

UNIVERSIDAD AUTÓNOMA DE MADRID
DEPARTAMENTO DE MEDICINA
FACULTAD DE MEDICINA



TESIS DOCTORAL

DESARROLLO DE UN ÍNDICE DE GRAVEDAD

EN LA ARTRITIS REUMATOIDE

M Esther Toledano Martínez

Madrid, 2017



Universidad Autónoma De Madrid

Departamento de Medicina

Facultad de Medicina

DESARROLLO DE UN ÍNDICE

DE GRAVEDAD EN LA ARTRITIS REUMATOIDE

Tesis Doctoral presentada por la licenciada **D^a M. Esther Toledano Martínez**, para optar al grado de Doctora en Medicina

DIRECTORES:

ISIDORO GONZÁLEZ ÁLVARO

Jefe de Sección de Reumatología

Instituto de Investigación Sanitaria Hospital Universitario de La Princesa

MARÍA JESÚS GARCÍA DE YÉBENES Y PROUST

Instituto de Salud Músculo-esquelética (INMUSC)

TUTOR:

JESÚS-ALBERTO GARCÍA-VADILLO

Jefe de Sección de Reumatología

Instituto de Investigación Sanitaria Hospital Universitario de La Princesa

*¿Eres tú, Guadarrama, viejo amigo,
la sierra gris y blanca,
la sierra de mis tardes madrileñas
que yo veía en el azul pintada?
Por tus barrancos hondos
y por tus cumbres agrias,
mil Guadarramas y mil sones vienen
cabalgando conmigo, a tus entrañas*

Camino de Valsaín / Campos de Castilla

(Antonio Machado)

A mi familia, mi luz y mi refugio

A Carlos

AGRADECIMIENTOS

Me gustaría agradecer a los reumatólogos del Hospital Universitario de La Princesa por acompañarme en mis primeros pasos en el campo de la Reumatología, durante los años de la residencia. Su visión de esta maravillosa especialidad sigue siendo mi guía en el quehacer diario.

En especial, quiero expresar mi más sincero agradecimiento al Dr. Alberto García Vadillo, mi tutor, ahora y siempre, por su acertada orientación y perenne disponibilidad. Con él me inicié en la investigación y ha sido inestimable su colaboración en este proyecto.

Les agradezco a mis compañeros del Servicio de Reumatología Hospital Clínico San Carlos su acogida cuando llegué. El día a día con ellos es una fuente de enriquecimiento personal y profesional. En especial, al Dr. Juan Ángel Jover por brindarme la oportunidad de participar en este proyecto de colaboración.

Quiero manifestar mi agradecimiento al grupo investigador con el que he tenido el placer de trabajar para realizar este proyecto. A la Dra. Loreto Carmona, que lo visualizó como una realidad; sus brillantes ideas han permitido alcanzar nuestro objetivo. A la Dra. María Jesús García de Yébenes, por su trabajo incansable y orientación metodológica. Al Dr. Isidoro González-Álvaro, por transmitirme su vocación investigadora. A todos ellos por su orientación, ayuda y estimulación constante en todos los aspectos de la tesis durante estos años. Agradecerles la confianza que me han demostrado, la dedicación y la atención que en todo momento me han ofrecido.

Gracias a la generosa aportación de todos los reumatólogos, que desinteresadamente, han participado en el panel de expertos y la encuesta de este proyecto.

Finalmente, me gustaría dar las gracias a todos los pacientes, la razón última de mi vocación y de este trabajo.

Índice General

ABREVIATURAS	12
RESUMEN	14
PARTE I. INTRODUCCIÓN	17
Introducción general	17
1. Conceptos actuales de “tight control” y “treat-to-target”	17
2. Necesidad de instrumentos de medición	19
3. Validación de biomarcadores de la AR	20
Justificación del estudio	21
Objetivos de la tesis	22
Material y métodos.....	22
1. Estudio retrospectivo registro PEARL	22
2. Identificación de instrumentos existentes.	22
3. Selección de dominios e ítems que reflejan la gravedad: Panel de expertos y estudio Delphi.....	23
4. Desarrollo del índice.	24
PARTE II. ESTUDIOS	27
Sub-estudio 1. Incremento del tratamiento con FAME a lo largo del tiempo en pacientes con artritis de reciente comienzo del estudio PEARL	27
1. Introducción.....	27
2. Objetivo.....	28

3.	Material y métodos.....	28
3.1	Pacientes	28
3.2	VARIABLES ANALIZADAS.....	29
3.3	Análisis estadístico.....	30
4.	Resultados.....	31
4.1	Descripción de las 3 subpoblaciones de pacientes con artritis de inicio.....	31
4.2	Intensidad del tratamiento con FAME.....	32
4.3	Resultados después de dos años de seguimiento	34
5.	Conclusión.....	39
	Sub-estudio 2. Identificación de instrumentos existentes	41
1.	Objetivos	41
2.	Material y métodos.....	41
2.1	Estrategia de búsqueda.....	41
2.2	Selección de estudios.....	42
2.3	Evaluación de la calidad de los índices	43
3.	Resultados.....	43
3.1	Identificación de instrumentos para evaluación de gravedad	43
3.2	PAS	45
3.3	Rheumatoid arthritis records-based index of severity (RARBIS)	48
3.4	Claims-based index for rheumatoid arthritis severity (CIRAS)	48
4.	Evaluación crítica de los índices.....	52

Sub-estudio 3: Selección de dominios y variables que reflejan la gravedad: panel de expertos y estudio Delphi.....	56
1. Objetivo.....	56
2. Material y métodos.....	56
3. Resultados.....	58
4. Conclusiones	65
Sub-estudio 4. Diseño y construcción de un índice compuesto de evaluación de la gravedad.....	67
1. Introducción.....	67
2. Descriptivo de la visita basal.....	68
3. Definiciones y construcción del índice.....	71
4. Análisis de los ítems.....	73
4.1 Población y datos	73
4.2 Comportamiento de los ítems	73
4.3 Estructura y dimensionalidad: Análisis factorial.....	73
4.4 Fiabilidad: consistencia interna	75
4.5 Capacidad de discriminación	76
5. Resultados.....	78
5.1 Análisis de los ítems	78
5.2 Estructura y dimensionalidad.	79
5.2.1 Análisis factorial A: índice completo	79
5.2.1.1 Número de factores a extraer	79
5.2.1.2 Rotación e interpretación de los factores.....	80
5.2.1.3 Correlaciones ítem-total y consistencia interna.....	81
5.2.1.4 Optimización de la dimensionalidad	82

5.2.2	Análisis factorial B: eliminación del ítem 5 (reactantes fase aguda)	82
5.2.3	Análisis factorial C: eliminación del ítem 14 (corticodependencia).....	84
5.2.4	Análisis factorial D: eliminación del ítem 12 (cirugía ortopédica)	87
5.2.5	Análisis factorial E: eliminación del ítem 13 (ingresos).....	89
5.2.6	Análisis factorial F: eliminación del ítem 7 (imposibilidad remisión)	91
5.2.7	Análisis factorial G: eliminación ítem 2 (grandes articulaciones)	92
5.2.8	Análisis factorial H: eliminación del ítem 1 (poliartritis).....	94
5.2.9	Análisis factorial I: eliminación del ítem 11 (ausencia de mejoría HAQ) ..	95
6.	Capacidad de discriminación	98
7.	Conclusión.....	100
PARTE III. DISCUSIÓN GENERAL		104
Resumen de resultados. Fortalezas y limitaciones		104
Diseño y construcción de un índice de gravedad.....		107
Perspectivas futuras.....		110
CONCLUSIONES		112
BIBLIOGRAFÍA		114

Índice de Figuras y Tablas

Figura 1. Descripción del tratamiento recibido por los pacientes de la cohorte PEARL en los periodos de tiempo 2000-2004, 2005-2009 y 2010-2014.....	33
Figura 2. Actividad de la enfermedad en los pacientes de la cohorte PEARL en los 3 periodos estudiados	34
Figura 3. Progresión radiológica en manos en los pacientes de la cohorte PEARL en los 3 periodos estudiados	37
Figura 4. Estrategia de búsqueda.	42
Figura 5. Diagrama de flujo de artículos.....	44
Figura 6. Solución factorial definitiva: factorización.	98
Tabla I. Características de las poblaciones del registro PEARL en los tres periodos de tiempo analizado.	31
Tabla II. Variables asociadas con actividad de la enfermedad al finalizar el seguimiento	35
Tabla III. Variables asociadas con los componentes de la progresión radiológica (erosión y estrechamiento del espacio articular) entre la visita basal y los dos años de seguimiento en manos mediante el método de Sharp con la modificación de van der Heijde.....	38
Tabla IV. Estudios excluidos tras selección.....	44
Tabla V. Características de los pacientes de las cohortes incluidas en la revisión.	46
Tabla VI. Características de los estudios incluidos en la revisión.....	47
Tabla VII. Descripción de las variables, dominios y puntuación de índices RARBIS, CIRAS y PAS.	50
Tabla VIII. Índice RARBIS. Descripción características relevantes.	52

Tabla IX. Índice CIRAS. Descripción características relevantes.	53
Tabla X. Composite score for RA severity index. Descripción características relevantes	53
Tabla XI. Calidad metodológica de la validación de los índices incluidos.	54
Tabla XII. Variables constructo gravedad 1ª ronda Delphi.....	59
Tabla XIII .Prioridad de las variables elegidas por el panel de expertos.	61
Tabla XIV. Acuerdo en la definición de las variables elegidas por el panel de expertos.	62
Tabla XV. Factibilidad de las variables elegidas por el panel de expertos.	63
Tabla.XVI. Definiciones y variables elegidas por el panel de expertos y posterior Delphi	64
Tabla XVII. Características clínicas basales: Variables cualitativas.....	69
Tabla XVIII. Características clínicas basales: Variables cuantitativas.	70
Tabla XIX. Características clínicas basales: Tratamientos.	71
Tabla XX. Definiciones e ítems.....	72
Tabla XXI. Distribución de las puntuaciones y patrón de valores perdidos.	79
Tabla XXII. Factorial A: factores, autovalores y varianza explicada.	80
Tabla XXIII. Factorial A: cargas factoriales tras rotación varimax.	81
Tabla XXIV. Factorial A: correlación ítem-total y Consistencia interna.....	82
Tabla XXV. Factorial B: factores, autovalores y varianza explicada.	83
Tabla XXVI. Factorial B: cargas factoriales tras rotación varimax.	84
Tabla XXVII. Correlación ítem-total y consistencia interna.	84
Tabla XXVIII. Factorial C: factores, autovalores y varianza explicada.	85
Tabla XXIX. Factorial C: cargas factoriales tras rotación varimax.....	86

Tabla XXX. Factorial C: Correlación ítem total y consistencia interna.	86
Tabla XXXI. Factorial D: factores, autovalores y varianza explicada.	87
Tabla XXXII. Factorial D: cargas factoriales tras rotación varimax.	88
Tabla XXXIII.-Factorial D: cargas factoriales tras rotación varimax.	88
Tabla XXXIV. Factorial E: factores, autovalores y varianza explicada.	89
Tabla XXXV. Factorial E: cargas factoriales tras rotación varimax.....	90
Tabla XXXVI. Factorial E: consistencia interna y correlación ítem-total.....	90
Tabla XXXVII. Factorial F: factores, autovalores y varianza explicada.....	91
Tabla XXXVIII. Factorial F: cargas factoriales tras rotación varimax.....	91
Tabla XXXIX.- Factorial F: Consistencia interna y correlación ítem-total.	92
Tabla XL. Factorial G: factores, autovalores y varianza explicada.....	92
Tabla XLI. Factorial G: cargas factoriales tras rotación varimax.....	93
Tabla XLII. Factorial G: Consistencia interna y correlación ítem-total.	93
Tabla XLIII. Factorial H: factores, autovalores y varianza explicada.....	94
Tabla XLIV. Factorial H: cargas factoriales tras rotación varimax.	94
Tabla XLV. Factorial H: Consistencia interna y correlación ítem-total.....	95
Tabla XLVI. Factorial I: factores, autovalores y varianza explicada.	96
Tabla XLVII. Factorial I: cargas factoriales tras rotación varimax.....	96
Tabla XLVIII. Factorial I: Consistencia interna y correlación ítem-total.....	96
Tabla XLIX. Análisis factorial: Resumen.	97
Tabla L. Análisis discriminante: clasificación cruzada.	99
Tabla LI. Regresión lineal múltiple.....	99

ABREVIATURAS

- ACPA:** anticuerpos anti-péptido citrulinado.
- ACR:** American College of Rheumatology.
- Anti-TNF:** anti-Tumoral Necrosis Factor.
- AR:** artritis reumatoide.
- AZA:** azatioprina.
- AUC:** area under the curve.
- CARC:** consulta de artritis de reciente comienzo.
- CDAI:** clinical disease activity index.
- C1-C2:** columna cervical niveles 1 y 2.
- CIRAS:** Claims-based index for rheumatoid arthritis severity.
- CsA:** ciclosporina.
- DAS:** disease activity score.
- DAS28:** disease activity score based on a 28-joint count.
- DE:** desviación estándar.
- FAME:** fármaco modificador de la enfermedad.
- FR:** factor reumatoide.
- HAQ:** Health Assessment Questionnaire.
- HQ:** hidroxicloroquina.
- HUPI:** Hospital Universitario La Princesa Index.
- IMC:** índice masa corporal.
- LF:** leflunomida.

MTX: metotrexato.

N: número.

NI: no incluido en el modelo final por no ser significativo.

NOC: neumonía organizativa criptogénica.

p50: percentil 50 o mediana.

PAS: Patient Activity Scale.

PCR: proteína C reactiva.

OR: odds ratio.

RARBIS: Rheumatoid arthritis records-based index of severity.

Ref: referencia.

RIC: rango intercuartilico.

SDAI: simple disease activity index.

SHS: método de Sharp con la modificación de van der Heijde.

SSZ: sulfasalazina.

TB: terapia biológica.

TC: estrategia "tight control".

T2T: estrategia "treat-to-target".

TFME: número de días de tratamiento con FAME, ponderado por tipo de fármaco.

VSG: velocidad de sedimentación globular.

VGEM: valoración global de la enfermedad por el médico.

RESUMEN

El objetivo de la presente investigación es desarrollar un índice que permita estimar la gravedad de la artritis reumatoide a los dos años de seguimiento, refleje adecuadamente el marco teórico de este constructo y pueda obtenerse en la práctica habitual a partir de los datos existentes en la historia clínica. En la primera parte se revisan los conceptos actuales de “tigh control” y “treat-to-target”, estrategias terapéuticas intensivas ampliamente validadas como herramientas que mejoran la evolución de la enfermedad. También se revisa la mejor respuesta terapéutica a cualquier fármaco modificador de la enfermedad (FAME) en estadios tempranos, concepto conocido como “ventana de oportunidad”. Estas estrategias han tenido, por su gran eficacia, un efecto inesperado en la investigación de biomarcadores. Puesto que en los últimos años los pacientes tienden a ser tratados de forma más intensiva, independientemente de su probable gravedad, las variables de desenlace clásicas no muestran diferencias entre pacientes con peor o mejor pronóstico de inicio. El trabajo de investigación presentado se desarrolla en estudios consecutivos. En el primer sub-estudio se analizará si el tratamiento con FAME y terapia biológica (TB) ha cambiado a lo largo del tiempo en el registro de artritis de inicio PEARL (Princesa Early Arthritis Register Longitudinal). Ante la necesidad de disponer de una variable de gravedad válida que permita evaluar la utilidad de nuevos biomarcadores en artritis de reciente comienzo, en el segundo sub-estudio se evaluarán las herramientas existentes, mediante una revisión sistemática de estudios que desarrollen o validen variables de gravedad en adultos con artritis reumatoide (AR). En el tercer sub-estudio, se convocará un panel de expertos/Delphi para definir los dominios relevantes del constructo “gravedad”, una vez diseñado un mapa conceptual. Una vez elegidas las variables, se realizará el desarrollo del índice de gravedad mediante un análisis

factorial exploratorio. Finalmente, se realiza una detallada discusión de los resultados. Las conclusiones finales ilustran la dificultad para encontrar un criterio externo que refleje el constructo “gravedad” y las perspectivas futuras del índice desarrollado en la investigación de biomarcadores.

PARTE I. INTRODUCCIÓN

Introducción general	17
1. Conceptos actuales de “tight control” y “treat-to-target”	17
2. Necesidad de instrumentos de medición	19
3. Validación de biomarcadores de la AR	20
Justificación del estudio	21
Objetivos de la tesis	22
Material y métodos	22
1. Estudio retrospectivo registro PEARL	22
2. Identificación de instrumentos existentes.	22
3. Selección de dominios e ítems que reflejan la gravedad: Panel de expertos y estudio Delphi.....	23
4. Desarrollo del índice.	24

PARTE I. INTRODUCCIÓN

INTRODUCCIÓN GENERAL

1. Conceptos actuales de “tight control” y “treat-to-target”

La artritis reumatoide (AR) es una enfermedad autoinmune caracterizada por dolor, rigidez matutina, tumefacción, deformidad articular y discapacidad funcional(1). Así mismo, los pacientes con AR presentan un incremento del riesgo de morbi-mortalidad fundamentalmente debida a infecciones, eventos cardiovasculares y cáncer(2, 3).

Las recomendaciones actuales para el manejo de la AR apoyan la puesta en funcionamiento de estrategias terapéuticas intensivas, ya que la evidencia científica demuestra que mejoran la evolución de la enfermedad(4),(5, 6). Estas estrategias intensivas se basan en dos pilares clave: a) la evaluación objetiva del proceso inflamatorio de forma estrecha (“tight control”; TC) que puede mejorar el curso de la enfermedad, al adaptar las decisiones terapéuticas a la evolución concreta del paciente de una forma más ecuánime(7-9); y b) el establecimiento de un objetivo terapéutico en base a índices compuestos de actividad con el ajuste del tratamiento para lograrlo, lo que se conoce como “treat-to-target” (t2t)(7, 9).

Por otra parte, el inicio precoz de tratamiento con FAME también constituye un evento clave, al haberse confirmado la hipótesis del concepto de “ventana de oportunidad”, que refleja la mejor respuesta al tratamiento en los estadios tempranos de la enfermedad, a diferencia de fases más avanzadas(10-12). De esta forma, el inicio precoz de un FAME puede enlentecer o evitar la progresión y modificar, subsidiariamente, el exceso de riesgo asociado.

La implantación de las Consultas de Artritis de Reciente Comienzo (CARC) responde, precisamente, a la intención de utilizar de forma temprana estrategias terapéuticas dirigidas a mejorar los resultados clínicos y radiológicos(13-16). Sin embargo, el tratamiento intensivo con FAME puede suponer también un aumento de la toxicidad(17). El reconocimiento, al inicio de la enfermedad, de las características relacionadas con un curso evolutivo desfavorable puede ayudar a mejorar la relación riesgo/beneficio y ajustar el abordaje terapéutico precoz e intensivo en los casos con peor pronóstico(18). Este es otro campo en el que las CARC han demostrado ser de utilidad, siendo la consecución de remisión o la evaluación de la progresión radiológica las variables de desenlace más frecuentemente empleadas para identificar factores de mal pronóstico.

Sin embargo, la puesta en marcha de las estrategias tight control y treat-to-target a lo largo de los últimos 10 años ha tenido, precisamente por su gran eficacia, una consecuencia inesperada en la investigación de biomarcadores. Puesto que los pacientes con peor pronóstico inicial son tratados de forma más intensiva, las variables de desenlace clásicas, como la consecución de remisión o la progresión radiológica a largo plazo, no muestran diferencias entre los pacientes con ACPA positivo y negativo(19), siendo los ACPA el biomarcador mejor validado como asociado a peor pronóstico en la AR. Para obviar este problema y poder explorar la utilidad de biomarcadores de mal pronóstico en pacientes con artritis de inicio, se han propuesto diferentes variables subrogadas de gravedad relacionadas con el uso de tratamiento más intenso, como la necesidad del uso de TB(20) o la cantidad acumulada de tratamiento(19, 21) para explorar factores pronósticos en pacientes con artritis de inicio, aunque la validez y propiedades métricas de estas variables está por dilucidar. No obstante, y a pesar de las limitaciones existentes, la esperanza de poder revertir el curso evolutivo de la enfermedad en sus fases precoces, ha promovido el uso de estrategias terapéuticas más intensas incluso en pacientes con artritis indiferenciada(22, 23). Este mensaje ha calado hondo entre los reumatólogos, lo que

ha provocado que, al menos en el entorno de las consultas de artritis de inicio, las diferencias entre la intensidad de tratamiento prescrita a los pacientes con ACPA positivo y negativo se hayan diluido durante los últimos años.

2. Necesidad de instrumentos de medición

Para establecer el pronóstico de la enfermedad y valorar su evolución se necesitan instrumentos de medida consensuados y validados. Los instrumentos disponibles e incorporados a la práctica clínica, hasta el momento, son índices de medición de la actividad. Inicialmente se desarrollaron el DAS([24](#), [25](#)), basado en opinión de expertos, el DAS en 28 articulaciones (DAS28)([26](#), [27](#)), y el DAS basado en la PCR (DAS28-PCR) [<http://www.das-score.nl>]([28](#)). Posteriormente han surgido otros índices de actividad que intentan mejorar las limitaciones de los anteriores: necesidad de calculadora, utilización de la VSG como reactante de fase (más inespecífico y menos sensible al cambio que la PCR)([29](#)), y sobreestimación de la actividad en mujeres debido a la presentación de recuentos más elevados de articulaciones dolorosas([30-32](#)). Los índices SDAI y CDAI, desarrollados por Smolen([33](#)), pretenden salvar estas limitaciones del DAS. Sin embargo, la simplicidad de ambos instrumentos, basados en la suma aritmética de variables clínicas (número articulaciones dolorosas -NAD- y número articulaciones tumefactas -NAT-), valoración global de la enfermedad por el médico y el paciente (VGEM y VGEP), y un reactante de fase aguda (PCR) sin inclusión en su fórmula de ponderaciones para los diferentes dominios del índice, reduce su utilidad y dificulta su interpretación en algunas situaciones clínicas([34](#), [35](#)). Nuestro grupo contribuyó al esfuerzo por una mejor evaluación de la actividad de la AR con el índice HUPI que evita el sesgo por género, pondera las variables y puede ser calculado con VSG, PCR o ambos reactantes de fase aguda([36](#)).

A pesar de las estrategias actuales de control estricto, del conocimiento de factores pronósticos, y del generoso desarrollo de índices en reumatología la realidad es que, por el momento, no existe un instrumento de medida de la gravedad de la AR que nos

permita identificar, con precisión, a los pacientes que van a necesitar estrategias terapéuticas más intensivas. Algunas variables aisladas, como la presencia de enfermedad erosiva o el uso de TB en los dos primeros años reflejan, en cierta medida, la gravedad de la AR y han sido utilizadas en algunos estudios como variables subrogadas de gravedad con la intención de detectar factores predictivos del pronóstico(20, 37, 38). Sin embargo, no se dispone de ningún índice compuesto que incluya otros aspectos diferentes de la gravedad de la enfermedad, como puedan ser la discapacidad o la aparición de manifestaciones extraarticulares.

3. Validación de biomarcadores de la AR

La evidencia disponible en el momento actual sobre la asociación entre biomarcadores y peor evolución de la AR, es bastante limitada a pesar de las esperanzas puestas en este campo y del enorme esfuerzo realizado por numerosos grupos de investigación. La búsqueda de biomarcadores pronósticos sigue siendo un reto en la actualidad. En primer lugar, la historia natural de la enfermedad ha cambiado como consecuencia de un tratamiento más precoz y eficaz(14, 39). En segundo lugar, por el sesgo de indicación que conlleva la determinación de ACPA. Nuestro grupo ha descrito que, actualmente, los pacientes con ACPA ya no presentan peor evolución desde el punto de vista de la actividad de la enfermedad, ni de la progresión radiológica, ya que reciben un tratamiento más intenso que los que no presentan ACPA(19). Hemos demostrado que para poner de manifiesto posibles biomarcadores pronósticos es necesario tener en cuenta estos sesgos(21, 40, 41). Por ello, hemos propuesto que se utilice la intensidad de tratamiento con FAME acumulado en los dos primeros años de seguimiento como una variable subrogada de gravedad(19). No obstante, esta variable que puede ser muy útil en un entorno homogéneo como pueden ser las CARC, puede presentar limitaciones importantes debidas, por ejemplo, a la variabilidad existente entre los reumatólogos en el manejo de FAME y que ha sido constatada en el estudio emARII(42). Por ello, para la búsqueda eficaz de biomarcadores pronósticos, sería

deseable disponer de una variable que reflejase el concepto de gravedad de la AR y que tuviese en cuenta todos los dominios relacionados con el desenlace de la enfermedad, como son la discapacidad, la progresión radiológica, la actividad inflamatoria y el tratamiento recibido entre otros. El ideal sería que esta variable se pudiese obtener tras un seguimiento de entre dos y cinco años. Esta limitación temporal evita el efecto de regresión a la media de seguimientos cortos y la dificultad de obtener cohortes de gran tamaño con seguimiento prolongado. Los resultados de un estudio preliminar no publicado, realizado por nuestro grupo en colaboración con el Hospital Clínico San Carlos, basado en la valoración global de estos factores por médicos implicados en la atención de pacientes con AR, han demostrado una importante variabilidad individual en la valoración de la gravedad. Por ello sería deseable utilizar una aproximación sistemática y metodológicamente correcta en el desarrollo de una variable de este tipo. Por otra parte, la utilidad de un biomarcador como medida de la gravedad, especialmente al inicio de la enfermedad, debe demostrarse en un estudio de validación. Una dificultad importante de este tipo de estudios es la necesidad de disponer de un criterio externo, llamado patrón oro, que refleje adecuadamente el concepto que pretendemos medir, en este caso la gravedad al inicio de la enfermedad, y que se toma como referencia para determinar la validez de la prueba o biomarcador en estudio. La realidad es que, por el momento, no se dispone de este patrón oro.

