Trends in Citations to Books on Epidemiological and Statistical Methods in the Biomedical Literature

Citation for published version:
Porta, M, Vandenbroucke, JP, Ioannidis, JPA, Sanz, S, Fernandez, E, Bhopal, R, Morabia, A, Victora, C & Lopez, T 2013, 'Trends in Citations to Books on Epidemiological and Statistical Methods in the Biomedical Literature' PLoS One, vol 8, no. 5, e61837. DOI: 10.1371/journal.pone.0061837

Digital Object Identifier (DOI):
10.1371/journal.pone.0061837

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
PLoS One

Publisher Rights Statement:
This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Trends in Citations to Books on Epidemiological and Statistical Methods in the Biomedical Literature

Miquel Porta1,2*, Jan P. Vandenbroucke3, John P. A. Ioannidis4,5, Sergio Sanz1, Esteve Fernandez6, Raj Bhopal7, Alfredo Morabia8,9, Cesar Victora10, Tomàs Lopez1

1 Hospital del Mar Institute of Medical Research (IMIM), Barcelona, CIBER de Epidemiología y Salud Pública (CIBERESP), and School of Medicine, Universitat Autònoma de Barcelona, Catalonia, Spain, 2 Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America, 3 Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands, 4 Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, 5 Stanford Prevention Research Center, Department of Medicine and Department of Health Research and Policy, Stanford University School of Medicine, and Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, California, United States of America, 6 Cancer Prevention and Control Programme, Institut Català d'Oncologia (ICO-IDIBELL), and School of Medicine, Universitat de Barcelona, L'Hospital de Lloretgat, Barcelona, Catalonia, Spain, 7 Centre for Population Health Sciences, University of Edinburgh, Edinburgh, Scotland, United Kingdom, 8 Mailman School of Public Health, Columbia University, New York, New York, United States of America, 9 School of Earth and Environmental Sciences and Center for the Biology of Natural Systems, Queens College, City University of New York, Flushing, New York, United States of America, 10 Universidade Federal de Pelotas, Pelotas, Brazil

Abstract

Background: There are no analyses of citations to books on epidemiological and statistical methods in the biomedical literature. Such analyses may shed light on how concepts and methods changed while biomedical research evolved. Our aim was to analyze the number and time trends of citations received from biomedical articles by books on epidemiological and statistical methods, and related disciplines.

Methods and Findings: The data source was the Web of Science. The study books were published between 1957 and 2010. The first year of publication of the citing articles was 1945. We identified 125 books that received at least 25 citations. Books first published in 1980–1989 had the highest total and median number of citations per year. Nine of the 10 most cited texts focused on statistical methods. Hosmer & Lemeshow's *Applied logistic regression* received the highest number of citations and highest average annual rate. It was followed by books by Fleiss, Armitage, et al., Rothman, et al., and Kalbfleisch and Prentice. Fifth in citations per year was Sackett, et al., *Evidence-based medicine*. The rise of multivariate methods, clinical epidemiology, or nutritional epidemiology was reflected in the citation trends. Educational textbooks, practice-oriented books, books on epidemiological substantive knowledge, and on theory and health policies were much less cited. None of the 25 top-cited books had the theoretical or sociopolitical scope of works by Cochrane, MeKean, Rose, or Morris.

Conclusions: Books were mainly cited to reference methods. Books first published in the 1980s continue to be most influential. Older books on theory and policies were rooted in societal and general medical concerns, while the most modern books are almost purely on methods.

Introduction

If one considers the book as the macro unit of thought and the periodical article the micro unit of thought, then…

Eugene Garfield (1955) [1]

Academic and professional books are both actors and witnesses of their corresponding disciplines. While some books move the frontiers of ignorance, others harvest and synthesize knowledge that seems established [2–5]. In principle, analyses of citations to books may shed light on how concepts and methods changed in the course of time whilst a discipline evolved as a field of practice and academic subject; such analyses are also relevant to explore the influence of a discipline on other fields [4–8].

Analyses of citations to scientific books are rare and, to our knowledge, there are no comprehensive analyses of citations to books of epidemiology and biostatistics, nor to sets of books in any other of the health and life sciences [6–10]. The vast majority of citation analyses involve citations to articles [11]. Studies on citations made by books (to other books or to articles) are also uncommon. In this study we will analyze citations to books made by scientific biomedical articles. Analyzing the bibliometric impact of books is also relevant at a time when the very nature of books...
and the whole publishing endeavour are experiencing enormous changes.

Analyses of citations to epidemiologic books may complement other analyses on the nature, evolution and endeavours of epidemiology, a particularly integrative science [12,15]; for decades, epidemiology has been useful to integrate knowledge, methods, reasoning and cultural referents from multiple health and social sciences, including medicine and public health. Hence, uses of epidemiologic and statistical reasoning, knowledge, and methods—foremost, uses in biomedical research—are of broad scientific interest. Indeed, today concepts and methods with strong epidemiological roots and properties seem fruitfully applied “within” and “outside” epidemiology [12–15].

As a consequence, there cannot be an exhaustive and fixed list of books on epidemiology and biostatistics. This does not preclude the analysis of an intellectually coherent set of books. Furthermore, since the present study is the first of its kind in epidemiology, a broad perspective is warranted. Therefore, rather than using narrow lists or definitions of what is a book on epidemiology or biostatistics, we will apply a wide, inclusive approach. Specifically, we will aim at including in the study books that have a clear biostatistic or epidemiologic component or dimension, relevant to epidemiology. As we shall see, while all books included are unequivocally on epidemiology or biostatistics, many focus exclusively on biostatistical methods and techniques, and a few others are fundamentally on public health, preventive medicine, clinical epidemiology or other epidemiologic specialties. The Appendix S1 includes a few texts (mostly, essays on health policy and philosophy of medicine) in which epidemiology has a secondary role, but which nevertheless are referents.

The objective of this study was to analyze the number and time trends of citations received from scientific biomedical articles by selected books of epidemiology, biostatistics and related disciplines, including public health and preventive medicine.

Methods

Selection of books

Based on academic lists of textbooks [3,4,16], books selected by The James Lind Library [17] and the People’s Epidemiology Library [18], publishers’ catalogues, books cited in other books, and our own teaching and research references, we first searched for citations to over 200 books on epidemiological and statistical methods and concepts. The books initially included were published from 1957 until 2010, and this is the main period of publication covered by the present study; nevertheless, we occasionally expanded the timeframe backwards to assess books published before 1957 that we deemed important texts for reference.

The primary aim was to include books that had a clear biostatistic or epidemiologic component or dimension, relevant to epidemiology and, foremost, to actual uses of epidemiologic methods in research and practice. We accepted books that focus exclusively on biostatistical methods and techniques, but did not consider books on bioinformatics or that focus on mathematical models without emphasising their potential biostatistical application. We also included books on public health, preventive medicine, clinical epidemiology and other epidemiologic specialties. We thus selected books that are pertinent to applied biostatistical and epidemiological methods and techniques; epidemiologic concepts and theory; health services and policies; and substantive epidemiology (e.g., nutritional epidemiology). Books were selected regardless of whether they were aimed at researchers, practitioners, postgraduate students or undergraduates.

If a book received less than 25 citations, it was finally not included in the study; examples of exclusions are given in the Table S1. Books included are listed in the References section in alphabetic order [19–304], except if already cited paragraphs [12,14]. If a book had many editions, only a selection is cited in the References, but we aimed at counting citations to all editions (see Appendix S1).

Bibliometric analyses

The data source for the study was the Web of Science, produced historically by the Institute for Scientific Information, Inc. (ISI), and in recent years by Thomson – Reuters [305]. The database includes: a) the Science Citation Index Expanded (SCI-Expanded), from 1945 to 2011; b) the Social Sciences Citation Index, from 1956 to 2011; and c) the Arts & Humanities Citation Index, from 1975 to 2011. Thus, while 1945 is the first year of publication of the citing articles, the year of publication of cited books is not restricted by 1945; since articles published from 1945 onwards could cite books published before 1945.

To measure the overall bibliometric impact of each book, we aimed at retrieving all citations received by all editions of the book from the year of publication until 31 December 2011. Thus, our primary measure was the total number of citations received by each book, unadjusted by the number of editions or time since first publication. A second indicator was the average number of citations received by the book per year since publication of the first edition. We also focused on books that received over 1000 citations since publication and books that received an average of more than 40 citations per year since publication. Thus, books published long ago had more time to accumulate citations. Books published in recent decades were a priori more likely to receive more than 40 citations per year, since the citing base of articles was larger. Using sampling of papers with random trigrams author searches, we estimated that in the Web of Science the number of papers (and by extrapolation, citations) indexed per year increased approximately 3-fold in the last 3 decades. Finally, we computed an annual citation rate, the number of citations received each year by the selected books at each corresponding time divided by the corresponding number of selected books. To allow for delays in the publication of citing articles, our last search was conducted on February 9, 2012 (always aiming at retrieving citations made by articles published until December 31, 2011). The present study was developed from a minor pilot study conducted in 2006 [16].

Through a ‘cited reference search’, we used combinations of ‘cited author’ and ‘cited work’, the latter referring to the title of a book or to the name of a journal, not to the title of an article. All citations were checked for accuracy as explained in Appendix S1. For each book several possible citation options were searched to allow for different abbreviations and for citation errors. Information on all books was verified against valid information on all editions of the book.

For a given book we included all directly related editions and printings. Citations to each book were searched independently by name of author or editor, and title; all authors or editors (not just the first one) were used in different searches. The data source does not allow to reliably identify chapters of books written by authors other than the book editors [306], unless the title of the book is unique or unambiguous (e.g., Modern Epidemiology [250–252], Oxford Textbook of Public Health [133–137]). Thus, if the title was unambiguous, the number of citations includes citations to individual chapters not authored by the book editors; if the title of the book was common, we could not include citations registered with the name of the first author of each chapter (when other than the editors), and some citations to chapters not written by the
editors had to be excluded (Appendix S1 and Figure S1). Citing articles were thus identified, and the year of publication of each article was analyzed (through the Web of Science option “analyze results”) to assess time trends in the number of citations to each book.

Results

General trends

The 125 books that received 25 citations or more were published between 1913 and 2004 and, therefore, the maximum number of years since publication was 98, and the minimum, 7 (median, 24 years); 100 books (80%) had been published for over 15 years (Table 1). Besides books published in the earlier periods, 46 books (37%) were published in 1980–1989, 35 (28%) in 1990–1999, and 14 (11%) in 2000–2004. We thus see a possible peak of influential books published around 1980–1989; most notably, works by Hosmer & Lemeshow, Rothman et al., Kalbfeisch & Prentice, Breslow & Day, Kleinbaum et al., Rosner, Sackett et al., Schlesselman & Stolley, Hennekens & Buring, Fletcher et al., Bland, Friedman et al., Hulley et al, Feinstein, or Kelsey et al. (Table 2).

