

A avaliação da produção científica nas subáreas da Saúde Coletiva: limites do atual modelo e contribuições para o debate

Evaluation of scientific production in different subareas of Public Health: limits of the current model and contributions to the debate

La evaluación de la producción científica en subáreas de la Salud Colectiva: límites del modelo actual y contribuciones para el debate

Jorge Alberto Bernstein Iriart ¹

Suely Ferreira Deslandes ²

Denise Martin ³

Kenneth Rochel de Camargo Jr. ⁴

Marília Sá Carvalho ⁵

Cláudia Medina Coeli ⁶

Abstract

The aim of this study was to discuss the limits of the quantitative evaluation model for scientific production in Public Health. An analysis of the scientific production of professors from the various subareas of Public Health was performed for 2010-2012. Distributions of the mean annual score for professors were compared according to subareas. The study estimated the likelihood that 60% of the professors in the graduate studies programs scored P50 (Very Good) or higher in their area. Professors of Epidemiology showed a significantly higher median annual score. Graduate studies programs whose faculty included at least 60% of Epidemiology professors and fewer than 10% from the subarea Social and Human Sciences in Health were significantly more likely to achieve a "Very Good" classification. The observed inequalities in scientific production between different subareas of Public Health point to the need to rethink their evaluation in order to avoid reproducing inequities that have harmful consequences for the field's diversity.

Evaluation of Research Programs and Tools; Researcher Performance Evaluation Systems; Scientific Publication Indicators

Resumo

O objetivo foi discutir os limites do modelo de avaliação quantitativa da produção científica na Saúde Coletiva. Foi realizada uma análise da produção científica dos docentes nas subáreas da Saúde Coletiva no triênio 2010-2012. Foram comparadas as distribuições da pontuação anual média de docentes segundo subáreas. Estimou-se a probabilidade de 60% dos docentes do programa de pós-graduação (PPG) terem pontuação igual ou acima do P50 da área (Muito Bom). A subárea Epidemiologia apresentou mediana da pontuação anual dos docentes significativamente maior. PPGs cuja composição inclui 60% ou mais de docentes da subárea Epidemiologia e menos de 10% da subárea Ciências Sociais e Humanas em Saúde apresentaram probabilidade significativamente maior de alcançar o nível Muito Bom. As desigualdades constatadas nas medidas da produção científica entre as subáreas da Saúde Coletiva apontam para a necessidade de se repensar a avaliação, de forma a não reproduzir iniquidades que gerem consequências nocivas para a diversidade do campo.

Avaliação de Programas e Instrumentos de Pesquisa; Sistemas de Créditos e Avaliação de Pesquisadores; Indicadores de Produção Científica

¹ Instituto de Saúde Coletiva, Universidade Federal da Bahia, Salvador, Brasil.

² Instituto Nacional de Saúde da Mulher, da Criança e do Adolescente Fernandes Figueira, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

³ Departamento de Medicina Preventiva, Universidade Federal de São Paulo, São Paulo, Brasil.

⁴ Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil.

⁵ Programa de Computação Científica, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

⁶ Instituto de Estudos em Saúde Coletiva, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.

Correspondência

J. A. B. Iriart

Instituto de Saúde Coletiva, Universidade Federal da Bahia.

Rua Basílio da Gama s/n, 2º andar, Salvador, BA 40110-040, Brasil.
iriart@ufba.br

“Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing that it is stupid” (autoria desconhecida).

Introdução

A Saúde Coletiva se constituiu como um campo de produção de conhecimentos científicos e práticas marcado pela interdisciplinaridade ¹. Tem, como seus eixos estruturantes, as subáreas da Epidemiologia, Ciências Sociais e Humanas em Saúde (CSHS) e Políticas, Planejamento e Gestão em Saúde (PPGS), recebendo contribuições diretas também das disciplinas da Biomedicina, Biologia, Estatística, dentre outras. A identidade do campo se construiu em torno do conceito de Saúde Coletiva em sua transcendência aos limites do conceito de Saúde Pública ². Como valores estruturantes, apresenta perspectiva crítica na abordagem das desigualdades e inequidades em saúde e o compromisso com a construção do Sistema Único de Saúde (SUS), valendo-se de um olhar ampliado sobre o processo saúde-doença, no qual a abordagem das várias subáreas (e disciplinas auxiliares) contribui para a construção de uma visão complexa, multidimensional e enraizada em contextos socioculturais, históricos, econômicos e políticos. Apesar de partilharem valores e confluem para objetivos comuns na produção de conhecimento, os três subcampos constituintes da Saúde Coletiva possuem diferentes perspectivas epistemológicas, teóricas e metodológicas para a análise da relação saúde-doença-atenção, gerando culturas acadêmicas distintas, com processos de trabalho e tempos diversos para a produção do conhecimento científico ³. Esta heterogeneidade, que caracteriza a Saúde Coletiva, é parte de sua riqueza e se revela fundamental para abordar a complexidade do objeto saúde. Todavia, as disputas internas se fazem presentes na delimitação dos rumos e perspectivas do próprio campo. Entendemos a Saúde Coletiva a partir da noção de campo científico, ou seja, como contexto social com suas intrínsecas relações de força, de monopólio e de lutas (para manutenção ou transformação dessas relações de poder). O que sempre estará em jogo é a capacidade técnica e o poder social de falar e agir de maneira autorizada e com autoridade sobre os objetos e temas, como um agente legitimado pela ciência ⁴.

O crescimento da pós-graduação no Brasil nas últimas décadas aponta o indiscutível amadurecimento e vigor do campo. Nesse âmbito, desenvolve-se parte extremamente significativa da produção acadêmica da Saúde Coletiva e cri-

térios importantes para definir padrões da produção acadêmica. Se, na década de 1970, havia menos de 10 programas de pós-graduação (PPG), em 2013 se contabilizavam 75, sendo 13 (17%) de mestrados acadêmicos, 32 (43%) mestrados profissionais, 28 (37%) mestrados e doutorados e dois programas de doutorado em associação. Contudo, os programas refletem forte concentração regional, sendo 51% deles sediados no Sudeste e apenas 2% no Norte ⁵.

