

Sophia: 29235

MCT/BPC&T

CNPq

***Diretório dos Grupos de Pesquisa
no Brasil - versão 4.0
2000***

**Estratificação dos
Grupos de Pesquisa**

Ministério da Ciência e Tecnologia (MCT)
Serviço de Biblioteca de Política em
Ciência e Tecnologia (BPC&T)
Código de barra N°: 29235

VERSÃO PRELIMINAR

Sujeita a modificações

001.891(81)

E82

2000

001.891(81)
E82
2000

Sumário

Apresentação	- 1
Introdução	- 3
Metodologia	- 4
Principais resultados	- 6
Relação entre Qualificação e Produtividade dos Grupos de Pesquisa	- 9
Anexo – Procedimentos utilizados para o cálculo dos escores de Qualificação (Q) e de Produtividade (P) dos grupos de pesquisa vinculados às IES	- 13

Apresentação

A construção de ferramentas analíticas da atividade científico-tecnológica está na raiz da concepção do Diretório dos Grupos de Pesquisa. Nesse contexto, desde 1995 vem sendo desenvolvido um algoritmo capaz de estratificar os grupos em bases quali-quantitativas. A versão beta do algoritmo foi desenvolvida a partir da base de dados da versão 2.0 (Lourenço, R.S. et alii – *Indicadores de Qualidade e Produtividade dos Grupos de Pesquisa no Brasil*. Educação Brasileira, 17(35):107-128, 1995) e a primeira versão foi publicada em suporte papel a partir dos dados da versão 3.0 (1997) (Coav/Sup/CNPq/MCT – *A Pesquisa no Brasil: perfil da pesquisa no Brasil e hierarquização dos grupos de pesquisa a partir dos dados do Diretório dos Grupos de Pesquisa no Brasil*. Brasília, 101p., 1998). Na versão 4.0, o algoritmo foi aperfeiçoado e seus resultados foram incorporados dinamicamente à página do Diretório na Internet, como um módulo específico.

Numa taxonomia de instrumentos de estratificação que pusessem em um extremo os processos integralmente cientométrico-produtivistas e em outro apreciações integralmente qualitativas, com baixo poder de comparabilidade e formalização, o algoritmo está situado numa posição intermediária. São três os aspectos que o fundamentam: (1) a dimensão complexa e qualitativa dos insumos básicos que orientam o processo discriminatório (julgamento de bolsas de pesquisa pelo CNPq e avaliação dos programas de pós-graduação pela Capes); (2) a parametrização realizada segundo os traços específicos de cada uma das grandes áreas da árvore do conhecimento do CNPq; (3) a utilização de ferramenta estatística simples e de fácil compreensão, mas suficiente para estabelecer uma formalização que permita a comparação entre os grupos.

Como um elemento de reforço do poder de discriminação do modelo conceitual do algoritmo, vale a pena ainda lembrar a relativa independência entre os três eventos fundamentais na construção do mesmo: (1) a coleta de dados a respeito dos grupos de pesquisa; (2) a avaliação dos projetos e currículos no processo de concessão de bolsas de pesquisa pelos comitês assessores no CNPq e; (3) a avaliação dos programas de pós-graduação pelos comitês de área e pelo CTC na Capes. São processos onde as unidades de análise são distintas (pesquisadores, grupos de pesquisa e programas de pós-graduação), que ocorrem em momentos diferentes, em instituições diferentes, lançando mão de critérios de avaliação distintos e realizados por indivíduos, em grande parte, diferentes.

Com o intuito de testar o poder discriminatório do algoritmo, desde a versão beta (1995) vêm sendo realizados testes de pertinência ou correlação entre a hierarquia por ele construída e outras hierarquias disponíveis. A primeira delas foi um exercício sobre a localização relativa nos estratos

produzidos pelo algoritmo dos pesquisadores constantes da "lista dos 170 pesquisadores mais produtivos do país", aparecida na Folha de São Paulo em maio de 1995. Desses pesquisadores, 102 atendiam ao critério de estarem registrados na versão 2.0 do Diretório e pertencerem a uma Instituição de Ensino Superior, e 94 (92,1%) localizavam-se no estrato superior do algoritmo.

Os testes mais abrangentes foram as medidas de associação entre os escores médios de qualidade (que servem de base para a construção dos estratos) produzidos pelo algoritmo e escores-testes médios de produtividade dos grupos. Essas medidas foram realizadas segundo parâmetros estabelecidos para cada uma das grandes áreas *per se* e os exercícios foram feitos com os dados das versões 2.0 e 3.0 (1995 e 1997). Com a exceção dos grupos das ciências agrárias da versão 3.0 (0,876), os coeficientes de correlação entre os escores médios componentes dos escores de qualidade (Q) e de produtividade (P) para todas as grandes áreas esteve sempre acima de 0,92.

Introdução

Este trabalho classifica, em cinco estratos, os grupos de pesquisa vinculados às Instituições de Ensino Superior (IES)¹ cadastrados na versão 4.0 do Diretório dos Grupos de Pesquisa no Brasil (ano de referência 2000), tomando como indicador a densidade de pesquisadores qualificados por dois sistemas de avaliação já consagrados no Brasil, baseados em julgamentos desenvolvidos pelos próprios pares, de forma relativamente independente em relação ao *Directório* e entre si: o sistema de avaliação *ex-ante* dos projetos de pesquisa e dos currículos dos pesquisadores candidatos às bolsas de pesquisa concedidas pelo CNPq e o sistema de avaliação dos programas de pós-graduação empreendido pela CAPES.

Coloca em evidência as concentrações geográfica e institucional da pesquisa desenvolvida no âmbito das IES. Ordena as instituições sob a ótica da pesquisa científica por grande área do conhecimento, tendo em conta os quantitativos de grupos de pesquisa classificados nos diferentes estratos, em termos absolutos e relativos.