All 10 top-cited books had an edition (not necessarily the first) during 1980–1989. This decade also stands out when we analyze the relative proportion of the study books that received over 1000 citations: such proportion peaked at near 15% in 1980–1989, as did the relative proportion of texts that received more than 40 citations per year since publication (Table 1).

The total number of citations received by the 125 books was 183,401, of which 52% were to books first published in 1980–1989. Books first published in this decade also had the highest mean and median number of citations per year since publication (Table 1). The crude number of citations to the selected books increased substantially from 1990 to the mid-1990s, and remained stable thereafter: there were 160 citations in 1970, 869 citations in 1980, 3980 citations in 1990, 7356 citations in 1997, and 8838 in 2008 (Figure 1); in such years the annual citation rate was 12.3, 26.3, 50.4, 68.7 and 70.7 citations/book, respectively.

Among the 50 top-cited books, 15 are completely or strongly on applied biostatistical methods, and the contents of at least a further three is heavily on quantitative methods (18 out of 50 = 36%). At least 33 of such 50 books (two-thirds) have strong methodological contents.

Nine of the top 10 books are general works on statistical or epidemiological methods: 7 of 10, including the top three, are strongly focused on statistics, while two address epidemiological methods as well (4th. and 8th. places, also with strong statistical contents). Ranking 10th. and 15th., we find texts on health measurement (Table 2).

Influential individual titles

Hosmer & Lemeshow’s Applied logistic regression [139] is the title that received the highest number of citations and the highest average annual rate since publication. It is followed by Fleiss [87], Armitage et al. [33], Rothman et al. [250], and Kalbfeisch & Prentice [155] (Table 2). None of the previous 5 texts is affected by limitations on retrieving citations to book chapters not authored by the book editors (see Methods). Except for Hosmer & Lemeshow’s, whose ranking does not change, the ranking of all other 9 books among the top 10 by total citations changes slightly when we use the average number of citations per year: ranking 2nd. is then Rothman et al., 3rd. is Fleiss, and 4th. is Armitage et al. Fifth in citations per year is Sackett et al. Evidence-based medicine. Of books that have been on the market for less than 15 years, only 9 have
### Table 2. Number of citations received by books on epidemiology, biostatistics and related fields.*

| Rank | Author(s) | Brief title & references | Edition, year | Years | Citations (total) | Citations per year |
|------|-----------|--------------------------|--------------|-------|-------------------|--------------------|
| 1    | Hosmer, Lemeshow | Applied logistic regression [139,140] | 1st, 1989 2nd, 2000 | 22    | 19 276            | 876.2              |
| 2    | Fleiss     | Statistical methods for rates and proportions [87–89] | 1st, 1973 2nd, 1981 3rd, 2003 | 38    | 17 173            | 447.4              |
| 3    | Armitage, Berry, Matthews | Statistical methods [33–36] | 1st, 1971 2nd, 1987 3rd, 1994 4th, 2002 | 40    | 16 204            | 405.1              |
| 4    | Rothman, Greenland, Lash | Modern epidemiology [250–252] | 1st, 1986 2nd, 1998 3rd, 2008 | 25    | 12 256            | 490.2              |
| 5    | Kalbfleisch, Prentice | Statistical analysis of failure time data [155,156] | 1st, 1980 2nd, 2002 | 31    | 8097              | 261.2              |
| 6    | Kleinbaum, Kupper, Muller, Nizam | Applied regression analysis [163–166] | 1st, 1978 2nd, 1988 3rd, 1998 4th, 2007 | 33    | 7662              | 232.2              |
| 7    | Breslow, Day | Statistical methods in cancer research. Vols I & II [59,60] | 1st, 1980 2nd, 1987 | 31    | 6584              | 212.4              |
| 8    | Kleinbaum, Kupper, Morgenstern | Epidemiologic research [167] | 1st, 1982 2nd, 1987 | 29    | 5167              | 178.2              |
| 9    | Rosner     | Fundamentals of biostatistics [245–249] | 1st, 1982 2nd, 1990 3rd, 1995 4th, 2000 5th, 2006 | 29    | 4612              | 159.0              |
| 10   | Streiner, Norman | Health measurement scales [273–276] | 1st, 1989 2nd, 1995 3rd, 2003 4th, 2008 | 22    | 4591              | 208.7              |
| 11   | Sackett, Richardson, Straus, et al. | Evidence-based medicine [258–260] | 1st, 1997 2nd, 2000 3rd, 2005 | 14    | 4081              | 291.5              |
| 12   | Sackett, Haynes, Tugwell, Guyatt | Clinical epidemiology [255–257] | 1st, 1985 2nd, 1991 3rd, 2005 | 26    | 4069              | 156.5              |
| 13   | Willett    | Nutritional epidemiology [302,303] | 1st, 1990 2nd, 1998 | 21    | 3968              | 189.0              |
| 14   | Schlesselman, Stolley | Case-control studies [261] | 1st, 1982 | 29    | 3512              | 121.1              |
| 15   | McDowell, Newell | Measuring health [190–192] | 1st, 1987 2nd, 1996 3rd, 2006 | 24    | 2522              | 105.1              |
| 16   | Pocock     | Clinical trials [224,225] | 1st, 1983 | 28    | 2431              | 86.8               |
| 17   | Hill       | Principles of medical statistics [128–130] | 1st, 1937 10th, 1977 12th, 1991 | 74    | 2123              | 28.7               |
| 18   | Last       | Dictionary of epidemiology [12,172–172] | 1st, 1983 2nd, 1988 3rd, 1995 4th, 2001 5th, 2008 | 28    | 2112              | 75.4               |
| 19   | Rose et al. | Cardiovascular survey methods [236–238] | 1st, 1968 2nd, 1982 3rd, 2004 | 43    | 2072              | 48.2               |
| 20   | Hennekens, Buring | Epidemiology in medicine [127] | 1st, 1987 | 24    | 2045              | 85.2               |
| 21   | Lilienfeld, Lilienfeld, Stolley | Foundations of epidemiology [176–178] | 1st, 1976 2nd, 1980 3rd, 1994 | 35    | 1947              | 55.6               |
| 22   | MacMahon, Pugh, Trichopoulus | Principles and methods [180,181] | 1st, 1960 2nd, 1996 | 51    | 1944              | 38.1               |
| Rank | Author(s) | Brief title & references | Edition, year | Years | Citations (total) | Citations per year |
|------|-----------|--------------------------|---------------|-------|------------------|--------------------|
| 23   | Pearl     | Causality [221]          | 1st., 2000    | 11    | 1913             | 173.9              |
| 24   | Fletcher, Fletcher, Wagner | Clinical epidemiology [90–93] | 1st., 1982 2nd., 1988 3rd., 1996 4th., 2005 | 29    | 1732             | 59.7               |
| 25   | APHA, Emerson, et al. | Communicable diseases [66,67] | 17th., 2000 19th., 2008 | 41    | 1623             | 39.6               |
| 26   | Bland     | Introduction medical statistics [56–58] | 1st., 1987 2nd., 1995 3rd., 2000 | 24    | 1586             | 66.1               |
| 27   | Friedman, Furberg, DeMets | Fundamentals of clinical trials [100–103] | 1st., 1981 2nd., 1985 3rd., 1996 4th., 2010 | 30    | 1498             | 49.9               |
| 28   | Hulley, Cummings | Clinical research [146,147] | 1st., 1988 2nd., 2001 | 23    | 1475             | 64.1               |
| 29   | Townsend, Davidson | Inequalities in health [293] | 1st., 1982 | 29    | 1364             | 47.0               |
| 30   | Berkman, Kawachi | Social epidemiology [49] | 1st., 2000 | 11    | 1311             | 119.2              |
| 31   | Blalock    | Causal inference in non-experimental research [52] | 1st., 1964 | 47    | 1269             | 27.0               |
| 32   | Cochrane   | Effectiveness and efficiency [69] | 1st., 1972 | 39    | 1264             | 32.4               |
| 33   | Clayton, Hills | Statistical models [68] | 1st., 1993 | 18    | 1260             | 70.0               |
| 34   | Hosmer     | Applied survival analysis [141,142] | 1st., 1999 2nd., 2008 | 12    | 1254             | 104.5              |
| 35   | Feinstein  | Clinical epidemiology [85] | 1st., 1985 | 26    | 1217             | 46.8               |
| 36   | Marmot, Wilkinson | Social determinants [186,187] | 1st., 1999 2nd., 2006 | 12    | 1214             | 101.2              |
| 37   | Kelsey, Whittmore, Evans, Thompson | Methods in observational epidemiology [157,158] | 1st., 1986 2nd., 1996 | 45    | 1196             | 26.6               |
| 38   | Petitti    | Meta-analysis [222,223] | 1st., 1994 2nd., 2000 | 17    | 1125             | 66.2               |
| 39   | McKeown    | The role of medicine [193] | 1st., 1976 | 35    | 1101             | 31.5               |
| 40   | Rose       | The strategy of preventive medicine [233] | 1st., 1992 | 19    | 1066             | 56.1               |
| 41   | Schottenfeld, Fraumeni | Cancer epidemiology [262–265] | 1st., 1975 1st., 1982 2nd., 1996 3rd., 2006 | 36    | 1055             | 29.3               |
| 42   | Kleinbaum, Klein | Logistic regression [168,169] | 1st., 1994 2nd., 2002 | 17    | 1001             | 58.9               |
| 43   | Holland, Detels, Knox, et al. | Oxford textbook of public health [132–137] | 1st., 1984 2nd., 1991 3rd., 1997 4th., 2002 5th., 2009 | 27    | 977              | 36.2               |
| 44   | Rosenau, Maxcy, Sartwell, Last, et al. | Preventive medicine [236–243] | 1st., 1913 7th., 1951 10th., 1973 11th., 1980 13th., 1992 | 98    | 931              | 9.5                |
| 45   | Doll, Peto | Causes of cancer [73] | 1st., 1981 | 30    | 912              | 30.4               |
| 46   | Kahn, Sempos | Statistical methods in epidemiology [154] | 1st., 1989 | 22    | 901              | 41.0               |
| 47   | Sudman, Bradburn, Wansink | Asking questions [281,282] | 1st., 1982 2nd., 2004 | 29    | 795              | 27.4               |
| 48   | Miettinen  | Theoretical epidemiology [195] | 1st., 1985 | 26    | 794              | 30.5               |
| 49   | Kleinbaum  | Survival analysis [170,171] | 1st., 1997 2nd., 2005 | 15    | 783              | 52.2               |
| 50   | Mausner, Bahn | Epidemiology [188,189] | 1st., 1974 2nd., 1985 | 37    | 767              | 20.7               |
| Rank | Author(s)b | Brief title & referencesb | Edition, yearc | Yearsd | Citations (total)e | Citations per yeare |
|------|------------|----------------------------|----------------|--------|-------------------|---------------------|
| 51   | Meinert    | Clinical trials [194]     | 1st., 1986     | 25     | 702               | 28.1                |
| 52   | Patrick, Erickson | Health status and health policy [220] | 1st., 1993 | 18     | 680               | 37.8                |
| 53   | Feinstein  | Clinimetrics [86]         | 1st., 1987     | 24     | 650               | 27.1                |
| 54   | Szklo, Nieto | Epidemiology. Beyond the basics [268] | 1st., 2000 | 11     | 642               | 58.4                |
| 55   | Gordis     | Epidemiology [108–110]    | 1st., 1996     | 15     | 641               | 42.7                |
| 56   | Checkoway, Pearce, Kriebel | Occupational epidemiology [64,65] | 1st., 1989 2nd., 2004 | 22     | 619               | 28.1                |
| 57   | Rothman    | Epidemiology: an introduction [254] | 1st., 2002 | 9      | 599               | 66.6                |
| 58   | Rosen      | A history of public health [234,235] | 1st., 1958 2nd., 1993 | 53     | 595               | 11.2                |
| 59   | Morris     | Uses of epidemiology [199–201] | 1st., 1957 2nd., 1964 3rd., 1975 | 54     | 555               | 10.3                |
| 60   | Khoury, Beaty, Cohen | Fundamentals of genetic epidemiology [159] | 1st., 1993 | 18     | 533               | 29.6                |
| 61   | Susser     | Causal thinking [283]     | 1st., 1973     | 38     | 475               | 12.5                |
| 62   | Adami, Hunter, Trichopoulos | Textbook of cancer epidemiology [28,29] | 1st., 2002 2nd., 2008 | 9      | 440               | 48.9                |
| 63   | Blalock    | Causal models in social sciences [53,54] | 1st., 1971 2nd., 1985 | 40     | 436               | 10.9                |
| 64   | Morrison   | Screening [202,203]       | 1st., 1985 2nd., 1992 | 26     | 353               | 13.6                |
| 65   | Margetts, Nelson | Nutritional epidemiology [182,183] | 1st., 1991 2nd., 1997 | 20     | 325               | 16.3                |
| 66   | Monson     | Occupational epidemiology [196,197] | 1st., 1980 2nd., 1990 | 31     | 297               | 9.6                 |
| 67   | Bailar, Mosteller, Hoaglin | Medical uses of statistics [40–42] | 1st., 1986 2nd., 1992 3rd., 2009 | 25     | 293               | 11.7                |
| 68   | Abramson   | Survey methods [19–24]    | 1st., 1974 2nd., 1979 3rd., 1984 4th., 1990 5th., 1999 6th., 2008 | 37     | 289               | 7.8                 |
| 69   | Strom      | Pharmacoepidemiology [277–280] | 1st., 1989 2nd., 1994 3rd., 2000 4th., 2005 | 22     | 264               | 12.0                |
| 70   | Elwood     | Causal relationships in medicine [78–80] | 1st., 1988 2nd., 1998 3rd., 2007 | 23     | 263               | 11.4                |
| 71   | Kahn       | Introduction to epidemiologic methods [153] | 1st., 1983 | 28     | 262               | 8.6                 |
| 72   | Gore, Altman | Statistics in practice [111] | 1st., 1982 | 29     | 249               | 8.6                 |
| 73   | Friedman   | Primer of epidemiology [95–99] | 1st., 1974 2nd., 1980 3rd., 1987 4th., 1994 5th., 1998 | 37     | 226               | 6.1                 |
| 74   | Jensen, et al. | Cancer registration [152] | 1st., 1991 | 20     | 208               | 10.4                |
| 75   | Schulze, Perera | Molecular epidemiology [266] | 1st., 1993 | 18     | 204               | 11.3                |
| 76   | Beagleshole, Bonita, Kjellström | Basic epidemiology [48] | 1st., 1993 2nd., 2006 | 18     | 190               | 10.6                |
| 77   | Robertson  | Injury epidemiology [230–232] | 1st., 1992 2nd., 1998 3rd., 2007 | 19     | 184               | 9.7                 |
| Rank | Author(s) | Brief title & references | Edition, year | Years | Citations (total) | Citations per year |
|------|-----------|--------------------------|--------------|-------|------------------|-------------------|
| 78   | Hulka, Wilcosky, Griffith | Biological markers in epidemiology [144] | 1st., 1990 | 21    | 184              | 8.8               |
| 79   | Teutsch, Churchill | Principles and practice of public health surveillance [287,288] | 1st., 1994 2nd., 2001 | 17    | 173              | 10.2              |
| 81   | Khoury, Little, Burke | Human genome epidemiology [161,162] | 1st., 2003 2nd., 2010 | 8     | 171              | 21.4              |
| 81   | Jekel, Elmore, Katz | Epidemiology, biostatistics and preventive medicine [148–150] | 1st., 1996 2nd., 2001 3rd., 2007 | 15    | 163              | 10.9              |
| 82   | Murphy | The logic of medicine [210,211] | 1st., 1976 2nd., 1997 | 35    | 152              | 4.3               |
| 83   | Greenwood | Epidemics and crowd diseases [118] | 1st., 1935 | 76    | 152              | 2.0               |
| 84   | Elliott, et al. | Geographical and environmental epidemiology [74] | 1st., 1992 | 19    | 151              | 7.9               |
| 85   | Greenberg, et al. | Medical epidemiology [113–116] | 1st., 1993 2nd., 1996 3rd., 2001 4th., 2005 | 18    | 147              | 8.2               |
| 86   | Polgar, Thomas | Introduction to research in the health sciences [224] | 1st., 1988 | 23    | 139              | 6.0               |
| 87   | Elston, Johnson | Essentials of biostatistics [75,76] | 1st., 1987 2nd., 1994 | 24    | 134              | 5.6               |
| 88   | Hill | Statistical methods [131] | 1st., 1962 | 49    | 131              | 2.7               |
| 89   | Marmot, Elliott | Coronary heart disease epidemiology [184,185] | 1st., 1992 2nd., 2005 | 19    | 128              | 6.7               |
| 90   | Barker, Rose | Epidemiology in medical practice [43–47] | 1st., 1976 2nd., 1979 3rd., 1984 4th., 1990 5th., 1998 | 35    | 128              | 3.7               |
| 91   | Fox, Hall, Elveback | Epidemiology [94] | 1st., 1970 | 41    | 125              | 3.0               |
| 92   | Doll | Prevention of cancer [71,72] | 1st., 1967 2nd., 2008 | 44    | 124              | 2.8               |
| 93   | Friis, Sellers | Epidemiology for public health practice [104–107] | 1st., 1996 2nd., 1999 3rd., 2004 4th., 2009 | 15    | 115              | 7.7               |
| 94   | Brownson, Baker, Leet, Gillespie | Evidence-based public health [61] | 1st., 2003 | 8     | 114              | 14.3              |
| 95   | Nelson, Williams, Graham | Infectious diseases epidemiology [212,213] | 1st., 2001 2nd., 2006 | 10    | 111              | 11.1              |
| 96   | Weiss | Clinical epidemiology [297,298] | 1st., 1986 2nd., 1996 | 25    | 111              | 4.4               |
| 97   | Rothman | Causal inference [253] | 1st., 1988 | 23    | 107              | 4.7               |
| 98   | Hartzema, Porta, Tilson, Chan | Pharmacoepidemiology [123–126] | 1st., 1988 2nd., 1991 3rd., 1998 4th., 2008 | 23    | 106              | 4.6               |
| 99   | Linet | The leukemias [179] | 1st., 1985 | 26    | 105              | 4.0               |
| 100  | Ahlbom, Norell | Introduction to modern epidemiology [30,31] | 1st., 1984 2nd., 1990 | 27    | 97               | 3.6               |
| 101  | Coggan, Rose, Barker | Epidemiology for the uninitiated [70] | 1st., 1993 | 18    | 92               | 5.1               |
| 102  | Holland | Screening [138] | 1st., 1990 | 21    | 87               | 4.1               |
| 103  | Gregg | Field epidemiology [119–121] | 1st., 1996 2nd., 2002 3rd., 2008 | 15    | 85               | 5.7               |
| 104  | Vineis, et al. | Metabolic polymorphisms [295] | 1st., 1999 | 12    | 83               | 6.9               |
| 105  | Abramson, Abramson | Making sense of data [25–27] | 1st., 1988 2nd., 1994 3rd., 2001 | 23    | 83               | 3.6               |
had more than 40 citations per year: as just mentioned, Sackett et al. Evidence-based medicine (ranking 11th. by total citations and 5th. by citations/year), Pearl (23rd. and 12th., respectively), Berkman & Kawachi (30th. and 16th.), Hosmer’s Survival analysis (34th. and 18th.), Marmot & Wilkinson (36th. and 19th.), Rothman’s Introduction (57th. and 25th.), Szkl & Nieto (54th. and 31st.), Kleinbaum’s Survival analysis (49th. and 34th.), and Adami et al. (62nd. and 36th.).

