

# Indicadores bibliométricos e econométricos para a avaliação de instituições científicas\*

Ronald Rousseau

## INTRODUÇÃO

Este artigo é dividido em duas partes. A primeira descreve a avaliação da LUC, uma pequena universidade de Flandres, Bélgica, baseada em dados cienciométricos e tendo por principal objetivo avaliar a pesquisa com visibilidade internacional. A segunda parte mostra como um método econométrico pode ser usado para incluir na avaliação o ensino e o levantamento de recursos, entre outros aspectos. As duas abordagens mostram como um corpo teórico combinando bibliometria, ciencimetria e econometria pode ser aplicado a problemas concretos (Van Raan<sup>1</sup>, Egghe e Rousseau<sup>2</sup>). Portanto, a bibliometria não é uma solução em busca de um problema (Wallace<sup>3</sup>).

## PARTE A: UM ESTUDO CIENCIOMÉTRICO DAS PUBLICAÇÕES CIENTÍFICAS DA LUC

Este estudo, autorizado pelo reitor e pelo Conselho de Pesquisa da LUC, investiga a posição internacional da Escola de Ciência e Medicina da LUC (Rousseau<sup>4</sup>), abordando, em particular, o período de 1981 a 1993:

- a dimensão do potencial de todos os grupos de pesquisa, medida em termos de equivalência pessoa-ano (person-year equivalents, PYE) dedicada à pesquisa científica;

\* Trabalho apresentado no Seminário sobre Avaliação da Produção Científica, realizado em São Paulo pelo Projeto SciELO, de 4 a 6 de março de 1998. O autor agradece aos organizadores, em particular a Abel Packer pelo convite. Traduzido e publicado na *Ciência da Informação* com a permissão do autor.

- os resultados da pesquisa, em particular a posição internacional evidenciada pelas publicações em revistas científicas; especialmente, a proporção em que essas publicações foram citadas pela literatura profissional internacional;
- as observações que os cientistas fizeram sobre os métodos cienciométricos e os seus resultados.

Para poder comparar os resultados da LUC com os resultados de outras instituições, o método usado foi o mesmo empregado na avaliação de outras universidades flamengas (De Bruin et al.<sup>5</sup>, De Bruin, Moed e Spruyt<sup>6</sup>).

## Premissas, objetivos e advertências

A aplicação dos métodos cienciométricos para a avaliação da pesquisa científica (Moed et al.<sup>7</sup>, Oppenheim<sup>8</sup>) é baseada nas seguintes premissas:

- a) O progresso é alcançado mediante o trabalho de cientistas.
- b) Esses cientistas constroem seu trabalho sobre as obras de colegas e precursores de sua área.
- c) Os resultados desse trabalho são publicados
- d) Portanto, submetidos à apreciação de pares.
- e) Em suas publicações, os cientistas mostram como construíram seus trabalhos a partir de obras anteriores mencionando-as em seus textos em uma lista de referências.

## Resumo

O artigo é dividido em duas partes. A primeira descreve a avaliação de uma pequena universidade, baseada em dados cienciométricos tendo por principal objetivo avaliar a pesquisa com visibilidade internacional. A segunda parte mostra como um método econométrico (DEA: data enveloping analysis) pode ser usado para incluir na avaliação o ensino e o levantamento de recursos, entre outros aspectos. As duas abordagens mostram como um corpo teórico combinando (bibliometria, ciencimetria e econometria) pode ser aplicado a problemas concretos.

## Palavras-chave

Bibliometria; Ciencimetria; Econometria.

f) As revistas científicas desempenham papel essencial na comunicação entre pesquisadores correlacionados. Essa é a principal razão do porquê de a literatura científica ser a representação da atividade científica e da rede de relações entre os campos.

g) O número de publicações de um grupo de pesquisa pode ser considerado como um indicador da sua produção científica.

h) O número de vezes que as obras de um grupo são citadas por outras publicações é a medida do impacto ou da visibilidade internacional dessas obras.

i) No sistema global de revistas, é possível distinguir entre as revistas internacionais "centrais", quer dizer, importantes, e as periféricas, em geral revistas com um perfil voltado a interesses regionais.

j) O Institute for Scientific Information (ISI) cobre, com a atuação do seu *Science Citation Index* (SCI), a maioria das revistas internacionais importantes nas áreas de ciências puras, aplicadas e médicas. O *Social Science Citation Index* (SSCI) tem a mesma função para a área de ciências sociais.

Os indicadores cienciométricos não objetivam substituir os especialistas, mas, ao contrário, tornar a pesquisa visível e analisável, de modo que os especialistas tenham ao seu alcance a informação adequada e, conseqüentemente, possam formular uma opinião mais bem fundamentada (Abelson<sup>9</sup>, Wallmark e Sedig<sup>10</sup>). Tais indicadores são mais eficazes em um nível de agregação mais alto (Garfield<sup>11</sup>).

Os cientistas que trabalham em universidades têm várias atribuições: além do *ensino* e da *pesquisa*, também têm um compromisso em relação à sociedade como um todo. O método cienciométrico considera somente o *aspecto da pesquisa* e, essencialmente, qual o alcance que as contribuições de um grupo visivelmente têm em relação ao desenvolvimento de novos conhecimentos na frente de pesquisa. Isso significa que muitas atividades normalmente feitas por pesquisadores não são levadas em consideração.

Os resultados desta investigação foram distribuídos em três versões. A primeira versão contém os resultados de cada grupo separadamente. Cada grupo de pesquisa recebeu um relatório contendo apenas os seus próprios resultados, como também a descrição dos resultados globais da LUC. A segunda versão contém os resultados globais da LUC. Essa versão foi apresentada aos grupos de pesquisa e ao Conselho de Pesquisa. Finalmente, há a versão oficial em holandês (depois traduzida para o inglês) publicada com a aprovação do Conselho de Pesquisa para que os interessados em seus métodos ou resultados pudessem consultá-la. Na versão oficial (Rousseau<sup>4</sup>), todos os dados foram mantidos no anonimato.

#### Indicadores de potencial

Estimamos a capacidade de pesquisa de cada grupo e da escola (doravante simplesmente denominada LUC), expressa em equivalência pessoa-ano em pesquisa (person-year equivalents research - PYE-Res). Para obter esses dados, observamos as seguintes normas. Um parâmetro foi introduzido para expressar a fração de horas trabalhadas em pesquisa, usando as seguintes porcentagens (da mesma forma como a OECD): para os membros da equipe pagos pela universidade, esse parâmetro foi estabelecido em 0.4 (40%); para os pesquisadores pagos pela National Science Foundation e outras organizações similares, o parâmetro foi estabelecido em 1.0 (100%); o pessoal acadêmico empregado por período menor que o integral — 100% — foi computado na proporção do seu regime de trabalho (quem trabalha meio período tem um parâmetro igual a 0.2). Se a porcentagem do regime de trabalho é menor que 20%, essas pessoas são consideradas apenas como professores, e seu parâmetro de pesquisa é 0.0. Usando esses parâmetros e o número de pessoas ativas em cada grupo de pesquisa, calculamos o indicador PYE-Res/YEAR, isto é, a média de equivalência pessoa-ano em pesquisa por ano no período de 1981 a 1993.

