Patents as indicators of corporate technological strength *

Francis NARIN and Elliot NOMA
CHI Research/Computer Horizons, Inc., Haddon Heights, NJ 08035, U.S.A.

Ross PERRY
Ross Perry Associates, Toronto, Ontario, Canada M5E 1V3

Quantitative indicators of the technological strengths of individual companies would be an important addition to the financial and economic data used in competitor assessments, merger/acquisition analyses, investment decisions, and corporate planning and management. This paper examines the links between corporate patent and patent citation data, and several other indicators of corporate performance: changes in sales and profits, research and development budgets, scientific productivity, and expert opinions of company technological strength. The study covers 17 US pharmaceutical companies for which financial, R&D and expert opinion data were readily available. For these pharmaceutical companies it was found that the patent data are an excellent indicator of overall corporate technological strength with (1) an overall correlation of 0.82 between expert opinion of pharmaceutical company technical strength, and the number of US patents granted to the companies, and (2) correlations, in the general range of 0.6 to 0.9, between increases in company profits and sales, and both patent citation frequency and concentration of company patents within a few patent classes.

1. Introduction

1.1. Science, technology and economic indicators

We are most pleased to dedicate this paper to the memory of a friend and valued colleague, Yvan Fabian. Yvan was a major force in bringing quantitative techniques, especially those of R&D input indicators, into the regular statistics of OECD: at the time of his passing he was working actively to bring the use of technology output indicators to the same level of attention at OECD.

In the last decade there has been nearly an exponential increase in the use of quantitative, "science indicators" techniques to measure and assess scientific activity. This utilization of science indicators has come relatively swiftly for two reasons: (1) the acceptance of the research paper as a prime output of science, and (2) the accessibility of large, computerized data banks of the scientific literature.

For example, the *Science Indicators* reports [12] are now a well accepted and heavily used compendia of objective data summarizing national scientific performance. Program managers in US scientific agencies such as the National Institutes of Health, [10] and the National Science Foundation, [9] use publication and citation data in the planning and evaluation of their research programs. Moreover, in the United Kingdom, the Science Policy Research Unit [16] has taken a leading position in the use of indicators techniques. Publication and citation data are also widely, if not publicly, used in the assessment of university performance, in assessing individual scientist performance, and in many other aspects of the assessment and management of science. Finally, OECD recently had a workshop [13] on the use of science and technology indicators in the higher education sector.

Despite the widespread use of quantitative assessment techniques for science, equivalent techniques for the assessment of technological performance are not nearly as well developed. Two factors have been critical in delaying technological performance methods: (1) until very recently, sui-

* Based on work performed under US National Science Foundation SBIR Grant No. SES-8460698. The opinions expressed herein are those of the authors, and not of the National Science Foundation.

Research Policy 16 (1987) 143–155
North-Holland

table databases of patent data did not exist, (2) patents measure technical productivity somewhat less directly than papers measure scientific productivity.

Clearly, however, there is need for measures of corporate technical productivity and strength. Quantitative indicators of company technological strengths would be of enormous value to the corporate and financial communities for competitor assessments, for merger/acquisition analysis, and in many other areas of corporate planning and management. Today, such indicators are almost non-existent.

Financial and economic data, of course, track the past performance and current status of a company. These data may also predict future corporate performance, since a successful company will probably continue to be so in the future. Past evidence of good management, quality R&D, effective sales force, and a well-run distribution network should lead to the expectation of continued good performance in the future. These data, however, are linked to past R&D performance, while future financial performance in normal times is probably dependent on current technological strength [15].

For many companies the fast pace of technological change means that long-term success depends on the continued introduction of new products, and on the continued application of new production methods. These pre-conditions of continued corporate success, however, do not appear in the standard indicators of growth such as market share or profit. (Though their effect may appear in the company's current market value [7].) In fact, they may even appear to have a negative short-term influence, since they are resource intensive and may decrease short-term profits.

Identifying the specific ideas leading to a successful product can only be done in hindsight. Even then, the large number of influences and the long time delays may make it impossible to determine all of the contributing factors. However, we would like to test if an evaluation of the overall direction, breadth, and quality of a company's research program could predict the potential for long-term corporate health. This paper explores the use of patent and patent citation data as a direct measure of company technological strength by comparing patent data with other measures of company performance such as R&D budgets, peer opinions of company strength, research paper production, and increases in company sales and profits.

Patenting seems to have a well established place in the product life cycle within a corporation: Corporate sales lead to corporate profits, which may be used in research and development to produce scientific innovations (scientific publications), which may lead to technological innovations (some of which will be patented), which is the catalyst for new products and more efficient processes, which will increase corporate profits. In this model, measures of inputs and outputs at each part of the cycle should predict (with some time lag) the activity levels on the next part of the cycle.

Unfortunately, the assumptions behind this model are not fully consistent with the empirical findings. In particular, patent counts are not neatly placed between R&D inputs and new products. First, increased commitments to development do not precede increased patenting, but rather are simultaneous with increased patenting [1]. This suggests that the number of patents is a better indicator of corporate commitment to pursue innovation than the actual amount of innovation. Second, the number of patents is more highly correlated with the inputs to development (e.g. research personnel) than to the rate of new product introduction [5]. Both findings may be explained by noting that the quality of inventions varies greatly and that patent counts are unable to distinguish amongst different corporate patenting strategies. Therefore, in this paper, we look both at the number of patents and their quality as measured by the number of citations each has received.

