This is an author produced version of a paper published in Schizophrenia Research. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the published paper:
Sten Theander, Lennart Wetterberg


Schizophrenia Research,
2009 Jun 2

http://dx.doi.org/10.1016/j.schres.2009.05.013

Access to the published version may require journal subscription.

Published with permission from: Elsevier
Schizophrenia in Medline 1950-2006: A bibliometric investigation

Sten S. Theander a, Lennart Wetterberg b

a Division of Psychiatry, Department of Clinical Sciences, Lund University, Sweden.
b Department of Clinical Neuroscience, Section of Psychiatry at St. Goran’s Hospital, the Karolinska Institute, 112 81 Stockholm, SE-Sweden

Corresponding author:
Professor Lennart Wetterberg, MD, PhD,
Department of Clinical Neuroscience
Karolinska Institutet at St. Goran’s Hospital
SE-112 81 Stockholm
Sweden
E-mail address: lennart.wetterberg@ki.se
Phone/Fax 0046 8 256970

Abstract

The aim was to perform a bibliometric study, and compare the quantity of publications on schizophrenia with the total medical literature in Medline during 57 years, 1950-2006.
The annual additions of literature to Medline are continually increasing and form the Medline growth curve. Comparisons of the numbers of publication on schizophrenia, or any other disease, to this curve, may be used to estimate the research activity.
Methods for the identification of relevant references to papers on schizophrenia were evaluated and three different samples were operationally defined, retrieved and counted.
During 1950-2006, 16.28 million references were added to Medline. Nearly 68 000, 0.42 %, references were related to schizophrenia. The percentage of papers on schizophrenia among the psychiatric literature decreased from 5.2 to 2.6 %. The present study indicates that the percentage of references on schizophrenia in Medline has followed the general increase of medical publications. This pattern differs compared to some other research fields such as dementia, HIV, and peptic ulcer.
Samples of references on schizophrenia may be retrieved in Medline by operational definitions of search methods. The quantity of schizophrenia research during 57 years has kept pace with the total medical literature. One interpretation of the results is that more resources are needed to enhance research activities on schizophrenia.

Keywords: bibliometrics; history; literature; Medline; research; schizophrenia
1. Introduction

Bibliometric methods and results can illustrate the development of research. Medline is a powerful tool for study of the medical literature. The first step in such a study is to design objective and reproducible methods for the retrieval of references. As a prerequisite, we have analyzed the search and retrieval system developed by the National Library of Medicine (NLM). The number of references in Medline is continuously increasing, reflecting the growth of medical research; therefore, any changes in the number of references on schizophrenia, for example, need to be evaluated relative to all references in Medline. An analysis of this kind will show, for any disease, whether the scientific literature on that disorder has increased faster, slower, or with similar pace as Medline. In this study we have analyzed the development of literature on schizophrenia, and some other diseases for comparison, for the 57 years covered by Medline up to year 2006.

1.1 Bibliometric methods in psychiatry

Bibliometrics has been characterized as “a science of science” (de Solla Price, 1986) and “research on research” (Pincus, Henderson, Blackwood et al., 1993). Many psychiatric fields have been studied with the help of bibliometric methods, e.g. the relative size of the literature for diagnostic categories (Sprock and Herrmann, 2000), bipolar disorder (Clement, Singh and Burns, 2003; Lopez-Munoz, Vieta, Rubio et al., 2006), the use of DSM and ICD for classification (Lopez-Munoz, Garcia-Garcia, Saiz-Ruiz et al., 2008), the use of SSRIs for treatment of depression (Lopez-Munoz, Alamo, Rubio et al., 2003), studies of ADHD (Lopez-Munoz, Alamo, Quintero-Gutiérrez et al., 2008), eating disorders (Theander, 2002; Theander, 2004), and schizophrenia (Morlino, Lisanti, Gogliettino et al., 1997). Bibliometrics uses principally two methods, a) counting numbers of articles as quantitative measure (Price, 1951), or b) use of citation analysis in the search for the value or impact of a paper, a journal, or a research group (Garfield, 1955; Lundberg, 2006; Garfield, 2007; Koskinen, Isohanni, Paajala et al., 2008). The present study is based on the quantitative method, counting the number of articles. Bibliometric methods are also used when data are collected as a basis for a review of a specific question as in systematic reviews (Cohrane Foundation, 2009) or evidence mapping (Katz, Williams, Girard et al., 2003).