Figure 2 shows that Hosmer & Lemeshow clearly became the annually most cited book from the mid 1990s on. Modern epidemiology also received a large number of citations, at a rate that nowadays remains upward, in contrast with the three other books shown in Figure 2. This view is complemented by Figure 3, where we see the evolution of citations received by Modern epidemiology and 4 other books with similar and different features. In such Figures there is no presence of books of the so-called “first

### Table 2. Cont.

| Rank | Author(s) | Brief title & references | Edition, year | Years | Citations (total) | Citations per year |
|------|-----------|--------------------------|---------------|-------|-------------------|-------------------|
| 106  | Patrick, Peach | Disablement [219] | 1st., 1989 | 22  | 72 | 3.3 |
| 107  | Khoury, Thomson | Genetics and public health [160] | 1st., 2000 | 11  | 69 | 6.3 |
| 108  | Aschengrau, Seage | Essentials of epidemiology [37,38] | 1st., 2003 2nd., 2008 | 8 | 67 | 8.4 |
| 109  | Elston, Olson, Palmer | Biostatistical genetics [77] | 1st., 2002 | 9 | 65 | 7.2 |
| 110  | Starfield, et al. | Effectiveness of medical care [270] | 1st., 1985 | 26 | 60 | 2.3 |
| 111  | Toniolo, et al. | Biomarkers in cancer epidemiology [292] | 1st., 1999 | 14 | 58 | 4.1 |
| 112  | Morton, Hebel, McCarter | A study guide to epidemiology and biostatistics (204–209) | 1st., 1979 2nd., 1984 3rd., 1990 4th., 1996 5th., 2001 6th., 2005 | 32 | 57 | 1.8 |
| 113  | Timmreck | Introduction to epidemiology [289–291] | 1st., 1994 2nd., 1998 3rd., 2002 | 17 | 55 | 3.2 |
| 114  | Silman, MacFarlane | Epidemiological studies [268,269] | 1st., 1995 2nd., 2002 | 16 | 48 | 3.0 |
| 115  | Bhopal | Concepts of epidemiology [50,51] | 1st., 2002 2nd., 2008 | 9 | 43 | 4.8 |
| 116  | Jenicek | Epidemiology [151] | 1st., 1995 | 16 | 40 | 2.5 |
| 117  | Greenland | Evolution of epidemiologic ideas [117] | 1st., 1987 | 24 | 34 | 1.4 |
| 118  | Olsen, Trichopoulous, Saracci | Teaching epidemiology [215–217] | 1st., 1992 2nd., 2001 3rd., 2010 | 19 | 32 | 1.7 |
| 119  | Morabia | History of epidemiologic methods [14] | 1st., 2004 | 7 | 31 | 4.4 |
| 120  | Buck, et al. | Readings [62] | 1st., 1988 | 23 | 30 | 1.3 |
| 121  | Farmer, Miller, Lawrenson | Lecture notes [81–84] | 1st., 1977 2nd., 1983 3rd., 1991 4th., 1996 | 34 | 30 | 0.9 |
| 122  | Holford | Multivariate methods in epidemiology [132] | 1st., 2002 | 9 | 27 | 3.0 |
| 123  | Armenian, Shapiro | Epidemiology and health services [32] | 1st., 1998 | 13 | 27 | 2.1 |
| 124  | Stolley, Laskey | Investigating disease patterns [272] | 1st., 1995 | 16 | 27 | 1.7 |
| 125  | White, Henderson | Epidemiology as a fundamental science [299,300] | 1st., 1976 2nd., 1980 | 35 | 25 | 0.7 |

