Bibliometric analysis of the literature of randomized controlled trials

By Ming-yueh Tsay, PhD
mytsay@nccu.edu.tw
Professor

Graduate Institute of Library, Information and Archival Studies
National Chengchi University
Taipei, 116
Taiwan, Republic of China

Yen-hsu Yang, MA
luga@tmu.edu.tw
Librarian

Taipei Medical University Library
No. 250, Wu-Hsing Street
Taipei, 110
Taiwan, Republic of China

Objective: Evidence-based medicine (EBM) is a significant issue and the randomized controlled trial (RCT) literature plays a fundamental role in developing EBM. This study investigates the features of RCT literature based on bibliometric methods. Growth of the literature, publication types, languages, publication countries, and research subjects are addressed. The distribution of journal articles was also examined utilizing Bradford's law and Bradford-Zipf's law.

Method: The MEDLINE database was searched for articles indexed under the publication type ''Randomized Control Trial,'' and articles retrieved were counted and analyzed using Microsoft Access, Microsoft Excel, and PERL.

Results: From 1990 to 2001, a total of 114,850 citations dealing with RCTs were retrieved. The literature growth rate, from 1965 to 2001, is steadily rising and follows an exponential model. Journal articles are the predominant form of publication, and the multicenter study is extensively used. English is the most commonly used language.

Conclusions: Generally, RCTs are found in publications concentrating on cardiovascular disease, cancer, asthma, postoperative conditon, health, and anesthetics. Zone analysis and graphical formulation from Bradford's law of scattering shows variations from the standard Bradford model. Forty-two core journals were identified using Bradford's law.

INTRODUCTION

Bradford Hill and others at the Medical Research Council conducted the first randomized clinical trials in 1948, on the use of streptomycin for pulmonary tuberculosis [1]. Since Bradford Hill's pioneering accomplishment, use of the randomized trial methodology has increased, and the number of reported randomized controlled trials (RCTs) has grown exponentially [2]. According to the 2003 edition of the Medical Subject Headings (MeSH), a clinical trial relates to a study involving one or more test treatments, at least one control treatment, determined outcome measures for measuring the studied intervention, and a bias-free method for assigning patients to the test treatment. Such treatments may include drugs, devices, or procedures studied for diagnostic, therapeutic, or prophylactic effectiveness. Control measures include placebos, active medicine, no treatment, dosage forms and regimens, and historical comparisons. When randomization using mathematical methods, such as the use of a random numbers table, is employed to specify patients to test or control treatments to use, the trial is described as an RCT [3]. Therefore, while not all controlled stud-

ies are randomized, all randomized trials are controlled [4].

In his influential book, *Effectiveness and Efficiency: Random Reflections on Health Services,* published in 1972, Cochrane [5] emphasized the significance of using evidence from RCTs because these were likely to offer much more reliable information than other sources of evidence [6]. This concept has become the fundamental basis for what is currently called ''evidence-based medicine'' and has been acclaimed as a paradigm shift in the approach to clinical decision making [1].

The National Library of Medicine (NLM) has been attempting to improve access to information related to clinical trials in MEDLINE. This effort has been underway since 1995 and involves the retrospective assignment of the specific ''Publication Type (PT)'' values of ''Randomized Controlled Trial'' or ''Controlled Clinical Trial'' to MEDLINE journal citations [7].

The volume of medical literature is increasing rapidly, as Price pointed out in 1981 [8]. Thus, an internal medicine physician would need to read about seventeen articles daily, each day of the year to remain a committed physician [9]. Recently, experienced physicians emphasize employing evidence-based literature found in medical journals. However, performing a comprehensive medical literature search every time a clinical question arises is inefficient. To identify and review recently published studies that are likely to affect clinical practice is extremely desirable [10]. Therefore, for physicians to choose a small number of core journals for regular browsing is exceedingly important.

Since Bradford Hill's report on RCTs in 1948, the randomized controlled trial has matured. RCTs are valuable for accomplishing standardization and efficiency in the quality of medicine and are important for implementing evidence-based medicine. RCTs have been extensively studied in recent decades, and the RCT literature has grown rapidly. Thus, studying the important features of the RCT literature is significant, especially the growth pattern, journal productivity, and key concepts.

Bibliometrics is the quantitative study of literature as reflected in bibliographies. This method provides evolutionary models of science, technology, and scholarship [11]. The most prominent model for the distribution of bibliographic items is the Bradford distribution. This model addresses the way a subject's literature is distributed among the journals that contain it. Bradford proposed the concepts of core and scatter. Core refers to the small number of journals that publish the most papers in a field; scatter refers to the spread of literature over many publications.

Applying the computer as an effective means of using bibliometric methods, this study intends to accomplish the following objectives:

1. explore the growth pattern of the RCT literature from 1965 to 2001
2. investigate the publication types of the RCT literature from 1990 to 2001
3. identify the key concepts of the literature that adopted the RCT methodology from 1990 to 2001
4. find the country and language distributions of the RCT literature from 1990 to 2001
5. determine a nucleus of primary journals that contains a substantial proportion of the total RCT journal literature from 1990 to 2001 and investigate the features of these core journals

The starting year covered by MEDLINE, 1965, is chosen as the beginning of the growth pattern study. A significant observation of growth pattern requires long coverage years. However, the coverage dates for the rest of the studies focus on 1990 to 2001. Otherwise, the database would be too large to handle. Characteristics of the RCT literature of the last decade are considered to be more valuable.

LITERATURE REVIEW

Pritchard defines bibliometrics as ''the application of mathematics and statistical methods to books and other media of communication'' [12], which is the most commonly quoted definition and serves as the basis of this investigation. The essential bibliometric principles are briefly explained as follows.