JUSTIFICACIÓN DEL ESTUDIO

El desarrollo de un índice de gravedad de la enfermedad, y la demostración de su validez y fiabilidad, permitirá afrontar los requisitos metodológicos de los estudios de utilidad de biomarcadores. Además, y en el caso concreto de nuestro grupo, la disponibilidad de este índice validado nos permitiría investigar otros posibles biomarcadores pronósticos en AR.

OBJETIVOS DE LA TESIS

El objetivo general es desarrollar un índice para estimar la gravedad de la AR a los 2 años de seguimiento, que refleje adecuadamente el marco teórico de este constructo y que pueda obtenerse en la práctica clínica habitual a partir de los datos existentes en la historia clínica. Para ello nos planteamos los siguientes objetivos operativos:

1. Identificar y evaluar los índices o instrumentos de medición desarrollados previamente con objetivos similares, mediante revisión sistemática de la literatura.
2. Diseñar y construir un índice compuesto de evaluación de la gravedad.

Estos objetivos se presentan como sub-estudios consecutivos.

MATERIAL Y MÉTODOS

1. Estudio retrospectivo registro PEARL

Con el fin de analizar si el tratamiento con FAME y TB ha cambiado a lo largo del tiempo, se realizó un análisis retrospectivo de los datos recogidos en el registro de artritis de inicio PEARL (Princesa Early Arthritis Register Longitudinal), dividiendo a los pacientes en tres subpoblaciones en función del año de inicio de los síntomas.

2. Identificación de instrumentos existentes.

Con el fin de valorar la existencia de herramientas de medición del concepto de gravedad y evaluar su utilidad para el estudio de biomarcadores se realizó una revisión sistemática de la literatura sobre estudios de desarrollo o validación de variables de gravedad en adultos con AR. Los índices encontrados se analizaron mediante la lista de verificación COSMIN (*CO*nsensus-based *S*tandards for the selection of health

Measurement Instruments(43). Finalmente, cada una de las variables se evaluó en función de las cualidades consideradas de especial relevancia para el grupo investigador, como son: claridad de la definición, viabilidad, presencia en los dos primeros años de evolución y dominio al que pertenecen.

3. Selección de dominios e ítems que reflejan la gravedad: Panel de expertos y estudio Delphi.

Posteriormente, y para poder seleccionar los dominios relevantes para la construcción de un índice de gravedad en la AR, se convocó a un panel de diez expertos clínicos con experiencia demostrada en artritis de inicio. Antes de celebrar las reuniones con los expertos, los investigadores diseñaron un mapa conceptual del constructo gravedad, en base a lo publicado y a las experiencias de los investigadores principales, con el fin de dirigir adecuadamente el grupo y poder extraer la información relevante. Se solicitó que definieran las dimensiones del constructo “gravedad”, entendido como “enfermedad temprana grave”. Se insistió en la premisa de que fuera un constructo discriminativo que identificara, de forma retrospectiva, a pacientes con diferente nivel de gravedad, y que estuviera formado por variables que pudieran ser medidas en los dos primeros años de evolución de la enfermedad.

Tras elaborar la lista de dominios y variables constituyentes, se realizó una encuesta Delphi para recoger las opiniones de otros reumatólogos, a los que se les solicitó su opinión sobre la relevancia de los dominios propuestos para valorar la gravedad en pacientes con AR, así como sus métodos de medición. De cada variable propuesta se recogieron, mediante escalas tipo Likert de 5 a 10 niveles, valoraciones referidas a: a) la prioridad; b) el grado de acuerdo en la definición; c) la factibilidad, o disponibilidad de la información en las historias clínicas o base de datos utilizadas. Se decidió seleccionar aquellas variables que más del 70% de los expertos consideró que debían ser incluidas y cuya obtención fuera factible. Las variables con una valoración intermedia se sometieron a una nueva evaluación en una segunda ronda. Tras esta

última ronda, se tuvieron en cuenta todas aquellas variables con un grado de acuerdo entre los expertos superior al 50%. La encuesta se dirigió a usuarios potenciales de la variable gravedad, con representatividad geográfica a nivel nacional.

4. Desarrollo del índice.

Una vez elegidas las posibles variables constituyentes del índice, se siguió una metodología en diferentes pasos para evaluar su comportamiento y analizar la estructura y dimensionalidad, con el fin de establecer la selección definitiva de ítems que garantizara una adecuada capacidad de discriminación entre pacientes con distintos niveles de gravedad. Se exploró la estructura y dimensiones del índice mediante análisis factorial exploratorio (AF). El AF es una técnica multivariable cuyo objetivo es cuantificar e identificar las dimensiones subyacentes del constructo gravedad y tratar de explicar la mayor parte de la varianza de este constructo con el mínimo número de factores posible; es decir, buscar y definir las dimensiones fundamentales latentes a partir de una matriz de correlaciones. Además, el AF permite reducir los datos, ya que facilita la eliminación de las variables que no tienen ninguna contribución a los factores. De esta forma, la información contenida en un número original de ítems se condensa en una cantidad menor de nuevas dimensiones (factores) con una pérdida de información mínima.

PARTE II. ESTUDIOS

Sub-estudio 1. Incremento del tratamiento con FAME a lo largo del tiempo en pacientes con artritis reumatoide de inicio del estudio pearl	27
Sub-estudio 2. Identificación de instrumentos existentes	41
Sub-estudio 3: Selección de dominios y variables que reflejan la gravedad: panel de expertos y estudio Delphi.....	56
Sub-estudio 4. Diseño y construcción de un índice compuesto de evaluación de la gravedad.....	67

Sub-estudio 1. Incremento del tratamiento con FAME a lo largo del tiempo en pacientes con artritis reumatoide de inicio del estudio PEARL

1.	Introducción.....	27
2.	Objetivo.....	28
3.	Material y métodos.....	28
3.1	Pacientes.....	28
3.2	Variables analizadas.....	29
3.3	Análisis estadístico.....	30
4.	Resultados.....	31
4.1	Descripción de las 3 subpoblaciones de pacientes con artritis de inicio.....	31
4.2	Intensidad del tratamiento con FAME.....	32
4.3	Resultados después de dos años de seguimiento.....	34
5.	Conclusión.....	39

PARTE II. ESTUDIOS

SUB-ESTUDIO 1. INCREMENTO DEL TRATAMIENTO CON FAME A LO LARGO DEL TIEMPO EN PACIENTES CON ARTRITIS DE RECIENTE COMIENZO DEL ESTUDIO PEARL

1. Introducción

El manejo de la AR ha mejorado a lo largo de los últimos 15 años, permitiendo un mejor control de la enfermedad. La disponibilidad de TB al inicio del presente siglo representó una revolución en nuestra capacidad de mejorar la actividad de la enfermedad y detener la progresión radiológica(44). Además, la intensificación terapéutica con FAME, mediante la estrategia “treat-to-target” también ha contribuido a un mejor control de la actividad(7-9).

Por otro lado, algunas evidencias iniciales obtenidas a finales del pasado siglo pasado sobre la existencia de una “ventana de oportunidad” en la AR(45) se han confirmado definitivamente a principios de este siglo(11). Estas evidencias han favorecido el desarrollo de nuevos criterios de clasificación de la AR(46) y la propuesta de realizar una intervención terapéutica precoz, habiéndose sugerido la posibilidad de iniciar el tratamiento incluso antes del diagnóstico en algunos pacientes con el fin de evitar la progresión de la enfermedad e incluso alcanzar la remisión libre de fármacos(4).

Sin embargo, se desconoce cómo han sido interiorizados estos conceptos entre los reumatólogos en la práctica clínica diaria.

2. Objetivo

El objetivo de este sub-estudio es analizar si el tratamiento con FAME y TB ha cambiado a lo largo del tiempo en un registro de artritis de inicio.

3. Material y métodos

3.1 Pacientes

Se realizó un análisis retrospectivo de los datos recogidos en el estudio PEARL (Princesa Early Arthritis Register Longitudinal study). Para ser derivados a esta consulta de referencia, los pacientes debían haber presentado 1 ó más articulaciones inflamadas durante al menos un año. Los pacientes diagnosticados de gota, artritis séptica o viral, artrosis, espondiloartritis o enfermedad del tejido conectivo durante el periodo de seguimiento se excluyeron del estudio. Sólo aquellos pacientes que cumplieron los criterios ACR de 1987(47) y los diagnosticados de artritis indiferenciada(48) después de 24 meses de seguimiento fueron incluidos en este trabajo. El registro incluye 5 visitas estructuradas (basal, 6, 12, 24 y 60 meses), en las que se recogen de sistemáticamente variables clínicas, de laboratorio, de tratamiento, radiológicas y muestras biológicas según lo establecido en el protocolo. Estas visitas son realizadas por dos reumatólogos, Ana María Ortiz (AMO) e Isidoro González-Álvaro (IG-A), pero sin protocolo terapéutico preestablecido, puesto que la decisión de cuándo y cómo tratar a los pacientes recae en el criterio de cada uno de los 11 reumatólogos pertenecientes al mismo Servicio, responsables de cada paciente. El registro se inició en el año 2000 y en la actualidad continúa el reclutamiento; los pacientes incluidos en este estudio son los que en diciembre de 2016 presentaban un periodo de seguimiento de 24 meses.

El estudio PEARL se rige por los principios de la Declaración de Helsinki de 1983 y ha sido aprobado por el Comité de Ética del Hospital Universitario de La Princesa (PI-518). Todos los pacientes firmaron un consentimiento informado al entrar en el estudio.

3.2 Variables analizadas

Con el fin de estudiar si la intensidad del tratamiento cambiaba a lo largo del tiempo, la población total fue dividida en tres subpoblaciones, en función del año del inicio de los síntomas: 2000-2004, 2005-2009 y 2010-2014.

El protocolo del registro PEARL establece una cuidadosa recogida de información en relación con el tratamiento con fármacos modificadores de la enfermedad (FAME), ya sean químicos o biológicos. Con el fin de valorar la intensidad del tratamiento (*intensity of DMARD treatment (IDT)*) se describió con anterioridad la variable calculada como la suma del número de días en tratamiento con cualquier FAME o TB, ponderada por el tipo de fármaco (1x si antipalúdicos, 1,5x si metotrexato -MTX-, leflunomida -LF-, sulfasalazina -SSZ-, sales de oro o ciclosporina -CsA- y 2x para TB) y normalizada por el número de días entre la visita basal y la visita de los 24 meses. En resumen, un IDT= 1 significa que el paciente ha recibido tratamiento con monoterapia de antipalúdicos durante todos los días de los dos primeros años de seguimiento o el 75% de los días con monoterapia de MTX, LF o SSZ, ó 50% de los días con monoterapia de un biológico. Se registró el uso de TB durante los primeros 24 meses de seguimiento. Las variables clínicas y de laboratorio recogidas por protocolo permitieron calcular el DAS28, SDAI y HUPI con el fin de valorar la actividad de la enfermedad en cada visita. Se estableció la remisión a los dos años de seguimiento en un SDAI \leq 3,3.

Además, se realizaron radiografías de manos en la visita basal, 1 y 2 años de seguimiento y fueron valoradas mediante el método de Sharp con la modificación de van der Heijde (SHS; puntuación máxima de erosión= 160 y puntuación máxima de estrechamiento del espacio articular= 120) por un evaluador experimentado, José Ivorra-Cortés (JI-C). El coeficiente de correlación intraclase se valoró mediante una segunda lectura del 10% de las radiografías y fue del 0,99.

3.3 Análisis estadístico.

Se realizó un análisis descriptivo mediante el cálculo de media y desviación estándar (DE) de variables cuantitativas con una distribución normal. Se calculó la mediana y el rango intercuartil (RIC) para aquellas variables con una distribución no normal. Se utilizó la estimación de proporciones para describir variables cualitativas. Se utilizó el test de Anova para comparar medias de variables con una distribución normal y el test de Kruskal-Wallis para variables que no presentaran una distribución normal. Se utilizó el test χ^2 para variables cualitativas.

Al considerar que había diferencias en algunas características en las tres poblaciones, se decidió realizar un análisis multivariable que permitiera valorar si estas diferencias podían introducir sesgos de asociación en los resultados analizados.

En el caso del análisis de la actividad de la enfermedad al final del seguimiento, se utilizó un modelo de regresión lineal multivariable mediante el comando *glm* de Stata. Se utilizó un modelo de regresión logística multivariable mediante el comando *logit* de Stata en el caso de la evaluación del porcentaje de pacientes en remisión al final del seguimiento. Por último, dado que la variable progresión radiológica presentaba un número de observaciones con valor 0 mayor de las que cabría esperar (modelo truncado por ceros), se decidió transformar esta variable continua en una variable categórica con tres opciones, siendo 0 no progresión radiológica, 1 baja progresión radiológica y 2 alta progresión radiológica. El punto de corte para discriminar entre baja y alta progresión radiológica se consideró la mediana de las puntuaciones de aquellos pacientes con progresión radiológica mayor que 0. Posteriormente, la progresión radiológica se analizó mediante un modelo de regresión logística ordenada multivariable mediante el comando *ologit*.

En todos los modelos multivariable se incluyeron en el modelo inicial todas aquellas variables que fueron diferentes de manera significativa en el análisis bivariable entre las tres poblaciones. Posteriormente, el modelo final se obtuvo mediante un

procedimiento manual de eliminación progresiva de variables con $p > 0,15$. Todos los análisis se llevaron a cabo con el programa Stata v12.1 (CollegeStation, Tex, USA).

4. Resultados

4.1 Descripción de las 3 subpoblaciones de pacientes con artritis de inicio

La tabla I muestra las características principales de las tres subpoblaciones en las que fue dividido el registro PEARL. Las principales diferencias observadas se refieren a: a) aumento progresivo de la edad de inicio de la enfermedad; b) disminución constante de la duración de la enfermedad en la visita basal; y c) menor porcentaje de pacientes ACPA positivo en los primeros 5 años del registro.

Tabla I. Características de las poblaciones del registro PEARL en los tres periodos de tiempo analizado.

	2000-2004 (n=148)	2005-2009 (n=167)	2010 -2014 (n=141)	P
Mujeres, n (%)	111 (75,0)	136 (81,4)	113 (80,1)	0,344
Edad (años; p50 [RIC])	50,6 [38,5-64,7]	54,2 [44,4-68,7]	58,7 [48,2-69,5]	0,001
Tabaquismo, n (%)				0,099
Nunca	88 (62,9)	87 (55,8)	74 (52,9)	
Ex-fumadores	23 (16,4)	29 (18,6)	39 (27,9)	
Fumadores actuales	29 (20,7)	40 (25,6)	27 (19,3)	
Duración enfermedad (meses; p50 [RIC])	6,6 [4,6-9,8]	5,2 [2,9-8,2]	4,1 [2,6-7,1]	<0,001
FR, n (%)	67 (45,3)	88 (52,7)	84 (59,6)	0,052
ACPA, n (%)	52 (35,6)	86 (53,4)	76 (53,9)	0,002
Criterios AR 2010 basal, n (%)	88 (59,5)	117 (70,1)	96 (68,1)	0,115
Criterios AR 1987 2 años, n (%)	103 (69,6)	113 (67,7)	106 (75,2)	0,335
Artritis indiferenciada después de 2 años, n (%)	45 (30,4)	54 (32,3)	35 (24,8)	0,335
DAS28 basal (p50 [RIC])	4,2 [3,2-5,6]	4,9 [3,6-5,7]	4,3 [3,3-5,5]	0,040
HAQ basal (p50 [RIC])	0,875 [0,5-1,625]	1 [0,625-1,625]	0,875 [0,5-1,75]	0,441

Abreviaturas: n: número; RIC: rango intercuartil; ACPA: anticuerpos anti-péptido citrulinado; p50: percentil 50 o mediano; DE: desviación estándar; IMC: índice masa corporal; AR: artritis reumatoide; FR: factor reumatoide; DAS28: disease activity score based on a 28-joint count; HAQ: Health Assessment Questionnaire.

4.2 Intensidad del tratamiento con FAME

No se observaron diferencias en la mediana de la variable IDT entre los pacientes incluidos en el registro durante los 5 primeros años respecto a los incluidos posteriormente (*figura 1A*). Sin embargo, se puede apreciar que el percentil 25 aumentó progresivamente a lo largo del tiempo (*figura 1A*). Como muestra la *figura 1B*, este hallazgo se relacionó con una intensificación significativa del tratamiento en pacientes con artritis indiferenciada a lo largo de los tres periodos de tiempo, mientras que la intensidad de tratamiento de los pacientes que cumplieron los criterios AR fue similar en las tres poblaciones.

Por otro lado, el porcentaje de pacientes que recibió TB durante los dos primeros años de seguimiento aumentó de forma significativa, de 4,7% en los pacientes incluidos en el registro PEARL entre los años 2000 a 2004 a 12,8% en los que entraron entre 2010 y 2014 ($p= 0,045$, *figura 1C*). En el grupo intermedio el porcentaje de pacientes que recibió tratamiento biológico fue de 7,2%.

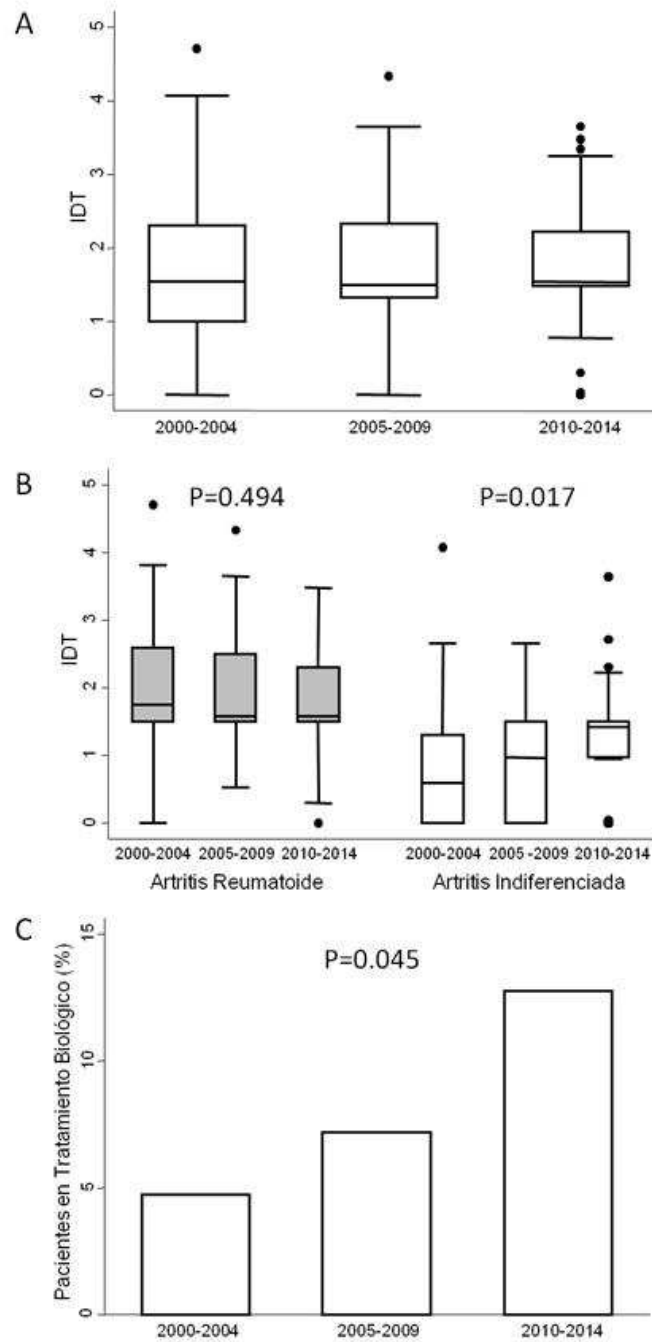


Figura 1. Descripción del tratamiento recibido por los pacientes de la cohorte PEARL en los periodos de tiempo 2000-2004, 2005-2009 y 2010-2014. A) Intensidad de Tratamiento (IDT: intensity of DMARD treatment, ver Métodos para definición) en la población completa; B) Diferenciación de la IDT entre pacientes que cumplen criterios de artritis reumatoide al final del seguimiento y los que se mantienen como artritis indiferenciada; C) Uso de terapia biológica. Los datos se muestran en A y B como la mediana de IDT (línea dentro de la caja) y los percentiles 25,75 (bordes inferior y superior de la caja), 10 y 90 (extremos de las líneas fuera de la caja). Los puntos representan los casos fuera de este rango. En C los datos se muestran en porcentaje.

4.3 Resultados después de dos años de seguimiento

La actividad de la enfermedad descendió de forma significativa en los tres grupos de pacientes, con una tendencia a presentar un DAS28 inferior lo largo del tiempo (*figura 2*, panel izquierdo). Datos similares se obtuvieron con el índice HUPI. Como se muestra en la Tabla II, (secciones media e izquierda), después de ajustar por edad al inicio de la enfermedad, género, positividad ACPA y actividad en la visita basal, los pacientes incluidos en el registro PEARL durante los periodos 2005-2009 y 2010-2014 tuvieron una actividad significativamente inferior, medida tanto con DAS28 como con HUPI, que los incluidos en el primer periodo. Además, se observó una tendencia al aumento del número de pacientes que entraban en remisión al final del seguimiento en la población incluida después de 2004 respecto a los que entraron en el periodo entre 2000 y 2004 (*figura 2*, panel derecho), lo que se confirmó después de ajustar por factores de confusión (tabla II, sección derecha).

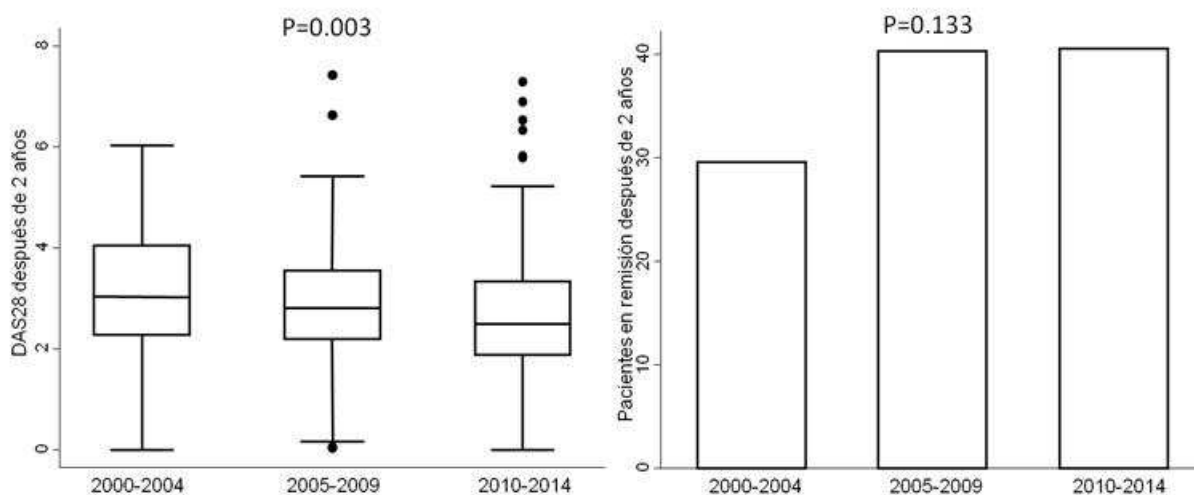


Figura 2. Actividad de la enfermedad en los pacientes de la cohorte PEARL en los 3 periodos estudiados. El panel izquierdo muestra la actividad de la enfermedad (medida según el DAS28). Los datos se muestran como la mediana del DAS28 (línea dentro de la caja) y los percentiles 25,75 (bordes inferior y superior de la caja), 10 y 90 (extremos de las líneas fuera de la caja). Los puntos representan casos fuera de este rango. El panel derecho muestra el porcentaje de pacientes en remisión definida como pacientes con SDAI < 3.

Tabla II. Variables asociadas con actividad de la enfermedad al finalizar el seguimiento.

