				,315				,825								,344	,333	,768				
	i5c		,341	,674						,319		,752								,333	,333	,759				
SM	i9b		,755									,333	,717								,761					
	i9c		,782										,848								,803					
	i9d		,798									,365	,425	,574							,604					
	i9f		,819										,863								,832					
	i9h		,571	,317									,308	,755							,563					,429

Tabla 22. Matriz factorial tras rotación Varimax para el *Pre-Test*, *Post-Test* y *Then-Test*.  
 Los valores de los factores de carga menores de 0,300 no se escriben para facilitar la interpretación.

Las matrices de componentes rotadas evidencian algunos cambios en lo que respecta a la distribución de las saturaciones de los ítems por las 8 componentes. Las saturaciones de los ítems pertenecientes a la dimensión **Funcionamiento Físico** poseen una estructura más variable en el *Then-Test* que para cualquiera de las demás aplicaciones, o sea, mirando hacia atrás, los pacientes dan respuestas más variables sobre la percepción que tiene de su Funcionamiento Físico. Recuérdese que habíamos registrado la presencia de cambios Beta para la subescala de Funcionamiento Físico.

Los ítems de la subescala **Desempeño Físico** cargan sobretodo el eje 3 en el momento del *Pre-Test*, mientras en las aplicaciones siguientes, surgen asociados sobre todo al eje 2. Se destaca la estructura semejante observada entre el *Post-Test* y el *Then-Test*, lo que tiene todo el sentido si tuviéremos en cuenta que estas dos aplicaciones son hechas en el mismo instante.

En los ítems relativos al **Dolor Corporal** también se registra un cambio de eje: mientras en el *Pre-Test* cargan el eje 5, en el *Post-Test* y el *Then-Test* pasan a cargar el eje 6. De nuevo se destaca la estructura semejante entre las dos últimas aplicaciones.

Los ítems de la dimensión **Salud General** distribuyen sus cargas por varios ejes distintos para todas las aplicaciones. Asimismo, es para el *Post-Test* y el *Then-Test* que los patrones observados se parecen más entre sí, ya que el ítem 1 (“*En general, usted diría que su salud es...*”) carga el eje 5 y los demás en el eje 7, exceptuando el ítem 11c (“*Creo que mi salud va a mejorar*”) que cambia de eje.

Las saturaciones de los ítems pertenecientes a la dimensión **Vitalidad** presentan una distribución bastante irregular, haciendo difícil la identificación de un patrón. Recuérdese que habíamos registrado la presencia de cambios Beta para esta subescala.

Los dos ítems relativos al **Funcionamiento Social** presentan saturaciones más elevadas para el eje 2 en el momento del *Pre-Test*, mientras para el *Post-Test* y el *Then-Test* reparten sus cargas por varios ejes.

Las cargas de los ítems relativos al **Desempeño Emocional** se concentran en el eje 3 tanto para el *Pre-Test* como para el *Post-Test*, y se cambian al eje 4 en el *Then-Test*.

Los ítems relativos de **Salud Mental** cargan el eje 2 en el *Pre-Test*, en el *Post-Test* reparten sus cargas entre los ejes 4 y 5, mientras en el *Then-Test* concentran sus cargas en el eje 3.

En suma, hemos observado cambios de los ítems de unas dimensiones a otras, indicando la posible ocurrencia de cambio. Sin embargo no hemos detectado una estructura suficientemente clara como para identificar el tipo de *response shift* ocurrido.

## 7.5. ANÁLISIS DEL SF-36 MEDIANTE EL STATIS DUAL

En una fase siguiente se sometieron las matrices relativas a las diferentes aplicaciones (*Pre-Test*, *Post-Test* y *Then-Test*) a un STATIS Dual de forma a comparar globalmente las aplicaciones en estudio y averiguar la existencia de una tipología común a las matrices de información en análisis que posibilite la descripción, en una única estructura, de las varias aplicaciones. Para este análisis hemos recurrido al *software* estadístico ADE-4. Los datos han sido previamente centrados.

Los análisis separados de las diferentes aplicaciones han conducido a los resultados presentados en las Tabla 23, para los 132 pacientes que han completado el estudio.

Los análisis separados obtenidos parecen evidenciar que las matrices consideradas poseen estructuras factoriales similares tanto en relación a los valores de la inercia total como en relación a los valores de la inercia asociada a cada eje (el primer plano factorial explica cerca de 45% de la inercia total para las diferentes aplicaciones). Sin embargo, al calcular el coeficiente de correlación vectorial *RV* de Hilbert Schmidt entre las diferentes aplicaciones (Tabla 24), se



obtienen coeficientes de correlación que no superan el valor 0,662. No siendo altos, estos valores traducen que el grado de similitud entre las nubes correspondientes al *Pre-Test*, *Post-Test* y *Then-Test* es relativamente reducido, indicando que se registraron alteraciones en la estructura de las diferentes aplicaciones y comprometiendo así la construcción de una estructura común entre las matrices en estudio.

Componentes	<i>Pre-Test</i>			<i>Post-Test</i>			<i>Then-Test</i>		
	Valor Propio	Inercia	Inercia Acumulada	Valor Propio	Inercia	Inercia Acumulada	Valor Propio	Inercia	Inercia Acumulada
1	0,397	0,326	0,326	0,450	0,381	0,381	0,374	0,317	0,317
2	0,154	0,127	0,453	0,112	0,096	0,477	0,130	0,110	0,427
3	0,096	0,078	0,531	0,085	0,073	0,550	0,091	0,077	0,504
4	0,071	0,058	0,589	0,068	0,059	0,609	0,077	0,065	0,570
5	0,060	0,050	0,639	0,053	0,046	0,654	0,066	0,056	0,626
6	0,050	0,041	0,680	0,050	0,043	0,697	0,056	0,048	0,674
7	0,042	0,035	0,714	0,043	0,037	0,733	0,049	0,042	0,715
8	0,038	0,031	0,745	0,035	0,030	0,764	0,041	0,035	0,750

Tabla 23. Análisis Separados para el *Pre-Test*, *Post-Test* y *Then-Test*, considerando los 8 primeros ejes.

Matriz	Pre-Test	Pos-Test	Then-Test
Pre-Test	1,000		
Post-Test	0,581	1,000	
Then-Test	0,538	0,662	1,000

Tabla 24. Matriz de coeficientes RV.

Sometiendo la matriz *RV* a un ACP, se obtuvo el **análisis de la interestructura** cuya imagen euclídea se presenta en la Fig.16.

El análisis de la imagen euclídea lleva igualmente a concluir que no es posible considerar la existencia de una estructura común a las varias aplicaciones en estudio ya que los ángulos formados por los vectores asociados a cada aplicación son grandes, indicando baja correlación.

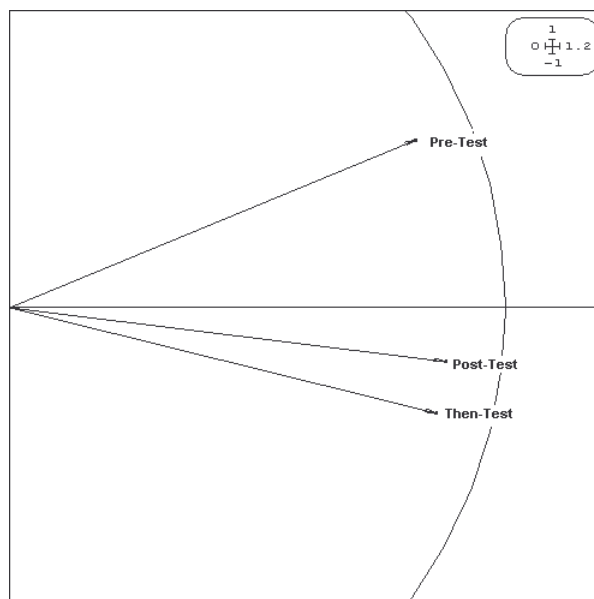


Fig.16. Gráfico de la Interestructura

Si no hubiera ocurrido cualquier tipo de cambio al nivel de la autopercepción de los pacientes, se esperaría que existiera un elevado grado de congruencia entre las matrices referentes al *Pre-Test* y *Then-Test*, ya que estas evaluaciones se reportan al mismo instante. Sin embargo, es interesante notar que el *Then-Test* posee una estructura de covariación más próxima al *Post-Test* que al *Pre-Test*. Esto se puede interpretar teniendo en cuenta que el *Post-Test* y el *Then-Test* han sido obtenidos bajo los mismos estándares de medida, mientras el *Pre-Test* se refiere a una evaluación obtenida bajo estándares distintos, evidenciando por lo tanto la ocurrencia de cambio.

## 7.6. ANÁLISIS DEL SF-36 MEDIANTE EL METABILOT

Como hemos visto, Krzanowski (1979) propone un procedimiento para comparar estructuras factoriales que está basado en la comparación de soluciones multidimensionales completas en vez de hacerse la comparación eje a eje. Este procedimiento permite comparar subespacios generados por componentes principales y es fácilmente generalizable a cualquier técnica factorial. En este apartado utilizamos representaciones Biplot para integrar información de varias matrices de datos, recurriendo al **Análisis Metabiplot** (Martín-Rodríguez *et al*, 2001). En el caso de congruencia de las estructuras en análisis, esta técnica posibilita la obtención de una configuración consenso en la cual sea posible comparar las configuraciones resultantes de análisis Biplot.

Las matrices de información relativas a las tres aplicaciones del SF-36 (*Pre-Test*, *Post-Test* y *Then-Test*) para los 132 pacientes que han completado el estudio, han sido sometidas a un Análisis Metabiplot, con el objetivo de comparar (y si posible integrar) las estructuras resultantes de los varios análisis multivariantes en un espacio consenso. El procedimiento consiste en buscar el ángulo mínimo entre las nuevas componentes del espacio consenso y las componentes principales definidas por cada grupo. Este análisis se ha llevado a efecto en el *software* Multbiplot que funciona en MatLab. Los datos han sido previamente centrados.

Al integrar las tres matrices de información en estudio en una configuración consenso, se observa que la absorción de inercia acumulada es similar para las distintas aplicaciones (Tabla 25). Para las tres aplicaciones se verifica que el primer eje absorbe la mayor parte de la inercia, siendo relativamente pequeña la variabilidad asociada a los demás ejes. Esto lleva a considerar la existencia de un espacio consenso.

Ejes								
	1	2	3	4	5	6	7	8
<b>Pre-Test</b>	40,842	47,527	55,491	61,485	67,214	70,496	72,625	75,301
<b>Post-Test</b>	45,497	50,716	56,955	61,475	66,378	70,22	73,855	76,597
<b>Then-Test</b>	45,48	50,307	56,608	62,526	66,796	71,929	75,159	77,464

**Tabla 25.** Porcentaje de inercia absorbida acumulada para cada aplicación por las componentes principales del espacio consenso.

Dado que los porcentajes de inercia absorbida presentan un patrón similar en las tres aplicaciones, se ha construido un **espacio consenso** que integra las tres matrices de información. O sea, se buscó el sistema de ejes ortogonales que mejor hace converger las principales direcciones de inercia, determinando los vectores propios que definen el espacio conjunto.

La Tabla 26 contiene los cuadrados de los cosenos de los ángulos que forman las componentes de cada aplicación por separado con las componentes del espacio consenso.

Componente								
	1	2	3	4	5	6	7	8
<b>Pre-Test</b>	0,998	0,985	0,976	0,946	0,945	0,834	0,385	0,621
<b>Post-Test</b>	0,997	0,990	0,980	0,964	0,949	0,861	0,838	0,659
<b>Then-Test</b>	0,998	0,985	0,980	0,973	0,956	0,956	0,795	0,310

**Tabla 26.** Cuadrado de los cosenos de los ángulos formados entre las componentes principales que definen cada grupo y los ejes de la representación conjunta.

Los cuadrados de los cosenos de los ángulos formados por las componentes que definen cada grupo y las componentes del espacio consenso, pueden interpretarse como una medida de similitud entre el subespacio de las componentes principales de cada grupo y las de la representación conjunta. Se observa que las tres aplicaciones presentan muy alta similitud con respecto a las componentes consenso, aunque la séptima componente consenso esté más alejada del *Pre-Test* y la octava componente consenso esté más alejada del grupo *Then-Test*.

Además, se ha buscado la similitud entre las componentes originales por separado y las nuevas componentes del espacio consenso observando los cosenos

de los ángulos entre estas componentes (Tabla 27). Se observa que, para las tres aplicaciones, la componente 1 presenta alta correlación con el eje 1 de la estructura consenso, lo que está de acuerdo con lo que ya habíamos encontrado para el STATIS Dual. Sin embargo, observando las demás componentes, se detectan cambios en el patrón observado. Nótese que los últimos ejes recogen menos variabilidad, sin embargo no tienen necesariamente menos importancia. En contextos de CVRS, aspectos claves pueden aparecer asociados a los últimos ejes. Estos resultados pueden interpretarse como fuerte evidencia de la ocurrencia de *response shift*.

		Componentes Comunes							
		1	2	3	4	5	6	7	8
Pre-Test	Componente 1	<b>0,976</b>	0,088	0,167	-0,063	-0,004	0,035	0,001	0,044
	Componente 2	-0,183	0,363	0,548	<b>-0,550</b>	0,358	0,200	0,024	0,039
	Componente 3	0,031	<b>-0,818</b>	0,284	0,001	0,460	-0,102	-0,017	0,008
	Componente 4	0,104	0,139	-0,648	-0,198	<b>0,626</b>	-0,008	0,089	-0,243
	Componente 5	0,020	-0,393	-0,303	<b>-0,554</b>	-0,364	0,445	0,067	0,032
	Componente 6	0,004	-0,058	0,201	-0,124	-0,23	-0,231	0,351	<b>-0,692</b>
	Componente 7	0,011	0,003	0,122	0,269	0,031	<b>0,627</b>	-0,187	-0,279
	Componente 8	-0,02	0,022	0,031	0,455	0,164	0,373	<b>0,463</b>	-0,026
Post-Test	Componente 1	<b>0,991</b>	0,043	0,107	0,037	0,014	0,010	-0,003	-0,023
	Componente 2	0,106	-0,581	<b>-0,693</b>	0,139	-0,101	0,019	0,206	0,207
	Componente 3	0,020	-0,065	-0,009	<b>-0,671</b>	0,529	0,352	0,236	0,108
	Componente 4	0,002	0,003	0,109	-0,289	<b>-0,754</b>	0,445	0,248	-0,097
	Componente 5	-0,031	-0,627	<b>0,648</b>	-0,053	-0,063	-0,266	0,140	0,209
	Componente 6	0,007	-0,357	0,023	-0,037	-0,002	0,389	<b>-0,764</b>	0,053
	Componente 7	-0,055	-0,106	0,220	<b>0,617</b>	0,279	0,536	0,270	-0,208
	Componente 8	-0,013	0,339	0,087	0,158	-0,088	0,168	-0,051	<b>0,710</b>
Then-Test	Componente 1	<b>0,994</b>	0,016	0,089	0,018	0,006	0,002	-0,019	0,003
	Componente 2	-0,041	0,366	0,541	<b>-0,611</b>	0,190	0,293	0,177	0,124
	Componente 3	0,066	0,028	<b>-0,679</b>	-0,275	0,131	0,634	-0,059	-0,040
	Componente 4	-0,023	<b>0,560</b>	0,059	0,514	-0,451	0,378	0,128	0,134
	Componente 5	-0,011	0,275	-0,066	0,407	<b>0,818</b>	-0,049	0,096	0,148
	Componente 6	-0,025	<b>-0,667</b>	0,342	0,276	0,100	0,522	0,182	0,130
	Componente 7	-0,049	0,036	0,216	0,096	0,077	0,180	<b>-0,832</b>	-0,190
	Componente 8	-0,002	0,122	0,216	0,093	0,118	0,131	-0,097	<b>-0,447</b>

Tabla 27. Cosenos de los ángulos entre las componentes originales y las componentes del espacio conjunto. Se han señalado los cosenos de valor absoluto superior a 0,5 y, para cada componente, se ha señalado a negrita el coseno correspondiente al mayor ángulo.

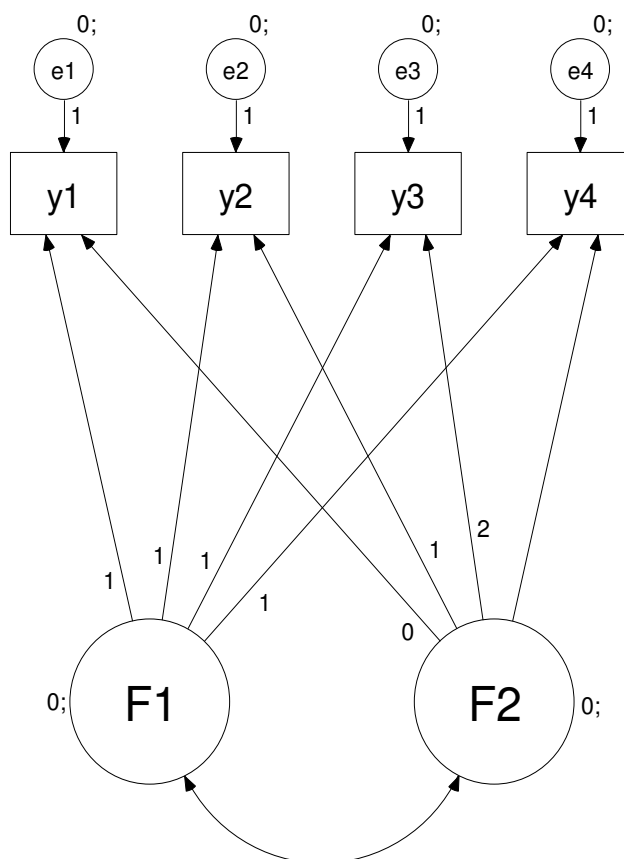
## 7.7. ANÁLISIS DEL SF-36 MEDIANTE UN MODELO DE CURVAS DE CRECIMIENTO LATENTE

Los Modelos de Curvas de Crecimiento son utilizados en estudios longitudinales para analizar las trayectorias de los individuos a lo largo del tiempo. En la literatura son varias las referencias de aplicación de estos modelos a contextos de Calidad de Vida. Beal *et al* (2007) han aplicado un Modelo de Curvas de Crecimiento Latente para evaluar las trayectorias al nivel de la depresión en un grupo de pacientes con esclerosis múltiple. Duncan & McAuley (1993) y McAuley *et al* (2000) han utilizado esta técnica para analizar las trayectorias de adultos mayores a participar en programas de actividad física, al nivel de sus sentimientos de bienestar, satisfacción con la vida y soledad a lo largo del tiempo. Westhovens *et al* (2006) ha utilizado el Análisis de Curvas de Crecimiento para evaluar cambio en la calidad de vida de pacientes con artritis reumatoide sometidos a un tratamiento específico.

En este estudio, se ajustó un modelo de Curva de Crecimiento Latente a cada una de las dimensiones del SF-36, con el objetivo de analizar el cambio en las diferentes subescalas a lo largo del tiempo, en una tentativa de aportar algo más a los análisis anteriores. Para el análisis subsecuente hemos considerado únicamente los 98 individuos que han completado el *Pre-Test*, 2ª Aplicación, *Post-Test* y *Then-Test*. Se han utilizado las puntuaciones medias no estandarizadas obtenidas en cada subescala. Se han sustituido los datos faltantes por la media obtenida en la subescala en análisis. El *software* utilizado ha sido el programa AMOS 5.

Se ha empezado por construir el modelo de curva de crecimiento latente presentado en la Fig.17. Las variables  $y_1$ ,  $y_2$ ,  $y_3$  y  $y_4$  representan las puntuaciones observadas para cada una de las dimensiones del SF-36 en estudio:

Funcionamiento Físico (FF), Desempeño Físico (DF), Dolor Corporal (DC), Salud General (SG), Vitalidad (VT), Funcionamiento Social (FS), Desempeño Emocional (DE) o Salud Mental (SM). Los índices 1, 2, 3 y 4 dicen respecto al *Pre-Test*, 2ª Aplicación, *Post-Test* y *Then-Test*, respectivamente.



**Fig.17.** Modelo de Curva de Crecimiento Latente para las dimensiones del SF-36. Las variables  $y_1$ ,  $y_2$ ,  $y_3$  y  $y_4$  representan las puntuaciones observadas en el *Pre-Test*, 2ª Aplicación, *Post-Test* y *Then-Test*, respectivamente.

La variable latente **F1** representa el **intercepto**, o sea, los valores iniciales. La variable latente **F2** representa la **pendiente**, o sea, la tasa de crecimiento. Se han fijado todos los coeficientes de regresión en 1 para el intercepto (F1). Mientras para la pendiente (F2) se han fijado en 0, 1 y 2 los coeficientes relativos al *Pre-Test*, 2ª Aplicación y *Post-Test*, de forma a que el coeficiente relativo al *Then-Test* pudiera ser libremente estimado. Las restantes restricciones se han impuesto por cuestiones de estimación. La **varianza de F1** se interpreta como la variabilidad

individual en el nivel inicial, mientras la **varianza de F2** traduce la variabilidad individual en las trayectorias dentro del grupo. Una covarianza positiva entre F1 y F2 significa que individuos con altos valores iniciales (interceptos) registran aumentos de puntuación (pendiente) más intensos, mientras los individuos que empiezan con puntuaciones más bajas tienden a registrar crecimientos más lentos en las puntuaciones y viceversa.

Los parámetros estimados al ajustar el modelo de curva de crecimiento para cada una de las dimensiones del SF-36 se encuentran en la Tabla 28. Los modelos resultantes presentan valores de RMSEA satisfactorios, indicando una buena calidad de ajuste.

Dim	Var(F1)	Var(F2)	Corr(F1,F2)	Var(e1)	Var(e2)	Var(e3)	Var(e4)	GL	$\chi^2$	RMSEA	ECVI
FF	21,88**	4,35**	-0,63*	3,29	17,67**	8,70	12,33**	6	3,33	0,000	0,199
DF	26,38**	4,15*	-0,83**	1,29	10,78**	18,28**	19,47**	6	0,56	0,000	0,171
DC	5,09**	1,12**	-0,68**	0,09	3,33**	2,55**	4,02**	6	1,77	0,000	0,163
SG	9,06**	1,46*	-0,49	1,66	2,17**	6,50**	10,06**	6	3,68	0,000	0,203
VT	10,79**	1,04	-0,36	4,45*	6,91**	5,63**	11,85**	6	2,54	0,000	0,191
FS	2,80**	0,28	-0,61	2,38**	1,93**	2,86**	2,75**	6	7,53	0,051	0,243
DE	10,00**	1,34	-0,67*	2,49*	8,36**	7,62**	9,29**	6	9,07	0,055	0,238
SM	18,06**	0,90	-0,47	4,79*	7,90**	10,95**	12,77**	6	5,99	0,000	0,206

**Tabla 28.** Parámetros estimados y bondad de ajuste para el modelo de Curva de Crecimiento Latente para las dimensiones del SF-36 (\*\*p-valor<0,01; \*p-valor<0,05)

La varianza de F1 es estadísticamente significativa para todas las dimensiones del SF-36, indicando variabilidad individual significativa en el *Pre-Test*, o sea, se ha registrado una variabilidad significativa en lo que respecta a la percepción que los pacientes tenían de su calidad de vida en todas las dimensiones. Relativamente a la varianza de F2, únicamente se obtienen valores estadísticamente significativos para las dimensiones asociadas a la Salud Física (FF, DF, DC y SG), indicando variabilidad significativa en la tasa de crecimiento de las puntuaciones, o sea, el cambio de puntuaciones a lo largo del estudio ha sido mucho más significativo en las dimensiones asociadas a la Salud Física que a las asociadas a la Salud Mental.



Para todas las dimensiones se obtienen valores negativos para la correlación entre F1 y F2, significando que, individuos que en el *Pre-Test* registran puntuaciones más bajas en las dimensiones del SF-36 (interceptos bajos), tienden a registrar variaciones más acentuadas. Por otra parte, individuos que empiezan con puntuaciones más elevadas, tienden a registrar variaciones de puntuación menos visibles. Se destaca esta tendencia en las dimensiones FF, DF, DC y DE para las cuales se obtienen correlaciones más elevadas y estadísticamente significativas. Estos resultados se explican teniendo en cuenta que, probablemente, estas dimensiones han sido las más visiblemente afectadas durante el período recogida de los datos. En el momento de las dos últimas aplicaciones, una gran parte de los pacientes se encontraba en fase de tratamientos, recibiendo radio o quimioterapia, haciendo con que se sintieran físicamente más débiles.

La varianza de los errores se interpreta como la cantidad de variabilidad que no es explicada por el modelo, asumiendo, en general, valores más elevados para el *Then-Test*. Si por una parte las evaluaciones iniciales tienden a ser homogéneas, a medida que el tiempo pasa, las evaluaciones de los pacientes se vuelven más heterogéneas.

El Análisis de Curvas de Crecimiento Latente puede constituir una herramienta eficaz para analizar la evolución de las puntuaciones asociadas a un conjunto de variables a lo largo del tiempo, posibilitando una evaluación global de las trayectorias de los pacientes relativamente a la percepción que poseen de su calidad de vida. Sin embargo, esta metodología no tiene en cuenta cualquier tipo de mecanismo de adaptación o *response shift* por parte de los individuos, muy frecuentes en estos contextos.

## 7.8. ANÁLISIS DEL SF-36 MEDIANTE EL MODELO DE ECUACIONES ESTRUCTURALES

Utilizando las puntuaciones obtenidas tras la aplicación del instrumento genérico, hemos recurrido a los Modelos de Ecuaciones Estructurales para detectar *response shift* a lo largo de las varias aplicaciones. En este análisis hemos considerado los 132 individuos que han completado el *Pre-Test* y *Post-Test* y hemos utilizado las puntuaciones medias obtenidas en cada subescala. Se han sustituido los datos faltantes por la media obtenida en la subescala en análisis. Para ajustar el Modelo de Ecuaciones Estructurales se ha utilizado el programa AMOS 5.

Hemos empezado con el cálculo de medidas de consistencia interna para las diferentes aplicaciones y subescalas y el índice *d* de Cohen (1988), atrás definido, cuyos resultados se muestran en la Tabla 29. Recordamos que un valor de *d* entre 0,2 y 0,5 indica un efecto pequeño, entre 0,5 y 0,8 moderado y mayor que 0,8 grande.

Subescala	Pre-Test			Post-Test			Índice <i>d</i> ( Post – Pre )
	Média	Desv. Est.	Alpha Cron.	Média	Desv. Est.	Alpha Cron.	
Funcionamiento Físico	2,39	0,54	0,926	2,31	0,54	0,924	-0,140
Desempeño Físico	3,30	1,37	0,955	2,89	1,32	0,955	-0,307*
Dolor Corporal	3,62	1,19	0,843	3,88	1,19	0,889	0,216*
Salud General	3,12	0,73	0,745	3,26	0,83	0,794	0,181
Vitalidad	3,17	1,01	0,842	3,23	1,01	0,861	0,064
Funcionamiento Social	3,92	1,19	0,825	4,06	1,08	0,811	0,120*
Desempeño Emocional	3,58	1,19	0,903	3,51	1,23	0,941	-0,057
Salud Mental	3,51	0,98	0,875	3,56	0,99	0,876	0,044

Tabla 29. Medias, desviación estándar, alpha de Cronbach y índice *d*, obtenidos en las diferentes aplicaciones de las subescalas del SF-36. Nota: \* $p < 0,05$  en el Test t para muestras pareadas.