#### Indicadores de produção, produtividade e impacto

A produção de um grupo de pesquisa (ou da LUC) é o número de suas publicações. A produtividade é expressa como o número de publicações por equivalência pessoa-ano em pesquisa. A medida do impacto é baseada no número de citações por publicação e relacionada a um padrão internacional (a média mundial).

#### Indicadores de produção (PA, Pt, P, Pf)

Para cada grupo de pesquisa, foram determinados o número total de publicações e o número de publicações incluído nos arquivos LUC-JCR (Journal Citation Reports), no período de 1981 a 1993. Esses indicadores são identificados pelos símbolos PA e Pt. O número de publicações incluídas no JCR no período entre 1981 e 1989 é identificado por P. Tomamos por base o *Science Citation Index* (SCI), que utiliza a seguinte classificação para as publicações: artigos (comuns/originais), cartas, notas, revisões, resumos de conferências, editoriais, resenhas de livros e correções. Somente as primeiras quatro categorias mencionadas são consideradas para os parâmetros Pt e P.

Com freqüência, um grupo de pesquisa publica artigos em colaboração com cientistas de outros grupos, tanto de dentro como de fora da universidade (Arunachalam<sup>12</sup>). É legítimo perguntar, então, se tais publicações podem ser integralmente atribuídas a cada grupo de pesquisa participante. Isso é um problema quando as publicações estão relacionadas a dados de insumo, como equivalência pessoa-ano em pesquisa. Na verdade, os artigos resultantes de tal colaboração são parcialmente baseados em esforços e recursos de outros grupos. Por essa razão, usamos dois esquemas de contagem: a usual e direta (uma publicação é uma unidade) e a fracionada (considerando a colaboração). Se, de acordo com os dados de publicação encontrados nos registros do SCI, um artigo é publicado por membros de dois grupos, cada grupo recebe metade do crédito. Se três grupos estão envolvidos, cada um recebe um terço do crédito, e assim por diante. As publicações computadas dessa forma no

LUC-JCR são identificadas por Pf. Esse indicador somente é usado para a avaliação da produtividade de um grupo de pesquisa.

Observe-se que os indicadores P, Pt e Pf somente estão relacionados aos artigos do LUC-JCR, ou seja, artigos originais, cartas, notas e revisões publicados em uma revista do JCR e que contenham a identificação da LUC.

### **Indicadores de produtividade (PA/PYE-Res por ano, Pf/PYE-Res por ano)**

A produtividade é medida usando-se as relações PA/PYE-Res e Pf/PYE-Res (o número total de publicações e o número fracionado de publicações no JCR por equivalência pessoa-ano em pesquisa) por ano. Como o objetivo deste relatório é determinar a posição internacional da LUC, seria desejável comparar a produtividade dos grupos de pesquisa e das escolas com um padrão internacional. Isso nos permitiria determinar se, de um ponto de vista global (internacional), a produtividade é alta ou baixa. Entretanto, até o momento não é possível obter dados confiáveis para traçar tal comparação.

A comparação da produtividade entre diferentes grupos de pesquisa geralmente não tem significado. Além do fato de que isso determinaria somente a posição relativa local de um grupo, há o aspecto complicador de que os hábitos de publicação diferem marcadamente entre campos e departamentos científicos. Por exemplo, tem pouco significado comparar a produtividade de um grupo experimental com a de um grupo aplicado. A principal função desse indicador, portanto, é acompanhar grupos de pesquisa com uma *produtividade extremamente baixa*, isto é, grupos que publicam muito pouco em relação à sua capacidade de pesquisa.

### **Indicadores de impacto (Ct, Ctex, C, Cex, CpP, CexpP, CpP/FCSm, CpP/JCSm, JCSm/FCSm, PSt, PS)**

Esses indicadores somente se referem aos artigos da LUC-JCR. Para cada cientista, determinamos o número total de citações durante o período 1981-1993 (e isso referente aos artigos publicados nesse período); seu símbolo é Ct. As autocitações também estão incluídas aí.

Para termos noção do número de autocitações, determinamos igualmente o número total de citações, excluindo as autocitações, e a porcentagem de autocitações (símbolos: Ctex, PSt). Além disso, esses indicadores foram determinados para um período de quatro anos após a publicação (símbolos: C, Cex, PS). Obtemos, ainda, a média do número de citações por artigo, com e sem autocitações (símbolos: CpP e CexpP).

Geralmente, um grupo de pesquisa publica artigos em determinadas revistas: é a coleção de revistas desse grupo. É possível obter um fator de impacto para cada revista dessa coleção, isto é, a média do número de vezes que um artigo publicado em uma revista é citado. Observe-se que, para o cálculo do fator de impacto de uma revista, todos os artigos são considerados, e não somente os publicados pelo grupo de pesquisa em avaliação. Neste relatório, esse fator de impacto é denominado JCS (Journal Citation Score – média de citações da revista). Usar fatores de impacto de revistas possibilita formular uma média-ponderada-do fator de impacto para a coleção de revistas na qual o grupo publica (símbolo: JCSm). Ilustramos essa formulação da seguinte maneira: supomos que um grupo tenha publicado, em um determinado ano, três artigos na revista A e dois na revista B. Supomos, ainda, que os artigos na revista A obtenham uma média de seis citações (JCS = 6), e que os da revista B obtenham duas (JCS = 2). A média-ponderada-do fator de impacto para a coleção de revistas desse grupo naquele determinado ano (JCSm) é:  $(3 \times 6 + 2 \times 2) : 5 = 4.4$ .

Os fatores de impacto podem ser calculados de diferentes maneiras. A que usamos neste relatório difere do método "oficial", ou seja, do ISI. A noção do fator de impacto sempre se refere a uma média do número de citações de artigos publicados em uma determinada revista. O fator de impacto do ISI se refere à média do número de citações obtidas *durante o segundo ou o terceiro ano* após a publicação. Mais precisamente: o fator de impacto ISI da revista X em 1998 é obtido dividindo-se a soma das citações no ano de 1998 de artigos publicados nessa revista em 1996 e 1997 pela soma do número de artigos publicados em X durante os anos de 1996 e 1997.

Limitamos nosso trabalho a um período predeterminado, chamado período de quatro anos (Rousseau<sup>13</sup>), que consiste do ano de publicação (ano um) e dos três anos seguintes. Essa escolha resulta no fato de que somente as publicações do período estudado (1981-1989) podem ser analisadas completamente.