The Bradford distribution is a bibliometric regularity first described in 1934. Bradford proposed the ''law of scattering,'' which states that

> If scientific journals are arranged in order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more predominantly devoted to the subject and several groups or zones containing the same number of articles as the nucleus, when the number of periodicals in the nucleus and succeeding zones will be as $1:n:n^2$. [13]

Bradford's law stresses that only a few sources in a field publish the majority of the literature, while the remainder is published over a vast range of sources consisting of only one or two articles each. Sources in the nucleus (zone one) form the core of the literature and are calculated to verify the most productive journals in any discipline. Bradford's original work was later modified and clarified by numerous studies. Perhaps the most attention has come from Brookes, who combined Bradford's law and Zipf's law and produced a Bradford-Zipf plot. He indicated that, when the cumulative number of articles is plotted against the natural logarithm of the cumulative number of journals, the resulting graph takes the form of an S-shaped curve, the central portion of which is linear. This exhibits the characteristics of three distinct regions: (1) a rapid rise for the first few points, (2) a major portion of linear relation between the two variables, and (3) a ''droop'' at the tail end of the distribution indicating the incompleteness of the bibliography. The few highly productive journals, especially devoted to the subject, account for the short rise [14].

The literature contains only a few bibliometric studies, based on Bradford's law and others, that investigate RCTs in a distinct disease area or in a single jour-

nal. Latronico et al. [15] systematically analyzed the number and quality of RCT reporting published in *Intensive Care Medicine* over 26 years. According to those results, 173 RCTs, 63% of which were from European countries, were analyzed. Kjaergard and Gluud [16] assessed whether trials with positive outcomes were cited more often than trials with negative outcomes. They reviewed 530 randomized clinical trials on hepato-biliary diseases published in 11 English-language journals indexed in MEDLINE from 1985 to 1996, concluding that positive trials were cited significantly more often than negative ones. Wu and Neuhauser [17] cited Balas's works, listing the journals most frequently reporting RCTs of organizational interventions in health care, including *Medical Care, British Medical Journal, American Journal of Public Health, JAMA, New England Journal of Medicine,* and so on.

Kljakovic [18] systematically reviewed cases in 9 general practice journals and 4 general medical journals and found that just over 7% of 10,607 publications involved single cases in both journal groups. Single cases were mainly published as reports or reviews in general practice journals and letters in general medical journals. Two percent of all single cases were published as original research papers in general medical journals, and none were published in general practice journals. Aoki [19] retrieved 4,487 articles with the publication type ''Practice Guideline'' from MEDLINE and analyzed them. The results showed that 108 articles were published in 1991 and 436 in 1992 for a 4-fold increase. Additionally, 55.8% of articles were from the United States and 82% were in English. The most common topics included HIV infection, breast neoplasms, mass screening, asthma, and hypertension.

Pratt [20] utilized MEDLINE to perform a bibliometric analysis of the literature of AIDS for the period of 1981 to 1990. That study reported growth statistics for AIDS literature, number of different languages, countries of publication, and number of periodical titles. The AIDS literature grew from fewer than 700 entries from 1981 to 1983 to a cumulative total of 29,077 entries by the end of 1990. The greatest relative expansion came in 1983 with a 24-fold increase compared to the previous year. Gillaspy and Huber [21] also employed Bradford's law to identify core journal publications for a collection focusing on AIDS in women. That study found that journal scatter for this subset of AIDS literature varied from the scatter in the general literature.

Hasbrouck et al. [22] examined the scientific literature by analyzing citation patterns of specific journal articles to and by the *American Journal of Epidemiology (AJE):* 178,396 journal citations to and 126,478 citations by *AJE* were made from 1983 through 1999. They sorted citations based on the subject category of the referencing or referenced journal. Clinical medical journals accounted for 50.6% of all citations combined (both referenced to and referenced by *AJE*). General and internal medicine (17.9%), cancer (10.4%), and cardiovascular (4.9%) journals had the highest number of citations. Not many citations to and by *AJE* were found in publications specializing in dermatology, gastroenterology, orthopedics, allergy, anesthesiology, surgery, rheumatology, and other areas.

Only a few bibliometric works investigate RCTs in separate disease areas or in single journals. Significantly, no study has been conducted on the RCT literature that attempts to encompass all areas of the health sciences. Therefore, this study is unique in examining RCTs based on a quantitative analysis of the methodology applied in medical research.

METHODOLOGY

In this study, the MEDLINE database, produced by NLM, was selected as the most important database of life sciences, medicine, and allied health. In 2002, MEDLINE included indexing for more than 4,600 journals worldwide (dating from 1966), in more than 40 different languages, and in over 12 million records. Substantive editorials, letters, and biographies are indexed in addition to articles. Moreover, the depth of indexing is a strong feature of MEDLINE as a bibliographic source, because the searcher does not need to rely on the title of an article alone to reveal the article's scope.

In a search conducted in September 2002, the search term ''Randomized Controlled Trial'' was limited to the ''Publication Type'' field to retrieve all pertinent RCT literature from 1965 to 2001. This search statement was developed to retrieve as complete a set of pertinent material as possible.

To study the characteristics of the RCT literature from 1990 to 2001, the five fields of publication type, source, language, publication country, and descriptor were analyzed. The descriptor field is indexed by using NLM's controlled vocabulary, Medical Subject Headings (MeSH). As indicated by Lancaster [23], a controlled vocabulary controls the synonyms, near synonyms, homographs, and related terms. A descriptor with an asterisk is a major descriptor that designates the main subject of the article. Each relevant bibliographic record was downloaded and then analyzed and processed using Microsoft Access database software, Microsoft Excel spreadsheet software, and the PERL programming language.