	HUPI		DAS28		Remisión (SDAI <3,3)	
	coeficiente β (IC 95%)	p	coeficiente β (IC 95%)	P	OR (IC 95%)	p
Edad al inicio enfermedad (años)						
<45	Ref.	-	Ref.	-	Ref.	-
45-65	0,62 (-0,04 – 1,27)	0,066	0,34 (0,03 – 0,65)	0,029	0,64 (0,37 -1,11)	0,116
>65	0,65 (-0,08 – 1,38)	0,080	0,16 (-0,18 – 0,50)	0,369	0,49 (0,26 – 0,91)	0,023
Mujeres	N.I.	N.S.	0,61 (0,30 – 0,91)	<0,001	0,49 (0,28 – 0,86)	0,012
Tabaquismo	N.I.	N.S.	N.I.	N.S.	N.I.	N.S.
Duración de enfermedad visita basal	N.I.	N.S.	N.I.	N.S.	N.I.	N.S.
FR	N.I.	N.S.	N.I.	N.S.	N.I.	N.S.
ACPA	0,54 (0,00 – 1,08)	0,048	N.I.	N.S.	N.I.	N.S.
Actividad enfermedad visita basal						
Remisión	Ref.	-	Ref.	-	Ref.	-
Baja	0,54 (-0,42 – 1,49)	0,269	-0,01 (-0,55 – 0,54)	0,981	0,57 (0,25 – 1,29)	0,177
Moderada	1,26 (0,31 – 2,21)	0,010	0,56 (0,12 – 0,99)	0,012	0,39 (0,17 – 0,89)	0,025
Grave	1,39 (0,47 – 2,31)	0,003	0,86 (0,41 – 1,31)	<0,001	0,37 (0,17 – 0,84)	0,017
Año de inicio síntomas						
2000-2004	Ref.	-	Ref.	-	Ref.	-
2005-2009	-0,89 (-1,54 – -0,24)	0,007	-0,32 (-0,62 – -0,01)	0,042	2,03 (1,15 – 3,60)	0,015
2010-2014	-1,26 (-1,92 – -0,60)	<0,001	-0,55 (-0,86 – -0,24)	<0,001	2,09 (1,16 – 3,76)	0,014

Abreviaturas: HUPI: Hospital Universitario Princesa Indez; DAS28: Disease Activity Score with 28 joint counts; SDAI: Simplified Disease Activity Index; OR: Odds ratio; Ref: Referencia; N.I.: no incluido en el modelo final por no ser significativo (N.S.); FR: factor reumatoide; ACPA: anticuerpos anti-peptido citrulinado; Análisis multivariable realizado mediante un modelo de regresión lineal para HUPI y DAS28 y mediante un modelo de regresión logística para remisión (Ver Métodos para información más detallada).

Con respecto a la progresión radiológica, a pesar de que más del 50% de los pacientes no presentaron un aumento del SHS medido en manos durante los dos años de seguimiento (*figura 3A*), la puntuación de erosión fue significativamente inferior en las poblaciones más recientes (*figura 3B*), mientras que hubo una tendencia no significativa a presentar una puntuación superior en el estrechamiento del espacio articular en el último grupo de pacientes (2010-2014; *figura 3C*). El análisis multivariable confirmó que, después de ajustar por género y edad, los pacientes incluidos en el registro PEARL durante el último periodo mostraron una puntuación inferior en la progresión de la erosión de forma estadísticamente significativa con respecto a aquellos incluidos en el periodo 2000 a 2004 (tabla III). Respecto al estrechamiento del espacio articular, sólo los pacientes con una edad superior a los 65 años mostraron una progresión significativa con respecto a los de menos de 45 años y esta fue la única variable asociada con este resultado (tabla III, sección izquierda).

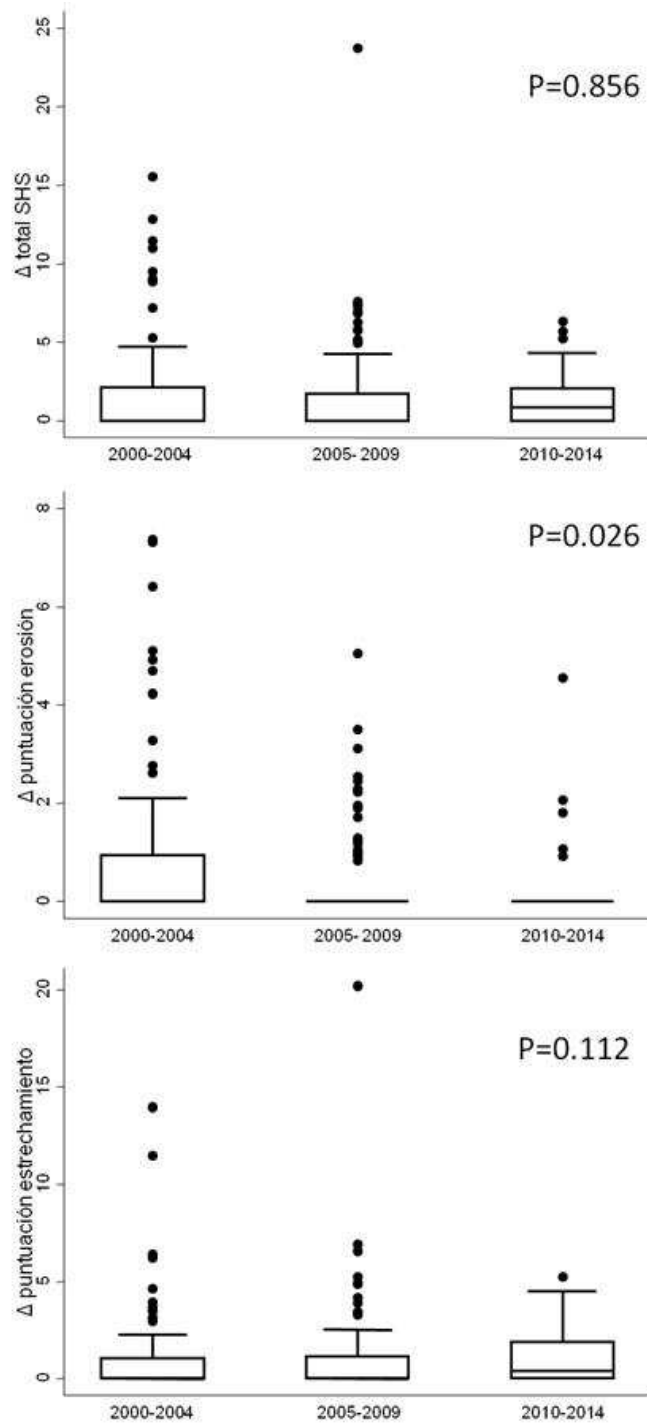


Figura 3. Progresión radiológica en manos en los pacientes de la cohorte PEARL en los 3 periodos estudiados. A) Incremento del índice Sharp con la modificación de van der Heijde (SHS) medido en manos durante los dos años de seguimiento. B) Puntuación de la subescala de erosión del SHS. C) Puntuación de la subescala de pinzamiento articular. Los datos se muestran como la mediana del SHS o sus subescalas (línea dentro de la caja) y los percentiles 25,75 (bordes inferior y superior de la caja), 10 y 90 (extremos de la líneas fuera de la caja). Los puntos representan casos fuera de este rango.

Tabla III. Variables asociadas con los componentes de la progresión radiológica (erosión y estrechamiento del espacio articular) entre la visita basal y los dos años de seguimiento en manos mediante el método de Sharp con la modificación de van der Heijde.

Edad al inicio enfermedad	Puntuación erosión		Puntuación estrechamiento	
	OR (IC 95%)	p	OR (IC 95%)	p
<45	Ref.	-	Ref.	-
45-65	2,81 (1,05 – 7,57)	0,040	1,75 (0,76 – 3,99)	0,186
>65	3,79 (1,36 – 10,57)	0,011	5,57 (2,42 – 12,86)	<0,001
Mujeres	0,58 (0,28 – 1,18)	0,134	N.I.	N.S.
FR	N.I.	N.S.	1,83 (0,92 – 3,64)	0,085
ACPA	N.I.	N.S.	0,54 (0,27 – 1,05)	0,071
Año de inicio síntomas				
2000-2004	Ref.	-	N.I.	N.S.
2005-2009	0,62 (0,31 – 1,26)	0,188	N.I.	N.S.
2010-2014	0,27 (0,09 – 0,78)	0,016	N.I.	N.S.
Puntos de corte				
No/baja progresión	1,27 (0,23 – 2,30)	-	1,41 (0,63 – 2,19)	-
Baja/alta progresión	2,21 (1,14 – 3,28)	-	2,52 (1,69 – 3,35)	-

Abreviaturas: OR: Odds ratio; Ref: Referencia; FR: factor reumatóide; ACPA: anticuerpos anti-peptido citrulinado; N.I.: no incluido en el modelo final por no ser significativo (N.S.).

5. Conclusión

En resumen, nuestros datos reflejan que el tratamiento con FAME y TB en pacientes con artritis de inicio durante los dos primeros años de la enfermedad se ha intensificado de manera gradual durante los últimos 15 años. Este abordaje ha permitido obtener mejores resultados clínicos, lo que a su vez dificulta la identificación de biomarcadores de gravedad.

Sub-estudio 2. Identificación de instrumentos existentes

1.	Objetivos	41
2.	Material y métodos.....	41
2.1	Estrategia de búsqueda.....	41
2.2	Selección de estudios.....	42
2.3	Evaluación de la calidad de los índices	43
3.	Resultados.....	43
3.1	Identificación de instrumentos para evaluación de gravedad	43
3.2	PAS	45
3.3	Rheumatoid arthritis records-based index of severity (RARBIS)	48
3.4	Claims-based index for rheumatoid arthritis severity (CIRAS)	48
4.	Evaluación crítica de los índices.....	52
5.	Conclusión.....	54

SUB-ESTUDIO 2. IDENTIFICACIÓN DE INSTRUMENTOS EXISTENTES

1. Objetivos

La necesidad de disponer de una variable de gravedad válida que permitiera evaluar la utilidad de biomarcadores en artritis de reciente comienzo, nos hizo plantearnos la necesidad de realizar una valoración de las herramientas existentes, con el objetivo de: 1) identificar índices o instrumentos que permitan discriminar entre pacientes con AR graves y no graves y; 2) evaluar su validez y aplicabilidad a partir de datos existentes en la historia clínica. Posteriormente, se estudió si los instrumentos identificados permitirían estimar la gravedad de pacientes con artritis de reciente comienzo en un momento suficientemente precoz del seguimiento como para poder realizar estudios de biomarcadores eficientes, para lo que establecimos como punto de corte los dos primeros años tras el diagnóstico.

2. Material y métodos

Se realizó una revisión sistemática de estudios sobre desarrollo o validación de índices o variables de gravedad en pacientes adultos con AR.

2.1 Estrategia de búsqueda

Un investigador con experiencia, Loreto Carmona (LC), realizó una búsqueda sensible en MEDLINE (1950-enero 2017), mediante términos Mesh y texto libre, sin restricción de idiomas. La estrategia incluyó sinónimos de “artritis reumatoide”, “estudios de validación” o “desarrollo de (“variables” o “índices” o “cuestionarios”) o “estudios pronósticos”, y del término “gravedad” La estrategia de búsqueda está disponible como material suplementario.

En la *figura 4* se muestra la estrategia de búsqueda.

"Psychometrics"[Mesh] Psychometrics/*methods[MH] clinimetric[TW] AND properties[TW] applicability[TW] "Reproducibility of results"[MH] reliability[TW] feasibility[TW] Validation Studies[PT] Evaluation Studies[PT] Prognosis[MH]
"Disease Progression"[MH] "Severity of Illness Index"[MH] (index[Text Word] AND severity[Text Word]) "risk appraisal"[Text Word] (poor[Text Word] AND progn*[Text Word]) severity[Title/Abstract]
((rheumatoid arthritis[Text Word]) OR arthritis, rheumatoid[MeSH Terms])

Figura 4. Estrategia de búsqueda.

2.2 Selección de estudios

La selección de estudios fue realizada por un revisor entrenado Esther Toledano (ET) en tres pasos: 1) selección por títulos; 2) selección por resúmenes y 3) recopilación del texto completo de las referencias seleccionadas y evaluación de los mismos,

efectuando la selección final en función de los criterios fijados con anterioridad. Se registró una razón para excluir un estudio en todos los casos no incluidos.

2.3 Evaluación de la calidad de los índices

Los índices o cuestionarios de gravedad encontrados en este proceso de búsqueda se analizaron mediante la lista de verificación de COSMIN (**CO**nsensus-based **S**tandards for the selection of health **M**easurement **I**nstruments)(43). Esta lista contiene normas para adecuar los requisitos de diseño y métodos estadísticos para estudios sobre propiedades psicométricas de instrumentos de medición de la salud y puede ser utilizada para determinar si un estudio cumple con los estándares de buena calidad metodológica. Contiene cuatro pasos que permiten evaluar la calidad metodológica de los distintos aspectos de la validación del instrumento de medida a estudio.

Finalmente, en cada una de las variables de los índices se analizaron las cualidades consideradas de especial relevancia para nuestro grupo de investigación, como son: claridad de la definición, viabilidad, presencia en los dos primeros años de evolución y dominio al que pertenecen.

3. Resultados

3.1 Identificación de instrumentos para evaluación de gravedad

La estrategia de búsqueda identificó 3.519 artículos (*figura 4*). Se excluyeron 3.505 tras lectura de título y resumen. El motivo de exclusión fundamental fue no tener como objetivo la validación o desarrollo de un índice o cuestionario de gravedad en la AR. De los 14 restantes, se excluyeron otros 7 por el mismo motivo tras lectura en detalle. Finalmente, se incluyeron 7 artículos en la revisión sistemática cuyo objetivo fue el desarrollo o validación de un índice de gravedad en la AR. En total se encontraron tres índices de gravedad en la AR: la PAS, el RARBIS y el CIRAS, cuyas características se describen a continuación. Los artículos excluidos tras la lectura en detalle y las causas de exclusión se detallan en la tabla IV.

Tabla IV. Estudios excluidos tras selección.

Estudio	Razón exclusión
Waltz (2000)(49)	No inclusión índice gravedad AR
Lassere (2001)(50)	No inclusión índice gravedad AR
Leeb (2003)(51)	No inclusión índice gravedad AR
Chiou (2006)(52)	No inclusión índice gravedad AR
Landewe (2006)(53)	No inclusión índice gravedad AR Artículo opinión
Harrison (2007)(54)	No inclusión índice gravedad AR
Gossec (2009)(55)	No inclusión índice gravedad AR

El diagrama de flujo se representa en la *figura 5*.

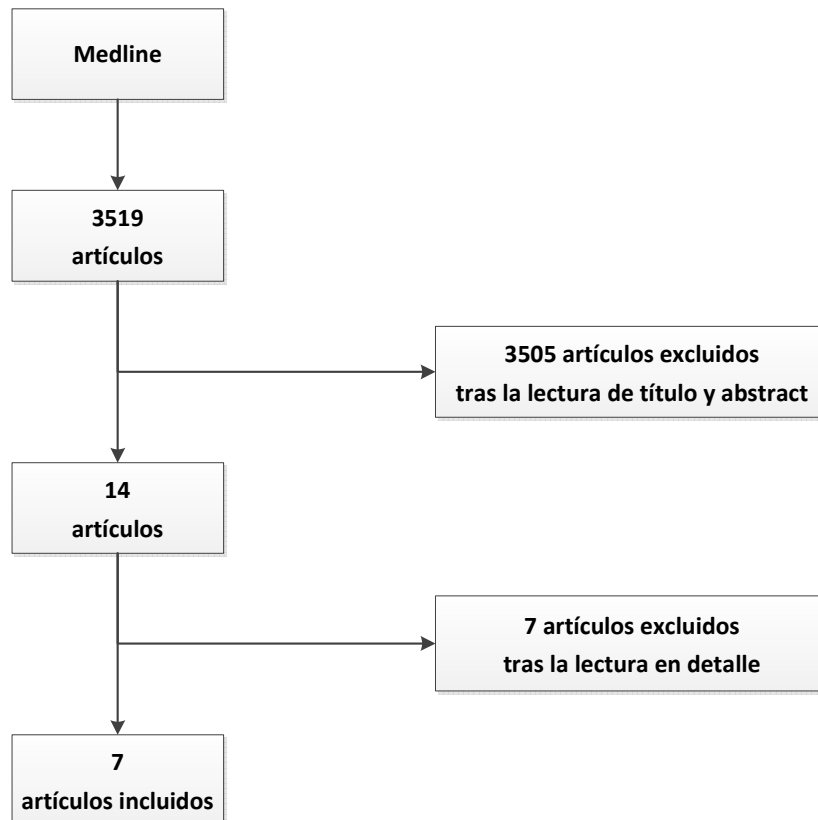


Figura 5. Diagrama de flujo de artículos.

3.2 PAS

La escala de actividad del paciente (PAS) fue una variable desarrollada por el Dr. Wolfe a partir de la media de las puntuaciones del *Health Assessment Questionnaire* (HAQ) (calibrado de 0 a 10), y dos escalas visuales analógicas, una para dolor y otra para la actividad total, ítems todos ellos proporcionados por el paciente. Esta escala cumplimentada por el paciente había demostrado una fuerte correlación con la actividad de la enfermedad en ensayos clínicos(56). El Dr. Wolfe la emplea como sinónimo de gravedad en un estudio en 2006(57) y la intentó validar evaluando su capacidad de discriminación para pacientes con distinta carga de tratamiento.

En este estudio de validación de criterio se incluyeron pacientes con AR pertenecientes al National Data Bank for Rheumatic Diseases (NDB), una cohorte longitudinal de AR. Los participantes de la cohorte NDB eran reclutados por reumatólogos que ejercían en EE.UU, y se les hacía un seguimiento prospectivo con cuestionarios semestrales, registro de información de variables demográficas y de tratamiento. Para dicho estudio de validación, los autores incluyeron pacientes con AR que hubieran completado al menos un cuestionario semestral entre enero de 2002 y diciembre de 2004, excluyendo los casos que participaban en registros de seguridad. Para cada paciente, se seleccionó de forma aleatoria una única visita en el intervalo de tiempo descrito.

Las características de los 7.541 pacientes incluidos en el estudio de validación de criterio se presentan en las tablas V y VI(58). De forma similar a los otros índices encontrados, no se ofrece información de variables basales que típicamente están asociadas con mayor gravedad. En el momento de la visita la mayoría de los pacientes (81,7%) utilizaba un FAME y/o TB, 74,2% un FAME, 33,1% un TB, y el 18,3% ni FAME ni TB. El valor de PAS medio fue 3,4. La puntuación mayor de la escala PAS se alcanzó en pacientes que no tomaban ninguna medicación (PAS 3,7; rango intercuartil (RIC) 1,7-5,7) y en los que no habían tomado con anterioridad FAME o TB (PAS 3,7; RIC 2,0-5,4).

Tabla V. Características de los pacientes de las cohortes incluidas en la revisión.

Características	Wolfe, 2006	Ting, 2008 Ting, 2005
Índice gravedad	<i>Composite score for RA severity index</i>	RARBIS CIRAS
Edad ,años (media±DE)	60,8±13,4	70,6±11,1
Mujeres	5882 (78)	11 (9)
Duración enfermedad (media±DE)	14,6	
Clase funcional ACR		
I		93 (78)
II		8 (7)
III		6 (5)
IV		4 (3)
Visitas reumatólogo	-	3 (2,1)
Hospitalizaciones	-	7 (6)
Rigidez matutina		
<1		70 (58)
1-4		25 (21)
>4		8 (7)
Brotos		
0		65 (54)
1		22 (18)
1-4		11 (9)
5+		3 (3)
Articulaciones inflamadas	-	64 (53)
Enfermedad erosiva	-	61 (51)
Estrechamiento espacio articular	-	74 (62)
Nódulos reumatoides	-	41 (34)
Vasculitis	-	1 (1)
Nódulos pulmonares	-	11 (9)
Subluxación atlo-axoidea	-	2 (2)
Glucocorticoides		
Intraarticular	2308 (30,6)	11 (9)
Intramuscular		1 (1)
Uso FAME	5595 (74,2)	-
Uso biológicos	2496 (33,1)	-
Uso FAME y/o biológicos	6161 (81,7)	-

Abreviaturas: DE: desviación estándar; ACR: American College of Rheumatology; FAME: fármacos modificadores de la enfermedad.

Tabla VI. Características de los estudios incluidos en la revisión.

Estudio	Índice estudiado	Comparador gravedad enfermedad	VARIABLES	Correlación RARBIS con subescala medicación	Correlación RARBIS sin subescala medicación
Ting, 2008	CIRAS	RARBIS	Edad y sexo	0,08	0,05
			Visitas reumatólogo	0,01	-
			Visitas rehabilitación	0,01	0,04
			Síndrome de Felty	0,01	0,03
			Nº marcadores inflamación	0,14	0,08
			Test FR	0,02	0,04
			Nº recuentos plaquetas	0,03	0,01
			Nº paneles bioquímicos	0,01	0,02
			CIRAS compuesto	0,56 (p< 0,0001)	0,51 (p< 0,0001)
Wolfe, 2006	PAS		Edad	AUC= 0,64	AUC= 0,70
			Sexo	Para detección valor PAS medio cohorte	Para detección 1º y 4º cuartil PAS cohorte
			Duración enfermedad		
			FAME y/o biológicos actuales		
			FAME y/o biológicos previos		

Abreviaturas: RARBIS: Rheumatoid arthritis records-based index of severity; CIRAS: Claims-based index for rheumatoid arthritis severity; PAS: Patient Activity Scale; FR: factor reumatoide; AUC: area under the curve.

Los autores valoraron la capacidad de las variables de tratamiento para predecir la puntuación PAS mediante el cálculo del área bajo la curva (ROC-AUC) y el porcentaje de clasificación correcta, obteniéndose un AUC de 0,64 y un porcentaje de clasificación correcta del 60,5% (tabla VI).

El Dr. Wolfe concluye que el uso de variables demográficas y de tratamiento en bases de datos administrativas no permite diferenciar grupos de pacientes en función de su gravedad con adecuada sensibilidad y/o especificidad.

3.3 Rheumatoid arthritis records-based index of severity (RARBIS)

El índice RARBIS fue concebido como un índice de gravedad de AR basado en registros médicos. Utiliza una serie de indicadores de gravedad previamente elegidos mediante metodología Delphi por un panel de 6 expertos reumatólogos(59). El RARBIS incluye 5 subescalas o dominios: cirugía, radiología, manifestaciones extraarticulares, clínica y variables de laboratorio. La subescala tratamiento fue finalmente excluida del RARBIS total. Aquellos dominios que habían sido valorados por el panel como muy relacionados con la gravedad se les asignó una puntuación mayor en el índice.

Posteriormente, se estudió el desarrollo del índice en una cohorte de veteranos y la validez de constructo mediante la comparación del RARBIS con la intensidad de tratamiento (baja -sin FAME ni TB-, moderada (hidroxicloroquina, sales de oro o sulfasalacina), alta (metotrexato, azatioprina, leflunomida o ciclosporina A) o muy alta (TB), demostrando una correlación con la intensidad de tratamiento de la AR (validez de constructo)(60), aunque la asociación era débil ($r= 0,35$, IC 95% 0,18-0,55).

En un estudio posterior se estableció la validez convergente con respecto al DAS28. Se seleccionaron para ello 100 pacientes de la cohorte BRASS (Brigham and Women's Hospital Rheumatoid Arthritis Sequential Study) con un DAS28 distribuido de forma equitativa en 4 cuartiles(61). La correlación entre el RARBIS total y el DAS28 fue débil ($r= 0,41$, IC 95% 0,23-0,56), así como la de las diferentes subescalas o dominios del índice RARBIS con el DAS28, que en algunos casos fue inexistente. No obstante, la correlación mejoró cuando los autores aumentaron el peso de la subescala clínica y disminuyeron el peso de la subescala cirugía ($r= 0,48$, IC 0,31-0,62)(61).

3.4 Claims-based index for rheumatoid arthritis severity (CIRAS)

El mismo grupo de investigación del RARBIS desarrolló con posterioridad el CIRAS(62), un índice de gravedad de la AR basado en indicadores potenciales de gravedad de diferentes dominios, tanto demográficos, como clínicos y, sobre todo, de utilización de servicios sanitarios. Para ello se incluyeron 120 pacientes de la mencionada cohorte de

veteranos del New England VA Health System que tenían al menos 2 visitas registradas con un diagnóstico de AR (CIE 9ª edición, código 714.0), visitas hospitalarias entre julio de 1.999 y junio de 2.001 y registros suficientes en relación con la enfermedad. Las características de los pacientes se detallan en las tablas V y VI(58).

Los autores desarrollaron el índice CIRAS mediante modelos de regresión lineal, en función de diferentes variables extraídas de datos administrativos. Utilizaron diversos procedimientos de selección con el fin de lograr el mejor modelo estadístico, según su R^2 . El CIRAS final incluyó las variables administrativas presentes en el mejor modelo: edad, sexo, síndrome de Felty, número de visitas al reumatólogo, visitas al rehabilitador, factor reumatoide (FR), recuento de plaquetas, número de marcadores inflamatorios y paneles bioquímicos. Cada una de las variables fue ponderada por su coeficiente de correlación en el análisis multivariable. Se evaluó la validez de constructo mediante análisis de correlación entre el CIRAS y el RARBIS, obteniéndose un $r = 0,56$ entre el índice CIRAS y el RARBIS con la subescala medicación, y un $r = 0,51$ sin la subescala medicación. La tabla VI muestra los coeficientes de correlación ajustados para cada variable propuesta.