A avaliação Capes (Coordenação de Aperfeiçoamento de Pessoal em Nível Superior) tem desempenhado papel importante no desenvolvimento da pós-graduação brasileira, mas, como qualquer processo avaliativo de longa duração, deve ser periodicamente repensada para o aprimoramento dos mecanismos de atribuição de valor e diferenciação entre objetos avaliados. Essa revisão é necessária, sobretudo, para que se evite a cristalização de parâmetros e critérios e para que se discutam criticamente as consequências que produzem, evitando distorções ^{6,7}. A avaliação tem um papel indutor importante no desenvolvimento da área, pois os critérios utilizados orientam os caminhos para a organização dos programas e apontam tendências.

A produtividade científica corresponde a 35% da nota na avaliação dos PPG em Saúde Coletiva e tem sido um indicador importante para diferenciar a nota entre eles, baseando-se fundamentalmente em critérios quantitativos. Como os valores de referência para atribuição de notas são definidos retrospectivamente ao final do período de avaliação, ao se totalizarem os indicadores de produção, há uma tendência em impulsionar, de forma crescente, a cobrança da produtividade científica dos pesquisadores, buscando antecipar uma meta sempre superior à do triênio anterior.

Nos últimos dois triênios, o crescimento da produção científica da área tem sido exponencial, e as medidas que servem de base para os indicadores utilizados para mensurar a produção científica (mediana da produção, percentil 80 e média da produção *per capita*) refletem esse crescimento. No triênio 2007-2009, a mediana de produção da área foi de 390 pontos, subindo para 620 no triênio 2010-2012. Vale ressaltar novamente que a mediana é uma informação fornecida *a posteriori*, ou seja, após a avaliação trienal. Dessa forma, os pesquisadores não têm conhecimento da mediana que resultará em ponto de corte para critérios de produtividade de produção intelectual. Esse dado quantitativo, juntamente com a produção *per capita* e o percentual de docentes no P80, resulta na avaliação da produção nos níveis Muito Bom, Bom, Regular, Fraco e Deficiente ⁵.

A diversidade da conformação dos PPG em Saúde Coletiva é outro aspecto que deve ser levado em conta na análise crítica sobre os critérios de quantificação adotados. Considerando-se a totalidade dos PPG nessa área, verifica-se que 47 possuem os três eixos do campo, 17 são temáticos e 11 são “disciplinares” (com apenas um dos eixos) ⁵.

Neste artigo, propomos discutir especificamente os limites das formas adotadas para a quantificação da produção científica na avaliação de uma área composta por subáreas heterogêneas. A avaliação quantitativa tem partido do pressuposto de que as diferenças existentes entre as subáreas da Saúde Coletiva não se refletem no quantitativo da produção científica dos docentes e programas, de modo que todos podem ser medidos pela mesma régua.

Nosso objetivo foi, com base na análise quantitativa da produção científica dos docentes nas subáreas da Saúde Coletiva, verificar se há diferenças significativas, ou não, nas medidas dessa produção entre elas. Nossa hipótese é que, dadas as distinções na forma de produção e divulgação científica entre as subáreas Epidemiologia e CSHS/PPGS, existe uma diferença importante na mediana da produção entre essas subáreas.

Método

Foi realizada uma análise seccional da produção científica dos docentes na pós-graduação senso estrito nas subáreas da Saúde Coletiva, no triênio 2010-2012. Para classificar todos os professores do quadro permanente dos PPG nas três subáreas constitutivas do campo, foi realizada uma consulta aos coordenadores de pós-graduação na área da Saúde Coletiva, solicitando que cada um classificasse os docentes de seu programa.

Os coordenadores foram orientados a classificá-los em quatro categorias: (1) Epidemiologia; (2) Políticas, Planejamento e Gestão em Saúde; (3) Ciências Sociais e Humanas em Saúde; e (4) Outras. Esta última opção deveria ser considerada nos casos em que o pesquisador não se reconhecesse em nenhuma das três subáreas propostas. Sugerimos que a classificação fosse feita principalmente pela subárea de atuação do docente: disciplinas ensinadas, projetos e produção científica. No caso dos pesquisadores que transitavam entre as áreas, pedimos para considerar, na classificação, a subárea predominante em que eles estavam produzindo/publicando. No triênio 2010-2012, 40 programas acadêmicos de pós-graduação em Saúde Coletiva foram avaliados pela comissão de avaliação da Capes. O coordenador de cada um desses programas re-

cebeu uma planilha com os nomes dos docentes permanentes que integraram o programa no triênio 2010-2012.

No mês de abril de 2014, recebemos 33 planilhas preenchidas. Somente sete programas não as devolveram. Um programa foi excluído da enquete por ter sido descredenciado na última avaliação, e os seis restantes foram classificados pelos pesquisadores por meio de consulta aos currículos Lattes. Considerou-se, na classificação, a área de atuação, ensino, pesquisa e produção científica do docente. A classificação foi realizada de forma independente por dois pesquisadores; nos casos em que não houve consenso, recorreu-se a um terceiro pesquisador. No total, na pesquisa, foram incluídos 39 dos 40 programas avaliados na trienal 2013, perfazendo um total de 748 docentes. Empregando essa classificação, calculamos a proporção de docentes segundo subárea, tanto no conjunto da área, como dentro de cada programa.

Com base nas planilhas com a pontuação anual dos docentes utilizadas na avaliação Capes, foi calculada a pontuação média de cada um no triênio. Essa média foi calculada dividindo-se o total de pontos pelo número de anos em que o professor fez parte do corpo docente permanente no triênio (variando de um a três anos). Os resultados que obtivemos são um pouco diferentes daqueles divulgados pela comissão da avaliação trienal, em função de termos usado método de cálculo ligeiramente diverso.

