

Trends in the investigation of social determinants of health: selected themes and methods

Tendências na investigação dos determinantes sociais da saúde: métodos e temas selecionados

Roger Keller Celeste ¹
 João Luiz Bastos ²
 Eduardo Faerstein ³

Abstract

We analyze bibliometric trends of topics relevant to the epidemiologic research of social determinants of health. A search of the PubMed database, covering the period 1985-2007, was performed for the topics: socioeconomic factors, sex, race/ethnicity, discrimination/prejudice, social capital/support, lifecourse, income inequality, stress, behavioral research, contextual effects, residential segregation, multilevel modeling, regression based indices to measure inequalities, and structural equation modeling/causal diagrams/path analysis. The absolute, but not the relative, frequency of publications increased for all themes. Total publications in PubMed increased 2.3 times, while the subsets of epidemiology/public health and social epidemiologic themes/methods increased by factors of 5.3 and 5.2, respectively. Only multilevel and contextual analyses had a growth over and above that observed for epidemiology/public health. We conclude that there is clearly room for wider use of established techniques, and for new methods to emerge when they satisfy theoretical needs.

Food Consumption; Diet Surveys; Epidemiologic Methods

Introduction

A nearly exponential growth in the scientific output on social epidemiology has been documented ¹. However, since Solla-Price's seminal work on the exponential growth of science ², it has been shown that scientific production doubles every ten to fifteen years ³. Hence, it remains unexplored whether the absolute growth described for general science would also be observed in relative terms for specific areas of knowledge. By examining scientific growth in specific areas, it is possible to highlight emerging themes, so as to indicate possible advances in the near future. The objective of the present study was to assess trends in scientific production of methods and themes in the investigation of social determinants of health, between 1985 and 2007.

Methods

This study consists of a bibliometric analysis of yearly trends in groups of publications indexed in PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) from 1985 to 2007. The main database in PubMed is MEDLINE, but the proportion of non-MEDLINE citations is unknown and a strict search of MEDLINE is not possible. In addition to MEDLINE, PubMed covers in-process citations, some papers before 1950 (OLDMEDLINE), and

¹ Faculdade de Odontologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.

² Centro de Ciências da Saúde, Universidade Federal de Santa Catarina, Florianópolis, Brasil.

³ Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil.

Correspondence

R. K. Celeste
 Departamento de Odontologia Preventiva e Social, Faculdade de Odontologia, Universidade Federal do Rio Grande do Sul, Rua Ramiro Barcellos 2492, 3ª andar, Porto Alegre, RS 90035-003, Brasil.
roger.keller@ufrgs.br

citations that are out-of-scope from MEDLINE journals.

Each journal article added to PubMed's collection is indexed by specialized staff under at least one MeSH (Medical Subject Headings) descriptor of the U.S. National Library of Medicine Thesaurus, which was established in 1960 and has been updated since. The MeSH Thesaurus contains a tree-like cross-referenced structure of controlled vocabulary, used to translate terminology employed in articles in different idioms to a "system language". Each "entry term" encompasses many synonyms, near-synonyms and related concepts, regardless of the idiom, wording and spelling used by authors. When a MeSH term is used, PubMed automatically searches on narrower descriptors further out on that branch of the tree. For instance, the use of the MeSH term "Socioeconomic Factors" automatically searches for "Education Status", "Income", "Occupation", "Inequality" and "Social Class".

Principles applied to build search strategies

First, whenever possible, a MeSH term was employed rather than non-MeSH "keywords" in all search strategies. If an appropriate MeSH term could not be identified, a search strategy was built with text words based on the authors' experience. Second, we favored the earliest MeSH terms – those added before 1985 – and the most general category of the MeSH tree hierarchy. Third, MeSH term definitions were scrutinized in order to choose only those terms with the desired meaning. Finally, based on a preliminary analysis, we excluded redundant terms in the search strategy, so as to keep it as parsimonious as possible. Our searches were conducted on May 17, 2008.