Esses indicadores básicos respondem às questões: Qual a frequência em que esses artigos são citados dentro do conjunto de revistas do SCI? E qual é a média ponderada do fator de impacto das revistas nas quais a pesquisa é publicada? No entanto, não respondem a outra questão: se um grupo é ou não mais citado que grupos similares; se a soma das citações do grupo é alta ou baixa vista de uma perspectiva internacional. Uma comparação direta entre os grupos ou departamentos de uma universidade, baseada nos números absolutos de citações (C), de citações por publicação (CpP) ou o fator de impacto de uma coleção de revistas (JCSm), geralmente não tem significado devido à variação dos hábitos de citação nos diferentes campos da ciência. Em áreas como imunologia, as listas de referências são muito mais extensas do que em matemática ou em ciências da computação. Portanto, optamos por uma abordagem na qual, pelo menos para a análise de impacto, um nível internacional de referência foi estabelecido. Os resultados dos grupos de pesquisa da LUC são, então, comparados com esse nível internacional de referência. Calculamos, por subcampo, o número médio de citações que um artigo desse campo obtém, que é chamado média global por campo, e identificado como FCS (Field Citation Score – média das citações no campo). Para a classificação das revistas nos subcampos, é utilizada as categorias do ISI. Assim, os resultados de citações da LUC são comparados com "o mundo", e, portanto, é possível estabelecer se os artigos de um grupo de pesquisa são mais ou menos citados que a média mundial do campo no qual o grupo trabalha. Geralmente, um grupo atua em mais de um campo, e, conseqüentemente, é usada uma média ponderada (símbolo: FCSm), similar ao cálculo da média ponderada do fator de impacto (JCSm). Ver também Van Hooydonk et al.<sup>14</sup>.

Três indicadores importantes são formulados baseados nas variáveis CpP, JCSm e FCSm:

a) A relação entre o impacto obtido e o impacto da coleção de revistas: CpP/JCSm.

Aqui, os resultados da média de citações de um grupo de pesquisa (CpP) são comparados com a média ponderada do fator de impacto das revistas nas quais o grupo publica (JCSm). Esse indicador permite determinar se o conjunto das obras do grupo – dentro do parâmetro de citações em quatro anos – é mais citado do que a média. O poder de avaliação desse indicador é limitado pelo fato de que o nível de citação da coleção de revistas não é considerado como um todo. Um grupo que publica em revistas menos importantes – com um fator de impacto baixo – e é citado tanto quanto a média de artigos da sua coleção de revistas é classificado, por esse indicador, no mesmo nível de um grupo que publica nas melhores revistas, apresenta um fator de impacto alto e obtém um resultado médio de citações igual ao da sua coleção de revistas.

b) A relação entre o impacto obtido e a média global de impacto dos campos nos quais o grupo atua: CpP/FCSm.

Esse indicador compara os resultados da média de citações do conjunto das obras do grupo com uma média internacional de citações dos campos nos quais o grupo trabalha. Uma taxa maior que 1 significa que o grupo é citado com mais frequência do que a média mundial dos seus subcampos. Esse indicador é mais apropriado para medir a posição internacional de um grupo de pesquisa.

c) A relação entre a média do fator de impacto da coleção de revistas e a média mundial de impacto dos campos nos quais o grupo atua: JCSm/FCSm.

Esse indicador compara a média ponderada do fator de impacto da coleção de revistas com a média mundial de citações dos campos nos quais o grupo trabalha. Uma taxa JCSm/FCSm maior que 1 significa que o grupo publica em revistas que têm um fator de impacto alto em relação à média mundial (no campo).

GRÁFICO 1  
Número de publicações por ano

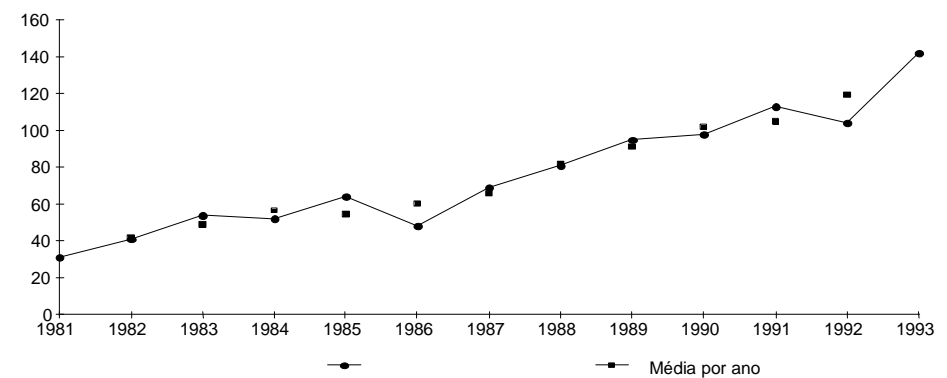
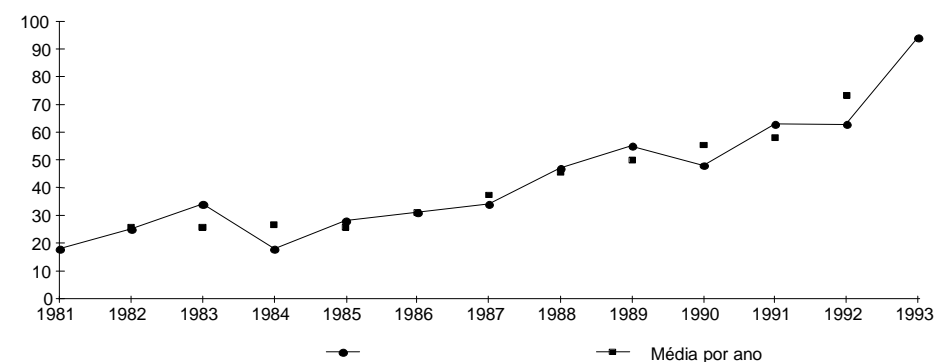


GRÁFICO 2  
Número de artigos JCR por ano



É claro que esses três indicadores não são independentes. O conhecimento de dois deles determina o terceiro.

#### Análises das tendências

Uma das limitações dos indicadores mencionados é o fato de que eles oferecem um valor médio que ultrapassa o período estudado. Assim, complementamos esse quadro com uma série de análises de tendências (Rousseau<sup>15</sup> aborda outro método para análises de tendências). Aqui, os indicadores são estudados ano por ano, a fim de conhecermos a evolução do impacto de um grupo. Os gráficos que seguem mostram a evolução anual e as médias de três anos. Considerando que a evolução ano a ano frequentemente ocorre de forma irregular, recomendamos observar com mais atenção a evolução das médias dos três anos.

#### 1. Número de publicações por ano

O gráfico 1 mostra o número total por ano dos trabalhos publicados por um grupo de pesquisa ou departamento. Todas as publicações são consideradas, incluindo os artigos publicados tanto em revistas não indexadas no JCR, como em livros e em anais de congressos.

#### 2. Número de artigos JCR por ano

O gráfico 2 computa somente os artigos publicados em revistas do JCR.

#### 3. Produtividade

Esse gráfico (não reproduzido neste trabalho) mostra a relação entre o número de publicações e o fator de equivalência pessoa-ano em pesquisa (PYE-Res).