Ao final, averigua a existência de correlação entre o grau de qualificação e a produtividade técnico-científica dos grupos de pesquisa vinculados às IES. O indicador de produtividade considera a produção C&T dos pesquisadores doutores referente ao triênio 1997-1999 cadastrada com o auxílio do currículo Lattes até 10 de novembro de 2000 (artigos, livros e capítulos de livros publicados, produção tecnológica desenvolvida, teses e dissertações defendidas sob orientação de pesquisadores pertencentes aos grupos).

¹ Considerou-se também como IES as instituições de pesquisa que ministram cursos de pós-graduação reconhecidos pela CAPES.

Metodologia

O trabalho circunscreve 10.772 grupos vinculados às 155 Instituições de Ensino Superior (englobando cerca de 92% do total de grupos cadastrados na versão 4.0 do *Diretório*). Não envolve, portanto, os grupos vinculados a Empresas/Institutos de Pesquisa e Institutos Tecnológicos (69 instituições) que em conjunto abarcam 988 grupos cadastrados (**Figura 1**). Os Institutos de Pesquisa que ministram cursos de pós-graduação reconhecidos pela CAPES foram considerados como IES.

Os grupos de pesquisa foram classificados em 5 estratos rotulados de "A" a "E". No primeiro estrato ("A") foram classificados os grupos que podem ser qualificados como de **Excelência**, tendo em vista a altíssima concentração de pesquisadores participantes de programas de pós-graduação chancelados pela CAPES com os graus mais elevados da escala de avaliação utilizada por aquela Agência e, ainda, apoiados pelo CNPq com bolsas de pesquisa de elevada categoria/nível. Na medida que caminhamos do primeiro ("A") para o quinto ("E") estrato, a densidade de pesquisadores qualificados vai se reduzindo sensivelmente. Neste último estrato, que podemos qualificar como grupos **Em Formação**, é nula ou muito rarefeita a presença de pesquisadores qualificados pelos dois sistemas de avaliação que escoram este trabalho, como poderá ser visto mais adiante. Foram denominados como **Em Consolidação** os grupos classificados nos estratos "C" e "D" e como **Consolidados** aqueles alojados nos estratos "A" e "B".

Os grupos de pesquisa foram dispostos nos cinco estratos observando os seguintes passos:

1º) os grupos vinculados às IES foram subdivididos em dois subconjuntos (**Figura 1**). O primeiro deles congrega 81% do total de grupos vinculados às IES: 8.773 grupos que contam com a presença de pelo menos um pesquisador com titulação máxima igual a doutorado e, ainda, pelo menos um bolsista de pesquisa ou docente de um curso avaliado com grau igual ou superior a 3. O segundo subconjunto comporta os demais grupos de pesquisa: 1.999 grupos, perfazendo cerca de 19% do total de grupos vinculados às IES.

2º) os grupos pertencentes ao primeiro subconjunto foram distribuídos nos cinco estratos referidos inicialmente, rotulados de "A" a "E", com base em um escore de qualificação (Q) que mede a densidade de pesquisadores classificados pelos dois sistemas de avaliação já citados, o sistema de bolsas de pesquisa do CNPq e o sistema de avaliação dos programas de pós-graduação da CAPES. Foram considerados os bolsistas em atividade em novembro de 2000 e o cadastro de docentes

1999 vinculados aos cursos/programas de pós-graduação reconhecidos pela CAPES até dezembro de 1999, bem como os resultados da avaliação concluída em 1998. Para calcular a densidade de pesquisadores qualificados em relação ao total de doutores cadastrados no grupo, tanto as categorias/níveis dos bolsistas de pesquisa¹ como os graus atribuídos aos programas de pós-graduação em que o pesquisador participa como docente foram transformados em uma escala numérica. Também foram classificados no estrato "E" todos os grupos de pesquisa participantes do segundo subconjunto, mencionado no parágrafo anterior, conforme ilustra a **Figura 2**.

O escore de qualificação (Q) é uma estatística padronizada, com média e desvio-padrão respectivamente iguais a 50 e 20, calculado a partir da média aritmética e do desvio-padrão do indicador de densidade correspondente aos grupos de igual porte no que se refere a número de doutores participantes do grupo (seis classes definidas: 1, 2, 3, 4, 5 e 6 ou mais doutores), da mesma grande área do conhecimento em que o grupo foi cadastrado. Assim, a estatística Q varia entre 0 e 100. Se aproxima de zero quando for muito rara a presença de pesquisadores qualificados no grupo e, no outro extremo, tende à marca de 100 quando for muito intensa a presença desses pesquisadores.

Foram selecionadas 4 separatrizes da estatística Q: 15, 35, 55 e 75. Tais limites foram fixados analisando-se os indicadores correspondentes aos grupos estratificados em função da estatística Q organizada em 29 faixas de amplitude igual a três. Assim, os grupos foram classificados nos diferentes estratos como segue:

Q	Estrato
Inferior a 15	E
15 --- 35	D
35 --- 55	C
55 --- 75	B
75 ou +	A

As **Tabelas 1 e 2** informam o número de grupos classificados por estrato, segundo as grandes áreas e áreas do conhecimento predominantes nas atividades do grupo, respectivamente.