En un estudio posterior (63) se evaluó la validez convergente del CIRAS con respecto al DAS28-PCR y al multi-dimensional health assessment questionnaire (MD-HAQ). Para ello se seleccionaron 315 pacientes, incluidos tanto en la cohorte BRASS como en Medicare, que cumplían el criterio de tener al menos una medida válida de DAS28-PCR en los 365 días previos al reclutamiento. La correlación entre el CIRAS y el DAS28-PCR ($r = 0,07$) y el MD-HAQ ($r = 0,008$) fue muy baja. Finalmente, los resultados de un modelo de regresión lineal con variables adicionales extraídas de datos administrativos presentes en la cohorte Medicare no incluidas en el CIRAS (afectación pulmonar intersticial, cirugía de mano, solicitud del test de la tuberculina, solicitud ACPA, corticoides, opioides, AINEs, número de fármacos biológicos y no biológicos) demostraron la escasa capacidad del CIRAS para explicar la variabilidad del DAS28-PCR ($R^2 = 0,23$).

La descripción de las variables, dominios y puntuación de los 3 índices incluidos en la revisión se describen en la tabla VII.

Tabla VII. Descripción de las variables, dominios y puntuación de índices RARBIS, CIRAS y PAS.

Ítem/variable	Dominio	Puntuación variable	Puntuación dominio	Puntuación total
Índice RARBIS				15
Fusión C1-C2	Cirugía	3	5	
Cualquier articulación mano	Cirugía	1	5	
Cualquier articulación pie	Cirugía	1	5	
Articulaciones grandes (caderas, rodillas, hombros, codos)	Cirugía	2 máx	5	
Subluxación C1-C2	Radiología	3	4	
Cualquier erosión	Radiología	1	4	
Vasculitis	Manifestaciones extraarticulares	1	1	
Nódulos pulmonares	Manifestaciones extraarticulares	1	1	
Brote artritis	Clínica	3 máx	3	
Valoración global enfermedad médico	Clínica	2	3	
Estado funcional	Clínica	3 máx	3	
Rigidez matutina	Clínica	2 máx	3	
FR > límite normalidad	Laboratorio	1	2	
VSG > edad/2 ó PCR > límite normalidad ó plaquetas > 450.000	Laboratorio	1	2	
Índice RARBIS con medicación				18
Medicación	Tratamiento	3	3	

Ítem/variable	Dominio	Puntuación variable	Puntuación dominio	Puntuación total
Índice CIRAS				6.5
Edad	Identificación/datos personales		-0,066	
Género 0: hombre 1: mujer	Identificación/datos personales		-0,092	
Nº marcadores inflamatorios 0: no 1: sí	Variables analíticas inflamación		0,60	
Visitas Rehabilitador 0: no 1: sí	Utilización servicios sanitarios		0,69	
FR 0: no 1: sí	Variables inmunológicas		2,1	
Síndrome Felty 0: no 1: sí	Manifestaciones extraarticulares		2,3	
Recuento Plaquetas 0: recuento plaquetas=0 1: recuento plaquetas=1 2: recuento plaquetas=2 3: recuento plaquetas=3 4: recuento plaquetas=4 5: recuento plaquetas≥5	Variables analíticas inflamación		0,42	
Paneles Bioquímicos 0: recuento bioquímico=0 1: recuento bioquímico=1 2: recuento bioquímico=2 3: recuento bioquímico=3 4: recuento bioquímico=4 5: recuento bioquímico≥5	Variables analíticas bioquímicas		-0,14	
Visitas Reumatólogo 0: visitas reumatólogo=0 1: visitas reumatólogo=1, 2, 3 ó 4 2: visitas reumatólogo>4	Utilización servicios sanitarios		0,52	
PAS	Actividad	0-10	0-10	10
HAQ	Calidad de vida	0-3	0-3	
VAS dolor	Dolor	0-10	0-10	
VAS global	Actividad/función/dolor	0-10	0-10	

Abreviaturas: C1-C2: cervical 1-2; VSG: velocidad sedimentación globular; PCR: proteína C reactiva; RARBIS: Rheumatoid arthritis records-based index of severity; CIRAS: Claims-based index for rheumatoid arthritis severity; PAS: Patient Activity Scale; FR: factor reumatoide; HAQ: Health Assessment Questionnaire; VAS: visual analog scale.

4. Evaluación crítica de los índices

Los aspectos de la validación estudiados, así como las cualidades consideradas de especial relevancia por el grupo investigador se describen para los tres índices en las tablas VIII, IX y X.

Tabla VIII. Índice RARBIS. Descripción características relevantes.

Ítem/variable	Claridad definición	Feasability (viabilidad)	2º año	Dominio
Fusión C1-C2	No definición Registrado en historia clínica	+++	No	Cirugía
Cualquier articulación mano	No definición Registrado en historia clínica	+++	No	Cirugía
Cualquier articulación pie	No definición Registrado en historia clínica	+++	No	Cirugía
Articulaciones grandes (caderas, rodillas, hombros, codos)	No definición Registrado en historia clínica	+++	No	Cirugía
Subluxación C1-C2	No definición Registrado en historia clínica	+++	No	Variables radiológicas
Cualquier erosión	No definición Registrado en historia clínica	+++	Sí	Variables radiológicas
Vasculitis	No definición Registrado en historia clínica	+++	Sí	Manifestaciones extraarticulares
Nódulos pulmonares	No definición Registrado en historia clínica	+++	Sí	Manifestaciones extraarticulares
Brote artritis	“flare up”, “ongoing”, “active”	+++	Sí	Manifestaciones extraarticulares
Valoración global enfermedad médico	No definición Registrado en historia clínica	++	Sí	Inflamación
Estado funcional	No definición Registrado en historia clínica Si no información, no limitación funcional	+++	Sí	Valoración global enfermedad
Rigidez matutina	Minutos	+++	Sí	Calidad vida
Variables analíticas	FR > límite superior normalidad VSG > edad/2 ó PCR > límite superior normalidad o plaquetas > 450.000	+++	Sí	Inflamación
Medicación	HQ, oro, SSZ, MTX, LFCF, AZA, CsA, anakinra, ADA, ETA, IFX	+++	Sí	Variables analíticas inflamación

Abreviaturas: C1-C2: cervical 1-2; HQ: hidroclicloroquina; SSZ: sulfasalacina; MTX: metotrexato; AZA: azatioprina; CsA: ciclosporina A; ADA: adalimumab; ETA: etanercept; IFX: infliximab. Viabilidad: +: pocas variables presentes; ++: bastantes variables presentes; +++: todas las variables presentes.

Tabla IX. Índice CIRAS. Descripción características relevantes.

Ítem/variable	Claridad definición	Feasability (viabilidad)	2º año	Dominio
Edad	Sí	+++	Sí	Identificación/datos personales
Género	Sí	+++	Sí	Identificación/datos personales
Nº marcadores inflamatorios	VSG o PCR No está definida la escala de medida de las variables Incluidas en una misma categoría	++	Sí	Variables analíticas inflamación
Visitas Rehabilitador	No definición Se contó cada visita terapia física/ocupacional como una visita	+++	Sí	Utilización servicios sanitarios
FR	No está definido el punto de corte	+++	Sí	Variables inmunológicas
Síndrome Felty	No definición Síndrome Felty	+		Manifestaciones extraarticulares
Recuento Plaquetas	No está definida la escala de medida de las variables	++	Sí	Variables analíticas inflamación
Paneles Bioquímicos	No definición parámetros incluidos	++	-	Variables analíticas bioquímicas
Visitas Reumatólogo	No definición	+++	Sí	Utilización servicios sanitarios

Abreviaturas: VSG: velocidad sedimentación globular; PCR: proteína C reactiva; FR: factor reumatoide). Viabilidad: +: pocas variables presentes; ++: bastantes variables presentes; +++: todas las variables presentes

Tabla X. Composite score for RA severity index. Descripción características relevantes.

Ítem/variable	Claridad definición	Feasability (viabilidad)	2º año	Dominio
PAS	Escala 0-10	+++	Sí	Actividad
HAQ	Escala 0-3	+++	Sí	Calidad de vida/función
VAS dolor	Escala 0-10	+++	Sí	Dolor
VAS global	Escala 0-10	+++	Sí	Actividad/función/dolor

Abreviaturas: PAS: Patient Activity Scale; HAQ: Health Assessment Questionnaire; VAS: visual analog scale. Viabilidad: +: pocas variables presentes; ++: bastantes variables presentes; +++: todas las variables presentes.

Únicamente se pudo evaluar la calidad metodológica de aquellos aspectos de la validación que fueron analizados para cada uno de los índices mediante *The COSMIN checklist*, según se describe en la tabla XI.

Tabla XI. Calidad metodológica de la validación de los índices incluidos.

Aspecto validez	PAS	RARBIS	CIRAS
Lógica	+	++	+
De contenido	NE	NE	NE
De constructo	+	++	++
De criterio	+++	NE	NE

Abreviaturas: PAS: Patient Activity Scale; RARBIS: Rheumatoid arthritis records-based index of severity; CIRAS: Claims-based index for rheumatoid arthritis severity; NE: no evaluado. Calidad metodológica según COSMIN: + baja; ++ moderada; +++: excelente.

En definitiva, los resultados de la presente revisión constatan que no existe ningún índice que permita estimar la gravedad de la AR de forma temprana en los 2 años de seguimiento (ni a largo plazo) y que se adapte a la estrategia actual de manejo terapéutico. Por todo lo expuesto creemos razonable el desarrollo de un nuevo índice de gravedad *ad hoc* para pacientes de una consulta de ARC en la actualidad.

Sub-estudio 3: Selección de dominios y variables que reflejan la gravedad: panel de expertos y estudio Delphi

1.	Objetivo.....	56
2.	Material y métodos.....	56
3.	Resultados.....	58

SUB-ESTUDIO 3: SELECCIÓN DE DOMINIOS Y VARIABLES QUE REFLEJAN LA GRAVEDAD: PANEL DE EXPERTOS Y ESTUDIO DELPHI

1. Objetivo

Seleccionar los dominios relevantes para la construcción de un índice de gravedad en la AR, gracias al criterio de un panel de expertos.

2. Material y métodos

Se convocó a un panel de expertos clínicos, con experiencia demostrada en artritis de inicio, y amplia representatividad a nivel local y nacional. El número de expertos convocados fue de diez en dos reuniones sucesivas atendiendo a su disponibilidad: Lydia Abásolo (LA, Hospital Clínico San Carlos), José Luis Andreu (JLA, Hospital Universitario Puerta de Hierro), Alejandro Balsa (AB, Hospital Universitario La Paz), Tatiana Cobo (TC, Hospital Infanta Sofía), Javier García (JG, Hospital Doce de Octubre), José de la Mata (JM, Instituto de Salud Osteoarticular –ARI–), Javier Orte (JO, Hospital Universitario Ramón y Cajal), Ana María Ortiz (AMO, Hospital Universitario La Princesa), Olga Sánchez Pernaute (OS-P, Fundación Jiménez Díaz) y Virginia Villaverde (VV, Hospital Universitario de Móstoles). La reunión fue moderada por una facilitadora de ideas, Loreto Carmona (LC), experimentada en dinámica de grupos focales.

Antes de ambas reuniones, los investigadores diseñaron un mapa conceptual del constructo gravedad, en base a lo publicado y a las experiencias de los investigadores principales, con el fin de dirigir adecuadamente el grupo y poder extraer la información relevante.

Al primer panel de expertos acudieron 6 reumatólogos (LA, JLA, JG, JM, JO y AMO). Tras una breve introducción sobre el objetivo, el procedimiento del proyecto y un

pequeño resumen sobre la revisión sistemática realizada para valorar los índices de gravedad validados hasta el momento, los dominios y variables constituyentes y las limitaciones clínicas y metodológicas que planteaban estos índices para el objetivo de nuestro estudio, se les pidió que definieran los dominios o dimensiones del constructo “gravedad”, entendido como “enfermedad temprana grave”. Se insistió en la premisa de que fuera un constructo discriminativo que identificara retrospectivamente pacientes distintos con respecto a la gravedad, formada por dominios/variables que pudieran ser medidos en los 2 primeros años de evolución de la enfermedad. El panel fue estimulado mediante técnicas de tormenta de ideas y preguntas dirigidas concretas.

En el segundo panel de expertos, que contó con la asistencia de 4 reumatólogos (AB, TC, OS-P y VV), se realizó un ejercicio de votación con el fin de profundizar sobre las variables identificadas por el primer panel de expertos. En este ejercicio también participaron los investigadores del proyecto (LC, IG-A y ET).

Tras elaborar la lista de dominios y variables constituyentes, se realizó una encuesta Delphi para recoger las opiniones de otros especialistas en Reumatología, a los que se les pidió opinión sobre la relevancia de los dominios propuestos para valorar la gravedad en pacientes con AR, así como sus métodos de medición. De cada variable propuesta se recogieron valoraciones con respuestas tipo Likert de 5 a 10 niveles (0: completamente en desacuerdo y 5-10: completamente de acuerdo, según la pregunta):

- La prioridad, mediante el grado de acuerdo sobre la importancia de la variable y su inclusión en un índice de gravedad (escala 0-10)
- El grado de acuerdo en la definición (escala 0-5).
- La factibilidad, o disponibilidad de la información en las historias clínicas o base de datos utilizadas (escala 0-5).

Se decidió seleccionar aquellas variables consideradas como necesarias en el índice por más del 70% de los expertos y cuya obtención fuera factible. Se decidió rechazar las que sólo contaran con la consideración de un 20% de expertos. Finalmente, las variables con una valoración intermedia se someterían a una nueva evaluación en una segunda ronda. Tras esta última ronda, se tuvieron en cuenta todas aquellas variables que contaran con un grado de acuerdo entre los expertos superior al 50%. El estudio Delphi se realizó de forma anónima. Excepcionalmente, se llevaría a cabo una tercera ronda en el caso de que hubiera desacuerdo, previa exploración de las causas del mismo mediante preguntas abiertas.

La encuesta se dirigió a usuarios potenciales de la variable gravedad, fundamentalmente investigadores y reumatólogos con representatividad geográfica a nivel nacional. El número adecuado de participantes se estimó entre 25 y 100.

3. Resultados

El panel de expertos, convocado en dos reuniones sucesivas, permitió identificar los siguientes **factores modificadores** de la gravedad:

- Apoyo socio familiar.
- Condiciones socio-económicas.
- Accesibilidad a los recursos sanitarios.
- Nivel cultural.
- Conocimiento de la enfermedad.
- Tipo de trabajo realizado.
- Adherencia al tratamiento.

El panel también identificó los siguientes **factores predictivos o coadyuvantes**:

- FR.
- ACPA.

Igualmente se identificaron como potenciales **criterios externos** para evaluar la validez de criterio:

- Valoración de gravedad por el médico: VGEM (valoración global enfermedad médico)
- Carga de tratamiento.
 - Líneas de tratamiento.
 - Número de fármacos utilizados.
 - Dosis de FAME y tratamiento biológico.
 - Dosis acumulada de corticoides.

El panel de expertos identificó 17 potenciales variables del constructo gravedad (tabla XII), que se sometieron al criterio del resto de los reumatólogos, mediante metodología Delphi.

Tabla XII. Variables constructo gravedad 1ª ronda Delphi.

Número	Variable
1	Poliartritis
2	Afectación de grandes articulaciones
3	Actividad elevada a dos años de seguimiento
4	Actividad mantenida a dos años de seguimiento
5	Reactantes de fase aguda elevados
6	Manifestaciones extraarticulares
7	Imposibilidad para conseguir la remisión
8	Suspensión de metotrexato por ineficacia
9	Ausencia de respuesta al metotrexato
10	Número de fármacos considerados ineficaces
11	Necesidad de tratamiento agresivo a dos años
12	Erosiones en el momento del diagnóstico
13	Número de erosiones a los dos años del diagnóstico
14	Ausencia de mejoría en el HAQ
15	Necesidad de cirugía ortopédica/prótesis
16	Ingresos hospitalarios
17	Corticoddependencia

Abreviaturas: HAQ: Health Assessment Questionnaire.

La encuesta Delphi fue enviada a 100 reumatólogos de toda España. Se obtuvo una tasa de respuesta del 61%. El 56% de los encuestados eran mujeres. En cuanto a la edad, el 3,5% tenía <35 años, el 26,7% entre 35-44 años, y el 25% entre 50-54 años, con una media de 20 años de experiencia excluyendo la formación MIR. El 98% desarrollaba su actividad asistencial en la sanidad pública y el 28% en la sanidad privada. Se obtuvo una amplia representatividad nacional (las provincias más representadas fueron La Coruña, Barcelona, Madrid y Sevilla).

En las tablas XIII, XIV y XV se detallan las respuestas de los encuestados con respecto a la prioridad, acuerdo en la definición y factibilidad de las variables elegidas previamente por el panel de expertos.

Tabla XIII. *Prioridad de las variables elegidas por el panel de expertos.*

Variable	0	1	2	3	4	5	6	7	8	9	10	Total	Promedio ponderado
1	1,64 1	1,64 1	6,56 4	0,00 0	3,28 2	1,64 1	6,56 4	19,67 12	22,95 14	19,67 12	16,39 10	61	8,38
2	1,64 1	0,00 0	1,64 1	8,20 5	3,28 2	11,48 7	8,20 5	21,31 13	24,59 15	13,11 8	6,56 4	61	7,77
3	0,00 0	0,00 0	1,67 1	0,00 0	0,00 0	3,33 2	0,00 0	6,67 4	23,33 14	33,33 20	31,67 19	60	9,70
4	0,00 0	0,00 0	0,00 0	1,67 1	0,00 0	3,33 2	3,33 2	10,00 6	26,67 16	26,67 16	28,33 17	60	9,48
5	0,00 0	0,00 0	0,00 0	1,64 1	1,64 1	0,00 0	4,92 3	13,11 8	36,07 22	24,59 15	18,03 11	61	9,23
6	0,00 0	0,00 0	1,67 1	0,00 0	1,67 1	0,00 0	3,33 2	6,67 4	23,33 14	35,00 21	28,33 17	60	9,62
7	0,00 0	0,00 0	5,00 3	3,33 2	0,00 0	6,67 4	10,00 6	13,33 8	16,67 10	21,67 13	23,33 14	60	8,68
8	0,00 0	3,33 2	3,33 2	3,33 2	5,00 3	13,33 8	15,00 9	25,00 15	20,00 12	10,00 6	1,67 1	60	7,38
9	0,00 0	0,00 0	6,67 4	1,67 1	3,33 2	8,33 5	13,33 8	15,00 9	21,67 13	21,67 13	8,33 5	60	8,10
10	0,00 0	0,00 0	1,67 1	0,00 0	1,67 1	10,00 6	10,00 6	21,67 13	35,00 21	35,00 21	18,33 11	60	9,20
11	0,00 0	0,00 0	1,72 1	1,72 1	5,17 3	0,00 0	3,45 2	8,62 5	20,69 12	39,66 23	18,97 11	58	9,22
12	0,00 0	0,00 0	0,00 0	0,00 0	0,00 0	5,36 3	8,93 5	10,71 6	25,00 14	50,00 28	50,00 28	56	10,05
13	0,00 0	1,79 1	0,00 0	1,79 1	0,00 0	1,79 1	5,36 3	3,72 2	23,21 13	32,14 18	30,36 17	56	9,52
14	0,00 0	0,00 0	3,51 2	1,75 1	0,00 0	10,53 6	3,51 2	22,81 13	14,04 8	28,07 16	15,79 9	57	8,68
15	3,57 2	1,79 1	0,00 0	0,00 0	3,57 2	3,57 2	7,14 4	17,86 10	32,14 18	26,79 15	26,79 15	56	9,05
16	1,75 1	0,00 0	0,00 0	0,00 0	7,02 4	8,77 5	8,77 5	31,58 18	26,32 15	15,79 9	15,79 9	57	8,96
17	0,00 0	1,75 1	3,51 2	1,75 1	5,26 3	5,26 3	3,51 2	17,54 10	21,05 12	22,81 13	17,54 10	57	8,54

Tabla XIV. Acuerdo en la definición de las variables elegidas por el panel de expertos.

Variable % n	1	2	3	4	5	Total	Promedio ponderado
1	1,64 1	8,20 5	11,48 7	31,15 19	47,54 29	61	4,15
2	1,64 1	1,64 1	11,48 7	52,46 32	32,79 20	61	4,13
3	5,00 3	15,00 9	23,33 14	45,00 27	11,67 7	60	3,43
4	8,20 5	18,03 11	27,87 17	32,79 20	13,11 8	61	3,25
5	1,64 1	8,20 5	16,39 10	40,98 25	32,79 20	61	3,95
6	0,00 0	10,00 6	25,00 15	36,67 22	28,33 17	60	3,83
7	8,47 5	13,56 8	33,90 20	32,20 19	11,86 7	59	3,25
8	5,00 3	20,00 12	30,00 18	35,00 21	10,00 6	60	3,25
9	7,02 4	14,04 8	36,84 21	35,09 20	7,02 4	57	3,21
10	6,67 4	26,67 16	20,00 12	35,00 21	11,67 7	60	3,18
11	3,45 2	32,67 19	36,21 21	22,41 13	5,17 3	58	2,93
12	1,75 1	15,79 9	5,26 3	38,60 22	38,60 22	57	3,96
13	3,64 2	20,00 11	10,91 6	49,09 27	16,36 9	55	3,55
14	0,00 0	5,26 3	5,26 3	36,84 21	52,63 30	57	4,37
15	5,26 3	10,53 6	21,05 12	40,35 23	22,81 13	57	3,65
16	1,79 1	8,93 5	17,86 10	33,93 19	37,50 21	56	3,96
17	16,07 9	23,21 13	25,00 14	30,36 17	5,36 3	56	2,86

Abreviaturas: n: número.

Tabla XV. Factibilidad de las variables elegidas por el panel de expertos.

Variable % n	1	2	3	4	5	Total	Promedio ponderado
1	0,00 0	0,00 0	6,56 4	36,07 22	57,38 35	61	4,51
2	0,00 0	1,67 1	11,67 7	36,67 22	50,00 30	60	4,35
3	3,39 2	6,78 4	22,03 13	45,76 27	22,03 13	59	3,76
4	3,28 2	13,11 8	26,23 16	39,34 24	18,03 11	61	3,56
5	0,00 0	0,00 0	5,00 3	25,00 15	70,00 42	60	4,65
6	1,69 1	6,78 4	18,64 11	25,42 15	47,46 28	59	4,10
7	6,90 4	12,07 7	29,31 17	32,76 19	18,97 11	58	3,45
8	1,67 1	6,67 4	18,33 11	43,33 26	30,00 18	60	3,93
9	5,08 3	5,08 3	16,95 10	50,85 30	22,03 13	59	3,80
10	1,69 1	8,47 5	15,25 9	45,76 27	28,81 17	59	3,92
11	5,26 3	21,05 12	21,05 12	40,35 23	12,28 7	57	3,33
12	1,75 1	7,02 4	15,79 9	35,09 20	40,35 23	57	4,05
13	10,71 6	14,29 8	16,07 9	37,50 21	21,43 12	56	3,45
14	3,51 2	5,26 3	12,28 7	38,60 22	40,35 23	57	4,07
15	1,75 1	1,75 1	15,79 9	40,35 23	40,35 23	57	4,16
16	0,00 0	1,79 1	8,93 5	35,71 20	53,57 30	56	4,41
17	3,51 2	14,04 8	22,81 13	29,82 17	29,82 17	57	3,68

Abreviaturas: n: número.

Se alcanzó el suficiente grado de acuerdo propuesto como objetivo en cuanto a prioridad y factibilidad en todas las variables en primera ronda. No obstante, se decidió someter a una segunda ronda las definiciones de cada una de las variables con los cambios propuestos por los expertos encuestados. En la tabla XVI se describe la definición final alcanzada de cada una de las variables tras la segunda votación. Según

el criterio de los expertos, los ítems propuestos se redujeron a 15 después de unificar las variables *suspensión de metotrexato por ineficacia, ausencia de respuesta al metotrexato y número de fármacos considerados ineficaces* dentro la misma variable denominada *refractariedad*.

Tabla.XVI. Definiciones y variables elegidas por el panel de expertos y posterior Delphi.