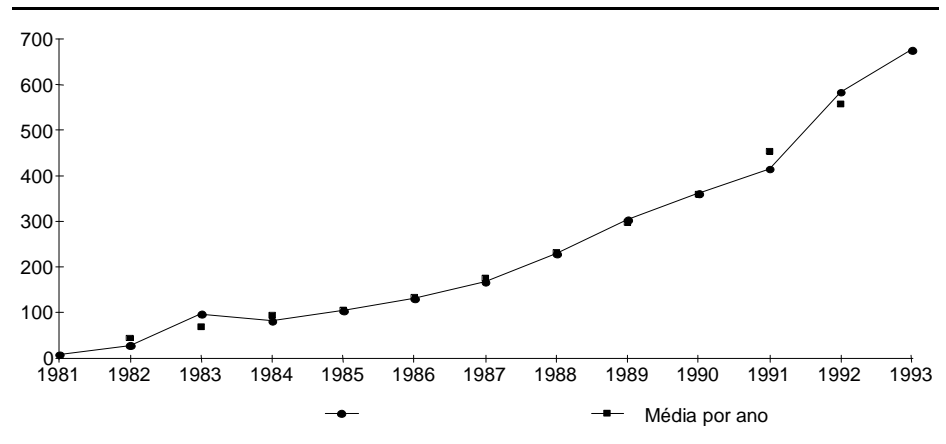
4. Produtividade JCR

Esse gráfico (não reproduzido aqui) é similar ao anterior, exceto que agora somente as publicações JCR são consideradas.

5. Número total de citações por ano

No gráfico 3, mostramos o número total de citações (em revistas JCR) que o grupo ou o departamento obteve por ano. Somente as citações de publicações no período 1981-1993 são consideradas. Como o número cumulativo de artigos citáveis aumenta ao longo do período, é natural que esse gráfico também aumente. Se isso não ocorrer, é um ponto negativo para o grupo.

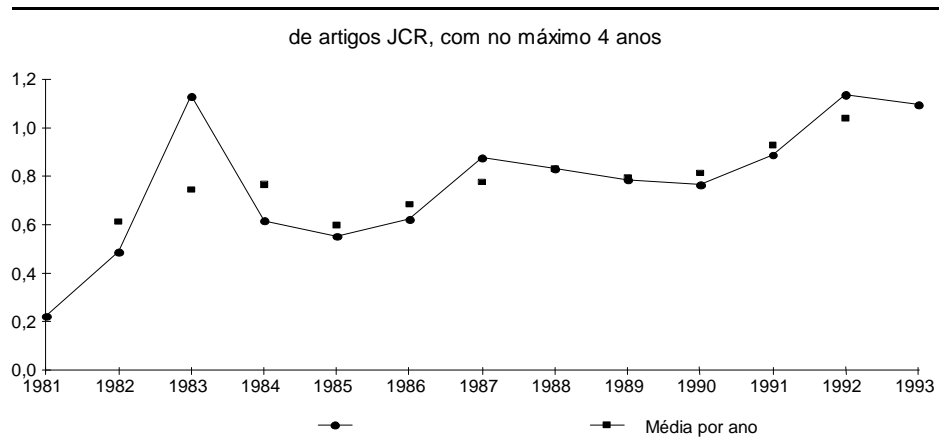
GRÁFICO 3  
Número total de citações por ano



6. Número de citações por publicação – os trabalhos foram publicados há quatro anos no máximo

Esse gráfico mostra os mesmos dados do gráfico anterior, com a diferença que as publicações têm no máximo quatro anos. Conseqüentemente, os primeiros três anos são incompletos (1981-1983).

GRÁFICO 4  
Número de citações por publicação



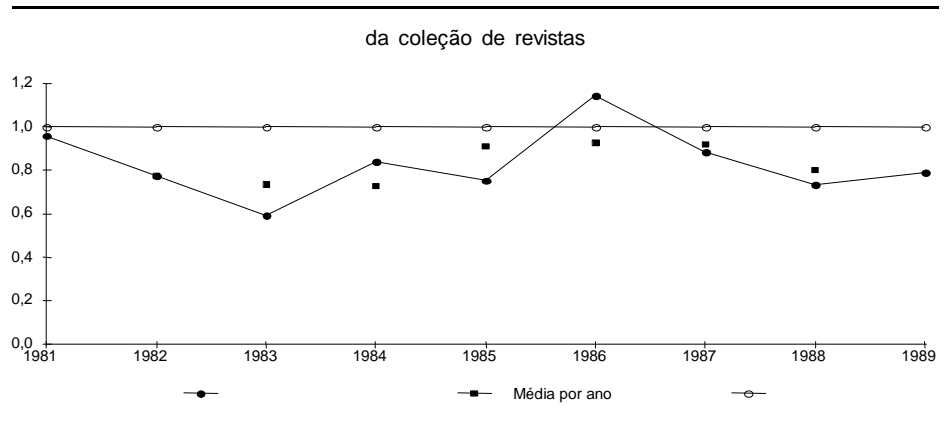
7. Número de citações de artigos JCR com, no máximo, quatro anos

O gráfico 4 é similar aos dois anteriores, considerando agora somente as publicações JCR. Assim, esse é um gráfico que mostra publicações por artigo, sendo que os três primeiros anos são incompletos.

GRÁFICO 5  
Impacto

8. Impacto das citações com quatro anos

Esse gráfico mostra, por artigo JCR, o número total de citações durante um período de quatro anos dos artigos publicados no primeiro ano desse período (o ano no eixo horizontal). Isso significa que os últimos três anos (1991-1993) são incompletos. Portanto, é normal que esses gráficos decresçam ao final! Antes desse período (1991-1993), um parâmetro de quatro anos foi usado.



9. Impacto com relação à coleção de revistas

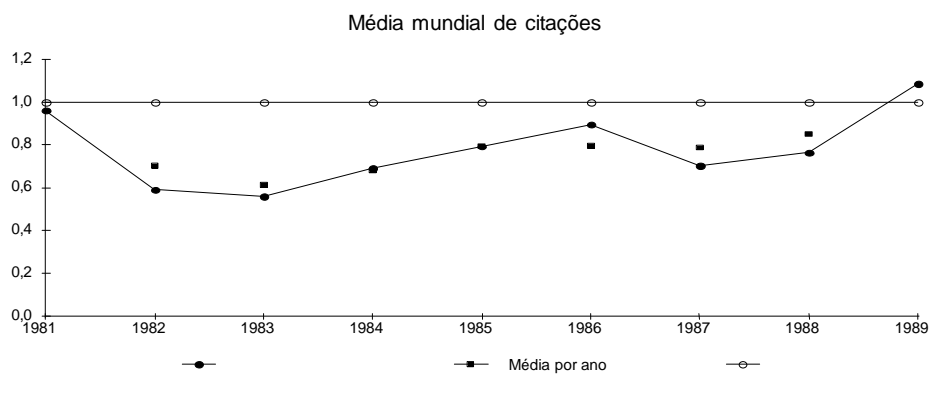
Os valores dos pontos do gráfico 5 são os quocientes dos números do gráfico anterior divididos pela média do número de citações obtido por um artigo publi-

cado na mesma revista durante um período de quatro anos. Se o grupo obtém os mesmos resultados da sua coleção de revistas, esse quociente é um.