1.2. An overview of patent citation analysis

In patent citation analysis, as in literature citation analysis, there is an underlying assumption that a patent which is highly cited – that is, is referred to by many subsequently issued patents – contains a technological advance of particular merit. The first quantitative study assessing citation rates to technologically important patents showed that patents associated with the 1970 Industrial Research and Development Awards (IR-100 Awards), given for outstanding new products by *Industrial Research* magazine, were 2.5 times as frequently cited as randomly chosen control patents from the same time frame [4]. Furthermore, much of the difference in citation rate was in the presence of a relatively small number of highly cited patents in the IR-100 set, and no highly cited patents amongst the control set of randomly chosen patents. That is, patents associated with outstanding new products were rather highly cited.

In a later validation study of patent citations as indicators of science and foreign dependence, relatively high correlations were found between peer and citation rankings of patent subclasses, in terms of foreign dependence and science dependence [3]. Specifically, a group of senior R&D managers, vice presidents of research, and so forth, were asked to rank 24 patent subclasses in terms of whether they felt that the subclasses were foreign dependent, and whether they felt the subclasses were science dependent. These peer rankings were then compared with bibliometric measures of these dependencies – specifically, the number of citations to the scientific literature as a measure of science dependence, and the number of citations to foreign origin material as a measure of foreign dependence. For both cases a high degree of association was found between the peer rankings and the bibliometric measures.

There is another interesting and important aspect of patent citation analysis which is also beginning to be quantified: namely, the fact that citations to the scientific literature, from patents, provide an indicator of linkage between technology and science. The first study of the strength of the linkage between basic research literature and patents showed that patents in highly scientific areas of technology contain as many citations to the scientific literature as do research papers [2]. A very recent study showed that the time lag for citations from patents to the literature is as short in some biotechnology areas as the time lag for cites from papers to papers [11], indicating that biotechnology and bioscience are contemporaneous.

1.3. Comparisons of scientific, technological and financial variables

When applying quantitative techniques to individual companies, the availability and accuracy of the data are major problems. When available, the corporate R&D budget and number of R&D personnel are probably the best indicators of overall corporate commitment to new ideas and technologies.

One of the visible output indicators of R&D productivity is the number and quality of scientific papers produced. By counting and classifying papers one can gain some insight into the main lines of company research and assess the quality of the laboratories which produce the ideas. For industries in which publication is the norm, such as pharmaceuticals, papers can be a valuable indicator.

In this paper various indicators of company research activity are compared. With the use of retrospective data it is shown that they correlate with overall company performance for certain technologically based companies. In particular, data obtained from US patents and citations to US patents were compared with other indicators of corporate performance; peer judgement of research performance, literature based indicators of research publication, and corporate financial performance data. For the particular group of companies evaluated, 17 primarily US based pharmaceutical companies listed in table 1, it is shown that the number and citation quality of US patents are strongly related, respectively, to peer judgments of research performance and increases in company profits and sales.

2. Data

2.1. Technical performance data

This study builds upon an earlier study by Koenig [8]. Koenig examined pharmaceutical research from a bibliometric perspective, comparing publication rates and citation rates for US pharmaceutical companies with various measures of successful pharmaceutical research including measures of new drug production, peer rankings of the pharmaceutical companies, and research budgets. In general, Koenig found modest (0.3–0.6) but statistically significant correlations between bibliometric measures of pharmaceutical company research and the other measures of research productivity. Koenig did not look at patent or patent citation data, nor did he report on the relationship between pharmaceutical company sales and profits, and the other measures of research perfor-

Table 1
Selected 1975 to 1983 sales and financial characteristics of pharmaceutical companies (companies ranked in descending order of drug sales as a percent of total corporate sales in 1979)

Company	World rank [a] (1975)	1979 Drug sales [b] ($mm)	1979 Drug sales% total sales	Annual % change net profits
1. *Drug-dependent (over 60% of sales are drugs)*				
Merck & Co.	4	2004	84	8.4
Syntex Corp.	33	n.a.	69 (1977)	6.4
SmithKline Corp.	31	862	64	38.0
Upjohn Co.	15	956	63	6.4
2. *Drug-dominant (between 40% and 60% of sales are drugs)*				
Sterling Drug Inc.	20	768	58	8.9
Schering-Plough Corp.	18	757	53	2.1
Pfizer Inc.	7	1430	52	13.8
G.D. Searle & Co.	29	n.a. [c]	51 (1977)	32.0
E.R. Squibb & Sons Inc.	16	900	50	5.8
Abbott Labs	28	830	49	19.6
Eli Lilly & Co.	12	1003	45	13.5
Hoffman La Roche Inc.	2	1374	44	n.a.
American Home Products Corp.	6	1448	43	12.9
3. *Multi-product (less than 40% of sales are drugs)*				
Bristol-Myers Co.	14	946	34	14.9
Warner-Lambert Co.	10	1045	32	n.m. [d]
American Cyanamid	27	n.a.	20 (1977)	n.a.
Johnson & Johnson	41	760	18 (1977)	16.2

[a] World rank is calculated by the percentage of total world pharmaceutical sales in 1975.
[b] 1979 pharmaceutical sales are in millions of current US dollars.
[c] n.a. = not available.
[d] n.m. = not meaningful.
Sources: S. Pradhan, *International Marketing of Pharmaceuticals* (1983); *Chemical Age*, 23 July 1976; OECD, *An Industry Like No Other* (1982); Standard and Poor's, *Industry Surveys*, various issues.

mance. Thus, for the study reported herein, a group of U.S. pharmaceutical companies was studied because:

- independent measures of the scientific excellence of these firms were available in the work of Koenig,
- reasonably complete company R&D figures, and company sales and profits data, are available,
- this industry has a high propensity to patent and relies on patent protection more than most other industries, and
- patents have a readily identifiable position in the chain of new drug development: between basic research and marketable drugs.