Análise de dados

Foram realizadas análises comparativas segundo subáreas e segundo programa. Para as análises segundo subáreas, os docentes que participaram em mais de um programa foram considerados apenas uma vez.

O teste de Kruskal-Wallis foi empregado para avaliar diferenças nas distribuições da pontuação média anual de docentes no triênio, segundo subárea. Por sua vez, o teste qui-quadrado foi empregado para avaliar diferenças nos indicadores, proporção de docentes com pontuação igual ou acima da mediana e do percentil 80 da área.

Para avaliar a influência da composição interna do programa segundo subáreas, na probabilidade de este obter o conceito Muito Bom no indicador distribuição de publicações qualificadas em relação ao corpo docente permanente do programa, isto é, ter pelo menos 60% dos seus docentes permanentes alcançando pontuação anual média igual ou acima da mediana da área, foram criadas duas classificações: (1) apresentar 60% ou mais de docentes classificados na subárea Epidemiologia e menos de 10% classificados

na subárea CSHS ou PPGS; (2) apresentar 60% ou mais de docentes classificados na subárea Epidemiologia e simultaneamente menos de 10% em CSHS e PPGS. Essa opção foi feita em virtude da grande variedade observada na composição interna dos programas, o que inviabilizou uma categorização mais refinada.

Os pontos de corte empregados para a construção da variável composição foram definidos a partir da inspeção da distribuição da proporção de docentes segundo subáreas, de forma que o valor 1 na variável viesse a expressar um programa com nítida predominância da subárea Epidemiologia e, ao mesmo tempo, apresentasse baixa participação de docentes da subárea CSHS e/ou PPGS, quando comparado aos demais programas da área. A associação entre a composição do programa e a obtenção do conceito Muito Bom no indicador distribuição de publicações qualificadas em relação ao corpo docente permanente do programa (ter pelo menos 60% dos seus docentes permanentes alcançando pontuação anual média igual ou acima da mediana da área) foi avaliada para cada classificação, empregando razões de chances estimadas por meio de modelos de regressão logística.

Todas as análises foram realizadas empregando-se o software R versão 3.0.2 (The R Foundation for Statistical Computing, Viena, Áustria; <http://www.r-project.org>).

Resultados

Dos 748 docentes do quadro permanente da pós-graduação em Saúde Coletiva no triênio 2010-2012, 48,9% (365) são da Epidemiologia, 19,5% (146) da subárea PPGS, 16,8% (126) da CSHS e 14,8% (111) pertencem à categoria Outras. Esta última reuniu aqueles que atuam em pesquisa clínica, biológica, farmácia, bioquímica, engenharia e meio ambiente, saúde de ecossistemas, toxicologia ocupacional e ambiental e saúde do trabalhador. Quase a metade pertence à subárea Epidemiologia; entretanto, quando analisamos essa distribuição dentro de cada programa, observamos uma grande variedade de arranjos, com alguns programas apresentando mais de 70% de docentes na subárea Epidemiologia, enquanto outros apresentam menos de 30% (Figura 1).

Considerando a área como um todo, os valores de mediana e do percentil 80 da pontuação média anual dos docentes permanentes foram iguais a, respectivamente, 195 e 405,3 pontos. Observa-se diferença significativa ($p < 0,001$) entre as subáreas em relação às distribuições da pontuação média anual dos docentes perma-

nentes (Tabela 1; Figura 2). Para todas as subáreas, observa-se assimetria positiva, embora esta seja mais acentuada na subárea Epidemiologia (Figura 2). Os valores de mediana e dos percentis 25 e 75 são bem mais elevados na subárea Epidemiologia, quando comparados aos valores observados nas demais subáreas (Tabela 1). A proporção de docentes da subárea Epidemiologia que atinge pontuação igual ou acima da mediana da área é cerca de 1,5 vez maior do que as proporções observadas nas demais subáreas ($p < 0,001$) (Tabela 1). Essas diferenças ainda se tornam mais marcantes quando se analisa a proporção de docentes que atinge pontuação igual ou acima do percentil 80 da área. O valor observado para a subárea Epidemiologia é aproximadamente o dobro, o triplo e o quádruplo dos valores observados, respectivamente, para as subáreas Outras, CSHS e PPGS ($p < 0,001$) (Tabela 1).

Onze programas apresentaram 60% ou mais de docentes da subárea Epidemiologia e menos de 10% classificados na subárea CSHS ou na subárea PPGS. Esses programas apresentaram 4,4 (IC95%: 0,99-19,56) vezes mais chance de atingir o nível Muito Bom, quando comparados aos demais. Seis programas atenderam ao critério mais restrito – Epidemiologia > 60% e CSHS e PPGS < 10% –; quando esse critério de classificação foi utilizado, a razão de chances de obter o conceito Muito Bom foi de 6,25 (IC95%: 0,96-40,80) (Tabela 2).