#### 10. Impacto com relação à média mundial de citações

Os valores dos pontos do gráfico 6 são os quocientes dos números do gráfico referente ao impacto das citações com quatro anos e a média do número de citações que um artigo no mesmo campo (isto é, na mesma categoria do JCR) obteve durante o mesmo período de quatro anos.

**GRÁFICO 6**  
**Impacto**



#### Cobertura do SCI

Os indicadores de impacto são inteiramente baseados nos artigos que trazem o endereço da LUC e publicados em revistas cobertas pelo SCI. A hipótese central é que o SCI é um instrumento adequado para selecionar artigos divulgados em revistas internacionais de importância (Rousseau e Spinak<sup>16</sup>). Nesta seção, analisaremos mais profundamente a validade dessa afirmação. Para tanto, elaboramos dois indicadores adicionais.

##### a) Pt/Pjtot

Esse indicador dá a porcentagem de artigos nos registros da LUC-JCR (Pt) em relação ao número total de artigos de revistas (Pjtot). Se essa porcentagem se aproxima de 100, isso significa que o grupo de pesquisa publica seus artigos principalmente nas revistas cobertas pelo JCR. Nesse caso, o SCI tem uma boa cobertura da produção desse grupo.

##### b) Pjtot/PA

Esse indicador dá a porcentagem de artigos de revistas em relação ao número total de publicações (PA). Se essa porcentagem se aproxima de 100, isso significa que o grupo publica principalmente artigos e que a importância de outros tipos de publicação é menor.

Como esses dados devem ser interpretados? Comparativamente falando, se um grupo de pesquisa publica muitos artigos em revistas não cobertas pelo JCR, isso pode ser interpretado de duas formas. Se partirmos do princípio de que

a cobertura do SCI é adequada e realmente abrange todas as revistas internacionais de importância, então concluiremos claramente que esse grupo não publica em revistas importantes. Nesse caso, a relação Pt/Pjtot é um indicador da posição internacional do grupo, como demonstra seu comportamento ao publicar, isto é, a escolha que faz das revistas para publicar artigos (ou, pode ser que os artigos desse grupo somente sejam aceitos por revistas menos importantes). Assim, uma taxa pequena significa uma fraca posição internacional.

Por outro lado, podemos partir do princípio de que a cobertura do SCI é inadequada e que muitas revistas importantes não são consideradas. Assumindo também que o grupo publica em revistas importantes, então a relação Pt/Pjtot torna-se um indicador da cobertura do SCI, sendo o grupo de pesquisa (departamento ou escola) um padrão.

Pode haver um problema com a adequação do SCI naqueles campos em que a publicação de alto nível não é restrita a um grupo relativamente pequeno das melhores revistas internacionais de língua inglesa, mas em que existe uma grande multiplicidade de revistas, muitas vezes com interesse local (ou seja, dirigidas a uma nação ou a um grupo lingüístico), que publicam artigos de elevada qualidade. Em geral, o problema não está nas ciências. Entretanto, esse aspecto deve ser considerado na interpretação dos resultados.

Observações análogas podem ser feitas para a relação Pjtot/PA. Se um grupo tem uma taxa Pjtot/PA baixa, isso pode significar que o grupo não quer ou não pode publicar em revistas. Porém, também pode significar que as revistas desse campo de pesquisa exercem um papel menor na disseminação do conhecimento científico.

#### Discussão dos resultados

A escola de Ciência e Medicina publicou 1038 trabalhos durante o período de 1981 a 1993, o que significa uma média de aproximadamente 80 publicações por ano, 82% das quais são artigos de revistas. Conseqüentemente, 18% se referem a outros tipos de publicações, demonstrando que, na escola de Ciência e Medicina, os anais de congressos e as contribuições incluídas em livros também são importantes. Dos artigos em periódicos, 72% se enquadram nas categorias do JCR: artigos originais, cartas, revisões e notas. Em geral, não há problemas de cobertura para o SCI, porque grande parte das outras publicações é composta por resumos nas revistas JCR.

Os 586 artigos da escola constantes do JCR são citados 2 915 vezes. A porcentagem de autocitações é igual a 38%, o que é muito alto. Considerando o impacto nesse curto espaço de tempo, observamos que esses artigos (publicados no período 1981-1989) são citados em média 3.38 vezes durante os primeiros quatro anos após a publicação (39% são autocitações).

A média do fator de impacto da coleção de revistas da escola é de 4.77, e a média da soma mundial de citações é de 4.83, mostrando que a escola como um todo encontra-se no nível da média. A média do impacto da escola está situada logo abaixo da média mundial ( $3.88/4.83 = 0.8$ ), considerando margens de erro. O impacto da coleção de revistas é de 0.81.

Análises das tendências mostram uma trajetória irregular, mas terminando de uma forma ascendente. Podemos dizer seguramente que em 1989 a escola de Ciências e Medicina alcançou a média mundial.

### Discussões com os cientistas

Consultamos os pesquisadores em dois encontros de validação. Os cientistas tiveram oportunidade de corrigir dados numéricos. Além disso, os resultados das análises cienciométricas foram confrontados com a imagem que os cientistas tinham de sua escola e de seu grupo de pesquisa. Durante as entrevistas, coletamos informações que tiveram importância concreta na correta interpretação dos resultados cienciométricos. Essas interpretações, encaminhadas ao Conselho de Pesquisa para deliberação, fazem deste relatório um instrumento válido para uma política de pesquisa da universidade. Os pesquisadores formularam uma série de questões importantes, às vezes críticas, mas nunca totalmente negativas.

Os debates que mantive com os cientistas foram, acredito, muito instrutivos para ambas as partes. Como mencionado anteriormente, apenas os próprios cientistas podem corretamente avaliar os resultados obtidos. O aspecto comum a todas as discussões foi o fato de que os cientistas ficaram um tanto desapontados com os resultados (as médias durante todo o período), mas satisfeitos com as análises de tendências que mostram o progresso que fizeram. Frequentemente, alegaram que os programas de pesquisa maiores precisam de um período de incubação para serem iniciados. Eles também acentuaram o fato de que, além das citações “objetivas”, também os contatos pessoais e os pedidos para orientação foram meios importantes de solicitude apreciados por seus pares.

Como respostas a questões mais específicas, registramos as seguintes observações:

- Os métodos cienciométricos são úteis, desde que as diferenças de meios e de potenciais humanos entre os grupos de pesquisa sejam consideradas.

- Visibilidade é um dos aspectos importantes da pesquisa científica. Um estudo de citação oferece interessante *feedback* para isso.

- Um estudo cienciométrico confiável informa os membros da universidade de modo que eles possam tentar melhorar sua visibilidade (conforme Cole e Cole<sup>17</sup>).