Los índices de consistencia interna son bastante satisfactorios para todas las subescalas. Los efectos detectados indican diferencias significativas entre el *Post-*

*Test* y el *Pre-Test* para las subescalas de Desempeño Físico y Dolor Corporal, aunque de baja magnitud. En particular, las diferencias observadas muestran un empeoro significativo al nivel del Desempeño Físico y una mejoría al nivel del Dolor Corporal percibido.

Para evaluar la sensibilidad de las subescalas del SF-36, Liang *et al* (1985) sugieren comparar el cuadrado del valor absoluto del estadístico *t* para las diferencias para cada una de las subescalas, relativamente a la subescala con mayor valor *t*. Esta comparación refleja en simultáneo la magnitud del cambio a lo largo del tiempo y la variabilidad en cada subescala. El estadístico de eficiencia relativa  $(T1/T2)^2$  descrito permite identificar cual de las subescalas, en cada instante, es más sensible. Los resultados obtenidos para nuestra muestra, se presentan en la Tabla 30. Para la diferencia entre el *Post-Test* y la el *Pre-Test*, la subescala que revela mayor sensibilidad es la de Desempeño Físico, siguiéndose las subescalas de Dolor Corporal y Salud General. La subescala que presenta menor sensibilidad es la de Salud Mental.

Subescala	Post-Test / Pre-Test
Funcionamiento Físico	0,286
Desempeño Físico	1,000
Dolor Corporal	0,491
Salud General	0,452
Vitalidad	0,053
Funcionamiento Social	0,156
Desempeño Emocional	0,035
Salud Mental	0,029

**Tabla 30.** Estadísticos de eficiencia relativa para la sensibilidad de las subescalas del SF-36 relativamente a las diferencias entre el y *Post-Test* y *Pre-Test*.

En un grupo de pacientes de osteoartritis, a los que se les aplicó el SF-36 antes y después de un tratamiento que ha durado entre 3 y 4 semanas, Angst *et al*

(2001b) han encontrado que la subescala de Dolor Corporal era la más sensible, siguiéndose la de Desempeño Físico.

Los fundamentos teóricos del SF-36 (Ware *et al*, 1993) y los resultados del Análisis Factorial Confirmatorio han permitido construir los modelos de medida para los diferentes tiempos en estudio. Las Fig.18 y Fig.19 contienen los diagramas causales (*path diagrams*) que han conducido a una mejor calidad de ajuste para el *Pre-Test* y *Post-Test*, respectivamente.

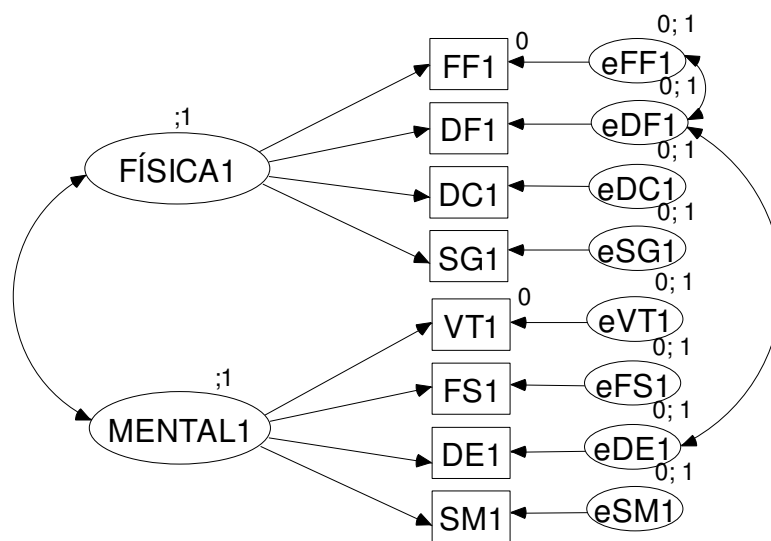


Fig.18. Diagrama Causal para la estructura de factores y relaciones en el SF-36 - *Pre-Test*.

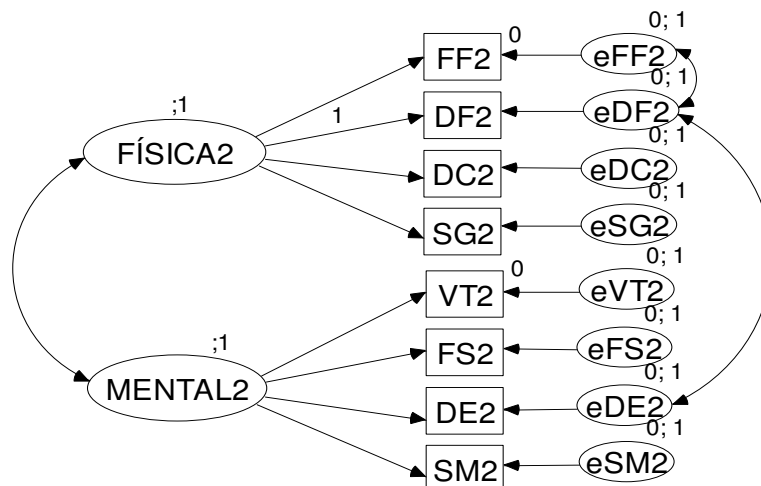


Fig.19. Diagrama Causal para la estructura de factores y relaciones en el SF-36 - *Post-Test*.

Como hemos dicho, los círculos representan variables no observadas, latentes y los cuadrados representan las variables observadas. Las variables latentes son:

- la Salud **FISICA**, que es medida por el Funcionamiento Físico (**FF**), Desempeño Físico (**DF**), Dolor Corporal (**DC**) y Salud General (**SG**);
- la Salud **MENTAL**, que es medida por la Vitalidad (**VT**), Funcionamiento Social (**FS**), Desempeño Emocional (**DE**) y Salud Mental (**SM**).

Las demás variables latentes son los factores residuales **eFF**, **eDF**,..., **eSM** que representan todo lo que es específico en FF, DF,..., SM, más las variaciones producidas por los errores aleatorios.

Para que el modelo resulte totalmente identificado, se ha asumido que todas las variables latentes tienen varianzas unitarias. Además, para el *Post-Test* se ha fijado a la unidad la carga de la variable latente FISICA relativa a la variable DF. El criterio utilizado ha sido siempre elegir el modelo que conducía a mejor calidad de ajuste desde que estuviera en concordancia con los principios teóricos conocidos del SF-36. Nótese que, ambos modelos contienen el tratamiento de la covarianza entre FF y DF y entre DF y DE, indicando que, tanto en el *Pre-Test* como en el *Post-Test*, la percepción que los pacientes tenían de su Desempeño Físico se correlaciona con la percepción relativa tanto al Funcionamiento Físico como al Desempeño Emocional.

Se estimaron los modelos de ecuaciones estructurales presentados a los diferentes instantes en estudio. Se han obtenido los modelos de medida de la Salud Física y Mental contenidos en las Fig.20 y Fig.21 donde se muestran las estimaciones de máxima verosimilitud para las cargas de los factores comunes y las covarianzas de los factores comunes. Las Tabla 31 y Tabla 32 contienen los

coeficientes de regresión no estandarizados para las variables latentes. Nótese que, para todos los casos, el valor de la razón crítica (C.R.) supera el valor 2,58, indicando que los coeficientes de regresión son altamente significativos ( $p < 0,01$ ).

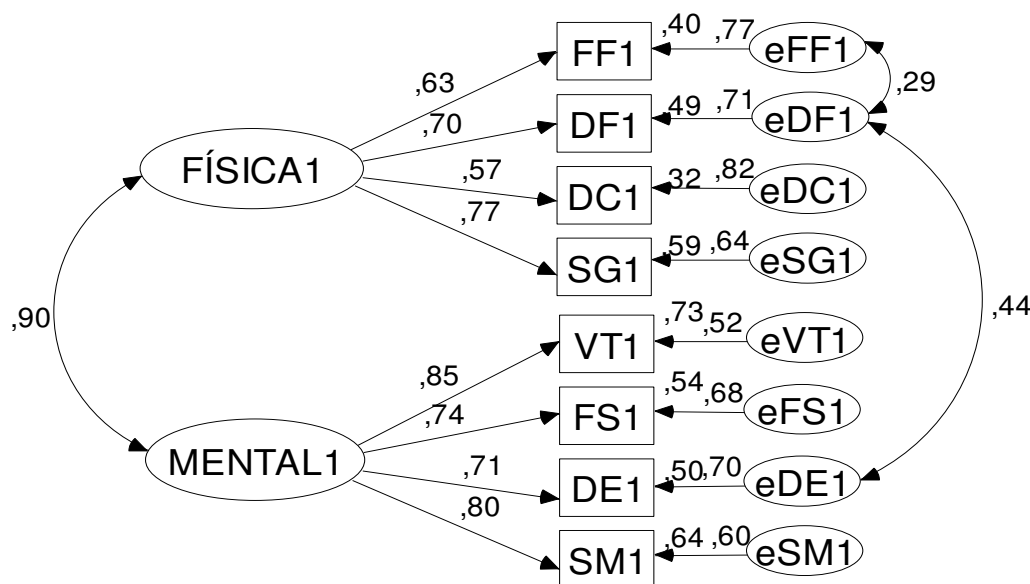


Fig.20. Modelo de Medida para el *Pre-Test* (RMSEA=0,088; Chi-square = 34,301; DF=17)

	Estimate	S.E.	C.R.	P	Label
DF1 <--- FÍSICA1	,949	,111	8,531	***	par_2
DC1 <--- FÍSICA1	,674	,103	6,559	***	par_3
SG1 <--- FÍSICA1	,560	,058	9,623	***	par_4
FF1 <--- FÍSICA1	,340	,046	7,378	***	par_5
DF1 <--- eDF1	,961	,070	13,770	***	par_6
DC1 <--- eDC1	,975	,066	14,733	***	par_7
SG1 <--- eSG1	,463	,040	11,486	***	par_8
VT1 <--- eVT1	,523	,052	10,111	***	par_9
FS1 <--- MENTAL1	,872	,094	9,311	***	par_10
DE1 <--- MENTAL1	,838	,094	8,879	***	par_11
SM1 <--- MENTAL1	,782	,074	10,498	***	par_12
VT1 <--- MENTAL1	,858	,074	11,561	***	par_13
FF1 <--- eFF1	,413	,030	13,743	***	par_14
SM1 <--- eSM1	,590	,048	12,250	***	par_15
FS1 <--- eFS1	,802	,060	13,415	***	par_16
DE1 <--- eDE1	,833	,060	13,869	***	par_17

Tabla 31. Coeficientes de Regresión no estandarizados relativos al *Pre-Test*.

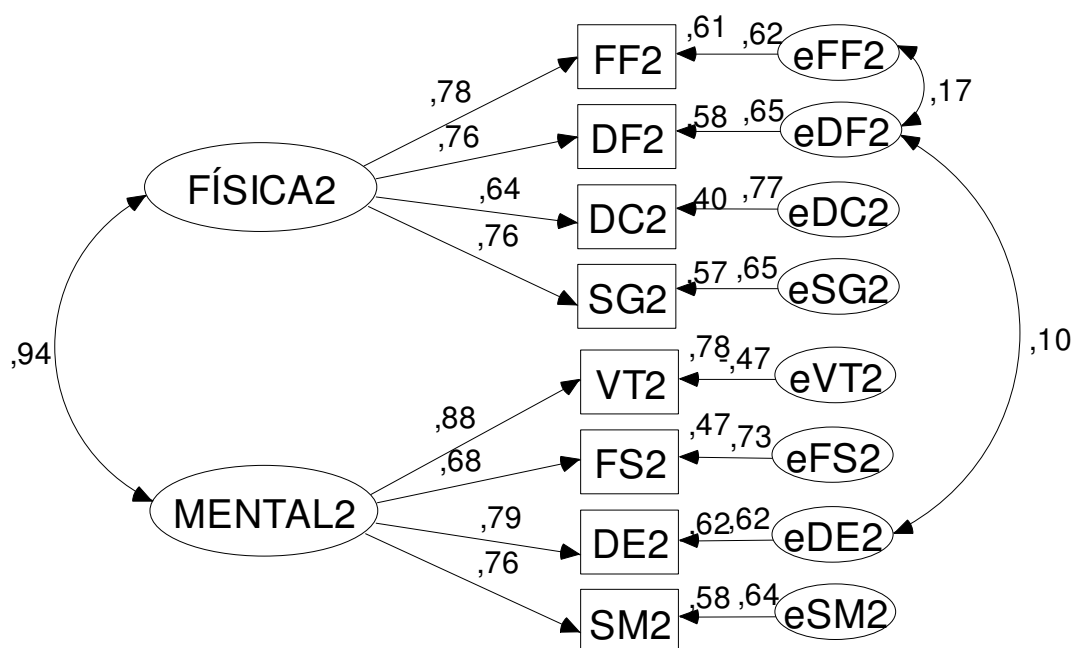


Fig.21. Modelo de Medida para el *Post-Test* (RMSEA=0,053; Chi-square = 24,635; DF=18).

	Estimate	S.E.	C.R.	P	Label
DF2 <--- FÍSICA2	1,000				
DC2 <--- FÍSICA2	,753	,094	8,046	***	par_2
SG2 <--- FÍSICA2	,622	,060	10,370	***	par_3
FF2 <--- FÍSICA2	,417	,036	11,620	***	par_4
DF2 <--- eDF2	,847	,066	12,857	***	par_5
DC2 <--- eDC2	,916	,062	14,776	***	par_6
SG2 <--- eSG2	,538	,041	13,102	***	par_7
VT2 <--- eVT2	-,470	,047	-10,074	***	par_8
FS2 <--- MENTAL2	,732	,082	8,943	***	par_9
DE2 <--- MENTAL2	,964	,085	11,386	***	par_10
SM2 <--- MENTAL2	,754	,071	10,624	***	par_11
VT2 <--- MENTAL2	,892	,065	13,777	***	par_12
FF2 <--- eFF2	,333	,027	12,189	***	par_13
SM2 <--- eSM2	,635	,046	13,762	***	par_14
FS2 <--- eFS2	,783	,053	14,646	***	par_15
DE2 <--- eDE2	,759	,057	13,200	***	par_16

Tabla 32. Coeficientes de Regresión no estandarizados relativos al *Post-Test*.

Como medidas de la calidad de ajuste, se presentan los valores de la prueba ji-cuadrado y el RMSEA. El ji-cuadrado es una medida de la calidad global de ajuste, mientras la raíz del error cuadrático medio de aproximación (RMSEA) se basa en la discrepancia de la población. La función de discrepancia se obtiene al ajustar el modelo a los momentos de la población en lugar de los momentos de la muestra. Browne & Cudeck (1993) consideran que valores de RMSEA próximos o inferiores a 0,05 indican un buen ajuste del modelo en relación a sus grados de libertad, valores próximos a 0,08 indican errores de aproximación razonables y valores iguales o mayores que 0,1 deben conducir a rechazar el modelo.

Observando los modelos de medida obtenidos, no se registran cambios significativos en los patrones de carga, lo que indica que, para las tres aplicaciones, los conceptos son definidos por las mismas variables. Una vez que no se observan cambios en el patrón de las cargas de los factores comunes no existe evidencia de haber ocurrido **reconceptualización**.

Además, se observa que, relativamente a la dimensión Salud Física, en el *Pre-Test*, la variable Salud General es la más indicativa, mientras en el *Post-Test* son las variables Funcionamiento Físico y Desempeño Físico, respectivamente. Esto nos lleva a considerar que, mientras en el *Pre-Test* era la dimensión Salud General la más importante en la definición de Salud Física, en el momento del *Post-Test*, las dimensiones Funcionamiento Físico y Desempeño Físico pasan a ser las más importantes.

Por otra parte, en lo que respecta a la dimensión Salud Mental, la variable Vitalidad es la más indicativa para las dos aplicaciones, siguiéndose la dimensión Salud Mental en el momento del *Pre-Test* y la Dimensión Desempeño Emocional para el *Post-Test*. Por lo tanto, mientras en el momento del *Pre-Test* los ítems relativos a Salud Mental asumen mayor importancia que los relativos al Desempeño Emocional en la medición del constructo Salud Mental, en los momentos de evaluación subsiguientes este orden se invierte. Estos resultados



podrán sugerir **repriorización** una vez que, si la magnitud de la carga de una determinada variable observada cambia de un tiempo a otro, entonces la variable ha pasado a ser más o menos indicativa del concepto en cuestión, o sea, la variable ha pasado a ser más o menos importante en la medición del concepto.

Teniendo en cuenta los modelos considerados para cada instante, se ha construido un **modelo longitudinal** donde se considera conjuntamente el *Pre-Test* y el *Post-Test*. En la Fig.22 se presenta el respectivo diagrama causal.

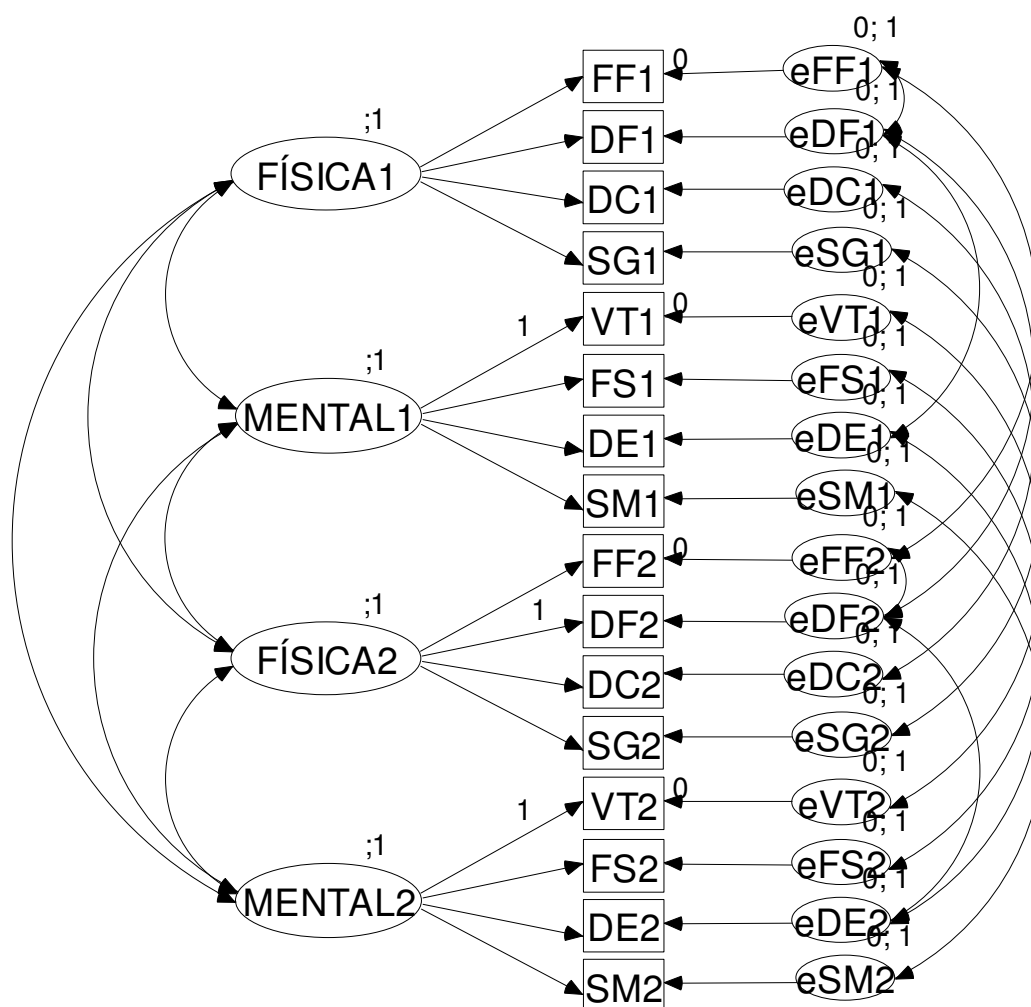


Fig.22. Diagrama Causal para la estructura de factores y relaciones en el SF-36 - *Pre-Test* y *Post-Test*.

Las estimaciones de máxima verosimilitud obtenidas para las cargas de los factores comunes y las covarianzas de los factores comunes del modelo longitudinal que considera conjuntamente el *Pre-Test* y el *Post-Test* se muestran en la Fig.23.

La Tabla 33 contiene los coeficientes de regresión no estandarizados para las variables latentes. Nótese que, para todos los casos, el valor de la razón crítica (C.R.) supera el valor 2,58, indicando que los coeficientes de regresión son altamente significativos ( $p < 0,01$ ).

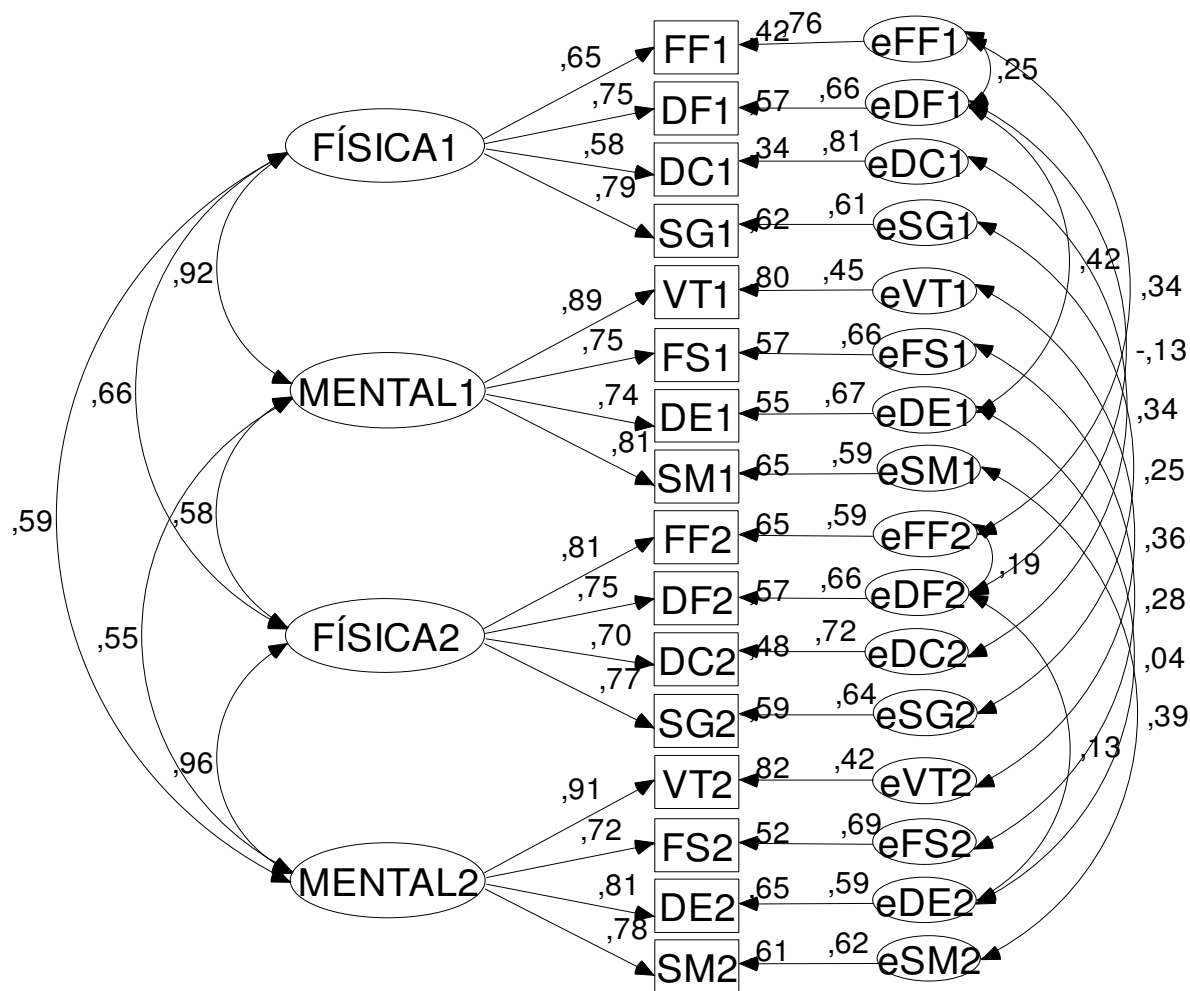


Fig.23. Modelo de Medida para el *Pre-Test* y *Post-Test* (RMSEA=0,074; Chi-square=152,995; DF=89)

	Estimate	S.E.	C.R.	P	Label
DF1 <--- FÍSICA1	1,055	,110	9,581	***	par_2
DC1 <--- FÍSICA1	,716	,106	6,774	***	par_3
SG1 <--- FÍSICA1	,599	,056	10,622	***	par_4
FF1 <--- FÍSICA1	,352	,046	7,712	***	par_5
SG1 <--- eSG1	,464	,039	12,041	***	par_6
FS1 <--- eFS1	,817	,060	13,720	***	par_7
DE1 <--- eDE1	,833	,060	13,991	***	par_8
FS1 <--- MENTAL1	,938	,092	10,253	***	par_9
DE1 <--- MENTAL1	,918	,094	9,712	***	par_10
SM1 <--- MENTAL1	,810	,069	11,816	***	par_11
VT1 <--- MENTAL1	1,000				
DF1 <--- eDF1	,925	,069	13,329	***	par_12
SM1 <--- eSM1	,597	,048	12,536	***	par_13
VT1 <--- eVT1	,506	,052	9,818	***	par_14
DC1 <--- eDC1	,995	,067	14,877	***	par_15
FF1 <--- eFF1	,410	,029	14,346	***	par_16
DF2 <--- FÍSICA2	1,000				
DC2 <--- FÍSICA2	,880	,096	9,164	***	par_18
SG2 <--- FÍSICA2	,651	,060	10,832	***	par_19
FF2 <--- FÍSICA2	,446	,035	12,659	***	par_20
SG2 <--- eSG2	,547	,040	13,521	***	par_21
FS2 <--- eFS2	,778	,053	14,647	***	par_22
DE2 <--- eDE2	,755	,057	13,261	***	par_23
FS2 <--- MENTAL2	,815	,081	10,071	***	par_24
DE2 <--- MENTAL2	1,029	,084	12,272	***	par_25
SM2 <--- MENTAL2	,804	,067	12,006	***	par_26
VT2 <--- MENTAL2	1,000				
DF2 <--- eDF2	,874	,065	13,439	***	par_27
SM2 <--- eSM2	,636	,046	13,737	***	par_28
VT2 <--- eVT2	,468	,047	10,034	***	par_29
DC2 <--- eDC2	,909	,062	14,758	***	par_30
FF2 <--- eFF2	,328	,026	12,471	***	par_31

**Tabla 33.** Coeficientes de Regresión no estandarizados relativos al modelo conjunto del *Pre-Test* con el *Post-Test*.

En la Tabla 34 se presentan las matrices de correlaciones relativas a las tres aplicaciones. La dimensión Vitalidad relativa al *Post-Test* es la que presenta correlaciones más elevadas con las demás variables, también relativas al *Post-Test*. Se destaca también que, en general, las variables relativas al *Post-Test* presentan correlaciones más fuertes entre sí que las relativas al primer instante de medida.