*Based on ISI - Thomson - Reuters Web of Science [305].
*Ranked by total number of citations received by all editions in all years (from year of publication until 31 December 2011).
*For identification purposes: only the first authors and an abbreviated title are mentioned; the full reference is provided in the list of references.
*All editions, except for Hill’s Principles of medical statistics.†
*Years elapsed from first edition until the end of 2011.
*Simple average from year of publication until 31 December 2011 (all editions).
*The data source does not allow to distinguish citations to each volume.
*While only the main editions are mentioned in the Table and References, citations to all editions were counted.
*The corresponding (identical) journal article [337] received 2369 citations.

doi:10.1371/journal.pone.0061837.t002
and second generations” [2] (e.g., by Greenwood, Hill, Morris, MacMahon & Pugh, Susser, Lilienfeld & Lilienfeld).

Texts on clinical epidemiology and evidence-based medicine were in positions 11th., 12th., 20th., 24th., 28th or 35th. (by total citations) (Table 2). While other texts on clinical epidemiology appeared earlier on, they were seldom cited. The bloom of clinical epidemiology—and later, evidence-based medicine—in the 1980s is apparent in Figure 4; the Figure suggests that the academic influence of clinical epidemiology was largely due to authors from North America. It should be noted that the third edition of Sackett et al. Clinical epidemiology [255–257] is actually a different book with different contents from the previous two editions, whereas all editions of Fletcher et al. Clinical epidemiology [90–93] are updated revisions of essentially the same text.

Willett’s book on nutritional epidemiology, also strongly focused on methods, ranked 13th.; it is the first book on that specialty. Figure 5 likely reflects an important part of the surfacing of nutritional epidemiology.

Besides introductory works on clinical epidemiology, introductory textbooks on epidemiology do not appear until positions 21st. (Lilienfeld & Lilienfeld) and 22nd. (MacMahon et al., “the first formal epidemiologic textbook ever published in the United States” [5]); the two texts have remarkably similar figures (Table 2). Figure 6 deliberately contrasts books of the “second generation”
[2] (MacMahon et al., Lilienfeld & Lilienfeld) (which we deem highly influential in the education of many epidemiologists and other health professionals, but rarely considered appropriate as a supporting citation in a scientific article) and of the “third generation” (Breslow & Day, and Schlesselman & Stolley) (commonly used in teaching but also acceptable as a citation in a scientific article). They are followed by what we may call the ‘fourth generation’, i.e., texts by Kleinbaum et al., Rothman et al., or Szko & Nieto. Figure 6 reflects the higher weight in the citing literature of research methods-oriented books over textbooks primarily used in the classroom.

Results in Table 2 permit to assess the evolution of citations received by different kinds of books on a similar subject. For instance, Figure 7 shows such evolution for the book by Breslow & Day (a text in two volumes which, though sponsored by the International Agency for Research on Cancer, is more centered on methods than on cancer control), Schottenfeld & Fraumeni Cancer epidemiology [262–265] (a classic example of a large reference volume synthesizing substantive knowledge on cancer, with relatively less attention to methods), Doll & Peto Causes of cancer [73] (a highly influential review, commissioned by the US Congress) [307], and Adami et al. Textbook of cancer epidemiology
(an emerging textbook summarizing substantive knowledge in a smaller format).

Other general introductory texts ranked 26th. (Bland), 50th. (Mausner & Bahn), 55th. (Gordis), 57th. (Rothman), 71st. (Kahn), 72nd. (Gore & Altman), 73rd. (Friedman) or 76th. (Beaglehole et al.). The text by Ahlbom & Norell (100th.) is also an example of a book used in introductory courses and not aiming to be highly cited in the scientific literature. The same applies to Coggon et al. (101st.) [70], a book stemming from a series of articles in a medical journal; also originating in series of articles published in medical journals are books by Hill [128], Bailar & Mosteller [40], and – already mentioned – Gore & Altman [111].

While many of the books included in the study discuss causal inference, books explicitly centered on causality are in positions 23rd. (Pearl), 31st. and 63rd. (Blalock), 61st. (Susser) (“the first book-length treatment of causal reasoning in epidemiology” [308]), and 70th. (Elwood) (Figures S2 and S3). Besides some books on clinical epidemiology, the first book clearly addressed to public health and clinical practice is the American Public Health Association (APHA) *Control of communicable diseases manual* (ranking 25th.). There were three books devoted solely to clinical trials (in positions 16th., 27th. and 51st.) (Table 2, and Figure S4). The top ranking books on social epidemiology were in positions 29th., 30th. and 36th. (time trends are shown in Figure S5).

---

**Figure 5. Evolution of the 3,968 citations received by W. Willett’s ‘Nutritional epidemiology’.**
doi:10.1371/journal.pone.0061837.g005

**Figure 6. Citations to 2 books preferentially used in research and to 2 books preferentially used in teaching (total, 13,987 citations).**
doi:10.1371/journal.pone.0061837.g006
None of the 25 top-cited books has the theoretical, conceptual, historical or sociopolitical approach of works by Cochrane (32nd.) [69], McKeown (39th.) [193], Rose (40th.) [233], Morris (59th.) [199], or Murphy (82nd.) [210] (Figure 8). Figures for the books by Cochrane, McKeown, and Rose (about 1000 total citations each, and 30–60 citations/year) are similar to figures for popular texts outside epidemiology by Kleinman (The illness narratives) [309], Fuchs (Who shall live?) [310], or Knowles (Doing better and feeling worse) [311] (Appendix S1, section 2).

The most cited books on historical aspects of epidemiology and public health were Rosen’s (58th., after 53 years since first publication; average of 11 citations/year) [234] and Morabia’s (119th. after 7 years; 4 citations/year) [14]. Some of the extensive textbooks (e.g., on cancer epidemiology [262–265], public health [133–137], and preventive medicine [236–244]) were in positions 41st., 43rd. and 44th., respectively. Their figures are particularly affected by current limitations on retrieving citations to book chapters not authored by book editors. Other texts on substantive specialties were in positions 56th. and 66th. (occupational epidemiology), 60th. (genetic epidemiology), 62nd. (cancer epidemiology), 64th. and 102nd. (screening), 65th. (nutritional epidemiology), 69th. (pharmacoepidemiology) or 77th. (injury epidemiology). Citations to the text on cardiovascular survey methods by Rose et al. [236] were estimated to be 2072 (ranking 19th.), and to
the book on coronary heart disease by Marmot and Elliott [184], 128 (89th). Six texts on molecular and genetic epidemiology ranked 75th. (Schulte & Perera), 78th. (Hulka et al.), 80th. and 107th. (Khoury et al.), 104th. (Vineis et al.), 109th. (Elston et al.), and 111th. (Toniolo et al.). The only book on meta-analysis in Table 2 is Pettiti's (38th., with over 1100 citations).

Other analyses and comparisons

The study aims did not include to assess the scholarly performance or influence of an author. It will hence suffice to illustrate how citations to books may complement other analyses on such performance. For instance, Figure S5 may add to existing views on the influence of the two books by H. Blalock selected for the study; citations peaked in the 1970s, and were still significant in subsequent decades. Another example: a priori both A. Feinstein and D. Sackett had a strong influence on clinical epidemiology and the methodology of clinical research; yet, a complementary perspective is gained when looking at the citations during the past 25 years to the 2 most cited books by Feinstein (1867 citations) and 2 books by Sackett et al. (8148 citations) (Figure 9). The evolution over time of the bibliographic impact of some books is also illustrated by the two graphs contrasting trends in citations

![Figure 9. Citations to 2 books by D. Sackett et al. (8,150 citations) and 2 books by A. Feinstein (1,867 citations).](https://doi.org/10.1371/journal.pone.0061837.g009)

![Figure 10. Citations to 3 books by K. Rothman et al. (12,962 citations) and 4 books by D. Kleinbaum et al. (14,613 citations).](https://doi.org/10.1371/journal.pone.0061837.g010)
received by 3 books by K. Rothman et al. (12,962 citations) and 4 books by D. Kleinbaum et al. (14,613 citations) (Figure 10). Other examples are given in Figures S6, S7, S8, S9, S10.

**Discussion**

The most prominent pattern in our findings was that the books were mainly cited to reference methods and statistical techniques. The number of citations was larger for books that provide unique sources for (general or specific) methods used in biomedical research papers than for books that innovated in theory, discuss concepts or are preferentially used in teaching. Nine of the 10 most cited texts focused on widely used statistical methods. Some books that have likely been influential –scientifically, theoretically, educationally– in biomedical research or health policy were not highly cited; notable examples include texts by Hill [128], Susser [283], Rose [233], Morris [199], Cochrane [69], or McKewon [193]. Their lower citation figures may partly be due to earlier publication, when all works received less citations because a lower number of articles was published than later on. For instance, Hill’s books ranked 17th. [128–130] and 88th. [131] (Table 2); the first book [128–130] was first published in 1937, decades ahead of all texts in the first 40 positions (Figure S6). Older books on theory and policies were rooted in societal and general medical concerns, while the most modern books are almost purely on methods, particularly statistical methods.

By study design, 1945 was the first year of publication of the citing articles; yet, plenty of books on statistics and epidemiology existed before 1945: several books by Major Greenwood (1880–1949) (e.g., his major opus, *Epidemics and Crowd Diseases*) [118]; Woods and Russels book on teaching statistics, which preceded A.B. Hill’s; the *Epidemiologic Essays* by Crookshank from 1930; *Epidemiology Old and New*, by Hamer, from 1928; or *The Natural History of Disease*, by Ryle, of 1948, which inspired countless public health professionals [14,15]. So, while we had to stick to a relatively limited time-frame we could also make some comparisons with a few selected classics, including the 3 books published between 1913 and 1949 (Table 1) [118,128,236].