Careful verification and editing of data were performed as discussed below. Several journal titles referred to the same journal if the name of the journal changed or if compilers of the database cited the journal inconsistently. In such cases, summing up the citations under the more recent or better-known name was appropriate. For example, the *British Journal of Clinical Practice* was used instead of its former title, *International Journal of Clinical Practice.* To determine which journals changed names, the complete list was compared with the list of name changes in the *List of Journals Indexed in Index Medicus* in each edition from 1991 to 2001. A similar situation arose when two journals merged. A citation to merged journals was counted under the new name. For example, the *Journal of Clinical Monitoring and Computing* resulted from the

Table 1
Annual production of randomized controlled trial (RCT) literature in MEDLINE, 1965–2001

Year	RCT articles (A)	RCT cumulative (C)	MEDLINE articles (B)	MEDLINE cumulative (D)	A/B (%)	C/D (%)
1960	0	0	6	6	—	—
1961	0	0	3	9	—	—
1962	0	0	22	31	—	—
1963	0	0	1,191	1,222	—	—
1964	0	0	7,687	8,909	—	—
1965	38	38	100,232	109,141	0.04	0.03
1966	201	239	175,077	284,218	0.11	0.08
1967	241	480	187,121	471,339	0.13	0.10
1968	292	772	204,166	675,505	0.14	0.11
1969	306	1,078	211,335	886,840	0.14	0.12
1970	369	1,447	213,209	1,100,049	0.17	0.13
1971	378	1,825	218,643	1,318,692	0.17	0.14
1972	475	2,300	222,837	1,541,529	0.21	0.15
1973	597	2,897	227,077	1,768,606	0.26	0.16
1974	667	3,564	230,427	1,999,033	0.29	0.18
1975	963	4,527	244,687	2,243,720	0.39	0.20
1976	1,264	5,791	248,967	2,492,687	0.51	0.23
1977	1,311	7,102	255,170	2,747,857	0.51	0.26
1978	1,530	8,632	264,954	3,012,811	0.58	0.29
1979	1,885	10,517	273,579	3,286,390	0.69	0.32
1980	2,169	12,686	271,804	3,558,194	0.80	0.36
1981	2,264	14,950	273,991	3,832,185	0.83	0.39
1982	2,535	17,485	284,902	4,117,087	0.89	0.42
1983	3,037	20,522	298,636	4,415,723	1.02	0.46
1984	2,811	23,333	307,297	4,723,020	0.91	0.49
1985	3,547	26,880	317,419	5,040,439	1.12	0.53
1986	3,859	30,739	329,849	5,370,288	1.17	0.57
1987	4,449	35,188	347,181	5,717,469	1.28	0.62
1988	4,530	39,718	364,285	6,081,754	1.24	0.65
1989	5,645	45,363	379,763	6,461,517	1.49	0.70
1990	6,773	52,136	387,646	6,849,163	1.75	0.76
1991	7,226	59,362	388,683	7,237,846	1.86	0.82
1992	7,640	67,002	390,713	7,628,559	1.96	0.88
1993	8,296	75,298	397,118	8,025,677	2.09	0.94
1994	9,535	84,833	405,937	8,431,614	2.35	1.01
1995	10,454	95,287	416,177	8,847,791	2.51	1.08
1996	10,270	105,318	421,818	9,269,609	2.43	1.14
1997	10,421	115,978	431,678	9,701,287	2.41	1.20
1998	10,793	126,771	445,691	10,146,978	2.42	1.25
1999	11,481	138,252	455,505	10,602,483	2.52	1.30
2000	11,159	149,411	487,349	11,089,832	2.29	1.35
2001	10,802	160,213	505,059	11,594,891	2.14	1.38

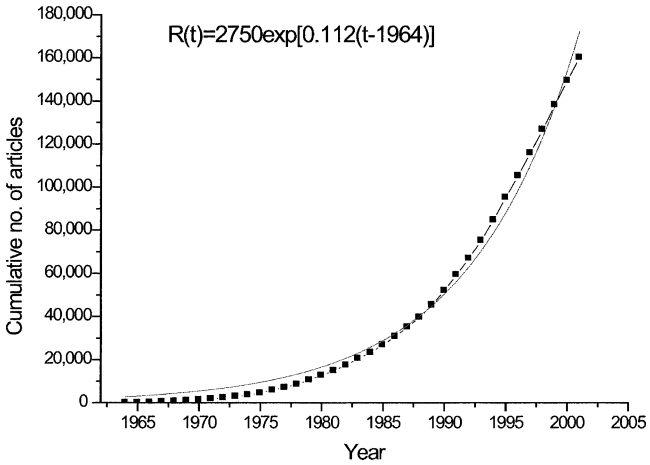
merger of two journals, *International Journal of Clinical Monitoring and Computing* and *Journal of Clinical Monitoring*. A citation to either of them was counted under the new name.

RESULTS

Literature growth

Tague et al. [24] indicated that three possible growth models were available: linear, exponential, and logistic. The authors retrieved 160,213 items on RCTs from MEDLINE from 1965 to 2001. Table 1 summarizes the annual number of articles and their accumulative values, along with the corresponding total items and percentages in MEDLINE. From 1965 to 1974, the number of articles published annually increased by no more than 100 items, except in 1966 and 1973. In 1966, RCT articles jumped from 38 to 201, and, in 1973, 122 more articles were published than in the previous year. The RCT literature grew rapidly after 1974 (667 articles). The number of articles reached 3,000 in 1983. The year 1984 showed a slight reduction (2,811). After 1984, the RCT literature kept rising, especially in 1989, 1990, and 1994, when the number of published articles showed sharp increases. In each of these years, at least 1,000 more articles were published than in the previous years. In Table 1, columns 6 and 7 show the yearly percentage and cumulative percentage of RCT articles in MEDLINE. From 1966 to 2001 (except 1984, 1996, 2000, and 2001), the percentage of RCT articles in MEDLINE each year went up gradually from 0.04% to a maximum of 2.52% in 1999 and fell to 2.14% in 2001. The cumulative percentage of RCT articles also demonstrated increasing growth in MEDLINE.