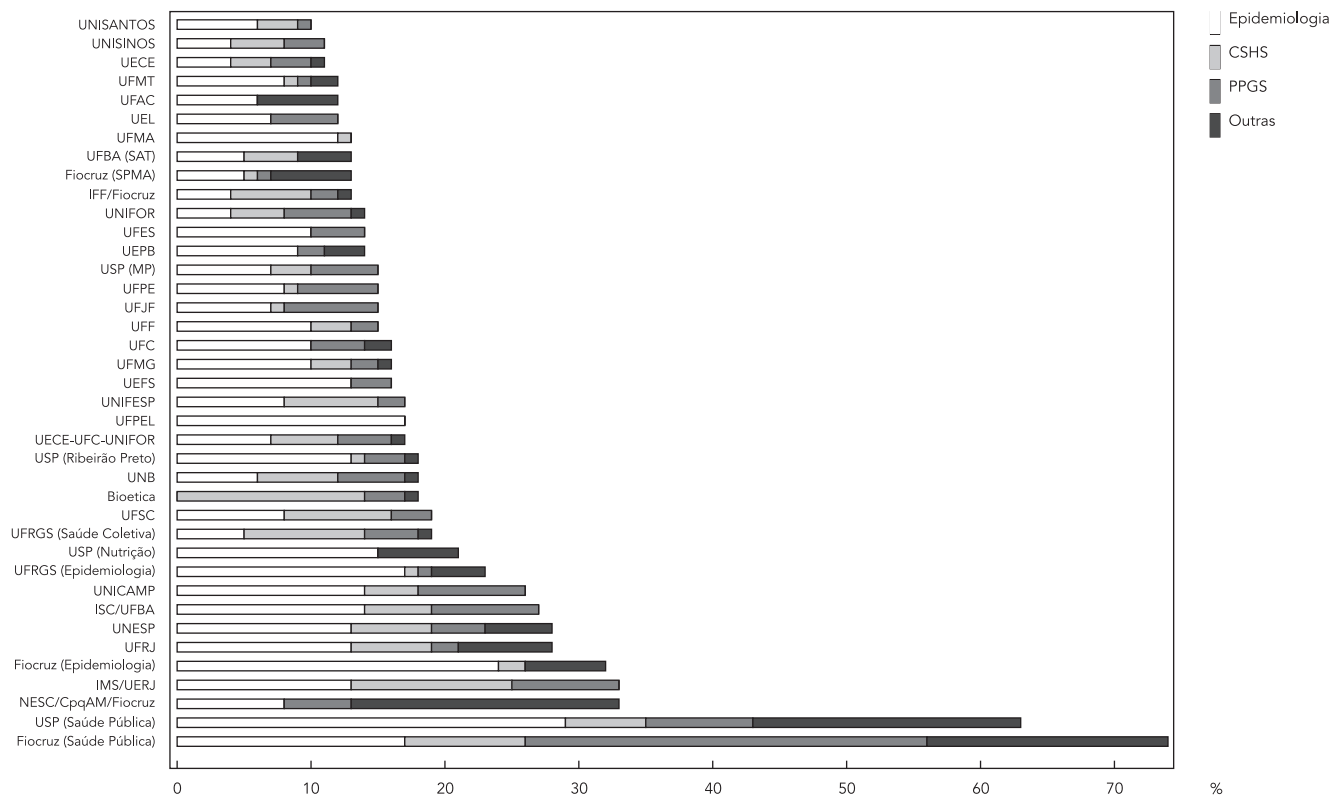
Discussão

Os resultados da pesquisa demonstram claramente a existência de desigualdades no quantitativo da produção científica entre as subáreas que constituem a Saúde Coletiva, em especial com relação à subárea Epidemiologia, que se destaca frente às demais. Tentamos estabelecer duas linhas de argumentação reflexiva: entender melhor as possíveis causas dessas disparidades e refletir sobre os efeitos não antecipados e indesejáveis que produzem, seja para os pesquisadores, seja para os PPG em Saúde Coletiva.

As diferenças encontradas na pontuação da produtividade das subáreas podem ser explicadas, sobretudo, pelas distintas formas de produção entre elas. O primeiro ponto a considerar é a diferença no número de coautores por artigo entre as subáreas. Camargo Jr. et al.⁸ mostraram que a mediana no número de coautores em artigos epidemiológicos era o dobro da mediana verificada em artigos da área de CSHS (4 e 2, respectivamente). Os padrões de autoria apontam diferenças nos processos de trabalho acadêmico segundo cada subárea, indicando formas dis-

Figura 1

Distribuição de docentes permanentes segundo programa e subárea da Saúde Coletiva no triênio 2010-2012.



CSHS: Ciências Sociais e Humanas em Saúde; Fiocruz: Fundação Oswaldo Cruz; IFF/Fiocruz: Instituto Nacional de Saúde da Mulher, da Criança e do Adolescente Fernandes Figueira/Fiocruz; IMS/UERJ: Instituto de Medicina Social/Universidade do Estado do Rio de Janeiro; ISC/UFBA: Instituto de Saúde Coletiva/UFBA; NESC/CpqAM/Fiocruz: Departamento de Saúde Coletiva/Centro de Pesquisas Aggeu Magalhães/Fiocruz; PPGS: Políticas, Planejamento e Gestão em Saúde; UECE: Universidade Estadual do Ceará; UEFS: Universidade Estadual de Feira de Santana; UEL: Universidade Estadual de Londrina; UEPB: Universidade Estadual da Paraíba; UFAC: Universidade Federal do Acre; UFBA: Universidade Federal da Bahia; UFC: Universidade Federal do Ceará; UFES: Universidade Federal do Espírito Santo; UFF: Universidade Federal Fluminense; UFJF: Universidade Federal de Juiz de Fora; UFMA: Universidade Federal do Maranhão; UFMG: Universidade Federal de Minas Gerais; UFMT: Universidade Federal do Mato Grosso; UPFE: Universidade Federal de Pernambuco; UFPel: Universidade Federal de Pelotas; UFRGS: Universidade Federal do Rio Grande do Sul; UFRJ: Universidade Federal do Rio de Janeiro; UFSC: Universidade Federal de Santa Catarina; UnB: Universidade de Brasília; UNESP: Universidade Estadual Paulista; UNICAMP: Universidade Estadual de Campinas; UNIFESP: Universidade Federal de São Paulo; UNIFOR: Universidade de Fortaleza; UNISANTOS: Universidade Católica de Santos; UNISINOS: Universidade do Vale do Rio dos Sinos; USP: Universidade de São Paulo.