The lower citation figures of books on concepts and policies than of books on methods do not necessarily mean that the citing literature pays less attention to theoretical than to methodological issues: such figures may reflect that authors of articles tend to take basic ideas and concepts for granted, as part of the core or background knowledge of most readers, and therefore not needing to be referenced [2,4,8,13]. Citation frequency is only one and modest proxy for influence [1]. A main reason why books on biostatistics top the list may be that they provide convenient citations for statistical techniques widely used in biomedical research. Most biomedical journals request such citations. An important issue is also that findings must be replicable and, therefore, authors cite methods in detail.

It was beyond the scope of the study to register: a) whether citations were to a single technique explained in a specific part of the cited book or a generic citation; b) who was citing (e.g., what epidemiologic or clinical specialties, substantive areas, academic groups); or c) the appropriateness of the citation. As citations to articles [1,13,312–315], citations to books may be shaped by convenience and habit rather than by a genuine need for scientific reference. Many citations to articles are non-specific, not-justified, or wrong [316]. We are not aware of similar studies on book citations, but we suspect that these problems may also be common with citations to books.

The total citing base (over 20 millions of articles with over half a billion citations) [317–319] certainly includes many more empirical studies (which require methodological and technical citations) than articles on theoretical and conceptual issues. However, a key question is whether current practices of referencing do justice to the needs of science. For instance, logistic regression has many variants and implementations; it is important that a research paper explains exactly what was done in the model selection, building, analysis, and validation in the specific application. If instead the paper provides no information on these issues, but just a citation to a book, then science is not well served. Future studies could assess how often the citation also included reference to specific pages or chapters in the book, and whether the citation was specific or just to the whole book.

Future research could use a random sample of the citing articles, their citations to the relevant books, and the 3 factors mentioned above; e.g., see the appropriateness and specificity of the citations, and whether they act as an alibi for not mentioning important methodological details. Such study will need retrieving and data extracting the articles; so, it will have to be limited to a few hundred references, to examine the sentence and context where the reference appears. A future study could also identify influential books that we may have overlooked, use other reliable data sources to measure citations received by books, or obtain data on annual book sales, which could be compared to citation data.

Books primarily meant for students, some of which have likely been highly sold and effective in classrooms—and most influential in coalescing epidemiologic thinking—collected relatively less citations. Important examples include books by Barker et al. [43], MacMahon et al. [180], the Lilienfelds [176], Gordis [108], or Szklö & Nieto [286]. Being used in research is not the primary purpose of such textbooks [2,3]. Knowledge sifted into a textbook has become normal, background, foundational knowledge [4]; there is no need to reference the material, even though it may be of great importance. A rather different case is a book not just addressed to students, but also to practicing physicians: Sackett’s et al. *Evidence-based medicine* [258,259], which gathered more than 4000 citations. An analysis of the development of professional and scientific movements as evidence-based medicine could use and expand some of the study findings [12,14,320,321]. It would in turn provide new insights into the actual influence of some books; e.g., Cochrane’s [69] and Sackett’s [258,259] roles in the creation of the Cochrane Collaboration and the UK National Institute for Health and Clinical Excellence (NICE) [8].

Analyzing citations received by books of epidemiology published over the last 50 years is also a way of tracing the history of the discipline and, to a lesser extent, of biomedical research. Citations contribute to analyze the evolution of the corpus of methods and concepts used by epidemiology and by the disciplines with which epidemiology interacts [2,8,12–15,322,323]. The rise of multivariate methods is a relevant example; it clearly explains a large portion of the citations made by papers published in the 1980s. Such methods were then not yet ‘normal science’ [4] in epidemiological and clinical research, and were hence in need of referencing.

The decade 1980–1989 stood out as the period in which more highly-cited books were published. Although we ignore the total number of books on epidemiology and biostatistics published in each period, missing books are unlikely to be that influential so as to change this finding. Overall, other than a few dozen highly-cited books, most books had limited or minimal citation impact. Books published in recent years do not seem to have the influence of textbooks published for the first time in the 1980s. Moreover, the total annual citations to books have been rather steady in the last 15 years, or even declining for many specific texts. Given the large increase in the annual number of citations registered in the
Web of Science, the relative citation influence of the study books is probably decreasing sharply, when seen as a proportion of total citations. The total number of citations to the study books (some 183 000 over half a century) probably represents a small fraction of the total citations to the epidemiological and statistical literature, which we estimate to be several million among the over half a billion citations indexed by the Web of Science. Even the few highly-cited books seem to have lost their relative impact in the last 2 decades, because between 1980 and 2011 the number of citing articles (and, correspondingly, citations) increased several-fold. It would next be necessary to perform some adjustment of the citation counts for the total number of citations each year. A comparison with books from other disciplines would then also be meaningful. In spite of its limitations (e.g., choice of citing journals, focus on research-oriented, largely Anglo-Saxon journals) the study data source has long enabled relevant analyses [1,6,312–315,324].

We are not aware of any typology of books that could be useful for the present study. Although it would have been possible to use narrower definitions of what is a book on epidemiology and biostatistics, we chose a wide and inclusive approach, which yields a broad perspective and enables a richer set of comparisons; it also allows readers to conduct other, more focused comparisons. The inclusive approach seems particularly appropriate at a time when no other analyses are yet available.

In spite of the study limitations, we think that just comparing the number and trends of citations to the books that we identified is of interest. The uses of books and their true impact on research and professional practices are questions that often seem to escape analysis. Yet books comprise an important part of scientific production, and evaluations of such production would benefit from considering books along scientific articles. Several projects are under development to register citations to books [11,325,326]. However, none has at present the scope in time and the qualitative features of the present study. Analyses of citations to books are just one approach to analyses of uses of books. An additional task for future research would be to analyze the relative importance of books and articles in the exercise of scientific leadership and social influence by selected authors. An example is shown in Figure S10, which shows that the increase in the number of citations to the book on logistic regression by Hosmer & Lemeshow did not come at the expense of the highly cited article by Mantel & Haenszel [327].

Several influential authors wrote more than one or two books. In the case of M. Susser, for instance, his Sociology in medicine [328] has accumulated over 300 citations. To value Susser’s oeuvre it is thus necessary to assess several other books of his [283,306,329–331]. J. Last, D. Sackett, O. Miettinen, J. Morris, W. Holland, D. Kleinbaum, A. Feinstein, K. Rothman, M. Khoury, H. Blalock, G. Friedman or J. Abramson are other –very different, obviously– examples of authors with more than one influential book [3,4,5,14,15,323,332,333]. Some books on research methods were less cited than articles on the same methods; for instance, some papers by Hill [334] or Miettinen [335] have each accumulated over 2400 and 1500 citations, respectively [6]. However, seminal papers published long ago received comparatively few citations; for instance, a paper by Jerome Cornfield [336], published over 60 years ago, received less than 450 citations. The review on the causes of cancer by Doll & Peto [307] provides a telling contrast between citations received by the book [73] (about 900 citations) and the –identical– journal article [337] (some 2300 citations); the case illustrates the need to exert caution when dealing with different formats and vehicles of scientific communication, an issue that is getting ever more complex with digital formats and the Internet [11,338]. Several other articles by Doll & Peto have each received over 1500 citations. The most cited paper by Willett [339] has received some 2000 citations, about half than his book [302]. But some sets of papers by him and other authors are, together, more cited than the reference book by the same author(s). The most cited papers by Cochrane [340,341] received some 500 citations, less than half his book [69]. The top-cited paper by Rose [342] received some 1000 citations, similar to his book [293]; and again, several other papers of his have been highly cited. All articles coauthored by K. J. Rothman received over 12400 citations, a similar amount than Modern epidemiology [250–252]. Just two of several books by David G. Kleinbaum combined –Applied regression analysis [163–166] and Epidemiologic research [167]– received over 17000 citations (Table 2), whereas his articles have some 3000 citations.

To conclude, results may reflect a considerably positive influence of epidemiologic and statistical science on biomedical research [139,140,199–201]. On the other hand, results on the smaller number of citations received by theoretical and policy-oriented books cannot discern whether the influence was weaker on the ideas and policies that so much affect the health and wellbeing of citizens.

Supporting Information
Appendix S1 Details on Materials and Methods. (DOCX)

Table S1 Examples of books that received less than 25 citations. (DOCX)

Figure S1 Two examples of books whose titles were recorded in different ways in the Web of Science, the Oxford Textbook of Public Health and Modern Epidemiology. (TIF)

Figure S2 Citations to 4 books on causality (total, 2,529 citations). (TIF)

Figure S3 Citations to 2 books by H. Blalock (1,705 citations). (TIF)

Figure S4 Citations to 3 books on clinical trials: S.J. Pocock (2,421 citations), L.M. Friedman et al. (1,498 citations), and C. Meinert (703 citations). (TIF)

Figure S5 Citations to 3 books on social epidemiology: Berkman & Kawachi (1,311 citations), Marmot & Wilkinson (1,214 citations) and Townsend & Davidson (1,364 citations). (TIF)

Figure S6 Evolution over time of the 2,123 citations received by A. B. Hill’s ‘Principles of medical statistics’. (TIF)

Figure S7 Evolution over time of the 475 citations received by Susser’s ‘Causal thinking in the health sciences’. (TIF)

Figure S8 Evolution over time of the 733 citations received by 3 books by M. Khoury. (TIF)
Figure S9  Citations to 2 books on occupational epidemiology: H. Checkoway et al. (619 citations) and R.R. Monson et al. (297 citations).

(TIF)

Figure S10  Trends of citations to the book by Hosmer & Lemeshow, and to the article by Mantel & Haenszel.

(TIF)

Acknowledgments

The authors gratefully acknowledge technical assistance from José Pumarega, Magda Gasull, and Yolanda Rovira.

