- (1955); Chem. Abstr., 50, 10124d (1956).
 (21) A. E. Arbusov and L. V. Nesterov, Dokl. Akad. Nauk SSSR, 92, 57 (1953); Chem. Abstr., 48, 10538b (1954).
 (22) (a) A. H. Ford-Moore and J. H. Williams, J. Chem. Soc., 1465 (1947); (b) G. M. Kosolapoff, J. Am. Chem. Soc., 70, 1971 (1948).
- (23) J. Kwiatek, U.S. Patent 2882311 (1959); Chem. Abstr., 53, 16964i
- (1959).
 (24) R. M. Davidson and G. L. Kenyon, J. Org. Chem., 42, 1030 (1977).
 (25) S. Inokawa, Y. Nakatsukasa, M. Horisaki, M. Yamashita, H. Yoshida, T. Ogata, and H. Inokawa, Synthesis, 179 (1977).

An Evaluation of the Leading Patent Equivalents Services[†]

TRISHA M. JOHNS*

G. D. Searle & Company, Skokie, Illinois

WILLIAM G. ANDRUS

The Upjohn Company, Kalamazoo, Michigan

STANLEY DE VOE

Medical Research Division, American Cyanamid Company, Pearl River, New York

JOHN MYERS

A. H. Robins Company, Richmond, Virginia

ROGER GRANT SMITH

Merck & Co., Inc., Rahway, New Jersey

OTOKAR UHLIR

Abbott Laboratories, North Chicago, Illinois

Received May 29, 1979

Patent equivalents services are useful tools for identifying patent families. A patent family consists of a group of patents in various countries which cover the same invention. The patent equivalents services provided by the three leading suppliers—Chemical Abstracts Service, Derwent Publications Ltd., and INPADOC—have been tested and evaluated for coverage, recall, and accuracy. Results are reported for a test sample consisting of 75 pharmaceutical patent families.

INTRODUCTION

Patent equivalents arise when a patentee (either the inventor or the inventor's assignee) applies for patent protection for an invention in a number of countries. The patentee may then be issued patents in several countries, all protecting the same invention. Since it would be misleading to treat each of the patents as if it pertained to a separate invention, it has become the practice of those who deal with patent information to label such patents as "equivalents", "counterparts", "cognates", "corresponding", or, collectively, a "patent family".

Why is it important to know if a given patent has equivalents in other countries? Some of the main reasons are the following:

- (1) Avoiding translation-Locating an equivalent in a familiar language is usually cheaper and faster than having a translation made of a patent in an unfamiliar language.
- (2) Determining extent of patent protection—Knowledge of the countries in which patent protection for a competing product has been applied for is useful to the business executive in planning marketing strategy and in assessing the competitor's confidence in the product. Newman and Hoegberg² stress the importance of such knowledge in making long range decisions.
- (3) Disclosure of patentability problems—A patent issued by an examining country³ will sometimes be delayed if there are legal impediments (e.g., an interference) to the issuance
- Presented at the PMA Science Information Subsection Annual Meeting, March 6, 1979.

- of the patent. The patent attorney can often detect such patentability problems by analyzing the time lags between the filing dates and issue dates of the examined members of a patent family. The attorney may also glean valuable information by comparing the claims of equivalent patent documents.⁴ Claims of questionable patentability will probably have been eliminated from the examined patents.
- (4) Avoiding double abstracting and indexing—It is far more efficient to abstract and index only one member of a patent family and relate the others by cross-reference than it is to abstract and index every member of the patent family individually. The amount of work saved becomes apparent when one considers the huge volume of new patent documents: the 24 major patent-issuing countries currently generate about 9500 documents per week, of which 5200 are found to be equivalents.⁵ By not having to abstract or index these equivalents, 50% of the effort required is saved. This is usually a concern of the large commercial services which abstract and index patents, although there are a number of private firms which do their own patent abstracting and indexing in specific areas.
- (5) Avoiding confusion during patent searching-This use of equivalents has been discussed by Valicenti.⁶ When more than one data base is used during a patent search, the same patent family may have been uncovered by each of the data bases, but because different members of the family are cited by each service, they superficially appear to be separate, distinct hits. The correlation of these equivalents will reduce the hit count and avoid confusing the recipient of the search.