The data used in this analysis can be divided into the following six categories:
(1) scientific publishing data,
(2) drug production data,
(3) research spending data,
(4) peer ratings of quality of company research,
(5) corporate patenting data,
(6) financial performance data.

Sixteen of Koenigs variables, covering the first four of these categories, were used in this analysis and are listed in table 2. Seventeen of the larger patenting companies of the 24 companies analyzed by Koenig were used in the analysis.

For these 17 companies, information on 12 additional variables was collected from the CHI Research/Computer Horizons, Inc. (CHI) databases on patent citations (PATENT CITATION INDICATORS Database), and biomedical publication (BID-MEDLINE Database). These 12 additional variables are listed in table 3.

It must be noted here that various inconsistencies arose because of the different ways in which publication, patent, and corporate financial data

Table 2
Koenig's indicators of scientific research by company

1 K ARTS	*Number of scientific articles published by researchers affiliated with the selected pharmaceutical companies, 1970–74*: these articles were found in the Corporate Index to the *Science Citation Index*.
2 STAR	*Number of star articles published by these same researchers, 1970–74*: articles with five or more citations received in the third year after publication.
3 K CITES	*Total citations received by identified scientific articles, 1973–77*: all citations in the third year after publication were summed.
4 K C/ART	*Mean number of citations per article, 1973–77.*
5 STAR C	*Number of star articles in clinical medical journals, 1973–77*: these included articles with five or more citations received in the third year after publication.
6 % STAR C	*Percentage of company clinical articles that were stars, 1973–77.*
7 % STAR	*Percentage of all company articles that were starts, 1973–77.*
8 BUDGET	*Average yearly research budget, 1965–78*; expresseds in millions of current dollars.
9 DRUGS	*Total number of new drug applications approved by the Food and Drug Administration, 1965–76.*
10 ITG	*Total number of new drug applications approved by the Food and Drug Administration and determined as "important therapeutic gains", 1965–76*; by FDA definition, these drugs are the most important advances and are one of the best indicators of quality in research output.
11 SCORE	*Production score*: a weighted measure of all qualifying drug output including new chemical entities, important therapeutic gains and non-ITG, with and without patent protection, 1965–76. Koenig calculated production score as follows: 10 (ITGP) + 2.5 (ITGN) + 4 (NDAP) + 1 (NDAN) with the following definitions of the corporate terms:
ITGP	Important therapeutic gain drug with patent protection.
ITGN	Important therapeutic gain without patent protection.
NDAP	Non-important therapeutic gain drug with patent protection.
NDAN	Non-important therapeutic gain drug without patent protection.
Koenig canvassed the members of the National Institutes of health (NIH) expert advisory panel on pharmacology on their views of corporate research (at the time of canvassing, experts were not shown the publication and other data collected within his study). Four criteria were used:	
12 CREATIV	"Creativity and innovativeness" (*creativity*) in the early 1980s.
13 CONTRIB	"Overall contribution to medical well being" (*contribution*) in the early 1980s.
14 COMMERC	"Commercial effectiveness in capitalizing upon pharmaceutical research" (*commercialization*) in the early 1980s.
15 BASIC R	"Success in pursuing basic biomedical research" (*basic research*) in the early 1980s.
16 EXPOP	A composite of variables 12 to 15 called *expert opinion* was calculated by summing the scores of creativity, contribution, commercialization and basic research.

are presented. Ayerst and Wyeth, for instance, are two of the pharmaceutical companies considered in Koenig's paper. With the scientific literature, affiliations typically identify authors as uniquely belonging to either Ayerst or Wyeth Laboratories. However, all financial data and patent data for Ayerst and Wyeth are carried under the parent company name American Home Products, so that for purposes of this research Ayerst and Wyeth were combined. Similarly, other subsidiaries of the 17 pharmaceutical companies were included, of necessity, in the patent and financial data.

To eliminate the confounding effects of the non-pharmaceutical components of the 17 companies, only their research publications in biomedicine and their patents in pharmaceutically related fields were included in the analysis.

