



INDICADORES BIBLIOMÉTRICOS EN LA EVALUACIÓN DE LA INVESTIGACIÓN

BRUNO MALTRÁS (*),
MIGUEL A. QUINTANILLA (**)
JAVIER VIDAL (***)

INTRODUCCIÓN

Una de las características de los proyectos de evaluación institucional impulsados por el Consejo de Universidades en España, desde el Plan Experimental de 1992, es que se ha pretendido siempre adoptar una perspectiva integral en la evaluación de la universidad, utilizando una metodología homogénea para la evaluación de todas sus actividades: docencia, investigación y prestación de otros servicios. Esta es la razón de que en la *Guía de Evaluación* que editó el Consejo de Universidades (1996) se incluyera un documento preparado explícitamente para la evaluación institucional de la investigación en los Departamentos e Institutos Universitarios¹. La lógica interna del modelo utilizado y algunos de los problemas derivados de su aplicación, se tratan en otra contribución de este mismo volumen, por lo que no insistiremos aquí sobre ello. Nuestro objetivo se ciñe, por el contrario, a explicar algunos

aspectos técnicos del uso de indicadores bibliométricos en el modelo de evaluación de la investigación que estamos comentando.

En el *Protocolo de Evaluación de la Investigación* se utilizan una serie de indicadores cuantitativos (en buena parte inspirados en los trabajos de Vidal, 1996 y Maltrás, 1996) que permiten describir de forma sintética algunos aspectos relevantes de la *actividad* científica que se realiza en los departamentos universitarios: nivel de actividad en proyectos, contratos, tesis, indicadores de éxito, evolución y concentración de esa actividad, etc. Por otra parte se utilizan también algunos indicadores para evaluar *resultados*. Entre ellos figuran los indicadores bibliométricos, a los que sólo se dedican un par de tablas. Para rellenar estas tablas se pide a los Comités de Autoevaluación que indiquen los tipos de publicaciones que han realizado los miembros del departamento durante el periodo evaluado y que los clasifiquen en

(*) Universidad de Salamanca.

(**) Universidad de Salamanca. Miembro del Comité Técnico del Plan Nacional de Evaluación de la Calidad de las Universidades.

(***) Universidad de León.

(1) El documento fue inicialmente redactado por Miguel A. Quintanilla y Javier Vidal. La versión final se benefició de las sugerencias de los miembros del Comité Nacional de Evaluación, en especial de Miguel Valcárcel y José María Luxán. Pero no se publicó de forma independiente, sino fragmentado y adaptado para poder ser incluido en la *Guía de Evaluación* editada por el Consejo.

tres grupos (A, B y C) según el juicio que les merezca su nivel de *calidad*. Como complemento, en el caso de las áreas de ciencias experimentales, tecnológicas y de la salud, se ha proporcionado a los comités un listado con la clasificación de las revistas del SCI en tres grupos, de modo que los comités puedan utilizar esta clasificación como referencia para su propia evaluación (Quintanilla y Maltrás, 1996).

En la *Guía* se insiste en que la evaluación debe entenderse de forma comprensiva e integral, y no reducirse al mero recuento y clasificación de publicaciones. Sin embargo, somos conscientes de que la utilización de indicadores bibliométricos suele suscitar más interés y más polémica de la que en principio cabría esperar. Por otra parte, el uso de estos indicadores en el protocolo específico que estamos comentando contrasta con algunas prácticas habituales en la evaluación de investigación que pueden desvirtuar los objetivos y la metodología utilizada en el PNECU. Finalmente la aplicación de esta metodología en la primera fase del Plan nos ha permitido detectar algunos problemas y dificultades que nos proponemos aclarar.

A continuación explicaremos las características de la clasificación en tres grupos de las revistas del SCI, las limitaciones y el alcance que este instrumento puede tener para la evaluación institucional. Después explicaremos los intentos fallidos que hemos realizado para obtener un instrumento similar en el caso de las ciencias sociales y humanas.

LOS INDICADORES BIBLIOMÉTRICOS EN LAS CIENCIAS EXPERIMENTALES

Como referencia para la evaluación de la producción científica de cada comité de autoevaluación se ha elaborado una clasificación de las revistas del SCI en tres grandes grupos. Nos proponemos ahora exponer la metodología seguida para la elaboración de esta lista y explicar su alcance y sus limitaciones para la evaluación institucional.