1.2 The Medline database

The searchable database PubMed at the NLM covers the period from 1950 and onwards. The great majority of all the references in PubMed are indexed in Medline. In April 2009, Medline contained about 17.7 million references. Medline covers mainly medical journals, but also many journals on sociology and psychology. About 5 000 journals are now being indexed. An elaborate system of keywords, the Medical Subject Headings (MeSH), has been developed at the NLM. The tag [mh] after a search word restricts the search to those references where the word is used as a MeSH term. Each reference in Medline contains a list of MeSH terms, usually between 8 and 15, but fewer during the 1950s. The list of MeSH terms summarizes the content. From 1975 an English abstract, when available, is included in the reference. References from the database OldMedline are here included among those from Medline.
Material and methods

2.1 Search methods
We have utilized the Medical Subject Heading system in Medline and its hierarchic tree structure which is based on 16 main categories. Psychiatry and Psychology are indexed in Category F (Table 1), where the tree Mental Disorders (F03) covers psychiatry and F01, F02, and F04 mainly psychological aspects. Category F may be used as a measure of the total psychiatric research. After the indexing of an article, the reference is added to Medline. The sum of all new references during a year creates the annual addition to Medline. Two MeSH terms were used in the search instructions: Schizophrenia and Schizophrenic Psychology. As an additional way to identify references on schizophrenia we have searched for ´schizophrenia´ as a text word, either in the title of the paper or in the abstract. By a search for the truncated word schizophren*, references containing words as schizophrenic (-s), or schizophreniform will also be retrieved. We have regarded references on schizophrenia in childhood as false positives (Piggott and Simson, 1975). Therefore, all search instructions were followed by "NOT schizophrenia, childhood".

2.2 Operational definitions of retrieved samples for schizophrenia
Three operational definitions for retrieved sampling were used to include the literature on schizophrenia in Medline, with the exception of childhood schizophrenia. The three defined samples in the present study will be referred to in the text by their labels “S”, “S+” and “S tw”.

2.2.1 Sample S
Sample S includes all references that are indexed with Schizophrenia. The search instruction was: Schizophrenia[mh]

2.2.2 Sample S+
Sample S+ is made wider by adding references that are indexed with Schizophrenic Psychology, and references not in sample S that have the truncated word schizophren* in the title of the article. The search instruction was (Schizophrenia[mh] OR Schizophren*[ti] OR Schizophrenic Psychology[mh]).

2.2.3 Sample S tw
Sample S tw is made still wider by a search for the truncated word schizophren* anywhere in the text. The search instruction was: schizophren*[tw] In addition to the previous samples, S tw includes references with schizophren* in the text of the abstract.

2.3 Sampling of other research areas in Medline for comparison with schizophrenia
The number of references during 57 years was also sampled for three other research areas; Dementia with a similar severity and costs to society as schizophrenia, human immunodeficiency virus (HIV) as example of a new severe infective agent, and peptic ulcer in which important discoveries have simplified treatment during the studied time period.
3. Results

3.1 Medline and schizophrenia 1950-2006

The number of references indexed in Medline during 1950-2006 is 16.28 million. The annual additions increased from 96 000 (mean of 1950-1952) to 621 900 (mean of 2004-2006), or with a factor of 6.5. The increase pattern of Medline for the entire period 1950-2006 is illustrated in Fig. 1, where the blue line shows how the annual additions have increased, while the green line shows the resulting increase of the total number of references in Medline. The references to papers on schizophrenia, red line, is represented in Fig. 1 by sample S+ (see 2.2.2), comprising 67 893 articles or a mean of 0.42 % of the annual increase in Medline.