Figure 1
Cumulative growth of the randomized controlled trial (RCT) literature, 1965–2001



The growth of the RCT literature is also illustrated in Figure 1 to reveal a possible growth model. The figure indicates that the RCT literature grows exponentially. The literature sustains a constant growth rate, and the development has not yet matured. A mature subject would be inclined to exhibit a logistic growth pattern, in other words, the growth rate becomes increasingly smaller and the accumulative literature approaches an asymptotic value as research on that subject is no longer active. The best fit of the curve gives $R(t) = 2750\exp[0.112(t - 1964)]$, with R(t) being the cumulative number of articles at year t. This suggests the yearly growth rate is about 11.2%.

Document type and publication type of journals

From 1990 to 2001, the most common document type is journal article, contributing about 98% of the total. News, letters, editorials, newspaper articles, and conference papers make up the remaining 2%. Letters are the 2nd most common document type.

Citations retrieved from MEDLINE for this study are those in which the publication type included ''Randomized Controlled Trial.'' Among 103,792 journal articles, 90,577 entries provide ''RCT'' as the only publication type, so no further analysis on these articles is possible. The remaining 13,215 items included additional publication type designations. The publication

Table 2
Publication types of journal articles in RCT literature, 1990–2001

Publication type	Articles	%
Multicenter study	15,335	73.63
Phase II clinical trial	1,121	5.38
Phase III clinical trial	1,112	5.34
Review	1,106	5.31
Review, tutorial	827	3.97
Phase I clinical trial	492	2.36
Others (meta-analysis, evaluation studies, review literature, etc.)	835	4.01
Total	*20,828	100.00

* An individual article may belong to one or more publication types; therefore, the total number of publication types is significantly greater than the total items analyzed (13,215). There are 103,792 journal articles in the study interval. However, only 13,215 items have detailed information about publication type.

type distribution for those remaining items is significant and meaningful. Table 2 displays the publication type statistics for these articles. Because an individual article might belong to more than 1 publication type, the total number of publication types (20,828) is significantly greater than the total items (13,215). Table 2 shows that about three-quarters employ the multicenter study (73.63%), indicating that a controlled study conducted by several cooperating institutions is a prominent approach. Phase II and phase III clinical trials are the 2nd and 3rd most applied approaches in these studies. They account for 5.38% and 5.34%, respectively. Phase II clinical trials are preplanned, usually controlled, clinical studies of the safety and efficacy of diagnostic, therapeutic, or prophylactic drugs, devices, or techniques based on several hundred volunteers, including a limited number of patients, and are conducted over roughly 2 years in either the United States or a foreign country. Phase III clinical trials resemble phase II clinical trials, except for the experimental group and observation period. Phase III clinical trials are based on a sufficiently large group of patients and are closely monitored by physicians for adverse response to long-term exposure over a period of about 3 years [25].

Article review types, such as review (5.31%) and tutorial review (3.97%), also have made significant contributions to the RCT literature. Both of them together summarize and critically annotate information on a special subject that has been published over a period of time. If they are compiled with expert knowledge, give adequate coverage of the literature, and are well written, with full bibliographical detail, their value is inestimable, as they can save much time in tracking down and consulting references [26]. Moreover, tutorial reviews often substitute as refresher courses for practitioners who want to update their awareness or as quick courses for students who are unfamiliar with a subject.

The multicenter study accounts for 73.63% of our sample. Phase II clinical trials, phase III clinical trials, and review type literature come next and contribute about 5%, each.

Country and language

Seventy-four countries are represented in the sample, and the United States is the predominant country (as expected because MEDLINE is a US-based database). About 39.9% of the journals and 50.6% of the articles have been published in the United States. England (15.8% of journals and 21.7% of articles) and Germany (6.5% of journals and 6.1% of articles) contribute the 2nd and 3rd most number of articles, followed by Denmark, Switzerland, and the Netherlands, each contributing 2.0% to 4.0% of the total journals and articles. Italy, Canada, Ireland, France, and Norway also significantly contribute to the RCT literature. The involvement of clinical medicine researchers from several countries clearly suggests that the RCT methodology has drawn the attention of medical practitioners and researchers worldwide.

Consistent with the countries of publication, English is the predominant language of articles on RCTs, constituting 92.9% of the total. Only 7% of the articles are not in English. The most common non-English language is German, which constitutes 2.2% of the total. Essentially, English is the only language for RCT literature, possibly owing to the fact that the United States and the United Kingdom are the predominant countries of publication and that MEDLINE is a US-based database. Moreover, English is the official language for most international conferences.

Subject analysis

RCTs offer the best evidence for the efficacy of medical interventions, provided that high standards of transparent reporting are used. This study analyzes the subject aspects of each RCT study as indicated by the major descriptors. There are 99,062 descriptors, including 42,457 major descriptors, assigned to the 114,850 RCT bibliographic records. RCT methods are applied to the major developments in several areas of clinical medicine. Table 3 clearly demonstrates that drug therapy for hypertension, constituting 2,291 articles, is the area that most often employs RCT methods. The medical area with 2nd most use of RCTs is the therapeutic use of anticancer drug combinations (2,140 articles). Drug therapy in asthma comes next, with 1,397 papers published in the form of RCT. Another 3 disciplines that adopt the RCT methodology heavily (more than 1,000 articles) are drug effects for hemodynamics, drug therapy in pain during the period after surgery, and prevention and control of pathologic processes that affect patients after a surgical procedure.