Para elaborar la lista se han empleado dos fuentes de datos complementarias: el *Science Citation Index* y el *Journal Citation Reports*, ambas producidas por el *Institute for Scientific Information*.

La base de datos del *Science Citation Index* (SCI) cataloga anualmente más de medio millón de documentos publicados en unas tres mil cuatrocientas revistas científicas de circulación internacional. El SCI incluye las referencias bibliográficas de todos los documentos publicados por las revistas que cubre, es decir, no realiza selección previa alguna. Tiene, además, dos características que hacen de ella una base de datos única en cuanto a su utilidad para estudios bibliométricos: por un lado, la referencia bibliográfica de cada documento contiene todos los autores del documento y todas sus instituciones respectivas; por otro lado, recoge todas las citas realizadas por cada documento que indica y dirigidas a trabajos anteriores, lo que le lleva a procesar anualmente varios millones de citas. Aunque su finalidad principal es la de ofrecer información a los propios científicos en relación con la literatura científica mundial, desde su creación se ha constituido como la base de datos principal para la realización de estudios bibliométricos, a pesar de ciertos sesgos geográficos en su cobertura que son bien conocidos. Puede asumirse que es una muestra amplia y representativa de lo que ha dado en llamarse el *main stream* o *corriente principal de la ciencia*, es decir, la ciencia con difusión y atención internacional (Vidal, 1993).

La segunda fuente de datos empleada es el *Journal Citation Reports* (JCR). Su contenido es una serie de datos básicos acerca de las revistas científicas, tales como número de documentos publicados por cada una, número de citas recibidas, etc. así como ciertos indicadores calculados sobre la base de estos datos básicos. De estos indicadores el más conocido es el llamado *factor de impacto*. Es importante señalar que en la base JCR se pueden distin-

guir dos grupos de revistas. Por una parte está el conjunto de revistas que se catalogan en el SCI. Éstas constituyen el núcleo de *revistas fuente* o *revistas citantes*, de las que se obtienen todas las referencias bibliográficas con sus citas. Pero, por otra parte, en los informes del JCR se recogen también aquellas otras revistas cuyos artículos han sido significativamente citados por las revistas fuente pero que no están incluidas en la base del SCI. Podemos denominar a este grupo como el de las *revistas no-fuente* o *sólo citadas*. Aunque estas revistas no figuran en la base SCI, sin embargo, en el JCR sí se recogen sus datos básicos, el número de citas recibidas y los indicadores de factor de impacto, etc. (con la desventaja de que, para estas revistas, no se computan las citas que en ellas se hacen a documentos publicados por ellas mismas).

El factor de impacto (FI) es la medida más ampliamente conocida y mencionada con relación a la importancia de las revistas. Sin embargo, y a pesar de su sencillez, es frecuentemente mal entendido y aplicado. *El FI es la media ponderada de citas que han recibido durante un año determinado los artículos publicados por una revista durante los dos años anteriores y que provienen de las revistas fuente del SCI.* Se trata de una estimación de la atención que han despertado los artículos de una revista y de la influencia que han podido tener en el proceso científico posterior. La estimación se basa en una parte de las citas totales que reciben los artículos, ya que sólo se toman las citas emitidas *durante un año* por las *revistas fuente* (y se descartan las citas que han recibido o recibirán esos artículos en otros años o por otras revistas). El FI se puede utilizar como un indicador indirecto de calidad de la producción científica, pero es preciso saber cuál es su alcance, su significación y sus limitaciones (Maltrás, 1996).

En primer lugar el FI es un indicador referido directamente a una *revista* científica, no a los trabajos individuales publicados en esa revista. Por consiguiente, cuando se

utiliza como un indicador de la calidad de un trabajo científico individual se puede estar cometiendo un grave error.

En segundo lugar el FI no es propiamente una medida de calidad científica sino de visibilidad, influencia o prestigio de una revista en el conjunto de la comunidad científica internacional. El hecho de que se pueda emplear como un indicador indirecto de calidad se debe a las prácticas que, se supone, siguen los científicos cuando publican un trabajo y los editores cuando lo aceptan para su publicación. En último término estas prácticas se basan en el juicio de pares sobre la calidad de cada trabajo científico y son determinantes del prestigio de la revista. Como norma general, cabe suponer que una revista de prestigio científico tendrá una fuerza de atracción mayor sobre los autores que publican en su área, lo que significa que aumentará la competencia para publicar en ella, lo que a su vez conduce a elevar los estándares de rigor y calidad que aplican los evaluadores internos de la revista para aceptar nuevos trabajos, etc.