¹ O tempo de titulação dos doutores não bolsistas também foi considerado (ver procedimento para o cálculo do escore de qualificação no *Saiba Mais*)

Principais resultados

Algumas variáveis selecionadas, como número de pesquisadores, doutores, bolsistas de pesquisa e docentes, em termos absolutos e relativos, são apresentados na **Tabela 3**, contrastando os diferentes perfis dos grupos de pesquisa classificados nos diferentes estratos. Dentre esses indicadores, cabe destacar a proporção de bolsistas de pesquisa em relação ao total de doutores cadastrados no grupo, bem como o percentual de docentes doutores vinculados a programas avaliados com diferentes graus em relação ao total de docentes que fazem parte do grupo. Na citada tabela foram utilizadas as seguintes legendas:

- Dout:** número de pesquisadores com titulação máxima igual a doutorado;
- BPQ:** bolsistas de pesquisa do CNPq em novembro de 2000 (modalidades PQ, PV, RD, DCR) cadastrados no *Diretório*;
- BPQ(I):** subconjunto de BPQ constituído pelos bolsistas de categoria/nível igual ou superior a 1C (ou seja, 1A, 1B ou 1C) ;
- BPQ(II):** subconjunto de BPQ constituído pelos bolsistas de categoria/nível igual ou inferior a 2A (ou seja, 2A, 2B ou 2C) ;
- DOC:** docentes cadastrados no *Diretório* com titulação máxima igual a doutorado, vinculados a programas de pós-graduação reconhecidos e avaliados pela CAPES com grau igual ou superior a 3;
- DOC(6;7):** subconjunto de DOC, formado por docentes vinculados a programas de pós-graduação avaliados com grau 6 ou 7;
- DOC(5):** subconjunto de DOC, constituído por docentes vinculados a programas de pós-graduação avaliados com grau 5;
- DOC(3;4):** subconjunto de DOC, integrado por docentes vinculados a programas de pós-graduação avaliados com grau 3 ou 4.

A **Tabela 4**, bem como as **Figuras 3 e 4**, detalham os indicadores relacionados aos bolsistas de pesquisa e aos docentes, por grande área do conhecimento, dando maior nitidez aos contrastantes perfis dos grupos classificados nos cinco estratos criados.

A já conhecida concentração geográfica dos grupos de pesquisa pode ser visualizada na **Tabela 5**. Nela, as unidades da federação (UF) foram ordenadas considerando os quantitativos dos grupos de pesquisa distribuídos nos diferentes estratos, de acordo com dois critérios. O primeiro deles leva em conta a soma do número de grupos incluídos nos estratos A e B, denominados como Consolidados [ver primeira coluna, (N)]. O segundo critério considera a proporção desses grupos no total de grupos localizados na mesma unidade da federação [ver segunda coluna, (N*)].

Pode-se constatar que 2/3 de todos os grupos vinculados às IES e 85% dos grupos consolidados se concentram em apenas quatro unidades da federação: São Paulo, Rio de Janeiro, Minas Gerais e Rio Grande do Sul.

Examinando-se os percentuais de grupos consolidados, pode-se visualizar a concentração desses grupos no eixo SP/RJ (46% e 43%, respectivamente). Logo a seguir, DF, MG e RS (39%; 38% e 30%) e depois SC, PE e BA (24%, 20% e 19%). Nessas 8 Unidades da Federação estão concentrados 80% dos grupos de pesquisa e 95% dos grupos mais qualificados (estratos "A" e "B").

A **Tabela 6** mostra a concentração dos grupos de pesquisa em termos institucionais. Nessa tabela foram listadas apenas 55 IES (de um total de 155 participantes da versão atual do *Directório*): todas aquelas que abrigam 50 ou mais grupos de pesquisa, além de todas as IES que contam com pelo menos 7 grupos de pesquisa classificados como consolidados em uma grande área do conhecimento. Atendo-se aos números relativos [percentuais da coluna (A+B)/T] destacam-se 14 instituições, em que a proporção dos grupos consolidados ultrapassa a marca de 50%: IMPA com 90%, FGV/EASP e IUPERJ com 82% e 78%, respectivamente; a seguir, UFV, PUC/RJ e UNIFESP (71%, 69% e 62%); depois, com percentuais decrescentes entre 59 e 51, encontramos: USP, UFRJ, UNICAMP, UFRGS, UFLA, FIOCRUZ, PUC/SP e UFMG.

Pode-se constatar ainda, que as 14 primeiras IES relacionadas na **Tabela 6** reúnem cerca de 50% do total de grupos cadastrados pelo *Directório* e quase 80% do subconjunto formado pelos grupos mais qualificados (estratos "A" e "B"). Dessas, apenas UFPE e UnB se localizam fora da região Sul/Sudeste. Dessas mesmas 14 IES, 13 são públicas, sendo 10 federais e 3 estaduais.

Ampliando o leque, pode ser observado, ainda na mesma tabela, que as primeiras 28 instituições aglutinam quase 2/3 do total de grupos vinculados às IES e mais de 90% do total de grupos classificados nos estratos "A" e "B".

O **Quadro 1** evidencia a participação percentual dos grupos consolidados, por IES, na esfera de cada uma das seis grandes áreas do conhecimento. Destacam-se, com percentual acima de 50:

- a) Agrárias: UFMG, UFV, UFRGS, UFSM, UFLA, UNICAMP e USP;
- b) Biológicas: UNIFESP, UFV, UFRJ, UFMG, UNICAMP, FIOCRUZ, USP, UFRGS e UFES;
- c) Exatas e da Terra: IMPA, UFSCar, USP, UNICAMP, UFRGS, UFPE e PUC/RJ;

d) Saúde: FIOCRUZ, UNIFESP, USP e UFRGS;

e) Humanidades: FGV/EASP, IUPERJ, PUC/RJ, USP, UFRJ, PUC/SP, UFRGS, UNICAMP, UnB, UFMG e UMESP;

f) Engenharias e Ciência da Computação: PUC/RJ, UFRJ, UNICAMP, UFSC, USP, UFMG e UFU;

Dentre as 22 Instituições acima relacionadas, 20 estão localizadas no Sudeste/Sul (16 no triângulo formado pelos Estados de São Paulo, Rio de Janeiro e Minas Gerais), apenas uma no Nordeste (UFPE) e outra no Centro-Oeste (UnB). São majoritariamente públicas (17 das 22 IES): 15 universidades (12 federais e 3 estaduais) e 2 institutos de pesquisa.