3.2 Schizophrenia in relation to psychiatry and psychology in Medline 1950-2006

The number of articles per year 1950-2006 on Psychiatry/Psychology (P/P), Mental Disorders and Schizophrenia are presented in Fig. 2 as percentage of all the references in Medline. Additional results, given as means of each decade, are presented in Table 2. During 1950-1959, the P/P literature comprises 5.6 % of all the references in Medline, compared with 13.2% during the 1970s and 17.7 % during 2000-2006 (Fig. 2). In the period 1950-1961 the frequency is low. From 1962 the percentage of P/P articles is increasing, although with considerable variations from one year to the next. This variation is less pronounced from the year 1970. Among the total P/P literature the percentage of references on Mental Disorders varies around 29 % during the period 1950-2006.

3.3 Literature on schizophrenia in Medline 1950-2006

References to papers on schizophrenia during the period 1950-2006 are for sample S: 61 893 (0.38 %), sample S+: 67 893 (0.42 %), and for Sample S tw: 75 437 (0.46 %). During the 1970s the annual addition of references on schizophrenia, sample S +, was 1 043 and during 2000-2006 2 536: an increase of 2.4. This increase is similar to the increase of Medline during the same period; the figures are 239 907 and 559 514: an increase of 2.3. From the 1960s and onwards the frequency of references on schizophrenia has been about 0.44 % (mean) in sample S+, with a tendency to a trough in the 1980s. The percentage of papers on schizophrenia in Medline during 1950-2006 is shown in Fig.3 and among the psychiatric literature in Table 2. From 5.2 % in the 1950s, a decrease is seen, and from the1980s to 2000-2006 the percentage is about 2.7 %. The number of references on schizophrenia in Medline according to the four MeSH-headings (F01-F04) for a specific year is exemplified for year 2004 in Table 1. It is clear from this table that the schizophrenia diagnosis or associate diagnoses may be found under several headings in Medline.

3.4 Schizophrenia compared to some other research fields in Medline 1950-2006

Examples of the research output of some other diseases relative to Medline for the period 1950-2006 are presented in Fig. 4 as mean of annual additions in the five-year periods. The literature on dementia e g Alzheimer’s disease has increased significantly. In another disease, peptic ulcer, the research output decreased notably some years after 1982. In new fields like HIV, which was discovered in the early 1980’s as a cause of the novel disease AIDS, there has been a sharp increase in research relative to Medline. In comparison with these diseases the literature in schizophrenia closely follows the Medline growth curve (Fig. 4).
Discussion

4.1 Bibliometric use of Medline for evaluating research in schizophrenia

The cause of schizophrenia is unknown. It has an early onset and there is no test to confirm the illness. Schizophrenia is a syndromal diagnosis based upon qualitative appearance of aberrant psychological phenomena which in addition may change markedly over time. Since the symptomatology is variable and the etiology unidentified, research on schizophrenia engages many scientific fields and methods where genetics, environmental factors, cognition, brain imaging and morphology, biochemistry and psychopharmacology, are involved. Research reports on schizophrenia may thus appear in a broad spectrum of scientific journals. Our aim was to study the development of research on schizophrenia in relation to the total medical literature. We have limited the investigation to Medline, a strategy which is an advantage in a quantitative study of this kind, as all comparisons are made between groups of references from the same database. In any bibliometric study it is problematic to identify the relevant references, not least in a heterogeneous syndrome like schizophrenia. In daily clinical work it may be difficult to differentiate between ‘core’ schizophrenia and milder or atypical syndromes, and this problem is also present in the indexing of scientific articles. The choices between suitable MeSH terms cannot be made according to operational definitions. Several MeSH terms describe syndromes that are bordering on schizophrenia, as well as symptoms and traits that are often met in schizophrenia. Usually, MeSH-terms of this kind are indexed together with Schizophrenia [mh], but this is far from a general rule. MeSH terms that are connected with schizophrenia are overlapping each other in many complex ways.