1. **Poliartritis**, definida como la presencia de ≥ 4 articulaciones tumefactas simultáneamente en más del 75% de las visitas durante los 2 primeros años de seguimiento objetivadas por un reumatólogo.
2. Afectación de **grandes articulaciones**, considerando como tales cualquier de las siguientes: hombros, caderas, rodillas, codos o tobillos.
3. **Actividad elevada a los dos años de seguimiento**, definida por un índice de actividad DAS28-VSG $> 3,2$ ó DAS-PCR $> 2,95$.
4. **Actividad mantenida durante los dos años de seguimiento**, definida como una actividad elevada en $>75\%$ de las visitas, medida mediante cualquier índice de actividad (DAS28-VSG ó DAS28-PCR) en los dos primeros años.
5. **Reactantes de fase aguda elevados de forma mantenida**: elevación de VSG y/o PCR por encima del valor de referencia del laboratorio (VSG: <15 en varones y <25 en mujeres; PCR $<0,8$) y en $>75\%$ de las visitas de los dos primeros años de seguimiento.
6. **Manifestaciones extraarticulares**: presencia de cualquiera de las siguientes manifestaciones sistémicas: afectación pulmonar, vasculitis, amiloidosis, Síndrome de Felty, y serositis.
7. **Imposibilidad para conseguir la remisión**: incapacidad para alcanzar los criterios EULAR de remisión booleana y una valoración global de la enfermedad por el médico ≥ 20 (VGEM de 0-100) en $>75\%$ de las visitas (4 de 5) de los 2 años de seguimiento.
8. **Refractariedad**: actividad elevada (DAS28-VSG $>3,2$ ó DAS-PCR $>2,95$) tras 6 meses de tratamiento con metotrexato (actividad elevada en cada visita con MTX a dosis superior a 15) o leflunomida en la visita previa.
9. **Tratamiento intensivo a lo largo de los 2 años**: necesidad de utilizar tratamiento con cualquier terapia combinada (metotrexato/leflunomida), cualquier combinación con fármacos biológicos o triple terapia (metotrexato/sulfasalazina/hidroxicloroquina) en la visita de los 2 años de seguimiento.
10. **Erosiones en el momento del diagnóstico**: presencia cualitativa de ≥ 1 erosión objetivada en el momento del diagnóstico (suma de erosiones en manos derecha e izquierda).
11. **Nuevas erosiones a los 2 años del diagnóstico**: aumento del número de erosiones a los 2 años del seguimiento con respecto a la visita basal (diferencias entre el número de erosiones entre los dos años de seguimiento y la basal).
12. Ausencia de mejoría de al menos 0,5 entre el **HAQ** basal y a los 2 años de seguimiento.
13. **Necesidad de cirugía ortopédica/prótesis**: realización de cualquier tratamiento quirúrgico/ortopédico de estabilización o reemplazo articular, con excepción de la sinovectomía, en los 2 primeros años de seguimiento.
14. **Ingresos hospitalarios** relacionados con la actividad inflamatoria de la enfermedad.
15. **Corticodependencia**: imposibilidad de suspender los corticoides en los 2 primeros años de seguimiento, independientemente de la dosis (cualquier dosis de corticoides en visita de 12 y 24 meses)

Abreviaturas: DAS28: Disease Activity Score with 28 joint counts; VSG: velocidad sedimentación globular; PCR: proteína C reactiva; VGEM: valoración global de la enfermedad por el médico; MTX: metotrexato; HAQ: Health Assessment Questionnaire.

4. Conclusiones

- Se alcanzó el suficiente grado de acuerdo para incluir dentro del constructo gravedad todas las variables propuestas por el panel de expertos en primera ronda.
- Tras una segunda ronda, las variables quedaron reducidas a 15, con una definición consensuada también por el panel: poliartritis, afectación de grandes articulaciones, actividad elevada a dos años de seguimiento, actividad mantenida durante los dos años de seguimiento, reactantes de fase aguda elevados de forma mantenida, manifestaciones extraarticulares, imposibilidad para conseguir la remisión, refractariedad, tratamiento intensivo a lo largo de los dos años, erosiones en el momento del diagnóstico, nuevas erosiones a los dos años del diagnóstico, HAQ, necesidad de cirugía ortopédica/prótesis, ingresos hospitalarios y corticodependencia.
- Estas variables se utilizaron para el diseño y construcción de un índice compuesto de evaluación de la gravedad.

Sub-estudio 4. Diseño y construcción de un índice compuesto de evaluación de la gravedad

1.	Introducción.....	67
2.	Descriptivo de la visita basal.....	68
3.	Definiciones y construcción del índice.....	71
4.	Análisis de los ítems.....	73
4.1	Población y datos.....	73
4.2	Comportamiento de los ítems.....	73
4.3	Estructura y dimensionalidad: Análisis factorial.....	73
4.4	Fiabilidad: consistencia interna.....	75
4.5	Capacidad de discriminación.....	76
5.	Resultados.....	78
5.1	Análisis de los ítems.....	78
5.2	Estructura y dimensionalidad.....	79
5.2.1	Análisis factorial A: índice completo.....	79
5.2.1.1	Número de factores a extraer.....	79
5.2.1.2	Rotación e interpretación de los factores.....	80
5.2.1.3	Correlaciones ítem-total y consistencia interna.....	81
5.2.1.4	Optimización de la dimensionalidad.....	82
5.2.2	Análisis factorial B: eliminación del ítem 5 (reactantes fase aguda).....	82
5.2.3	Análisis factorial C: eliminación del ítem 14 (corticodependencia).....	84
5.2.4	Análisis factorial D: eliminación del ítem 12 (cirugía ortopédica).....	87
5.2.5	Análisis factorial E: eliminación del ítem 13 (ingresos).....	89
5.2.6	Análisis factorial F: eliminación del ítem 7 (imposibilidad remisión).....	91
5.2.7	Análisis factorial G: eliminación ítem 2 (grandes articulaciones).....	92
5.2.8	Análisis factorial H: eliminación del ítem 1 (poliartritis).....	94
5.2.9	Análisis factorial I: eliminación del ítem 11 (ausencia de mejoría HAQ) ..	95
6.	Capacidad de discriminación.....	98
7.	Conclusión.....	100

SUB-ESTUDIO 4. DISEÑO Y CONSTRUCCIÓN DE UN ÍNDICE COMPUESTO DE EVALUACIÓN DE LA GRAVEDAD

1. Introducción

El objetivo general de este proyecto es desarrollar un índice de gravedad de la artritis reumatoide (AR), a los 2 años de seguimiento, que refleje adecuadamente el marco teórico de este constructo y que pueda obtenerse en la práctica clínica habitual a partir de los datos existentes en la historia clínica.

La primera fase del proyecto ha consistido en una revisión sistemática de la literatura para identificar posibles índices similares, realización de grupos de expertos y estudio Delphi. La información recogida en esta fase se transformó en una serie de definiciones e ítems a incluir en el índice. El objetivo de esta segunda fase es construir el índice, a partir de estas definiciones e ítems, y analizar sus propiedades psicométricas.

La muestra de estudio es la cohorte PEARL de artritis de reciente comienzo del Hospital de La Princesa. Esta cohorte incluye todos los pacientes con al menos 1 articulación tumefacta durante un periodo menor de 1 año, y diagnosticados de AR (criterios revisados ACR 1987)([47](#)) o artritis inflamatoria indiferenciada([48](#)). El protocolo de seguimiento clínico está formado por 4 visitas durante los 2 primeros años desde el diagnóstico (basal, 6, 12 y 24 meses).

El análisis de esta segunda fase se llevará a cabo mediante los siguientes procesos:

- Descriptivo de la visita basal.
- Definiciones y construcción del índice.
- Análisis de los ítems.
- Estructura y dimensionalidad: Análisis factorial.
- Fiabilidad: consistencia interna.

- Capacidad de discriminación entre grupos con diferente gravedad: Análisis discriminante.

2. Descriptivo de la visita basal

En el momento de realizar el estudio, la cohorte PEARL estaba formada por 683 pacientes, mayoritariamente mujeres (78%), de alrededor de 50 años, raza caucásica (80%) y ámbito residencial urbano (57%). Los criterios diagnósticos de AR más frecuentes fueron la artritis de manos (90%), afectación de más de 3 articulaciones (73%), con rigidez matutina (65%) y distribución simétrica (63%). El 43% presentaba factor reumatoide positivo y el 10% tenía antecedentes familiares. La mayoría de los pacientes mostraba algún nivel de actividad (baja, moderada, alta), en porcentajes variables según se utilizara DAS28 o DAS28-PCR. Por otra parte, los criterios de remisión ACR-EULAR sólo se cumplían en un 10% y 7% de los casos según SDAI y criterios booleanos, respectivamente(46) (tabla XVII).

El tiempo de evolución promedio era de $6,7 \pm 5,2$ meses. La valoración global de la enfermedad, por parte de paciente y médico, alcanzaba valores medios de $40,7 \pm 23,4$ y $35,1 \pm 23,6$, respectivamente; con EVA de dolor de $41,8 \pm 2,6$; recuentos de articulaciones dolorosas y tumefactas de $5,3 \pm 6,2$ y $4,6 \pm 4,6$, correspondientemente, rigidez matutina durante $41,4 \pm 52,0$ minutos, VSG y PCR de $28,9 \pm 21,9$ mm/1ª hora y $1,6 \pm 3,2$ mg/dl; valores DAS28 de $4,2 \pm 1,5$ y de DAS28-PCR de $3,8 \pm 1,4$, con SDAI de $19,7 \pm 16,2$; CDAI de $6,6 \pm 5,3$ e índice de Sharp de $3,8 \pm 6,9$ (tabla XVIII). En relación con el tratamiento, los fármacos más utilizados eran metotrexato (9%), a dosis media de $12,5 \pm 4$ mg/sem, y sulfasalazina (1.714 ± 393 mg/d). Los FAME se usaban habitualmente en monoterapia (11,4%) (tabla XIX).

De los 683 pacientes que realizaron la visita basal, 474 llegaron a la de los 6 meses, 456 a la de los 12 meses y 466 a la de los 2 años. Por consiguiente, durante el seguimiento se perdieron 217 casos, lo que supone el 31,7%. Estos 466 pacientes constituyen la muestra de análisis para la construcción y validación del índice de gravedad.

Tabla XVII. Características clínicas basales: Variables cualitativas.

Variable	N (%)
Sexo mujer (n=683)	538 (78,8%)
Raza caucásica (n=683)	544 (79,6%)
Lugar de residencia urbano (n=682)	646 (57,4%)
Criterios diagnósticos de AR (n=683)	
• Artritis de manos	614 (89,9%)
• Radiología compatible	84 (12,3%)
• Rigidez matutina	447 (65,4%)
• Artritis en ≥ 3 articulaciones	496 (72,6%)
• Nódulos reumatoides	19 (2,8%)
• Artritis simétrica	431 (63,1%)
Factor reumatoide (n=683)	295 (43,2%)
Antecedentes familiares de AR (n=683)	67 (9,8%)
Estado civil (n=681)	
• Casado	392 (57,6%)
• Separado	54 (7,9%)
• Soltero	164 (24,1%)
• Viudo	71 (10,4%)
Situación laboral (n=683)	
• Activo	387 (56,7%)
• Ama de casa	111 (16,2%)
• Estudiante	2 (0,3%)
• Jubilado	142 (20,8%)
• En el paro	23 (3,4%)
• Desconocido	18 (2,6%)
Incapacidad laboral temporal (n=683)	71 (10,4%)
Incapacidad laboral permanente (n=683)	7 (1,0%)
Actividad según DAS28 (n=634)	
• Remisión	98 (15,5%)
• Baja	71 (11,2%)
• Moderada	277 (43,7%)
• Alta	188 (29,6%)
Actividad según DAS28-PCR (n=649)	
• Remisión	113 (17,4%)
• Baja	231 (35,6%)
• Moderada	153 (23,6%)
• Alta	152 (23,4%)
Remisión ACR-EULAR: SDAI (n=649)	67 (10,3%)
Remisión ACR-EULAR: Booleano (n=649)	44 (6,8%)

Abreviaturas: n: número; AR; artritis reumatoide; DAS28: Disease Activity Score with 28 joint counts; PCR: proteína C reactiva.

Tabla XVIII. Características clínicas basales: Variables cuantitativas.

Variable	N	Media ± DE	Mediana (RIQ)
Peso	629	67,2±14,1	65,0 (58,0-75,7)
Talla	630	1,60±0,2	1,6 (1,5-1,6)
Índice de masa corporal	629	26,5±4,9	26,1 (23,3-29,1)
Edad al inicio de la enfermedad	679	53,6±16,5	53,6 (41,1-65,9)
Edad en la primera visita	683	54,2±16,4	54,0 (41,9-66,1)
Valoración global de la enfermedad (paciente)	679	40,7±23,4	42,0 (23,0-56,0)
Valoración global de la enfermedad (médico)	683	35,1±23,6	30,0 (16,0-50,0)
Escala visual analógica (dolor)	678	41,8±26,0	43,5 (20,0-60,0)
Health Assessment Questionnaire (HAQ)	675	1,0±0,7	0,9 (0,4-1,5)
Número de articulaciones dolorosas	683	5,3±6,2	3,0 (0-8,0)
Número de articulaciones tumefactas	683	4,6±4,6	3,0 (1,0-7,0)
Duración de la rigidez matutina (minutos)	677	41,4±52,0	20,0 (0-60,0)
Velocidad de sedimentación globular	641	28,9±21,9	22,0 (13,0-37,0)
Proteína C reactiva	653	1,6±3,2	0,5 (0,2-1,4)
Hemoglobina (mg/dl)	666	13,4±1,4	13,4 (12,6-14,3)
DAS-28	634	4,2±1,5	4,1 (3,1-5,4)
DAS28-PCR	649	3,8±1,4	3,7 (2,7-4,8)
DAS-28 (3 variables)	638	4,1±1,4	4,0 (3,1-5,2)
DAS28-PCR (3 variables)	653	3,6±1,4	3,5 (2,5-4,5)
SDAI	649	19,7±16,2	15,1 (7,0-27,4)
CDAI	683	6,6±5,3	5,2 (2,8-8,5)
Índice de Sharp	251	3,8±6,9	1,0 (0,0-4,0)
Tiempo de evolución (meses)	466	6,7±5,2	5,4 (3,1-8,5)

Abreviaturas: DE: desviación estándar; RIQ: rango intercuartil; HAQ: Health Assessment Questionnaire; DAS28: Disease Activity Score with 28 joint counts; PCR: proteína C reactiva; SDAI: Simple Disease Activity Index; CDAI: Clinical Disease Activity Index.

Tabla XIX. Características clínicas basales: Tratamientos.

Fármaco	N (%)	
Metotrexato (n=683)	64 (9,4%)	
Leflunomida (n=683)	5 (0,7%)	
Antipalúdicos (n=683)	23 (3,4%)	
Sulfasalazina (n=683)	7 (1,0%)	
Sales de oro (n=683)	1 (01%)	
Ciclosporina (n=683)	-	
FAME (n=683)		
• Monoterapia	78 (11,4%)	
• Terapia combinada	11 (1,6%)	
Anti-TNF (n=683)	-	
Dosis	Media ± DE	Mediana (RIQ)
Corticoides (mg/d)	2,7±6,9	0 (0-0)
Metotrexato (n=64)	12,5±4,0	10,0 (10,0-15,0)
Leflunomida (n=5)	16,0±5,5	20,0 (10,0-20,0)
Antipalúdicos (n=23)	282,6±114,4	200 (200-400)
Sulfasalazina (n=7)	1714±393	2000 (1500-2000)
Sales de oro (n=1)	200	200

Abreviaturas: N: número; FAME: fármaco modificador de la enfermedad; anti-TNF: anti Tumoral Necrosis Factor; DE: desviación estándar; RIQ: rango intercuartil; d: día.

3. Definiciones y construcción del índice

Las definiciones de los ítems a incluir en el índice se obtuvieron de los resultados del estudio Delphi. Uno de los ítems seleccionados fue el de **Tratamiento intensivo a lo largo de los 2 años**, definido como *“la necesidad de utilizar tratamiento con cualquier terapia combinada (metotrexato/leflunomida), cualquier combinación con fármacos biológicos, o triple terapia (metotrexato, sulfasalazina, hidroxicloroquina) en la visita de los 2 años de seguimiento”*. A pesar de haber sido seleccionado en el estudio Delphi, se decidió no incluirlo en el índice inicial debido a que esta variable mide una actitud del médico y no una característica del paciente.

Por consiguiente, se creó un índice de 14 ítems de respuesta dicotómica (si/no) (tabla XX).

Tabla XX. Definiciones e ítems

1. **Poliartritis**, definida como la presencia de ≥ 4 articulaciones tumefactas simultáneamente en más del 75% de las visitas durante los 2 primeros años de seguimiento objetivadas por un reumatólogo.
2. Afectación de **grandes articulaciones**, considerando como tales cualquiera de las siguientes: hombros, caderas, rodillas, codos o tobillos.
3. **Actividad elevada a los dos años de seguimiento**, definida por un índice de actividad DAS28-VSG $>3,2$ ó DAS-PCR $>2,95$.
4. **Actividad mantenida durante los dos años de seguimiento**, definida como una actividad elevada en $>75\%$ de las visitas, medida mediante cualquier índice de actividad (DAS28-VSG ó DAS28-PCR) en los dos primeros años.
5. **Reactantes de fase aguda elevados de forma mantenida**: elevación de VSG y/o PCR por encima del valor de referencia del laboratorio (VSG: <15 en varones y <25 en mujeres; PCR $<0,8$) y en $>75\%$ de las visitas de los dos primeros años de seguimiento.
6. **Manifestaciones extraarticulares**: presencia de cualquiera de las siguientes manifestaciones sistémicas: afectación pulmonar, vasculitis, amiloidosis, Síndrome de Felty, y serositis.
7. **Imposibilidad para conseguir la remisión**: incapacidad para alcanzar los criterios EULAR de remisión booleana y una valoración global de la enfermedad por el médico ≥ 20 (VGEM de 0-100) en $>75\%$ de las visitas de los 2 años de seguimiento.
8. **Refractariedad**: actividad elevada (DAS28-VSG $>3,2$ ó DAS-PCR $>2,95$) tras 6 meses de tratamiento con metotrexato (actividad elevada en cada visita con MTX a dosis superior a 15) o leflunomida en la visita previa.
9. **Erosiones en el momento del diagnóstico**: presencia cualitativa de ≥ 1 erosión objetivada en el momento del diagnóstico (suma de erosiones en manos derecha e izquierda).
10. **Nuevas erosiones a los 2 años del diagnóstico**: aumento del número de erosiones a los 2 años del seguimiento con respecto a la visita basal (diferencias entre el número de erosiones entre los dos años de seguimiento y la basal).
11. Ausencia de mejoría de al menos 0,5 entre el **HAQ** basal y a los 2 años de seguimiento.
12. **Necesidad de cirugía ortopédica/prótesis**: realización de cualquier tratamiento quirúrgico/ortopédico de estabilización o reemplazo articular, con excepción de la sinovectomía, en los 2 primeros años de seguimiento.
13. **Ingresos hospitalarios** relacionados con la actividad inflamatoria de la enfermedad.
14. **Corticodependencia**: imposibilidad de suspender los corticoides en los 2 primeros años de seguimiento, independientemente de la dosis (cualquier dosis de corticoides en visita de 12 y 24 meses)

Abreviaturas: DAS28: Disease Activity Score with 28 joint counts; VSG: velocidad sedimentación globular; PCR: proteína C reactiva; VGEM: valoración global de la enfermedad por el médico; MTX: metotrexato; HAQ: Health Assessment Questionnaire.

4. Análisis de los ítems

Una vez diseñado el índice se utilizó una metodología con pasos sucesivos para la evaluación de su comportamiento y el análisis de la estructura y dimensionalidad, con el fin de establecer la selección definitiva de ítems que garantizara una adecuada capacidad de discriminación entre pacientes con distintos niveles de gravedad.

4.1 Población y datos

Los 14 ítems seleccionados se recogieron en 466 pacientes de artritis de inicio, mediante revisión de historias clínicas. Sin embargo, en la muestra de análisis sólo se incluyeron aquellos casos con información completa en todos los ítems del índice, ya que la metodología a utilizar no permite la inclusión de valores perdidos ni la imputación de datos.

4.2 Comportamiento de los ítems

Se analizaron las puntuaciones de los ítems individuales, así como el número y patrón de valores perdidos. Con el fin de descartar posibles sesgos de selección relacionados con la pérdida de información se compararon los pacientes, con y sin valores perdidos, en la visita basal.

4.3 Estructura y dimensionalidad: Análisis factorial

Se exploró la estructura y dimensiones del índice mediante análisis factorial (AF). El AF es una técnica multivariante cuyo objetivo es definir la estructura subyacente de una matriz de datos. Es el método más utilizado para examinar la dimensionalidad de los ítems, ya que analiza la estructura de las interrelaciones (correlaciones) existentes en un elevado número de variables (ítems de un cuestionario), definiendo un número de dimensiones subyacentes denominadas factores. El principio básico es que los ítems que muestran elevada correlación entre ellos se agrupan en un único factor, mientras que los que integran un determinado factor muestran baja correlación con los que

pertencen a otros factores. El objetivo del análisis factorial es cuantificar e identificar las dimensiones subyacentes a un constructo (gravedad en este caso), y tratar de explicar la mayor parte de la varianza con el mínimo número de factores posible; es decir, buscar y definir los constructos o dimensiones fundamentales latentes a partir de una matriz de correlaciones. Por otra parte, el AF también sirve para la reducción de datos, ya que permite eliminar los ítems que no tienen ninguna contribución a los factores. De esta forma la información contenida en un número original de variables se condensa en una cantidad menor de nuevas dimensiones (factores) con una pérdida de información mínima. La perspectiva exploratoria del AF es la más utilizada como método de reducción de datos.

Al encontrarnos en la fase de diseño del índice se utilizó un AF exploratorio, puesto que no se conocen las posibles dimensiones subyacentes en los ítems seleccionados para medir la gravedad de la AR. La contribución de cada variable a cada factor se denomina carga factorial. El método de extracción de las cargas factoriales y de obtención de la mayor cantidad de varianza explicada fue el de componentes principales.

La rotación de los factores es una herramienta que facilita su interpretación cuando el número de factores es superior a 1. La rotación implica un cambio de los ejes de referencia sobre el origen hasta alcanzar otra posición. El efecto de la rotación es redistribuir la varianza para conseguir un patrón de factores con mayor significado. El criterio varimax permite obtener unas cargas más extremas (más cercanas al -1 y al +1) y otras más cercanas al 0. El resultado es una interpretación más sencilla de los factores al indicar, más claramente, una asociación positiva o negativa entre la variable y el factor. Por consiguiente, se utilizó el método de rotación varimax, una vez comprobado que los factores extraídos eran ortogonales; es decir, independientes.

No existe una prueba estadística para la inclusión de variables en los factores identificados tras el AF. No obstante, habitualmente se utiliza una carga factorial

mínima de 0,5 ó superior, junto con una nítida separación entre factores. Por consiguiente, para la exclusión de los ítems se emplearon los siguientes criterios:

- Ítems que no cargaban en ningún factor (correlaciones inferiores a 0,5 tras rotación).
- Baja comunalidad (proporción de la variable explicada por los factores).
- En el caso de solapamiento, o presencia de cargas muy similares y superiores a 0,5 en varios factores, la variable a incluir en el factor fue la más relacionada con el factor desde el punto de vista conceptual.
- Ítems con correlaciones ítem-total de 0,3 o inferiores.

En el caso de variables dicotómicas, como las que nos ocupan, el análisis factorial se realiza sobre una matriz de correlaciones tetracóricas. Estas correlaciones se basan en la asunción de una serie de variables inobservables o latentes, de naturaleza continua, sobre las que se construyen las variables observables dicotómicas (ítems). De acuerdo con este supuesto, las opciones de respuesta de las observables establecerían un tipo de umbral sobre el continuo de las latentes. De este modo, los dos valores de las variables dicotómicas simbolizarían ambos polos del continuo.

4.4 Fiabilidad: consistencia interna

La consistencia interna se define como el grado de interrelación entre los ítems. Es una medida del grado con el que los ítems valoran un mismo constructo. Es evidente que los componentes de un mismo factor tendrán una elevada correlación entre ellos. Además de evaluar la homogeneidad del índice, el análisis de la consistencia interna también es una técnica de reducción de ítems.

- Correlaciones inter-ítem e ítem-total: Las correlaciones entre los ítems e ítem-total indican si el ítem forma parte de la escala o no. Los ítems incluidos en una misma dimensión o factor deben mostrar una correlación inter-ítem entre 0,2 y 0,5. Este rango es bastante amplio, pero depende fundamentalmente de la amplitud del constructo a medir. En caso de constructos amplios es posible que

las correlaciones inter-ítem sean más bajas que en constructos más delimitados. Por otro lado, si la correlación entre dos ítems es elevada (por encima de 0,7 ó 0,8) podemos suponer que ambos están midiendo prácticamente lo mismo, por lo que podría eliminarse uno de ellos. La correlación ítem-total es un parámetro de discriminación ya que indica si los ítems pueden discriminar pacientes con diferentes magnitudes del constructo en estudio. En general, se considera que los ítems con una correlación ítem-total inferior a 0,3 no tienen capacidad de discriminación para diferenciar entre pacientes con mayor o menor nivel de gravedad, por lo que pueden eliminarse. Sin embargo, es importante retener los ítems que se consideran importantes aunque sus propiedades psicométricas no sean las mejores.

- Consistencia interna. El alfa de Cronbach es el parámetro utilizado para valorar la consistencia interna de una escala. En el caso de variables dicotómicas el equivalente del alfa es el estadístico de Kuder-Richardson. En test con respuestas 1 ó 0 (respuesta correcta o incorrecta) el coeficiente utilizado es el de Kuder-Richardson 20. Con test muy breves (entre 10 y 15 ítems) una fiabilidad de 0,50 puede considerarse suficiente, pero para tomar decisiones individuales sobre los sujetos la fiabilidad debería ser mucho mayor (en torno a 0,8). Otro factor del que depende la fiabilidad es de las diferencias reales entre los sujetos; si son muy homogéneos la fiabilidad tiende a bajar, sin que ello signifique que la prueba sea mala o produzca malos resultados(64). Los ítems que presentan una correlación ítem-total inferior a 0,3 pueden no añadir poder explicativo al cuestionario. Se suele presentar la consistencia interna de cada dominio, en lugar de la global.