References

1. Garfield E (1955) Citation indexes for science. A new dimension in documentation through association of ideas. Science 122: 100-111.
2. Zhang FF, Michaels DC, Mathema B, Kaushal S, Chatterjee A, et al. (2004) Evolution of epidemiologic methods and concepts in selected textbooks of the 20th century. Soc Praventivmed 49: 97–104.
3. Bhopal R (1997) Which book? A comparative review of 25 introductory epidemiology textbooks. J Epidemiol Community Health 51: 612–622.
4. Bhopal R (1999) Paradigms in epidemiology textbooks: In the footsteps of Thomas Kuhn. Am J Public Health 89: 1162–1165.
5. Krueger N (1998) Epidemiology and the web of causation: has anyone seen the spider? Soc Sci Med 39: 807–903.
6. Glaser J (2004) Why are the most influential books in Australian sociology not necessarily the most highly cited ones? J Sociol 40: 261–262.
7. Porta M, Fernandez E, Bolu´mar F (2006) The “bibliographic impact factor” and the still uncharted sociology of epidemiology. Int J Epidemiol 35: 1130–1135.
8. Alvarez-Dardet C, Ruiz MT (1993) Thomas McKeown and Archibald Maclean: a journey through the diffusion of their ideas. BMJ 306: 1252–1255.
9. Nederhof AJ, van Leeuwen TN, van Raan AJ (2010) Highly cited non-journal publications in political science, economics and psychology: a first exploration. Scientometrics 83: 363–374.
10. Astron P (2011) Comparing citation patterns in entrepreneurship research articles in subject handbooks and Web of Science journals. In: Noyons E, Nguêhe P, Leto J, editors. Proceedings of the 13th Conference of the International Society for Scientometrics and Informetrics (ISSI) (Durban, South Africa, July 4–7, 2011), Vols. 1 and 2. Proceedings of the International Conference on Scientometrics and Informetrics. pp. 89–96.
11. Kousha K, Thelwall M, Rezaie S (2011) Assessing the citation impact of books: the role of Google Books, Google Scholar, and Scopus. J Am Soc Inform Sci Technol 62: 2147–2164.
12. Porta M. (2008) A dictionary of epidemiology, 5th edition. New York: Oxford University Press.
13. Leydesdorff L, Amsterdamska O (1990) Dimensions of citation analysis. Sci Technol Human Values 15: 305–335.
14. Morabia A (2004) A history of epidemiologic methods and concepts. Basel: Birkhäuser/Springer.
15. Amsterdamska O (2005) Demarcating epidemiology. Sci Technol Human Values 30: 17–51.
16. Porta M, Fernandez E, Puigdomènech E (2006) Citations received by books: an approach to the influence of epidemiologic thought within the academic communities? Rev Saúde Pública 40 (Supl): 50–56.
17. The James Lind Library. Available: http://www.jameslindlibrary.org/ Accessed 20 March 2013.
18. People’s Epidemiology Library. Available: http://www.epidemiology.ch/history/PeopleEpidemiologyLibrary.html Accessed 20 March 2013.
19. Abramson JH (1974) Survey methods in community medicine, 1st edition. London: Churchill Livingstone.
20. Abramson JH (1979) Survey methods in community medicine, 2nd edition. London: Churchill Livingstone.
21. Abramson JH (1984) Survey methods in community medicine, 3rd edition. London: Churchill Livingstone.
22. Abramson JH (1990) Survey methods in community medicine, 4th edition. London: Churchill Livingstone.
23. Abramson JH, Abromson ZH (2008) Research methods in community medicine, 6th edition. Chichester, West Sussex: John Wiley and Sons.
24. Abramson JH (1998) Making sense of data. A self-instruction manual of the interpretation of epidemiological data. New York: Oxford University Press.
25. Abramson JH (1994) Making sense of data. A self-instruction manual of the interpretation of epidemiological data, 2nd edition. New York: Oxford University Press.
26. Abramson JH, Abramson ZH (2001) Making sense of data. A self-instruction manual of the interpretation of epidemiological data, 3rd edition. New York: Oxford University Press.
27. Abramson JH, Abramson ZH (2002) Textbook of cancer epidemiology. New York: Oxford University Press.
28. Adami HO, Hunter D, Trichopoulos D, eds. (2002) Textbook of cancer epidemiology. New York: Oxford University Press.
29. Adami HO, Hunter D, Trichopoulos D, eds. (2008) Textbook of cancer epidemiology, 2nd ed. New York: Oxford University Press.
30. Ahlbom A, Norr¨e S (1984) Introduction to modern epidemiology, 1st ed. Chestnut Hill, MA: Epidemiology Resources.
31. Ahlbom A, Norr¨e S (1990) Introduction to modern epidemiology, 2nd ed. Chestnut Hill, MA: Epidemiology Resources.
32. Armenian HK, Shapiro S (1998) Epidemiology and health services. New York: Oxford University Press.
33. Armitage P (1971) Statistical methods in medical research, 1st ed. Oxford: Blackwell.
34. Armitage P, Berry G (1987) Statistical methods in medical research, 2nd ed. Oxford: Blackwell.
35. Armitage P, Berry G (1994) Statistical methods in medical research, 3rd ed. Oxford: Blackwell.
36. Armitage P, Berry G, Matthews JNS (2002) Statistical methods in medical research, 4th ed. Oxford: Blackwell.
37. Aschengrau A, Seage GR (2003) Essentials of epidemiology in public health. Sudbury: Jones & Bartlett.
38. Aschengrau A, Seage GR (2008) Essentials of epidemiology in public health, 2nd ed. Sudbury: Jones & Bartlett.
39. Ashton J, ed. (1994) The epidemiological imagination: A reader. Philadelphia, PA: Open University Press.
40. Bailar JC, Mosteller F, eds. (1986) Medical uses of statistics, 1st ed. Waltham, MA: New England Journal of Medicine Books.
41. Bailar JC, Mosteller F, eds. (1990) Medical uses of statistics, 2nd ed. Waltham, MA: New England Journal of Medicine Books.
42. Bailar JC, Hoaglin DC, eds. (2009). Medical uses of statistics, 3rd ed. Hoboken, NJ: John Wiley & Sons.
43. Barker DJP, Rose GA (1976) Epidemiology in medical practice, 1st ed. Edinburgh: Churchill Livingstone.
44. Barker DJP, Rose GA (1979) Epidemiology in medical practice, 2nd ed. Edinburgh: Churchill Livingstone.
45. Barker DJP, Rose GA (1984) Epidemiology in medical practice, 3rd ed. Edinburgh: Churchill Livingstone.
46. Barker DJP, Rose GA (1990) Epidemiology in medical practice, 4th ed. Edinburgh: Churchill Livingstone.
47. Barker DJP, Cooper C, Rose GA (1998) Epidemiology in medical practice, 3rd ed. Edinburgh: Churchill Livingstone.
48. Beaglesløe R, Bonita R, Kjellstrom T (1993) Basic epidemiology. Geneva: World Health Organization.
49. Berkman LF, Kawachi I (2000) Social epidemiology. New York: Oxford University Press.
50. Bhogal R (2002) Concepts of epidemiology. An integrated introduction to the ideas, theories, principles and methods of epidemiology. New York: Oxford University Press.
51. Bhogal R (2008) Concepts of epidemiology. An integrated introduction to the ideas, theories, principles and methods of epidemiology, 2nd ed. New York: Oxford University Press.
52. Blalock HM (1964) Causal inference in non-experimental research. Chapel Hill, NC: University of North Carolina Press.
53. Blalock HM Jr, ed. (1971) Causal models in the social sciences, 1st ed. New York: Aldine.
54. Blalock HM Jr, ed. (1985) Causal models in the social sciences, 2nd ed. Chicago: Aldine.
55. Blalock HM Jr, ed. (1985) Causal models in panel and experimental designs. New York: Aldine.
56. Bland M (1987) An introduction to medical statistics, 1st ed. Oxford: Oxford University Press.
57. Bland M (1995) An introduction to medical statistics, 2nd ed. Oxford: Oxford University Press.
58. Bland M (2000) An introduction to medical statistics, 3rd ed. Oxford: Oxford University Press.

Author Contributions

Conceived and designed the experiments: MP. Performed the experiments: MP SS TL. Analyzed the data: TL SS MP. Contributed reagents/materials/analysis tools: MP JPV JFP SS TL. Wrote the paper: MP JPV JFP EF RB AM CV TL SS. Are aware of, and accept responsibility for, the manuscript; read and approved the present version of the article; and made a significant intellectual contribution to the research and gave permission to submit the manuscript for possible publication: MP JPV JFP SS EF RP AM CV TL.
312. Garfield E (1972) Citation analysis as a tool in journal evaluation. Science 178: 471–479.
313. Garfield E (1985) Uses and misuses of citation frequency. Current Contents 43: 3–9.
314. Seglen PO (1992) How representative is the journal impact factor? Research Evaluation 2: 143–149.
315. Porta M (1996) The bibliographic ‘impact factor’ of ISI, Inc.: how relevant is it really for public health journals? J Epidemiol Community Health 50: 606–610.
316. Toronzo A, Bonito MG, Ioannidis JP (2007) Persistence of contradicted claims in the literature. JAMA 298: 2517–2526.
317. Evans JA (2008) Electronic publication and the narrowing of science and scholarship. Science 321: 395–399.
318. Ioannidis JP (2006) Concentration of the most-cited papers in the scientific literature: analysis of journal ecosystems. PLoS One 1: e3.
319. Lariviere V, Gingras Y, Archambault E (2009) The decline in the concentration of citations, 1990–2007. J Am Soc Inf Sci Technol 60: 858–862.
320. Vandenbroucke JP (1998) Clinical investigation in the 20th century: the ascendancy of numerical reasoning. Lancet 352 (Suppl 2): 12–16.
321. Vandenbroucke JP (1998) Medical journals and the shaping of medical knowledge. Lancet 352: 2001–2006.
322. Porta M, Álvarez-Dardet C (1998) Epidemiology: bridges over (and across) roaring levels. J Epidemiol Community Health 52: 605.
323. Bolu´mar F, Porta M (2004) Epidemiologic methods: beyond clinical medicine, beyond epidemiology. Eur J Epidemiol 19: 733–735.
324. Garfield E (1987) The 101 most-cited papers from the British Medical Journal highlight the important role of epidemiology in medicine. Current Contents 7: 3–11.
325. Thomson Reuters. Book Citation Index. Available: http://wokinfo.com/products_tools/multidisciplinary/bookcitationindex/. Accessed 20 March 2013.
326. Asham J, Testa J (2011) Thomson Reuters Book Citation Index. In: Noyons E, Nguhüe P, Leta J. Proceedings of the 13th Conference of the International Society for Scientometrics and Informetrics (ISSI) [Durban, South Africa, July 4–7, 2011], Vols. 1 and 2. Proceedings of the International Conference on Scientometrics and Informetrics. pp. 13–18.
327. Mantel N, Haenszel W (1959) Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 22: 719–748.
328. Sussel M, Watson W, Hopper K, eds. (1985) Sociology in medicine, 3rd ed. New York: Oxford University Press.
329. Sussel M (1987) Epidemiology, health and society. Selected papers by Mervyn Sussel. New York: Oxford University Press.
330. Sussel M, Stein Z (2009) Eras in epidemiology: the evolution of ideas. New York: Oxford University Press.
331. Porta M (2006) Things that kept coming to mind while thinking through Sussel’s South African memoir. J Epidemiol Community Health 60: 539–541.
332. Krieger N (2007) Ways of asking and ways of living: reflections on the 50th anniversary of Morris’ ever-useful Uses of Epidemiology. Int J Epidemiol 36: 1173–1180.
333. Black N, Neuhauser D (2006) Books that have changed health services and health care policy. J Health Serv Res Policy 11: 180–183.
334. Hill AB (1965) Environment and disease – association or causation. Proc Royal Soc Med London 58: 295–300.
335. Miettinen O (1976) Estimability and estimation in case-referent studies. Am J Epidemiol 103: 226–233.
336. Cornfield J (1951) A method of estimating comparative rates from clinical data – applications to cancer of the lung, breast, and cervix. J Natl Cancer Inst 11: 1269–1275.
337. Doll R, Peto R (1981) The causes of cancer –quantitative estimates of avoidable risks of cancer in the United States today. J Natl Cancer Inst 66: 1191–1308.
338. Peter J (2008) Testing the calculation of a realistic h-index in Google Scholar, Scopus, and Web of Science for F. W. Lancaster. Library Trends 56: 784–815.
339. Willet WC, Sampson L, Stampfer MJ, Rosner B, Bain C, et al. (1985) Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 122: 51–65.
340. St Leger AS, Cochrane AL, Moore F (1979) Factors associated with cardiac mortality in developed countries with particular reference to the consumption of wine. Lancet 1: 1017–1020.
341. Elwood PC, Cochrane AL, Burr ML, Sweetnam PM, Williams G, et al. (1974) A randomized controlled trial of acetyl salicylic acid in the secondary prevention of mortality from myocardial infarction. Br Med J 1: 436–440.
342. Rose G (1983) Sick individuals and sick populations. Int J Epidemiol 14: 32–38.
Supplementary data.