En cualquier caso, la aplicación indiscriminada del FI como medida objetiva de la calidad de las revistas presenta varios problemas bien conocidos. Uno de ellos es el de la dependencia respecto a las costumbres de citar de cada área científica, costumbres que, además, pueden cambiar con el tiempo. Por ejemplo: el que una revista tenga un FI superior a otra de otra área puede esconder el hecho de que la segunda sea la mejor de su disciplina y la primera sea una revista de nivel medio dentro de la suya.

Otro problema frecuentemente olvidado es el de la métrica de los valores del FI: no se puede asumir sin más que si una revista tiene un FI doble que otra se pueda por ello valorar el doble que ésta. No hace falta entrar en muchos detalles para comprender este punto, pero sí puede ser útil decir aquí que las distribuciones de los valores del FI en las distintas áreas presentan diferencias significativas en cuanto a los

máximos y mínimos y también en cuanto a las diferencias y frecuencias intermedias de los valores.

Para afrontar estos problemas hemos propuesto en otras ocasiones utilizar dos indicadores normalizados de calidad, para cada área científica, basados en el FI. Se trata de los indicadores P10 (puntuación decílica) y %SUP (decil superior). Ver Maltrás y Quintanilla, 1995.

El indicador P10 se construye ordenando todos los documentos de un área de la base completa SCI por el factor de impacto de las revistas en que se publican (excluidas las revistas con FI=0) y dividiendo el total en diez partes o deciles. A partir de aquí se sustituye el FI de cada revista por la puntuación decílica que le corresponde (1 al decil con FI más bajo, 10 al más alto). Cuando los documentos de una revista se distribuyen en dos o más deciles, se asigna a la revista la puntuación que le corresponda según la proporción de documentos que tenga en cada decil. Por construcción de la escala, la puntuación media para el total de documentos de cualquier área es 5,5 que coincide con el valor teórico de la mediana.

El indicador %SUP representa el porcentaje de documentos de un grupo que se incluyen en el decil 10 de la producción mundial de la correspondiente área científica. Se puede interpretar como un indicador del nivel de excelencia.

La clasificación de revistas del SCI que hemos realizado para el PNECU se basa en estos indicadores. Para construirla partimos de la clasificación en áreas científicas de las revistas del SCI (en torno a 140 áreas diferentes, aunque el número puede variar según los años), y distribuimos las revistas de cada área en tres grupos, según el FI y el número de documentos publicados por

cada revista, de acuerdo con los siguientes criterios:

1. *El primer grupo (A)* contiene las revistas con un FI más elevado dentro de su área, hasta completar el 10% del total de los documentos del área ($10 \geq P10 > 9$).

2. *En el segundo (B)* se incluyen las revistas con más impacto de las restantes hasta completar otro 40% de los documentos del área ($9 \geq P10 > 5$).

3. *El tercer grupo (C)*, constituido por el resto de las revistas del área indexadas por el SCI, suma el 50% de los documentos publicados en las revistas con menor FI ($5 \geq P10 \geq 1$).

4. Cuando una revista pertenece a varias áreas científicas (lo que ocurre en la cuarta parte de la base, aproximadamente) y no coincide en todas ellas el grupo al que se asigna siguiendo los criterios anteriores, se la incluye en el grupo con FI más elevado.

Esta clasificación inicial la realizamos con los datos del factor de impacto correspondientes a los años 1992, 1990 y 1988. Para confeccionar la lista final se añadió el siguiente criterio complementario:

5. Si a una misma revista le corresponden grupos diferentes en distintos periodos, se incluye en el grupo con factor de impacto más elevado.