Fonte: dados provenientes da avaliação trienal (2010-2012) da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

Nota: docentes com atuação em mais de um programa foram contabilizados em cada programa aos quais são vinculados.

tintas de organização dos respectivos grupos. Como já foi observado, a coautoria potencializa para cada pesquisador o número de artigos publicados⁸.

Além das diferenças epistemológicas, teóricas e metodológicas, o formato de apresentação dos artigos entre as subáreas também se diferencia. Nas CSHS, é fundamental explicitar o quadro teórico e descrever o contexto social e cultural

em que estão inseridos os sujeitos da pesquisa, mostrando na análise a sua complexidade. A limitação de conteúdo, dada pela restrição do número máximo de palavras por artigo pelas revistas científicas, pode, em alguns casos, ser nociva à compreensão de determinados estudos qualitativos. Para efeito de exemplo, a renomada revista *Social Science and Medicine* aceita artigos com até 8 mil palavras, enquanto muitas revistas

Tabela 1

Indicadores de produção científica segundo subárea da Saúde Coletiva no triênio 2010-2012.

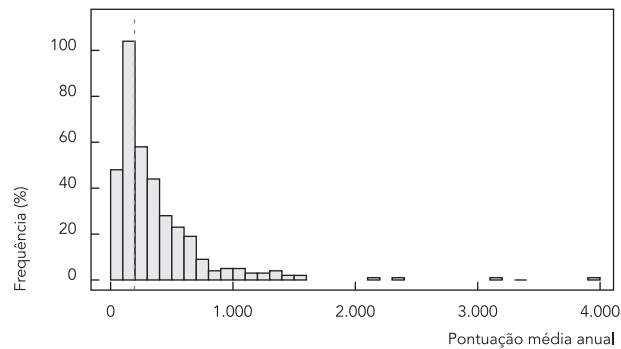
Indicador de produção	Epidemiologia	Políticas, Planejamento e Gestão em Saúde	Ciências Sociais e Humanas em Saúde	Outras
Total de docentes	365	146	126	111
Pontuação média anual				
Mediana	245,0	165,9	155,0	178,3
Percentil 25	145,0	103,7	72,1	90,0
Percentil 75	465,0	248,3	263,3	295,9
Proporção de docentes com pontuação igual ou acima				
Mediana da área (195 pontos)	61,1%	41,8%	38,9%	40,5%
Percentil 80 da área (405,3 pontos)	30,1%	7,5%	10,3%	14,4%

Fonte: dados provenientes da avaliação trienal (2010-2012) da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

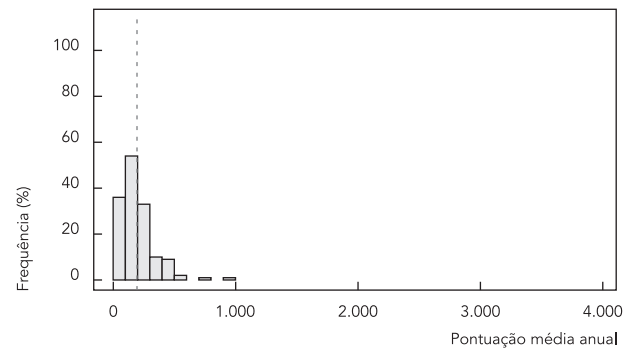
Figura 2

Distribuição da pontuação média anual de docentes permanentes segundo subárea da Saúde Coletiva no triênio 2010-2012.

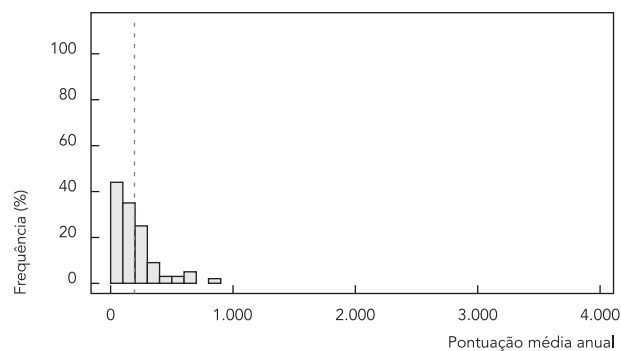
2a) Epidemiologia



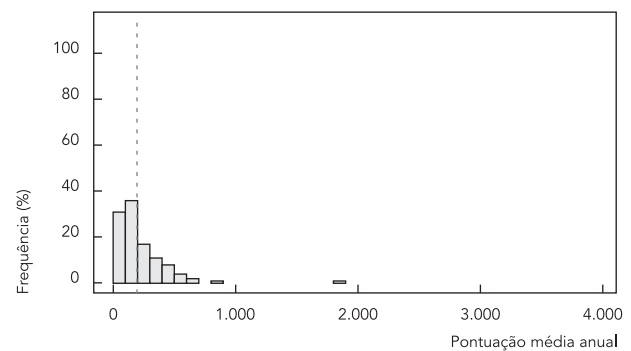
2b) Políticas, Planejamento e Gestão em Saúde



2c) Ciências Sociais e Humanas em Saúde



2d) Outras



Fonte: dados provenientes da avaliação trienal (2010-2012) da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

Qualis A na área da Saúde estabelecem limite de 3.500. Dessa forma, nem todas as revistas de Saúde Pública se mostram receptivas à natureza dos estudos de CSHS.

A produção científica na Epidemiologia tem se caracterizado pela inserção em grandes redes de pesquisa, com divisão da produção em artigos que focam em resultados específicos de um projeto de pesquisa. O desenho dos estudos epidemiológicos permite que os resultados sejam separados e discutidos individualmente em vários artigos. Críticas sobre as consequências do *salami slicing* (prática de “fatiar” um corpo de dados em peças menores para publicação em casos nos quais um artigo completo seria mais adequado) já são feitas há muito⁹. A pressão pela quantidade de publicações tem ônus graves para todas as áreas: baixa inovação, baixa criatividade, reprodução de mesmos resultados em revistas diferentes, superficialidade de análise e mesmo aumento de fraudes^{10,11,12,13,14}.