In general, RCTs are found in publications specializing in cardiovascular disease, cancer, asthma, postoperative condition, health, and anesthetics. Cardiovascular disease includes hypertension, myocardial infarction, coronary artery operations, congestive heart failure, and blood lipids. Anticancer drug combinations and breast cancer are the two cancer topics that have adopted the RCT method the most. Exercise, depressive disorder, and quality of life related to health

Table 3
Top 20 subject distribution of RCT literature, 1990–2001

Rank	Major descriptor	Articles
1	Hypertension/dt [Drug Therapy]	2,291
2	Antineoplastic Agents, Combined/tu [Therapeutic Use]	2,140
3	Asthma/dt [Drug Therapy]	1,397
4	Hemodynamics/de [Drug Effects]	1,043
5	Pain, Postoperative/dt [Drug Therapy]	1,012
6	Postoperative Complications/pc [Prevention & Control]	1,000
7	Blood Pressure/de [Drug Effects]	990
8	Antihypertensive Agents/tu [Therapeutic Use]	964
9	Myocardial Infarction/dt [Drug Therapy]	889
10	Exercise/ph [Physiology]	870
11	Coronary Artery Bypass	855
12	Angiotensin-Converting Enzyme Inhibitors/tu [Therapeutic Use]	757
13	Anti-Inflammatory Agents, Non-Steroidal/tu [Therapeutic Use]	756
14	Breast Neoplasms/dt [Drug Therapy]	751
15	Pain, Postoperative/pc [Prevention & Control]	751
16	Depressive Disorder/dt [Drug Therapy]	760
17	Quality of Life	709
18	Heart Failure, Congestive/dt [Drug Therapy]	653
19	Lipids/bl [Blood]	623
20	Anesthetics, Local/ad [Administration & Dosage]	584

also heavily employ the RCT research method. Additionally, prevention and control of pathological processes that affect patients after a surgical procedure and drug therapy for postoperative pain typically use RCT studies. Among the top twenty subjects, almost every area involved drug therapy. This observation agrees with Greenhalgh's report that the RCT study is applied mainly to drug treatments and surgeries [27].

Bradford's law and journal literature

As discussed previously, the journal article is the single most common form of publication. Our sample for 1990 to 2001 includes 2,848 journals. Of these, 481 journals have published only 1 paper that employed an RCT. Table 4 lists the nucleus and the successive zones of journals. Four zones, each publishing approximately 26,000 RCT articles, constitute the most specific subdivisions of these data for which the Bradford hypothesis is valid. The nucleus of journals (zone 1) consists of 42 journals, followed by 105 titles (zones 2), 287 titles (zone 3), and 2,414 titles (zone 4). The ratio of journal number among these 4 zones is 42:105:287:2,414 = 1:2.5:6.8:57.4, which is quite close to 1:2.5:2.5^2(6.25):2.5^3(15.6), except for zone 4, which is much larger than Bradford's law predicts. This may be due to the widespread nature of RCT articles in a vast number of journals.

The 2nd test of Bradford's law plots the cumulative number of articles on an arithmetic scale against the journal rank on a logarithmic scale. This typically produces a graph (Figure 2) with an initial curve followed by a central linear portion and a culminating deviation from linearity known as the Groos droop. Figure 2 illustrates the Bradford-Zipf plot—the cumulative number of papers for each journal against the logarithm of its ranks—for the RCT journal literature. Noticeably, the figure conforms very well to the typical Bradford-Zipf pattern. The approximately linear portion appears after the journal rank of about 42. The top 42 journals may be considered the core journals in the RCT literature. The final droop portion begins approximately at the journal rank of 500.

Hawkins [28] suggests that the droop might be due to the dispersion of the literature on the subject under study. This finding corresponds to the large number of journals in zone 4. Table 4 also lists the cumulative number of articles and the percentage of each zone. The 147 journals in the 1st and 2nd zones cover 50% of literature, and the 42 journals in the 1st zone cover 25% of the literature. This finding implies that, while 50% of the literature is concentrated in only the first 147 journals, the remaining 50% is scattered in 2,701 journals, showing remarkable scattering of RCT literature. The final droop portion reveals that the RCT literature has been extensively spread to many different journals. The scattering of information poses a problem in the complete retrieval of relevant information.

CORE JOURNALS

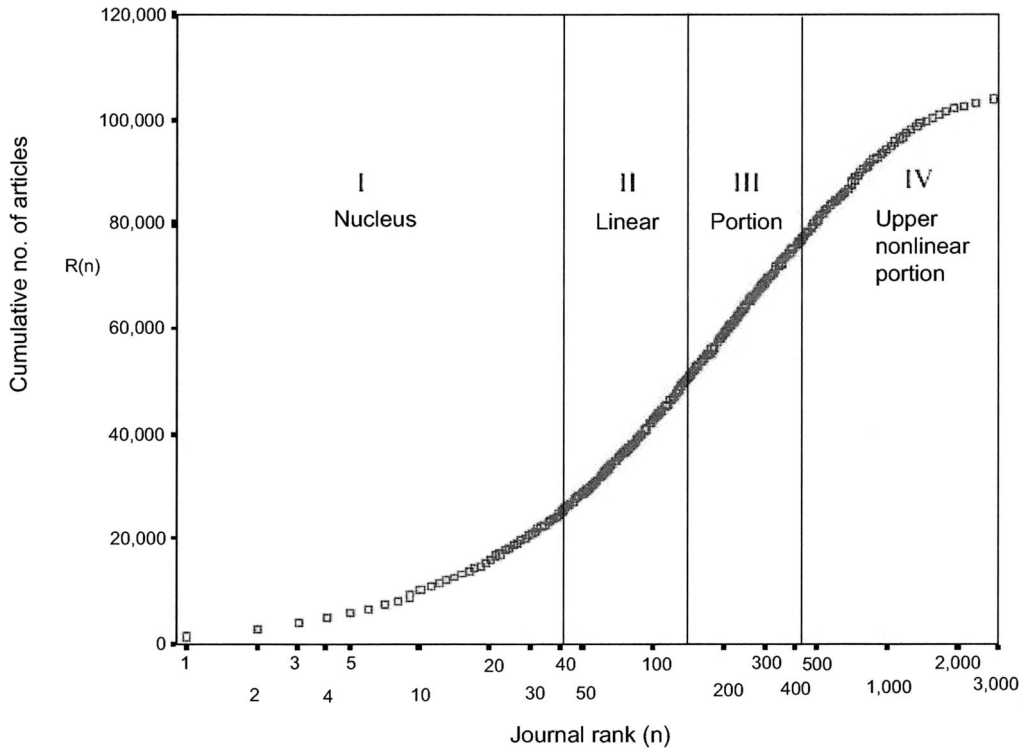
To more closely examine the forty-two core journals, Table 5 lists the number of journal articles in descending order, cumulative articles and percentage, publication frequency, subject field, and ISI's impact factor, taken from *List of Journals Indexed in Index Medicus* [29], *Ulrich's International Periodical Directory* [30], and *Journal Citation Reports* [31]. Among forty-two core journals, twenty-five (indicated by the symbol ''*'' on the journal title in Table 5) were selected by NLM [30] as ''core clinical journals.''