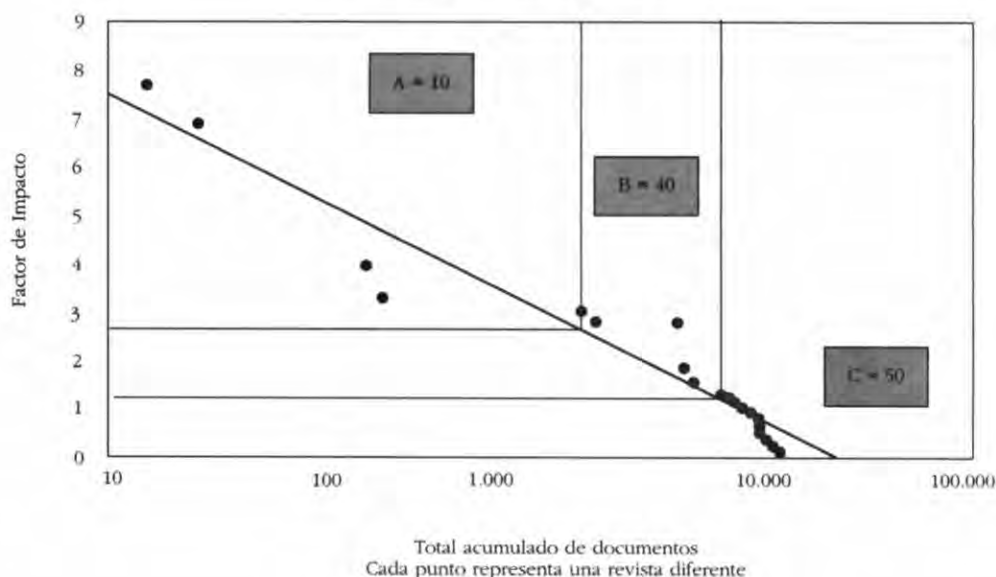
Una vez establecidos los grupos de revistas, la evaluación de un determinado departamento se realiza identificando la proporción de documentos que los investigadores del departamento publican en las revistas de cada grupo durante el periodo considerado. El grupo A se puede interpretar como un grupo de excelencia, que contiene como mínimo ² el 10 % superior de

(2) La precisión no puede ser absoluta, ya que la escala utilizada puede ser demasiado gruesa en algunas áreas con un número reducido de revistas científicas y, por otra parte, los criterios 4 y 5 utilizados hacen que los grupos superiores sean en la práctica algo más numerosos que lo que correspondería a una aplicación estricta de los criterios de la puntuación P10.

cada área en la producción mundial; si un documento concreto se clasifica como de grupo A significa que aproximadamente el 90 % de los documentos del área han aparecido en revistas con un FI menor que la suya. En el caso del grupo B, significa que aproximadamente el 50 % de las publicaciones del área están en revistas con menor FI. El gru-

po C incluye al resto de las publicaciones. La estimación de la calidad de un departamento cualquiera se realiza a partir del análisis de la distribución de sus publicaciones en esos tres grupos. En el gráfico se puede observar la distribución correspondiente a las 28 revistas del área de Química Orgánica del SCI en 1992.

*Distribución de artículos y revistas por factor de impacto
Área de Química Orgánica*



El uso de esta lista de referencia para evaluar la calidad de la producción científica de un departamento o de una institución tiene ventajas, pero también limitaciones. Entre las ventajas, podemos señalar las siguientes:

- Se simplifica mucho el proceso de evaluación y se posibilita su realización por parte de los propios agentes interesados. Éstos no han de disponer de un entrenamiento específico en técnicas bibliométricas. El único dato

que hay que manejar es el nombre de las revistas en las que se ha publicado la producción científica de un departamento en el periodo considerado.

- Facilidad en la interpretación. La interpretación es sencilla, sitúa la producción de cualquier institución en relación con la producción mundial del área y con la de cualquier otra institución.
- Fiabilidad. Se pueden comparar directamente diversos grupos dentro de una institución o diferentes insti-

tuciones. Los resultados son homogéneos y significan lo mismo en todas partes.