Adicionalmente, o processo de qualificação de periódicos adotado pela Capes – o Qualis – também contribui para as distorções. Resumidamente, embora existam critérios diversos para a classificação de revistas, todos estão baseados em indicadores de citação. Sem repetir aqui todas as críticas já feitas a estes indicadores, o que seria exaustivo, chamamos a atenção para alguns aspectos relevantes para essa discussão.

Em primeiro lugar, há distinções inerentes aos padrões de citação em diferentes áreas do conhecimento, o que se traduz em divergências nas faixas de valores alcançados por periódicos de diferentes disciplinas¹². No caso dos periódicos nos quais os docentes dos programas brasileiros publicam, revistas biomédicas têm, tipicamente, indicadores mais altos que os demais, sendo seguidas pelas revistas de Epidemiologia; já as revistas de ciências sociais e humanas tendem a ter indicadores mais baixos. Como a classificação nos diferentes estratos é feita pelo ranqueamento dos indicadores, isso significa que as revistas nas quais docentes da área de Epidemiologia usualmente publicam são qualificadas em estratos mais elevados do que aquelas nas quais os docentes das outras duas áreas publicam habitualmente.

Em segundo lugar, a estrutura das diferentes faixas do Qualis é restrita, a priori, por critérios definidos pela Diretoria de Avaliação da Capes. Por exemplo, são estabelecidos percentuais máximos de inclusão de periódicos em determinados níveis, o que limita a autonomia das áreas na classificação de suas revistas nas faixas mais elevadas.

Em seu conjunto, tais peculiaridades do sistema de classificação de periódicos acentuam

ainda mais a desigualdade entre as subáreas. Esses pontos merecem reflexão, dada sua sinergia. As características intrínsecas de uma determinada cultura científica acabam por representar uma posição desfavorável nesse mercado, cujas regras de acumulação de capital acadêmico são influenciadas e mantidas, em sua lógica de reprodução, pelas desiguais ofertas de publicação nos meios valorizados segundo critérios previamente definidos. Tal quadro remete diretamente ao que Bourdieu define como relações de dominação entre os agentes de um determinado campo científico, por meio de vários dispositivos, estratégias e regras, que se impõem como orientadoras do campo e servem para a conservação de sua estrutura de poder¹⁵.

Esses critérios de quantificação da produtividade geram ainda o efeito de modelo, influenciando diversas agências de fomento, que acabam por adotar os mesmos padrões. Na prática, um pesquisador com boa quantificação terá maiores oportunidades para inserção em PPG, obtenção de bolsas e recursos. Mais uma vez, os critérios de acumulação do capital acadêmico se evidenciam, configurando o que Robert Merton, o sociólogo da ciência, denominou de “efeito Mateus”¹⁶, que leva à concentração de prestígio e recursos em cada vez menos pesquisadores e instituições de alto reconhecimento.

A desigualdade na avaliação quantitativa da produção científica observada na avaliação dos PPG pela Capes repercute no desenvolvimento do campo da Saúde Coletiva, em face do seu poder indutor. As exigências cada vez maiores de produção quantitativa levam os coordenadores de pós-graduação a aumentarem os escores de pontos de produtividade para credenciamento de docentes, tornando cada vez mais difícil o credenciamento de docentes permanentes das áreas que menos pontuam. A predominância de docentes da subárea Epidemiologia no total dos programas pode se dever a esse processo de seleção interna. Caso não se coloque um limite na corrida pela pontuação, é possível que, a cada triênio, continuemos a observar o aumento da mediana da produção, tendo como resultado o crescimento da desigualdade na distribuição dos docentes entre as subáreas da Saúde Coletiva. Teme-se que esse desequilíbrio venha a comprometer a diversidade interna da área, levando à progressiva diminuição da participação percentual de docentes das subáreas de CSHS e PPGS. Esta seria uma consequência nefasta do atual modelo de avaliação, que comprometeria a riqueza constitutiva da área.

É importante frisar, no entanto, que, ao abordarmos as diferenças na produção científica entre as subáreas, nossa intenção não é, de modo

Tabela 2

Associação entre a composição do programa de pós-graduação (PPG) e a obtenção do conceito Muito Bom no indicador distribuição de publicações qualificadas em relação ao corpo docente permanente do PPG, segundo composição do PPG. Triênio 2010-2012.

Classificação segundo composição do PPG	Total de PPG	Conceito Muito Bom (%)	OR	IC95%
Epidemiologia > 60% e CSHS ou PPGS < 10%				
Sim	11	54,5	4,4	0,99-19,56
Não	28	21,4		
Epidemiologia > 60% e CSHS e PPGS < 10%				
Sim	6	66,7	6,3	0,96-40,80
Não	33	24,2		

CSHS: Ciências Sociais e Humanas em Saúde; IC95%: intervalo de 95% de confiança; OR: *odds ratio*; PPGS: Políticas, Planejamento e Gestão em Saúde.

Nota: Conceito Muito Bom – ter pelo menos 60% dos seus docentes permanentes alcançando pontuação anual média igual ou acima da mediana da área.

algun, desconsiderar as dificuldades e a complexidade da pesquisa epidemiológica, nem deixar de reconhecer o mérito da publicação de artigos epidemiológicos em revistas conceituadas. As exigências crescentes de aumento da produção científica geram consequências indesejadas para todas as subáreas, mas as atingem de forma diferenciada e desigual, e os atuais critérios de avaliação não contemplam essas distinções. Este é o ponto que queremos ressaltar.