How do we know when patents are equivalent? The usual method is by examining the bibliographic data which appear at the beginning of each patent. Such data include the name of the patentee, country, date, and application number of the priority application. If the priority information is identical for two patents, the patents are assumed to be equivalent. Thus, patent families may be built up by matching priority data. The organizations which provide comprehensive equivalents services have developed efficient data-processing techniques for performing this task with thousands of patents.

Unfortunately, the equivalent determination process is not always so straightforward. Typographical errors can cause mismatches. Confusion can result when a single patent is based on two or more priority applications. Confusion can also result when two or more legally related applications are filed on one invention, often as the result of provisions in the local patent laws. (Some countries allow the applicant to file divisions, continuations, or continuations-in-part of pending applications. In some patent offices, already issued patents may be reissued or be enhanced by patents of addition.) The ultimate ambiguity arises when a patent family is filed nonconvention, that is, without claiming a priority application.⁹ Unless a time-consuming search is made via patentee and subject matter (and even that does not always bear fruitful results), it is likely that the members of the patent family will never be linked together and will appear to be separate, isolated

At the present time in the field of chemistry there are three main services which provide information on patent families: Chemical Abstracts Service (CAS) of Columbus, Ohio, U.S.A.; Derwent Publications Ltd. of London, England; and the International Patent Documentation Center (INPADOC) of Vienna, Austria. CAS has published a printed Patent Concordance since 1962. Derwent has published a printed Equivalents-to-Basic¹⁰ Index since 1963.¹¹ Both CAS and Derwent have also introduced on-line equivalents services corresponding to their printed indexes. CAS' service is called CA Patent Concordance and is available via Lockheed Information Service's DIALOG system. Derwent's service is called WPI and is available via System Development Corporation's ORBIT System. INPADOC's service, called the Patent Family Service (PFS), is available only on microfiche. 12 INPA-DOC's comprehensive file goes back to 1973; a less complete file goes back to 1968.

Details of the coverages of these three services are given in Appendixes A, B, and C.¹³⁻¹⁵ Their specific methodologies for determining equivalents have been described in the literature.^{5,16,17} Other equivalents services also exist but they are limited either in their country coverage or subject matter coverage. Some of these have been briefly described by Duffey.¹⁸

RELATED LITERATURE

According to Cleverdon, ¹⁹ there are six important criteria to be considered in assessing the efficiency of an information retrieval system: coverage, recall, precision, response time, user effort, and form of output. Previous studies have touched on some but not all of these criteria.

Smith, Anderson, and Jackson²⁰ compared three on-line files—Derwent, CAS, and the IFI ClaimsTM Service—for gross number of equivalents found and unique equivalents. Derwent's WPI service was found to be superior; however, the sample size (one patent family) was intended for illustration rather than in-depth analysis.

Kaback²¹ compared the same three on-line files in terms of coverage, timeliness of file update, and ease of use. Of the three files, Kaback stated that Derwent's was clearly superior. Kaback's conclusions were based upon his extensive experience

in the patent information field rather than experimental data.

The question of timeliness was also addressed by Oppenheim and Sutherland²² in a comparison of CAS and Derwent. For a sample of 147 metallurgy patents, Derwent was found to be twice as fast as CAS (107 vs. 277 days) at getting new patent citations incorporated into its files.

To the best of the authors' knowledge, there has been no comparison of INPADOC's patent family service with any other equivalents file.

BACKGROUND OF THE STUDY

The need for a meaningful comparative study of the three services was recognized by the authors. Since the authors are all employees of pharmaceutical companies, it was felt to be appropriate to concentrate on patents in the pharmaceutical field. It was decided to limit the study to two of the performance criteria of Cleverdon, i.e., coverage and recall, since these were considered to be the most vital requirements of an equivalents service. The other criteria, though important, would have required a considerable additional expenditure of effort which the authors did not feel was justifiable at this time. Thus, the study was limited to a statistical comparison of the coverage and recall of pharmaceutical patents as covered by the three main services.