1. Details on the search of citations

In our searches, we used combinations of ‘cited author’ and ‘cited work’ as, for example, ‘(H*er AND Ap* log*)’ for *Applied logistic regression* (by David W. Hosmer and Stanley Lemeshow). Such truncation gives results as ‘HASMER D APPL LOGISTIC REGRES - HOASMER DW – APPL LOGISTIC REGRES’, ‘HOEMER DW – APPL LOGISTIC REGRES’, ‘HOFMER DW – APPL LOGISTIC REGRES’, ‘HOSMER D – APPL LOG REGRESSION’. Books with more than one author or editor required searches for all authors or editors. In the previous example, we completed the search with ‘Lemes* AND Ap* log*’.

Errors in citations or in the transcription of citations from the articles to the data source abounded. For instance, the name of Fleiss (author of *Statistical methods for rates and proportions*) appears as Feiss, Feliss, Fleis, Fless, Fliss, Leiss, Lfeiss, Meiss, Reiss, Sleiss, Fklesiss, Fleiss or Fleise; and the title of the book (besides author ‘Fleiss’ or similar names), appears as ‘Stat Methods Weight P’, ‘Stat Methods Dates P’, ‘Stat Methods Goes Ra’ or Stat Meth Jbr Rat’. Such errors in citations require that searches are performed looking individually for many possible combinations. By looking at the actual citing articles we verified that sometimes the error in the citation appeared in the article, and sometimes in the data base.

Information on cited books in the data source was verified against valid information on all editions of the book; for instance, the year of publication that appeared in a citation in the data source was checked against information from the publisher’s web, from online catalogues of libraries (e.g., the British Library, Harvard’s Countway Library of Medicine and Hollis catalogues), from book pages available through search engines (e.g., Google Scholar) and websellers (e.g., Amazon).

Other examples of searches are:

– *Statistical methods* (Armitage et al.). Searched as ‘A*mi* AND st*’ it appears recorded as, for example, ‘AMITAGE – STAT METHODS MED RES’, ‘ARMI – STAT METHODS MED RES’, ‘ARMIGATE P – STATISTICAL METHODS’, ‘ARMIGTAGE PBG – STAT METHODS MED RES’, ‘ARMITAGE P – STA METHODS MED RES’. With this search we obtained some results of citations to articles published by the journal *Statistics in Medicine*, which we excluded, for instance, ‘ARMITAGE P – STAT MED – 1993 – [volume] 12 – [page] 1533’. We also searched the other editors of the book (e.g., as ‘Berry AND st* met*’ and ‘Matth* AND st* met*’).

– *Applied regression analysis* (Kleinbaum et al.). We searched for citations to ‘K* AND ap* reg*’. Thus, for instance, we got Cited Authors in results as ‘KEIBAUN DG – KELIMBAUM DG’. As in previous examples we searched for all the editors of the book: ‘Kupp* AND ap* reg’, ‘Mull* AND ap* reg’, ‘Niz* AND ap* reg’.

– The book *Preventive medicine and hygiene* (Rosenau) raises another issue: it has changed title and editors in different editions; e.g., Maxcy KF, ed., *Rosenau Preventive medicine and hygiene*, 7th. ed.; Sartwell PE, ed., *Maxcy-Rosenau Preventive medicine and public health*, 10th. ed.; Last JM, ed., *Maxcy-Rosenau Public health and preventive medicine*, 11th. ed.; Last JM, Wallace RB, eds., *Maxcy-Rosenau-Last Public health and preventive medicine*, 13th. ed; Wallace RB, ed. Maxcy-Rosenau-Last Public health and preventive medicine. 15th ed. New York: McGraw-Hill, 2008. We counted all citations to all editions taking into account all variants and possible errors mixing authors and titles. Thus, for instance, searched as ‘Rose* AND prev*’ we got results as ‘ROSENAU MJ – PREVENTIVE MED HYGIE’, ‘ROSENAU – PREVENTIVE MED HYGIE’ or ‘ROSENAU – PREVENTIVE
MEDICINE – 1956’. The last result seems to be a citation to the second revision by Maxcy of Rosenau’s work, but the citing article actually considered Rosenau the only author. Further examples of the searches for this book are: ‘[Cited author] Ma*cy AND [Cited work] Rosen*’, ‘Ma*cy AND prev*’, ‘Last AND pub* h*’, ‘[Cited author] Last AND [Cited work] Maxcy*’, ‘[Cited author] Sar*w* AND [Cited work] Ma*cy*’; ‘Sartw* AND prev*’.

ISI - Thomson - Reuters data does not actually allow to reliably identify chapters of books written by authors other than the book editors, unless the title of the book is unique or unambiguous. The present study does include citations to chapters that were cited by the citing articles using the name of the book editor, the book title, and the pages corresponding to the chapter.

We next offer examples of books whose title was unambiguous and for which, therefore, the number of citations included in the present study includes all citations (i.e., it also includes citations to individual chapters not authored by the book editors):

– Oxford Textbook of Public Health (Holland et al.). When searched as ‘(oxf* t*x* p* h*) OR (oxf* t*x* pub*) ’ and variants, it appears recorded in the data base as, for example, ‘OXFORD TXB PUBLIC HL’, ‘OXFORD TXB PULIC HLT’, ‘OXFORD TXB PUBL HLTH’ ‘OXFORD TEXTBOOK PUBL’ See also Supplementary data, Figure 1.

– Modern Epidemiology (Rothman et al.). When searched as ‘mod* epi*’, it appears recorded in the data base as, for example, ‘MODERN EPIDEMIOLOGY’. The truncation ‘mod* epi*’ required us to discard one by one citations to books registered in the data base as, for instance, ‘MODELING AIDS EPIDEM’, ‘MODELLING EPISODES A’ or ‘MODELS SEIZURES EPIL’. Citations to articles whose titles begin with words similar to the initial title of the book of interest were also easily excluded during this phase; in the ISI - Thomson - Reuters data base it is easy to differentiate a citation to an article and a citation to a book. A longer truncation as, for instance, ‘modern epi*’ would miss citations with mistakes as ‘MODEN EPIDEMIOLOGY’.

– Teaching epidemiology (Olsen et al.). Searched as ‘tea* epi*’, it appears recorded in the data base as, for example, ‘OLSEN J – TEACHING EPIDEMIOLOG – 1992’ or as ‘FLOREY CD – TEACHING EPIDEMIOLOG – 1992 – [page] 23’. When a citation mistake was obvious, the citation was counted (e.g., we did include one citation to ‘OLSEN J – TEACHING EPIDEMIOLOG – 1993’, instead of 1992). By checking each possible citation there was little doubt that certain citations should not be counted (e.g., ‘ROLLESTON TW – TEACHING EPICTETUS B – 1888’ or ‘VOS G – TEACHING EPIST HEBRE – 1956’).

For the following books it was feasible to search and count citations to individual chapters, either because the title was unequivocal or because we specifically searched for citations that included the name of the first author of each and all chapters plus the title of the book; we also started by identifying all citations that included the name of the book, looked at the names of the authors of such citations, verified which names where actually authors of chapters of the book, and counted such citations as citations to the book (please, see full citations in references of main text):

– Rothman KJ. Modern epidemiology (1986). Rothman KJ, Greenland S, eds. Modern epidemiology. 2nd. ed. (1998). Rothman KJ, Greenland S, Lash TL, eds. Modern epidemiology. 3rd. ed. (2008).

– Holland WW, Detels R, Knox, eds Oxford Textbook of Public Health (1984) and 2nd ed. (1991). Detels R, McEwen J, Beaglehole T, Tanaka H, eds Oxford Textbook of Public Health 3rd ed.(1997) and 4th ed. (2002).Detels R, Beaglehole T, Lansang MA, Gulliford M, eds. Oxford Textbook of Public Health. 5th. ed. (2009).

– Jensen O, Parkin DM, MacLennan R, Muir CS, Skeet R. Cancer registration (1991).
– Adami HO, Hunter D, Trichopoulos D, eds. Textbook of cancer epidemiology. 1st. ed. (2002), 2nd. ed. (2008).
– Nelson KE, Williams CM, Graham NMH. *Infectious diseases epidemiology. Theory and practice*. 1st ed. (2001) and 2nd ed. (2007).
– Khoury MJ, Little J, Burke W. *Human genome epidemiology* 1st ed. (2003). Khoury MJ, Bedrosian S, Gwinn M, et al., eds *Human genome epidemiology* 2nd ed. (2010).
– Gregg MB, ed. *Field epidemiology*. 1st ed. (1996), 2nd ed. (2002), 3rd ed. (2008).
– Olsen J, Trichopoulos D, eds. *Teaching epidemiology*. 1st ed. (1992). Olsen J, Saracci R, Trichopoulos D, eds. *Teaching epidemiology*. 2nd ed. (2001), 3rd ed. (2010).
– Buck C, Llopis A, Nájera E, Terris M, eds. *The challenge of epidemiology* (1988).
– Elston R, Olson J, Palmer L, eds. *Biostatistical genetics and genetic epidemiology* (2002).
– Toniolo P, Boffetta P, Shuker DEG, et al., eds. *Application of biomarkers in cancer epidemiology* (1997).
– Armenian HK, Shapiro S. *Epidemiology and health services* (1998).
– Ashton J, ed. *The epidemiological imagination: A reader* (1994).
– Wild C, Vineis P, Garte S, eds. *Molecular epidemiology of chronic diseases* (2008).
– Carrington M, Hoelzel R. *Molecular epidemiology* (2001).
– Rebbeck TR, Ambrosone CB, Shields PG, eds. *Molecular epidemiology: applications in cancer and other human diseases* (2008).
– Hartzema AG, Porta MS, Tilson HH, eds. *Pharmacoepidemiology. An introduction*. 2nd ed (1991), 3rd ed (1998). Hartzema AG, Tilson HH, Chan KA, eds *Pharmacoepidemiology and therapeutic risk management* (2008). In this case we did as explained at the start of this paragraph, additionally taking into account that several authors of chapters had written chapters in the book edited by Brian Strom (Pharmacoepidemiology); it was possible to distinguish between the two books through the year of publication, by reading the reference in the citing article, through the Digital object identifier (doi), PubMed or the citing journal’s web page. In addition, several citations cited articles published by authors of Hartzema’s et al. book in the journal *Pharmacoepidemiology and drug safety*, which it was feasible to identify (and not count as citations to the book) since such citations included as source ‘PHARMACOEPI DEM DR S’, ‘PHARMACOEPI DEMIOL S1’ and variants.