A desigualdade produzida pelos critérios aplicados repercute também na avaliação dos programas, pois os docentes não estão homogeneamente distribuídos pelos PPG. Há, como demonstramos, uma grande variedade de arranjos entre as subáreas nos diversos programas, observando-se ora predominância quase absoluta da Epidemiologia, ora a presença minoritária desta, com vários graus intermediários de composição parcial. Os programas majoritariamente formados por epidemiologistas, especialmente se também apresentarem baixa participação de docentes das subáreas CSHS e PPGS, têm nítida vantagem na pontuação quantitativa quando comparados aos programas mais abrangentes, os quais apresentam representação relevante de docentes de várias subáreas da Saúde Coletiva.

A exigência de produtividade científica não é exclusividade da ciência brasileira e vem sendo questionada e discutida em várias áreas do conhecimento, em diversos países. Shore¹¹ descreve a sociedade de auditoria como aquela em que as pessoas são interpeladas como auditadas, a prestação de contas é mesclada com elaborados mecanismos de controle para submeter o desempenho individual ao olhar de especialistas externos e na qual todo aspecto de trabalho deve ser classificado e avaliado em relação a metas econômicas e pontos de referência burocráticos. O autor chama a atenção sobre como, gradativa-

mente, ela molda nossas vidas, nossas relações, nossas identidades profissionais e a maneira como nos conduzimos na vida acadêmica. Citando os acadêmicos britânicos, Shore comenta que, cada vez mais, eles projetam suas pesquisas pensando na burocracia da avaliação de qualidade e em criar evidência mensurável de desempenho que produzirá contagens positivas no próximo exercício de avaliação. O argumento do autor é que esses novos sistemas de auditoria não são, como alegam, apenas práticas neutras, destinadas a promover “transparência” ou eficiência. Ao contrário, são tecnologias disciplinares – ou “técnicas do *self*” – com o objetivo de conformar novas normas de conduta na força de trabalho.

Como limites do estudo, não podemos deixar de citar a dificuldade em operar as classificações taxonômicas segundo subárea, pois os pesquisadores não se encontram necessariamente em uma área específica; muitos circulam entre as áreas, outros trabalham interdisciplinarmente. É importante ressaltar, ainda, a diversidade existente nas formas de produção científica no interior das subáreas. Por fim, a variação no tamanho dos programas, sendo a maioria constituída por menos de 20 docentes, prejudica as comparações entre eles. Todavia, esse exercício, mesmo com as limitações apontadas anteriormente, permitiu evidenciar empiricamente diferenças entre subáreas, que precisam ser consideradas no processo de avaliação de programas e pesquisadores.

Considerações finais e propostas para a avaliação

É necessário refletir como estamos conduzindo a produção científica em nosso campo. Estaríamos priorizando a criatividade, inovação, impacto social, ou meramente visando a uma maior adequação às regras de produtivismo vigentes e incorporados no modelo de avaliação? Posto em melhores termos, como pensar maneiras para que a avaliação dos programas, que representa um processo fundamental para o desenvolvimento e legitimidade da área, não se torne um exercício de produção de desigualdade com mecanismos revestidos de uma suposta neutra aferição?

Tal tarefa é tão urgente quanto legítima. A própria comissão de avaliação demonstrou, no seu relatório, preocupação com o processo corrente, afirmando que *“o atual modelo de avaliação está esgotado tendo completado um ciclo muito positivo para o desenvolvimento do SNPG e da própria ciência brasileira. É necessário, portanto, um novo esforço de formulação que contemple as mudanças e facilite o processo de acompanhamento dos cursos”*¹⁷, concluindo com a seguinte proposta: *“A avaliação deve ser predominantemente realizada em bases qualitativas ainda que indicadores quantitativos possam ser utilizados como balizadores para as decisões da comissão. Os cursos devem ser analisados em sua integridade e não como a resultante de um conjunto fragmentado de aspectos”*¹⁷.

Embora, a rigor, mudanças verdadeiramente revolucionárias fossem necessárias para modificar substancialmente a situação vigente, sua baixa exequibilidade política, ao menos a curto prazo, deve ser reconhecida. Mudanças incrementais em alguns aspectos da sistemática de avaliação poderiam garantir critérios menos indutores de iniquidades, sem alterar substancialmente o processo num primeiro momento. Basicamente, é necessário que se garanta que comparações sejam feitas, de fato, entre similares, respeitando as heterogeneidades. Isso se aplica tanto à produção de diferentes programas, quanto à classificação de periódicos.

Iniciando pela segunda comparação, é necessária uma revisão do Qualis que garanta ao menos a valoração adequada de revistas dentro do seu escopo disciplinar. Desse modo, em vez de simplesmente ranquear todas as revistas com publicações de docentes da área, para, a partir daí, determinar os diferentes estratos, seria desejável que o ranqueamento fosse feito, ao menos, conforme grandes áreas disciplinares, garantindo que os periódicos com melhores indicadores de cada área estivessem representados nos es-

tratos superiores. Concretamente, isso significaria que, por exemplo, uma revista de ciências sociais poderia ser classificada como A1, mesmo tendo indicadores menores que o de revistas biomédicas.

Quanto à comparação entre programas, seria necessário considerar, de modo diferenciado, programas disciplinares, especialmente os dedicados majoritariamente à Epidemiologia, e os abrangentes. Isso requer que haja uma definição mínima do que é um “programa abrangente”, não sendo suficiente a mera presença de um ou outro docente das subáreas PPGS e CSHS para caracterizar essa abrangência. Para essa definição, seria importante desenvolver critérios taxonômicos qualitativos, análogos aos que utilizamos para classificar as áreas de atuação de docentes individuais.

Seria desejável, por fim, conter a corrida desenfreada pela publicação. Uma alternativa seria considerar patamares máximos, e não só mínimos, para os indicadores de produção – considere-se que isso já é feito com relação ao número de orientandos que um docente pode ter no período de avaliação. O estímulo à publicação desenfreada está se tornando danoso para a própria produção científica; portanto, talvez seja o caso de se estabelecer alguma pactuação sobre o que seria, afinal, uma produção razoável. Como afirmam Wager et al.¹⁸, *“ao medir a produtividade, instituições e financiadores deveriam estar alertas não apenas para pesquisadores improdutivos, mas também para aqueles que são inviavelmente prolíficos”*.