- Muitas pessoas enfatizaram a natureza objetiva dos dados sobre citação, mas também observaram a sua unilateralidade: ser um membro da universidade exige muito mais do que pesquisa científica como a que é publicada em revistas do JCR.

- Um grupo de pesquisa afirmou que uma publicação deve ser o resultado natural da pesquisa, e não vice-versa. Nesse mesmo sentido, outro grupo declarou que publicar *mais* não deve ser o alvo de uma política de publicação, mas sim a obtenção de resultados de pesquisa que possam ser publicados.

- O estudo cienciométrico foi visto como um incentivo – para o Conselho de Pesquisa – para valorizar o trabalho realizado pelos grupos de pesquisa.

- Foram encontrados problemas de cobertura nas áreas de ciências da computação e estatística. Para estatística, o problema reside em algumas revistas não cobertas pelo ISI; para ciências da computação, o problema se refere mais ao fato de que, nesse campo, os cientistas tendem a publicar mais em anais de congressos que, com frequência, têm critérios de seleção muito restritos.

- Alguns cientistas quiseram saber como, em um futuro próximo, as revistas eletrônicas e as citações nelas contidas serão avaliadas (Collins e Berge<sup>18</sup>).

- Um grupo de pesquisa não vê sentido em um estudo cienciométrico, em particular no caso de uma pequena universidade como a LUC, em que os grupos de pesquisa são muito pequenos. Para esse grupo, o uso indiscriminado dos resultados dessa análise para propósitos administrativos está fora de questão.

- Parece ser mais difícil obter bons resultados quando se publica em revistas com alto impacto.

Observação (minha): Essas revistas de alto impacto, em geral, pertencem a campos da ciência conhecidos como “grande ciência”. Particularmente nesses campos, o fato de a LUC ser uma pequena universidade representa um aspecto adverso. Isso explica, pelo menos parcialmente, os resultados decepcionantes de alguns grupos de pesquisa.

### Conclusão da Parte A

Estamos convencidos de que, para aqueles que usam estudos de avaliação da ciência de uma forma responsável objetivando uma política de pesquisa, este relatório prova ser um instrumento útil. Para os cientistas da LUC, foi uma forma de *feedback* em relação a uma de suas principais atividades.

### PARTE B: DEA – DATA ENVELOPMENT ANALYSIS COMO INSTRUMENTO PARA A CONSTRUÇÃO DE INDICADORES CIENCIOMÉTRICOS

O objetivo desta parte é apresentar a técnica DEA – Data Envelopment Analysis, ou análise que inclui, encapsula os dados, e mostrar sua viabilidade em aplicações cienciométricas. Como exemplo, iremos examinar os resultados de desempenho dos departamentos da universidade.

## Exposição do problema e sua solução: DEA – Norman e Stoker<sup>19</sup>

Os principais problemas ao avaliarmos a eficiência de instituições públicas, como as universidades, são a falta de um padrão absoluto e uma boa estimativa da função da produção, no sentido econômico da palavra. A função da produção é a relação funcional entre insumos e produtos, definindo quais e quantos insumos são necessários para gerar um produto específico. Outro problema é a necessidade de considerar tanto os múltiplos insumos quanto os múltiplos produtos. Nesse contexto, a expressão Unidades de Tomada de Decisão (Decision Making Units – DMUs) será usada para instituições, governos etc. Uma primeira e bastante informal definição de eficiência poderia ser o uso feito de recursos para a realização de objetivos, levando-se em consideração fatores socioeconômicos.

A descoberta ocorreu durante uma pesquisa desenvolvida por Charnes, Cooper e Rhodes<sup>20</sup>, o primeiro trabalho que usou a técnica DEA. Sua definição formal de eficiência é:

“100% de eficiência são alcançados por (uma unidade) somente quando:

(a) Nenhum de seus produtos pode aumentar sem que um ou mais de seus insumos aumentem ou que outros produtos diminuam.

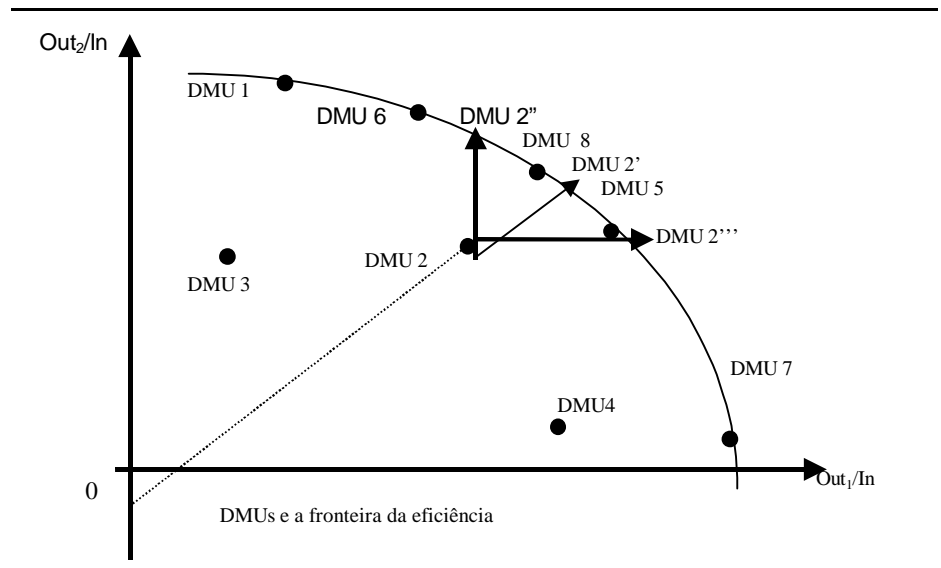
(b) Nenhum de seus insumos pode diminuir sem que alguns de seus produtos diminuam ou que outros insumos aumentem.”

Já que não há maneira de estabelecer um padrão absoluto de eficiência, a seguinte definição de eficiência relativa foi introduzida por Charnes e Cooper: “100% de eficiência relativa são alcançados por qualquer (unidade) somente quando comparações com outras (unidades) relevantes não ofereçam evidência de ineficiência no uso de qualquer insumo ou produto.”

O uso da DEA tem duas vantagens:

- É um método não paramétrico, quer dizer, não é necessário conhecer a relação funcional entre insumos e produtos.

Figura 1



• O método conduz a resultados de desempenho relativos. Não é necessário definir um padrão externo de desempenho absoluto. Os membros do grupo são comparados somente com outros membros desse grupo (as unidades de tomada de decisão – DMUs).

DEA tem sido usada, por exemplo, em estudos sobre administração de hospitais, na organização do Comando de Recrutamento da Marinha dos Estados Unidos, em distritos escolares, em análises sobre a lucratividade de diferentes setores da mesma empresa, em países, universidades e departamentos universitários (Arcelus e Coleman<sup>21</sup>, Beasley<sup>22</sup>, Degraeve et al.<sup>23</sup>).