Search strategies and data analysis

We tallied the annual number of publications identified using the following search strategies (details of search strategy can be obtained from the authors): all publications in the PubMed database (search strategy number #1); publications in epidemiology/public health (#2); and selected themes in social epidemiology (#14). We further refined the bibliographic analysis of (#14) by selecting publications focusing on 11 themes: socioeconomic factors (#3), sex (#4), race/ethnicity (#5), prejudice/discrimination (#6), social capital/support (#7), lifecourse (#8), income inequality (#9), stress (#10), behavioral research (#11), contextual factors (#12), and residential segregation (#13). The number of publications identified in the epidemiology/public health search (#2)

served as the denominator to calculate the proportion represented by each of the 11 social epidemiology themes (i.e. strategies from #3 to #13 were divided by #2). The total number of publications in PubMed (#1) was used as the denominator to calculate the proportion of publications in epidemiology/public health (#2 divided by #1) and in social epidemiology (#14 divided by #1).

We also examined trends for three types of data analysis methods: multilevel modeling (#15), regression-based indices to measure inequalities (#16), and structural equation modeling/causal diagrams/path analysis (#17). The absolute number of publications of these data analysis techniques among the eleven social epidemiology themes were determined for each year between 1985 and 2007 and plotted. Their proportion in relation to the total number of articles in the 11 themes (i.e. the number of publications from strategies #15 to #17 divided by the number of publications from strategy #14) was also determined and plotted for the period 1985-2007.

Results

Between 1985 and 2007, there was a 2.3-fold increase in the annual number of citations added to the PubMed: from 329,263 in 1985 to 759,698 in 2007. In the same period, articles indexed under epidemiology/public health headings (search strategy #2) increased 5.3 times (from 48,719 in 1985 to 256,892 in 2007) and, among the selected themes in social epidemiology (search strategy #14), there was a 5.2-fold increase (from 9,349 to 49,052).

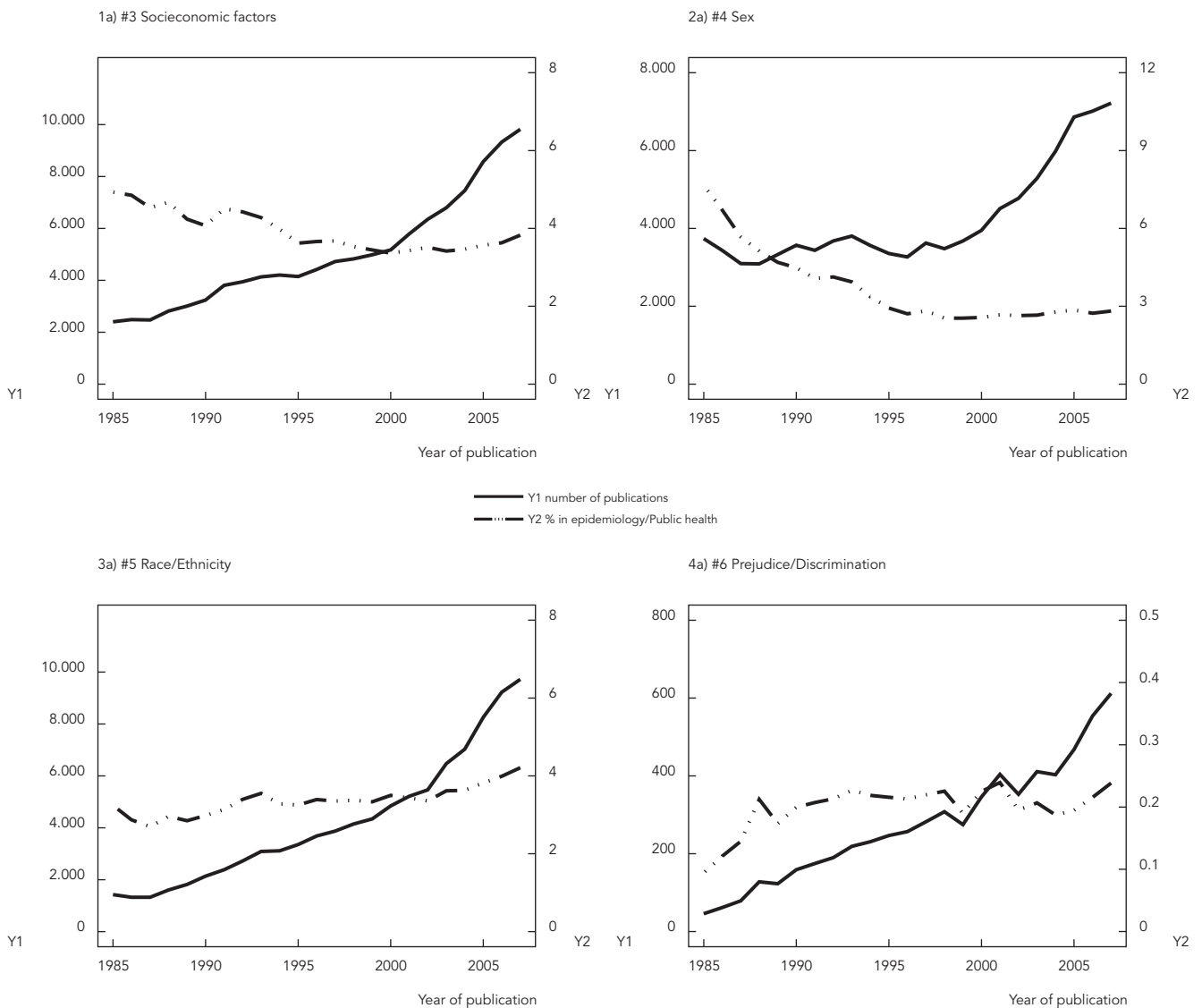
In 2007, more than 30% of the scientific output indexed in this bibliographic database had at least one of the descriptors used to identify citations in the area of epidemiology/public health. In contrast, the relative contribution of social epidemiology increased moderately, compared to epidemiology/public health, reaching 6.5% in 2007.