	VT2	SM2	DE2	FS2	FF2	SG2	DC2	DF2	VT1	SM1	DE1	FS1	FF1	SG1	DC1	DF1
<b>VT2</b>	1,00															
<b>SM2</b>	,672	1,00														
<b>DE2</b>	,671	,625	1,00													
<b>FS2</b>	,599	,550	,595	1,00												
<b>FF2</b>	,674	,509	,567	,541	1,00											
<b>SG2</b>	,639	,569	,533	,388	,629	1,00										
<b>DC2</b>	,552	,446	,535	,370	,447	,504	1,00									
<b>DF2</b>	,674	,547	,611	,455	,664	,564	,482	1,00								
<b>VT1</b>	,451	,423	,275	,301	,394	,408	,278	,303	1,00							
<b>SM1</b>	,318	,519	,289	,250	,388	,306	,286	,273	,697	1,00						
<b>DE1</b>	,327	,441	,338	,358	,365	,307	,244	,281	,570	,584	1,00					
<b>FS1</b>	,218	,332	,152	,339	,317	,235	,164	,219	,583	,625	,583	1,00				
<b>FF1</b>	,319	,196	,207	,265	,511	,329	,216	,266	,479	,378	,465	,393	1,00			
<b>SG1</b>	,403	,395	,310	,211	,463	,487	,294	,343	,661	,530	,436	,486	,515	1,00		
<b>DC1</b>	,116	,118	-,037	,078	,100	,168	,320	,145	,473	,318	,401	,421	,389	,417	1,00	
<b>DF1</b>	,430	,351	,416	,382	,485	,362	,368	,326	,595	,432	,686	,458	,629	,501	,421	1,00

Condition number = 53,181; Eigenvalues: 7,246 2,365 1,009 ,886 ,807 ,578 ,480 ,452 ,413 ,382 ,344 ,254 ,248 ,213 ,188 ,136

**Tabla 34.** Matriz de Correlaciones Muestrales de las variables endógenas para el modelo conjunto del *Pre-Test* y *Post-Test*.

Considerando los datos relativos al *Pre-Test* y *Post-Test*, se ha aplicado el procedimiento descrito por Oort (2005) y Oort *et al* (2005) para detectar *response shift* y medir cambios Alpha, anteriormente sintetizado en la Tabla 2. Siguiendo los pasos especificados, se utilizó el Modelo de Ecuaciones Estructurales para investigar las hipótesis de invarianza que han sido anteriormente definidas (Tabla 1).

***Paso 1 - Establecer el Modelo de Medida***

El primer paso consiste en establecer el modelo. Los únicos requisitos son que el modelo ajuste bien las medias y covarianzas observadas y que tenga una interpretación clara. Oort (2005) se refiere a este modelo como Modelo 1. Las restricciones que proporcionan escalas y orígenes para las variables no observadas (o sea, los factores comunes FÍSICA y MENTAL)  $Diag(\Phi_{11})= Diag(\Phi_{22})=I$ .

Si los patrones de las cargas factoriales en las matrices  $\Gamma_1$  y  $\Gamma_2$  del Modelo 1 fueran tan distintos que no sea posible identificar cualesquiera factores que mantengan su significado a lo largo de los tiempos, entonces se debería concluir que la **reconceptualización** ha sido tan fuerte que no es posible seguir probando

otros tipos de cambio. Sin embargo, si podemos encontrar un modelo con buena calidad de ajuste, con patrones  $\Gamma_1$  y  $\Gamma_2$  que en gran medida son los mismos y con factores interpretables, entonces, se puede seguir investigando *response shift*.

En el Modelo de Medida presentado en la Fig.22 todos los parámetros son libres para ser estimados (excepto los que se han fijado por razones de identificación del modelo). Este modelo corresponde al **Modelo 1** de Oort (2005). En este modelo no se imponen restricciones a lo largo del tiempo. La prueba ji-cuadrado de bondad de ajuste resulta significativa y el RMSEA encontrado indica una calidad de ajuste razonable (RMSEA=0,074).

Los índices obtenidos para el Modelo 1 indican buen ajuste. Los patrones de carga para ambos los períodos son similares; por lo tanto, ambos los tiempos miden los mismos conceptos y se puede entonces evaluar la ocurrencia de *response shift*.

### ***Paso 2 - Prueba Global de Response Shift***

El segundo paso consiste en ajustar un modelo "no *response shift*" que Oort (2005) refiere como Modelo 2, y realizar una prueba global de "no *response shift*", comparando los ajustes de los Modelos 1 y 2. En el Modelo 2 impusimos todas las restricciones de las hipótesis de invarianza que están asociadas con *response shift* (Tabla 2).

Si el Modelo 2 no es significativamente peor que el Modelo 1, entonces, aparentemente, todos los cambios en las medias observadas  $\mu$  y covarianzas  $\Sigma$  pueden explicarse totalmente por cambios en las medias ( $\alpha$ ) y covarianzas ( $\Phi$ ) de los factores comunes. La diferencia en las medidas  $\chi^2$  de ajuste de los Modelos 1 y 2 pueden pues ser utilizadas como una prueba global de "no *response shift*". Si esta diferencia no resulta significativa, se puede concluir que no ha ocurrido *response shift* y saltar el Paso 3.

El diagrama causal para este modelo se encuentra definido en la Fig. 24. Las estimaciones de máxima verosimilitud para las cargas de los factores comunes y las covarianzas de los factores comunes se muestran en Fig. 25. El modelo encontrado no es satisfactorio (RMSEA=0,207). El hecho de haber encontrado un modelo significativamente peor indica que el modelo no es similar para ambos los periodos en estudio, indicando la presencia de *response shift*.

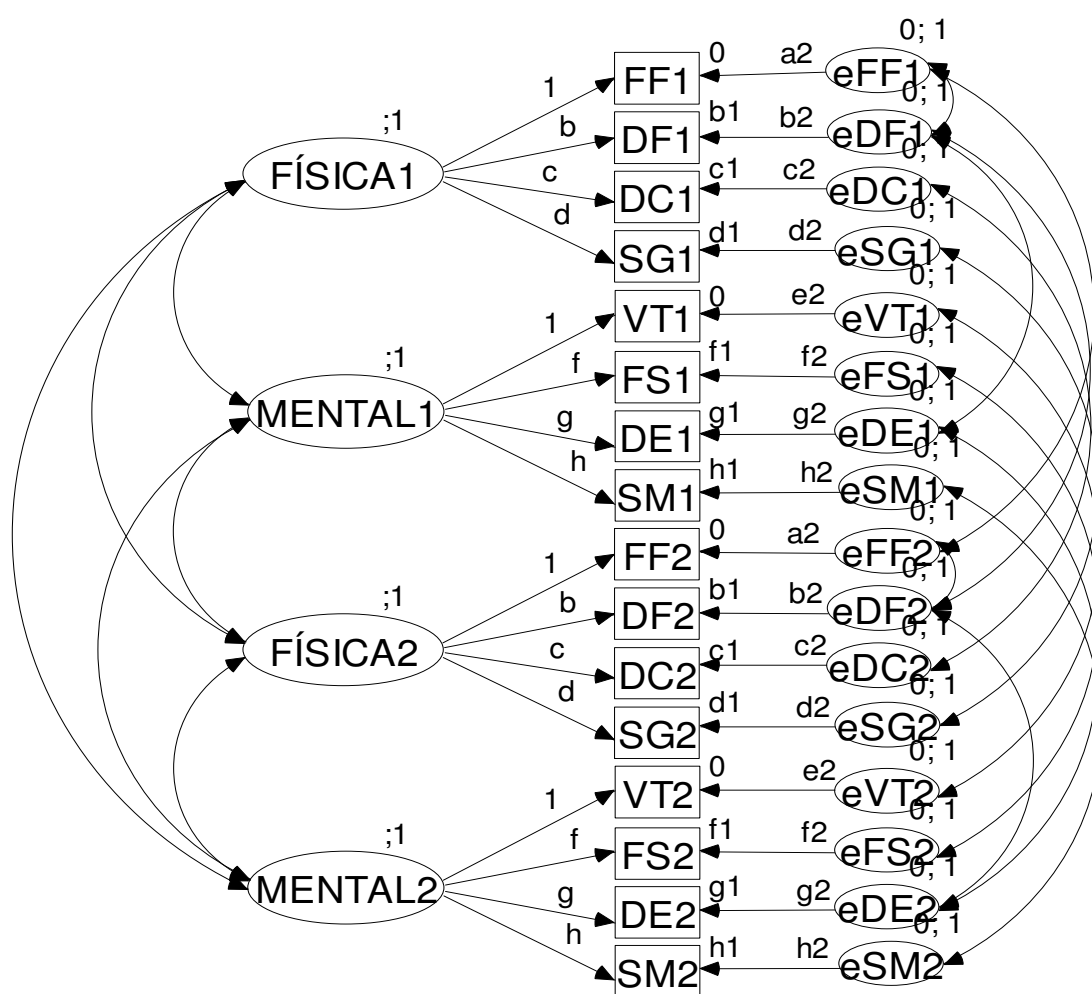


Fig. 24. Diagrama Causal para la estructura de factores y relaciones en el SF-36 para el *Pre-Test* y *Post-Test* - Modelo Longitudinal (Modelo 2), factores de carga e interceptos invariantes.

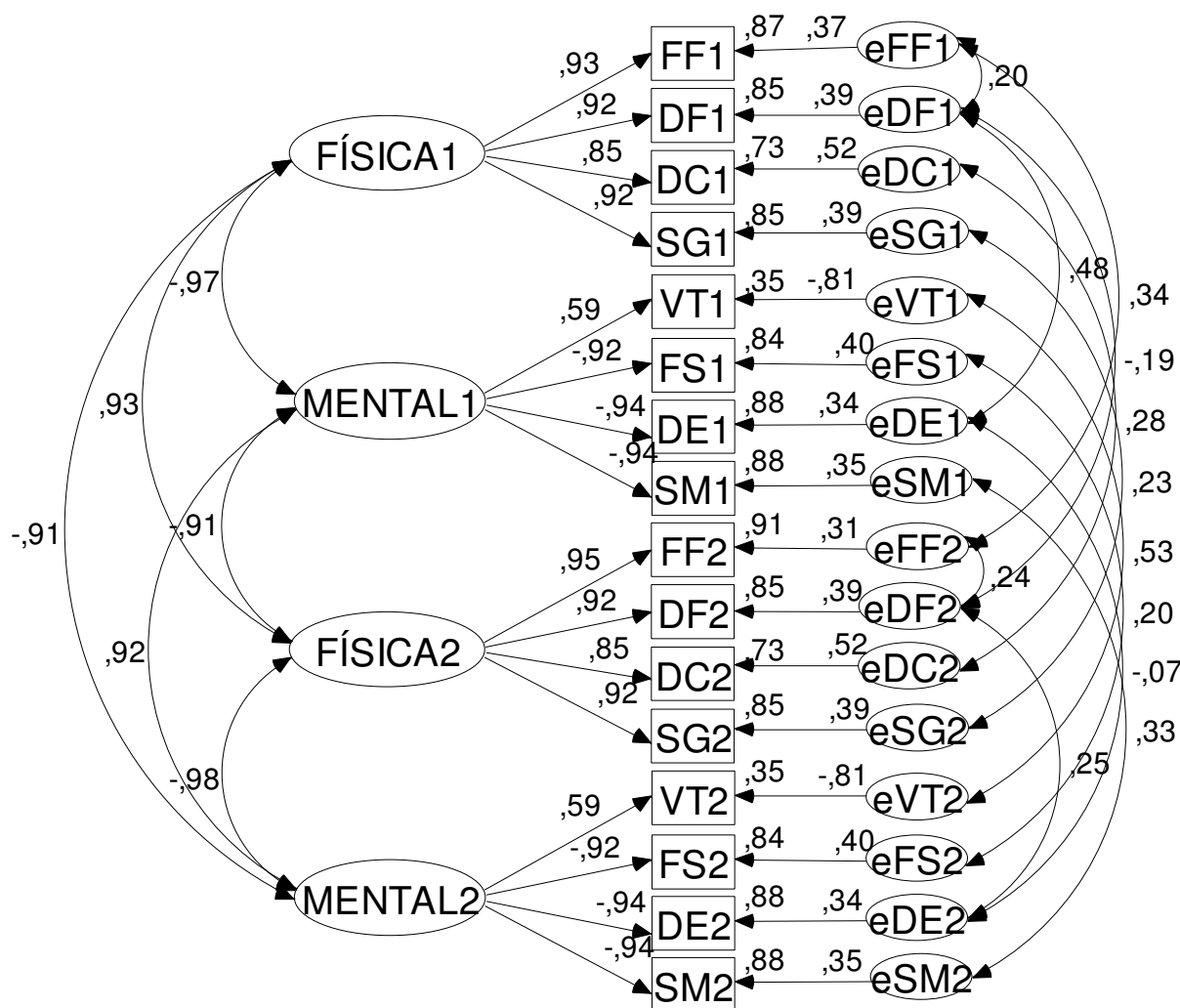


Fig. 25. Modelo de Medida para el *Pre-Test* y *Post-Test* - Modelo Longitudinal (Modelo 2), factores de carga e interceptos invariantes (RMSEA=0,207; Chi-square=720,961; DF=109).

### Paso 3 - Detección de Response Shift

En el tercer paso, se detecta *response shift* chequeando si la calidad de ajuste del Modelo 2 puede ser significativamente mejorada removiendo, a lo largo de los tiempos, las restricciones sobre los elementos individuales dentro de  $\Gamma$ ,  $\tau$  y  $\text{Diag}(\Omega)$ . Paso a paso, la eliminación de todas las restricciones insostenibles asegura un modelo en el cual todos los fenómenos de *response shift* estén considerados. Oort (2005) se refiere a este modelo como Modelo 3 (Tabla 2).

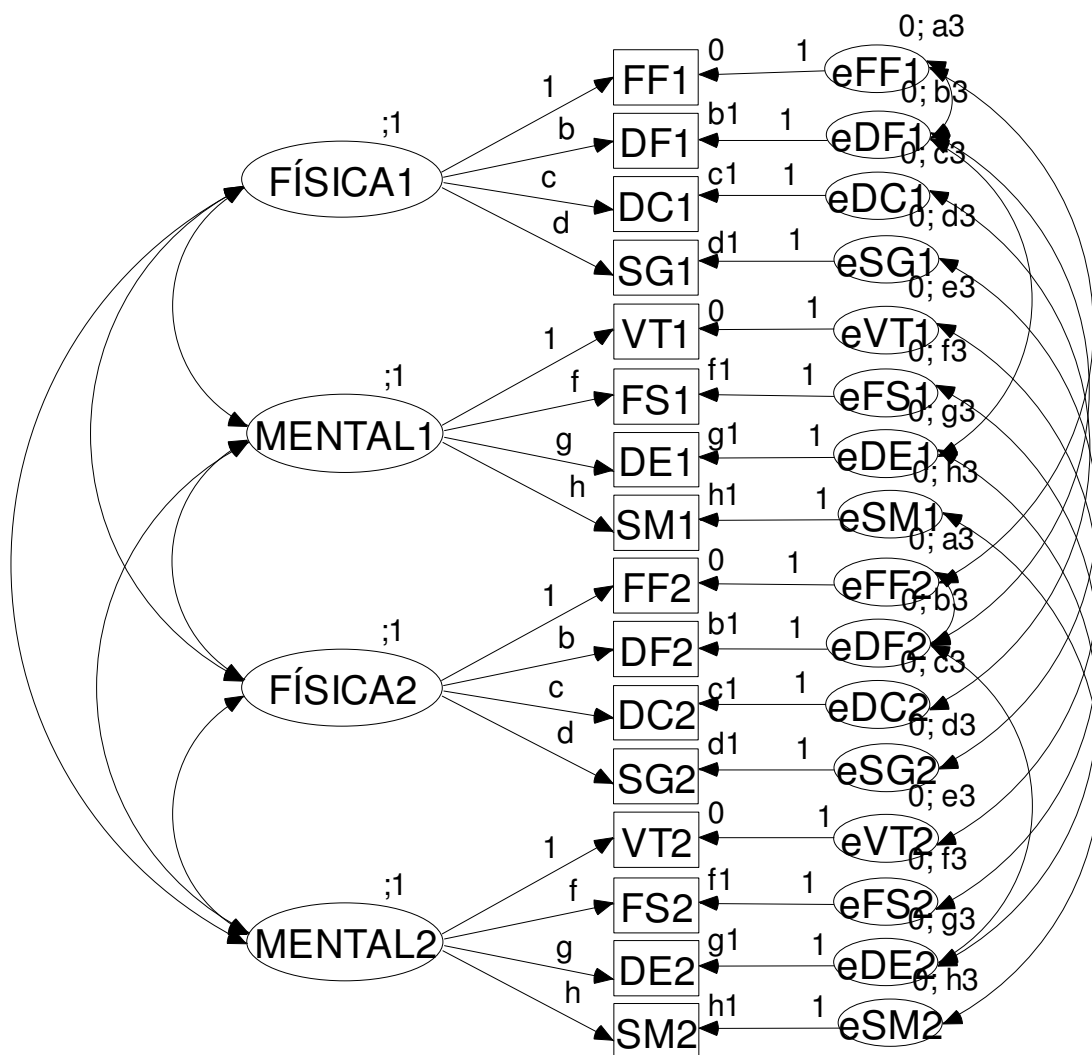
La especificación del Modelo 3 puede basarse en los Tests de Multiplicadores de Lagrange, cambios en los parámetros esperados, Tests de Wald o en la inspección de las discrepancias estandarizadas entre medias, varianzas y covarianzas observadas. Consecuentemente, la mejoría del ajuste en cada paso de modificación puede testarse a través del Test  $\chi^2$  para las diferencias. Asimismo, hay que tener en cuenta que la especificación de un modelo no debe ser únicamente orientada por el cálculo de estadísticos y que no se deben modificar los modelos de forma que no tengan sentido.

Las diferencias a lo largo de los tiempos entre las cargas de los factores comunes (en  $\Gamma$ ), los interceptos (en  $\tau$ ) y varianzas residuales (en  $\Omega$ ) pueden interpretarse de acuerdo con la operacionalización de *response shift* (Tabla 1). O sea, las cargas de los factores comunes que cambian de cero a no cero (o viceversa) indican reconceptualización, otros cambios en las cargas de los factores comunes indican repriorización, cambios en los interceptos indican recalibración uniforme y cambios en las varianzas residuales indican recalibración no uniforme. En el Modelo 3, con todos los fenómenos de *response shift* considerados, el cambio en las medias de los factores comunes, representadas por  $\alpha_2 - \alpha_1$  constituye una medida válida de Cambio Alpha.

El primer paso ha sido restringir las varianzas de los factores residuales, según el diagrama presentado en la Fig. 26. Se ha obtenido el modelo de medida de la Fig. 27. Aunque se haya registrado una mejoría relativamente al modelo anterior (RMSEA=0,124), la calidad de ajuste no es todavía satisfactoria. Diferencias a lo largo de los tiempos entre las varianzas residuales (o sea, dentro de  $\text{Diag}(\Omega)$ , entre  $\text{Diag}(\Omega_1)$  y  $\text{Diag}(\Omega_2)$ ) son consideradas como evidencia de **recalibración no uniforme**.



Por otra parte, diferencias a lo largo de los tiempos entre los interceptos  $\tau$  (o sea, dentro de  $\tau$ , entre  $\tau_1$  y  $\tau_2$ ) pueden interpretarse como evidencia de **recalibración uniforme**. Por lo tanto, en la fase siguiente, se han levantado paso a paso las restricciones inicialmente impuestas sobre los interceptos hasta conseguir el modelo con mejor calidad de ajuste.



**Fig. 26.** Diagrama Causal para la estructura de factores y relaciones en el SF-36 para el *Pre-Test* y *Post-Test* - Modelo Longitudinal, factores residuales invariantes (Modelo 3).

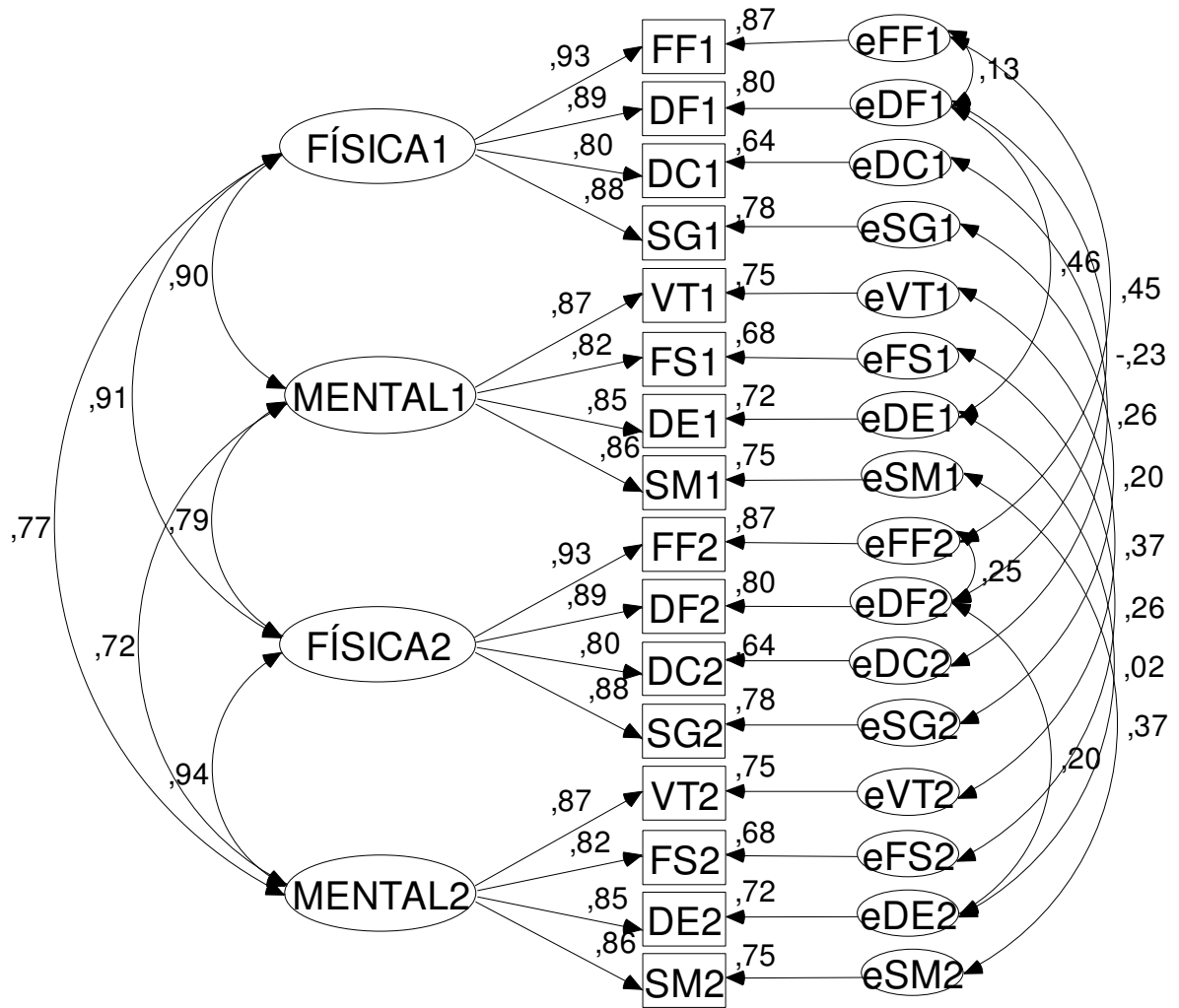
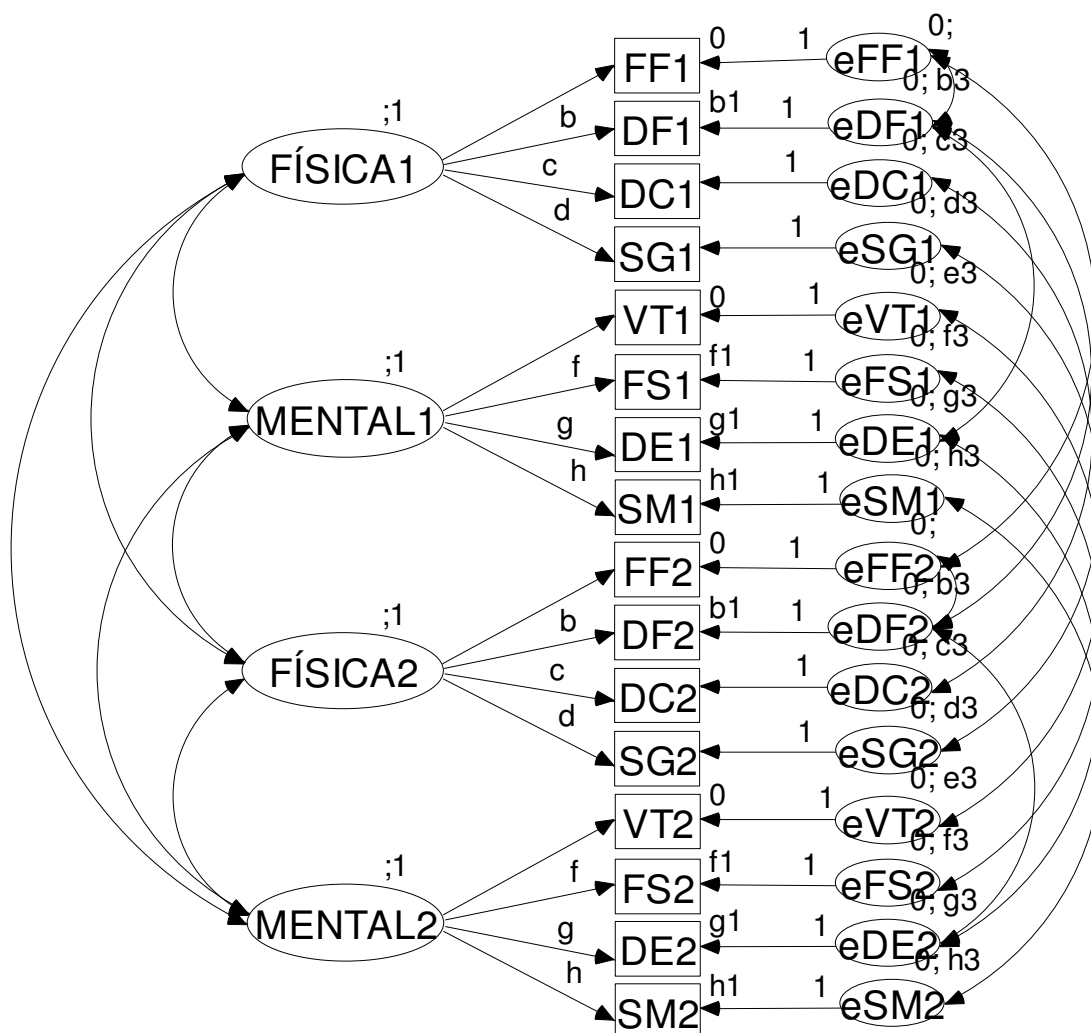


Fig. 27. Modelo de Medida para el *Pre-Test* y *Post-Test* - Modelo Longitudinal, factores residuales invariantes (Modelo 3; RMSEA=0,124; Chi-square=330,489; DF=110).

El paso siguiente ha sido observar los índices de modificación asociados a las relaciones establecidas entre las variables con el objetivo de disminuir la discrepancia. El Índice de Modificación (Sörbom, 1989) deriva de la prueba del Multiplicador de Lagrange y mide cuánto se espera que el valor de la ji-cuadrado disminuya si un parámetro particular se deja libre y el modelo se estima nuevamente. Por lo tanto, el valor de este índice es aproximadamente igual a la diferencia de la ji-cuadrado entre el modelo con el parámetro fijo y el modelo con el parámetro libre. La modificación del modelo paso a paso ha conducido al

diagrama que se presenta en la Fig. 28 y al respectivo modelo de medida de la Fig. 29.