Relação entre Qualificação e Produtividade dos Grupos de Pesquisa

Como já foi observado, o conceito básico da estratificação proposta reside na densidade de bolsistas de pesquisa e docentes vinculados à pós-graduação. Por outro lado, seu traço metodológico mais marcante é de construir-se com base em sistemas independentes de avaliação (CAPES e CNPq). Não obstante, o Diretório busca informações sobre a produtividade dos grupos de pesquisa.

Como se verá mais adiante, as correlações entre a qualificação e a medida da produtividade dos grupos é fortíssima, seja para a totalidade dos grupos, seja para conjuntos segundo as grandes áreas. Ante esta evidência, poderia ser perguntado qual a necessidade de construirmos um escore de qualificação que não utiliza a medida direta dessa produtividade. Ocorre que para conjuntos relativamente grandes de unidades (pesquisadores, grupos, etc.), há, na literatura, evidência suficiente no sentido de que a produtividade avalia, com correção, a qualificação. Os problemas da produtividade aparecem à medida que as populações analisadas vão diminuindo de tamanho, em virtude da crescente interferência da variância da produtividade conforme diminui o número de unidades no conjunto analisado. Esses problemas atingem seu limite máximo quando pretendemos localizar um grupo individualmente numa escala de qualidade através da medida simples da produtividade. Daí a necessidade de incorporar-se ao modelo de avaliação um conjunto maior de variáveis, mais além da medida pura e simples da produtividade. Portanto, o exercício desenvolvido não deve ser interpretado como uma validação do escore "Q" do algoritmo de estratificação que está sendo apresentado *vis-à-vis* uma medida de produtividade. Almeja pura e simplesmente demonstrar que para médias de conjuntos relativamente grandes de grupos, aquele escore "Q" e a medida da produtividade correlacionam-se fortemente.

Foi então construído um escore de produtividade (P) de cada um dos grupos de pesquisa vinculados às IES. Como ponto de partida foi construído um escore de produtividade de cada pesquisador-doutor com currículo Lattes cadastrado no CNPq até 10/11/2000, considerando dez tipos de produto técnico-científico selecionados de acordo com a sua natureza e agrupados em cinco subconjuntos relativamente homogêneos, como segue:

- 1) ART - Artigos publicados em periódicos especializados com corpo editorial e sistema de *referees*.
 - a) An = artigos de circulação nacional (considerados como tal aqueles publicados em português ou sem informação sobre o idioma da publicação);

b) Ae = artigos de circulação internacional (considerados como tal aqueles publicados em outro idioma que não o português);

2) TCE – Trabalhos completos publicados em anais;

3) L&C – Livros e Capítulos de livro publicados:

a) Li = livros;

b) Cl = capítulos de livro.

4) PTC – Produção técnica desenvolvida:

a) Sf = softwares;

b) Pd = produtos tecnológicos;

c) Pc = processos tecnológicos.

5) T&D – Teses e dissertações defendidas sob orientação de pesquisadores pertencentes ao grupo:

a) Te = teses de doutorado;

b) Di = dissertações de mestrado.

Inicialmente, foi calculado um indicador de produtividade relacionado a cada um dos cinco subconjuntos de produtos formados, para cada pesquisador-doutor pertencente ao grupo ponderando-se os diferentes tipos de trabalho de acordo com os valores apresentados abaixo:

Natureza	Tipo	Ponderação (v)
ART	An	0,30
	Ae	0,70
L&C	Li	0,60
	Cl	0,40
PTC	Sf	0,33
	Pd	0,33
	Pc	0,33
T&D	Te	0,70
	Di	0,30

Em uma segunda etapa, para cada pesquisador, foram calculados cinco escores padronizados de produtividade (**S**), distintos para cada um dos cinco subconjuntos de produtos. A estes escores atribuiu-se média e desvio-padrão iguais a 50 e 20, respectivamente, tomando por base a média aritmética e o desvio-padrão dos indicadores de produtividade relacionados àqueles pesquisadores que desenvolveram pelo menos um dos tipos dos produtos pertencentes ao subconjunto, no âmbito de uma mesma grande área do conhecimento.

Na terceira etapa, os cinco escores padronizados (**S**) foram agregados em um único índice, ponderando diferenciadamente valorizações conforme a grande área do conhecimento. As ponderações (**w**) utilizadas nesta etapa são resumidas a seguir.

Grande Área	ART	TCE	L&C	PTC	T&D
Agrárias	0,30	0,10	0,15	0,10	0,35
Biológicas	0,50	0,05	0,05	0,05	0,35
Exatas e da Terra	0,50	0,05	0,05	0,05	0,35
Engenharias	0,25	0,20	0,05	0,15	0,35
Humanidades	0,30	0,05	0,25	0,05	0,35
Saúde	0,35	0,05	0,20	0,05	0,35

Em seguida, um escore de produtividade (**p**) foi calculado para cada um dos pesquisadores, de forma padronizada, com média e desvio-padrão iguais a 50 e 20, respectivamente, levando em conta a média aritmética e o desvio-padrão do índice agregado de produtividade (**U**) da grande área do conhecimento relacionada à sua área de atuação (a primeira por ele indicada no currículo).

Finalmente, o escore de produtividade do grupo (**P**) foi calculado tomando-se como *proxi* a média aritmética simples dos escores individuais (**p**) dos pesquisadores pertencentes ao grupo de pesquisa.