2.2. *BID-MEDLINE publication database*

The publication and citation data for the papers of the 17 companies were taken from the BID-MEDLINE database, maintained by CHI for the National Institutes of Health. This databases consists of, essentially, all papers in a central core of approximately 240 highly influential biomedical

Table 3
Twelve patent and BID-MEDLINE variables

1	C PAPERS	*Number of clinical medical papers, 1973–80*; Koenig's period of examination (1970–74) was thereby extended.
2	B PAPERS	*Number of biomedical research papers, 1973–80.*
3	TDEC C	*Number of clinical medical papers which achieved the top decile in terms of citations, 1973–78*; top decile performance is defined as the fraction of a set of papers that are amongst the most highly cited 10 percent of papers in a field; this statistic focuses on the very important papers.
4	TDEC B	*Number of biomedical research papers in the top decile, 1973–78.*
5	PATS	*Number of company pharmaceutical patents, 1975–82.*
6	PCITES	*Total number of citations to company pharmaceutical patents, 1975–83.*
7	PATSCORE	*Standard patent citation score developed from the company patent profile*; this is a measure of deviation from expert citation performance – high scores indicate comparative technological strength.
8	C/PAT	*Mean number of citations per patent, 1975–82.*
9	TOTPUB	*Total number of publications (clinical and biomedical), 1973–80.*
10	PLIT	*Total number of patent citation to the scientific literature, 1975–82.*
11	LC/PAT	Mean number of literature citations per patent, 1975–83.
12	GINIPC	*Gini coefficient for main patent classes for each company, 1975–82*; this measures the degree of concentration of patenting activity by a company in selected pharmaceutical patent classes.

journals [6]. These journals cover 60–70 percent of all of the research supported by the US National Institute of Health and contain approximately 70,000 papers per year. The papers by pharmaceutical company scientists in this database were identified from the corporate author data on the *Science Citation Index* corporate tapes.

The BID-MEDLINE database also contains citation data including counts of essentially all citations received from the 240 biomedical journals during the period 1973 to 1980. Note that this means that earlier years will have a longer citation history: 1973 has eight years of citation data, 1974 has seven years, etc. Since the one or two times a 1979 paper is cited by 1980 papers is not statistically meaningful, citation counts to papers published in the latest years are omitted from the summary statistics. This insures that the citation statistics, both number of cites per paper, number of papers in the "top decile", and so forth, in the publication data are based typically on four or five years of citations.

2.3. CHI's PATENT CITATION INDICATORS database

For the patent data there were similar confounding problems, and a decision had to be made about which patents to include in the patent analysis. CHI's PATENT CITATION INDICATORS database contains essentially a million US patents

Table 4
US patent classes used for identification of pharmaceutical patents

Class number	Official designation
260	Chemistry, carbon compounds
424	Drug, bio-affecting & body treatment
435	Chemistry, molecular biology & microbiology
436	Chemistry, analytical & immunological testing
536	Organic compounds: carbohydrates
542	Organic compounds: intercyclic acyclic-CH-containing
544	Organic compounds: 6-membered hetero ring, at least one nitrogen
546	Organic compounds: 6-membered hetero ring, one nitrogen & 5 carbons
548	Organic compounds: 5- membered hetero ring, at least one nitrogen
549	Organic compounds: hetero ring with sulfur
556	Organic compounds: silicon containing
560	Organic compounds: esters
562	Organic compounds: radical chalcogen-H acids
564	Organic compounds: amino nitrogen containing
568	Organic compounds: boron, phosphorus, sulfur, oxygen, containing

issued since 1971. For the years 1975 onward the database contains all US patents issued, approximately 60–70,000 patents per year, spread across 335 major US Patent Office classes. To make the patent data as comparable as possible to the literature data, it was decided to restrict the analysis to patents for the formulas of medical substances and drugs used in the treatment and prevention of disease. As an approximation to this category, 15 US Patent Office classes, in which the pharmaceutical companies patent heavily, were identified. These classes are listed in table 4. The percentage of company patents in those 15 classes, out of the total number of patents assigned to the company, appears in table 5, along with counts of the number of papers, and the number of patents. Note, however, that because the US Patent Office classification is an art classification, and the chemical classes are based on chemical structure, there are many patents for chemicals which are not pharmaceutically related in those 15 classes. It is unlikely, however, that very many of the pharmaceutically related patents would not be within those 15 classes.

From table 5 we see that most pharmaceutically intensive companies have 90 percent or more of their patents in the selected 15 classes. By contrast a company such as Johnson & Johnson has most of its patents outside of the selected classes, thereby supporting the classification of Johnson & Johnson as a multi-product company.

Table 5
Seventeen pharmaceutical company names'[a] for publications and patents, with basic data

	# PUBS 1973–80	# PATS 1975–82	Patent names	Publication names	% PATS in selected classes
1	151	810	Squibb Corp.	Squibb Corp.	96
2	157	592	SmithKline	SmithKline	95
			SmithKline & French	SmithKline & French	
			Allergan Pharmaceuticals, Inc.		
			Menley & James Lab.		
			Norden Labs.		
3	85	225	Schering-Plough Corp.	Schering-Plough Corp.	94
4	69	108	Mead Johnson (subs. of Bristol Meyer)	Mead Johnson	94
5	81	281	Bristol Meyer	Bristol Meyer	
			Westwood Pharmaceuticals	Westwood Pharmaceuticals	92
6	398	1125	Merck & Co.	Merck & Co.	91
7	122	625	Pfizer, Inc.	Pfizer, Inc.	90
8	1047	1171	F. Hoffmann La Roche	Hoffmann La Roche	89
			Roche Inst.	Roche Inst.	
9	381	807	Eli Lilly & Co.	Eli Lilly & Co.	
				Greenfield Labs	89
10	123	235	G.D. Searle & Co.	G.D. Searle & Co.	89
11	456	1609	UpJohn Corp.	UpJohn Corp.	88
12	135	494	American Home Products	American Home Products	86
				Ayerst Labs	
				Wyeth Labs	
13	125	328	Syntex (USA)	Syntex Research	86
14	81	368	Sterling Drug Corp.	Sterling Drug Corp.	78
15	200	362	Warner Lambert	Warner Lambert	
			Parke-Davis	Parke-Davis	75
16	139	349	Abbott Laboratories	Abbott Laboratories	69
17	82	899	American Cyanamid	American Cyanamid	59
18	71	218	Johnson & Johnson	Johnson & Johnson	40
			McNeil Labs	McNeil Labs	
			McNeillab	Ortho Pharm Corp.	
			Ortho Pharm Corp.		