**Fig. 28.** Diagrama Causal para la estructura de factores y relaciones en el SF-36 para el *Pre-Test* y *Post-Test* - Modelo Longitudinal modificado en que no todos los interceptos se han mantenido invariantes (Modelo 3.1).

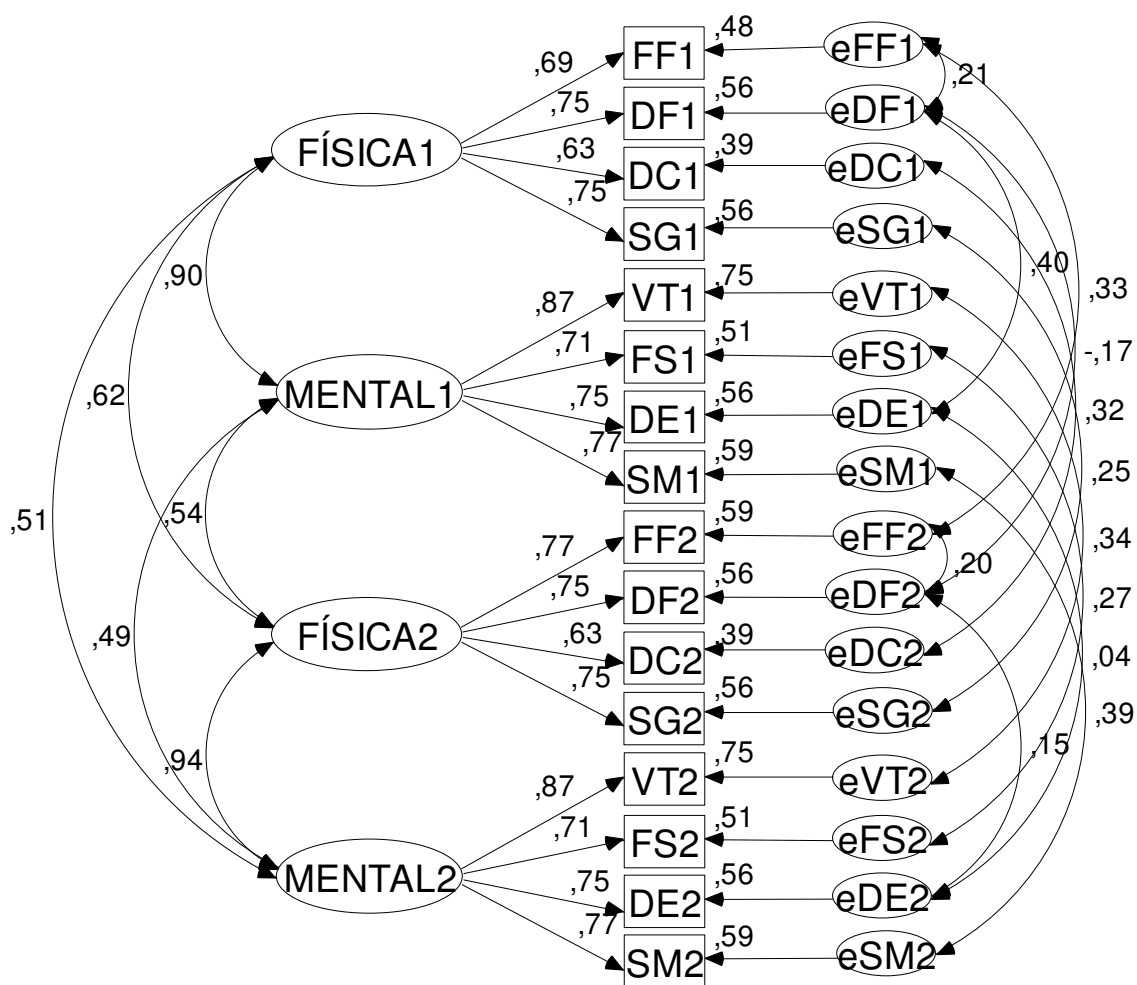


Fig. 29. Modelo de Medida para el *Pre-Test* y *Post-Test* - Modelo Longitudinal modificado en que no todos los interceptos se han mantenido invariantes (Modelo 3.1; RMSEA=0,067; Chi-square=163,833; DF=103).

Del análisis de los índices de modificación ha resultado todavía el tratamiento de la covarianza entre eDC1 y eFF2 así como entre eDC1 y eDE2. Esto indica que la percepción que los pacientes tenían de su Dolor Corporal en el momento del *Pre-Test* está relacionada con la percepción que tienen tanto de su Funcionamiento Físico como de su Desempeño Emocional en el momento del *Post-Test*. Se ha obtenido el modelo de medida presentado en la Fig.30 que presenta ahora un nivel de ajuste más satisfactorio (RMSEA=0,056).

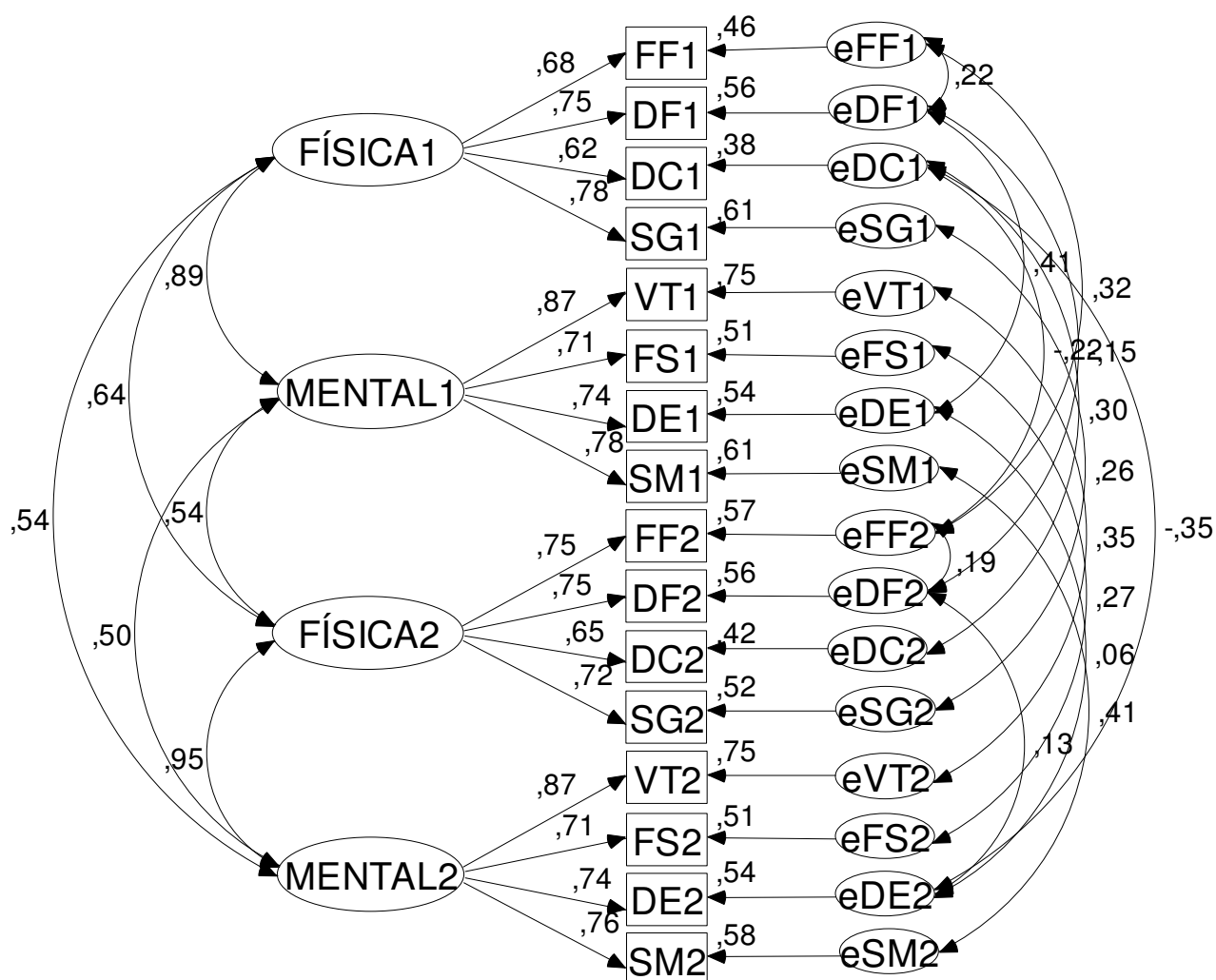


Fig.30. Modelo de Medida para el Pre-Test y Post-Test - Modelo Longitudinal modificado con el tratamiento de la covarianza entre los factores residuales (Modelo 3.2; RMSEA=0,056; Chi-square=142,110; DF=101).

#### Paso 4 - Evaluación de Cambios Alpha

En el cuarto paso se testa la ocurrencia de Cambio Alpha con el Test  $\chi^2$  para las diferencias. Si la hipótesis nula de invarianza de las medias de los factores comunes ( $\alpha_2=\alpha_1$ ) es rechazada, entonces las diferencias entre las medias de los factores comunes a lo largo de los tiempos, se puede tomar como medida del verdadero cambio. Además, otros tipos de cambio pueden ser probados añadiendo restricciones a lo largo de los tiempos en  $\text{Diag}(\Phi), \Phi^*$  y, posiblemente,

$\Omega^*$  (Tabla 1). Únicamente las restricciones que se han demostrado ser sostenibles son consideradas en el modelo final, que Oort (2005) refiere como Modelo 4 (Tabla 2).

Al imponer medias de los factores comunes invariantes (Modelo 4.1), la calidad de ajuste ha empeorado (RMSEA=0,062; Chi-square=154,856; DF=101), llevando a concluir que se ha producido verdadero cambio en las medias (Cambio Alpha). Por otra parte, al imponer apenas la media del factor Salud Mental invariante (Modelo 4.1.1), se ha obtenido un modelo con mejor calidad de ajuste (RMSEA=0,055; Chi-square=142,400; DF=102). Finalmente, se han impuesto varianzas de los factores comunes invariantes (Modelo 4.2), correlaciones entre los factores comunes invariantes (Modelo 4.3) y correlaciones entre los factores residuales invariantes (Modelo 4.4), lo que ha conducido a modelos con calidad de ajuste similar, indicando que no se ha producido un cambio verdadero ni en las varianzas de los factores comunes (cambio verdadero en las varianzas), ni en la correlación entre ellos (reconceptualización o repriorización de nivel alto), ni en la correlación entre los factores residuales (reconceptualización o repriorización de nivel bajo).

Se acepta pues el Modelo 4.1.1 como modelo final. El Modelo 4.1.1 no contiene las restricciones de los parámetros cuyo Índice de Modificación reveló que no debían ser restringidos a través del tiempo. En particular, se liberaron las cargas de las variables observadas FF1 y FF2, los interceptos de DC1, SG1, DC2 y SG2, las varianzas de los factores residuales eFF1 y eFF2 y las medias de FÍSICA1 y FÍSICA2.

En la Tabla 35 se presenta la comparación de los índices obtenidos para los diferentes modelos ajustados. Además de la ji-cuadrado y el RMSEA, se presenta el Índice de Validación Cruzada Esperada (ECVI) para los varios modelos. El ECVI es una medida de información teórica que considera tanto la calidad de

ajuste como la complejidad del modelo. Este índice decrece cuando se añaden parámetros al modelo y toma el valor más bajo para el mejor modelo; a partir de este punto el valor del ECVI se incrementa al añadir más parámetros. Por lo tanto, modelos simples y con buen ajuste presentan valores bajos de ECVI, mientras modelos complejos y con pobre ajuste presentan valores más elevados. De esta forma, este estadístico posibilita la comparación de modelos.

Modelo	Descripción	GL	$\chi^2$	RMSEA	ECVI
Modelo 1	Modelo sin restricciones	89	152,995	0,074 (0,054; 0,094)	2,130 (1,898; 2,273)
Modelo 2	Modelo “no response shift” (factores de cargas e interceptos invariantes)	109	720,961	0,207 (0,193; 0,222)	6,160 (5,540; 6,837)
Modelo3	Varianzas de los factores residuales invariantes	110	330,489	0,124 (0,109; 0,139)	3,164 (2,776; 3,610)
Modelo 3.1	Modelo Corregido (no todos interceptos invariantes)	102	160,382	0,067 (0,047; 0,086)	1,999 (1,762; 2,296)
Modelo 3.2	Modelo Corregido (tratamiento de covarianza)	101	142,110	0,056 (0,032; 0,076)	1,863 (1,653; 2,135)
Modelo 4.1	Medias de los factores comunes invariantes	101	154,856	0,062 (0,041; 0,081)	1,930 (1,705; 2,216)
Modelo 4.1.1	Medias del factor común Salud Mental invariante	102	142,400	0,055 (0,031; 0,075)	1,850 (1,641; 2,122)
Modelo 4.2	Varianzas de los factores comunes invariantes	102	144,571	0,056 (0,032; 0,076)	1,852 (1,640; 2,125)
Modelo 4.3	Correlación entre los factores comunes invariantes	102	143,087	0,055 (0,032; 0,076)	1,856 (1,645; 2,128)
Modelo 4.4	Correlaciones entre los factores residuales invariantes	103	148,038	0,058 (0,035; 0,078)	1,878 (1,662; 2,156)

**Tabla 35.** Bondad de Ajuste para los modelos obtenidos según el procedimiento de Oort (2005) para detectar *Response Shift*.

Aceptando el Modelo 4.1.1 como modelo final, hemos analizado con más detalle los parámetros que han cambiado entre el *Pre-Test* y el *Post-Test*. La Tabla 36 contiene la síntesis de esto análisis.

Cargas Factoriales ( $\Gamma$ )									
Pre-Test				Pos-Test					
	Salud Física	Salud Mental		Salud Física	Salud Mental		Salud Física	Salud Mental	
FF1	0,37		FF2	0,39					
DF1	1,01		DF2	1,01					
DC1	0,78		DC2	0,78					
SG1	0,58		SG2	0,58					
VT1		0,87	VT2		0,88				
FS1		0,81	FS2		0,81				
DE1		0,87	DE2		0,87				
SM1		0,75	SM2		0,75				
Interceptos ( $\tau$ )		FF	DF	DC	SG	VT	FS	DE	SM
Pre-Test		0	-3,18	-1,37	-0,61	0	1,01	0,35	0,80
Post-Test		0	-3,18	-0,78	-0,24	0	1,01	0,35	0,80
Varianzas de los Factores Residuales (Diag( $\Omega$ ))									
	eFF	eDF	eDC	eSG	eVT	eFS	eDE	eSM	
Pre-Test	0,16	0,78	0,90	0,26	0,25	0,64	0,63	0,38	
Post-Test	0,11	0,78	0,90	0,26	0,25	0,64	0,63	0,38	
Correlaciones entre los Factores Comunes ( $\Phi^*$ )									
	FÍSICA 1	MENTAL 1	FÍSICA 2	MENTAL 2					
FÍSICA 1	1								
MENTAL 1	,898	1							
FÍSICA 2	,633	,543	1						
MENTAL 2	,532	,499	,943	1					
Medias de los Factores Comunes ( $\alpha$ )									
	FÍSICA 1	MENTAL 1	FÍSICA 2	MENTAL 2					
	6,41	3,67	5,99	3,67					

Tabla 36. Parámetros estimados (no estandarizados) para el Modelo 4.1.1.

Del análisis de todo el proceso de detección de *response shift* destacamos que:

- No existe evidencia de **reconceptualización** ya que no se han registrado cambios significativos en el patrón de las cargas de los factores comunes, o sea, los conceptos evaluados en el *Pre-Test* y el *Post-Test* siguen siendo definidos por las mismas variables.



- No existe evidencia de **repriorización**, ya que no se han registrado cambios en lo que respecta a la importancia relativa de las variables indicadoras relativamente a las dimensiones Salud Física e Salud Mental.
- Hemos encontrado diferencias para los interceptos relativos a las variables **Dolor Corporal** y **Salud General**, indicando la ocurrencia de **recalibración uniforme** en lo que respecta a estas dimensiones. Ha ocurrido un cambio relativamente al significado que los individuos atribuyen a la escala de respuesta de estos ítems, o sea, un determinado nivel de Dolor Corporal o Salud General en el instante del *Pre-Test*, asume un significado distinto en el *Post-Test*. Siendo uniforme, esta recalibración ha afectado todas las opciones de respuesta en la misma dirección y extensión.
- Hemos encontrado diferencias en lo que respecta a las varianzas de los factores residuales de la variable **Funcionamiento Físico**, indicando la ocurrencia de **recalibración no uniforme**. También aquí se concluye que ha ocurrido un cambio relativamente al significado que los individuos atribuyen a la escala de respuesta de los ítems relativos al Funcionamiento Físico, pero no según la misma dirección y/o extensión.
- Se ha registrado un cambio verdadero (**cambio Alpha**) en la media del factor común **Salud Física**. En particular, se ha observado una disminución de la puntuación media entre el *Pre-Test* y el *Post-Test*, lo que significa que, en el momento del *Post-Test*, los pacientes perciben una CVRS inferior en lo que respecta a su Salud Física.

El hecho de que no se haya detectado repriorización puede estar relacionado con el bajo tamaño muestral, Oort(2005) refiere que para tal, el tamaño mínimo necesario sería 200. Cualquiera de estos fenómenos, para ser detectado, tendría que ser experimentado por una parte sustancial de los pacientes.

Nótese que, mediante la comparación entre el *Pre-Test* y el *Then-Test*, las técnicas utilizadas anteriormente, habían evidenciado la ocurrencia de cambios Beta o Gamma en las dimensiones de Funcionamiento Físico y Vitalidad. El Modelo de Ecuaciones Estructurales ha posibilitado la integración del *Pre-Test* y *Post-Test*, identificando la ocurrencia de cambios Beta al nivel de las dimensiones de Dolor Corporal y Salud General (recalibración uniforme) y Funcionamiento Físico (recalibración no uniforme).

### *Análisis de la Dimensión Funcionamiento Físico*

En el análisis efectuado mediante el Modelo de Ecuaciones Estructurales sólo hemos considerado las puntuaciones obtenidas en cada una de las dimensiones y no las obtenidas en cada ítem. Con el objetivo de analizar el comportamiento de los ítems individualmente, hemos elegido la dimensión Funcionamiento Físico por ser aquella donde habíamos detectado cambio Beta tanto con el *Then-Test* como mediante el Procedimiento de Oort (2005).

Después de establecer el modelo de medida, ha resultado el modelo longitudinal cuyo diagrama causal se presenta en la Fig. 31.

Hemos empezado por imponer todas las restricciones (modelo “no *response shift*”). El modelo “no *response shift*” se ha revelado inviable y, paso a paso, hemos mejorado la calidad de ajuste del modelo removiendo las restricciones sobre los parámetros, chequeando la calidad de ajuste y observando los Índices de Modificación. Los parámetros estimados para el modelo final se encuentran en la Tabla 37.

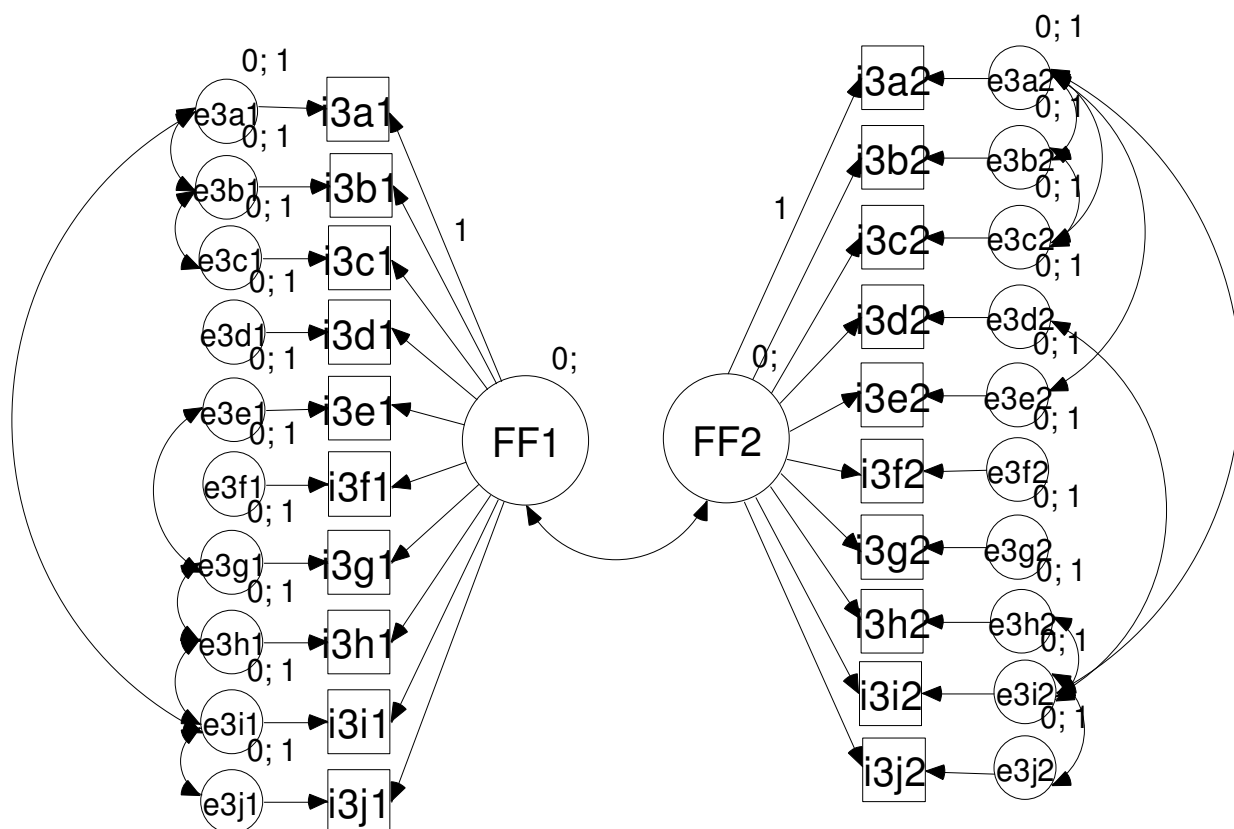


Fig. 31. Diagrama Causal para la estructura de factores y relaciones en la subescala de Funcionamiento Físico del SF-36 para el *Pre-Test* y *Post-Test*.

Ítems	3a	3b	3c	3d	3e	3f	3g	3h	3i	3j
<b>Cargas Factoriales (<math>\Gamma</math>)</b>										
Pre-Test	0,44	0,52	0,52	0,65	0,58	0,63	0,69	0,56	0,47	0,26
Post-Test	0,23	0,52	0,52	0,65	0,58	0,54	0,61	0,61	0,47	0,32
<b>Interceptos (<math>\tau</math>)</b>										
Pre-Test	1,93	2,37	2,48	2,46	2,70	2,48	2,48	2,60	2,86	2,87
Post-Test	1,61	2,23	2,35	2,46	2,80	2,48	2,48	2,60	2,76	2,87
<b>Varianzas de los Factores Residuales (Diag(<math>\Omega</math>))</b>										
Pre-Test	0,38	0,26	0,35	0,21	0,08	0,23	0,19	0,25	0,19	0,18
Post-Test	0,38	0,26	0,35	0,21	0,08	0,23	0,28	0,25	0,19	0,18
<b>Medias de los Factores Comunes (<math>\alpha</math>)</b>										
	FF (Pre-Test)					FF (Post-Test)				
	2,16					2,06				

Tabla 37. Parámetros estimados (no estandarizados) para el Modelo de Ecuaciones Estructurales relativo a la dimensión Funcionamiento Físico del SF-36 para el *Pre-Test* y *Post-Test* (RMSEA=0,079; Chi-square=292,442; DF=160).

Observando las cargas factoriales, destacamos que el **ítem 3a** (“Su salud actual, ¿le limita para hacer esfuerzos intensos, tales como correr, levantar objetos pesados, o participar en deportes agotadores?”) es el que pierde más importancia entre el *Pre-Test* y el *Post-Test*. Pierden también importancia los **ítems 3f** (“Su salud actual, ¿le limita para agacharse o arrodillarse?”), **3g** (“Su salud actual, ¿le limita para caminar un kilómetro o más?”), mientras cobran más importancia los **ítems 3h** (“Su salud actual, ¿le limita para caminar varias manzanas (varios centenares de metros)?”) y **3j** (“Su salud actual, ¿le limita para bañarse o vestirse por sí mismo?”). Se destaca también que los **ítems 3f** (“Su salud actual, ¿le limita para agacharse o arrodillarse?”) y **3h** (“Su salud actual, ¿le limita para caminar varias manzanas (varios centenares de metros)?”) invierten su orden de importancia: en el *Pre-Test* el **3f** es más importante que el **3h**, mientras en el *Post-Test* este orden se invierte.

Observando los interceptos, se detectan cambios para el **ítem 3a** (“Su salud actual, ¿le limita para hacer esfuerzos intensos, tales como correr, levantar objetos pesados, o participar en deportes agotadores?”), **3b** (“Su salud actual, ¿le limita para hacer esfuerzos moderados, como mover una mesa, pasar la aspiradora, o caminar más de una hora?”), **3c** (“Su salud actual, ¿le limita para coger o llevar la bolsa de la compra?”), **3e** (“Su salud actual, ¿le limita para subir un solo piso por la escalera?”) y **3i** (“Su salud actual, ¿le limita para caminar una sola manzana (unos 100 metros)?”). Cambios en los interceptos indican **recalibración uniforme**, o sea, la escala de medida de estos ítems ha sido afectada en la misma dirección y magnitud.

Las varianzas de los factores residuales cambiaron para el **ítem 3g** (“Su salud actual, ¿le limita para caminar un kilómetro o más?”), indicando **recalibración no uniforme**, o sea la escala de medida asociada a este ítem no ha sido afectada según la misma dirección y/o magnitud.

Por último, se observa un cambio real en la media del factor común Funcionamiento Físico. En particular, se ha observado una disminución de la

puntuación media entre el *Pre-Test* y el *Post-Test*, lo que se interpretaría como un **cambio Alpha**, o sea, el Funcionamiento Físico de estos pacientes se ha efectivamente deteriorado.

Los resultados encontrados en nuestro estudio, son compatibles con las descripciones existentes en la literatura. En realidad, muchos otros estudios confirman la presencia y la importancia del fenómeno de *response shift* en estudios longitudinales con pacientes de cáncer (Adang *et al*, 1998; Ahmed *et al*, 2001; Hagedoorn *et al*, 2002; Schwartz *et al*, 1999; Rees *et al*, 2002; Michielsen *et al*, 2007; Sharpe *et al*, 2005; Rees *et al*, 2005; Korfage *et al*, 2007; Pucciarelli *et al*, 2008; Andrykowski *et al*, 2009; Westerman *et al*, 2008; Santana *et al*, 2009).

## CONCLUSIONES

1.- En la literatura científica se ha registrado un aumento notable del interés por el tema de la Calidad de Vida relacionada con la Salud. En particular, ha aumentado el número de estudios longitudinales que evalúan CVRS en pacientes pertenecientes a diferentes grupos de diagnóstico.