A **Figura 5** ilustra a relação entre as *médias* dos escores de qualificação (**Q**) e de produtividade (**P**) dos grupos de pesquisa vinculados às Instituições de Ensino Superior. Cada ponto do gráfico refere-se à *média* dos escores correspondentes a pequenos subconjuntos formados pelos grupos tabulados na distribuição de freqüências da estatística Q, com amplitude de classe igual a 5 (primeira classe: grupos com Q inferior a 20, segunda classe: Q igual ou superior a 20 mas inferior a 25, e assim, sucessivamente, até a última classe, onde estão agrupados os grupos com Q igual ou superior a 85). A fortíssima correlação entre as *médias* dos dois escores pode ser constatada visualmente. O coeficiente de correlação linear entre as médias é igual a 0,998, indicando uma relação muito forte entre as variáveis estudadas, qualificação e produtividade dos grupos de pesquisa vinculados às IES, em termos gerais. Gráficos semelhantes (**Figura 6**) são apresentados para cada uma das seis grandes áreas do conhecimento.

A **Tabela 7** reúne o total da produção técnico-científica dos grupos de pesquisa vinculados às IES, em números absolutos, por estrato, bem como a produtividade anual média por pesquisador com titulação máxima igual a doutorado.

No quadro inferior da mesma tabela é apresentado um índice simplificado de produtividade em que a média anual de cada tipo de produto foi igualada a 100. Examinando-se esses índices simplificados pode-se chegar à mesma constatação apontada anteriormente, ou seja: em geral, a produtividade dos grupos está correlacionada positivamente à qualificação dos grupos de pesquisa – quanto mais elevada a qualificação do grupo de pesquisa maior a produtividade e vice-versa.

Considerando isoladamente alguns tipos de produto, chega-se à conclusão que a diferença de produtividade entre os grupos classificados nos estratos criados seguem a mesma tendência, de forma muito mais expressiva. A produtividade anual média dos pesquisadores-doutores, em termos de artigos publicados em periódicos de circulação internacional, correspondente aos grupos classificados no estrato "A", por exemplo, representa quase o quádruplo daqueles alojados no estrato "E" (1,24/0,32). Proporções semelhantes encontramos para dissertações de mestrados concluídas sob orientação de pesquisadores do grupo (quase 6 para 1). Tal distância é ainda mais expressiva no que se refere às teses de doutorado: cerca de 30:1. A **Figura 7** ilustra esse diferencial de produtividade entre os estratos.

Anexo

Procedimentos utilizados para o cálculo dos escores de Qualificação (Q) e de Produtividade (P) dos grupos de pesquisa vinculados às IES

Procedimentos utilizados para o cálculo dos escores de Qualificação (Q) e de Produtividade (P) dos grupos de pesquisa vinculados às IES

1. Procedimentos para o cálculo do escore de Qualificação (Q)
2. Procedimentos para o cálculo do escore de Produtividade (P)
3. Exemplo para o cálculo do escore (Q)
4. Exemplo para o cálculo do escore (P)

1. Procedimentos para o cálculo do escore de Qualificação (Q)

O cálculo do escore de (Q), que procura medir o grau de qualificação dos grupos de pesquisa vinculados às IES, toma como ponto de partida as seguintes variáveis:

$$X_i = \sqrt{\frac{B_i}{\mu(B_i)} + \frac{D_i}{\mu(D_i)}}$$

A estatística X_i foi calculada somente para os grupos pertencentes ao subconjunto constituído pelos grupos de pesquisa que contam com a presença de pelo menos um doutor e, ainda, pelo menos um bolsista de pesquisa do CNPq ou um docente vinculado a um programa de pós-graduação com grau igual ou superior a 3, sendo:

$$B_i = \frac{\sum_{j=1}^n b_j \cdot w_h}{n}$$

$$D_k = \frac{\sum_{j=1}^n d_j \cdot v_g}{n}$$

$i = 1, 2, \dots, q$

q representa o total de grupos de pesquisa, de uma determinada grande área do conhecimento, que contam com a presença de pelo menos um pesquisador doutor e, ainda, pelo menos um bolsista de pesquisa ou docente de um programa de pós-graduação avaliado pela CAPES com grau igual ou superior a 3;

$k = 1, 2, \dots, r$

r representa o total de grupos de pesquisa, de uma determinada grande área do conhecimento, que contam com a presença de pelo menos um docente doutor, vinculado a um programa de pós-graduação avaliado pela CAPES com grau igual ou superior a 3;

$n = n^o$ de pesquisadores cadastrados no Diretório com titulação máxima igual a doutorado;

$b_j = n^o$ de pesquisadores doutores;

$w_h =$ ponderação, segundo a condição do pesquisador, como segue:

i) doutores bolsistas de pesquisa do CNPq (novembro de 2000):

Categoria/Nível ₁	1A	1B	1C	2A	2B	2C
w_h	1,00	0,90	0,80	0,60	0,50	0,40

ii) doutores não-bolsistas do CNPq:

Tempo de titulação (anos completos em dezembro de 2000)	w_h
10 ou +	0,40
5 --- 10	0,20
Menos de 5	0,10

d_j = nº de pesquisadores doutores, docentes cadastrados pela CAPES (1999) em programas de pós-graduação avaliados com grau igual ou superior a 3 ($j = 1, 2, \dots, n$);

v_g = ponderação, segundo o conceito do programa de pós-graduação, atribuído pela CAPES (1998), como segue:

Conceito	7	6	5MD	5M	4	3
v_g	1,00	0,90	0,75	0,65	0,45	0,20

5MD: programas de pós-graduação com conceito 5, que ministram cursos de mestrado e de doutorado, ou apenas doutorado;

5M: programas de pós-graduação com conceito 5, que ministram cursos de mestrado;

As ponderações v e w foram multiplicadas pelo fator f nos casos em que o pesquisador doutor participa de mais de 2 grupos de pesquisa, como segue:

$$f = \frac{1}{\sqrt{g-1}} \quad \text{onde,}$$

g = número de grupos em que o pesquisador doutor participa (somente para $g > 2$)

$$\text{Assim sendo, } 0 < B_i \leq 1 \quad \text{e} \quad 0 < D_k \leq 1$$

¹ Os bolsistas da modalidade DCR, categoria 3/níveis A e C, foram computados como 2C; os da modalidade PV categoria 1/níveis D, E e F, como 1A, 1B e 1C, respectivamente, e os de categoria 2/níveis D, E e F, como 2A, 2B e 2C, respectivamente; os da modalidade RD, categoria 2/nível F, como 2C.