4.2 Search strategies and samples

Different search methods may be used for retrieval of references to papers on schizophrenia. A broad and extensive search will result in many non-relevant references being retrieved (false positives). Such references will “dilute” the sample and give misleading results. A strict and narrow search method, on the other hand, will leave many relevant references unidentified (false negatives). One of our samples, the wider schizophrenia diagnosis "S+", is a result of a middle way between the broad and the narrow methods. Our investigation is entirely based on search methods, defined by specific search instructions. Thus, the groups of references are operationally defined, and the findings possible to replicate. The indexing of papers at the NLM is a great help in retrieving relevant references, but the use of MeSH terms is not without problems. During the 57 years that are covered by the present study, different classifications and diagnoses have been used in psychiatry and such changes will also be reflected in the indexing practice. The indexing of a paper cannot be entirely objective. We have observed a problem in the use of the MeSH term Schizophrenic Psychology (SP). During the studied period, a total of 21 400 references have been indexed with SP. The majority were also indexed with Schizophrenia [mh] and is thus included in sample S. Almost 5 000 references, however, are indexed with SP, but not with Schizophrenia and was missed by a search for Schizophrenia [mh]. These references were included in sample S+. The described problem is most common in references from the 1950s and 1960s. Also a search for schizophrenia as a text word, Sample S tw, entails some problems: As abstracts were not added until 1975, a search for text words has less effect earlier than that year. Also after 1974 the effect will differ among references with and without abstracts. The wider retrieval will doubtless increase the number of false positives. We give some examples when authors might
include schizophren* in the abstract of a paper that is not dealing with schizophrenia: in a study of alcoholics, hallucinations are compared "with those in schizophrenia"; in a study on affective disorders "schizophrenics were excluded"; in an experimental study it is said: "our findings might be of interest in the research on schizophrenia". In experimental studies, laboratory animals are often used. Thus, a study may be considered as experimental when 'Animal' occurs in the abstract. S tw has a larger percentage of Animal compared with S+: 1990-1999 S+: 0.1 %, S tw: 2.3 %. 2000-2006: S+ 2 %, S tw: 4.8 %. S tw has also had a greater increase in recent years compared with S and S+. We have here presented some arguments in favor of sample S+ being the most representative: more complete than Sample S, and probably containing less false positives than sample S tw.

4.3 Previous studies of schizophrenia
One previous quantitative bibliometric study has focused on schizophrenia (Morlino, Lisanti, Gogliettino et al., 1997). The authors presented the percentage of papers on schizophrenia and the direction of research in each of three psychiatric journals during 1980-1994. Some differences between the journals were noted, but hardly any obvious trends over time. The following recommendation was given: "Given the importance of schizophrenia in psychiatric research and practice, it will be useful to regularly monitor the research trends in this special field."

4.4 How representative is Medline?
The new references indexed in Medline during a year reflect the scientific literature in Medicine of that year. A basic question for the present study is to what extent Medline is representative of the total medical literature, and especially of all the literature on psychiatry and psychology. The main factor for a database to be representative is the criteria for the selection of journals for indexing. The NLM is describing the selection in a fact sheet (NLM, 2007), and further information is given in a FAQ paper (NLM 2008). An important statement is: "scientific merit of a journal's content is the primary consideration in selecting journals for indexing". We have some concerns about the representation of psychiatric and psychological literature in Medline during the period 1950-1969. The percentage during the earliest 12 years is low, and is followed by a high although variable increase (Fig. 2).