Trends in selected themes in social epidemiology are depicted in Figures 1 and 2. Absolute frequencies tended to increase over the studied period, and more markedly in recent years. Themes like socioeconomic factors, sex, race/ethnicity, behavioral research, stress, contextual factors, lifecourse and prejudice/discrimination are good examples in this regard. Due to small numbers, however, we could not be confident about the trend patterns for income inequality (#9) and residential segregation (#13).

A different picture emerges when relative frequencies are considered. When the count of all epidemiology/public health publications serves

Figure 1

Trends in absolute/relative number of publications in socioeconomic factors, sex, race/ethnicity; prejudice/discrimination; social capital/support, life course, income inequality, and stress research in PubMed from 1985 to 2007.



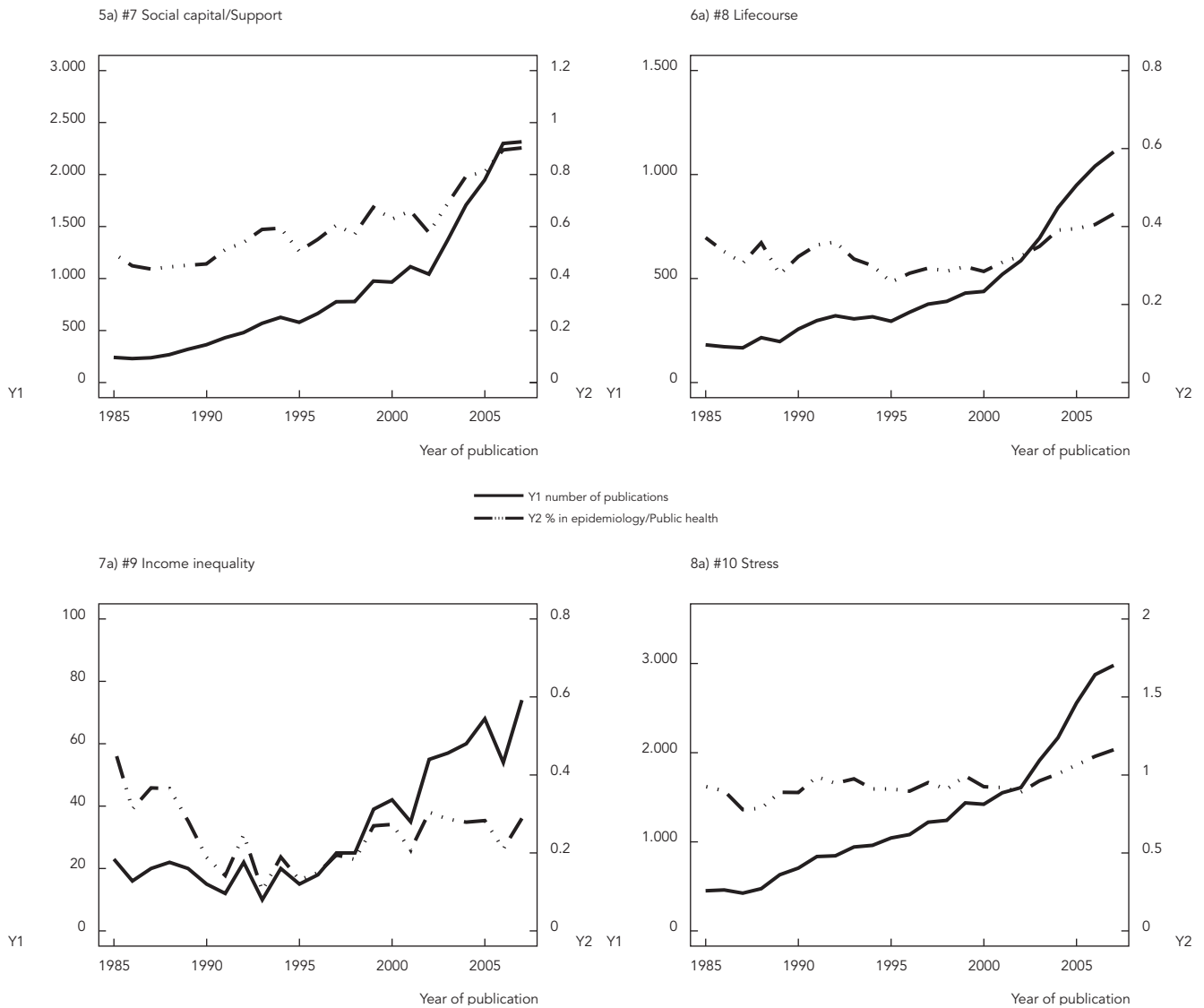
(continues)