2.- La mayoría de las técnicas estadísticas a las que habitualmente se recurre para analizar datos de CVRS no tienen en cuenta la presencia del fenómeno conocido como *response shift* con lo cual la evaluación del verdadero cambio resulta sesgada. En la literatura son muy pocas las soluciones propuestas para detectar *response shift*.

3.- Métodos factoriales poco utilizados en el contexto de la CVRS como el STATIS o el Procedimiento de Krzanowski son una contribución importante para el análisis y detección del Cambio. Sin embargo, no posibilitan una identificación eficaz de las componentes del *response shift*: recalibración, repriorización y reconceptualización.

4.- En el contexto médico, el *Then-Test* es la herramienta más utilizada para evaluar el Cambio. Esta metodología ha sido objeto de críticas por diversos autores, sobre todo en lo que se refiere a problemas de memoria por parte de los pacientes. Además, la utilización del *Then-Test* únicamente permite detectar la ocurrencia de Cambio y no aislar las diferentes componentes del *response shift*.

5.- El Modelo de Zwinderman es la única propuesta conocida que permite la cuantificación del Cambio, aislando el cambio debido al tratamiento y el cambio

debido al tiempo. Sin embargo, este autor ni siquiera plantea la existencia del fenómeno de *response shift*. En este trabajo se propone complementar la información obtenida mediante la aplicación de este modelo a través de una técnica factorial como el HJ-Biplot para analizar los cambios ocurridos al nivel de las relaciones entre las variables y así descartar, o no, la ocurrencia de *response shift*.

6.- El procedimiento propuesto por Oort está basado en los Modelos de Ecuaciones Estructurales y puede constituir una herramienta útil y eficaz en la evaluación del Cambio y, en particular, en la detección de las varias componentes del fenómeno de *response shift*, sin que para ello sea necesario la obtención de datos adicionales como en el *Then-Test*.

7.- En el estudio multicéntrico realizado con pacientes oncológicos, el *Then-Test* permitió detectar Cambio Beta en las dimensiones Funcionamiento Físico y Vitalidad. En el resto de las dimensiones, una vez descartada la presencia de Cambio Beta o Gamma, se ha detectado Cambio Alpha en las dimensiones Desempeño Físico, Dolor Corporal y Funcionamiento Social.

8.- El Modelo de Zwinderman ha permitido cuantificar el cambio en la CVRS global de los pacientes separando la parte debida al tratamiento y la parte debida al tiempo. Ambas estimaciones traducen un empeoramiento global de la CVRS de los pacientes, que es todavía más acentuado en el grupo de pacientes que se encontraba bajo tratamiento. Considerando el cambio debido al tratamiento, este deterioro de la calidad de vida de los pacientes es más evidente para la dimensión Dolor Corporal; mientras considerando el cambio debido al tiempo, el empeoramiento de la calidad de vida es más visible en la dimensión de Desempeño Físico.

9.- El análisis del SF-36 mediante un Modelo de Curvas de Crecimiento ha evidenciado que, a lo largo de las diferentes aplicaciones, el cambio ha sido mucho más significativo en las dimensiones asociadas a la Salud Física que en las asociadas a la Salud Mental.

10.- En el Análisis STATIS Dual y METABILOT se han evidenciado cambios en la estructura factorial de los diferentes tiempos, lo que se ha interpretado como evidencia de que ha ocurrido Cambio Beta o Gamma.

11.- El análisis del conjunto de datos en estudio mediante un Modelo de Ecuaciones Estructurales y, en particular, la aplicación de la propuesta metodológica de Oort, han permitido detectar recalibración uniforme en las dimensiones Dolor Corporal y Salud General y recalibración no uniforme en la dimensión Funcionamiento Físico. Esto pone de manifiesto que el *Then-Test* no permite detectar Cambio Beta cuando la recalibración ocurrida es no uniforme. Además, el análisis mediante este procedimiento ha evidenciado Cambio Alpha en la dimensión general Salud Física.



---

---

## REFERENCIAS

---

---

---

## REFERENCIAS

Aaronson N.K.; Bakker W.; Stewart A.L., van Dam, F.S.A.M.; van Sándwich, N.; Yarnold, J.R. & Kirkpatrick, A. (1987). Multidimensional approach to the measurement of quality of life in lung cancer clinical trials. In: Aaronson NK, Beckmann J, editors. *The Quality of Life of Cancer Patients* (pp. 63-82) New York: Raven Press.

Aaronson, N.K.; Ahmedzai, S.; Bergman, B.; Bullinger, M.; Cull, A.; Duez, N.J.; Filiberti, A.; Flechtner, H.; Fleishman, D.B.; De Haes, J.C.J.M.; Kaasa, S.; Klee, M.; Osoba, D.; Razavi, D.; Rofe, P.; Schraub, S.; Sneeuw, K.; Sullivan, M. & Takeda, F. (1993). The European Organization for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology. *Journal of the National Cancer Institute*, **85**(5), 365-376.

Adang, E.M.M.; Kootstra, G.; Engel, G.L.; van Hooff, J.P. & Merckelback HLGJ (1998). Do retrospective and prospective quality of life assessments differ for pancreas-kidney transplant recipients? *Transplant International*, **11**, 11-15.

Addington-Hall, J. & Kalra, L. (2001). Who should measure quality of life? *British Medical Journal*, **322**(9), 1417-1420.

Ahmavaara, Y. (1954). Transformation analysis of factorial data. *Annals of the Academy Science Fennicae. Series B*, **881**(2), 54-59.

Ahmed, S.; Mayo, N.E.; Wood-Dauphinee, S. & Hanley, J. (2001). Response shift in the assessment of health-related quality of life (HRQL) post-stroke [Abstract]. *Quality of Life Research*, **10**, 204.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic and Control*, **19**, 716-723.

Albrecht, G.I., & Devlieger, P.J. (1999). The disability paradox: high quality of life against all odds. *Social Science and Medicine*, **48**(8), 977-988.

Allen, M.J., & Yen, W.M. (2002). *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press.

- Anderson, C.; Laubscher, S. & Burns, R. (1996). Validation of the Short Form 36 (SF-36) health survey questionnaire among stroke patients. *Stroke*, **27**(10), 1812-1816.
- Andrykowski, M.A.; Donovan, K.A. & Jacobsen, P.B. (2009). Magnitude and correlates of response shift in fatigue ratings in women undergoing adjuvant therapy for breast cancer. *Journal of Pain and Symptom Management*, **37**(3), 341-351.
- Angst, F.; Aeschlimann, A. & Stucki, G. (2001a). Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. *Arthritis & Rheumatism*, **4**, 384-391.
- Angst, F.; Aeschlimann, A.; Steiner, W. & Stucki, G. (2001b). Responsiveness of the WOMAC osteoarthritis index as compared with the SF-36 in patients with osteoarthritis of the legs undergoing a comprehensive rehabilitation intervention. *Annals of the Rheumatic Diseases*, **60**, 834-840.
- Aranda, S.; Schofield, S.P.; Weih, L.; Yates, P.; Milne, D.; Faulkner, R. & Voudouris, N. (2005). Mapping the quality of life and unmet needs of urban women with metastatic breast cancer. *European Journal of Cancer Care*, **14**, 211-222.
- Armenakis, A., Field, H., Wilmoth, J. (1977). An algorithm for assessing factor structure congruence. *Educational and Psychological Measurement*, **37**, 213-214.
- Arndt, V.; Merx, H.; Stegmaier, C.; Ziegler, H. & Brenner, H. (2004). Quality of Life in Patients With Colorectal Cancer 1 Year After Diagnosis Compared With the General Population: A Population-Based Study. *Journal of Clinical Oncology*, **22**, 4829-4836.
- Arndt, V.; Merx, H.; Stegmaier, C.; Ziegler, H. & Brenner, H. (2005). Persistence of Restrictions in Quality of Life from the First to the Third Year after Diagnosis in Women With Breast Cancer. *Journal of Clinical Oncology*, **23**, 4945-4953.
- Avis, N.E.; Crawford, S. & Manuel, J. (2005). Quality of life among younger women with breast cancer. *Journal of Clinical Oncology*, **23**, 3322-3330.
- Bacalá, N. (2004). *Contribuciones al Análisis de Matrices de Datos Multivía: Tipología de las Variables*. Tesis Doctoral. Universidad de Salamanca. España.
- Badia, X. & Carné, X. (1998). La evaluación de la calidad de vida en el contexto del ensayo clínico. *Medicina Clínica*, **110**(14), 550-556.

- Balain, B; Ennis, O.; Kanis, G.; Singhal, R; Roberts, S.N.; Rees, D. & Kuiper, J.H. (2009). Response shift in self-reported functional scores after knee microfracture for full thickness cartilage lesions. *Osteoarthritis and Cartilage*, **17**(8), 1009-1013.
- Baltes, P.B. & Nesselroade, J. R. (1973). The developmental analysis of individual differences on multiple measures. In J. R. Nesselroade & H. W. Reese (Eds.), *Life span developmental psychology: Methodological issues*. New York: Academic Press.
- Barclay-Goddard, R.; Epstein, J.D. & Mayo, N.E. (2009). Response shift: a brief overview and proposed research priorities. *Quality of Life Research*, **18**, 335–346.
- Barroso, S. (2007). *A Pessoa com Cancro*, Lisboa: Sociedade Portuguesa de Oncologia.
- Bartholomew, S. (2002). *Implicit theories and beta change in longitudinal evaluations of training effectiveness: an investigation using Item Response Theory*. PhD Disertation. Virginia Polytechnic Institute and State University.
- Beal, C.C.; Stuifbergen, A.K., Sands, D.; Brown A. (2007). Depression in multiple sclerosis: a longitudinal analysis. *Archives of Psychiatric Nursing*, **21**, 181–191.
- Bedeian, A.; Armenakis, A. & Randolph, W. (1988). The significance of congruence coefficients: a comment and statistical test. *Journal of Management*, **14**, 559-566.
- Benyamini, Y.; Leventhal, E.A. & Leventhal, H. (2000). Gender differences in selfassessment of health. *Psychosomatic Medicine*, **62**(3), 354–64.
- Bergner, M.; Bobbitt, R.A.; Carter, W.B. & Gilson, B.S. (1981). The Sickness Impact Profile: Development and Final Revision of a Health Status Measure. *Medical Care*, **19**, 787-805.
- Bernhard, J.; Zahrieh, D.; Coates, A.S.; Gelber, R.D.; Castiglione-Gertsch, M.; Murray, E.; Forbes, J.F.; Pery, L.; Collins, J.; Snyder, R.; Rudenstam, C.; Crivellari, D.; Veronesi, A.; Thürlimann, B.; Fey, M.F.; Price, K.N.; Goldhirsch, A. & Hürny, C. for the International Breast Cancer Study Group (IBCSG) (2004). Quantifying trade-offs: quality of life and quality-adjusted survival in a randomised trial of chemotherapy in postmenopausal patients with lymph node-negative breast cancer. *British Journal of Cancer*, **91**, 1893-1901.
- Bezjak, A.; Ng, P.; Skeel, R.; DePetrillo, A.D.; Comis, R. & Taylor, K.M. (2001). Oncologists' use of quality of life information: Results of a survey of Eastern Cooperative Oncology Group physicians. *Quality of Life Research*, **10**(1), 1-13.

- Bishop, M. (2005). Quality of life and psychosocial adaptation to chronic illness and acquired disability: a conceptual and theoretical synthesis. *Rehabilitation Counseling Bulletin*, **48**(4), 219–231.
- Bland, J.M. & Altman, D.G. (1997). Statistics notes: Cronbach's alpha. *British Medical Journal*, **314**, 572.
- Boini, S.; Briançon, S.; Guillemin, F.; Galan, P. & Herberg, S.(2004). Impact of cancer occurrence on health-related quality of life: a longitudinal pre-post assessment. *Health and Quality of Life Outcomes*, **2**, 4. Consultado a 15 de marzo de 2009 en la World Wide Web: <http://www.hqlo.com/content/2/1/4>.
- Bono, R.; Arnau, J. & Vallejo, G. (2008). Técnicas de análisis aplicadas a datos longitudinales en psicología y ciencias de la salud: período 1985-2005. *Papeles del psicólogo: revista del Colegio Oficial de Psicólogos*, **29**(1), 136-146.
- Bordeleau, L.; Szalai, J.; Ennis, M.; Leszcz, M.; Specca, M.; Sela, R.; Doll, R.; Chochinov, H.M.; Navarro, M.; Arnold, A.; Pritchard, K.I.; Bezjak, A.; Llewellyn-Thomas, H.A.; Sawka, C.A. & Goodwin, P. J. (2003) Quality of life in a randomized trial of group psychosocial support in metastatic breast cancer: overall effects of the intervention and an exploration of missing data. *Journal of Clinical Oncology*, **21**, 1944–1951.
- Breckler, S.J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin*, **107**(2), 260-273.
- Broberger, E.; Sprangers, M. & Tishelman, C. (2006). Do Internal. Standards of Quality of Life Change in Lung Cancer. Patients? *Nursing Research*, **55**(4), 274-282.
- Brody, H. & Sobel, D.S. (1979). A systems view of health and disease. In D. S. Sobel (Ed.), *Ways of health: Holistic approaches to ancient and contemporary medicine* (pp. 87-104) New York: Harcourt Brace Jovanovich.
- Brook, R.; Ware, J.; Davies-Avery, A.; Stewart, A.; Donald, C.; Rogers, W.; Williams, K. & Johnston, S. (1979). Overview of Adult Health Status Measures Fielded in Rand's Health Insurance Study. *Medical Care*, **17**(7) (Supplement), 1-131.
- Brooks, R. & The EuroQol Group (1996). EuroQol: the current state of play. *Health Policy*, **37**, 53-72.
- Brossart, D. F., Clay, D. L., & Willson, V. L. (2002). Methodological and statistical considerations for threats to internal validity in pediatric outcome data: Response shift in self-report outcomes. *Journal of Pediatric Psychology*, **27**(1), 97–107.

- Browne, J. P.; O'Boyle, C. A.; McGee, H. M.; McDonald; N. J., & Joyce, C. R. (1997). Development of a direct weighting procedure for quality of life domains. *Quality of Life Research*, **6**(4), 301-309.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). London: SAGE.
- Brunelli, C.; Constantini, M.; Di Giulio, P.; Gallucci, M.; Fusco, F.; Miccinesi, G.; Paci, E.; Peruselli, C.; Morino, P.; Piazza, M.; Tamburini, M. & Toscani, F. (1998). Quality-of-life evaluation: When do terminal cancer patients and health-care providers agree? *Journal of Pain and Symptom Management*, **15**(3), 151-158.
- Burgess, A.P.; Carretero, M.; Elkington, A.; Pasqual-Marsettin, E.; Lobaccaro, C.; Catalán, J. (2000). The role of personality, coping style and social support in health-related quality of life in HIV infection. *Quality of Life Research*, **9**(4), 423-437.
- Caffo, O.; Amichetti, M.; Ferro, A.; Lucenti, A.; Valduga, F. & Galligioni, E. (2003). Pain and quality of life after surgery for breast cancer. *Breast Cancer Research and Treatment*, **80**(1), 39-48.
- Calman, K.C. (1984). Quality of Life in Cancer Patients – an hypothesis. *Journal of Medical Ethics*, **10**(3), 124-127.
- Camfield, L. & Skevington, S.M. (2007). On Subjective Well-Being and Quality of Life. *Journal of Health Psychology*, **13**, 764-775.
- Campolina, A.G. & Ciconelli, R.M. (2006). Qualidade de vida e medidas de utilidade: parâmetros clínicos para as tomadas de decisão em saúde. *Revista Panamericana de Salud Publica*, **19**(2), 128-136.
- Carlson, L. & Bultz, B. (2003). Benefits of psychosocial oncology care: Improved quality of life and medical cost offset. *Health and Quality of Life Outcomes*. **1**, 8. Consultado a 29 de enero de 2009 en la World Wide Web: <http://www.hqlo.com?content/1/1/8>.
- Carlson, L.E.; Angen, M.; Cullum, J., Goodey, E.; Koopmans, J.; Lamont, L., MacRae, J.H.; Martin, M.; Pelletier, G.; Robinson, J.; Simpson, J.S.A.; Speca, M.; Tillotson, L. & Bultz, B.D. (2004). High levels of untreated distress and fatigue in cancer patients. *British Journal of Cancer*, **90**, 2297-2304.
- Carvalho, T. (2007). *Aprender sobre o Cancro*, Lisboa: Sociedade Portuguesa de Oncologia.

---

Cella, D.; Tulsky, D.; Gray, G.; Sarafian, B.; Linn, E.; Bonomi, A.; Silberman, M.; Yellen, S.B.; Winicour P. & Brannon, J. (1993). The Functional Assessment of Cancer Therapy Scale: development and validation of the general measure. *Journal of Clinical Oncology*, **11**, 570-579.

Cella, D.F. & Cherin, E.A. (1988). Quality of Life during and after cancer treatment. *Comprehensive Therapy*, **14**(5), 69-75.

Chow, E.; Chiu, H. ; Doyle, M., Hruby, G.; Holden, L.; Barnes, E.A.; Tsao, M.; Mallia, G.; Harris, K. & Danjoux, C. (2007). Patient expectation of the partial response and response shift in pain score. *Supportive Cancer Therapy*, **4**(2), 110-118.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences*. New Jersey: Lawrence.

Comité Consultivo para la Prevención del Cáncer (1999). *Recomendaciones sobre el Cribado del Cáncer en la Unión Europea*, Viena.

Couvreur, C. (2001). *A Qualidade de Vida: Arte para Viver no Século XXI*, Loures: Lusociência.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**(3), 297-334.

Curran, P. J., & Hussong, A. M. (2001). Structural equation modeling of repeated measures data. In D. Moskowitz & S. Hershberger (Eds.), *Modeling intraindividual variability with repeated measures data: Methods and applications*. New York: Erlbaum.

Curran, P.J. & Hussong, A.M. (2003). The use of latent trajectory models in psychopathology research. *Journal of Abnormal Psychology*, **112**, 526-544.

Curran, P.J.; Bauer, D.J. & Willoughby, M.T. (2004). Testing and probing main effects and interactions in latent curve analysis. *Psychological Methods*, **9**(2), 220-237.

De Bock, G.H.; Bonnema, J.; Zwaan, R.E.; Van de Velde C.J.H.; Kievit, J. & Stiggelbout, A.M. (2004). Patient's needs and preferences in routine follow-up after treatment for breast cancer. *Breast Journal Cancer*, **90**(6), 1144-1150.

De Haes J.C.J.M.; Olschewski, M.; Fayers, P.M.; Visser, M.R.M.; Cull, A.; Hopwood, P. & Sanderman, R. (1996). *Measuring the quality of life of cancer patients with the Rotterdam Symptom Checklist (RSCL): a manual*, Northern Centre for Healthcare Research: Groningen.

- De Haes, J.C.J.M.; Van Knippenberg, F.C.E. & Neijt, J.P. (1990). Measuring psychological and physical distress in cancer patients: structure and application of the Rotterdam Symptom Checklist. *British Journal of Cancer*, **62**, 1034-1038.
- Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, B*, **39**, 1-38.
- Dibb, B. & Yardley, L. (2006). Factors important for the measurement of social comparison in chronic illness: a mixed-methods study. *Chronic Illness*, **2**(3), 219-320.
- Donovan, K.A.; Jacobsen, P.B.; Andrykowski, M.A.; Winters, E.; Balducci, L.; Malik, U.; Kenady, D. & McGrath, P. (2004). Course of fatigue in women receiving chemotherapy and/or radiotherapy for early stage breast cancer. *Journal of Pain and Symptom Management*, **28**(4), 373-380.
- Duncan, T.E. & McAuley, E. (1993). Social support and efficacy cognitions in exercise adherence: A latent growth curve analysis. *Journal of Behavioral Medicine*, **16**(2), 199-218.
- Echteld, M.A.; van Zuylen, L.; Bannink, M.; Witkamp, E. & der Rijt, C.C. (2007). Changes in and correlates of individual quality of life in advanced cancer patients admitted to an academic unit for palliative care. *Palliative Medicine*, **21**, 199-205.
- Efficace, F.; Therasse, P.; Piccart, M.J.; Coens, C.; van Steen, K.; Welnicka-Jaskiewicz, M.; Cufer, T.; Dyczka, J.; Lichinitser, M.; Shepherd, L.; de Haes, H.; Sprangers, M.A. & Bottomley, A. (2004). Health-related quality of life parameters as prognostic factors in a non-metastatic breast cancer population. An international multicenter study. *Journal of Clinical Oncology*, **22**, 3381-3388.
- Efficace, F.; Osoba, D.; Gotay, C.; Sprangers, M.; Coens, C. & Bottomley, A. (2007). Has the quality of health-related quality of life reporting in cancer clinical trials improved over time? Towards bridging the gap with clinical decision making. *Annals of Oncology*, **18**(4), 775 - 781.
- EMA - European Medicines Agency (2005). Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQOL) measures in the evaluation of medicinal products' European Medicines Agency Pre-authorisation Evaluation of Medicines for Human Use, Committee for Medical Products for Human Use (CHMP), London. Consultado a 29 de junio de 2009 en la World Wide Web: <http://www.emea.eu.int/pdfs/human/ewp/13939104en.pdf>



- 
- Emmerich, W. (1968). Personality development and concepts structure. *Child Development*, **39**, 671-690.
- Engel, J.; Kerr, J.; Schlesinger-Raab, A. *et al* (2003). Predictors of Quality of Life of Breast Cancer Patients. *Acta Oncologica*, **42** (7), 710-718.
- Escoufier, Y. (1973). Le traitement des variables vectorielles. *Biometrics*, **29**, 750-760.
- Escoufier, Y. (1976). Opérateur associé à un tableau de donnés. *Annales de l'INSEE*, **22-23**, 165-178.
- Fairclough, DL. (1997). Summary measures and statistics for comparison of quality of life in a clinical trial of cancer therapy. *Statistics in Medicine*, **16**(11), 1197-209.
- Fayers, P.M. & Machin, D. (2000). *Quality of Life: Assessment, Analysis and Interpretation*. Chichester: John Wiley & Sons.
- Fayers, P.M. & Sprangers, M.A.G. (2002). Understanding self-rated health. *The Lancet*, **359**, 187-188.
- Fayers, P.M., Hand, D.J., Bjordal, K., & Groenvold, M. (1997). Causal indicators in quality of life research. *Quality of Life Research*, **6**, 393-406.
- Fayers, P.M.; Langston, A.L; Robertson, C. (2007). Implicit self-comparisons against others could bias quality of life assessments. *Journal of Clinical Epidemiology*, **60**, 1034-1039.
- FDA (Food and Drug Administration) (2006). *Guidance for industry: Patient reported outcomes: Use in medical product development to support labeling claims: Draft guidance*. Rockville, MD: FDA.
- Ferreira P.L. & Santana P. (2003). Percepção de estado de saúde e de qualidade de vida da população activa: contributo para a definição de normas portuguesas. *Revista Portuguesa de Saúde Pública*, **21**(2), 15-30.
- Ferreira P.L. (2000a). Criação da versão portuguesa do MOS SF-36. Parte I – Adaptação cultural e linguística. *Acta Médica Portuguesa*, **13**, 55-66.
- Ferreira P.L. (2000b). Criação da versão portuguesa do MOS SF-36. Parte II – Testes de validação. *Acta Médica Portuguesa*, **13**, 119-127.
- Figueiredo, A.; Araújo, P. & Figueiredo, P. (2006). Qualidade de Vida do Doente Oncológico. *Revista de Enfermagem Oncológica*, **36**(9), 15-22.
- Frisch, M.B. (2000). *Improving mental and physical health care through Quality of Life*
-

---

*therapy and assessment*. In: Diener, E. & Rahtz, D. (2000). *Advances in Quality of Life theory and research*. Dordrecht: Kluwer Academic Publishers.

Froberg, D.G. & Kane, R.L. (1989a). Methodology for Measuring Health-State Preferences-I: Measuring Strategies. *Journal of Clinical Epidemiology*, **42**(4), 345-354.

Froberg, D.G. & Kane, R.L. (1989b). Methodology for Measuring Health-State Preferences-II: Scaling Methods. *Journal of Clinical Epidemiology*, **42**(5), 459-471.

Froberg, D.G. & Kane, R.L. (1989c). Methodology for Measuring Health-State Preferences-III: Population and Context Effects. *Journal of Clinical Epidemiology*, **42**(6), 585-592.

Froberg, D.G. & Kane, R.L. (1989d). Methodology for Measuring Health-State Preferences-IV: Progress and a Research Agenda. *Journal of Clinical Epidemiology*, **42**(7), 675-685.

Galindo, M.P. (1986). Una alternativa de representación simultánea: HJ-Biplot. *Questiio*, **10**(1), 13-23.

Galindo, P. y Vicente-Galindo, M.P. (2004). Calidad de vida relacionada con la Salud: un constructo multidimensional dinámico. *Metodología de las Ciencias del Comportamiento*, Suplemento 2004, 219-227.

Ganz, P.A.; Guadagnoli, E., Landrum, M.B., Lash, T.L., Rakowski, W. & Silliman, R.A. (2003). Breast cancer in older women: quality of life and psychosocial adjustment in the 15 months after diagnosis. *Journal of Clinical Oncology*, **21**(21), 4027-4033.

Ganz, P.A.; Kwan, L.; Stanton, A.L.; Krupnick, J.L.; Rowland, J.H.; Meyerowitz, B.E.; Bower, J.E. & Belin, T.R. (2004). Quality of life at the end of primary treatment of breast cancer: First results from the moving beyond cancer randomized trial. *Journal of the National Cancer Institute*, **5**, 376-387.

Gervaz, P; Bucher, P; Konrad, B; Morel, P.; Beyeler, S.; Lataillade, L. & Allal, A. (2008). A Prospective Longitudinal Evaluation of Quality of Life After Abdominoperineal Resection, *Journal of Surgical Oncology*, **97**, 14–19.

Gestel, Y.R.B.M; Voogd, A.C.; Vingerhoets, A.J.J.M. Mols, F.; Nieuwenhuijzen, G. A. P.; Repelaer, O. J.; Van Berlo, C. L.H.; & Van De Poll-Franse, L.V. (2007). A comparison of quality of life, disease impact and risk perception in women with invasive breast cancer and ductal carcinoma in situ. *European Journal of Cancer*, **43**(3), 549-556

- 
- Gibbons, F.X. (1999). Social comparison as a mediator of response shift. *Social Science and Medicine*, **48**(11), 1517-1530.
- Glimelius, B.; Hoffman, K.; Graf, W.; Pahlman, L. & Sjoden, P.O. (1994). Quality of life during chemotherapy in patients with symptomatic advanced colorectal cancer. *The Nordic Gastrointestinal Tumor Adjuvant Therapy Group*, **73**, 556–562.
- Golembiewski, R. T., Billingsley, K. & Yeager, S. (1976). Measuring change and persistence in human affairs: Types of changes generated by OD designs. *Journal of Applied Behavioural Science*, **12**, 133-157.
- Goodwin, P.J.; Black, J.T.; Bordeleau, L.J. & Ganz, P.A. (2003) Health-Related Quality-of-Life Measurement in Randomized Clinical Trials in Breast Cancer—Taking Stock. *Journal of the National Cancer Institute*, **95**(4), 263-281.
- Gotay, C.C. (2004). Assessing cancer-related quality of life across a spectrum of applications. *Journal of the National Cancer Institute Monographs*, **33**, 126-133.
- Guilhemin, F.; Bombardier, C. & Beaton, D. (1993). Cross-cultural adaptation of Health related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology*, **12**, 1417-1432.
- Guyatt, G.H., Kirschner, B. & Jaeschke, R. (1992). Measuring health status: what are the necessary measurement properties. *Journal of Clinical Epidemiology*, **45**, 1341-1345.
- Guyatt, G.H.; Feeny, D.H. & Patrick, D.L. (1993). Measuring health-related quality of life: basic sciences review. *Annals of Internal Medicine*, **118**, 622-629.
- Guyatt, G.H.; Walter, S. & Norman, G. (1987). Measuring Change Over Time: Assessing the Usefulness of Evaluative Instruments. *Journal of Chronic Diseases*, **40**(2), 171-178.
- Hagedoorn, M.; Sneeuw, K.C. & Aaronson, N.K. (2002). Changes in physical functioning and quality of life in patients with cancer: Response shift and relative evaluation of one's condition. *Journal of Clinical Epidemiology*, **55**, 176-183.
- Hagelin, C.L.; Wengström, Y.; Ahsberg, E. & Fürst, C.J. (2009). Fatigue dimensions in patients with advanced cancer in relation to time of survival and quality of life. *Palliative Medicine*, **23**, 171–178.
- Hays, R.D. & Morales, L.S. (2001). The RAND-36 measure of health-related quality of life. *Annals of Medicine*, **33**(5), 350-357.