4.5 Research productivity relative to the Medline growth curve
The publication growth curve that is based on the annual additions to Medline may be regarded as a measure to which other growth curves can be compared. An increase of the percentage signifies that the literature on a disease is increasing faster than Medline. As shown, research in peptic ulcer decreased after the discovery of helicobacter in 1982. In other medical fields the research development has increased relative to Medline: An intense research effort will often arise in novel diseases like AIDS. In comparison with many other diseases research in schizophrenia has followed the Medline growth curve during the studied period.

4.6 Conclusion
The present investigation of publications in Medline offers background information for future studies on schizophrenia. Bibliometric methods have e.g been used in a recent study on early detection and intervention in psychosis (Schultze-Lutter, Ruhrmann, and Klosterkötter, 2008). On the one hand, when the research activity on a disease is keeping pace with Medline, as is
the case with schizophrenia, this may be evaluated as a satisfactory result. On the other hand, schizophrenia is a severe global disease, affecting about 1 percent of the population worldwide and strikes people at an early age. The loss of productivity also adds to the societal cost. Funding has an impact on research output and the priorities should also reflect societal needs (May, Smedby and Wetterberg, 1986). More research resources about the etiology, treatment, prevention and other aspects of schizophrenia are needed; and research development in dementia is an example to follow.
Role of Funding Sources

Funding of this study was provided by Sjöbring Foundation, University of Lund, Sweden. The foundation had no further role in this study.

Contributors

The first author performed most of the data collection from Medline, produced diagrams and wrote the early draft of the manuscript. The corresponding author with extensive experience in schizophrenia research participated in every phase of the study.

Conflict of interest

Both authors declare that they have no conflicting interests.

Acknowledgements

We thank the staff of the University and Medical Faculty libraries, University of Lund, Sweden for advice, Lars Gårding and Paul Kester for fruitful discussions, and Gunilla Theander Kester and Daniel Kester for help with the English language.
References


Theander, S.S., 2002. Literature on eating disorders during 40 years: increasing number of papers, emergence of bulimia nervosa. Eur Eat Disord Rev. 10 386-398.

Table 1. Schizophrenia in Medline describes the four Medical Subject Headings (MeSH)-trees (F01-F04) in the Psychiatry/Psychology Category.\textsuperscript{a} Below each tree some examples are given of MeSH terms that deal with, or are related to schizophrenia. The unique tree numbers are included. Note the hierarchical structure (see tree F03) with new subheadings successively being entered into the system. After each MeSH term the number of references, published in 2004 are presented in \textit{italics}.

**F01  Behavior and Behavior Mechanisms  \( N = 67\,535 \)**

- Catatonia  \([F01.145.126.156]\) \( N = 33 \)
- Delusions  \([F01.145.126.200]\) \( N = 145 \)
- Depersonalization  \([F01.145.126.300]\) \( N = 23 \)
- Paranoid Behavior  \([F01.145.126.962]\) \( N = 4 \)
- Schizophrenic Language  \([F01.145.126.975]\) \( N = 7 \)

**F02  Psychological Phenomena and Processes\textsuperscript{b}  \( N = 42\,920 \)**

- Mental Health  \([F02.418]\) \( N = 623 \)
- Psychological Theory  \([F02.739]\) \( N = 2231 \)
- Mental Processes  \([F02.463]\) \( N = 23\,827 \)

**F03  Mental Disorders  \( N = 29\,320 \)**

- Schizophrenia and Disorders with Psychotic Features  \([F03.700]\) \( N = 3\,243 \)
  - Capgras Syndrome  \([F03.700.300]\) \( N = 5 \)
  - Psychotic Disorders  \([F03.700.675]\) \( N = 850 \)
  - Paranoid Disorders  \([F03.700.450]\) \( N = 30 \)
  - Schizophrenia  \([F03.700.750]\) \( N = 2642 \)
    - Schizophrenia, Catatonic  \([F03, 700, 750, 300]\) \( N = 4 \)
    - Schizophrenia, Disorganized  \([F03, 700, 750, 350]\) \( N = 4 \)
    - Schizophrenia, Paranoid  \([F03, 700, 750, 600]\) \( N = 92 \)
    - Shared Paranoid Disorder  \([F03.700.750.700]\) \( N = 4 \)