Para clareza, usamos a figura 1 para explicar a técnica da DEA, em que vemos uma representação gráfica da eficiência relativa de várias DMUs imaginárias. Essas unidades empregam um insumo para gerar dois produtos. Os eixos horizontal e vertical representam a relação entre cada produto ( $OUT_1$  e  $OUT_2$ ) e a quantidade de insumo ( $IN$ ) usado. Na DEA, o desempenho de cada DMU é comparado com os das outras unidades consideradas no caso. As unidades que têm melhor desempenho otimizam mais seus insumos do que as demais. Novamente, a relatividade dessa abordagem deve ser enfatizada. As unidades mais otimizadas formam uma fronteira, chamada “fronteira da eficiência”. As unidades com menor desempenho precisam de mais insumos para produzir a mesma quantidade de produ-

tos e estão, portanto, situadas a certa distância da fronteira. A distância radial representa sua ineficiência. É evidente que todas as unidades situadas na fronteira da eficiência terão uma taxa de desempenho relativo igual a 1 (são 100% eficientes).

A fronteira da eficiência inclui todas as unidades de tomada de decisão, como pode ser visto na figura. É por isso que a técnica é a análise que inclui, encapsula os dados. Para aumentar o desempenho de uma DMU fraca, como a DMU 2, uma nova estratégia deve objetivar uma evolução em direção à fronteira da eficiência. A técnica DEA sugere uma evolução radial da origem ao alvo na fronteira, que representa o que a unidade poderia atingir. Dessa forma, o ótimo desempenho é alcançado na unidade DMU 2'. Observe-se que DMU 2' está situada entre as ótimas unidades DMU 8 e DMU 5, que formam um grupo de referência para a DMU 2. O grupo de referência permite-nos calcular as reduções de insumos e, ocasionalmente, as reduções de produtos que são necessárias para obter uma unidade com desempenho eficiente e nos mostra onde e como atuar. Além da radial, existem também outras formas de tornar certa unidade eficiente, por exemplo, uma evolução puramente vertical conduz à unidade eficiente DMU 2''.

Agora, para solucionar um problema determinado, temos que calcular o desempenho de cada DMU  $k$  ( $k=1..n$ ), comparado com as melhores unidades



do grupo. Portanto, precisamos maximizar a taxa de produção ponderada ( $O_{rk}$ ;  $r=1..s$ ) com os insumos ponderados ( $I_{ik}$ ;  $i=1..m$ ) sujeitos à condição de que taxas similares para cada DMU sejam menores ou iguais à unidade. Mais precisamente, isso resulta no seguinte modelo matemático: para cada  $k$  fixo ( $k = 1, \dots, n$ ), encontrar

$$MAX h_k = \frac{\sum_{r=1}^s v_{rk} * O_{rk}}{\sum_{i=1}^m u_{ik} * I_{ik}} \quad (1)$$

sujeito a

$$\frac{\sum_{r=1}^s v_{rk} * O_{rj}}{\sum_{i=1}^m u_{ik} * I_{ij}} \leq 1 \quad \text{para } j = 1 \dots n \quad (2)$$

$$v_{rk} \geq \epsilon \quad r = 1 \dots s \quad (3)$$

$$u_{ik} \geq \epsilon \quad i = 1 \dots m \quad (4)$$

As variáveis de decisão desse modelo são os coeficientes ponderados ( $v_{rk}$  e  $u_{ik}$ ): um grupo para cada unidade. Eles são construídos dessa forma e geram resultados mais benéficos para a unidade em questão. Isso significa que valores altos são atribuídos aos melhores resultados e que valores baixos são atribuídos aos insumos mais freqüentemente usados. O resultado mais elevado possível é, claro, um. A unidade é então chamada eficiente. As equações (3) e (4) definem por razões técnicas uma margem dentro da qual os valores não podem atingir. Usamos  $\epsilon = 0.000001$ .

Esse modelo DEA é não-linear e não-convexo. Já que não existe uma maneira fácil para solucionar esse tipo de problema, iremos transformá-lo em um modelo de programação linear, igualando-se o denominador a um. O novo modelo permite uma solução fácil empregando-se softwares existentes, como o LINGO. Assim, obtemos o seguinte modelo linear e convexo:

$$MAX h_k = \sum_{r=1}^s v_{rk} * O_{rk}$$

sujeito a

$$\sum_{r=1}^s v_{rk} * O_{rj} - \sum_{i=1}^m u_{ik} * I_{ij} \leq 0 \quad \text{com } j = 1 \dots n$$

$$\sum_{i=1}^m u_{ik} * I_{ik} = 1$$

$$v_{rk} \geq \epsilon$$

$$u_{ik} \geq \epsilon$$

Pode ser demonstrado que a DEA nunca irá superestimar o potencial de aperfeiçoamento relativo àquele identificado por qualquer outro método ponderado. Observe-se que as comparações são sempre feitas com relação ao desempenho alcançado, nunca com relação a algum objetivo hipotético, que poderia não ser atingido.

### Departamentos de Universidade (Degraeve et al.<sup>23</sup>)

#### • Seleção de resultados e recursos

Os resultados respondem à questão: Quais são os objetivos de um departamento de universidade? Os recursos respondem à seguinte pergunta: Quem ou o que permite a um departamento alcançar seus objetivos?

#### Resultados:

Os esforços em **educação** medidos como o número de estudantes atendidos pelo departamento. Os estudantes são classificados de acordo com o seu nível (graduação ou pós-graduação); e se precisam equipamentos de laboratório (caros) – uma ou duas variáveis.

#### Pesquisa básica e aplicada.

Recursos obtidos para realizar esse tipo de pesquisa (Degraeve et al.<sup>23</sup>) – uma ou duas variáveis.

Avaliações dos departamentos feitas por um comitê externo (como em Beasley<sup>22</sup>).

#### Recursos:

Pessoal (FTE) pago com recursos do governo, ou despesas com salários.

Outros recursos da universidade subvencionados pelo governo (incluindo a compra de equipamentos e também verbas para a biblioteca).

Os resultados da Catholic University of Leuven são: 36 departamentos – 13 dos quais atingem um grau de eficiência igual a 1 – são 100% eficientes: lei eclesiástica, filosofia, criminologia, economia aplicada, história, psicologia e antropologia, física, química, eletrônica, arquitetura, ciências biológicas aplicadas, medicina e cinestesia. Apenas para registro, mencione os mais fracos: construção (46%), geologia e geografia (42%), odontologia (27%) e educação física (ginástica e esportes: 23%). Esses resultados claramente mostram que não há inclinações em favor das humanidades ou das ciências ou das ciências aplicadas.

#### • Eficiência e eficácia

Para obter um grau de eficiência alta, a DMU é livre para usar seus recursos da forma que quiser. Os departamentos podem usar todos os seus insumos para obter um único produto/resultado (por exemplo, o ensino) e ainda conseguir o grau 1 de eficiência. E mais, se uma unidade possui várias funções, ela pode não ser considerada como realmente eficiente, se não tiver um bom desempenho em cada função. Isso leva à noção de eficácia. A eficácia tem sido observada dentro do quadro DEA, através da imposição de limites aos fatores de medida (Pedraja-Chaparro et al.<sup>24</sup>).