as the denominator, the exponential growth pattern seen with the absolute count is no longer observed. Increases or stationary trends are seen for behavioral research, race/ethnicity, stress, contextual factors, lifecourse, prejudice/discrimination and social capital/support. Among these, contextual factors showed the steepest relative increase in the period, rising from almost 4% in 1985 to 7% of all epidemiology/public health publications in 2007. In contrast, socioeconomic fac-

tors and gender showed declining relative trends. Residential segregation and income inequality each exhibited a fluctuating pattern, which likely reflects the small number of publications.

Publications using one of the three analysis methods – multilevel modeling, structural equation modeling/causal diagrams/path analysis, and regression-based indices – all showed steep increases in the absolute number of articles, as well as in their relative frequencies (Figure 2).

Figure 1 (continued)



Among these, multilevel modeling emerged as the most employed method.

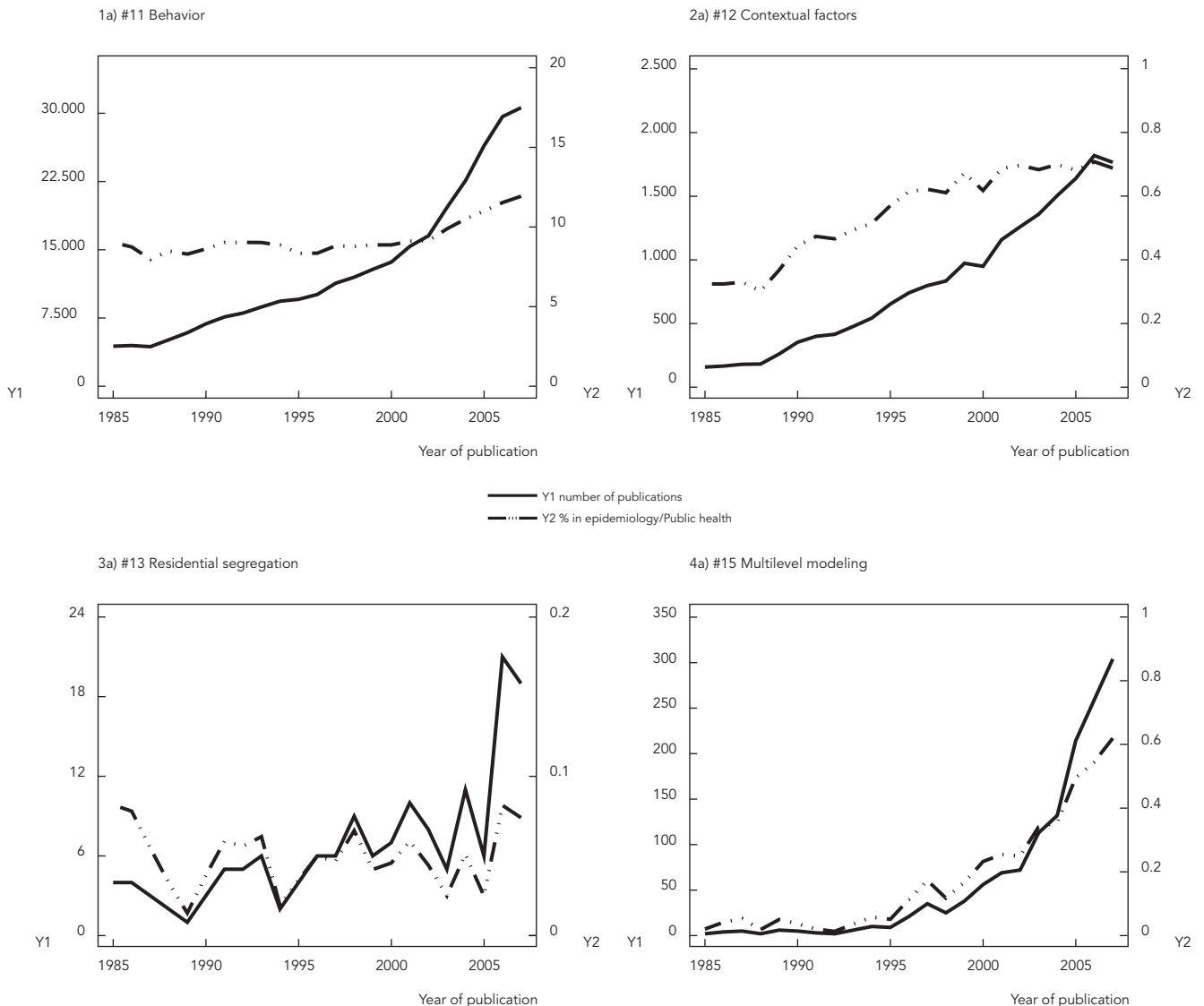
Discussion