$\mu(\mathbf{B}_i)$ = média aritmética da estatística **B** correspondente ao subconjunto formado pelos grupos de pesquisa que contam com a presença de pelo menos um docente doutor, vinculado a um programa de pós-graduação avaliado pela CAPES com grau igual ou superior a 3;

$\mu(\mathbf{D}_i)$ = média aritmética da estatística **D** correspondente ao subconjunto constituído pelos grupos de pesquisa que contam com a participação de pelo menos um docente doutor vinculado a um programa de pós-graduação avaliado com grau igual ou superior a 3;

As médias $\mu(\mathbf{B}_i)$ e $\mu(\mathbf{D}_i)$, calculadas para cada grande área do conhecimento, são apresentadas na Tabela 1 , a seguir:

Tabela 1 – Médias aritméticas das estatísticas **B_i** e **D_i** , por grande área do conhecimento

Grande Área	Média (B)	Média (D)
Agrárias	0,3288	0,3723
Biológicas	0,4040	0,4351
Exatas	0,4048	0,4434
Engenharias	0,3750	0,4150
Humanidades	0,4110	0,3968
Saúde	0,3449	0,3413

Em uma segunda etapa, o escore de qualificação (**Q**) de cada grupo foi calculado padronizando a variável **X** a ele referente, tomando por base a média aritmética e o desvio-padrão correspondentes aos grupos de igual porte (em termos do número de doutores pertencentes ao grupo, 6 estratos foram fixados: 1, 2 , 3, 4, 5 e 6 ou mais doutores por grupo) e da mesma grande área do conhecimento em que foi classificado, de acordo com as fórmulas apresentadas a seguir:

$$Q_i = 50 + 20z_i$$

sendo,

$$z_i = \frac{X_i - \mu(X)_s}{\sigma(X)_s}$$

$$-2,5 \leq z_i \leq 2,5$$

$$0 \leq Q_i \leq 100$$

$\mu(\mathbf{X})_s$ e $\sigma(\mathbf{X})_s$: média aritmética e desvio-padrão, respectivamente, da estatística **X**, referentes ao **s**ésimo estrato (porte em termos do nº de doutores pertencentes ao grupo) da grande área do conhecimento predominante em que o grupo foi classificado. Esses parâmetros são apresentados na Tabela 2, a seguir.

Tabela 2 – Parâmetros utilizados para padronizar a estatística "X" no cálculo do escore de qualificação (Q).

Dout./Grupo	Ciências Agrárias		Ciências Biológicas		C. Exatas e da Terra	
	Média (X)	Desvio-Padrão (X)	Média (X)	Desvio-Padrão (X)	Média (X)	Desvio-Padrão (X)
1	1,5162	0,3457	1,5662	0,3412	1,5753	0,3104
2	1,3314	0,3311	1,3746	0,3285	1,3731	0,3413
3	1,3950	0,3329	1,2898	0,3053	1,2939	0,3309
4	1,3110	0,3100	1,2554	0,3027	1,2569	0,3240
5	1,2930	0,2927	1,2038	0,3074	1,1947	0,2710
6 ou +	1,2923	0,2844	1,2180	0,2434	1,2447	0,2718

Dout./Grupo	Engenharias e C. da Computação		Humanidades		Ciências da Saúde	
	Média (X)	Desvio-Padrão (X)	Média (X)	Desvio-Padrão (X)	Média (X)	Desvio-Padrão (X)
1	1,4406	0,3967	1,3888	0,3787	1,4217	0,3498
2	1,3219	0,3460	1,2640	0,3395	1,2881	0,3229
3	1,2920	0,3509	1,2414	0,3095	1,2683	0,3082
4	1,3094	0,2966	1,1927	0,2986	1,2207	0,2929
5	1,2702	0,2766	1,1891	0,3071	1,2239	0,2601
6 ou +	1,3003	0,2818	1,2069	0,2797	1,2551	0,2570

2. Procedimentos para o cálculo do escore de Produtividade (P)

O escore de produtividade (p) foi calculado, primeiramente, para cada um dos pesquisadores doutores com currículo cadastrado na Plataforma Lattes/CNPq (até 10 novembro de 2000), agrupando dez tipos de produto técnico-científico de diferentes naturezas em 5 subconjuntos relativamente homogêneos, como segue:

- 1) ART - Artigos publicados em periódicos especializados com corpo editorial e sistema de *referees*:
 - a) An = artigos de circulação nacional (considerados como tal aqueles publicados em português ou sem informação sobre o idioma da publicação);
 - b) Ae = artigos de circulação internacional (considerados como tal aqueles publicados em outro idioma que não o português);
- 2) TCE – Trabalhos completos publicados em anais;
- 3) L&C – Livros e Capítulos de livro publicados:
 - a) Li = livros;

b) Cl = capítulos de livro.

4) PTC – Produção técnica desenvolvida:

a) Sf = softwares;

b) Pd = produtos tecnológicos;

c) Pc = processos tecnológicos.

5) T&D – Teses e dissertações defendidas sob orientação de pesquisadores pertencentes ao grupo:

a) Te = teses de doutorado;

b) Di = dissertações de mestrado.

Para cada subconjunto foi calculado um indicador de produtividade por pesquisador doutor pertencente ao grupo de pesquisa, conforme a fórmula apresentada a seguir, ponderando-se os diferentes tipos de trabalho pertencentes ao subconjunto formado, de acordo com os valores apresentados na Tabela 3, adiante:

$$Y = \sqrt{\text{Log} (1 + v \cdot T)}$$

Log = logaritmo decimal;

v = ponderação entre os diferentes tipos de trabalho da mesma natureza (Tabela 3);

T = tipos de trabalho da mesma natureza (ver Tabela 3);

Dessa forma, a estatística Y_1 , correspondente aos Artigos publicados em periódicos especializados – ART, por exemplo, foi calculada como segue:

$$Y_1 = \sqrt{\text{Log} [1 + (0,3 An + 0,7 Ae)]}$$

An representa o número de artigos publicados em periódicos de circulação nacional e **Ae** o de circulação internacional².