By contrast, the title of some books is unspecific or similar to a journal’s name; hence, it was not possible to identify all citations using the name of the first author of chapters not authored by the book editors. Examples include:
– Schottenfeld D, Fraumeni FJ Jr, eds. *Cancer epidemiology and prevention* (1st ed.: 1982).
– Rosenau, Maxcy, Sartwell, Last, eds. *Preventive medicine* (1st ed.: 1913).
– Hulley SB, Cummings SR, eds. *Designing clinical research* (1st ed.: 1988).
– Strom BL, ed. *Pharmacoepidemiology* (1st ed.: 1989).
– Sackett DL et al, ed. *Evidence-based medicine: how to practice and teach EBM* (1st ed.: 1997). The title of this book partly coincides with the name of the journal *Evidence-based medicine for primary care and internal medicine*. To distinguish between citations to the book and to articles in the journal by the book authors we looked at the references in the citing articles; again, the articles were identified through the Digital object identifier (doi) or PubMed. It was often possible to distinguish between citations to the book and to articles in the journal because the year of publication in the citation was not one of the years of publication of any of the editions of the book, or because the citation included an article ID, although none of these two circumstances was required to make the distinction. Citations that started as ‘EVIDENCE BASED MED H’ were safely attributable to the book. Citations to works in “EVIDENCE BASED MED” could be to the book or the journal articles, and were checked as mentioned; most were citations to the book. Citations to other books presenting similar issues were treated in the same way.

Certainly, titles of some books may be thought to fall within the two groups mentioned above (i.e., books with unambiguous and common titles); for instance:

– Jensen O, Parkin DM, MacLennan R, Muir CS, Skeet R. *Cancer registration: principles and methods*. Lyon: International Agency for Research on Cancer, 1991. We searched for it with ‘canc*
reg*’ OR ‘princi* meth*’ and checked the year of publication (as well as the authors, of course). When we used ‘canc* reg* p*’, 9 citations were missed, including 8 to ‘JENSEN OM – CANCER REGISTRATION – 1991’. As mentioned above, it was feasible to search and count citations to individual chapters.

Additional examples of books that could not be assessed for inclusion are:

– IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, IARC Handbooks of Cancer Prevention, and the series of *Cancer Incidence in Five Continents*.
– White KL, Frenk J, Ordóñez C, et al., eds. *Health services research: An anthology*. Washington, DC: Pan American Health Organization; 1992.

In summary, if the title was unambiguous, the number of citations included in the present study includes citations to individual chapters not authored by the book editors; if the title of the book was common, we tried to include all citations registered with the name of the first author of each chapter, as explained above.

At present, the data source does not permit to include citations to books made in other books.

2. Referents: books not focused on epidemiology and biostatistics

We next provide a few examples of texts (mostly, on health policy, philosophy of medicine and related areas) in which epidemiology and biostatistics have a marginal role, but which nevertheless may act as referents of the study main findings:

– Paul Starr. The social transformation of American medicine (1st. ed., 1982): 3,121 citations.
– Ivan Illich. Medical nemesis (1st. ed.; 1975); Limits to medicine: Medical nemesis (2nd. ed.: 1976) (1977-1982): 2,024 citations.
– Arthur Kleinman. The illness narratives. Suffering, healing and the human condition (1st. ed., 1988): 1,640 citations.
– Victor Fuchs. Who shall live? Health, economics, and social choice (editions of 1974, 1975, 1983, 1986, 2001): 628 citations.
– John H. Knowles. Doing better and feeling worse. Health in the United States (1st. ed., 1977): 325 citations (including citations to chapters).
– McMichael AJ. Human frontiers, environments and diseases. Past patterns, uncertain futures. Cambridge: Cambridge University Press (1st. ed., 2001): 154 citations.
– Siegrist J, Marmot M. Social inequalities in health. New York: Oxford University Press (1st. ed., 2006): 140 citations (including citations to chapters).
– Kawachi I, Kennedy BP. The health of nations. Why inequality is harmful to your health. New York: The New York Press (1st. ed., 2002): 94 citations (including citations to chapters).
– Kawachi I, Wamala S. Globalization and health. New York: Oxford University Press (1st. ed., 2006): 22 citations.
– Milos Jenicek. Foundations of evidence-based medicine (1st. ed., 2003): 34 citations.
– Jenicek M, Hitchcock DL. Evidence-based practice. Logic and critical thinking in medicine. Chicago: American Medical Association (AMA Press) (1st. ed., 2005): 20 citations.
– Jenicek M. A physician's self-paced guide to critical thinking. Chicago: American Medical Association (AMA Press) (1st. ed., 2006): 9 citations.
3. Books not included in the study

If a book is not included in this article (e.g., it does not appear in Table 2) in spite of appearing to fulfill inclusion criteria, chances are that the number of citations it received was less than 25. Examples of such books are given in Supplementary Table S1.
Supplementary data.

**Supplementary Table S1. Examples of books that received less than 25 citations.***

| Author(s)                   | Brief title                                | Edition, year | Years | Citations (total) | Citations per year |
|----------------------------|--------------------------------------------|---------------|-------|-------------------|--------------------|
| Blalock                    | Causal models in panel and experimental designs | 1st., 1985 | 26    | 23                | 0.9                |
| Wild, Vineis, Garte        | Molecular epidemiology                     | 1st., 2008   | 3     | 22                | 7.3                |
| Prince, et al.             | Practical psychiatric epidemiology         | 1st., 2003   | 8     | 22                | 2.8                |
| Page, Cole, Timmreck       | Basic epidemiological methods              | 1st., 1995   | 16    | 22                | 1.4                |
| Roberts                    | Epidemiology for clinicians               | 1st., 1977   | 34    | 21                | 0.6                |
| Norell                     | Workbook of epidemiology                  | 1st., 1995   | 16    | 18                | 1.1                |
| Ashton                     | Epidemiological imagination               | 1st., 1994   | 17    | 12                | 0.7                |
| Rebeck, Ambrosone, Shields | Molecular epidemiology                    | 1st., 2008   | 3     | 9                 | 3.0                |
| Unwin, Carr, Leeson        | Introductory study guide                   | 1st., 1997   | 14    | 9                 | 0.6                |
| Sydenstricker              | Health and environment                     | 1st., 1933, 2nd., 1972 | 78 | 9 | 0.1 |
| Hudson, et al.             | Clinical preventive medicine               | 1st., 1988   | 23    | 8                 | 0.3                |
| Schuman                    | Practice-based epidemiology                | 1st., 1986   | 25    | 8                 | 0.3                |
| Morrell                    | Epidemiology in general practice           | 1st., 1988   | 23    | 6                 | 0.3                |
| Carrington, Hoelzel        | Molecular epidemiology                     | 1st., 2001   | 10    | 5                 | 0.5                |
| Harkness                   | Epidemiology in nursing practice           | 1st., 1995   | 16    | 3                 | 0.2                |
| Steenland                  | Case studies in occupational epidemiology  | 1st., 1993   | 18    | 2                 | 0.1                |
| Gray, Fowler               | Essentials of preventive medicine          | 1st., 1984   | 27    | 2                 | 0.1                |
| Waning, Montagne           | Pharmacoepidemiology                       | 1st., 2000   | 11    | 1                 | 0.1                |
| Loue                       | Case studies in forensic epidemiology      | 1st., 2002   | 9     | 1                 | 0.1                |
| Winklestein, French, Lane  | Basic readings in epidemiology             | 1st., 1970   | 41    | 1                 | 0.0                |

*Please see book references, and footnotes to Table 2 in the article’s main text.
Cited Reference Search  (Find the articles that cite a person's work)

**Step 2:** Select cited references and click "Finish Search."

Hint: Look for cited reference variants (sometimes different pages of the same article are cited or papers are cited incorrectly).

| Select References | Cited Author | Cited Work [SHOW EXPANDED TITLES] | Year | Volume | Page | Article ID | Citing Articles | View Record |
|-------------------|--------------|-----------------------------------|------|--------|------|------------|----------------|-------------|
| ROTHAM KJ         | MODERN EPIDEMIOLOGY | 1986 | 1 | 1986 | 1 |
| ROTHAN KJ         | MODERN EPIDEMIOLOGY | 1986 | 1 | 1986 | 1 |
| ROTHEMAN KJ       | MODERN EPIDEMIOLOGY | 1986 | 1 | 1986 | 1 |
| ROTHAMER K        | MODERN EPIDEMIOLOGY | 1986 | 1 | 1986 | 1 |
| ROTHHAM KJ        | MODERN EPIDEMIOLOGY | 1986 | 1 | 1986 | 1 |
| ROTHAMK K         | MODERN EPIDEMIOLOGY | 1986 | 1 | 1986 | 1 |
| ROTHMA K          | MODERN EPIDEMIOLOGY | 1986 | 1 | 1986 | 1 |
| ROTHMAM K         | MODERN EPIDEMIOLOGY | 1986 | 1 | 1986 | 1 |
| ROTHMAN K         | MODERN EPIDEMIOLOGY | 1986 | 1 | 1986 | 1 |
| ROTHMANA          | MODERN EPIDEMIOLOGY | 1986 | 1 | 1986 | 1 |
Citations to four books on causality (total, 2,529 citations)

- Pearl. Causality.
- Susser. Causal thinking.
- Rothman. Causal inference.
- Greenland. Evolution of epidemiologic ideas.

Figure S2
Citations to two books by H. Blalock (1,705 citations)
Figure S4

Citations to three books on clinical trials: Pocock (2,421 citations), Friedman et al. (1,498 citations), and Meinert (703 citations)

- Pocock
- Friedman et al.
- Meinert
Citations to 3 books on social epidemiology:
Berkman & Kawachi (1,311 citations), Marmot & Wilkinson (1,214 citations) and Townsend & Davidson (1,364 citations)

Figure S5

- Berkman & Kawachi.
- Marmot & Wilkinson.
- Townsend & Davidson.
Evolution over time of the 2,123 citations received by Austin B. Hill’s ‘Principles of medical statistics’.

Figure S6
Evolution over time of the 475 citations received by Mervyn Susser’s ‘Causal thinking in the health sciences’.

Figure S7
Evolution over time of the 733 citations received by 3 books by Muin J. Khoury
Citations to two books on occupational epidemiology:
Checkoway et al. (619 citations) and Monson et al. (297 citations)

Figure S9
Trends of citations to the book by Hosmer & Lemeshow, and to the article by Mantel & Haenszel.

Figure S10

- Hosmer & Lemeshow. Logistic regression.
- Mantel & Haenszel. Statistical aspects of the analysis of data from retrospective studies of disease (J Natl Cancer Inst. 1959).