**F04  Behavioral Disciplines and Activities  \( N = 21\,267 \)**

- Schizophrenic Psychology\textsuperscript{c}  \([F04.824]\) \( N = 709 \)
- Psychiatric Status Rating Scales  \([F04.586]\) \( N = 2089 \)

\textsuperscript{a} Many of the MeSH terms in this diagram are overlapping. The sum of references from all the F-trees is 161 042, but a search for F01 OR F02 OR F03 OR F04, which search method does not retrieve any duplicates, will retrieve 106 281 references. Thus, 54 761 references are overlapping between the trees. \textsuperscript{b} The group Mental Processes in tree F02 contains a very large number of references. In this group most of the references that deal with psychology are indexed, e.g. papers on cognition, learning, perception, and thinking. \textsuperscript{c} There is of course a large overlap between Schizophrenia (S) and Schizophrenic Psychology (SP), but many references that are indexed with SP are not indexed with S., which may be a source of error in the retrieval.
Table 2. Literature on Schizophrenia, Psychiatry and Psychology (P/P), and Mental Disorders in Medline during 1950 - 2006 (five 10-year periods and one 7 year period.). Schizophrenia is defined as sample S+ (see 2.2.2). P/P-literature is defined as references indexed as belonging to any of the four MeSH-headings F01-F04.

<table>
<thead>
<tr>
<th>Period</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
<th>90-99</th>
<th>00-06</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Annual additions (N) to Medline</strong></td>
<td>103 369</td>
<td>159 772</td>
<td>239 907</td>
<td>318 271</td>
<td>414 800</td>
<td>559 513</td>
</tr>
<tr>
<td>Psychiatry and Psychology</td>
<td>5 810</td>
<td>15 623</td>
<td>31 623</td>
<td>43 056</td>
<td>63 447</td>
<td>99 082</td>
</tr>
<tr>
<td>Mental Disorders</td>
<td>1 629</td>
<td>4 733</td>
<td>8 967</td>
<td>12 744</td>
<td>17 967</td>
<td>27 614</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>300</td>
<td>771</td>
<td>1 043</td>
<td>1 227</td>
<td>1 672</td>
<td>2 536</td>
</tr>
</tbody>
</table>

**Schizophrenia relative to Medline and Mental Disorders**

| Schizophrenia as % of Medline | .29 | .48 | .43 | .39 | .40 | .45 |
| Schizophrenia as % of Mental Disorders | 5.2 | 4.9 | 3.3 | 2.8 | 2.6 | 2.6 |
Fig. 1. Increase of Medline and Schizophrenia during 57 years. Various scales have been used to depict the results in the same diagram.

Units of publications as stated below the diagram

- Annual Additions to M in units of 5,000
- All references in Medline at end of year in units of 100,000
- Sample S+ in units of 50
Fig. 2. Literature on Schizophrenia, Mental Disorders and Psychiatry/Psychology in Medline during 57 years as percentage of all references in Medline. The full length of the bars represents all references in the Psychiatry/Psychology Category. The red parts at the top of the bars represent the literature on schizophrenia with mean of 0.4% of all references in Medline.

% of all references in Medline.
Fig. 3. Literature on schizophrenia (Sample S+) in Medline as percent of all references in Medline 1950-2006.

% x 100
Fig 4. The literature on schizophrenia and some other research fields during 57 years, 1950-2006. Data are given as the mean number of references per year as percent of the annual additions to Medline. The diagram shows 11 five-year periods, plus one two-year period for HIV as an example of a new serious virus infection causing AIDS, peptic ulcer, and dementia.

100 = 1%.