4.5 Capacidad de discriminación

El objetivo de esta fase del análisis es evaluar la capacidad para discriminar entre pacientes con gravedad baja y elevada. Para ello se utilizará un análisis discriminante seguido de un análisis de regresión lineal.

El análisis discriminante es una técnica estadística multivariante cuya finalidad es fundamentalmente clasificatoria. A partir de una variable dependiente cualitativa y un conjunto de una o más variables independientes cuantitativas, el AD permite clasificar a los individuos o casos en alguno de los grupos establecidos por la variable dependiente. Lo específico de esta técnica es que los grupos ya estén constituidos de antemano debiendo identificar lo específico de cada uno de ellos para poder asignar a los individuos o casos a cada uno de ellos. En esta finalidad clasificatoria va implícita, pues, una finalidad descriptiva. Las variables que más y mejor caracterizan y diferencian a los grupos se denominan variables canónicas o discriminantes, se presentan como combinaciones lineales de las variables originales y se expresan por una función discriminante.

En nuestro caso, el objetivo del análisis discriminante es determinar hasta qué grado contribuye cada factor a la discriminación entre los dos grupos con diferente nivel de gravedad. Para ello se utilizará como variable dependiente o de clasificación el número de días de tratamiento con FAME, ponderado por tipo de fármaco, denominada TFME (IDT o *intensity of DMARD treatment* en el sub-estudio 1) y como variables independientes la combinación de las cargas factoriales de cada factor multiplicadas por cada uno de los ítems respectivos.

A partir de los resultados del AD se calculará la función discriminante para cada paciente de la muestra multiplicando los coeficientes no estandarizados por el valor de cada uno de los factores. Esta función discriminante será la variable dependiente en un análisis de regresión lineal múltiple con selección por pasos hacia delante de los ítems extraídos en el análisis factorial. El objetivo es determinar qué ítems explican la mayor parte de la función discriminante. En el índice final sólo se incluirán aquellos ítems seleccionados por el procedimiento por pasos que expliquen, en su conjunto, más del 95% de la varianza de la variable dependiente y contribuyan significativamente a la misma. Las ponderaciones de los ítems se obtendrán a partir de los coeficientes de regresión del modelo lineal final con las variables previamente seleccionadas.

5. Resultados

5.1 Análisis de los ítems

De los 466 pacientes integrantes de la cohorte de artritis de inicio, sólo 156 mostraban información completa en los 14 ítems seleccionados inicialmente a partir de los resultados del estudio Delphi. Estos 156 constituyen la muestra de análisis. La comparación de pacientes con y sin información completa en el índice demostró la ausencia de diferencias por sexo, edad de inicio, reactantes de fase aguda (VSG y PCR), e índices de actividad (DAS28PCR, SDAI y CDAI). Por el contrario, el tiempo de evolución de la enfermedad al inicio de seguimiento fue ligeramente mayor en los pacientes con información completa que en los perdidos (7,3 vs 6,4 meses; $p=0,039$).

En la tabla XXI se muestra la distribución de los 14 ítems, inicialmente seleccionados, junto con sus valores perdidos. La mayor pérdida de información se produjo en ítems relacionados con la realización de radiología [erosiones al diagnóstico (47% perdidos) y aumento del número de erosiones a los dos años (52%)], afectación de grandes articulaciones (33%), ingresos por actividad inflamatoria (21%), y corticodependencia (12%).

Tabla XXI. Distribución de las puntuaciones y patrón de valores perdidos.

Item	Definición	Valores perdidos N (%)	Puntuaciones No-Sí
1	Poliartritis en más del 75% de las visitas	-	90%-10%
2	Afectación de grandes articulaciones	33%	23%-44%
3	Actividad elevada a los dos años de seguimiento	4%	59%-36%
4	Actividad mantenida durante los dos años de seguimiento	-	58%-42%
5	Elevación mantenida de reactantes de fase aguda	-	67%-33%
6	Manifestaciones extra-articulares	4%	94%-2%
7	Imposibilidad para conseguir la remisión	-	70%-30%
8	Refractariedad	-	74%-26%
9	Erosiones en el momento del diagnóstico	47%	38%-15%
10	Nuevas erosiones a los 2 años del diagnóstico	52%	37%-11%
11	Ausencia de mejoría en el HAQ los 2 primeros años	2%	43%-55%
12	Necesidad de cirugía ortopédica en los 2 primeros años	4%	95%-0,4
13	Ingresos por actividad inflamatoria	21%	74%-5%
14	Corticoddependencia en los 2 primeros años	12%	72%-17%

Abreviaturas: HAQ: Health Assessment Questionnaire.

5.2 Estructura y dimensionalidad.

5.2.1 Análisis factorial A: índice completo

Los resultados del análisis factorial se presentan los siguientes apartados:

5.2.1.1 *Número de factores a extraer*

La extracción de los factores subyacentes a los datos se realizó mediante el criterio del valor propio de Kaiser (varianza explicada por cada factor superior a 1). En la tabla XXII se presentan todos los factores, junto con sus valores propios y la proporción de varianza explicada, tanto por cada factor, como acumulada. El número de factores con autovalor superior a 1 es de 5 y entre ellos explican un total del 82,1% de la varianza total.

Tabla XXII. Factorial A: factores, autovalores y varianza explicada.

Factor	Autovalor	Varianza explicada (%)	Varianza acumulada (%)
1	5,541	39,6	39,6
2	2,018	14,4	54,0
3	1,606	11,5	65,5
4	1,310	9,4	74,8
5	1,016	7,3	82,1
6	0,870	6,2	88,3
7	0,518	3,7	92,0
8	0,473	3,4	95,4
9	0,293	2,1	97,5
10	0,258	1,8	99,3
11	0,057	0,4	99,7
12	0,040	0,3	100
13	0,000	0,00	100
14	-0,000	-0,00	100

5.2.1.2 Rotación e interpretación de los factores

Al existir 5 factores, y tras comprobar que son ortogonales mediante el análisis de sus correlaciones, se llevó a cabo una rotación ortogonal varimax con el fin de encontrar soluciones más fáciles de interpretar. La rotación mejora el significado y facilita la interpretación de los factores extraídos al maximizar las cargas de las variables con mayor asociación con un factor y minimizar las que tienen una asociación más débil. La varianza total explicada es la misma, pero el porcentaje explicado por cada factor puede cambiar. En la tabla XXIII se presentan las cargas factoriales tras rotación varimax y la comunalidad o proporción de varianza de cada ítem explicada por la solución factorial. En el primer factor cargan los ítems de actividad elevada a los dos años ($r=0,897$), y mantenida a lo largo de los 2 años ($r=0,946$), imposibilidad de conseguir la remisión ($r=0,776$), refractariedad ($r=0,726$), y cirugía ortopédica ($r=0,592$); en el segundo factor cargan las erosiones al diagnóstico ($r=0,822$), su aumento a los dos años ($r=0,811$), y los ingresos ($r=0,628$); en el tercero se agrupan la poliartritis ($r=0,742$) y la falta de mejoría del HAQ ($r=-0,921$). En el cuarto la afectación

de grandes articulaciones ($r=0,684$) y las manifestaciones extra-articulares ($r=0,894$); y finalmente, en el quinto factor cargan la elevación de reactantes de fase aguda ($r=0,586$), y la corticodependencia ($r=0,820$). El ítem con menor comunalidad (55%) es el 5, referido a los reactantes de fase aguda.

Tabla XXIII. Factorial A: cargas factoriales tras rotación varimax.

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Comunalidad
1 Poliartritis	0,553		0,742			95,9%
2 Grandes articulaciones				0,684		83,9%
3 Actividad a los 2 años	0,897					88,5%
4 Actividad mantenida	0,946					93,8%
5 Reactantes de fase aguda					0,586	54,8%
6 Manifestaciones extraarticulares				0,894		91,9%
7 Imposibilidad remisión	0,776					76,8%
8 Refractariedad	0,726					68,1%
9 Erosiones al diagnóstico		0,822				77,5%
10 Nuevas erosiones a los 2 años		0,811				69,5%
11 Ausencia de mejoría HAQ			-0,921			92,3%
12 Cirugía ortopédica	0,592	0,526				97,1%
13 Ingresos (actividad/inflamación)		0,628		0,531		74,8%
14 Corticodependencia					0,820	84,3%

5.2.1.3 Correlaciones ítem-total y consistencia interna

El primer factor tiene una elevada consistencia interna ($KR=0,737$), que va disminuyendo en los factores subsiguientes, formados también por un número menor de ítems.

Tabla XXIV. Factorial A: correlación ítem-total y Consistencia interna.

	Correlación ítem-total
FACTOR 1	
3 Actividad a los 2 años	0,618
4 Actividad mantenida	0,723
7 Imposibilidad remisión	0,499
8 Refractariedad	0,489
12 Cirugía ortopédica	0,128
Consistencia interna KUDER RICHARDSON	0,737
FACTOR 2	
9 Erosiones al diagnóstico	0,376
10 Nuevas erosiones a los 2 años	0,300
13 Ingresos (actividad/inflamación)	0,243
Consistencia interna KUDER RICHARDSON	0,478
FACTOR 3	
1 Poliartritis	-0,250
11 Ausencia de mejoría HAQ	-0,250
Consistencia interna KUDER RICHARDSON	-0,441
FACTOR 4	
2 Grandes articulaciones	0,121
6 Manifestaciones extraarticulares	0,121
Consistencia interna KUDER RICHARDSON	0,147
FACTOR 5	
5 Reactantes de fase aguda	0,293
14 Corticodependencia	0,293
Consistencia interna KUDER RICHARDSON	0,462

Abreviaturas: HAQ: Health Assessment Questionnaire.

5.2.1.4 Optimización de la dimensionalidad

Para la reducción de ítems se plantea, en primer lugar, eliminar el ítem 5 sobre Reactantes de fase aguda, ya que tiene baja carga factorial y comunalidad y poca capacidad de discriminación.

5.2.2 Análisis factorial B: eliminación del ítem 5 (reactantes fase aguda)

La eliminación de los reactantes de fase aguda produce una solución factorial de cuatro factores que explica el 78% de la varianza (tabla XXV). En el primero de ellos

cargan los ítems de actividad, en el segundo la alteración estructural junto con los ingresos y la cirugía ortopédica, en el tercero la poliartritis y la ausencia de mejoría en el HAQ y en el cuarto la afectación de grandes articulaciones y las manifestaciones extra-articulares. La corticodependencia no carga en ningún factor y tiene una baja comunalidad (tabla XXVI).

La consistencia interna de los factores es de 0,781; 0,461; -0,441; y 0,146, respectivamente (tabla XXVII).

Estos resultados plantean la necesidad de eliminar el ítem 14 sobre corticodependencia, ya que no carga en ningún factor y presenta baja comunalidad (52%).

Tabla XXV. Factorial B: factores, autovalores y varianza explicada.

Factor	Autovalor	Varianza explicada (%)	Varianza acumulada (%)
1	5,253	40,4%	40,4%
2	1,982	15,2%	55,6%
3	1,602	12,3%	68,0%
4	1,306	10,0%	78,0%
5	0,974	7,5%	85,5%
6	0,741	5,7%	91,2%
7	0,482	3,7%	94,9%
8	0,298	2,3%	97,2%
9	0,266	2,0%	99,3%
10	0,057	0,4%	99,7%
11	0,040	0,3%	100%
12	0,000	0,0%	100%
13	-0,000	-0,0%	100%

Tabla XXVI. Factorial B: cargas factoriales tras rotación varimax.

Item	Factor 1	Factor 2	Factor 3	Factor 4	Comunalidad
1 Poliartritis			0,774		93,4%
2 Grandes articulaciones			0,548	0,557	72,4%
3 Actividad a los 2 años	0,923				88,1%
4 Actividad mantenida	0,928				91,4%
6 Manifestaciones extraarticulares				0,947	91,8%
7 Imposibilidad remisión	0,739				69,8%
8 Refractariedad	0,777				62,1%
9 Erosiones al diagnóstico		0,842			76,2%
10 Nuevas erosiones a los 2 años		0,804			68,8%
11 Ausencia de mejoría HAQ			-0,895		82,1%
12 Cirugía ortopédica	0,590	0,602			94,9%
13 Ingresos (actividad/inflamación)		0,628		0,505	71,1%
14 Corticodependencia					52.2%

Tabla XXVII. Correlación ítem-total y consistencia interna.

	Correlación ítem-total
FACTOR 1	
3 Actividad a los 2 años	0,619
4 Actividad mantenida	0,725
7 Imposibilidad remisión	0,496
8 Refractariedad	0,487
Consistencia interna KUDER RICHARDSON	0,781
FACTOR 2	
9 Erosiones al diagnóstico	0,377
10 Nuevas erosiones a los 2 años	0,307
12 Cirugía ortopédica	0,233
13 Ingresos (actividad/inflamación)	0,269
Consistencia interna KUDER RICHARDSON	0,461
FACTOR 3	
1 Poliartritis	-0,250
11 Ausencia de mejoría HAQ	-0,250
Consistencia interna KUDER RICHARDSON	-0,441
FACTOR 4	
2 Grandes articulaciones	0,121
6 Manifestaciones extraarticulares	0,121
Consistencia interna KUDER RICHARDSON	0,146

Abreviaturas: HAQ: Health Assessment Questionnaire.

5.2.3 Análisis factorial C: eliminación del ítem 14 (corticodependencia)

La no inclusión en el índice de la corticodependencia vuelve a ofrecer una solución factorial de 4 factores que explican el 81,3% de la varianza total (tabla XXVIII). El reparto de cargas es similar al caso anterior, aunque en este caso la cirugía ortopédica es el ítem con cargas más bajas que solapan en los factores 1 y 2 (tabla XXIX). Los valores de consistencia interna de los 4 factores son similares a los anteriores y en cuanto a la capacidad de discriminación de los ítems, se observa que la cirugía ortopédica no aporta nada a la clasificación de los grupos, por lo que es el siguiente ítem a eliminar.

Tabla XXVIII. *Factorial C: factores, autovalores y varianza explicada.*

Factor	Autovalor	Varianza explicada (%)	Varianza acumulada (%)
1	5,029	41,9	41,9
2	1,842	15,3	57,3
3	1,591	13,3	70,5
4	1,299	10,8	81,3
5	0,834	6,9	88,3
6	0,554	4,6	92,9
7	0,445	3,7	96,6
8	0,289	2,4	99,0
9	0,070	0,6	99,6
10	0,047	0,4	100
11	0,000	0,0	100
12	-0,000	-0,0	100

Tabla XXIX. Factorial C: cargas factoriales tras rotación varimax.

Item	Factor 1	Factor 2	Factor 3	Factor 4	Comunalidad
1 Poliartritis			0,780		0,080
2 Grandes articulaciones			0,522	0,637	0,204
3 Actividad a los 2 años	0,921				0,125
4 Actividad mantenida	0,927				0,084
6 Manifestaciones extraarticulares				0,948	0,070
7 Imposibilidad remisión	0,739				0,288
8 Refractariedad	0,777				0,385
9 Erosiones al diagnóstico		0,843			0,233
10 Nuevas erosiones a los 2 años		0,831			0,277
11 Ausencia de mejoría HAQ			-0,898		0,181
12 Cirugía ortopédica	0,601	0,574			0,042
13 Ingresos (actividad/inflamación)		0,637		0,515	0,268

Abreviaturas: HAQ: Health Assessment Questionnaire.

Tabla XXX. Factorial C: Correlación ítem total y consistencia interna.

	Correlación ítem-total
FACTOR 1	
3 Actividad a los 2 años	0,618
4 Actividad mantenida	0,723
7 Imposibilidad remisión	0,499
8 Refractariedad	0,489
12 Cirugía ortopédica	0,128
Consistencia interna KUDER RICHARDSON	0,737
FACTOR 2	
9 Erosiones al diagnóstico	0,376
10 Nuevas erosiones a los 2 años	0,300
13 Ingresos (actividad/inflamación)	0,243
Consistencia interna KUDER RICHARDSON	0,478
FACTOR 3	
1 Poliartritis	-0,250
11 Ausencia de mejoría HAQ	-0,250
Consistencia interna KUDER RICHARDSON	-0,441
FACTOR 4	
2 Grandes articulaciones	0,121
6 Manifestaciones extraarticulares	0,121
Consistencia interna KUDER RICHARDSON	0,147

5.2.4 Análisis factorial D: eliminación del ítem 12 (cirugía ortopédica)

La varianza explicada por los 4 factores obtenidos es del 80,6% (tabla XXXI). El primer factor sigue agrupando variables de actividad (mantenida, elevada a los 2 años, imposibilidad de remisión y refractariedad), el segundo lesión estructural e ingresos; el tercero grandes articulaciones y manifestaciones extra-articulares; y el cuarto poliartritis y capacidad funcional (tabla XXXII).

La consistencia interna oscila entre 0,146 para el factor 3 (formado sólo por dos ítems) y 0,781 para el factor 1 (tabla XXIII). El ítem a eliminar en el paso siguiente serán los ingresos por presentar cargas factoriales bajas (<0,6) y solapadas (factores 2 y 3), y baja capacidad de discriminación (correlación ítem-total=0,243) (tabla XXIV).

Tabla XXXI. Factorial D: factores, autovalores y varianza explicada.

Factor	Autovalor	Varianza explicada (%)	Varianza acumulada (%)
1	4,105	37,3	37,3
2	1,955	17,8	55,1
3	1,478	13,4	68,5
4	1,331	12,1	80,6
5	0,781	7,1	87,8
6	0,579	5,3	93,0
7	0,405	3,7	96,7
8	0,291	2,6	99,3
9	0,073	0,7	100
10	0,000	0,0	100
11	-0,000	-0,0	100

Tabla XXXII. *Factorial D: cargas factoriales tras rotación varimax.*

Item	Factor 1	Factor 2	Factor 3	Factor 4	Comunalidad
1 Poliartritis	0,511			0,764	96,0%
2 Grandes articulaciones			0,679		78,5%
3 Actividad a los 2 años	0,921				87,2%
4 Actividad mantenida	0,951				96,3%
6 Manifestaciones extraarticulares			0,955		93,4%
7 Imposibilidad remisión	0,716				74,1%
8 Refractariedad	0,771				65,4%
9 Erosiones al diagnóstico		0,859			79,3%
10 Nuevas erosiones a los 2 años		0,776			61,6%
11 Ausencia de mejoría HAQ				-0,896	82,6%
13 Ingresos		0,597	0,572		72,7%

Abreviaturas: HAQ: Health Assessment Questionnaire.

Tabla XXXIII.-*Factorial D: cargas factoriales tras rotación varimax.*

	Correlación ítem-total
FACTOR 1	
3 Actividad a los 2 años	0,619
4 Actividad mantenida	0,725
7 Imposibilidad remisión	0,496
8 Refractariedad	0,487
Consistencia interna KUDER RICHARDSON	0,781
FACTOR 2	
9 Erosiones al diagnóstico	0,376
10 Nuevas erosiones a los 2 años	0,300
13 Ingresos	0,243
Consistencia interna KUDER RICHARDSON	0,478
FACTOR 3	
2 Grandes articulaciones	0,121
6 Manifestaciones extraarticulares	0,121
Consistencia interna KUDER RICHARDSON	0,146
FACTOR 4	
1 Poliartritis	-0,250
11 Ausencia de mejoría HAQ	-0,250
Consistencia interna KUDER RICHARDSON	-0,441

Abreviaturas: HAQ: Health Assessment Questionnaire.

5.2.5 Análisis factorial E: eliminación del ítem 13 (ingresos)

Los cuatro factores extraídos explican el 82,8% de la varianza total (tabla XXIV). La distribución de cargas varía sensiblemente ya que en el factor 2 cargan ahora la poliartritis y la capacidad funcional, mientras que en el 3 aparecen las lesiones estructurales (erosiones) y en el 4 las grandes articulaciones y las manifestaciones extraarticulares (tabla 19). Los valores de consistencia interna se mueven en el rango de 0,147 para el factor 4 (dos ítems) hasta 0,781 para el factor 1 (tabla XXV).

La imposibilidad de remisión y la actividad mantenida son variables redundantes, por lo que se decide eliminar el primero de estos dos ítems en el siguiente análisis.

Tabla XXXIV. Factorial E: factores, autovalores y varianza explicada.

Factor	Autovalor	Varianza explicada (%)	Varianza acumulada (%)
1	3,872	38,7	38,7
2	1,663	16,4	55,4
3	1,453	14,5	69,9
4	1,288	12,9	82,8
5	0,737	7,4	90,1
6	0,568	5,7	95,8
7	0,290	2,9	98,7
8	0,116	1,2	99,9
9	0,010	0,1	100
10	-0,000	-0,0	100

Tabla XXXV. Factorial E: cargas factoriales tras rotación varimax.

Item	Factor 1	Factor 2	Factor 3	Factor 4	Comunalidad
1 Poliartritis	0,503	0,776			97,2
2 Grandes articulaciones				0,663	78,3
3 Actividad a los 2 años	0,923				87,9
4 Actividad mantenida	0,951				95,8
6 Manifestaciones extraarticulares				0,981	98,1
7 Imposibilidad remisión	0,754				74,4
8 Refractariedad	0,769				64,7
9 Erosiones al diagnóstico			0,859		80,4
10 Nuevas erosiones a los 2 años			0,828		69,2
11 Ausencia de mejoría HAQ		-0,886			81,5

Abreviaturas: HAQ: Health Assessment Questionnaire.

Tabla XXXVI. Factorial E: consistencia interna y correlación ítem-total.

	Correlación ítem-total
FACTOR 1	
3 Actividad a los 2 años	0,619
4 Actividad mantenida	0,725
7 Imposibilidad remisión	0,496
8 Refractariedad	0,487
Consistencia interna KUDER RICHARDSON	0,781
FACTOR 2	
1 Poliartritis	-0,250
11 Ausencia de mejoría HAQ	-0,250
Consistencia interna KUDER RICHARDSON	-0,441
FACTOR 3	
9 Erosiones al diagnóstico	0,307
10 Nuevas erosiones a los 2 años	0,307
Consistencia interna KUDER RICHARDSON	0,481
FACTOR 4	
2 Grandes articulaciones	0,121
6 Manifestaciones extraarticulares	0,121
Consistencia interna KUDER RICHARDSON	0,147

Abreviaturas: HAQ: Health Assessment Questionnaire.

5.2.6 Análisis factorial F: eliminación del ítem 7 (imposibilidad remisión)

Esta solución factorial explica el 84,5% de la varianza total (tabla XXXVII) y su distribución es similar al análisis anterior (tabla XXXVIII). El siguiente ítem a eliminar es la afectación de grandes articulaciones, ya que es el que tiene cargas factoriales menores y nula capacidad de discriminación (correlación ítem-total=0,121) (tabla XXXIX).

Tabla XXXVII. Factorial F: factores, autovalores y varianza explicada.

Factor	Autovalor	Varianza explicada (%)	Varianza acumulada (%)
1	3,201	35,6	35,6
2	1,658	18,4	54,0
3	1,463	16,3	70,2
4	1,285	14,3	84,5
5	0,661	7,3	91,9
6	0,563	6,3	98,1
7	0,134	1,5	99,6
8	0,034	0,4	100
9	-0,000	-0,0	100

Tabla XXXVIII. Factorial F: cargas factoriales tras rotación varimax.

Item	Factor 1	Factor 2	Factor 3	Factor 4	Comunalidad
1 Poliartritis		0,806			96,8
2 Grandes articulaciones		0,515	0,653		79,1
3 Actividad a los 2 años	0,944				91,3
4 Actividad mantenida	0,914				91,4
6 Manifestaciones extraarticulares			0,983		98,6
8 Refractariedad	0,808				70,3
9 Erosiones al diagnóstico				0,857	79,6
10 Nuevas erosiones a los 2 años				0,837	71,0
11 Ausencia de mejoría HAQ		-0,890			82,8

Abreviaturas: HAQ: Health Assessment Questionnaire.

Tabla XXXIX.- Factorial F: Consistencia interna y correlación ítem-total.

	Correlación ítem-total
FACTOR 1	
3 Actividad a los 2 años	0,645
4 Actividad mantenida	0,628
8 Refractariedad	0,513
Consistencia interna KUDER RICHARDSON	0,769
FACTOR 2	
1 Poliartritis	-0,250
11 Ausencia de mejoría HAQ	-0,250
Consistencia interna KUDER RICHARDSON	-0,441
FACTOR 3	
2 Grandes articulaciones	0,121
6 Manifestaciones extraarticulares	0,121
Consistencia interna KUDER RICHARDSON	0,147
FACTOR 4	
9 Erosiones al diagnóstico	0,307
10 Nuevas erosiones a los 2 años	0,307
Consistencia interna KUDER RICHARDSON	0,481

Abreviaturas: HAQ: Health Assessment Questionnaire.