[a] Minor variants, e.g., SKF, SK&F and subsidiary institutes are included automatically.

3. Patents as measures of corporate technological strength

To determine the nature of the relationships between scientific research, expert judgment, drug registrations, and patent activities, Pearson product moment correlations were computed between all variables in tables 2 and 3. The correlations are summarized in table 6.

Before discussing the results, a number of important analytical issues inherent in this approach should be mentioned. First, there is the question of "synchronicity and lag" [8, p. 24], when can one activity such as publishing during one time period be meaningfully compared against another activity such as patenting, measured during a slightly different time period? Our, and Koenig's thesis is that the quantity and quality of corporate research at all levels changes relatively slowly over time. Therefore, it is sufficient to get measures that summarize the productivity at each level of product development, without being unduly concerned if the time periods for each measure are slightly different, or that the measures are not tracking a single idea through to fruition. Further, since patents, papers and sales all report specific points on an integrated continuum of R&D, there is always an inherent indeterminacy as to the actual timing of specific events.

Secondly, the limited number of observations, both in numbers of companies and numbers of years limits the statistical power of the correlational methods, but will not obscure the more powerful of the relations found.

Third, although the companies examined in this report account for the majority of US sales of pharmaceutical products, the group did not include two types of firms of growing interest and importance: non-US firms and the emerging small biotechnology firms such as Cetus and Genentech. Koenig omitted these types of firms from this study, because they were not important in the late 1960s and early 1970s.

The correlations in table 6 show the following:

- On a company-by-company basis the total number of patents "PATS" was relatively highly correlated, in the range 0.6–0.8, with the percent of highly cited clinical articles, with the number of new drugs registered, important new drugs approved, the composite drug output score, "creativity", "scientific contribution", "basic research excellence", the composite of expert opinion, and research papers. This shows that patent counts can be considered, a distillation of many aspects of research and development. These findings are also consistent with other reports of the correlation between R&D expenditures and patenting activity within a technological field [7].
- Koenig's peer ranking of the companies, EXPOP is highly correlated with number of patents (PATS; 0.82) and the number of patent citations to the scientific literature (PLIT; 0.84)

Table 6
Correlations between Koenig's data and patent data [a] ($N = 17$)

	PATS	TDEC C	TDEC B	PCITES	PATSCORE	PLIT	EXPOP
K ARTS	0.685**	0.434	0.523*	0.558*	−0.352	0.712**	0.740**
STARS	0.597*	0.440	0.624**	0.513*	−0.274	0.635**	0.693**
K CITES	0.598*	0.464	0.630**	0.514*	−0.279	0.633**	0.685**
STAR C	0.718**	0.491*	0.511*	0.577*	−0.321	0.766**	0.439
BUDGET	0.529*	0.541*	0.050	0.419	−0.214	0.665**	0.601*
DRUGS	0.640**	0.420	0.220	0.472	−0.288	0.633**	0.487*
ITG	0.573*	0.411	0.190	0.410	−0.295	0.532*	0.394
SCORE	0.606**	0.379	0.066	0.374	−0.320	0.556*	0.423
CREATIV	0.800**	0.374	0.339	0.763**	−0.046	0.825**	0.985**
CONTRIB	0.780**	0.339	0.291	0.810**	0.066	0.804**	0.975**
COMMERC	0.344	0.306	0.343	0.476	0.192	0.460	0.703**
BASIC R	0.813**	0.321	0.363	0.690**	−0.073	0.854**	0.964**
EXPOP	0.822**	0.346	0.347	0.777**	−0.038	0.837**	–
% STAR C	0.351	0.548*	0.476	0.417	0.076	0.505*	0.761**
K C/ART	0.360	0.501*	0.605**	0.450	0.028	0.466	0.685**

[a] * $P < 0.05$; ** $P < 0.01$.

and total number of citations to the company patents (PCITES; 0.78). Since both PLIT and PCITES increase as the number of patents grows, all three variables are highly correlated. The most reasonable interpretation is that experts are in actuality rating the *size* of company's R&D output.

- EXPOP has lower correlations with literature variables than with patent variables, and the partial correlation of the literature variables and EXPOP controlling for the patent variables is not significant.

This argues that both expert opinion and patent variables are measuring the outputs of the R&D process, while publications can be reviewed as intermediate measures of activity within the process.

In summary, patent counts, counts of patent citations, and scientific references in patents are highly correlated with expert opinion of corporate excellence in science and technology.

4. Patents as predictors of financial performance

The next step was the expansion of Koenig's analysis to assess the linkage between company financial performance and the underlying scientific and patent performance data. For 16 of the companies in the previous analysis, 13 measures of patenting and publishing data were correlated with six measures of financial performance. The six measures are defined in table 7, and, essentially, measure increases in company profits, sales, and equity.