Para os demais produtos foi utilizado o mesmo procedimento, com base nas ponderações apresentadas na Tabela 3:

² An: artigos de circulação nacional (considerados como tal aqueles publicados em português ou sem informação sobre o idioma da publicação); Ae: artigos de circulação internacional (considerados como tal aqueles publicados em outro idioma que não o português).

Tabela 3 - Ponderações internas aos subconjuntos de produtos formados, segundo os diferentes tipos de produto pertencentes ao subconjunto

Natureza	Tipo	Ponderação (v)
ART	An	0,30
	Ae	0,70
L&C	Li	0,60
	Cl	0,40
PTC	Sf	0,33
	Pd	0,33
	Pc	0,33
T&D	Te	0,70
	Di	0,30

Em uma segunda etapa, foram calculados cinco escores padronizados de produtividade (**S**) por pesquisador doutor, distintos para cada um dos cinco subconjuntos de produtos, com base na média aritmética e no desvio-padrão da variável **Y** relacionada aos pesquisadores que desenvolveram pelo menos um dos tipos dos produtos pertencentes ao subconjunto, no âmbito da mesma grande área do conhecimento (relacionada à primeira especialidade de atuação informada pelo pesquisador em seu currículo). Tais parâmetros são apresentados na Tabela 4, adiante.

Os escores de produtividade padronizados foram calculados utilizando as seguintes fórmulas:

$$S_i = 50 + 20z_i$$

onde, $i = 1, 2, \dots, m$

m representa o nº de pesquisadores classificados na grande área do conhecimento que geraram, durante o período de janeiro de 1997 a dezembro de 1999, pelo menos um dos tipos de produto pertencentes ao subconjunto de que faz parte;

$$z_i = \frac{Y_i - \mu(Y_i)}{\sigma(Y_i)} \quad -2,5 \leq z_i \leq 2,5$$

Tabela 4 – Parâmetros para o cálculo dos escores padronizados de produtividade (S), por subconjunto de produtos da mesma natureza e grande área do conhecimento

Grande área	Parâmetros 1/	Artigos em periódicos especializados	Trabalhos completos em eventos	Livros e capítulos de livro	Produção Técnica	Teses e dissertações orientadas
Ciências Agrárias	M	0,6859	0,5461	0,5343	0,4331	0,5875
	D	0,1923	0,2812	0,1438	0,1092	0,1721
Ciências Biológicas	M	0,7326	0,2798	0,5083	0,4299	0,5497
	D	0,1892	0,3002	0,1419	0,1186	0,1587
Ciências Exatas e da Terra	M	0,7361	0,5182	0,4848	0,4268	0,5313
	D	0,2024	0,2640	0,1170	0,1085	0,1554
Engenharias e C. da Computação	M	0,6264	0,8850	0,4855	0,4472	0,5712
	D	0,1865	0,1869	0,1225	0,1211	0,1701
Humanidades	M	0,5493	0,5347	0,5621	0,4337	0,5638
	D	0,1592	0,2250	0,1470	0,1123	0,1667
Ciências da Saúde	M	0,6894	0,3252	0,5607	0,4195	0,5662
	D	0,2000	0,3094	0,1620	0,1165	0,1606

1/ M: média aritmética(Y); D: desvio-padrão(Y).

A seguir, os cinco escores padronizados (S_i) foram agregados em uma única estatística (U), ponderando-os de forma diferenciada, de acordo com a grande área do conhecimento em que o pesquisador foi classificado, como segue. As ponderações adotadas (w) estão na Tabela 5, logo adiante.

$$U_i = \sum S_i * w_j$$

$$i = 1, 2, \dots, m \quad j = 1, 2, 3, 4, 5.$$

m representa o nº de pesquisadores classificados em uma grande área do conhecimento e que geraram, durante o período de janeiro de 1997 a dezembro de 1999, pelo menos um dos dez tipos de produto selecionados.

Tabela 5 - Ponderações utilizadas para o cálculo da estatística (U)

Grande Área	ART	TCE	L&C	PTC	T&D
Agrárias	0,30	0,10	0,15	0,10	0,35
Biológicas	0,50	0,05	0,05	0,05	0,35
Exatas e da Terra	0,50	0,05	0,05	0,05	0,35
Engenharias	0,25	0,20	0,05	0,15	0,35
Humanidades	0,30	0,05	0,25	0,05	0,35
Saúde	0,35	0,05	0,20	0,05	0,35

A seguir, um escore de produtividade (**p**) foi calculado para cada um dos pesquisadores doutores, com o auxílio da seguinte fórmula:

$$p_i = 50 + 20z_i$$

onde:

$$z_i = \frac{U_i - \mu(U)}{\sigma(U)}$$

$$-2,5 \leq z_i \leq 2,5$$

$$0 \leq p_i \leq 100$$

$\mu(\mathbf{U})$ e $\sigma(\mathbf{U})$: média aritmética e desvio-padrão, respectivamente, do índice agregado de produtividade (**U**) correspondente à grande área do conhecimento em que o pesquisador foi classificado. A Tabela 6, traz esses parâmetros.