5.2.7 Análisis factorial G: eliminación ítem 2 (grandes articulaciones)

Se obtiene una solución factorial de 4 factores que explican un 88,6% de la varianza (tabla XL), en la que la poliartritis es el ítem con menores cargas factoriales y solapadas entre los factores 3 y 4 (tabla XLI) que también presenta baja capacidad de discriminación (tabla XVII).

Tabla XL. Factorial G: factores, autovalores y varianza explicada.

Factor	Autovalor	Varianza explicada (%)	Varianza acumulada (%)
1	2,943	36,8	36,8%
2	1,586	19,8	56,7%
3	1,477	18,5	75,1%
4	1,079	13,5	88,6%
5	0,659	8,2	96,9%
6	0,134	1,7	98,5%
7	0,117	1,5	100%
8	-0,000	-0,0	100%

Tabla XLI. Factorial G: cargas factoriales tras rotación varimax.

Item	Factor 1	Factor 2	Factor 3	Factor 4	Comunalidad
1 Poliartritis			-0,594	-0,606	99,1%
3 Actividad a los 2 años	0,937				91,6%
4 Actividad mantenida	0,904				88,5%
6 Manifestaciones extraarticulares			0,986		99,0%
8 Refractariedad	0,845				76,5%
9 Erosiones al diagnóstico		0,856			83,5%
10 Nuevas erosiones a los 2 años		0,846			72,9%
11 Ausencia de mejoría HAQ				0,986	97,6%

Abreviaturas: HAQ: Health Assessment Questionnaire.

Tabla XLII. Factorial G: Consistencia interna y correlación ítem-total.

	Correlación ítem-total
FACTOR 1	
3 Actividad a los 2 años	0,645
4 Actividad mantenida	0,628
8 Refractariedad	0,513
Consistencia interna KUDER RICHARDSON	0,769
FACTOR 2	
9 Erosiones al diagnóstico	0,307
10 Nuevas erosiones a los 2 años	0,307
Consistencia interna KUDER RICHARDSON	0,481
FACTOR 3	
6 Manifestaciones extraarticulares	-
Consistencia interna KUDER RICHARDSON	-
FACTOR 4	
1 Poliartritis	- 0,250
11 Ausencia de mejoría HAQ	- 0,250
Consistencia interna KUDER RICHARDSON	0,441

Abreviaturas: HAQ: Health Assessment Questionnaire.

5.2.8 Análisis factorial H: eliminación del ítem 1 (poliartritis)

La eliminación del ítem 1 (poliartritis) produce una solución de tres factores que explican el 80% de la varianza. En el primer factor cargan los ítems de actividad (a los dos años, mantenida y refractariedad), en el segundo las lesiones estructurales (erosiones al diagnóstico y a los 2 años); y en el tercero las manifestaciones extraarticulares y la falta de mejoría en el HAQ (tabla 27 y 28). Los factores 1 y 2 tienen valores del coeficiente de Kuder Richardson de 0,769 y 0,481, respectivamente. La consistencia interna del tercer factor es muy baja (KR 0,101) y los dos ítems constituyentes carecen de capacidad de discriminación (correlación ítem-total inferior a 0,3) (tabla XLV). Por consiguiente se opta por eliminar el ítem 11 o ausencia de mejoría en el HAQ.

Tabla XLIII. Factorial H: factores, autovalores y varianza explicada.

Factor	Autovalor	Varianza explicada (%)	Varianza acumulada (%)
1	2,559	36,6	36,6
2	1,695	24,2	60,8
3	1,202	17,2	80,0
4	0,840	12,0	90,0
5	0,451	6,4	96,4
6	0,135	1,9	98,3
7	0,117	1,7	100

Tabla XLIV. Factorial H: cargas factoriales tras rotación varimax.

Item	Factor 1	Factor 2	Factor 3	Comunalidad
3 Actividad a los 2 años	0,932			90,6
4 Actividad mantenida	0,918			85,4
6 Manifestaciones extraarticulares			0,777	70,4
8 Refractariedad	0,859			78,4
9 Erosiones al diagnóstico		0,882		85,4
10 Nuevas erosiones a los 2 años		0,813		67,7
11 Ausencia mejoría HAQ			-0,819	67,8

Abreviaturas: HAQ: Health Assessment Questionnaire.

Tabla XLV. Factorial H: Consistencia interna y correlación ítem-total.

	Correlación ítem-total
FACTOR 1	
3 Actividad a los 2 años	0,645
4 Actividad mantenida	0,628
8 Refractoriedad	0,513
Consistencia interna KUDER RICHARDSON	0,769
FACTOR 2	
9 Erosiones al diagnóstico	0,307
10 Nuevas erosiones a los 2 años	0,307
Consistencia interna KUDER RICHARDSON	0,481
FACTOR 3	
6 Manifestaciones extraarticulares	-0,094
11 Ausencia mejoría HAQ	-0,094

Abreviaturas: HAQ: Health Assessment Questionnaire.

5.2.9 Análisis factorial I: eliminación del ítem 11 (ausencia de mejoría HAQ)

Este nuevo análisis produce una solución de tres factores que explican el 86% de la varianza total (tabla XLVI). Los tres factores pueden identificarse con la actividad de la enfermedad, la lesión estructural y las manifestaciones sistémicas, respectivamente (tabla XLVII). La consistencia interna del primer factor (constituido por actividad a los 2 años, actividad mantenida y refractoriedad), es de 0,769; la del segundo (erosiones al diagnóstico y aumento de su número a los 2 años) es más baja, con un coeficiente de Kuder-Richardson de 0,481, aunque éste factor está integrado únicamente por 2 ítems. En el tercer factor sólo cargan las manifestaciones extra-articulares, por lo que no puede calcularse su consistencia interna (tabla XLVIII). Esta solución factorial se da por definitiva. En la tabla XLIX se presenta un resumen de todos los análisis realizados y en la *figura 6* la representación gráfica de la solución factorial final.

Tabla XLVI. Factorial I: factores, autovalores y varianza explicada.

Factor	Autovalor	Varianza explicada (%)	Varianza acumulada (%)
1	2,559	42,6	42,6
2	1,607	26,8	69,4
3	1,005	16,7	86,2
4	0,547	9,1	95,3
5	0,165	2,8	98,0
6	0,117	1,9	100,0

Tabla XLVII. Factorial I: cargas factoriales tras rotación varimax.

Item	Factor 1	Factor 2	Factor 3	Comunalidad
3 Actividad a los 2 años	0,929			87,6
4 Actividad mantenida	0,918			85,4
6 Manifestaciones extraarticulares			0,963	92,9
8 Refractariedad	0,862			74,9
9 Erosiones al diagnóstico		0,686	0,622	86,7
10 Nuevas erosiones a los 2 años		0,944		89,6

Tabla XLVIII. Factorial I: Consistencia interna y correlación ítem-total.

	Correlación ítem-total
FACTOR 1	
3 Actividad a los 2 años	0,645
4 Actividad mantenida	0,628
8 Refractariedad	0,513
Consistencia interna KUDER RICHARDSON	0,769
FACTOR 2	
9 Erosiones al diagnóstico	0,307
10 Nuevas erosiones a los 2 años	0,307
Consistencia interna KUDER RICHARDSON	0,481
FACTOR 3	
6 Manifestaciones extraarticulares	-
Consistencia interna KUDER RICHARDSON.	-

Tabla XLIX. Análisis factorial: Resumen.

Factorial	Item eliminado (número de ítems en el índice)	Factores	Varianza explicada	Consistencia interna
A	Ninguno (n=14)	5	82%	Fc 1 (n=5) = 0,737 Fc 2 (n=3) = 0,478 Fc 3 (n=2) = 0,441 Fc 4 (n=2) = 0,147 Fc 5 (n=2) = 0,462
B	Reactantes de fase aguda (n=13)	4	78%	Fc 1 (n=4) = 0,781 Fc 2 (n=4) = 0,461 Fc 3 (n=2) = 0,441 Fc 4 (n=2) = 0,146
C	Corticodependencia (n=12)	4	81%	Fc 1 (n=4) = 0,737 Fc 2 (n=4) = 0,478 Fc 3 (n=2) = 0,441 Fc 4 (n=2) = 0,147
D	Cirugía ortopédica (n=11)	4	81%	Fc 1 (n=4) = 0,781 Fc 2 (n=3) = 0,478 Fc 3 (n=2) = 0,146 Fc 4 (n=2) = 0,441
E	Ingresos (n=10)	4	83%	Fc 1 (n=4) = 0,781 Fc 2 (n=2) = 0,441 Fc 3 (n=2) = 0,481 Fc 4 (n=2) = 0,147
F	Imposibilidad de remisión (n=9)	4	84%	Fc 1 (n=4) = 0,769 Fc 2 (n=2) = 0,441 Fc 3 (n=2) = 0,147 Fc 4 (n=2) = 0,481
G	Afectación de grandes articulaciones (n=8)	4	89%	Fc 1 (n=3) = 0,769 Fc 2 (n=2) = 0,481 Fc 3 (n=1) = ----- Fc 4 (n=2) = 0,441
H	Poliartritis (n=7)	3	80%	Fc 1 (n=3) = 0,769 Fc 2 (n=2) = 0,481 Fc 3 (n=2) = 0,101
I	Ausencia de mejoría en el HAQ (n=6)	3	86%	Fc 1 (n=3) = 0,769 Fc 2 (n=2) = 0,481 Fc 3 (n=1) = -----

Abreviaturas: HAQ: Health Assessment Questionnaire.

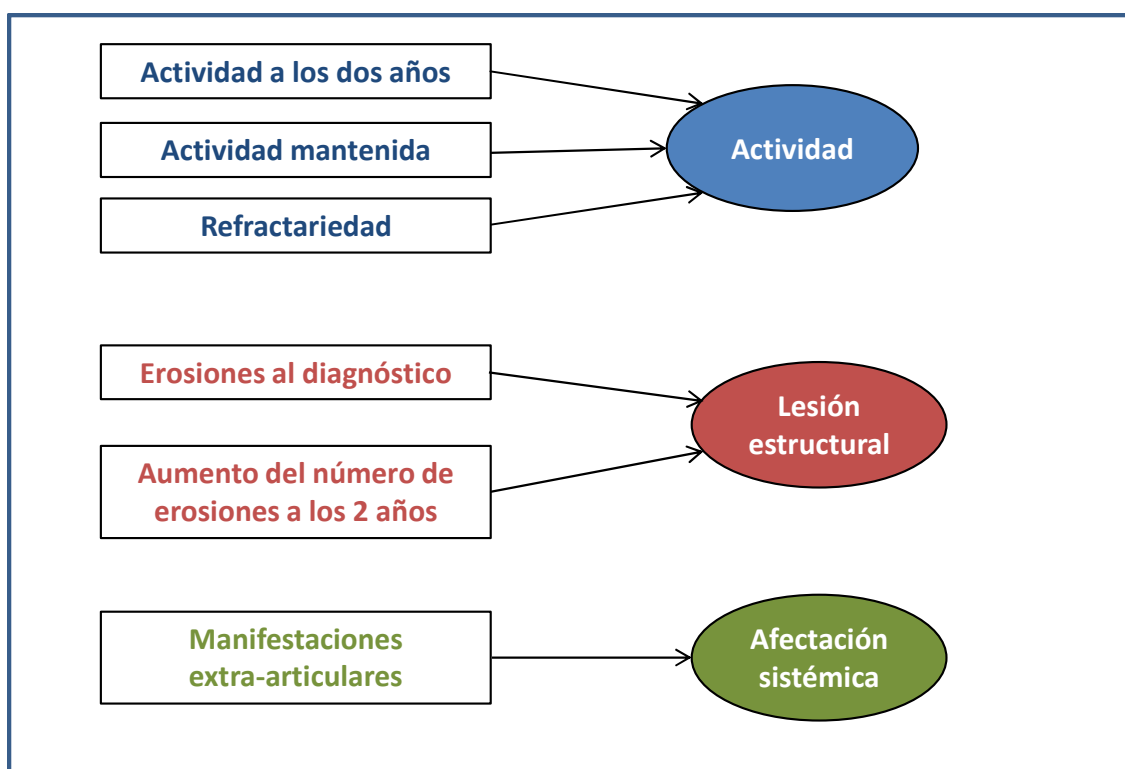


Figura 6. Solución factorial definitiva: factorización.

6. Capacidad de discriminación

La variable de clasificación externa de la gravedad de los pacientes (variable dependiente) se construyó a partir del número de días de tratamiento con FAME dividido por el tiempo de seguimiento de cada paciente. Esta variable se ponderó por el tipo de FAME. De esta forma, su valor es 1 si ha tomado antipalúdicos todos los días de su seguimiento; 1'5 si ha recibido todos los días MTX, LF o SSZ; 2 si ha estado todos los días con un biológico; 3 si tomó MTX y LF en combinación todos los días; y 3'5 si estuviera con un anti-TNF en combinación todos los días. Las puntuaciones de esta variable en la población estudiada oscilan entre 0 y 5, por lo que se decidió que a partir de esta variable continua, los pacientes se clasificaran en dos grupos:

- Leves: si el valor de la variable anterior era de $<2,5$
- Graves: si el valor de la variable anterior era $\geq 2,5$

Esta categorización del tratamiento con FAME fue la variable patrón externo o variable de clasificación utilizada como variable dependiente en el análisis discriminante.

Como variables independientes se introdujo la combinación lineal de los tres factores extraídos después de multiplicar, en cada uno de ellos, el valor de los ítems constituyentes por sus respectivas cargas factoriales.

Esta combinación clasificó correctamente al 55% de los pacientes leves y al 63% de los graves en la tabla de validación cruzada (clasificación de cada caso tras dejarlo fuera del cálculo de la función discriminante) (tabla 34), con una tasa de error global del 39%.

Tabla L. *Análisis discriminante: clasificación cruzada.*

Clasificación real	Clasificación Análisis Discriminante		Total
	Leves	Graves	
Leves	67 (55,4%)	54 (44,6%)	121 (100%)
Graves	13 (37,1%)	22 (62,9%)	35 (100%)
Total	80 (51,3%)	76 (48,7%)	156 (100%)

A pesar de la elevada tasa de error de clasificación obtenida en este análisis, se realizó el análisis de regresión lineal múltiple con la función lineal discriminante para calcular los pesos o ponderaciones de los ítems seleccionados, como aproximación preliminar al índice de gravedad. El modelo final explicó un 97,4% de la varianza de la función discriminante (R^2 ajustado= 0,97) e incluyó los ítems 3, 4, 8, 10 y 6. El ítem 9, referido a la presencia de erosiones en el momento del diagnóstico, es eliminado del modelo ya que no contribuye a la función discriminante (tabla LI).

Tabla LI. *Regresión lineal múltiple.*

Función discriminante	Coefficiente β	Valor p	IC 95%
Item 3: Actividad elevada a los dos años de seguimiento	0,66	<0,0001	0,59;0,73
Item 4: Actividad mantenida durante los dos años	0,68	<0,0001	0,62;0,75
Item 8: Refractoriedad	0,53	<0,0001	0,46;0,59
Item 9: Erosiones en el momento del diagnóstico	-	-	-
Item 10: Aumento del número de erosiones a los dos años	0,65	<0,0001	0,59;0,71
Item 6: Manifestaciones extra-articulares	-3,59	<0,0001	-3,75;-3,42
Constante	-0,88	<0,0001	-0,92;-0,84

A la vista de estos resultados, y tomando como ponderaciones de los ítems los coeficientes de la regresión, el índice propuesto, denominado EASI (*Early Arthritis Severity Index*) vendría representado por la siguiente ecuación:

$$\text{EASI} = -0,88 + 0,66*\text{item3} + 0,68*\text{item4} + 0,53*\text{item8} + 0,65*\text{item10} - 3,59*\text{item6}$$

Donde EASI tendría una puntuación entre 0 (extremadamente leve) y 1 (extremadamente grave).

7. Conclusión

Se ha obtenido una solución factorial con sentido clínico. Según esta solución, la estructura del constructo “gravedad” está formada por 3 dominios que son actividad, lesión estructural y manifestaciones sistémicas. Esta estructura mantiene unas buenas propiedades psicométricas, con un 86% de la varianza total y adecuada consistencia interna (0,769 para el primer dominio y 0,481 para el segundo, aunque éste último sólo contiene dos ítems).

A pesar de la plausibilidad biológica de esta estructura y de sus adecuadas propiedades psicométricas, esta solución factorial no clasifica correctamente más que al 61% de los pacientes en función de su gravedad, establecida por un criterio externo. De hecho, clasifica correctamente al 55% de los pacientes leves y al 63% de los graves según la intensidad de tratamiento con FAME.

Dados las buenas propiedades psicométricas de la solución factorial obtenida, los resultados del análisis discriminante podrían explicarse por el carácter “blando” (poco rotundo) de la variable de clasificación utilizada. Es decir, la clasificación de los pacientes en diferentes niveles de gravedad en función del tiempo de tratamiento con FAME, no deja de ser una variable artificial, y con cierto carácter de subjetividad, en la que no sólo interviene la situación clínica del paciente, sino también la variabilidad propia del médico, y otros factores dependientes del paciente y no siempre

relacionados con la gravedad de la enfermedad, como posible toxicidad o intolerancia farmacológica, comorbilidad o deseo de embarazo.

No obstante, y a pesar de las limitaciones de la variable de clasificación, se optó por realizar una aproximación preliminar a la construcción del índice mediante el cálculo de los pesos o ponderaciones de los ítems seleccionados. A pesar de su carácter preliminar y sus limitaciones metodológicas, los resultados de la ecuación tienen sentido clínico, ya que la lesión estructural que define gravedad no es la presencia de erosiones en el momento del diagnóstico (ítem eliminado del modelo), sino el aumento del número de ellas a los dos años de seguimiento.

El hallazgo más sorprendente al realizar la construcción del índice fue el signo negativo del coeficiente del ítem 6 que indica que la presencia de manifestaciones extraarticulares tiende a clasificar a los pacientes como menos graves. Ante esta incongruencia desde el punto de vista teórico, y con el fin de interpretar adecuadamente este resultado, se analizaron las características clínicas y el comportamiento de la variable dependiente en los cuatro pacientes que presentaron manifestaciones extraarticulares:

- El paciente ARC499, de 59 años de edad, presentó como manifestación sistémica una neumopatía intersticial no específica (NINE), que precisó tratamiento con ciclofosfamida durante 6 meses (fármaco no incluido para el cálculo de la variable IDT), recibiendo una dosis acumulada de glucocorticoides de 7,9 gr de prednisona en 2 años (muy por encima del promedio del resto de la población). El valor de la variable IDT utilizada como variable externa de clasificación fue de 2,53, por lo que entró en el grupo de pacientes graves.
- El paciente ARC383, de 70 años de edad, presentó como manifestación sistémica una neumonía organizativa criptogenética (NOC) que se trató con glucocorticoides, con una dosis acumulada de 6,6 gr de prednisona en 2 años. El valor de la variable IDT fue de 1,52, por lo que entró en el grupo de pacientes no graves.

- El paciente ARC391, de 76 años de edad, presentó como manifestación extraarticular una NINE leve que fue tratada con glucocorticoides, con una dosis acumulada de 6,61 gr de prednisona en 2 años. El valor de la variable IDT fue de 1,26, por lo que se clasificó en el grupo de pacientes no graves.
- El paciente ARC87, de 67 años de edad, presentó una serositis como manifestación sistémica que fue tratada con glucocorticoides con una dosis acumulada de 3,8 gr de prednisona en 2 años. El valor de la variable IDT fue de 1,78, por lo que fue clasificado como paciente no grave.

Por tanto, las limitaciones de la variable de clasificación propuesta, especialmente en lo que concierne a los pacientes que presentan manifestaciones sistémicas, obligan a considerar como preliminar el índice propuesto y subrayan la necesidad de seguir trabajando para conseguir nuevas variables de clasificación de la gravedad, además de evaluar la evolución de los pacientes estudiados en los próximos años.

PARTE III. DISCUSIÓN GENERAL

Resumen de resultados. Fortalezas y limitaciones	104
Diseño y construcción de un índice de gravedad	107
Perspectivas futuras.....	110

PARTE III. DISCUSIÓN GENERAL

RESUMEN DE RESULTADOS. FORTALEZAS Y LIMITACIONES

Los datos mostrados en el **sub-estudio 1** indican que durante los últimos 15 años los reumatólogos del Servicio de Reumatología del Hospital Universitario de La Princesa han realizado un tratamiento más intensivo de los pacientes clasificados como artritis indiferenciada. Aquellos pacientes que cumplieron los criterios de AR fueron tratados de manera más agresiva incluso en los primeros años del registro PEARL. La explicación para este hallazgo es coherente con el hecho de haber interiorizado el concepto de ventana de oportunidad y creer que tratar de forma precoz a nuestros pacientes mejora el pronóstico. De hecho, a lo largo de estos 15 años ha mejorado el porcentaje de pacientes en remisión y los datos de la progresión radiológica.

En nuestra opinión, y como consecuencia del cambio en el abordaje terapéutico de los pacientes con artritis de inicio, nuestra capacidad de conseguir resultados firmes en el estudio/detección de nuevos biomarcadores de gravedad es peor que hace 15-20 años. La progresión radiográfica ha sido considerada hasta ahora el patrón oro para analizar la gravedad de la enfermedad en AR. Sin embargo, la posibilidad de retrasar e incluso detener completamente la progresión radiológica ha mejorado a lo largo de los últimos 15 años. En este sentido, en el registro PEARL hemos demostrado que, en ausencia de un protocolo terapéutico preestablecido, los pacientes ACPA positivos, un biomarcador validado de gravedad, no presentan mayor progresión radiológica que los ACPA negativos. Esta observación se relaciona con un tratamiento más intenso en el primer grupo.

Considerando tanto la situación descrita como la complejidad para valorar de forma sistemática la progresión radiológica en grandes estudios observacionales, hemos utilizado como variable subrogada de gravedad la dosis acumulada de FAME durante los dos primeros años de tratamiento. Esta estrategia permite detectar algunos parámetros que podrían ser utilizados como biomarcadores de gravedad. Sin embargo, los resultados descritos en el presente trabajo sugieren que podría ser necesario desarrollar un índice compuesto para valorar globalmente la gravedad en pacientes con artritis de inicio. Esta herramienta tendría que incluir diferentes aspectos de la enfermedad (progresión radiológica, intensidad de tratamiento, presencia de complicaciones sistémicas, discapacidad,...) y permitiría realizar una búsqueda más eficiente de biomarcadores que predijeran gravedad.

La búsqueda de marcadores o factores pronósticos precisa de medidas de desenlace fiables. En enfermedades cuyas variables de desenlace son extremadamente objetivas y robustas, como el cáncer (muerte o progresión tumoral) o la osteoporosis (aparición de nueva fractura o variación de la densidad mineral ósea), la medicina personalizada o, al menos, ajustada a perfiles de paciente parece cada día más viable. Sin embargo, y por el momento, en la AR carecemos de estas variables objetivas o robustas. El principal “patrón oro” disponible hasta hace unos años, la progresión radiológica, ha sucumbido ante las estrategias de tratamiento intensivas(8) y la aparición de fármacos capaces de frenar la progresión radiológica, incluso en pacientes que no alcanzan una respuesta terapéutica mínima(65).

El **sub-estudio 2** tenía por objetivo identificar índices compuestos de gravedad, basados en diferentes dominios, que pudiesen ser aplicados a la búsqueda de factores predictivos de gravedad en la AR. La búsqueda sistemática realizada hasta 2017 sólo permitió identificar tres herramientas, los índices PAS(57), el RARBIS(60) y el CIRAS(62). Su desarrollo se inició poco tiempo después del inicio de la era biológica en el campo de la Reumatología, hace aproximadamente 15 años, y supone un considerable esfuerzo por encontrar un índice capaz de estimar la gravedad de la AR.

El más antiguo de ellos, RARBIS, se basa en datos obtenidos de las historias clínicas y su principal defecto es que, por su antigüedad, pondera circunstancias como la presencia de síndrome de Felty, que en la actualidad prácticamente no vemos en nuestros pacientes. Algunas otras variables incluidas en este índice, especialmente en los dominios cirugía y radiología, no están presentes en los 2 primeros años de enfermedad. Por último, su pobre correlación con las necesidades de tratamiento o la actividad clínica de los pacientes llevó a sus autores a replantearse el acercamiento al problema.

El Dr. Wolfe(57) utilizó el PAS como variable de gravedad sin haber realizado previamente un estudio de validación, sino intercambiando simplemente el concepto de actividad por el de gravedad, sin obtener una buena capacidad de discriminación para el criterio “tratamiento con FAME y biológico”.