No obstante, el procedimiento también tiene sus *limitaciones*, que deben tenerse en cuenta. En especial las siguientes:

- La lista se ha confeccionado sobre la base de los FI correspondientes al periodo 1988-1992. Aunque la clasificación es bastante estable, a medida que el periodo analizado se aleje de estos años de referencia, la clasificación será menos fiable. Para evaluaciones futuras debe actualizarse la lista hasta cubrir por lo menos el bienio anterior al último año evaluado.
- Para que los resultados de la evaluación sean significativos, se necesita que la producción científica del departamento evaluado en el periodo considerado supere un número mínimo (nosotros utilizamos 15 documentos como límite mínimo).
- En la interpretación no debe olvidarse que lo que se evalúa con estos criterios no es la calidad intrínseca de la producción científica de uno o varios investigadores, sino más bien la posición de la institución o grupo (departamento, área de conocimiento, universidad, país, región, etc.) en la comunidad científica internacional.
- Sólo se utilizan las revistas indizadas por el SCI, no todas las que aparecen en el JCR. Se supone que aquéllas son suficientemente representativas de cada área científica y que añadir más ítems no haría variar drásticamente el resultado de la clasificación, puesto que, por lo general, la base SCI incluye las revistas con más impacto de cada área. Sin embargo, en ocasiones puede suceder que alguna de las revistas

que figuran en el JCR con un FI elevado dentro de su área (y que sean reconocidas como importantes) no sea indizada por una u otra razón en la base SCI.

- Debe tenerse en cuenta también que el grado de cobertura de la producción científica española por el SCI es muy diferente según las áreas científicas. Por eso en la evaluación institucional debe utilizarse la clasificación de las revistas SCI como una referencia, pero no como el único instrumento de evaluación.

Aparte de estas limitaciones, que podemos considerar genéricas para cualquier proceso de evaluación que utilice indicadores bibliométricos, existen algunos *otros problemas específicos* que debemos comentar.

- La cobertura del JCR y del SCI debería ser inclusiva, pero en la práctica algunas revistas del SCI no pudimos localizarlas en el JCR (alrededor de 50 para todo el periodo). Esto pudo deberse a los cambios de título y de abreviatura, pero también a algún fallo de cobertura de la base en la versión CDROM que utilizamos.
- En un caso particular se produjo una curiosa confusión en las abreviaturas de las revistas: la abreviatura ACTA MATH DURJSHOLM correspondía a la revista *Acta Mathematica* (la abreviatura es más larga que el nombre completo), mientras ACTA MATH correspondía a *Acta Mathematica Hungarica*. Las personas encargadas de volcar los datos intercambiaron esas revistas y la primera, de reconocida calidad y con un elevado FI, pasó a figurar en un grupo inferior. Este error fue detectado durante la aplicación por expertos, aunque por suerte afectaba a un solo documento español publicado en esa revista durante todo el período analizado.

Globalmente pensamos que el uso de la lista de revistas SCI clasificadas en tres grupos, ha resultado útil en los procesos de evaluación institucional de las universidades que se han centrado en departamentos de ciencias experimentales. Así lo han entendido muchos comités de autoevaluación, aunque ha habido también bastantes desviaciones en el uso de este instrumento. Las dos desviaciones más comunes han consistido o bien en aplicar mecánicamente la clasificación SCI sin entrar en más detalles o bien —lo que todavía es peor— en aplicar directamente el FI para obtener un listado organizado de la producción científica de todos los departamentos de una universidad. En nuestra opinión la actitud correcta consiste, como hemos dicho, en usar esta clasificación como una referencia para la evaluación, siendo conscientes de todas sus limitaciones y sabiendo apreciar, en su justo término, sus indudables ventajas.

LOS INDICADORES BIBLIOMÉTRICOS EN CIENCIAS SOCIALES, JURÍDICAS Y HUMANIDADES

Si los criterios de calidad de las publicaciones generan problemas en su aplicación a determinadas áreas de las ciencias naturales, muchos más problemas surgen en las ciencias sociales, jurídicas o humanidades.

Al igual que para el resto de las áreas, sólo las bases de datos del ISI (*Institute for Scientific Information*) disponen de indicadores de calidad de las publicaciones basados en citas. El ISI dispone de una base de datos de ciencias sociales, SSCI (*Social Science Citation Index*), y otra de humanidades, A&HCI (*Arts & Humanities Citation Index*). Pero la imposibilidad de utilizar estas bases de datos como criterio de calidad se apoya en el hecho de que no recogen de forma suficiente los trabajos

publicados por investigadores españoles en estas áreas. Para ilustrar esta idea se puede señalar que mientras el SCI recoge 4.969 documentos de la Universidad de Barcelona para el período 1980-1990, el SSCI recoge tan sólo 192, y el A&HCI 171 (Bellavista, 1993). Y si tenemos en cuenta que esta universidad tenía más de 2.000 profesores en Ciencias Sociales, Jurídicas y Humanidades en ese período, se puede deducir que estas bases de datos sólo recogen una pequeña parte de lo publicado en estas áreas.