Tabela 6 – Parâmetros da estatística (U) utilizados no cálculo do escore (p)

Grande área de atuação do pesquisador	Média (U)	Desvio-padrão (U)
Ciências Agrárias	31,3953	18,0033
Ciências Biológicas	33,7140	19,8012
Ciências Exatas e da Terra	33,0516	19,4411
Engenharias e C. da Computação	31,3301	18,1376
Humanidades	31,2117	18,9388
Ciências da Saúde	31,7339	18,5350

Finalmente, um escore de produtividade (**P**) foi calculado para cada grupo de pesquisa tomando-se a média-aritmética simples dos escores (**p**) correspondentes aos pesquisadores doutores pertencentes ao grupo de pesquisa, com currículos cadastrados na Plataforma Lattes/CNPq até 10 de novembro de 2000.

3. Exemplo para o cálculo do escore padronizado Q de um grupo hipotético

Dados básicos:

- a) Grupo constituído por 7 pesquisadores, dos quais 5 com titulação máxima igual a doutorado (n = 5);
- b) Dos 5 doutores, 3 são bolsistas de pesquisa do CNPq e 2 não são bolsistas;
- c) Dos 5 doutores, 4 estão cadastrados pela CAPES como docentes da pós-graduação em cursos com grau igual a 3 ou +;
- d) Grande área predominante do grupo: Ciências Agrárias.

Condição 1/	(w)	Nº de doutores (b)	(g) 2/	(b)*(w) RQ[(g)-1]
BPQ 1A	1,00	1	2	1,000
BPQ 1B	0,90			-
BPQ 1C	0,80			-
BPQ 2A	0,60	1	3	0,424
BPQ 2B	0,50			-
BPQ 2C	0,40	1	1	0,400
Dout. 10 ou +	0,40	1	1	0,400
Dout. 5;10	0,20			-
Dout. -5	0,10	1	1	0,100
Soma		5		2,324

Conceito	(v)	Nº de docentes (d)	(g) 2/	(d)*(v) RQ[(g)-1]
7	1,00	1	2	1,000
6	0,90	1	1	0,900
5MD	0,75	1	3	0,530
5M	0,65	1	5	0,325
4	0,45			-
3	0,20			-
Soma		4		2,755

1/ Condição: BPQ: Doutores apoiados com bolsa de pesquisa pelo CNPq;

Dout.: Doutores não apoiados com bolsa de pesquisa, segundo o tempo de titulação em anos completos (dezembro de 2000); Dout. 10 ou + = dez ou mais anos completos de titulação;

Dout. 5;10 = entre cinco anos completos e menos de 10); Dout. -5 = menos de cinco anos.

2/ g = nº de grupos em que o pesquisador atua (utilizado quando g>2).

$$B = 2,324/5 = 0,4649$$

$$D = 2,755/5 = 0,5511$$

Tabela 1:

Média(B) Agr: 0,3288

Média(D) Agr: 0,3723

$$X = \text{Raiz}[B/\text{Média}(B) + D/\text{Média}(D)] = \text{Raiz}(0,4649/0,3288 + 0,5511/0,3723) = 1,7012$$

Tabela 2:

Ciências Agrárias		
Doutores	Média(X)	DP(X)
5	1,293	0,2927

$$z = [X - \text{Média}(X)]/\text{DP}(X) = (1,7012 - 1,2930)/0,2927 = 1,394$$

$$Q = 50 + 20z = 50 + 20 * 1,394 = 77,9$$

Legenda: DP = desvio-padrão; RQ = raiz quadrada.

4. Exemplo para o cálculo do escore padronizado de produtividade (p) de um pesquisador hipotético

Tipo de Trabalho	(w) 1/	Nº de Trabalhos (T)	(w) * (T)
1. ART - Artigos publicados em periódicos especializados			
An	0,3	1	0,30
Ae	0,7	3	2,10
$Y_1 = \text{raiz}\{\text{Log}[1+\text{Soma}(wT)]\} =$	0,7290		2,40
2. TCE - Trabalhos completos apresentados em eventos e publicados			
Tc		4	
$Y_3 = \text{raiz}\{\text{Log}[1+T]\} =$	0,8360		
3. L&C - Livros e capítulos de livros publicados			
Li	0,60	1	0,60
Cl	0,40	2	0,80
$Y_3 = \text{raiz}\{\text{Log}[1+\text{Soma}(wT)]\} =$	0,6166		1,40
4. PTC - Produção técnica			
Sf	0,33	1	0,33
Pd	0,33		-
Pc	0,33		-
$Y_4 = \text{raiz}\{\text{Log}[1+\text{Soma}(wT)]\} =$	0,3533		0,33
5. T&D - Teses e Dissertações orientadas			
Te	0,7	1	0,70
Di	0,3	3	0,90
$Y_5 = \text{raiz}\{\text{Log}[1+\text{Soma}(wT)]\} =$	0,6442		1,60

1/ Ver Tabela 3.

Tipo trabalho		Y_i	Média(Y)agr 2/	DP(Y)agr 2/
1.	Y_1	0,7290	0,6859	0,1923
2.	Y_2	0,8360	0,5461	0,1438
3.	Y_3	0,6166	0,5343	0,1438
4.	Y_4	0,3533	0,4331	0,1092
5.	Y_5	0,6442	0,5875	0,1721

2/ Ver Tabela 4.

S	z	S_i
S_1	0,224	54,49
S_2	2,016	90,33
S_3	0,572	61,45
S_4	(0,731)	35,39
S_5	0,329	56,59

Tipo trabalho	S_i	w(agr) 3/	w(agr) * S_i 3/
1.	54,49	0,30	16,3456
2.	90,33	0,10	9,0326
3.	61,45	0,15	9,2172
4.	35,39	0,10	3,5387
5.	56,59	0,35	19,8056
U			57,9397

3/ Ver Tabela 5.

Ciências Agrárias:	
Média(U)	DP(U)
31,3953	18,0033

Ver Tabela 6

$$z = [X - \text{Média}(U)] / \text{DP}(U) = (57,9397 - 31,3953) / 18,0033 = 1,474$$

$$p = 50 + 20z = 50 + 20 * 1,474 = 79,5$$