First, the conventional bibliometric patent indicators for the period 1975–82 were used, including: number of patents, total citations to those patents, average cites received per patent, and changes in the number of patents and citations received, therein, over the seven year period. Scientific papers and changes in publishing activity were also measured against financial performance. Three new patent-based indicators of technological concentration were also included in this analysis, as shown in table 8.

Table 8
Research concentration variables

% 424	% of patents in class 424 drugs, bio-affecting and body treatment
% 424/260	% of patents in class 424 and class 260, Chemicals, carbon compounds
% Top 3	% of patents in class 424 plus class 260, plus the other most important patent class in the company's portfolio

The financial measures are from the same time period as the patent activity measures, and the analysis examines the linkage between patents and overall corporate performance. This analysis was restricted to 16 of the 17 drug companies analyzed previously, because financial data for Hoffmann La Roche was not available.

As with the publication and patent data, there is a major problem in defining the pharmaceutical aspects of a company's finances, since the financial data represent the sales and profits for the entire firm, and there is no way to segregate just the pharmaceutical aspects.

To overcome this problem, the analysis of the relationship between patents and economic activ-

Table 7
Six financial variables

1 C SALES	Average annual percent *change in net sales (current dollars)* by company, 1973–82; the source of data for this measure and for the next four indicators was Standard and Poor's *Industry Surveys*, various editions.
2 C PROF	Average annual percent *change in net pre-tax profits (current dollars)* by company, 1973–82.
3 C EARN	Average annual percent *change in earnings per common share (current dollars)* by company, 1973–82.
4 C D/SN	Average annual percent *change in dividends per common share (current dollars)* by company, 1973–82.
5 C BK/SH	Average annual percent *change in book value per common share (current dollars)* by company, 1973–82.
6 C C EQTY	Average annual percent *change in common equity* by company, first quarter 1974 to first quarter 1985; this data was obtained from *Business Week's* annual scoreboard of financial performance published in the spring of each year.

Table 9
Correlation of financial (FINANC) and ten non-financial variables using data from sixteen pharmaceutical companies [a]

PATs	−0.230
PCITES	0.035
C/PAT	0.628 **
% 424	0.566 *
% 424/260	0.484
% TOP 3	0.512 *
C PATS	−0.094
C PCITES	−0.056
TOTPUBS	−0.061
C PAPS	0.087

[a] * $P < 0.05$; ** $P < 0.01$.

ity was carried out after dividing the 16 companies into three groups: (1) drug-dependent, (2) drug-dominant, and (3) multi-product firms, as shown previously in table 1. As a further simplification, because the 6 financial variables are highly inter-correlated, and all are in terms of percent change, they can be summarized by the mean of all six: a single financial change factor FINANC (the high intercorrelation amongst these variables guarantees very similar results if any were used in place of FINANC). The means of this change variable for the three groups of pharmaceutical companies were:

drug-dependent: +14%/year/company,
drug-dominated: +10%/year/company,
multi-product: +8%/year/company.

showing that the more drug-oriented companies grew faster in this time period.

FINANC was then correlated with all patent and publication data. Table 9 shows some of the results of this analysis. Of the variables, only C/PAT (mean number of cites received per patent) and the three variables indicating concentration in certain patent classes (%424, %424/260, and %TOP 3) have positive and meaningful correlations: the number of patents, PATS, is weakly and negatively correlated with financial performance. Citation to the patents, (PCITES), change in number of patents (C PATS), changes in citation to the patents (C PCITE), total publication (TOTPUBS), and changes in publication (C PAPS) are all uncorrelated with financial performance.

This shows that, for the pharmaceutical companies as a whole, the more traditional indicators: patent counts, publication counts, and changes in these counts, had no particular linkage with financial growth.

This in fact is consistent with previous research showing that "patent statistics are more likely to indicate the total size of the research effort than the rate of significant innovation" [5, p. 397].

These results can also be compared to those published by Scherer [15] which seem to show a relation between growth and patenting. This correlation, however only holds during years of high growth: no correlation was found during years of recession or recovery. Therefore, it is not surprising that over a period of years (in this case from 1973 to 1983), which included both economic upturns and downturns, there will be little correlation between patent counts and growth.

The patterns of high correlation between financial performance and citations received per patent – an indicator of innovativeness or quality – and of high correlation between financial performance and concentration, suggest that the fastest growth in pharmaceuticals comes to companies with highly innovative, concentrated research efforts. Citations per patent are a particularly interesting measure: patents appear to be highly cited for two reasons, which are sometimes interrelated, but not always.

First, the important seminal patents in any field are normally cited many times, as the technology is built from that original invention. Thus a company having highly cited patents will, per se, probably have a disproportionate share of the very leading edge technology. The association between originating technology and business results is reasonable in the context of the observed life cycle of pharmaceuticals: drug companies which are originators of technology are likely to dominate the early peak of a new product sales curve.

Second, highly cited patents, especially patents that are highly cited in a short period of time, are often heavily cited by subsequent patents owned by the same company. That is, highly cited patents are often part of a very tightly interlocked stream of inventions, and having highly cited patents may, in this sense, also be indicative of coherence, concentration and direction in a company's technological and patenting strategy.

In order to illustrate specifically the coherence of patent citation data, fig. 1 shows a patent citation diagram for all of the patents in the US patent system which are citing to a 1976 Smith-

Kline and French Laboratories patent number 3950333 for "pharmacologically active guanidine compounds". This is one of the patents related to Tagamet.