- 
- Heinrich, R.L.; Schag, C.C. & Ganz, P.A. (1984). Living with cancer: the cancer inventory of problem situations. *Journal of Clinical Psychology*, **40**, 972-980.
- Hill, L.G. & Betz, D.L. (2005). Revisiting the Retrospective Pretest, *American Journal of Evaluation*, **26**(4), 501 - 517.
- Howard, G.S., & Dailey, P.R. (1979a). Response shift bias: a source of contamination of self-report measures. *Journal of Applied Psychology*, **64**, 144-150.
- Howard, G.S.; Ralph, K.M.; Gulanick, N.A.; Maxwell, S.E.; Nance, D.W. & Gerber, S.K. (1979b). Internal invalidity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. *Applied Psychological Measurement*, **3**, 1-23.
- Howard, G.S.; Dailey, P.R. & Gulanick, N.A. (1979c). The feasibility of informed pretests in attenuating response-shift bias. *Applied Psychological Measurement*, **3**, 481-494.
- Hunt, S.M.; McKenna, S.P.; McEven, J.; Williams, J. & Papp, E. (1981). The Nottingham Health Profile: subjective health status and medical consultations. *Social Science Medicine*, **15**[A], 221-229.
- Hyland, M.E. (2003). A brief guide to the selection of quality of life instrument. *Health and Quality of Life Outcomes*, **1**, 24. Consultado a 29 de enero de 2009 en la World Wide Web: <http://www.hqlo.com/content/1/1/24>
- Idler, E. L. & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior*, **38**(1), 21-37.
- Jaeschke R.; Singer J. & Guyatt G. (1989). Measurement of health status: Ascertaining the minimal clinically important difference. *Controlled Clinical Trials*, **10**, 407-415.
- Jansen, S.J.; Stiggelbout, A.M.; Nooij, M.A.; Noordijk, E.M. & Kievit, J. (2000). Response shift in quality of life measurement in early-stage breast cancer patients undergoing radiotherapy. *Quality of Life Research*, **9**(6), 603-15.
- Janz, N.K.; Mujahid, M.; Lantz, P.M.; Fagerlin, A.; Salem, B.; Morrow, M.; Deapen, D. & Katz, S.J. (2005). Population-based study of the relationship of treatment and sociodemographics on quality of life for early stage breast cancer. *Quality of Life Research*, **14**(6), 1467-1479.
- Jones, B. L.; Nagin, D. S. & Roeder, K. (2001). A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methods Research*, **29**(3), 374-393.

- Jong, N.; Candel, M.J.J.M.; Schouten, H.C.; Huijjer Abu-Saad, H. & Courtens, A.M. (2004). Prevalence and course of fatigue in breast cancer patients receiving adjuvant chemotherapy. *Annals of Oncology*, **15**(6), 896-905
- Jöreskog, K. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294-316). Newbury, CA: Sage.
- Jöreskog, K. & Sörbom, D. (1981). *LISREL V: Analysis of Linear Structural Relationships by Maximum Likelihood and Least Square Methods*. University of Uppsala, Department of Statistics, Uppsala, Sweden.
- Kaarlola, A.; Pettila, V. & Kekki, P. (2003). Quality of life six years after intensive care. *Intensive Care Medicine*, **29**, 1294-1299.
- Kagan, J. (1980). Perspectives on continuity. In O.G. Brim, Jr. & J. Kagan (Eds.), *Constancy and change in human development* (26-74). Cambridge, MA: Harvard University Press.
- Kaplan G.A. & Camacho T. (1983). Perceived health and mortality: a nine-year follow-up of the human population laboratory cohort. *American Journal of Epidemiology*, **117**(3), 292-304.
- Kaplan, R.M. & Bush, J.W. (1982). Health-related quality of life measurement for evaluation research and policy analysis. *Health Psychology*, **1**, 61-80.
- Keller, S.; Ware, J.; Bentler, P.; Aaronson, N.; Alonso, J.; Apolone, G.; Bjorner, J.; Brazier, J.; Bullinger, M.; Kaasa, S.; Leplège, A.; Sullivan, M. & Gandek B. (1998). Use of structural equation modeling to test the construct validity of the SF-36 Health Survey in ten countries: Results from the IQOLA Project. *Journal of Clinical Epidemiology*, **51**(11), 1179-1188.
- Kerr, J.; Engel, J.; Schlesinger-Raab, A.; Sauer, H. & Hölzel, D. (2003). Communication, quality of life and age: results of a 5-year prospective study in breast cancer patients. *Annals of Oncology*, **14**(3), 421-427.
- Kline, R.B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guildford Press.
- Korfage, I.J.; Koning, H.J. & Essink-Bot, M.L. (2007). Response shift due to diagnosis and primary treatment of localized prostate cancer: a then-test a then-test and a vignette study. *Quality of Life Research*, **16**(10), 1627-1634.
- Korth, B. & Tucker, L. (1975). The distribution of chance congruence coefficients from simulated data. *Psicometrika*, **40**, 361-372.

- Krabbe, P.F.M; Salomon, J.A. & Murray, C.J L. (2007). Quantification of health states with rank-based nonmetric multidimensional scaling. *Medical Decision Making*, **27**(4), 395-405.
- Krzanowski, W. J. (1979). Between-groups comparison of principal components. *Journal of the American Statistical Association*, **74**(367), 703-707.
- L'Hermier des Plantes (1976). *Structuration des tableaux a trois indices de la statistique*. Thèse de 3 ème cycle, Université de Montpellier.
- Lavit, Ch. (1988). *Analyse conjointe de tableaux quantitatifs*. Masson: Paris.
- Lavit, Ch., Escoufier, Y., Sabatier, R. and Traissac, P. (1994). The ACT (Statis Method). *Computational Statistics and Data Analysis*, **18**, 97-119.
- Leendertse, F.; Van Rockel, G.; Sparreboom, T. & Zuidema, M. (1979). *Slaapstoornissen bij kankerpatiënten in Malignant Disease*, Watson, M. & Greer, S. (eds). Pergamon Press: Oxford.
- Levine, M.; Guyatt, G.H.; Gent, M., De Pauw, S.; Goodyear, M.D.; Hryniuk, WM; Arnold, A.; Findlay, B; Skillings, J.R. & Bramwell, V.H. (1988). Quality of life in stage II breast cancer, an instrument for clinical trials. *Journal of Clinical Oncology*, **6**, 1798-1810.
- Levine, S. & Croog, S.H. (1984). *What constitutes Quality of Life? A conceptualizacion of the dimensions of life quality in healthy populations and patients with cardiovascular diseases*. New York: Le Jacq Publishing.
- Levine, S. (1987). The Changing Terrains in Medical Sociology: Emergent Concern with Quality of Life. *Journal of Health and Social Behavior*, **28**(1), 1-6.
- Liang, M. H., Larson, M. G., Cullen, K. E., & Schwartz, J. A. (1985). Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis & Rheumatism*, **28**, 542-547.
- Liga Portuguesa Contra o Cancro (2009). *Rastreio de Cancro da Mama*. Consultado a 19 de agosto de 2009 en la World Wide Web: <http://www.ligacontracancro.pt/>
- Linszen A.C.G.; Van Dam, F.S.A.M.; Engelsman, E.; Van Benthem, J. & Hanewald, G.J.F.P. (1979). Leven met cytostatica. *Pharmaceut Weekbl*, **114**, 501-15.
- Luteijn, F.; Kok, A.R.; Hamel, L.F. & Poiesz, A. (1979). Enige ervaringen met een klachtenlijst. *Ned Tijdschr Psychol*, **34**, 167-179.

- Mangione, C.M.; Goldman, L.; Orav, E.J.; Marcantonio, E.R.; Pedan, A.; Ludwig, L.E.; Donaldson, M.C.; Sugarbaker, D.J.; Poss, R. & Lee, T.H. (1997). Health-related quality of life after elective surgery. *Journal of General Internal Medicine*, **12**, 686–697.
- Marchioro, G.; Azzarello, G.; Checchin, F.; Perale, M.; Segati, R.; Sampognaro E.; Rosetti, F.; Franchin, A.; Pappagallo, G.L. & Vinante, O. (1996). The impact of a psychological intervention on quality of life in non-metastatic breast cancer. *European Journal of Cancer*, **32A**, 1612-1615.
- Martín-Rodríguez, J., Galindo-Villardón, M.P. & Vicente-Villardón, J.L. (2001). Comparison and integration of subspaces from a biplot perspective. *Journal of Statistical Planning and Inference*, **102**(2), 1-13.
- Martins, J.C. (2009). Satisfação do doente com a informação sobre a doença e morbidade. *Revista Latino-Americana de Enfermagem*, **17**(3), 335-340.
- Mayo, N.E., Scott, S.C., Dendukuri, N., Ahmed, S., & Wood-Dauphinee, S. (2008). Identifying response shift statistically at the individual level. *Quality of Life Research*, **17**, 627–639.
- McArdle, J. J. & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development*, **58**, 110-133.
- McAuley, E.; Blissmer, B.; Marquez, D.X.; Jerome, G.J.; Kramer, A.F. & Katula, J. (2000). Social Relations, Physical Activity, and Well-Being in Older Adults. *Preventive Medicine*, **31**(5), 608-617.
- McClimans, L. (2006) *The problem of purpose in quality of life research*. Technical Report 02/06. London School of Economics and Political Science, CPNSS, London, UK.
- McHorney, C.; Ware, J. & Raczek, A. (1993). The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care*, **31**(3), 247-263.
- Meredith, W. & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, **55**, 107-122.
- Michalos, A.C.; Zumbo, B.D. & Hubley, A. (2000). Health and Quality of Life. *Social Indicators Research*, **51**, 245-286.
- Michielsen, H.J.; Van der Steeg, A.F.; Roukema, J.A. & De Vries, J. (2007). Personality and fatigue in patients with benign or malignant breast disease. *Support Care Cancer*, **15**(9), 1067-1073.

- Middel, B.; Goudriaan, H.; de Greef, M.; Stewart, R.; van Sonderen, E.; Bouma, J. & de Jongste, M. (2006). Recall bias did not affect perceived magnitude of change in health-related functional status. *Journal of Clinical Epidemiology*, **59**, 503-511.
- Millsap, R.E. & Hartog, S.B. (1988). Alpha, beta and gamma change in evaluation research: A structural equation approach. *Journal of Applied Psychology*, **73**(3), 574-584.
- Mols, F.; Vingerhoets, A.J.J.M.; Coebergh, J.W.W. & Poll-Franse, L.V. van de (2005). Quality of live among long-term breast cancer survivors: A systematic review. *European Journal of Cancer*, **41**(17), 2613-2619.
- Montazeri, A. (2008). Health-related quality of life in breast cancer patients: A bibliographic review of the literature from 1974 to 2007. *Journal of Experimental & Clinical Cancer Research*, **27**(32). Consultado a 29 de mayo de 2009 en la World Wide Web: <http://www.jeccr.com/content/27/1/32>.
- Morris, J.; Perez, D. & McNoe, B. (1998). The use of quality of life data in clinical practice. *Quality of Life Research*, **7**(1), 85-91.
- Mortimer, J.T.; Finch, M.D. & Kumka, D. (1982). *Persistence and change in development: The multidimensional self-concept*. In P. B. Baltes & O.G. Brim, Jr. (Eds), *Life span development and behavior* (4, 263-313). New York: Academic Press.
- Mosey, J.M., & Shapiro, E. (1982). Self-rated health: A predictor of mortality among the elderly. *American Journal of Public Health*, **72**, 800-808.
- Muldoon, M.F.; Barger, S.D.; Flory, J.D., & Manuck, S.B. (1998). What are quality of life measurements measuring? *British Medical Journal*, **316**, 542-545.
- Muthén, B. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In L. Collins & J. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 1-17). Washington, DC: American Psychological Association.
- Nagin, D. S. & Tremblay, R. E. (2001). Analyzing developmental trajectories of distinct but related behaviors: A groupbased method. *Psychological Methods*, **6**(1), 18-34.
- Nieuwkerk, P.T.; Tollenaar, M.S.; Oort, F.J. & Sprangers, M.A. (2007). Are retrospective measures of change in quality of life more valid than prospective measures? *Medical Care*, **45**(3), 199-205.



---

Northouse, L.L.; Mood, D.W.; Schafenacker, A.; Montie, J.E.; Sandler, H.M.; Forman, J.D.; Hussain, M.; Pienta, K.J.; Smith, D.C.; Sanda, M.G. & Kershaw, T. (2007). Living with Prostate Cancer: Patients' and Spouses' Psychosocial Status and Quality of Life *Journal of Clinical Oncology*, **25**(27), 4171-4177.

OCDE (Organisation de Coopération et de Développement Économique) (1995). *Panorama de la Santé – Les Indicateurs de L'OCDE 1995*, Logiciel Éco-Santé. Paris: OCDE.

OCDE (Organisation de Coopération et de Développement Économique) (2003). *Panorama de la Santé – Les Indicateurs de L'OCDE 2003*, Logiciel Éco-Santé. Paris: OCDE.

OMS (Organización Mundial de Salud) (1948). *Manual of the International Statistical Classification of Diseases. Injuries and Causes of Death*. Geneva: OMS.

OMS (Organización Mundial de Salud) (1993). Study Protocol for the World Health Organization project to develop Quality of Life assessment instrument (WHOQOL). *Quality of Life Research*, **2**(2), 153-159.

Oort, F.J. (2005). Using structural equation modeling to detect response shifts and true change. *Quality of Life Research*, **14**, 587-598.

Oort, F.J.; Visser, M.R.M. & Sprangers, M.A.G. (2005). An application of structural equation modeling to detect response shifts and true change in quality of life data from cancer patients undergoing invasive surgery. *Quality of Life Research*, **14**, 599-609.

Osborne, R.H.; Hawkins, M. & Sprangers, M.A. (2006) Change of perspective: a measurable and desired outcome of chronic disease self-management intervention programs that violates the premise of preintervention / postintervention assessment. *Arthritis & Rheumatism*, **55**(3), 458-465.

Paci, E. (1992). Assessment of validity and clinical application of an Italian version of the Rotterdam Symptom Checklist. *Quality of Life Research*, **1**, 129-134.

Padilla, G.V.; Present, C.; Grant, M.M.; Metter, G.; Lipsett, J. & Heidi, F. (1983). Quality of Life Index for patients with cancer. *Res. Nurs. Health*, **83**, 117-126.

Patrick, D.L. & Deyo, R.A. (1989). Generic and disease-specific measures in assessing health status and quality of life. *Medical Care*, **27**(3), S217-S232.

Pimentel, F.L. (2002). Survey of Portuguese Physicians on Quality of Life Information. *Quality of Life Research*, **11**(7), 654.

- 
- Pimentel, F.L. (2004). *Qualidade de Vida em Oncologia*, Lisboa: Permanyer Portugal.
- Pinson, D.M.K; Ottens, A.J & Fisher, T.A. (2009). Women Coping Successfully With Multiple Sclerosis and the Precursors of Change. *Qualitative Health Research*, **19**(2), 181-193.
- Postulart, D. & Adang, E.M. (2000). Response shift and adaptation in chronically ill patients. *Medical Decision Making*, **20**, 186–193.
- Potthoff, R. F. & Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, **51**, 313-326.
- Preacher, K. & Hayes, A. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, **40**(3), 879-891.
- Pucciarelli, S.; Del Bianco, P.; Toppan, P.; Serpentine, S.; Efficace, F.; Pasetto, L. M.; Friso, L.M.; De Salvo, G.L. & Nitti, D. (2008). Health-Related Quality of Life Outcomes in Disease-Free Survivors of Mid-Low Rectal After Curative Surgery. *Annals of Surgical Oncology*, **15**(7), 1846–1854.
- Rapkin, B.D. & Schwartz, C.E. (2004). Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift. *Health and Quality of Life Outcomes*, **2**, 14. Consultado a 15 de marzo de 2009 en la World Wide Web: <http://www.hqlo.com/content/2/1/14>.
- Rash, G. (1961). On general laws and the meaning of measurement in psicology, *Berkeley symposium on mathematical statistics and probability*, University of California.
- Razmjou, H.; Schwartz, C.E.; Yee, A. & Finkelstein, J.A. (2009). Traditional assessment of health outcome following total kneearthroplasty was confounded by response shift phenomenon. *Journal of Clinical Epidemiology*, **62**, 91-96.
- Reavley, N., Pallant, J.F. & Sali, A. (2009). Evaluation of the Effects of a Psychosocial Intervention on Mood, Coping, and Quality Of Life in Cancer Patients. *Integrative Cancer Therapies*, **8**(1), 47-55.
- Rees, J.; Clarke, M.G.; Waldron, D.; O'Boyle, C.; Ewings, P. & MacDonagh, R.P. (2005). The measurement of response shift in patients with advanced prostate cancer and their partners. *Health and Quality of Life Outcomes*, **3**, 21. Consultado a 15 de marzo de 2009 en la World Wide Web: <http://www.hqlo.com/content/3/1/21>

---

Rees, J.E.; Waldron, D.; O'Boyle, C.A. & MacDonagh, R.P. (2002): Response shift in individualized quality of life in patients with advanced prostate cancer [abstract]. *Clinical Therapeutics*, **24**(Supplement B), 33-34.

Resolução do Conselho de Ministros nº129/2001, Diário da República – I Série-B, Nº190, 17 de Agosto de 2001.

Reynolds, F. & Prior, S. (2003). "Sticking jewels in your life": Exploring women's strategies for negotiating an acceptable quality of life with multiple sclerosis. *Qualitative Health Research*, **13**(9), 1225-1251.

Ribeiro, J.P. (2005). *O importante é a saúde: Estudo de adaptação de uma técnica de avaliação da percepção do estado de saúde*. Lisboa: Fundação Merck.

Ridder, D.; Geenen, R.; Kuijer, R. & Van Middendorp, H. (2008). Psychological adjustment to chronic disease. *Lancet*, **372**, 246-255.

Ring, L.; Höfer, S.; Heuston, F.; Harris, D. & O'Boyle, C.A (2005). Response shift masks the treatment impact on patient reported outcomes (PROs): the example of individual quality of life in edentulous patients. *Health and Quality of Life Outcomes*, **3**, 55. Consultado a 15 de marzo de 2009 en la World Wide Web: <http://www.hqlo.com/content/3/1/55>

Rios, R. (2006). *Aportaciones al Análisis Dinámico de Tablas de Tres Vías*. Tesis Doctoral. Universidad de Salamanca. España.

Rios, R.; Galindo, P. & Vicente, S. (2007). Evaluación de distintos tipos de cambio y response shift mediante un modelo de ecuaciones estructurales: una aplicación al Panel de Hogares de la Unión Europea. *Estadística Española*, **49**(165), 349-372.

Robertson, C; Langston, A.L.; Stapley, S.; McColl, E.; Campbell, M.K.; Fraser, W.D.; Maclennan, G.; Selby, P.L.; Ralston, S.H., Fayers, P.M. & PRISM Trial Group (2009). Meaning behind measurement: self-comparisons affect responses to health-related quality of life questionnaires. *Quality of Life Research*, **18**, 221-230.

Rothwell, P.M.; Mc Dowell Z.; Wong, C.K. & Dorman, P.J. (1997). Doctors and patients don't agree: Cross sectional study of patients and doctors perceptions and assessments of disability in multiple sclerosis. *British Medical Journal*, **314**, 1580-1583.

Sajid, M.S.; Rimpel, J.; Iftikhar, M. & Baig, M.K. (2007). Use of health related quality of life tools in colorectal surgery. *Acta Chirurgica Belgica*, **107**, 623-629.

Santana, M.J.; Au, H.J.; Dharma-Wardene, M.; Hewitt, J.D.; Dupere, D.; Hanson, J.; Glosch, S. & Feeny, D. (2009). Health-related quality of life measures in routine clinical care: Can FACT-Fatigue help to assess the management of fatigue in cancer patients? *International Journal of Technology Assessment in Health Care*, **25**(1), 90–96.

Sá, E.S. (2003). Adaptação Mental à Doença Oncológica. *Nursing*, **181**, 23-29.

Santos, M.J.H., Costa, F.L., Watson, M., Geer, S., Haes, J., Knippenberg, F.V., Pruyn, J., & Borne, B.V. (1994). Adaptação psicológica e qualidade de vida em doentes oncológicos: Escalas de avaliação. *Revista de Psiquiatria Clínica*, **15**(1), 25-36.

Schipper, H; Clinch, J; McMurray, A & Levitt, M. (1984). Measuring the quality of life of cancer patients. The Functional Living Index-Cancer: Development and Validation. *Journal of Clinical Oncology*, **2**(5), 472-83.

Schmitt, N. (1982). The use of analysis of covariance structures to assess beta and gamma change. *Multivariate Behavioral Research*, **17**, 343–358.

Schmitt, N., Pulakos, E. & Lieblein, A. (1984). Comparison of three techniques to assess group-level beta and gamma change. *Applied Psychological Measurement*, **8**, 249-260.

Schroevers, M.J.; Ranchor, A.V. & Sanderman, R. (2004). The role of age at the onset of cancer in relation to survivors' long-term adjustment: a controlled comparison over an eight-year period. *Psychooncology*, **13**, 740–752.

Schünemann, H. & Guyatt, G. (2005). Commentary - GoodbyeM(C)ID! Hello MID, Where Do You Come From? *Health Services Research*, **40**, 593-597.

Schwartz, C.E. & Rapkin, B.D. (2004). Reconsidering the psychometrics of quality of life assessment in light of response shift and appraisal. *Health and Quality of Life Outcomes*, **2**, 16. Consultado a 15 de marzo de 2009 en la World Wide Web: <http://www.hqlo.com/content/2/1/16>.

Schwartz, C.E. & Sendor, R.M. (1999). Helping others helps oneself: response shift effects in peer support. *Social Science and Medicine*, **48**(11), 1563-1575.

Schwartz, C.E. & Sprangers, M.A. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science and Medicine*, **48**(11), 1531-1548.

Schwartz, C.E. & Sprangers, M.A. (2002). Adaptation to changing health: Response shift in quality of life research. *Quality of Life Research*, **11**, 185–187.

---

Schwartz, C.E.; Andresen, E.M.; Nosek, M.A., Krahn, G.L. & The RRTC Expert Panel on Health Status Management (2007). Response shift theory: important implications for measuring Quality of Life in people with disability. *Archives of Physical Medicine and Rehabilitation*, **88**(4), 529-536.

Schwartz, C.E.; Coulthard-Morris, L.; Cole, B. & Vollmer, T. (1997). The quality-of-life effects of Interferon-Beta-1b in multiple sclerosis: An Extended Q-TWiST analysis. *Archives of Neurology*, **54**, 1475-1480.

Schwartz, C.E.; Feinberg, R.G., Jilinskaia, E. & Applegate, J.C. (1999). An evaluation of a psychosocial intervention for survivors of childhood cancer: Paradoxical effects of response shift over time. *Psycho-Oncology*, **8**, 344-354.

Schwartz, C.E.; Sprangers, M.A.G.; Carey, A. & Reed, G. (2004). Exploring response shift in longitudinal data. *Psychology and Health*, **19**, 51-69.

Scientific Advisory Committee (2002). Assessing health status and quality of life instruments: Attributes and review criteria. *Quality of Life Research*, **11**, 193-205.

Selby, P.J.; Chapman, J.A.; Etazadi-Amoli, J.; Dalley, D. & Boyd, N.F. (1984). The development of a method for assessing the quality of life of cancer patients. *British Journal of Cancer*, **50**(1), 13-22.

Sharpe, L.; Butow, P.; Smith, C.; McConnell, D. & Clarke, S. (2005). Changes in quality of life in patients with advanced cancer: evidence of response shift and response restriction. *Journal of Psychosomatic Research*, **58**, 497-504.

Shimozuma, K.; Imai, H.; Kuroi, K.; Ohsumi, S. & Ono, M. (2007). Recent Topics of Health Outcomes Research in Oncology. *Breast Cancer*, **14**(1), 60-65.

Shun, S.C.; Beck, S.L., Pett, M.A. & Richardson, S.J. (2007). Assessing Responsiveness of Cancer-Related Fatigue Instruments: Distribution-Based and Individual Anchor-Based Methods. *The Oncologist*, **12**(4), 495-504.

Sinclair, G. & Blackburn, D.S. (2008). Adaptive coping with rheumatoid arthritis: the transforming nature of response shift. *Chronic Illness*, **4**, 219-230.

Singer, J. D. & Willett, J. B. (2005). *Longitudinal research: Current status and future prospects*. Consultado a 20 de mayo de 2008 en la World Wide Web: c

Singer, J. D. y Willett, J. B. (2006). *Longitudinal data analysis: Present status; future prospects*. Consultado a 20 de mayo de 2008 en la World Wide Web: <http://gseweb.harvard.edu/~faculty/singer/Presentations/Longitudinal%20research,%20GCP%202006.ppt>.

---

Sloan, J.; Symonds, T.; Vargas-Chanes, D. & Brooke, F. (2003). Practical guidelines for assessing the clinical significance of health-related quality of life changes within clinical trials. *Drug Information Journal*, **37**, 23–21.

Sörbom, D. (1989). Model modification. *Psychometrika*, **54**, 371-384.

Spitzer, W.O.; Dobson, A.J.; Hall, J.; Hall, J.; Chesterman, E.; Levi, J.; Shepherd, R.; Battista, R.N. & Catchlove, B.R. (1981). Measuring the Quality of Life of Cancer Patients: a Concise QL-Index for Use by Physicians. *Journal of Chronic Disease*, **34**, 585-597.

Sprangers, M.A. & Schwartz, C.E. (1999a). Integrating response shift into health-related quality of life research: a theoretical model. *Social Science and Medicine*, **48**, 1507-1515.

Sprangers, M.A. & Schwartz, C.E. (1999b). The challenge of response shift for quality-of-life-based clinical oncology research. *Annals of Oncology*, **10**(7), 747-749.

Sprangers, M.A.G.; van Dam, F.S.A.M; Broersen, J.; Lodder, L.; Wever, L.; Visser, M.R.M.; Oosterveld, P. & Smets, E.M.A. (1999). Revealing response shift in longitudinal research on fatigue: The use of the then-test approach. *Acta Oncologica*, **38**, 709-718.

Stewart, B.W. & Kleihues, P. (2003). *World Cancer Report*. Lyon: International Agency Research on Cancer.