El CIRAS se basa fundamentalmente en la obtención de información electrónicamente registrada sobre la utilización de servicios sanitarios. Aunque las variables incluidas en este índice sí pueden recogerse en los 2 primeros años de evolución de los pacientes con artritis de reciente comienzo, la incorporación de las estrategias “TC” y “treat-to-target” en la práctica clínica habitual hace que el número de visitas al reumatólogo y la realización de múltiples determinaciones de laboratorio no tengan por qué estar relacionadas con una mayor gravedad y sí con la puesta en práctica de esta estrategia terapéutica. Por otro lado, las visitas al rehabilitador, muy frecuentes en épocas anteriores, son excepcionales en la era de la TB, gracias al mejor control de la actividad de la enfermedad en los primeros años de evolución. Además el índice CIRAS incluye variables que, aunque relacionadas con la utilización de los servicios sanitarios (solicitud de marcadores de inflamación, de recuento de plaquetas y de otros parámetros bioquímicos), permiten poner en práctica la estrategia “treat-to-target” sin que eso signifique un aumento de la gravedad. Otras variables incluidas, como el Síndrome de Felty, tienen una presencia clínica excepcional.

Por otro lado, la validación de estos índices plantea algunos problemas metodológicos. Aunque el desarrollo del RARBIS se basó en un panel de expertos, las subescalas (dominios) y los indicadores (variables) se ponderaron de acuerdo con el criterio del panel según la correlación con la gravedad. Para valorar la posible interdependencia o contribución relativa de las variables seleccionadas no se utilizó ninguna técnica de reducción de ítems que evitara posibles redundancias. En el CIRAS las variables incluidas se eligen como parte del mejor modelo identificado. No en función de aspectos relevantes del constructo, sino en función de la disponibilidad de recogida en registros de salud, sin plantearse si reflejan todos los aspectos relevantes del constructo “gravedad”. En ambos índices únicamente se contempla la validez de constructo (validez convergente), sin tener en cuenta otros aspectos de la validación ni la viabilidad. La correlación con el constructo utilizado en ambos casos fue débil, tal y como se ha explicado con anterioridad. Además, por el hecho de utilizar la misma cohorte de pacientes para el desarrollo de ambos índices, algunas correlaciones encontradas eran esperables. El desarrollo del PAS plantea dos limitaciones importantes: la primera de ellas es la utilización de medidas autorregistradas por el paciente, lo que puede originar un sesgo en la evaluación de la gravedad; la segunda es que sólo contempló la validez de criterio, con un bajo poder predictivo de pacientes graves.

DISEÑO Y CONSTRUCCIÓN DE UN ÍNDICE DE GRAVEDAD

El ambicioso objetivo planteado en el **sub-estudio 4** es el desarrollo de un índice de gravedad de la AR a los dos años de seguimiento, que reflejase adecuadamente el marco teórico de este constructo y pudiera obtenerse en la práctica clínica habitual. La información relevante obtenida en los dos paneles de expertos convocados y posterior encuesta mediante metodología Delphi del **sub-estudio 3** se transformó en definiciones de variables de respuesta dicotómica, con las que se trabajó para el

desarrollo del índice. Desde esta fase del estudio en la que pudimos recoger la opinión y la experiencia de expertos en artritis reumatoide buscamos de forma incesante un patrón oro o criterio externo con el fin de valorar la validez de criterio de nuestro de índice gravedad, con serias dificultades para encontrarlo.

El panel identificó varios potenciales criterios externos, como son: la valoración global de gravedad por el médico, con un componente subjetivo inherente que puede introducir un sesgo en la valoración del paciente. Otra posibilidad barajada sería la carga de tratamiento, en cuanto a líneas terapéuticas, número o dosis de fármacos modificadores de la enfermedad o terapia biológica y dosis acumulada de corticoides. Es bien sabido que en el manejo terapéutico individual de un paciente no sólo interviene la situación clínica. También la variabilidad propia del médico y factores dependientes del paciente, no siempre relacionados con la gravedad, son importantes. La toxicidad o intolerancia farmacológica, las comorbilidades, el deseo gestacional o los temores e ideas preconcebidas de la enfermedad y/o del tratamiento, están presentes en la toma de decisiones, entre otras muchas circunstancias condicionantes.

Además, el panel identificó algunos factores modificadores de la enfermedad como son el apoyo social del paciente, las condiciones socio-económicas, la accesibilidad al sistema sanitario, el nivel cultural, el nivel del conocimiento de la enfermedad, el tipo de trabajo realizado y la adherencia terapéutica.

Tampoco se encontró en la literatura científica un patrón oro con el que valorar la validez de criterio de nuestro índice de gravedad, ni siquiera un patrón bronce, menos rotundo como criterio externo pero que reflejara el constructo gravedad.

Al realizar el análisis factorial exploratorio, a pesar de que el estudio Delphi seleccionó la variable *tratamiento intensivo a lo largo de los dos años*, como uno de los ítem a incluir en el índice, el grupo investigador decidió finalmente no incluirlo por indicar una actitud del médico y no una característica del paciente.

La solución factorial obtenida tiene sentido clínico. Según esta solución, la estructura del constructo “gravedad” está formada por tres dominios: actividad inflamatoria de la enfermedad, lesión estructural y manifestaciones sistémicas. Esta estructura mantiene unas buenas propiedades psicométricas, con un 86% de la varianza total explicada y adecuada consistencia interna (0,769 para el primer dominio y 0,481 para el segundo, aunque éste último sólo contiene dos ítems).

La variable de clasificación externa de la gravedad de los pacientes se construyó a partir del número de días de tratamiento con FAME dividido por el tiempo de seguimiento de cada paciente. Esta variable se ponderó por el tipo de FAME, según se ha explicado en el **sub-estudio 4**, y fue utilizada como variable dependiente en el análisis discriminante (IDT). Esta combinación clasificó correctamente al 55% de los pacientes leves y al 63% de los graves según el tiempo de tratamiento con FAME, con una tasa de error global del 39%.

La fórmula para calcular índice EASI, que se presenta en esta tesis, solo puede considerarse como una aproximación preliminar. El signo negativo del coeficiente del ítem 6 indica que la presencia de manifestaciones extra-articulares tiende a clasificar a los pacientes como menos graves. Esta situación no tiene fundamento lógico y se explica porque los pacientes con afectación sistémica reciben más corticoides e inmunosupresores como la ciclofosfamida, azatioprina o micofenolato de mofetilo, en vez de los FAME y TB que se incluyen en la variable utilizada como criterio externo de clasificación. El análisis de las características clínicas y de la variable dependiente IDT en los cuatro pacientes que presentaron manifestaciones extraarticulares corrobora esta explicación.

Dadas las buenas propiedades psicométricas de la solución factorial obtenida, los resultados del análisis discriminante podrían explicarse por el carácter poco rotundo de la variable de clasificación utilizada, según hemos explicado. Esta clasificación en diferentes niveles de gravedad en función del tratamiento con FAME es artificial y con cierto grado de subjetividad. Sin embargo, no hemos encontrado un patrón oro más

adecuado con el que comparar el índice, a pesar de nuestros esfuerzos. Esto, que en sí mismo en una limitación del estudio, no deja de ser un reflejo de la complejidad del “constructo gravedad”, cuyos dominios pueden estar interferidos por factores dependientes del médico (42) (la práctica individual del médico puede hacer a una enfermedad refractaria si no utiliza suficiente carga de tratamiento, la variabilidad interpersonal en el uso del corticoide), del paciente (comorbilidades) o del tratamiento (efectos secundarios, intolerancias), que pueden dificultar el manejo de la enfermedad. Por último, algunos factores psico-sociales (apoyo social, condiciones socioeconómicas, accesibilidad al sistema sanitario), educativos (nivel cultural, nivel de conocimiento de la enfermedad), laborales, o inherentes al paciente(66) (características psicológicas que condicionan la adherencia terapéutica, ideas preconcebidas de la enfermedad y los fármacos) pueden comportarse como factores modificadores de la gravedad. Estos últimos, aunque relevantes para el curso evolutivo y gravedad de la enfermedad, son difíciles de recolectar en la práctica clínica habitual, incluso en entornos de consultas monográficas de artritis de reciente comienzo.

Aun reconociendo las dificultades y limitaciones comentadas, se decidió realizar una primera aproximación al índice EASI, mediante el cálculo de los pesos de los ítems. Los resultados de la regresión lineal de la función discriminante, aún siendo preliminares, tienen sentido clínico ya que se excluyen las erosiones en el momento del diagnóstico y la presencia de manifestaciones extra-articulares tiende a clasificar a los pacientes como menos graves, circunstancia que ya se ha explicado que se deriva de que recibían más corticoides y menos FAME.

PERSPECTIVAS FUTURAS

Las limitaciones encontradas en la clasificación de los pacientes subrayan la necesidad de seguir avanzando en la obtención de una nueva variable de clasificación de gravedad, con mayor robustez que la utilizada en este estudio.

Se intentará una nueva estimación de la fórmula del índice EASI, basada en las siguientes premisas:

- Encontrar una mejor variable de clasificación de los pacientes graves.
- Incrementar la n de los pacientes consiguiendo aumentar el número de pacientes con lectura radiológica (principal factor limitante).
- Explorar nuevas formas de captar la contribución de las manifestaciones sistémicas al índice.

Igualmente, el curso de los pacientes estudiados durante los próximos años será otra forma de valorar evolutivamente la capacidad de discriminación y utilidad de este índice preliminar para la identificación de biomarcadores.

CONCLUSIONES

Los resultados de nuestra investigación nos han conducido a realizar las siguientes conclusiones:

- Nuestros datos reflejan que el tratamiento con FAME en pacientes con artritis de inicio durante los dos primeros años de la enfermedad se ha intensificado de manera gradual durante los últimos 15 años. Este abordaje ha permitido mejores resultados en el control de la enfermedad pero obstaculiza la identificación de biomarcadores de gravedad.
- Hasta el momento no existe ningún índice que permita estimar la gravedad de la AR de forma temprana en los dos primeros años de seguimiento y que se adapte a la estrategia actual de manejo terapéutico de la enfermedad.
- Una vez realizado el desarrollo del índice de gravedad, se ha obtenido una solución factorial con sentido clínico. Según esta solución, la estructura del constructo “gravedad” está formada por 3 dominios que son actividad, lesión estructural y manifestaciones sistémicas. Esta estructura mantiene unas buenas propiedades psicométricas, con un 86% de la varianza total y adecuada consistencia interna (0,769 para el primer dominio y 0,481 para el segundo, aunque éste último sólo contiene dos ítems).
- A pesar de la plausibilidad biológica de esta estructura y de sus adecuadas propiedades psicométricas, esta solución factorial no clasifica correctamente más que al 61% de los pacientes en función de su gravedad, establecida por un criterio externo. De hecho, clasifica correctamente al 55% de los pacientes leves y al 63% de los graves según el tiempo de tratamiento con FAME.
- Dadas las buenas propiedades psicométricas de la solución factorial obtenida, los resultados del análisis discriminante podrían explicarse por el carácter

“blando” (poco rotundo) de la variable de clasificación utilizada. Es decir, la clasificación de los pacientes en diferentes niveles de gravedad en función del tiempo de tratamiento con FAME, no deja de ser una variable artificial, y con cierto carácter de subjetividad, en la que no sólo interviene la situación clínica del paciente, sino también la variabilidad propia del médico, y otros factores dependientes del paciente y no siempre relacionados con la gravedad de la enfermedad, como posible toxicidad o intolerancia farmacológica, comorbilidad, características de la personalidad o deseo de embarazo.

- La ponderación negativa de las manifestaciones sistémicas podría explicarse por las características clínicas y de tratamiento de los cuatro pacientes con este tipo de afectación clínica.
- El índice propuesto sólo debe considerarse una aproximación preliminar ya que necesita ser evaluado con otras variables de clasificación de los pacientes más robustas y depuradas.
- Por consiguiente, es necesario seguir avanzando en la obtención de un patrón externo de gravedad más robusto, así como evaluar la evolución de los pacientes estudiados en los próximos años.

BIBLIOGRAFÍA

1. Carmona L, Cross M, Williams B, Lassere M, March L. Rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. 2010;24(6):733-45.
2. Firestein GS. Evolving concepts of rheumatoid arthritis. *Nature*. 2003;423(6937):356-61.
3. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. *Lancet*. 2010;376(9746):1094-108.
4. Combe B, Landewe R, Daien CI, Hua C, Aletaha D, Alvaro-Gracia JM, et al. 2016 update of the EULAR recommendations for the management of early arthritis. *Ann Rheum Dis*. 2017;76:948-959.
5. Sanmarti R, Garcia-Rodriguez S, Alvaro-Gracia JM, Andreu JL, Balsa A, Caliz R, et al. [2014 update of the Consensus Statement of the Spanish Society of Rheumatology on the use of biological therapies in rheumatoid arthritis]. *Reumatol Clin*. 2015;11:279-94.
6. Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol*. 2016;68(1):1-26.
7. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet*. 2004;364(9430):263-9.
8. Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, van Zeben D, Kerstens PJ, Hazes JM, et al. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized, controlled trial. *Arthritis Rheum*. 2005;52(11):3381-90.

9. Verstappen SM, Jacobs JW, van der Veen MJ, Heurkens AH, Schenk Y, ter Borg EJ, et al. Intensive treatment with methotrexate in early rheumatoid arthritis: aiming for remission. Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA, an open-label strategy trial). *Ann Rheum Dis.* 2007;66(11):1443-9.
10. Boers M. Understanding the window of opportunity concept in early rheumatoid arthritis. *Arthritis Rheum.* 2003;48(7):1771-4.
11. van Nies JA, Krabben A, Schoones JW, Huizinga TW, Kloppenburg M, van der Helm-van Mil AH. What is the evidence for the presence of a therapeutic window of opportunity in rheumatoid arthritis? A systematic literature review. *Ann Rheum Dis.* 2014;73(5):861-70.
12. van Nies JA, Tsonaka R, Gaujoux-Viala C, Fautrel B, van der Helm-van Mil AH. Evaluating relationships between symptom duration and persistence of rheumatoid arthritis: does a window of opportunity exist? Results on the Leiden Early Arthritis Clinic and ESPOIR cohorts. *Ann Rheum Dis.* 2015;74(5):806-12.
13. Bakker MF, Jacobs JW, Welsing PM, Vreugdenhil SA, van Booma-Frankfort C, Linn-Rasker SP, et al. Early clinical response to treatment predicts 5-year outcome in RA patients: follow-up results from the CAMERA study. *Ann Rheum Dis.* 2011;70(6):1099-103.
14. Descalzo MA, Carbonell J, Gonzalez-Alvaro I, Sanmarti R, Balsa A, Hernandez-Barrera V, et al. Effectiveness of a clinical practice intervention in early rheumatoid arthritis. *Arthritis Care Res.* 2012;64(3):321-30.
15. Quinn MA, Emery P. Are early arthritis clinics necessary? *Best Pract Res Clin Rheumatol.* 2005;19(1):1-17.
16. Young A. What have we learnt from early rheumatoid arthritis cohorts? *Best Pract Res Clin Rheumatol.* 2009;23(1):3-12.
17. Verstappen SM, Bakker MF, Heurkens AH, van der Veen MJ, Kruize AA, Geurts MA, et al. Adverse events and factors associated with toxicity in patients with early

rheumatoid arthritis treated with methotrexate tight control therapy: the CAMERA study. *Ann Rheum Dis.* 2010;69(6):1044-8.

18. Huizinga TW. Personalized medicine in rheumatoid arthritis: is the glass half full or half empty? *J Intern Med.* 2015;277(2):178-87.

19. Gonzalez-Alvaro I, Ortiz AM, Seoane IV, Garcia-Vicuna R, Martinez C, Gomariz RP. Biomarkers predicting a need for intensive treatment in patients with early arthritis. *Curr Pharm Des.* 2015;21(2):170-81.

20. Verstappen SM, Lunt M, Bunn DK, Scott DG, Symmons DP. In patients with early inflammatory polyarthritis, ACPA positivity, younger age and inefficacy of the first non-biological DMARD are predictors for receiving biological therapy: results from the Norfolk Arthritis Register. *Ann Rheum Dis.* 2011;70(8):1428-32.

21. Gonzalez-Alvaro I, Ortiz AM, Alvaro-Gracia JM, Castaneda S, Diaz-Sanchez B, Carvajal I, et al. Interleukin 15 levels in serum may predict a severe disease course in patients with early arthritis. *PloS one.* 2011;6(12):e29492.

22. Emery P, Durez P, Dougados M, Legerton CW, Becker JC, Vratsanos G, et al. Impact of T-cell costimulation modulation in patients with undifferentiated inflammatory arthritis or very early rheumatoid arthritis: a clinical and imaging study of abatacept (the ADJUST trial). *Ann Rheum Dis.* 2010;69(3):510-6.

23. Saleem B, Mackie S, Quinn M, Nizam S, Hensor E, Jarrett S, et al. Does the use of tumour necrosis factor antagonist therapy in poor prognosis, undifferentiated arthritis prevent progression to rheumatoid arthritis? *Ann Rheum Dis.* 2008;67(8):1178-80.

24. van der Heijde DM, van 't Hof M, van Riel PL, van de Putte LB. Development of a disease activity score based on judgment in clinical practice by rheumatologists. *J Rheumatol.* 1993;20(3):579-81.

25. van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis:

first step in the development of a disease activity score. *Ann Rheum Dis.* 1990;49(11):916-20.

26. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum.* 1995;38(1):44-8.

27. van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum.* 1998;41(10):1845-50.

28. Wells G, Becker JC, Teng J, Dougados M, Schiff M, Smolen J, et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. *Ann Rheum Dis.* 2009;68(6):954-60.

29. Castrejon I, Ortiz AM, Garcia-Vicuna R, Lopez-Bote JP, Humbria A, Carmona L, et al. Are the C-reactive protein values and erythrocyte sedimentation rate equivalent when estimating the 28-joint disease activity score in rheumatoid arthritis? *Clin Exp Rheumatol.* 2008;26(5):769-75.

30. Ahlmen M, Svensson B, Albertsson K, Forslind K, Hafstrom I. Influence of gender on assessments of disease activity and function in early rheumatoid arthritis in relation to radiographic joint damage. *Ann Rheum Dis.* 2010;69(1):230-3.

31. Sokka T, Toloza S, Cutolo M, Kautiainen H, Makinen H, Gogus F, et al. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. *Arthritis Res Ther.* 2009;11(1):R7.

32. Castrejon Fernandez I, Martinez-Lopez JA, Ortiz Garcia AM, Carmona Ortells L, Garcia-Vicuna R, Gonzalez-Alvaro I. [Influence of gender on treatment response in a

cohort of patients with early rheumatoid arthritis in the area 2 of Madrid]. *Reumatol Clin.* 2010;6(3):134-40.

33. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford).* 2003;42(2):244-57.

34. Leeb BF, Andel I, Sautner J, Bogdan M, Maktari A, Nothnagl T, et al. Disease activity measurement of rheumatoid arthritis: Comparison of the simplified disease activity index (SDAI) and the disease activity score including 28 joints (DAS28) in daily routine. *Arthritis Rheum.* 2005;53(1):56-60.

35. Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol.* 2005;23(5 Suppl 39):S100-8.

36. Castrejon I, Carmona L, Ortiz AM, Belmonte MA, Martinez-Lopez JA, Gonzalez-Alvaro I. Development and validation of a new disease activity index as a numerical sum of four variables in patients with early arthritis. *Arthritis Care Res.* 2013;65(4):518-25.

37. Knevel R, Krabben A, Brouwer E, Posthumus MD, Wilson AG, Lindqvist E, et al. Genetic variants in IL15 associate with progression of joint destruction in rheumatoid arthritis: a multicohort study. *Ann Rheum Dis.* 2012;71(10):1651-7.

38. Krabben A, Huizinga TW, Mil AH. Biomarkers for radiographic progression in rheumatoid arthritis. *Curr Pharm Des.* 2015;21(2):147-69.

39. Gonzalez-Alvaro I, Descalzo MA, Carmona L. Trends towards an improved disease state in rheumatoid arthritis over time: influence of new therapies and changes in management approach: analysis of the EMECAR cohort. *Arthritis Res Ther.* 2008;10(6):R138.

40. Lamana A, Balsa A, Rueda B, Ortiz AM, Nuno L, Miranda-Carus ME, et al. The TT genotype of the STAT4 rs7574865 polymorphism is associated with high disease activity and disability in patients with early arthritis. *PloS one*. 2012;7(8):e43661.
41. Martinez C, Ortiz AM, Juarranz Y, Lamana A, Seoane IV, Leceta J, et al. Serum levels of vasoactive intestinal peptide as a prognostic marker in early arthritis. *PloS one*. 2014;9(1):e85248.
42. Maese J, Garcia De Yebenes MJ, Carmona L, Hernandez-Garcia C. Management of rheumatoid arthritis in Spain (emAR II). Clinical characteristics of the patients. *Reumatol Clin*. 2012;8(5):236-42.
43. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCW. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* (2010);19(4):539–49.
44. Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez-Olivo MA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. *CMAJ*. 2009;181(11):787-96.
45. Anderson JJ, Wells G, Verhoeven AC, Felson DT. Factors predicting response to treatment in rheumatoid arthritis: the importance of disease duration. *Arthritis Rheum*. 2000;43(1):22-9.
46. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis Rheum*. 2011;63(3):573-86.
47. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum*. 1988;31(3):315-24.

48. Verpoort KN, van Dongen H, Allaart CF, Toes RE, Breedveld FC, Huizinga TW. Undifferentiated arthritis--disease course assessed in several inception cohorts. *Clin Exp Rheumatol*. 2004;22(5 Suppl 35):S12-7.
49. Waltz M. The disease process and utilization of health services in rheumatoid arthritis: the relative contributions of various markers of disease severity in explaining consumption patterns. *Arthritis Care Res*. 2000;13(2):74-88.
50. Lassere MN, van der Heijde D, Johnson KR, Boers M, Edmonds J. Reliability of measures of disease activity and disease damage in rheumatoid arthritis: implications for smallest detectable difference, minimal clinically important difference, and analysis of treatment effects in randomized controlled trials. *J Rheumatol*. 2001;28(4):892-903.
51. Leeb BF, Sautner J, Andel I, Rintelen B. SACRAH: a score for assessment and quantification of chronic rheumatic affections of the hands. *Rheumatology (Oxford)*. 2003;42(10):1173-8.
52. Chiou CF, Suarez-Almazor ME, Sherbourne CD, Chang CH, Reyes C, Dylan M, et al. Development and validation of a preference weight multiattribute health outcome measure for rheumatoid arthritis. *J Rheumatol*. 2006;33(12):2409-11.
53. Landewe R, van der Heijde D. The validity of a rheumatoid arthritis medical records-based index of severity compared with the DAS28. *Arthritis Res Ther*. 2006;8(3):107.
54. Harrison MJ, Hassell A, Dawes PT, Scott DL, Knight SM, Davis MJ, et al. The overall status in rheumatoid arthritis (OSRA) measure--further evidence to support its use in clinical practice. *Rheumatology (Oxford)*. 2007;46(5):849-55.
55. Gossec L, Dougados M, Rincival N, Balanescu A, Boumpas DT, Canadello S, et al. Elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: a EULAR initiative. *Ann Rheum Dis*. 2009;68(11):1680-5.

56. Wolfe F, Michaud K, Pincus T. A composite disease activity scale for clinical practice, observational studies, and clinical trials: the patient activity scale (PAS/PAS-II). *J Rheumatol.* 2005;32(12):2410-5.
57. Wolfe F, Michaud K, Simon T. Can severity be predicted by treatment variables in rheumatoid arthritis administrative data bases? *J Rheumatol.* 2006;33(10):1952-6.
58. Vinet E, Kuriya B, Widdifield J, Bernatsky S. Rheumatoid arthritis disease severity indices in administrative databases: a systematic review. *J Rheumatol.* 2011;38(11):2318-25.
59. Cabral Danielle KJN, Weinblatt Michael E, Ting Gladys, Avorn Jerry, Solomon Daniel H Development and Assessment of Indicators of Rheumatoid Arthritis Severity: Results of a Delphi Panel. *Arthritis & Rheumatism (Arthritis Care Res).* 2005;53(1):61-6.
60. Ting G, Schneeweiss S, Katz JN, Weinblatt ME, Cabral D, Scranton RE, et al. Performance of a rheumatoid arthritis records-based index of severity. *J Rheumatol.* 2005;32(9):1679-87.
61. Sato M, Schneeweiss S, Scranton R, Katz JN, Weinblatt ME, Avorn J, et al. The validity of a rheumatoid arthritis medical records-based index of severity compared with the DAS28. *Arthritis Res Ther.* 2006;8(3):R57.
62. Ting G, Schneeweiss S, Scranton R, Katz JN, Weinblatt ME, Young M, et al. Development of a health care utilisation data-based index for rheumatoid arthritis severity: a preliminary study. *Arthritis Res Ther.* 2008;10(4):R95.
63. Desai RJ, Solomon DH, Weinblatt ME, Shadick N, Kim SC. An external validation study reporting poor correlation between the claims-based index for rheumatoid arthritis severity and the disease activity score. *Arthritis Res Ther.* 2015;17:83.
64. Kehoe J. Basic item analysis for multiple-choice tests. *PARE* 1995;4(10).
65. Smolen JS, Aletaha D. Challenges of predicting treatment response in patients with rheumatoid arthritis. *Nat Clin Pract Rheumatol.* 2005;1(2):62-3.

66. Leon L, Redondo M, Garcia-Vadillo A, Perez-Nieto MA, Rodriguez-Rodriguez L, Jover JA, et al. Influence of patient personality in the treatment of rheumatoid arthritis. *Rheumatol Int.* 2016;36(11):1549-55.