On that diagram all patents from the most highly citing company, in this case SmithKline itself, are placed in the center column. For example, counting up from the bottom, there were 2

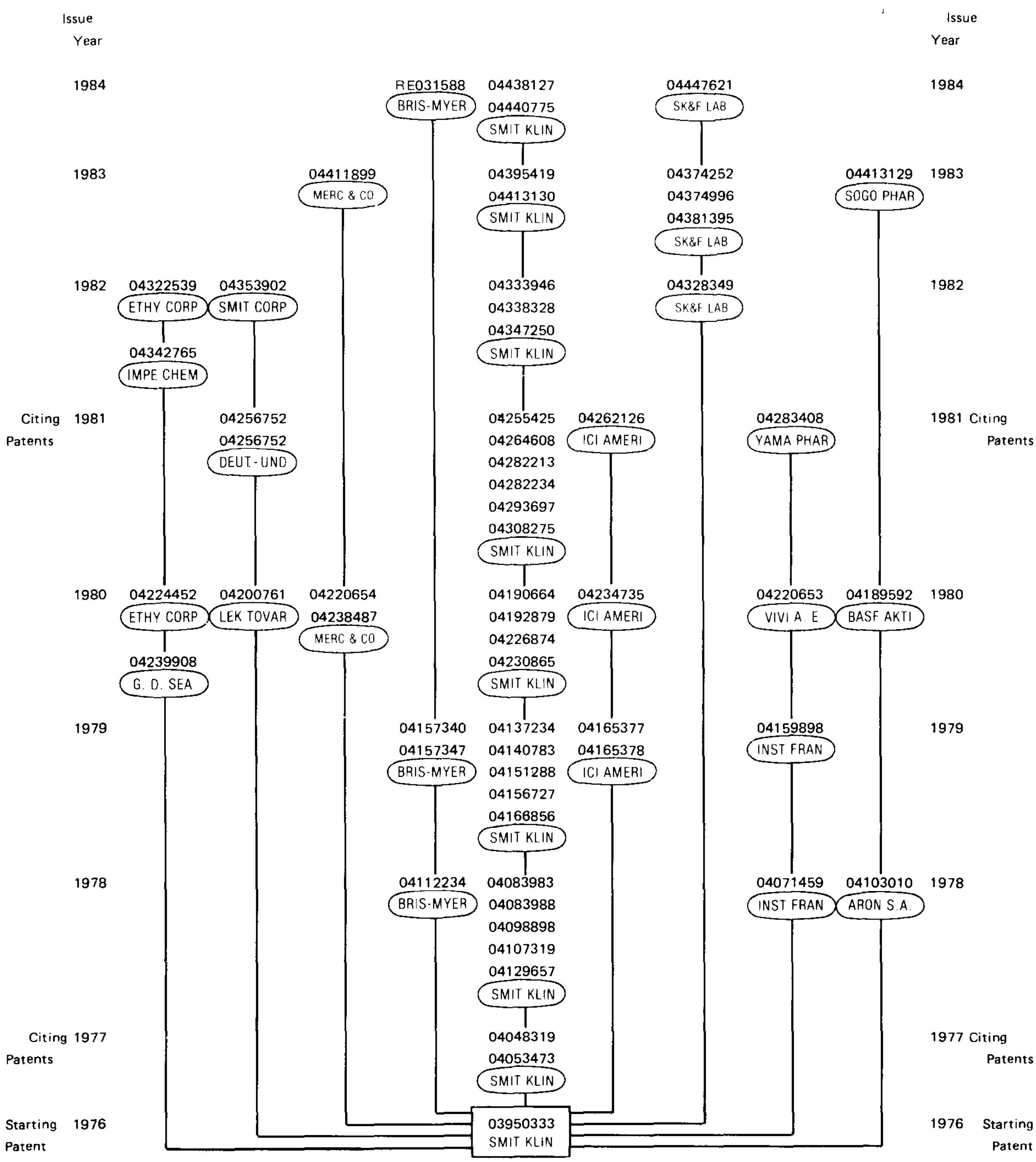

Fig. 1. US patent number: 3950333. Pharmacologically active guanidine compounds. Assignee: Smith, Kline & French Laboratories.

1977 SmithKline patents that cited 3950333, 5 1978 SmithKline patents citing 3950333, 5 other citing patents in 1979, and so on, for a total of 29 other later SmithKline patents that cite this 1976 starting patent.

In that diagram, computer generated company name abbreviations are shown in ovals and the actual number of the citing patents printed out. The patents in the second column to the right of center identified as SK&F Lab are also Smith-Kline patents, but generally with inventors in the US, while almost all of the patents in the center column are from the UK subsidiary of Smith-Kline.

Other companies whose patents cite this patent include 4 citations from ICI America patents, 4 cites from Bristol–Myers patents as well as cites from Merck, and from various German, Japanese, French, and other companies. The one patent listed there under Smit Corp is also a SmithKline and French Laboratories patent, for which there was some minor variance in the company name.

The main point of this diagram, however, is to show that this patent was cited by a large number of other patents, many of which are assigned to SmithKline, but many of which are also from other pharmaceutical companies throughout the world active in the research and development closely aligned to this very, very highly cited SmithKline patent. A patent cited this frequently is amongst the top fraction of a percent of most highly cited patents in the US patent system.