Stiefel, F.; Krenz, S.; Zdrojewski, C.; Stagno, D.; Fernandez, M.; Bauer, J.; Fucina, N.; Lüthi, F.; Leyvraz, S.; Borasio G. D. & Fegg, M. (2008). Meaning in life assessed with the "Schedule for Meaning in Life Evaluation" (SMiLE): a comparison between a cancer patient and student sample. *Support Care Cancer*, **16**, 1151-1155.

Stiggelbout, A.M.; Vogel-Voogt, E.; Noordijk, E.M. & Vlieland, T.P.M. (2008). Individual quality of life: adaptive conjoint analysis as an alternative for direct weighting? *Quality Life Research*, **17**, 641–649.

Streiner, D. L. & Norman, G. R. (1995). *Health Measurement Scales: a practical guide to their development and use*. New York: Oxford University Press.

Sullivan, M.; Karlsson, J. & Ware, J.E. (1995). The Swedish SF-36 health survey. I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. *Social Science and Medicine*, **41**, 1349-1358.

Suls, J.; David, J.P. & Harvey, J.H. (1996). Personality and Coping: Three Generations of Research. *Journal of Personality*, **64**, 711-735.

- Taguchi, R; Yamazaki, Y; Takayama, T & Saito, M. (2008). Life-lines of relapsed breast cancer patients: a study of post-recurrence distress and coping strategies. *Japanese Journal of Health and Human Ecology*, **74**(5), 217-235.
- Taris, T. (2000). *A Primer in Longitudinal Data Analysis*. London: Ed. Sage.
- Testa, M. & Simonson, D.C. (1996). Assessment of quality-of life outcomes. *The New England Journal Medicine*, **334**(13), 835-840.
- Thoma, A.; Cornacchi, S.; Lovrics, P. & Goldsmith, C. (2008). User's Guide to the Surgical Literature: How to Assess an Article on Health-Related Quality of Life. *Canadian Journal of Surgery*, **51**(3), 215-224.
- Torrance, G.W. & Feeny, D.H. (1989). Utilities and quality-adjusted life years. *International Journal of Technology Assessment in Health Care*, **5**(4), 559-575.
- Torrance, G.W., Feeny, D. & Furlong, W. (2001) Visual analog scales: Do they have a role in the measurement of preferences for health states? *Medical Decision Making*, **21**(4), 329-334.
- Tucker, L.R. (1951). *A method for synthesis of factor analysis studies*. Report 984. Department of the Army, Personnel Research Section. Washington DC.
- Ubel, P.; Loewenstein, G. & Jepson, C. (2003). Whose Quality of Life? A Commentary Exploring Discrepancies between Health State Evaluations of Patients and the General Public. *Quality Life Research*, **12**, 599-607.
- Ubel, P.; Loewenstein, G.; Hershey, J.; Baron, J.; Mohr T. & Asch, D. (2001). Do Nonpatients Underestimate the Quality of Life Associated with Chronic Health Conditions because of a Focusing Illusion? *Medical Decision Making*, **21**, 190-199.
- Velenik, V. Oblak, I. & Anderluh, F. (2008). Quality of life in patients after combined modality treatment of rectal cancer: Report of a prospective phase II study. *Radiology and Oncology*, **42**(4), 207-214.
- Velikova, G.; Booth, L.; Smith, A.B.; Brown, P.M.; Lynch, P.; Brown, J.M. & Selby PJ. (2004). Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *Journal of Clinical Oncology*, **22**(4), 714-724.
- Ventegodt, S.; Merrick, J. & Andersen, N.J. (2003). Quality of Life Theory I. The IQOL Theory: An Integrative Theory of the Global Quality of Life Concept. *The Scientific World Journal*, **3**, 1030-1040.

---

Vicente-Galindo, P. (2003). *Contribuciones al Análisis de Datos de Calidad de Vida relacionada con la Salud*. Tesis Doctoral. Universidad de Salamanca. España.

Vicente-Galindo, P. (2007). El Metodo Statis en la Detección de Cambio Beta y/o Gamma, en Estudios de Calidad de Vida, *XII Reunión Científica del GAB y I Encuentro Argentino-Chileno de Biometría*, San Martín de los Andes - Patagonia Argentina.

Vicente-Villardón, J.L.; Galindo-Villardón, M.P & Vicente-Galindo, M.P. (2004). Modelos factoriales para datos longitudinales con estructuras comunes y específicas. *Metodología de las Ciencias del Comportamiento*, Suplemento 2004, 619-627.

Visser, M.R., Smets, E.A.; Sprangers, M.A. & Haes, H.J. (2000). How Response Shift may affect the measurement of Change in Fatigue. *Journal of Pain and Symptom Management*, **20**(1), 12-18.

Visser, M.R.M.; Oort, F.J. & Sprangers, M.A.G. (2005). Methods to detect response shift in quality of life data: A convergent validity study. *Quality of Life Research*, **14**, 629–639.

Vriens, M. & Melton, E. (2002). Managing Missing Data. *Marketing Research*, **14**, 12–17.

Walters, S.J. & Brazier, J.E. (2003). What is the relationship between the minimally important difference and health state utility values? *Health and Quality of Life Outcomes*, **1**, 4. Consultado a 29 de enero de 2009 en la World Wide Web: <http://www.hqlo.com/content/1/1/4>

Ware, J., Snow, K., Kosinski, M., & Gandek, B. (1993). *SF-36 Health Survey: Manual and interpretation guide*. Boston: The Health Institute, New England Medical Centre, Boston, MA.

Watson, M.; Law, M.; Maguire, P.; Robertson, B.; Greer, S.; Bliss, J.M. & Ibbotson, T. (1992). Further development of a quality of life measure for cancer patients: the Rotterdam Symptom Checklist (revised). *Psycho- Oncology*, **1**, 35-44.

Westerman, M.J.; Hak, T.; Echteld, M.A. Groen, H.J. & van der Wal, G. (2007b). Change in what matters to palliative patients: eliciting information about adaptation with SEIQoL-DW. *Palliative Medicine*, **21**, 581–586.

Westerman, M.J.; Hak, T.; Sprangers, M.A.; Groen, H.J.; van der Wal, G. & The, A.M. (2008). Listen to their answers! Response behaviour in the measurement of physical and role functioning. *Quality of Life Research*, **17**, 549–558.



- Westerman, M.J.; Sprangers, M.A.; Groen, H.J.; van der Wal, G. & Hak, T. (2007a). Small-cell lung cancer patients are just 'a little bit' tired: response shift and self-presentation in the measurement of fatigue. *Quality of Life Research*, **16**, 853–861.
- Westhovens, R.; Cole, J.C.; Li, T.; Martin, M.; Maclean, R.; Lin, P.; Blaisdell, B.; Wallenstein, G.V.; Aranda, R. & Sherrer, Y. (2006). Improved health-related quality of life for rheumatoid arthritis patients treated with abatacept who have inadequate response to anti-TNF therapy in a double-blind, placebo-controlled, multicentre randomized clinical trial. *Rheumatology*, **45**(10), 1238-1246.
- WHO (World Health Organization) (2009). *Cancer*. Fact sheet n° 297. Consultado a 19 de agosto de 2009 en la World Wide Web: <http://www.who.int/mediacentre/factsheets/fs297/en/index.html>
- Wiel, H.; Geerts, E. & Hoekstra-Weebers, J. (2007). Explaining inconsistent results in cancer quality of life studies: the role of the stress–response system. *Psycho-Oncology*, **17**, 174–181.
- Wiles, C. (2008). Physiotherapy and related activities in multiple sclerosis. *Multiple Sclerosis*, **14**, 863-871.
- Willett, J. B. & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, **116**, 363-381.
- Wilson, I.B. & Clearly, P.D. (1995). Linking clinical variables with health related quality of life: a conceptual model of patient outcomes. *Journal of the American Medical Association*, **273**(1), 59-65.
- Winter, L.; Moss, M.S. & Hoffman, C. (2009). Affective Forecasting and Advance Care Planning: Anticipating Quality of Life in Future Health Statuses. *Journal of Health Psychology*, **14**(3), 447–456.
- Wrosch, C. & Scheier, M.F. (2003). Personality and quality of life: The importance of optimism and goal adjustment. *Quality of Life Research*, **12**, 59-72.
- Wu, C.H. (2009). Enhancing quality of life by shifting importance perception among life domains. *Journal of Happiness Studies*, **10**, 37-47.
- Wyrwich, K.W.; Tierney W.M. & Wolinsky, F.D. (2002). Using the standard error of measurement to identify important changes on the Asthma Quality of Life Questionnaire. *Quality of Life Research*, **11**(1), 1-7.

---

Yates, P.; Aranda, S.; Hargraves, M.; Mirolo, B.; Clavarino, A.; McLachlan, S.A. & Skerman, H. (2005). Randomized Controlled Trial of an Educational Intervention for Managing Fatigue in Women Receiving Adjuvant Chemotherapy for Early-Stage Breast Cancer. *Journal of Clinical Oncology*, **23**(25), 6027-6036.

Yeh, C.H. (2002). Health-related quality of life in pediatric patients with cancer: A structural equation approach with the Roy Adaptation Model. *Cancer Nursing*, **25**(1), 74-80.

Zee, B. (1998). Growth Curve Model Analysis for Quality of Life Data. *Statistics in Medicine*, **17**, 757-766.

Zittoun, R. (1986). Objectives of the study of the quality of life in oncology. *Bull. Cancer*, **35** (1), 79-85.

Zwinderman, A. (1990). The measurement of change of quality of life in clinical trials. *Statistics in Medicine*, **9**, 931-942.

---

## **ANEXO I – SF-36**

---



28608

Código

Data

Aplicação nº

Preencha o questionário de acordo com o exemplo, utilizando **caneta azul ou preta.**

Preencha

assim

assim não

## ESTADO DE SAÚDE SF-36 V2

**INSTRUÇÕES:** As questões que se seguem pedem-lhe a sua opinião sobre a sua saúde, a forma como se sente e sobre a sua capacidade de desempenhar as suas actividades habituais.

Pedimos que leia com atenção cada pergunta e que responda o mais honestamente possível. Se não tiver a certeza sobre a resposta a dar, dê-nos a que achar mais apropriada e, se quiser, escreva um comentário a seguir à pergunta.

**Para as perguntas 1 e 2, por favor assinale a opção que melhor descreve a sua saúde.**

1. Em geral, diria que a sua saúde é:

Óptima    Muito Boa    Boa    Razoável    Fraca

2. Comparando com o que acontecia há um ano como descreve o seu estado geral actual:

Muito Melhor    Com algumas Melhoras    Aproximadamente igual    Um pouco pior    Muito Pior

As perguntas que se seguem são sobre actividades que executa no seu dia-a-dia. Será que a sua saúde o limita nestas actividades? Se sim, quanto?

	Sim, muito limitado	Sim, um pouco limitado	Não, nada limitado
3. <b>Actividades violentas</b> , tais como correr, levantar pesos, participar em desportos extenuantes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. <b>Actividades moderadas</b> , tais como deslocar uma mesa ou aspirar a casa, trabalhar no jardim	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. <u>Levantar ou pegar nas compras de mercearia</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. <u>Subir vários lanços de escada</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. <u>Subir um lanço de escada</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. <u>Inclinar-se, ajoelhar-se ou baixar-se</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. <u>Andar mais de 1 km</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. <u>Andar várias centenas de metros</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. <u>Andar uma centena de metros</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. <u>Tomar banho ou vestir-se sozinho</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Durante as últimas 4 semanas teve, no seu trabalho ou actividades diárias, algum dos problemas apresentados a seguir como consequência do seu estado de saúde físico?

Quanto tempo, nas <b>últimas 4 semanas</b>	Sempre	A maior parte do tempo	Algum tempo	Pouco tempo	Nunca
13. Diminuiu o <b>tempo gasto</b> a trabalhar ou noutras actividades?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. Fez <b>menos</b> do que queria?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. Sentiu-se limitado no <b>tipo</b> de trabalho ou outras actividades?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. Teve <b>dificuldade</b> em executar o seu trabalho ou outras actividades (por exemplo, foi preciso mais esforço)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Durante as últimas 4 semanas, teve com o seu trabalho ou com as suas actividades diárias, algum dos problemas apresentados a seguir devido a quaisquer problemas emocionais (tal como sentir-se deprimido ou ansioso) ?

Quanto tempo, nas últimas 4 semanas	Sempre	A maior parte do tempo	Algum tempo	Pouco tempo	Nunca
17. Diminuiu o <b>tempo gasto</b> a trabalhar ou noutras actividades?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. Fez <b>menos</b> do que queria?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. Executou o seu trabalho ou outras actividades <b>menos cuidadosamente</b> do que era costume?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Para cada uma das perguntas 20, 21 e 22, por favor assinale a que melhor descreve a sua saúde.**

20. Durante as últimas 4 semanas, em que medida é que a sua saúde física ou problemas emocionais interferiram no seu relacionamento social normal com a família, amigos, vizinhos ou outras pessoas?

Absolutamente nada    Pouco    Moderadamente    Bastante    Imenso

21. Durante as últimas 4 semanas teve dores?

Nenhumas    Muito Fracas    Ligeiras    Moderadas    Muito fortes

22. Durante as últimas 4 semanas, de que forma é que a dor interferiu com o seu trabalho normal (tanto fora de casa como o trabalho doméstico) ?

Absolutamente nada    Pouco    Moderadamente    Bastante    Imenso

As perguntas que se seguem pretendem avaliar a forma como se sentiu e como lhe correram as coisas nas últimas 4 semanas.

Quanto tempo, nas últimas 4 semanas	Sempre	A maior parte do tempo	Algum tempo	Pouco tempo	Nunca
23. Se sentiu cheio de vitalidade?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. Se sentiu muito nervoso?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. Se sentiu tão deprimido que nada o animava?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. Se sentiu calmo e tranquilo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. Se sentiu com muita energia?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. Se sentiu deprimido?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29. Se sentiu estafado?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30. Se sentiu feliz?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31. Se sentiu cansado?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

32. Durante as últimas 4 semanas, até que ponto é que a sua saúde física ou problemas emocionais limitaram a sua **actividade social** (tal como visitar amigos ou familiares próximos)?

Sempre    A maior parte do tempo    Algum tempo    Pouco tempo    Nunca

Por favor, diga em que medida são verdadeiras ou falsas as seguintes afirmações.

	Absolutamente verdade	Verdade	Não sei	Falso	Absolutamente Falso
33. Parece que adoço mais facilmente do que os outros	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34. Sou tão saudável como qualquer outra pessoa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35. Estou convencido que a minha saúde vai piorar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36. A minha saúde é óptima	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**MUITO OBRIGADO**

---

## **ANEXO II – RSCL**

---



45114

Código Data Aplicação nº 

Preencha o questionário de acordo com o exemplo, utilizando **caneta azul ou preta**.

Preencha

assim



assim não



## ROTTERDAM SYMPTOM CHECKLIST (RSCL)

Este questionário contém perguntas acerca dos seus sintomas. Por favor, leia cada uma e faça uma cruz no quadrado associado à resposta que mais se aproxime de como se tem sentido **durante a última semana**.

- |  |                                   |                                       |  |  |
|--|-----------------------------------|---------------------------------------|--|--|
| 1. Falta de Apetite                    | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 2. Irritabilidade                      | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 3. Cansaço                             | <input type="checkbox"/> Nenhum   | <input type="checkbox"/> Pouco        | <input type="checkbox"/> Algum         | <input type="checkbox"/> Bastante        |
| 4. Preocupações                        | <input type="checkbox"/> Nenhumas | <input type="checkbox"/> Poucas       | <input type="checkbox"/> Algumas       | <input type="checkbox"/> Bastantes       |
| 5. Dores Musculares (músculos doridos) | <input type="checkbox"/> Nenhumas | <input type="checkbox"/> Poucas       | <input type="checkbox"/> Algumas       | <input type="checkbox"/> Bastantes       |
| 6. Depressão                           | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 7. Falta de energia                    | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 8. Dores                               | <input type="checkbox"/> Nenhumas | <input type="checkbox"/> Poucas       | <input type="checkbox"/> Algumas       | <input type="checkbox"/> Bastantes       |
| 9. Nervosismo                          | <input type="checkbox"/> Nenhum   | <input type="checkbox"/> Pouco        | <input type="checkbox"/> Algum         | <input type="checkbox"/> Bastante        |
| 10. Náuseas (enjoo)                    | <input type="checkbox"/> Nenhumas | <input type="checkbox"/> Poucas       | <input type="checkbox"/> Algumas       | <input type="checkbox"/> Bastantes       |
| 11. Esperança no futuro                | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 12. Dificuldade em adormecer           | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 13. Dores de cabeça                    | <input type="checkbox"/> Nenhumas | <input type="checkbox"/> Poucas       | <input type="checkbox"/> Algumas       | <input type="checkbox"/> Bastantes       |
| 14. Vômitos                            | <input type="checkbox"/> Nenhum   | <input type="checkbox"/> Poucos       | <input type="checkbox"/> Alguns        | <input type="checkbox"/> Bastantes       |
| 15. Tonturas                           | <input type="checkbox"/> Nenhumas | <input type="checkbox"/> Poucas       | <input type="checkbox"/> Algumas       | <input type="checkbox"/> Bastantes       |
| 16. Falta de interesse sexual          | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 17. Sentimento de solidão              | <input type="checkbox"/> Nenhum   | <input type="checkbox"/> Pouco        | <input type="checkbox"/> Algum         | <input type="checkbox"/> Bastante        |
| 18. Tensão nervosa                     | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 19. Ansiedade                          | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 20. Obstipação (prisão de ventre)      | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 21. Diarreia                           | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 22. Azia, arrotos                      | <input type="checkbox"/> Nenhum   | <input type="checkbox"/> Poucos       | <input type="checkbox"/> Alguns        | <input type="checkbox"/> Bastantes       |
| 23. Calafrios (arrepios)               | <input type="checkbox"/> Nenhum   | <input type="checkbox"/> Poucos       | <input type="checkbox"/> Alguns        | <input type="checkbox"/> Bastantes       |
| 24. Formigueiros nas mãos/pés          | <input type="checkbox"/> Nenhum   | <input type="checkbox"/> Poucos       | <input type="checkbox"/> Alguns        | <input type="checkbox"/> Bastantes       |
| 25. Acordar sobressaltado              | <input type="checkbox"/> Nunca    | <input type="checkbox"/> Poucas vezes | <input type="checkbox"/> Algumas vezes | <input type="checkbox"/> Bastantes vezes |
| 26. Dor ao engolir                     | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 27. Queda de cabelo                    | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 28. Ardor nos olhos                    | <input type="checkbox"/> Nenhum   | <input type="checkbox"/> Pouco        | <input type="checkbox"/> Algum         | <input type="checkbox"/> Bastante        |
| 29. Falta de ar                        | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |
| 30. Secura da boca                     | <input type="checkbox"/> Nenhuma  | <input type="checkbox"/> Pouca        | <input type="checkbox"/> Alguma        | <input type="checkbox"/> Bastante        |



Apresenta-se uma lista de actividades.  
Não pretendemos saber se as faz habitualmente, mas se é capaz de as fazer presentemente.  
Por favor, assinale a resposta que se aplica melhor à sua situação na **última semana**.

	Incapaz	Só com ajuda	Com dificuldade Sem ajuda	Sem ajuda
1. Cuidar de mim (lavar-me, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Andar por casa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Lida da casa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Subir escadas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Tarefas ocasionais (consertos, tricot, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Andar na rua	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Ir às compras	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Ir trabalhar	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9. Tendo em conta estes aspectos, como descreve a sua qualidade de vida durante a semana passada?

- Excelente
- Boa
- Moderadamente Boa
- Nem boa nem má
- Moderadamente má
- Má
- Muito má

Por favor, verifique se respondeu a todas as perguntas.  
Muito Obrigada pela sua colaboração.

**A preencher pelo Profissional de Saúde**

Avaliação da Qualidade de Vida do Paciente:

- Excelente  Moderadamente boa  Muito má
- Boa  Moderadamente má
- Nem boa nem má  Má



---

## **ANEXO III – EL MÉTODO STATIS**

---

## EL MÉTODO STATIS

Los métodos STATIS y STATIS Dual posibilitan la exploración simultánea de  $T$  matrices  $X_t$  de dimensión  $I \times J$ ,  $t=1, \dots, T$ , de datos cuantitativos obtenidos en diferentes tiempos sobre los mismos individuos pudiendo eventualmente las variables ser diferentes (STATIS) o sobre las mismas variables pudiendo eventualmente los individuos ser diferentes (STATIS Dual). Nótese que, el método STATIS enfatiza las posiciones relativas de los individuos, mientras el STATIS Dual enfatiza las relaciones entre las variables.

Estos métodos se desarrollan en cuatro etapas esenciales:

- 1) Estudio de la **interestructura**, cuyo objetivo es proceder a una comparación global de la estructura de las  $T$  matrices de datos;
- 2) Búsqueda de una estructura común a las matrices en estudio - la **matriz compromiso**;
- 3) Estudio de la **intraestructura** que consiste en representar la estructura de cada matriz de datos en un espacio de baja dimensión;
- 4) Interpretación de las **trayectorias** individuales en el diagrama factorial.

Veamos en qué consisten estas cuatro etapas:

### *1) Análisis de la Interestructura*

En el caso del STATIS, la estructura de cada matriz  $X_t$  es captada a través de la configuración

$$W_t = X_t M_t X_t'$$

que es la matriz de productos escalares entre individuos, siendo  $M_t$  la métrica del espacio de individuos. Cada matriz  $W_t$  define por tanto las distancias entre

individuos.  $M_t$  es habitualmente la matriz identidad de orden  $J$ , sin embargo, si el número de variables en las matrices  $X_t$  es diferente, es necesario introducir un término de compensación tomando los elementos de la diagonal iguales a  $\frac{1}{J_t}$ .

En el caso del STATIS Dual la estructura de cada matriz  $X_t$  es captada a través de la configuración

$$C_t = X_t' D_t X_t$$

que es la matriz de varianzas-covarianzas entre variables, siendo  $D_t$  la métrica del espacio de las variables.

A fin de hacer posible la comparación de las  $T$  matrices de datos y averiguar la existencia de una eventual estructura común a los individuos (en el caso del STATIS) o a las variables (en el caso del STATIS Dual), se define el producto interno de Hilbert-Schmidt entre configuraciones, el cual, para el STATIS viene

$$\langle W_t | W_{t'} \rangle = \text{Tr}(W_t D W_{t'}' D)$$

donde  $D$  es una matriz diagonal de orden  $I$ , que define la métrica en el espacio de las variables (pesos atribuidos a los individuos) y cuyos elementos de la diagonal son iguales a  $\frac{1}{I}$ .  $\langle W_t | W_{t'} \rangle$  traduce pues la similitud entre las matrices  $W_t$  y  $W_{t'}$ .

En el caso del STATIS Dual el producto interno de Hilbert-Schmidt entre configuraciones se define como

$$\langle C_t | C_{t'} \rangle = \text{Tr}(C_t M C_{t'}' M)$$

donde  $M$  define la métrica en el espacio de los individuos.

Los productos definidos inducen a una norma que es definida por  $\|W_t\|_{HS}^2 = \langle W_t | W_t \rangle$  (STATIS) o  $\|C_t\|_{HS}^2 = \langle C_t | C_t \rangle$  (STATIS Dual), respectivamente y, e por tanto, a una distancia:

$$d_{HS}(W_t | W_{t'}) = \|W_t - W_{t'}\|_{HS} = \sqrt{\|W_t\| + \|W_{t'}\| - 2\langle W_t | W_{t'} \rangle_{HS}} \quad (\text{STATIS}) \text{ o}$$

$$d_{HS}(C_t | C_{t'}) = \|C_t - C_{t'}\|_{HS} = \sqrt{\|C_t\| + \|C_{t'}\| - 2\langle C_t | C_{t'} \rangle_{HS}} \quad (\text{STATIS Dual}).$$

En el caso en que las configuraciones poseen normas distintas, es conveniente trabajar con las configuraciones normadas.

En el análisis de la interestructura se recurre habitualmente al Coeficiente de Correlación Vectorial entre dos estudios  $t$  y  $t'$  propuesto por Escoufier (1973) e definido de la siguiente forma:

$$RV(t, t') = \left\langle \frac{W_t}{\|W_t\|_{HS}} \middle| \frac{W_{t'}}{\|W_{t'}\|_{HS}} \right\rangle_{HS} = \frac{\text{Tr}(W_t D W_{t'} D)}{\sqrt{\text{Tr}(W_t D)^2} \sqrt{\text{Tr}(W_{t'} D)^2}} \quad (\text{STATIS}) \text{ o}$$

$$RV(t, t') = \left\langle \frac{C_t}{\|C_t\|_{HS}} \middle| \frac{C_{t'}}{\|C_{t'}\|_{HS}} \right\rangle_{HS} = \frac{\text{Tr}(C_t M C_{t'} M)}{\sqrt{\text{Tr}(C_t M)^2} \sqrt{\text{Tr}(C_{t'} M)^2}} \quad (\text{STATIS Dual}).$$

Los coeficientes  $RV$  surgen organizados en una matriz  $RV$ , cuadrada de orden  $T$  y posibilitan una interpretación fácil de la interestructura en la medida en que son no negativos y varían entre 0 e 1. Si  $RV(t, t')=1$ , entonces la distancia entre los estudios  $t$  y  $t'$  es nula y las estructuras en cuestión son coincidentes. La distancia entre dos estudios normados  $t$  y  $t'$ , es dada por:

$$d_{HS} \left( \frac{W_t}{\|W_t\|_{HS}} \middle| \frac{W_{t'}}{\|W_{t'}\|_{HS}} \right) = \sqrt{2(1 - RV(t, t'))} \quad (\text{STATIS}) \text{ o}$$

$$d_{HS} \left( \frac{C_t}{\|C_t\|_{HS}} \mid \frac{C_{t'}}{\|C_{t'}\|_{HS}} \right) = \sqrt{2(1 - RV(t, t'))} \text{ (STATIS Dual).}$$

Con el objetivo de comparar globalmente la estructura de las  $T$  matrices de datos, es posible obtener una representación euclídea de las mismas sometiendo la matriz  $RV$  a un Análisis en Componentes Principales (ACP). En el espacio de las Componentes Principales, cada matriz viene representada por un punto y la distancia entre puntos en el plano refleja el grado de semejanza entre matrices, posibilitando el análisis de la interestructura. Los coeficientes  $RV$  se interpretan como los cosenos de los ángulos entre los vectores que representan las matrices y aproximan la correlación vectorial entre las mismas, por lo que, si hay una estructura común, los ángulos son pequeños y la mayor parte de la variabilidad es explicada por el primer eje de la representación. Por eso, uniendo los puntos de la imagen euclídea con el origen del sistema de ejes, se obtiene una estimación gráfica de la correlación entre las estructuras mediante el ángulo que forman los vectores obtenidos. En esta imagen euclídea se pueden incluir matrices como elementos suplementarios, las cuales no intervienen en la determinación de los ejes pero se desea observar su posicionamiento en la interestructura.

## 2) *Matriz Compromiso*

Si el análisis de la interestructura permite concluir que, efectivamente las matrices tienen estructura similar, se procede a la construcción de una matriz que resuma la información proveniente de las configuraciones en estudio - la matriz compromiso. La matriz compromiso no es más que una media ponderada de las configuraciones  $W_t$  o  $C_t$  siendo, por tanto, la más correlacionada con todas las matrices y viene dada por:

$$W = \sum_{t=1}^T \alpha_t W_t \text{ (STATIS)} \quad \text{o} \quad C = \sum_{t=1}^T \alpha_t C_t \text{ (STATIS Dual)}$$

siendo los coeficientes  $\alpha_t$  las componentes del primer vector propio resultante de la diagonalización de la matriz  $RV$ .