Our results showed that absolute and relative trends can provide different conclusions, but regardless of how we plotted them, epidemiology and social epidemiology grew over and above the growth of general health science. One could

argue that the larger growth of social epidemiologic themes in relation to the total citations in PubMed could be determined in part by an increased indexation of epidemiologic/public health journals in the database over the study period. However, this is not likely to be the case; epidemiologic/public health journals account for only 1% (n = 370) of the 37,665 journals indexed in PubMed and 2.2% of all publications (data from PubMed Journal Database), such that an expressive number of journals would have to

Figure 2

Trends in absolute/relative number of publications in behavior research, contextual factors, residential segregation, multilevel modeling, regression based indices to measure inequalities, and structural equation in PubMed from 1985 to 2007.



(continues)

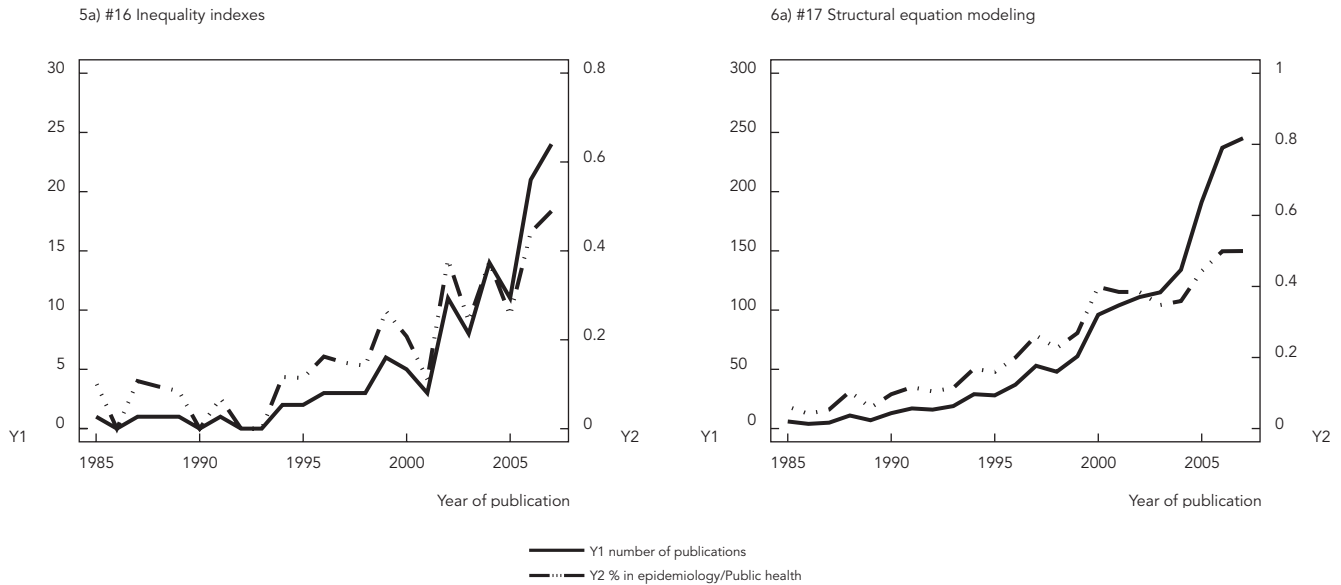
be included in the database in order to artificially influence the observed trends.