To ensure that these correlations are not merely artifacts of the between group differences in financial growth, the correlation between FINANC and the four variables (C/PAT, %424, %424/260, and %TOP 3) was compared for each group as shown in table 10. Table 10 clearly shows that the predictive power of these four patent measures is retained within all three company groups. There is a general trend of high correlations within group 1 and lower correlations in groups 2 and 3. However, the pattern of a decrease of group 3 as compared with group 2 is not consistent, and the small sample sizes preclude any meaningful statistical analysis.

Using C/PAT as a predictor variable of FINANC, the three regression lines are:

(group 1) FINANC = 7.2 * C/PAT + 3.2
(group 2) FINANC = 3.8 * C/PAT + 5.1
(group 3) FINANC = 10.3 * C/PAT − 4.2.

Table 10
Correlation of FINANC and four patent variables. For the sixteen pharmaceutical companies classified into three groups

	Group 1 (drug-dependent) $N = 4$	Group 2 (drug-dominant) $N = 8$	Group 3 (multi-product) $N = 4$
C/PAT	0.847	0.338	0.422
% 424	0.732	0.374	0.397
% 424/260	0.807	0.378	0.090
% TOP 3	0.849	0.236	0.765

Interestingly, these equations indicate that a small change in the mean number of cites per patent (C/PAT) predicts the largest increase in change in financial performance (FINANC) for the least drug-dependent companies (group 3). This is, however, due to a statistical artifact: the larger range of C/PAT and FINANC within group 3.

5. Discussion

Patent counts and patent citation counts may reveal two different aspects of the research and development cycle, with patent counts indicating the size of research inputs and patent citation counts indicating the quality or impact of the research outputs. This was indicated by the finding that number of patents correlate with expert opinions, budgets, and publication but not with financial performance, while cites per patent correlates with financial performance, but not with budget, expert opinion, or publication. In addition, the concentration of company patents was highly correlated with financial performance.

It is important to point out that the concentration and patent citation variables are linked conceptually as well as statistically, in that both relate to the presence of a small number of highly cited patents in company portfolios. These highly cited patents are often associated with major drug innovations. For example, SmithKline has had a very major increase in sales and profits, due to the anti-ulcer drug Tagamet, and a number of SmithKline's highly cited patents are Tagamet patents. These Tagamet patents are cited by many other SmithKline patents, as well as by many patents assigned to other companies. Thus, the Tagamet patents are a specific example of highly interlinked company patents in a specific technological area, high citation performance, and financial rewards for the discovery.

The results reported herein have been shown for a relatively small number of companies and in a field that is uniquely dependent upon patent protection to impose imitation delays upon the competition [14]. Nevertheless, they suggest that the use of patent citation data is one way to disentangle some of the company-to-company differences in patenting policies from the quality of company research programs. In research just starting at CHI Research, this methodology will be extended and tested for companies in the electronics, chemicals and auto-parts industries.

References

[1] B. Basberg, Technological change in the Norwegian Whaling Industry: A Case Study of the Use of Patent Statistics as a Technology Indicator, *Research Policy,* 11 (1982) 163–171.

[2] M.P. Carpenter, M. Cooper and F. Narin, Linkage between Basic Research Literature and Patents, *Research Management*, 13 (2) (March 1980) 30–35.

[3] M.P. Carpenter and F. Narin, Validation Study: Citations as Indicators of Science and Foreign Dependence, *World Patent Information*, 5 (3) (1983) 180–185.

[4] M.P. Carpenter, F. Narin and P. Woolf, Citation Rates to Technologically Important Patents, *World Patent Information*, 3 (4) (1981) 160–163.

[5] W.S. Comanor and F.M. Scherer, Patent Statistics as a Measure of Technical Change, *Journal of Political Economy*, 77 (3) (1969) 392–398.

[6] J. Davidson Frame and Francis Narin, NIH Funding and Biomedical Publication Output, *Federation Proceedings*, 35 (14) (December 1976) 2529–2532.

[7] Z. Griliches, Market Value, R&D, and Patents, *Economic Letters*, 7 (1981) 183–187.

[8] E. Koenig, A Bibliometric Analysis of Pharmaceutical Research, *Research Policy*, 12 (1983) 15–36.

[9] F. Narin, Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity, Contract NSF C-627, National Science Foundation (31 March, 1976). NTIS Accession No. PB252339/AS.

[10] F. Narin, NIH Program Evaluation Report: Subjective vs. Bibliometric Assessment of Biomedical Research Publications, US Department of Health and Human Services, Public Health Service, National Institutes of Health (1983).

[11] F. Narin and E. Noma, Is Technology becoming Science? *Scientometrics*, 7 (3) (1985) 369–381.

[12] National Science Board, *Science Indicators 1972, 1974, 1976, 1978, 1980, 1982* (National Science Foundation, Washington, D.C.).

[13] OECD Workshop on Science and Technology Indicators in the Higher Education Sector, Summary Record, Organisation for Economic Co-operation and Development, Directorate for Science, Technology and Industry. (Paris, 18 November, 1985).

[14] A. Phillips, *Technology and Market Structure: A Study of the Aircraft Industry* (Lexington Books, Lexington, MA, 1971).

[15] F.M. Scherer, Corporate Inventive Output, Profits and Growth, *Journal of Political Economy, 73* (1965) 290–297.

[16] Science Policy Research Unit (SPRU) Annual Report University of Sussex, UK (1984/1985).