### 3) *Análisis de la Intraestructura*

Esta etapa consiste en someter la matriz compromiso  $W$  o  $C$  a un ACP, permitiendo así obtener una imagen euclídea compromiso de los individuos o variables en estudio. Las coordenadas de los objetos de la  $t$ -ésima matriz en el espacio de las  $p$  componentes principales (espacio compromiso), son dadas por

$$T_t = W_t L E \text{ (STATIS)} \quad \text{o} \quad T_t = C_t L E \text{ (STATIS Dual)}$$

siendo  $E$  una matriz diagonal  $f \times f$  cuyos elementos principales son el inverso de las raíces cuadradas de los valores propios compromiso y  $L$  una matriz de dimensión  $I \times f$  que contiene los *scores* del ACP sobre la matriz compromiso.

En el espacio compromiso, la distancia entre puntos se interpreta como la distancia compromiso entre los individuos o variables por ellos representados. Por otra parte, los ángulos definidos entre las variables y los ejes, se interpretan en términos de correlación, lo que permite, a su vez, interpretar las posiciones de los individuos relativamente a los ejes compromiso.

### 4) *Interpretación de las Trayectorias*

Proyectando las líneas de las matrices originales en el espacio compromiso como elementos suplementarios, es posible analizar la evolución de las trayectorias a lo largo de los distintos grupos. Las trayectorias obtenidas definen los cambios en la posición de los individuos o variables en la configuración consenso a lo largo del tiempo, permitiendo por tanto analizar la forma como los individuos y variables contribuyen para las distancias entre las configuraciones.

Así, trayectorias de gran amplitud reflejan cambios en la estructura de los individuos o variables, mientras trayectorias envolventes describen evoluciones medias.

Al trabajar con datos de CVRS, dispondremos de  $T$  matrices referentes a  $T$  tiempos con  $I$  filas (pacientes) y  $J$  columnas (ítems). Dado que el objetivo reside en la comparación de las estructuras de covariación entre los ítems, en la detección de *response shift*, se enfatiza el STATIS Dual, que puede sintetizarse en el esquema presentado en la Fig.A1.1.

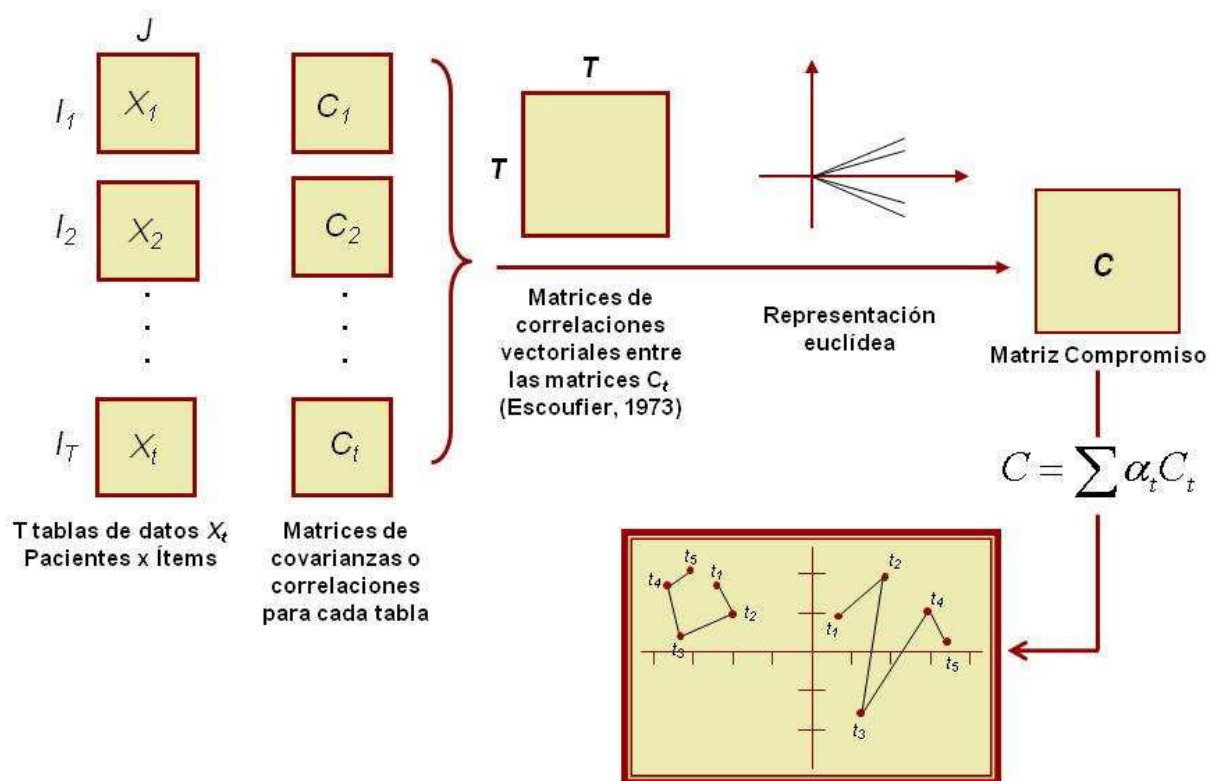


Fig.A1.1. El Método STATIS Dual.

---

**ANEXO IV – EL MODELO DE  
ZWINDERMAN**

---



## EL MODELO DE ZWINDERMAN

Empezaremos por ver el Modelo de Zwinderman para dos instantes de medida y luego su extensión a más que dos tiempos.

### *Modelo de Zwinderman para dos instantes de medida*

Considérese un modelo de medida en el que  $X_v$  es la suma de las puntuaciones relativas al paciente  $v$  sobre todos los ítems relativos a una determinada variable latente (por ejemplo, la movilidad). Además, sea

$$X_v = \tau_v + \varepsilon_v$$

donde  $\tau_v$  es el valor verdadero de una determinada variable latente (por ejemplo, la verdadera movilidad) del paciente  $v$  y  $\varepsilon_v$  una componente de error. La fiabilidad (*reliability*) se define como el cociente entre la varianza de las verdaderas componentes  $\sigma_\tau^2$  y la varianza observada  $\sigma_X^2$  en la muestra:

$$REL = \frac{\sigma_\tau^2}{\sigma_X^2}.$$

Obviamente *REL* depende de la varianza del verdadero valor de la variable latente (por ejemplo, movilidad) en la muestra (por ejemplo, si pacientes que tienen su movilidad seriamente comprometida abandonan el estudio, *REL* disminuye).

Consideremos ahora las respuestas de  $N$  pacientes a  $k$  ítems en dos momentos en el tiempo. Sea  $X_{vi1}$  la respuesta del paciente  $v$  al ítem  $i$  en el instante  $T_1$  y sea  $X_{vi2}$  la respuesta del paciente  $v$  al ítem  $i$  en el instante  $T_2$ .

Se asume una variable latente  $\theta_i$  para cada ítem y  $\theta_{vi}$  representa la posición del paciente  $v$  en el *continuum* latente  $\theta_i$  en el instante  $T_1$ . Debido a que se trata de un proceso temporal que envuelve uno o más tratamientos,  $\theta_{vi}$  cambia

segundo una cantidad  $\delta_v$ . Luego, la posición del paciente  $v$  en el *continuum* latente  $\theta_i$  en el instante  $T_2$  es dada por  $\theta_{vi2} = \theta_{vi} + \delta_v$ .

Según el modelo de Rash, la probabilidad condicional de observar una respuesta positiva del paciente  $v$  al ítem  $i$  en el instante  $T_1$  viene modelizada por la curva logística:

$$p(X_{vi1} = 1 | \theta_{vi}) = \frac{\exp(\theta_{vi})}{1 + \exp(\theta_{vi})} = f_{vi}. \quad (1)$$

De forma equivalente, la probabilidad condicional de observar una respuesta positiva del paciente  $v$  al ítem  $i$  en el instante  $T_2$  es

$$p(X_{vi2} = 1 | \theta_{vi2}) = \frac{\exp(\theta_{vi} + \delta_v)}{1 + \exp(\theta_{vi} + \delta_v)} = g_{vi}. \quad (2)$$

El cambio  $\delta_v$  es modelizado como una función del tiempo y del tratamiento que el paciente  $v$  recibe:

$$\delta_v = q_v \mu + \tau \quad (3)$$

donde  $\tau$  es el cambio inherente al proceso temporal y  $\mu$  el cambio debido al tratamiento. El peso  $q_v$  toma el valor 1 si  $v$  recibe el tratamiento y 0 caso contrario. Los efectos de otros tratamientos pueden también influenciar eventuales cambios. Todavía, tales efectos, no pueden ser separados de los efectos del proceso temporal.

De una forma más general, el parámetro de cambio  $\delta_v$  puede modelizarse como

$$\delta_v = \sum_{j=1}^m q_{vj} \mu_j + \tau \quad (4)$$

donde  $\mu_j$  representa el efecto del tratamiento  $j$  y  $q_{vj}$  denota el peso del tratamiento  $j$  para el paciente  $v$  (el peso del tratamiento puede ser dicotómico pero puede también ser continuo, denotando dosis o intensidad de tratamiento).

Asumiendo independencia condicional, la verosimilitud del modelo especificado en (1) y (2) puede escribirse como

$$L(\theta, \mu, \tau) = \prod_{v=1}^N \prod_{i=1}^k f_{vi}^{x_{vi1}} (1-f_{vi})^{1-x_{vi1}} g_{vi}^{x_{vi2}} (1-g_{vi})^{1-x_{vi2}} \quad (5)$$

donde  $Q$  es la matriz de pesos y  $X_1$  y  $X_2$  las respuestas en los instantes  $T_1$  y  $T_2$ , respectivamente. La verosimilitud es una función de los parámetros  $\theta$ ,  $\mu$  y  $\tau$  y puede ser vista como la verosimilitud de un modelo de regresión logística.

Hay que hacer la distinción entre parámetros incidentales y estructurales. Si se incluyen en la muestra más pacientes, el número de parámetros  $\theta$  crece rápidamente, mientras el número de parámetros efecto  $\mu$  y  $\tau$  se mantiene constante. En tales circunstancias,  $\theta$ ,  $\mu$  y  $\tau$  no pueden ser estimados de forma consistente maximizando (5).

Los parámetros  $\theta$  pueden considerarse parámetros perturbadores<sup>4</sup>. En este caso, se centra la atención en los parámetros efecto  $\mu$  y  $\tau$  y luego se busca una cadena de estimación condicional independiente de  $\theta$ . La única forma de obtener independencia exacta del parámetro perturbador es usar estadísticos suficientes<sup>5</sup> para  $\theta$ .

La suma de las puntuaciones  $x_{i1} + x_{i2}$  es un estadístico suficiente mínimo<sup>6</sup> para  $\theta_{ij}$ . Además, se prueba que los parámetros estructurales pueden ser estimados de forma consistente maximizando la verosimilitud condicional de las respuestas a los ítems dados los estadísticos suficientes para los parámetros incidentales. Sea  $X_1$  ( $N \times k$ ) la matriz de datos relativa al instante  $T_1$  y  $X_2$  ( $N \times k$ ) la matriz de datos relativa al instante  $T_2$ . Entonces, la verosimilitud condicional de  $X_1$  y  $X_2$  dado  $X_1 + X_2$  es dada por

<sup>4</sup> **Parámetro Perturbador** (*Nuisance Parameter*) – parámetro que, no siendo de interés inmediato, tiene que ser tomado en consideración en el análisis y estimación de los otros parámetros (ej. en una distribución Normal,  $\sigma^2$  es el parámetro perturbador cuando  $\mu$  es el parámetro de interés).

<sup>5</sup> **Estadístico Suficiente** (*Sufficient Statistic*) – Un estadístico  $T(X)$  es suficiente para  $\theta$  si la distribución de probabilidad condicional de los datos  $X$  dado el estadístico  $T(X)$  no depende de  $\theta$ .

<sup>6</sup> **Estadístico Suficiente Mínimo** (*Minimal Sufficient Statistic*) – Un estadístico suficiente para un parámetro  $\theta$  se dice estadístico suficiente mínimo si puede expresarse como función de algún estadístico suficiente para  $\theta$ .

$$p(X_1, X_2 | X_1 + X_2) = \prod_{v=1}^N \prod_{i=1}^k p(X_{vi1} = x_{vi1}, X_{vi2} = x_{vi2} | x_{vi1} + x_{vi2}) \quad (6)$$

que, después de alguna álgebra viene

$$L_c(\mu, \tau) = \prod_{v=1}^N \prod_{i=1}^k \left[ \frac{\exp(x_{vi2} \delta_v)}{1 + \exp(\delta_v)} \right]^{(x_{vi1} - x_{vi2})^2}. \quad (7)$$

Los estimadores de máxima verosimilitud condicional  $\hat{\mu}$  y  $\hat{\tau}$  se obtienen igualando a cero las derivadas parciales de (7) con respecto a  $\mu$  y  $\tau$ :

$$\frac{\partial L(\mu, \tau)}{\partial \mu} = 0 \quad \text{y} \quad \frac{\partial L(\mu, \tau)}{\partial \tau} = 0.$$

Los errores estándar asintóticos de  $\hat{\mu}$  y  $\hat{\tau}$  pueden ser obtenidos a través de la inversa de la matriz de información<sup>7</sup> respectiva.

Solamente se utilizan las respuestas para las cuales  $x_{vi1} \neq x_{vi2}$ . Por lo tanto, la verosimilitud en (7) puede interpretarse como

$$L_c = \prod_{v=1}^N \prod_{i=1}^k p(X_{vi2} = 1 | x_{vi2} \neq x_{vi1}). \quad (8)$$

Además, bajo la suposición de que el tratamiento no tiene efecto ( $\mu = 0$  y luego  $\delta_v = \tau$ ), el estimador de máxima verosimilitud de  $\tau$  viene

$$\hat{\tau} = \text{logit} \left\{ \frac{\sum_{i=1}^k n_{01}^{(i)}}{\sum_{i=1}^k n_{01}^{(i)} + n_{10}^{(i)}} \right\}, \quad (9)$$

siendo

$n_{0i}^{(i)}$  = número de pacientes con respuesta cero al ítem  $i$  en  $T_1$  y respuesta positiva en  $T_2$ ;

<sup>7</sup> La **Matriz de Información** (*Information Matrix*) es dada por  $-E[H]$  donde  $H$  es la matriz hesiana (matriz de las segundas derivadas).

$n_{10}^{(i)}$  = número de pacientes con respuesta positiva al ítem  $i$  en  $T_1$  y respuesta negativa en  $T_2$ .

En el caso en que se considera solamente un ítem, se puede utilizar el Test de McNemar para testar  $H_0 : \tau = 0$ .

Se asume que el proceso temporal y el tratamiento poseen igual efecto sobre todos los ítems y el “parámetro cambio” es independiente de  $i$ . Si, por el contrario, el proceso temporal y el tratamiento tienen un efecto diferenciado<sup>8</sup> sobre los ítems, se modifica (3) para

$$\delta_u = \alpha_i(q_v\mu + \tau), \quad (10)$$

donde  $\alpha_i$  denota la sensibilidad del ítem  $i$  al cambio.

#### *Modelo de Zwinderman para más que dos instantes de medida*

Zwinderman extiende el modelo a la situación en que los datos de CVRS dicen respecto a más que dos instantes. Consideremos las respuestas de  $N$  pacientes a  $k$  ítems medidos en  $p$  instantes ( $p \geq 2$ ). Sea  $X_{vit}$  la respuesta del paciente  $v$  al ítem  $i$  en el instante  $t$ . Entonces, la probabilidad condicional de observar una respuesta positiva del paciente  $v$  al ítem  $i$  en el instante  $t$  viene modelizada por la curva logística

$$p(X_{vit} = 1 | \theta_{vit}) = \frac{\exp(\theta_{vit})}{1 + \exp(\theta_{vit})} \quad (11)$$

donde  $\theta_{vit}$  denota la posición del paciente  $v$  en la variable latente medida por el ítem  $i$ ,  $\theta_i$ , en el instante  $t$ . Sea  $\theta_{vi}$  la posición del paciente  $v$  en  $\theta_i$  en el primero instante, entonces

<sup>8</sup> Muchos tratamientos tienen efectos beneficiosos sobre un determinado ítem pero presentan efectos colaterales que modifican la respuesta a otros ítems. Por ejemplo, la administración de APD a pacientes con cáncer ejerce un efecto benéfico sobre la movilidad. Luego, la probabilidad de obtener puntuaciones elevadas (adverso) en los ítems relativos a la movilidad disminuye para aquellos pacientes que reciben APD. Sin embargo, los efectos colaterales gástrico-intestinales del APD aumentan la probabilidad de respuesta a los ítems relacionados con la toxicidad.

$$\theta_{vit} = \theta_{vi} + \delta_{vt} = \theta_{vi} + \sum_{a=1}^{t-1} \delta_v^{(a,a+1)} = \theta_{vi} + \delta_v^{(1,2)} + \dots + \delta_v^{(t-1,t)} \quad (12)$$

donde  $\delta_{vt}$  denota el cambio registrado en el paciente  $v$  hasta el instante  $t$  y  $\delta_v^{(a,a+1)}$  el cambio de  $\theta_{vi}$  entre  $a$  y  $a+1$ . De forma similar a la situación en que tenemos dos instantes,  $\delta_v^{(a,a+1)}$  viene modelizado por

$$\delta_v^{(a,a+1)} = \sum_{j=1}^m q_{vj}^{(a,a+1)} \mu_j^{(a,a+1)} + \tau^{(a,a+1)} \quad (13)$$

donde  $\mu_j^{(a,a+1)}$  es el cambio debido al tratamiento  $j$  entre  $a$  y  $a+1$ . El peso  $q_{vj}^{(a,a+1)}$  indica la presencia o ausencia de dosis del tratamiento  $j$  para el paciente  $v$ , y puede depender del intervalo  $(a,a+1)$ .

Como se ha referido, la suma de las puntuaciones  $X_{v\bullet} = \sum_{t=1}^p X_{vit}$  es un estadístico suficiente para  $\theta_{vi}$ . Luego, los parámetros  $\mu$  y  $\tau$  pueden ser estimados de forma consistente maximizando la verosimilitud condicional de las respuestas a los ítems dados los estadísticos suficientes para  $\theta_u$ :

$$L_c = \prod_{v=1}^N \prod_{i=1}^k \prod_{t=1}^p p(X_{vit} = x_{vit} | X_{v\bullet} = x_{v\bullet}) \propto \prod_{v=1}^N \frac{\prod_{t=1}^p \varepsilon_{vt}^{x_{vt}}}{\prod_{i=1}^k \Gamma_{x_{vi\bullet}}(\varepsilon_{vt})} \quad (14)$$

donde  $x_{v\bullet} = \sum_{i=1}^k X_{vit}$ ,  $\varepsilon_{vt} = \exp(\delta_{vt})$  y  $\Gamma_{x_{vi\bullet}}(\varepsilon_{vt})$  es la función simétrica básica de orden  $x_{vi\bullet}$  de los parámetros  $\varepsilon_{vt}$  ( $t=1, \dots, p$ ). Las funciones simétricas básicas son definidas como:

$$\begin{aligned} \Gamma_0 &= 1 \\ \Gamma_1 &= \varepsilon_{v1} + \varepsilon_{v2} + \dots + \varepsilon_{vp} \\ \Gamma_2 &= \varepsilon_{v1}\varepsilon_{v2} + \varepsilon_{v1}\varepsilon_{v3} + \dots + \varepsilon_{vp-1,p} \\ &\dots \end{aligned}$$

$$\Gamma_p = \varepsilon_{v1}\varepsilon_{v2} + \dots + \varepsilon_{vp-1}\varepsilon_{vp}.$$

La verosimilitud condicional (14) es una generalización de la verosimilitud condicional del modelo de Rash.

El modelo general tal como se ha especificado en (11), (12) y (13) puede ser modificado de varias formas. Nótese que, igual se ha establecido anteriormente, se asume que el cambio es igual para todas las variables latentes en el modelo. El parámetro cambio  $\delta_u$  no depende de los ítems. Sin embargo, el modelo puede ser modificado de forma que permita cambios diferenciados con respecto a los ítems, introduciendo parámetros de sensibilidad en (13) de forma análoga que en (10). Por otra parte, el modelo especificado en (12) permite diferentes contrastes entre instantes. De forma alternativa, se puede asumir que el efecto del tratamiento es constante a lo largo del tiempo. Si los instantes son fijos e iguales para todos los pacientes y los pesos  $q_{vj}$  no cambian a lo largo del tiempo, el parámetro cambio  $\delta_u$  puede modelizarse como

$$\delta_{ut} = (t-1)\delta_v \quad \delta_v = q_v\mu + \tau. \quad (15)$$

El modelo de cambio especificado en (12) y (13) permite relaciones curvilíneas a lo largo del tiempo, mientras el modelo restringido es lineal.

Al Modelo de Zwinderman (1990) están subyacentes las siguientes **suposiciones:**

- **Independencia condicional** – se exige que las respuestas a los ítems sean independientes dado un nivel fijo de  $\theta_i$ :

$$p(X_{vit}, X_{vij} | \theta_{vi}) = p(X_{vit} | \theta_{vi})p(X_{vij} | \theta_{vi})$$

Esto significa que las probabilidades de respuesta no dependen de las respuestas a los otros ítems.

- **Suficiencia de la suma de las puntuaciones** – la suma de las puntuaciones a lo largo del tiempo es un estadístico suficiente mínimo para los valores individuales iniciales  $\theta_{vi1}$ . La suficiencia exige la inexistencia de interacción paciente-tiempo dentro de los grupos de tratamiento. Esto implica que, dentro de un determinado grupo de tratamiento, el cambio a lo largo del tiempo es el mismo para todos los pacientes.

Estas dos suposiciones se pueden testar recurriendo una razón de verosimilitudes:

$$X^2 = -2 \log \left[ \frac{L_{tot}(\mu, \tau | X)}{\prod_{g=1}^G L_g(\mu, \tau | X)} \right],$$

donde  $L_{tot}$  es el máximo de la verosimilitud condicional (14) calculado en la muestra completa y  $L_g$  es el máximo de la verosimilitud condicional para el  $g$ -ésimo subgrupo disjunto de  $n_g$  pacientes. La distribución asintótica de  $X^2$  es una ji-cuadrado con  $(G-1)(m+1)$  grados de libertad si  $n_g$  es grande en todos los  $G$  subgrupos. Como los parámetros  $\mu$  y  $\tau$  son estimados independientemente de la muestra de sujetos, las mismas estimaciones son validas en los  $G$  subgrupos. Luego se espera que el producto de las  $G$  verosimilitudes de los  $G$  subgrupos disjuntos se aproxime bastante de la verosimilitud global.

Los estadísticos de razón de verosimilitudes también pueden ser utilizados para testar hipótesis estructurales en el efecto de los tratamientos:  $H_0 : \mu_j = 0$ .

Para testar  $H_0$  se puede también utilizar  $z = \frac{\hat{\mu}}{SE(\hat{\mu})}$ , que tiene distribución

asintótica Normal estándar. Además, los estadísticos de razón de verosimilitudes pueden ser utilizados para contrastar el modelo de cambio (10) contra el (4).



El modelo presentado por Zwinderman en (11), (12) y (13) posibilita el análisis de observaciones de CVRS dicotómicas repetidas. El modelo puede generalizarse para medidas nominales, aunque, en este caso, la verosimilitud condicional (14) se vuelve difícil de calcular. Sin embargo, si se procede a la ordenación de las categorías, no se pierde mucha información cuando las observaciones son dicotómicas.

---

**ANEXO V – CUESTIONARIO  
SOCIODEMOGRÁFICO**

---



40308

Preencha o questionário de acordo com o exemplo, utilizando **caneta azul ou preta**.

Preencha

assim   
 assim não

Código Data 

## QUESTIONÁRIO SÓCIO-DEMOGRÁFICO

1. **Sexo:**  Masculino  Feminino

2. **Idade:**

3. **Estado Civil:**  Solteiro  Casado  União de Facto  Divorciado  Viúvo

4. **Tem Filhos?**  Não  Sim **Quantos?**  1  2  3  4  5 ou mais

5. **Actividade Actual:**  Tempo Completo  Tempo Parcial  Reformado  Desempregado

6. **Nível de Estudos:**  Sem Estudos  Primários  Secundários  Universitários

7. **Diagnósticos psiquiátricos anteriores:**  Não  Sim

Se respondeu Sim, refira quais: \_\_\_\_\_

8. **Diagnóstico de doença oncológica:** \_\_\_\_\_

9. **Data (ano) do diagnóstico:** \_\_\_\_\_

10. **Antes de lhe ter sido efectuado o diagnóstico, tinha contactado com alguém com cancro?**

Não  Sim Se Sim, quem? \_\_\_\_\_

11. **Está a efectuar algum tratamento?**  Não  Sim

Se respondeu Sim, refira qual: \_\_\_\_\_

12. **Realizou algum tratamento no passado?**  Não  Sim

Se respondeu Sim, refira qual e há quanto tempo terminou: \_\_\_\_\_

13. **Teve acompanhamento psicológico ao longo do processo de doença?**  Não  Sim

**Nota:** Esta secção será preenchida com ajuda do profissional de saúde de forma a incluir qualquer variável de interesse que não tenha sido considerada.

14. **Peso nos últimos 6 meses**

Estável  Perdeu menos de 5kg  Perdeu entre 5kg e 10 kg  Perdeu mais de 10kg

15. **Estadiamento**  Local  Loco - regional  Metástases

16. **Tipo de Cirurgia:** \_\_\_\_\_

17. **Terapêutica**

Cirurgia  Radioterapia  Quimioterapia  Radio+Quimioterapia  Hormonoterapia

Obs: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

---

**ANEXO VI – CONSENTIMIENTO  
INFORMADO**

---

## Carta de Consentimento Informado

A investigação na qual irá participar pretende dar um contributo no estudo da Qualidade de Vida em diagnósticos oncológicos, no âmbito da tese de doutoramento em Estatística Aplicada pela Universidade de Salamanca da doutoranda Sara Nunes sob orientação da Professora Doutora Maria Purificación Galindo. Em particular, pretendem-se avaliar alterações na autopercepção de pacientes em tratamento.

A sua participação consiste na resposta a dois questionários sobre Qualidade de Vida (*Medical Outcomes Study - SF-36 e Rotterdam Symptom Checklist - RSCL*) em vários momentos distintos e um questionário sócio-demográfico. Todos os dados serão tratados de forma confidencial. A qualquer momento pode decidir abandonar o estudo, sem qualquer tipo de consequência.

Se concordar em participar por favor assine no espaço abaixo. O nosso agradecimento por aceitar dar a sua importante contribuição para este estudo.

Eu, \_\_\_\_\_, tomei conhecimento do objectivo do estudo e do que tenho de fazer para participar no mesmo. Fui esclarecido(a) sobre todos os aspectos que considero importantes e as perguntas que coloquei foram respondidas. Fui informado(a) que tenho direito a recusar participar e que a minha recusa em fazê-lo não trará consequências para mim. Assim, declaro que aceito participar na investigação.

Data \_\_\_\_ / \_\_\_\_ / \_\_\_\_

O Participante \_\_\_\_\_