One of the limitations of the present study is that its findings are based only on PubMed; however, this is recognized as the largest and best-known database in the field of health sciences. Another limitation is that not all retrieved publications necessarily strictly fit the definition of social epidemiology, that is, explicitly incorporating so-

cial theory in the article's analytical framework⁴. This concern is tempered by our use of MeSH terms, which increased the sensitivity and the specificity of the search strategies. Because the terms used were constant over time, the results are likely to reflect real trends in social epidemiology publications indexed in PubMed.

Overall, from the results presented, it can be concluded that the branch of social determinants

Figure 2 (continued)



of health has been growing fast, and that this growth was seen in nearly all of the 11 sub-areas. It is important to emphasize that the magnitude of absolute increases in some sub-areas might be misinterpreted as outpacing others, when, in fact, relative figures reveal increases that were actually modest. Although the number of publications in social epidemiology increased more than the average growth of the total publications in PubMed,

epidemiology/public health did too, and the only themes in social epidemiology growing over and above trends in epidemiology/public health were those which lent themselves to multilevel or contextual analysis. This is a good example where methodological advances met theoretical needs. However, there is clearly room for wider use of established techniques, and for new methods to emerge and satisfy theoretical needs.

Resumo

Analisam-se tendências bibliométricas de tópicos relevantes na pesquisa epidemiológica de determinantes sociais da saúde. Realizaram-se buscas no PubMed no período 1985-2007, sobre: fatores socioeconômicos, sexo, raça/etnia, discriminação/preconceito, capital/suporte social, curso de vida, desigualdade de renda, estresse, pesquisa comportamental, efeitos contextuais, segregação residencial, modelo multinível, indicadores de desigualdade baseados em regressão e equações estruturais/análise de caminhos/diagramas causais. A frequência absoluta, e não a relativa, de publicações cresceu em todos os temas. O total de publicações aumentou 2,3 vezes no período, enquanto que o conjunto epidemiologia/saúde pública e os temas de epidemiologia social aumentaram 5,3 e 5,2 vezes, respectivamente. Efeitos contextuais e modelagem multinível apresentaram crescimento relativo acima do observado para epidemiologia/saúde pública. Conclui-se que existe espaço para a ampliação do uso das técnicas de análise existentes e para que novos métodos surjam, atendendo a necessidades teóricas específicas da área.

Consumo de Alimentos; Inquéritos sobre Dietas; Métodos Epidemiológicos

Contributors

R. K. Celeste collaborated in the design, write-up, data collection, analysis and interpretation and final revision. J. L. Bastos collaborated in the design, write-up, analysis and interpretation of data, and revision of the final version. E. Faerstein collaborated in the design and data interpretation, as well as in the write up and revision of the final version.

References

1. Kaplan GA. What's wrong with social epidemiology, and how can we make it better? *Epidemiol Rev* 2004; 26:124-35.
2. Solla-Price DJ. *Science since Babylon: enlarged edition*. New Haven: Yale University Press; 1975.
3. Fernández-Cano A, Torralbo M, Vallejo M. Reconsidering Price's model of scientific growth: an overview. *Scientometrics* 2004; 61:301-21.
4. Kaufman JS. Social epidemiology. In: Rothman KJ, Greenland S, Lash TL, editors. *Modern epidemiology*. Philadelphia: Lippincott Williams & Wilkins; 2008. p. 532-48.

Submitted on 31/Dec/2009

Final version resubmitted on 30/Jun/2010

Approved on 11/Aug/2010