Otra diferencia importante que existe entre estas áreas y las áreas experimentales y naturales es el tipo de documentos que se publican. Si bien es cierto que la mayor parte de los documentos se pueden clasificar en las categorías habituales (artículo, libro, volumen colectivo o actas de congreso) en muchas ocasiones hay dificultades para distinguir entre publicaciones de investigación, docentes o de divulgación. Por ejemplo, en el ámbito de las ciencias jurídicas, existen numerosas publicaciones, algunas de ellas diarias, en donde pueden aparecer desde estudios importantes a breves comentarios sobre sentencias o leyes. En las distintas áreas de filología, la publicación de una traducción, edición o adaptación de un texto puede ser parte de una gran labor académica o de investigación.

A pesar de este conjunto de dificultades conocidas, el Comité Técnico del PNECU nos encargó que exploráramos la posibilidad de obtener una clasificación de *medios de publicación* (revistas y editoriales) en los distintos ámbitos de la Ciencias Sociales, Jurídicas y Humanidades. Se intentaba que pudiera servir en estas áreas como referencia equivalente a la lista del SCI en las áreas de ciencias, utilizando como fuente primaria para esta clasificación los resultados de la 6ª evaluación del profesorado universitario (1994) que realizó la Comisión Nacional Evaluadora de la Actividad Investigadora (CNEAI) en los

campos 07 al 11. Estos campos son los correspondientes a las ciencias sociales, jurídicas y humanidades (Vidal y Quintanilla, 1996).

Por motivos de confidencialidad, sin embargo, la CNEAI sólo nos pudo facilitar la relación de las publicaciones (sólo el nombre de la revista o la editorial de cada registro) presentadas como mérito por los investigadores en su solicitud para obtener un tramo de investigación, indicando para cada ítem si correspondía a una solicitud concedida o denegada. No se pudo contar, por lo tanto, ni con el nombre del autor, ni con el área de conocimiento o campo científico al que pertenecía. Por consiguiente, la única información relevante de que se podía obtener era el número de veces que una determinada revista o editorial figuraba en solicitudes de tramos concedidos y denegados.

Los datos informatizados facilitados por la CNEAI se encontraban en formato texto, por lo que fue necesario realizar las transformaciones correspondientes para convertirla en una base de datos que finalmente incluyó 7.709 registros con los siguientes campos:

- Nombre de la publicación (revista, editorial, etc.).
- Tipo de publicación (revista, editorial comercial, entidad responsable de la publicación cuando no es una editorial comercial, acta de congreso, otros).
- Atributo «concedida» o «denegada» de la solicitud de evaluación en la que se incluyó ese ítem.

En la siguiente tabla se recoge el resumen de la información contenida en la base.

TIPO	ÍTEMS CORRESPONDIENTES A TRAMOS			TOTAL ÍTEMS DIFERENTES
	CONCEDIDOS	DENEGADOS	TOTAL	
Acta	73	101	174	160
Editorial	1.630	622	2.252	669
Entidad	1.703	810	2.513	632
Revista	1.956	810	2.766	1.280
Otros	2	1	3	4
Total	5.365	2.344	7.709	2.745

Se han realizado varias pruebas para encontrar un indicador que permitiese realizar una clasificación en tres grupos equivalente a la realizada con el SCI, explicada anteriormente. El supuesto básico es que aquellas revistas, editoriales, etc. en que aparecen con mayor frecuencia publicaciones de investigadores cuyas solicitudes han sido evaluadas positivamente pueden considerarse en principio de mayor calidad.

Para obtener una clasificación en tres grupos se ha calculado, por una parte, el porcentaje de tramos concedidos a una publicación sobre el total de tramos concedidos para su tipo de publicación (p.e.,

revistas) y, por otra parte, el mismo porcentaje para los tramos denegados. El indicador final utilizado ha sido la diferencia entre ambos porcentajes, que hemos llamado *Difer*.