Only one of the forty-two core journals has changed its name. *American Review of Respiratory Disease* changed its name to *American Journal of Respiratory and Critical Care* in 1995. Thirty-two journals are published in the United States. The United Kingdom publishes seven journals. Germany, Denmark, and Canada publish one journal each. All these journals are published

Table 4
Bradford zones of scatter for RCT literature, 1990–2001

Zones	No. of journals	No. of articles	Cumulative no. (%)	Description
1	42	25,899	25,899 (25.0)	Producing > 383 and < 1,384 articles
2	105	25,997	51,896 (50.0)	Producing > 154 and < 379 articles
3	287	25,842	77,738 (74.9)	Producing > 52 and < 153 articles
4	2,414	26,054	103,792 (100.0)	Producing ≥ 1 and < 51 articles
Total	2,848	103,792		

Figure 2
The Bradford-Zipf plot of RCT journal literature



in English, except *Canadian Journal of Anaesthesia,* which is an English-French publication.

In general, the more frequently a journal is published, the more productive it is. For example, eleven of the most productive journals are published weekly or biweekly, while twenty-seven journals are published monthly. Four of the weekly journals are the *New England Journal of Medicine, British Medical Journal, Lancet,* and *Circulation*. Only three journals published fewer than twelve issues per year, and the lowest publication frequency is eight times a year. No quarterly, semiannually, or annually published journals are included.

All core journals are included in Science Citation Index with impact factors from 0.568 to 29.065. The impact factor was defined by Journal Citation Reports (JCR) as the ''number of citations received in year 3 by articles published in years 1 and 2 divided by the number of articles published in years 1 and 2'' [32]. Journals with impact factors greater than 10 are related to general and internal medicine, such as the *New England Journal of Medicine* (29.065), *Lancet* (13.251), *JAMA* (17.569), and *Annals of Internal Medicine* (11.13). Two journals with impact factors less than 1 also appeared in the core. They are *Transplantation Proceedings* (0.568) and *Journal of Clinical Anesthesia* (0.724).

Although the core RCT literature is concentrated in a small number of journals, the journals are quite diversified in their subject coverage. As indicated in Table 5, the forty-two core journals can be categorized into nineteen subjects. RCTs are mainly published in anesthesia journals and pharmacology and pharmacy journals. Each of these two subject areas includes seven journals. Four anesthesia journals—*Anesthesia and Analgesia, British Journal of Anaesthesia, Anesthesiology,* and *Anaesthesia*—are ranked in the top ten core journals. The second subject field that implements RCT methods most is cardiac and cardiovascular systems; six journals belong to this category. The prominent cardiology journals are *American Journal of Cardiology* (no. 3), *Circulation* (no. 7), and *Journal of the American College of Cardiology* (no. 9). General and internal medicine comes third, with five journals.

SUMMARY AND CONCLUSIONS

This study investigated the growth of RCT literature, based on the MEDLINE database, and explored the various features of the literature using well-established bibliometric methods. The results are summarized as follows:

1. The RCT literature from 1965 to 2001 grew exponentially, indicating that the growth of the literature using RCTs maintains a constant rate for the period of the study. The best fit of data reveals that the yearly growth rate is about 11.2%.
2. The single most common form of publication covered in MEDLINE is the journal article, which contributes about 98% of the total RCT literature.
3. Analyzing the publication type demonstrated that, for articles where that level of detail is indicated, the

Table 5
Number of journal articles, publication frequency, impact factor, and subject field of 42 core journals