Impondo restrições aos instrumentos de medidas, acrescentando as seguintes condições:

$$v_{rk} - c_1 \sum_{l=1}^s v_{rl} \leq 0 \quad (8) \quad r = 1..s$$

e

$$u_{ik} - c_2 \sum_{j=1}^m u_{ij} \leq 0 \quad (9) \quad i = 1..m$$

em que  $c_1$  e  $c_2$  são constantes. Limites de 40% e 60% para cada insumo e de 25% a 40% para cada produto conduzem a uma redução geral nos graus de eficiência. Criminologia cai para 50% e física, para 46%. Contudo, lei eclesástica, cinestesia e história permanecem em 100%. Por outro lado, odontologia permanece em 26% e educação física, em 22%, persistindo como os mais fracos.

### Conclusão da Parte B

Demonstramos que eficiência e eficácia podem ser consideradas quando se compara o desempenho de uma universidade. DEA é um método adequado para atingir esse objetivo. Impor restrições aos valores dos insumos e dos produtos é uma forma de considerar a eficácia. Este trabalho (como outros similares) mostrou claramente o potencial de análise da DEA em estudos de desempenho (Johnes e Johnes<sup>25</sup>, Rousseau e Rousseau<sup>26,27</sup>).

### Bibliometric and econometric indicators for the evaluation of scientific institutions

#### Abstract

This article consists of two parts: the first part describes the evaluation of a small university. It is based on scientometric data and aims mainly at the evaluation of internationally visible research. The second part shows how an econometric method (DEA: data enveloping analysis) can be used to include teaching and other aspects, e.g. fund raising, into the evaluation. Both approaches show how a body of bibliometric, scientometric and econometric theory can be applied to real world problems.

#### Keywords

Bibliometrics; Scientometrics; Econometrics.

#### BIBLIOGRAFIA

1. VAN RAAN, A.F.J. (ed.). *Handbook of quantitative studies of science and technology*. Amsterdam: Elsevier, 1988.
2. EGGHE, L., ROUSSEAU, R. *Introduction to informetrics: quantitative methods in library, documentation and information science*. Amsterdam: Elsevier, 1990.
3. WALLACE, D.P. A solution in search of a problem: bibliometrics & libraries. *Library Journal*, n. 112, p. 43-7, May, 1987.
4. ROUSSEAU, R. *A scientometric study of the scientific publications of LUC. Period 1981-1993: report*. [Flanders]: [LUC], 1995.
5. DE BRUIN, R., KINT, A., LUWEL, M. et al. A study of research evaluation and planning: the University of Ghent. *Research Evaluation*, n. 3, p. 25-41, 1993.
6. DE BRUIN, R., MOED, H.F., SPRUYT, E. *Antwerpse analyses*. Antwerpen: CWTS, Leiden en UIA, 1993.
7. MOED, H.F., BURGER, W.J.M., FRANKFORT, J.G. et al. The use of bibliometric data for the measurement of university research performance. *Research Policy*, n. 14, p. 131-49, 1985.
8. OPPENHEIM, C. The correlation between citation counts and the 1992 research assessment exercise ratings for British library and information science university departments. *Journal of Documentation*, n. 51, p. 18-27, 1995.
9. ABELSON, P. Mechanisms for evaluating scientific information and the role of peer review. *Journal of the American Society for Information Science*, n. 41, p. 216-22, 1990.
10. WALLMARK, J.T., SEDIG, K.G. Quality of research measured by citation method and by peer review - a comparison. *IEEE Transactions on Engineering Management*, n. 33, p. 218-22, 1986.
11. GARFIELD, E. How to use citation analysis for faculty evaluations and when is it relevant, Part I. *Current Contents*, n. 44, Oct. 1983.
12. ARUNACHALAM, S. Science in the European Community - the extent of collaboration. *Current Science*, n. 63, p. 56-8, 1992.
13. ROUSSEAU, R. Citation distribution of pure mathematics journals. In: EGGHE, L., ROUSSEAU, R. (eds.). *Informetrics 87/88*. Amsterdam: Elsevier, 1988. p. 249-62.
14. VAN HOOYDONK, G., GEVAERT, R., MILIS-PROOST, G. et al. A biblioeconomic analysis of the impact factors of scientific disciplines. *Scientometrics*, n. 30, p. 65-81, 1994.
15. ROUSSEAU, R. Evolution d'importantes revues pharmacologiques. *Revue Française de Bibliométrie*, n. 5, p. 102-17, 1989.
16. ROUSSEAU, R., SPINAK, E. Do a field list of internationally visible journals and their journal impact factors depend on the initial set of journals? A research proposal. *Journal of Documentation*, n. 52, p. 449-56, 1996.
17. COLE, S., COLE, J.R. Scientific output and recognition: a study in the operation of the reward system in science. *The American Sociological Review*, n. 32, p. 377-90, 1967.
18. COLLINS, M.P., BERGE, Z.L. IPCT Journal: a case study of an electronic journal on the Internet. *Journal of the American Society for Information Science*, n. 45, p. 771-6, 1994.
19. NORMAN, M., STOKER, B. *Data envelopment analysis: the assessment of performance*. Chichester: Wiley & Sons, 1991.
20. CHARNES, A., COOPER, W.W., RHODES, E. Measuring the efficiency of decision making units. *European Journal of Operations Research*, v. 2, n. 6, p. 429-44, 1978.
21. ARCELUS, F.J., COLEMAN, D.F. An efficiency review of university departments. *International Journal of Systems Science*, v. 28, n. 7, p. 721-9, 1997.
22. BEASLEY, J.E. Determining teaching and research efficiencies. *Journal of the Operational Research Society*, n. 46, p. 441-52, 1995.
23. DEGRAEVE, Z., LAMBRECHTS, M., VAN PUYMBROECK, V. Een vergelijkende prestatie studie van de departementen van de Katholieke Universiteit Leuven. *Tijdschrift voor Economie en Management*, n. 41, p. 165-93, 1996.
24. PEDRAJA-CHAPARRO, F., SALINAS-JIMENEZ, J., SMITH, P. On the role of weight restrictions in data envelopment analysis. *Journal of Productivity Analysis*, n. 8, p. 215-30, 1997.
25. JOHNES, G., JOHNES, J. Apples and oranges; the aggregation problem in publications analysis. *Scientometrics*, n. 25, p. 353-65, 1992.
26. ROUSSEAU, S., ROUSSEAU, R. Data envelopment analysis as a tool for constructing scientometric indicators. *Scientometrics*, n. 40, p. 45-56, 1997.
27. ROUSSEAU, S., ROUSSEAU, R. Measuring the scientific wealth of European nations: taking effectiveness into account. *Scientometrics*, n. 42, p. 75-87, 1998.

#### Ronald Rousseau

KHBO, Zeedijk 101, Oostende, Belgium  
UIA, Universiteitsplein 1, Wilrijk, Belgium.

ronald.rousseau@kh.khbo.be

Tradução do inglês de Irati Antonio