Se han probado distintos procedimientos para establecer estadísticamente los grupos buscados, pero sin ningún resultado válido. Por ejemplo, al aplicar un análisis de clusters basado en las diferencias de medias da como resultado un grupo A de 849 revistas y un grupo B de 383, siendo significativas las diferencias entre los grupos. Sin embargo, este resultado no es adecuado a los objetivos buscados.

Como alternativa, se probaron dos nuevos criterios, teniendo en cuenta que la distribución del indicador utilizado es

una distribución normal. Los criterios para la asignación de grupos son los siguientes:

CRITERIO I	GRUPO	Criterios de asignación
	A	Difer > media + σ
B	Media + σ <= Difer = > media σ	
C	Resto	
CRITERIO II	A	Difer > media + σ
	B	Media + σ <= Difer = > media
	C	Resto

Utilizando estos criterios, se hizo un análisis con detenimiento de las publicaciones del tipo REVISTA, dado que son las que

más garantías ofrecen *a priori* (mayor número de ítems y mayor consistencia). Los resultados para este tipo han sido los siguientes:

N.º DE REVISTAS	CRITERIO I				
	1	2	3	TOTAL	%
1	64	0	0	64	5%
2	0	784	0	784	61%
3	0	332	100	432	34%
Total	64	1.116	100	1.280	100%
%	5%	87%	8%	100%	

Las distribuciones obtenidas parecen encontrarse dentro de los márgenes de lo esperable y el análisis de la varianza da como resultado diferencias significativas entre los grupos. Sin embargo, no podemos considerar el resultado satisfactorio para nuestros objetivos. Hay dos razones para esta conclusión:

En primer lugar, la imposibilidad de distinguir entre campos científicos, combinada con el tipo de distribución que hemos obtenido, resulta fatal para nuestros propósitos. En efecto, los ítems del grupo A (5 % del total) pertenecen todos a un reducido número de áreas científicas, de manera que la clasificación no nos permitiría discriminar niveles de calidad en la mayoría de las especialidades de las ciencias sociales y las humanidades.

Por otra parte, los datos que tenemos respecto a la distribución de los tramos concedidos y denegados por grandes áreas científicas nos hacen sospechar que para conseguir nuestros objetivos deberíamos disponer de información referida a varios años y convocatorias. En efecto, como puede comprobarse en la Tabla I, se producen grandes variaciones entre la convocatoria analizada y la anterior en el porcentaje de solicitudes denegadas y concedidas, por una parte, y entre los distintos campos, por otra.

CONCLUSIONES

Consideramos que los indicadores bibliométricos pueden ser un instrumento útil para la evaluación institucional de la

TABLA I
*Datos de dos convocatorias de la evaluación de la actividad investigadora.
 CNEAI, 1996*

CAMPO	SOLICITANTES				TRAMOS		
	CONCEDIDOS	CON. PARCIAL	DENEGADOS	% DENEGADOS	SOLICITADOS	CONCEDIDOS	% CONCEDIDOS
<i>5.ª evaluación 1993</i>							
7	237	12	98	39,36%	394	277	70,30%
8	93	14	55	51,40%	187	112	59,89%
9	169	5	30	17,24%	245	202	82,45%
10	200	12	70	33,02%	313	231	73,80%
11	168	8	117	66,48%	311	182	58,52%
Total	867	51	370	40,31%	1.450	1.004	69,24%
<i>6.ª evaluación 1994</i>							
7	307	9	238	75,32%	586	323	55,12%
8	207	6	119	55,87%	359	227	63,23%
9	243	3	75	30,49%	364	285	78,30%
10	269	7	102	36,96%	405	286	70,62%
11	419	15	109	25,12%	584	453	77,57%
Total	1.445	40	643	43,30%	2.298	1.574	68,49%

Diferencias entre las dos convocatorias

CAMPO	% SOLICITUDES DENEGADAS			% TRAMOS CONCEDIDOS		
	93	94	94-93	93	94	94-93
7	39,36%	75,32%	35,96	70,30%	55,12%	-15,19
8	51,40%	55,87%	4,47	59,89%	63,23%	3,34
9	17,24%	30,49%	13,25	82,45%	78,30%	-4,15
10	33,02%	36,96%	3,94	73,80%	70,62%	-3,18
11	66,48%	25,12%	-41,36	58,52%	77,57%	19,05
Total	40,31%	43,30%	2,99	69,24%	68,49%	-0,75

investigación. Pero debe ponerse el máximo cuidado en que se utilicen correctamente evitando tanto simplificaciones excesivas como inútiles refinamientos metodológicos que hacen impracticable su aplicación por los comités de autoevaluación.