Rank	Title	No. of articles	Cumulative no. (%)	Frequency	Impact factor	Subject
1	*Anesthesia & Analgesia**	1,384	1,384 (1.3)	Monthly	2.279	Anesthesiology
2	*Lancet**	1,361	2,745 (2.6)	51 times/year	13.251	Medicine, general and internal
3	*American Journal of Cardiology**	1,077	3,822 (3.7)	Semi-monthly	2.637	Cardiac and cardiovascular systems
4	*British Journal of Anaesthesia*	1,018	4,840 (4.7)	Monthly	2.205	Anesthesiology
5	*New England Journal of Medicine**	925	5,765 (5.6)	Weekly	29.065	Medicine, general and internal
6	*Journal of Clinical Oncology*	915	6,680 (6.4)	24 times/year	8.53	Oncology
7	*Circulation**	877	7,557 (7.3)	50 times/year	10.517	Cardiac and cardiovascular systems; hematology
8	*Anesthesiology**	727	8,284 (8.0)	Monthly	3.381	Anesthesiology
9	*Journal of the American College of Cardiology**	725	9,009 (8.7)	14 times/year	6.374	Cardiac and cardiovascular systems
10	*Anaesthesia**	719	9,728 (9.4)	Monthly	2.379	Anesthesiology
11	*American Journal of Clinical Nutrition**	705	10,433 (10.1)	Monthly	5.021	Nutrition and dietetics
12	*British Journal of Clinical Pharmacology*	693	11,126 (10.7)	Monthly	2.213	Pharmacology and pharmacy
13	*European Journal of Clinical Pharmacology*	681	11,807 (11.4)	Monthly	1.922	Pharmacology and pharmacy
14	*Chest**	631	12,438 (12.0)	Monthly	2.48	Cardiac and cardiovascular systems; respiratory system
15	*Canadian Journal of Anaesthesia*	623	13,061 (12.6)	10 times/year	1.143	Anesthesiology
16	*Acta Anaesthesiologica Scandinavica*	622	13,683 (13.2)	11 times/year	1.435	Anesthesiology
17	*BMJ**	610	14,293 (13.8)	weekly	6.629	Medicine, general and internal
18	*Journal of Clinical Endocrinology & Metabolism**	593	14,886 (14.3)	Monthly	5.16	Endocrinology and metabolism
19	*Journal of Clinical Pharmacology*	576	15,462 (14.9)	Monthly	2.167	Pharmacology and pharmacy
20	*Alimentary Pharmacology & Therapeutics*	571	16,033 (15.4)	Monthly	3.9	Gastroenterology; hepatology; pharmacology and pharmacy
21	*Clinical Pharmacology & Therapeutics**	568	16,601 (16.0)	Monthly	5.061	Pharmacology and pharmacy
22	*American Journal of Respiratory & Critical Care Medicine**	557	17,158 (16.5)	Monthly	5.956	Critical care medicine; respiratory system
23	*American Journal of Obstetrics & Gynecology**	524	17,682 (17.0)	Monthly	2.871	Obstetrics and gynecology
24	*Fertility & Sterility*	504	18,186 (17.5)	Monthly	2.96	Obstetrics and gynecology; reproductive biology
25	*Pediatrics**	487	18,673 (18.0)	Monthly	3.708	Pediatrics
26	*Journal of Cardiovascular Pharmacology*	479	19,152 (18.5)	Monthly	1.553	Cardiac and cardiovascular systems; pharmacology and pharmacy
27	*Obstetrics & Gynecology**	470	19,622 (18.9)	Monthly	2.196	Obstetrics and gynecology
28	*Journal of the American Medical Association**	467	20,089 (19.4)	48 times/year	17.569	Medicine, general and internal
29	*Journal of Infectious Diseases**	463	20,552 (19.8)	Semi-monthly	4.91	Infectious diseases
30	*Journal of Pediatrics**	458	21,020 (20.2)	Monthly	3.536	Pediatrics
31	*Diabetes Care*	456	21,466 (20.7)	Monthly	5.404	Endocrinology and metabolism
32	*Antimicrobial Agents & Chemotherapy*	428	21,894 (21.1)	Monthly	4.562	Antibiotics microbiology; pharmacology and pharmacy
33	*Transplantation Proceedings*	413	22,307 (21.5)	8 times/year	0.568	Immunology; surgery
34	*Clinical Therapeutics*	410	22,717 (21.9)	12 times/year	2.721	Pharmacology and pharmacy
35	*Human Reproduction*	406	23,123 (22.3)	Monthly	2.987	Obstetrics and gynecology; reproductive biology
36	*Annals of Internal Medicine**	405	23,528 (22.7)	Semi-monthly	11.13	Medicine, general and internal
37	*Journal of Clinical Anesthesia*	404	23,932 (23.1)	8 times/year	0.724	Anesthesiology
38	*Neurology**	402	24,334 (23.4)	24 times/year	5.212	Clinical neurology
39	*Cancer**	401	24,735 (23.8)	30 times/year	3.909	Oncology
40	*European Heart Journal*	395	25,130 (24.2)	24 times/year	5.153	Cardiac and cardiovascular systems
41	*Critical Care Medicine**	386	25,516 (24.6)	Monthly	3.486	Critical care medicine
42	*American Journal of Gastroenterology*	383	25,899 (25.0)	Monthly	3.549	Gastroenterology and hepatology

* Core clinical journal selected by National Library of Medicine in the Index Medicus.

multicenter study is the most widely employed (73.63%), followed by phase II clinical trials (5.38%) and phase III clinical trials (5.34%).

4. The United States of America is the predominant publishing country in the RCT literature (about 39.9% of the journals and 50.6% of the articles). English is the most common language. English articles constitute 92.9% of the total.

5. The subject areas are diverse and widely dispersed. The areas that employed RCT methods the most include drug therapy for hypertension, therapeutic use of combined antineoplastic agents, and drug therapy in asthma.

6. Forty-two core journals containing 25% of the RCT journal literature can be identified from Bradford zone analysis and the Bradford-Zipf plot. However, the total journal literature is widely spread among many different journals. About half of literature is concentrated in 147 journals, while the remaining half is scattered in 2,701 journals. Moreover, 481 of the journals in the study published only 1 paper.

7. The analysis of the core journals indicates that (1) the more frequently a journal is published, the more productive it is likely to be; (2) most of the core journals deal with anesthesia, pharmacology and pharmacy, cardiac and cardiovascular systems, and general and internal medicine; and (3) journals with impact factors greater than ten are related to general and internal medicine.

The information on literature growth, key concepts, and journal features provided by the present study should be of significant interest for understanding the

development of RCTs and the design of modern information retrieval systems. For example, the study of journal characteristics facilitates establishing a baseline for librarians in making decisions about journal subscriptions and cancellations in the RCT area. The nucleus journals (with a high productivity) identified by Bradford's law usually contain more relevant articles in the area, and, thus, subscriptions to such journals would be worthwhile. Inclusion of these journals in indexing and abstracting services would be justified scientifically. Furthermore, physicians may focus on these core journals as they are usually very busy, and these core journals would cover 25% of the literature on RCTs from 1990 to 2001. Subscribing to journals with low productivity would be unnecessary, and these journals could be discarded. Additionally, core journals can help direct readers to those journals that publish highly relevant articles. From the key concepts and publication types, the intellectual structure and development in the area of RCT research can be examined.