Concluindo, é forçoso reconhecer que o sonho de um sistema que privilegie o desenvolvimento solidário da área, avaliado qualitativamente, e não um ranqueamento competitivo baseado em critérios quantitativos duvidosos, ainda está distante. Mas é possível, dentro das restrições vigentes, praticar uma política de “redução de danos”, que, ao menos, não piore a desigualdade e a fragmentação da área, impedindo a exclusão de partes significativas do campo que chamamos de Saúde Coletiva.

Resumen

El objetivo de este trabajo fue discutir los límites del modelo de evaluación cuantitativa de la producción científica en la Salud Colectiva. Se realizó un análisis de la producción científica de los docentes en subáreas de la Salud Colectiva durante el trienio 2010-2012. Se compararon las distribuciones de la puntuación anual media de los docentes según subáreas. Se estimó la probabilidad de un 60% de los docentes del programa de posgrado (PPG) que tuvieran una puntuación igual o superior al P50 del área (Muy Buena). La subárea Epidemiología presentó una media de puntuación anual de los docentes significativamente mayor. PPGs cuya composición incluye un 60% o más de docentes del subárea Epidemiología, y menos de un 10% del subárea Ciencias Sociales y Humanas en Salud, presentaron una probabilidad significativamente mayor de alcanzar el nivel Muy Bueno. Las desigualdades constatadas en las medidas de producción científica entre las subáreas de la Salud Colectiva apuntan la necesidad de reexaminar la evaluación, de manera que no se reproduzcan inequidades que generen consecuencias nocivas para la diversidad del campo.

Evaluación de Programas e Instrumentos de Investigación; Sistemas de Créditos e Avaliação de Pesquisadores; Indicadores de Produção Científica

Colaboradores

J. A. B. Iriart, S. F. Deslandes e D. Martin contribuíram na concepção do estudo, coleta de dados, análise dos resultados, discussão e redação do artigo. K. R. Camargo Jr., M. S. Carvalho e C. M. Coeli contribuíram na análise estatística, análise dos resultados, discussão e redação do artigo.

Agradecimentos

Agradecemos aos coordenadores dos programas de pós-graduação por sua contribuição no preenchimento das planilhas de classificação dos docentes.

Referências

1. Paim JS, Almeida Filho N. A crise da saúde pública e a utopia da saúde coletiva. Salvador: Casa da Saúde; 2000.
2. Birman J. Apresentação: a interdisciplinaridade na Saúde Coletiva. *Physis* (Rio J.) 1996; 6:7-13.
3. Nunes ED, Ferreto LE, Oliveira ALO, Nascimento JL, Barros NF, Castellanos MEP. O campo da Saúde Coletiva na perspectiva das disciplinas. *Ciênc Saúde Coletiva* 2010; 15:1917-22.
4. Bourdieu P. O campo científico. In: Ortiz R, organizador. *Pierre Bourdieu: sociologia*. São Paulo: Ática; 1983. p. 122-55.
5. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Documento de área 2013. http://www.capes.gov.br/images/stories/download/avaliacaotrienal/Docs_de_area/Saude_Coletiva_doc_area_e_comiss%C3%A3o_att08deoutubro.pdf.

6. Franceschet M, Costantini A. The first Italian research assessment exercise: a bibliometric perspective. *J Informetr* 2011; 5:275-91.
7. Moed H. UK research assessment exercises: informed judgments on research quality or quantity? *Scientometrics* 2008; 74:153-61.
8. Camargo Jr. KR, Coeli CM, Caetano R, Maia VR. Produção intelectual em saúde coletiva: epistemologia e evidências de diferentes tradições. *Rev Saúde Pública* 2010; 44:394-8.
9. Broad WJ. The publishing game: getting more for less. *Science* 1981; 211:1137-9.
10. Fischer J, Ritchie EG, Hanspach J. Academia's obsession with quantity. *Trends Ecol Evol* 2012; 27:473-4.
11. Shore C. Cultura de auditoria e governança liberal: universidades e a política da responsabilização. *Mediações – Revista de Ciências Sociais* 2009; 14:24-53.
12. Diaz-Isenrath C. Notas a propósito de los mecanismos de acreditación del conocimiento y de la reconfiguración de lo social. *Mediações – Revista de Ciências Sociais* 2009; 14:54-68.
13. Adler R, Ewing J, Taylor P. Citation statistics: a report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS). Joint Committee on Quantitative Assessment of Research. Berlin: International Mathematical Union (IMU); 2009.
14. Gastaldo D, Bosi MLM. ¿Qué significa tener impacto? Los efectos de las políticas de productividad científica en el área de salud. *Enferm Clín* 2010; 20:145-6.
15. Bourdieu P. Os usos sociais da ciência. Por uma sociologia do campo científico. São Paulo: Editora Unesp; 2004.
16. Merton RK. The Matthew effect in science. *Science* 1968; 159:56-63.
17. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Relatório de avaliação 2010-2012. Trienal 2013. <https://docs.google.com/viewer?a=v&pid=sites&srcid=Y2FwZXMuZ292LmJyfHRyaWVuYWwtMjAxM3xneDo3ZWRIImUwNzBhOGJlMWNl>.
18. Wager E, Singhvi S, Kleinert, S. Too much of a good thing? A study of prolific authors. In: International Congress on Peer Review and Biomedical Publication. http://www.peerreviewcongress.org/abstracts_2013.html.

Recebido em 23/Abr/2015

Versão final reapresentada em 04/Ago/2015

Aprovado em 09/Set/2015