La clasificación de las listas del SCI en tres grupos de acuerdo con los criterios ex-

puestos cumple todas las condiciones para ser un buen instrumento para la evaluación institucional. Pero no debe usarse ni para evaluar la actividad científica de los investigadores individuales ni para reducir la evaluación institucional a un simple cálculo automático de porcentajes de publicaciones en cada grupo. Por el contrario hay

que insistir en que sólo se trata de una referencia para la evaluación interna de la propia institución. Además, debe tenerse en cuenta siempre el hecho de que en determinadas áreas científicas la base SCI puede no ser tan representativa como en otras. En todo caso la lista debe actualizarse periódicamente.

Un caso especial es el amplio grupo de las Ciencias Sociales y Humanidades que en conjunto suponen más de la mitad del sistema universitario español, pero que no se pueden considerar, en modo alguno, bien representadas en las bases de datos internacionales del ISI. Los intentos realizados hasta el momento para construir un sistema de referencia para evaluación de la calidad en este campo hay que considerarlos fracasados. En el caso concreto que hemos expuesto, la conclusión es que la clasificación obtenida no tenía las garantías suficientes para ser utilizada dentro del Plan Nacional de Evaluación de la Calidad de las Universidades, siendo los motivos fundamentales el no haber podido separar la información por campos científicos y el disponer de una sola convocatoria. No debe olvidarse la repercusión que una clasificación de este tipo tendría en todos los ámbitos, lo que exige las máximas garantías en su elaboración.

Las dificultades para la elaboración de una clasificación del tipo de la aquí propuesta y la necesidad que existe de disponer de ella, nos lleva a sugerir que se inicie lo antes posible el estudio con las condiciones necesarias para poder hacerla con las garantías que requiere algo que vaya a tener gran repercusión. Para ello, se pueden utilizar el mismo tipo de datos analizados en este informe, pero con una distinción de campos suficientemente desagregada y con 2 ó 3 convocatorias diferentes. Se sugiere así mismo que, se utilice un procedimiento de validación externa por parte de los expertos en distintas materias. Así mismo, el propio PNECU puede ser un excelente sistema de validación a corto plazo.

BIBLIOGRAFÍA

- BELLAVISTA, J. et al.: *Política científica y tecnológica. Evaluación del I+D en la Universitat de Barcelona*. Barcelona, Universitat de Barcelona, 1993.
- COMISIÓN NACIONAL EVALUADORA DE LA ACTIVIDAD INVESTIGADORA: *Seis años de evaluaciones*. Madrid, MEC, 1996.
- CONSEJO DE UNIVERSIDADES: *Guía de Evaluación del Plan Nacional de Evaluación de la calidad de las Universidades*. Madrid, Consejo de Universidades, Secretaría General, 1996.
- MALTRÁS, B.: *Los indicadores bibliométricos en el estudio de la ciencia. Fundamentos conceptuales y aplicación a la política científica. Tesis de doctorado*. Salamanca, Universidad de Salamanca, Facultad de Filosofía, Departamento de Filosofía, Lógica y Filosofía de la Ciencia, 1996.
- MALTRÁS, B. y QUINTANILLA, M. A.: *Spain Indicators on Scientific Production, 1986-1991*. Madrid, CSIC, 1995.
- QUINTANILLA, M. A. y MALTRÁS, B.: *Lista de revistas SCI*. Programa Nacional de Evaluación Institucional de las Universidades. Madrid, Consejo de Universidades (1996).
- VIDAL, J.: «El Science Citation Index y la evaluación de la investigación en las universidades». *Política Científica*, 37, pp. 57-59, 1993.
- *La caja gris. Microanálisis de la actividad investigadora, su gestión y evaluación en una institución universitaria. Estudio del caso de la Universidad de León*. Salamanca, Ediciones Universidad de Salamanca, 1996.
- VIDAL, J. y QUINTANILLA, M. A.: *Clasificación de las publicaciones en las áreas de ciencias sociales y humanidades*. Informe al Consejo de Universidades, septiembre, 1996.