ACKNOWLEDGMENT

The authors thank the National Science Council of Taiwan, Republic of China, for financially supporting this research under contract number NSC91-2413-H-032-009.

REFERENCES

1. Doll R, Peto R, Clarke M. First publication of an individually randomized trial. Control Clin Trials 1999 Aug; 20(4):367–8.
2. Enkin MW. [Foreword]. In: Jadad AR, ed. Randomized controlled trials: a user's guide. London, UK: BMJ Books, 1998:viii.
3. US National Library of Medicine. Medical Subject Headings: 2003 MeSH descriptor data. [Web document]. The Library. [cited 12 Feb 2003]. <http://www.nlm.nih.gov/mesh/>.
4. UK Cochrane Centre. The Cochrane reviewer's handbook glossary, 4.1.4. Oxford, UK: UK Cochrane Centre, 2001:8.
5. Cochrane AL. Effectiveness and efficiency: random reflections on health services (reprint 1972 and introduction by Silagy C). London, UK: Royal Society of Medicine, 1999.
6. The Cochrane Collaboration. The name behind the Cochrane Collaboration. [Web document]. The Collaboration. [cited 28 Oct 2002]. <http://www.cochrane.org/cochrane/archieco.htm>.
7. US National Library of Medicine. FAQ Cochrane collaboration. 2002. [Web document]. Bethesda, MD: The Library. [cited 25 Oct 2002]. <http://www.nlm.nih.gov/services/cochrane.html>.
8. Price DS. The development and structure of the biomedical literature. In: Warren KS, ed. Coping with the biomedical literature: a primer for the scientist and the clinician. New York, NY: Praeger, 1981.
9. Davidoff F, Haynes B, Sackett D, Smith R. Evidence-based medicine: a new journal to help doctors identify the information they need. BMJ 1995 Apr 29;310(6987):1085.
10. Mendelson T. Keeping up with medical literature. In: Friedland DJ, ed. Evidence based medicine: a framework for clinical practice. Stamford, CT: Appleton & Lange, 1998:146–7.
11. White HD, McCain KW. Bibliometrics. Annu Rev Inform Sci Technol 1989;24:119.
12. Pritchard A. Statistical bibliography or bibliometrics. J Doc 1969;25(4):348–9.
13. Bradford SC. Documentation. London, UK: Crosby Lockwood, 1948:154.
14. Brookes BC. Bradford's law and the bibliography of science. Nature 1969 Dec 6;224(223):953–6.
15. Latronico N, Botteri M, Minelli C, Zanotti C, Bertolini G, Candiani A. Quality of reporting of randomised controlled trials in the intensive care literature. a systematic analysis of papers published in Intensive Care Medicine over 26 years. Intensive Care Med 2002 Sep;28(9):1316–23.
16. Kjaergard LL, Gluud C. Citation bias of hepato-biliary randomized clinical trials. J Clin Epidemiol 2002 Apr;55(4): 407–10.
17. Wu GW, Neuhauser D. The literature on quality of life and organizational randomized clinical trials. Med Care 1997 Dec;35(12):1171–2.
18. Kljakovic M. Single cases in general practice and general medical journals. Aust Fam Physician 2002 Jul;31(7):669–73.
19. Aoki M. Analysis of the literature on practice guidelines: a bibliometric study. J Jpn Med Libr Assoc 2002 Mar;49(1): 50–8.
20. Pratt GF. A decade of AIDS literature. Bull Med Libr Assoc 1992 Oct;80(4):380–1.
21. Gillaspy ML, Huber JT. The literature of women and the Acquired Immunodeficiency Syndrome (AIDS): implications for collection development and information retrieval. Med Ref Ser Q 1996 Winter;15(4):21–38.
22. Hasbrouck LM, Taliano J, Hirshon JM, Dannenberg AL. Use of epidemiology in clinical medical publications, 1983–1999: a citation analysis. Am J Epidemiol 2003 Mar 1; 157(5):399–408.
23. Lancaster FW. Vocabulary control for information retrieval. 2nd ed. Arlington, VA: Information Resources Press, 1986.
24. Tague J, Beheshti J, Rees-Potter L. The law of exponential growth: evidence, implications and forecasts. Libr Trends 1981 Summer;30(1):125–49.
25. US National Library of Medicine. Medical Subject Headings: publication types—scope notes—2002. [Web document]. Bethesda, MD: The Library. [cited 14 Apr 2002]. <http://www.nlm.nih.gov/mesh/pubtypes2002.html>.
26. Morton LT, Godbolt S, ed. Information sources in medical sciences. 4th ed. London, UK: Bowker-SAUR, 1992: 61.
27. Greenhalgh T. How to read a paper. getting your bearings (deciding what the paper is about). BMJ 1997 Jul 26; 315(7102):243–6.
28. Hawkins DT. Bibliometrics of the online information retrieval literature. Online Rev 19782(4):345–51.
29. US National Library of Medicine. List of journals indexed in Index Medicus, 2002. [Web document]. Washington, DC: The Library, 2002. [9 Jul 2002]. <http://www.nlm.nih.gov/tsd/serials/lji.html>.
30. Ulrich's international periodicals directory. New York, NY: Bowker, 2002.
31. Journal citation reports. [Web document]. Philadelphia, PA: Thomson ISI, 2002. [cited 8 Mar 2002]. <http://isi3.isiknowledge.com/portal.cgi/jcr/>.
32. Lancaster FW. Bibliometric methods in assessing productivity and impact of research. Bangalore, India: SARADA Ranganathan Endowment for Library Science, 1991:20.

Received June 2003; accepted November 2004