PROCEDURE

The evaluation is based on a comparison of patent families owned by the participating companies with those reported by the three equivalents services. Each contributor selected 12 to 15 of his company's inventions in the pharmaceutical area that had been patented in at least five countries. Patents chosen for the study were published from 1973 to 1977, with families evenly distributed over this time, when possible. An additional criterion, that at least one of the family be published in France, Germany, Great Britain, or the United States, increased the likelihood that the invention would be included in all the services. This particular patent was identified as the "starting-point patent" for the equivalents search in each service.

To ensure that all participants use the systems optimally and consistently, procedures were drawn up and submitted to the suppliers for review. At the same time, the suppliers were asked to verify country coverage lists for their services which would be needed in the final evaluation.

Although the study deals exclusively with nonproprietary information, confidentiality with respect to patent numbers and subject coverage was assured by the fact that only statistical results were reported. This evaluation is, therefore, a composite of individual results of the contributors.

For each patent selected, in-house patent files were consulted to identify published equivalents. These were noted on the Confidential Work Sheets (Figure 1), beginning with the starting-point patent (i.e., the first family member published in France, Germany, Great Britain, or the U.S.). The starting-point patent for each invention was then searched in Derwent, CAS, and INPADOC, and all equivalents found were checked off on the work sheet.

Since each of the three files is organized differently, a separate procedure was required for each service. For both Derwent and CAS, the on-line method of searching was chosen over the manual method for the sake of convenience. It was assumed that both of these files are the same as their hard-copy counterparts.

Using the Derwent on-line patent file (SDC's "File WPI"), the starting-point patent was entered, and the accession number, priority information, and family members were printed. All additional priorities obtained were searched in the same way, and any new hits were printed. This procedure

Table I. Derwent Pharmaceutical Patent Coverage

country	from	to	comments
Argentina	Feb 6, 1975 (W40)	Oct 31, 1975	
		(Y03)	
Australia	1963	1969	
Austria	March 15, 1975 (W15)	present	
Belgium		-	gap in coverage from Dec 30, 1969 to Jan 29, 1970
Brazil	Dec 16, 1975 (X01)	present	unexamined specs only
Canada	1963	present	
Czechoslovakia	March 28, 1975 (W20)	present	excepted specs only
Denmark	Oct 14, 1974 (V45)	present	specs OPI 18 months only
Finland	Sept 30, 1974 (V45)	present	specs OPI 18 months only
France	1963	present	FR 2 000 000 thru 2 003 800 not covered
Germany, East	1963	present	DL65031 to DL69361 omitted from Plasdoc, Sections A, B
		-	and C only thru R11 to R23
Germany, West	1963	present	Offenlegenschrift with numbers less than 1 800 000 not covered from Dec 18, 1969 to Sept 14, 1972
Great Britain	1963	present	•
Hungary	May 28, 1975 (W26)	present	covers applications OPI with deferred examinations (H nos. and applications OPI with complete examinations if not c pletely OPI with the first ones (T nos.)
Ireland	1963	1969	• • •
Israel	March 13, 1975 (W15)	present	
Italy	Sept 30, 1977 (A01)	•	IT 1 018 001 onwards
Japan	1963	present	chemical only
Netherlands	1963	present	gap in OPI (green) from Dec 31, 1969 to Jan 30, 1970 gran (whites), NL 128629 thru NL 130812 not covered
Norway	Nov 4, 1974 (V48)	present	unexamined only
Portugal	Oct 4, 1974 (V52)	present	•
Rumania	June 1, 1975 (W32)	present	
South Africa	1963	present	gap at start of CPI, Dec 1969 and Jan 1970 journals omitted
Sweden	Sept 23, 1974 (V42)	present	unexamined only
Switzerland	1963	present	gap at start of CPI, specs published Jan 1, 1970 through Fel 26, 1970 omitted, Sections A, B & C only covered to wee \$06
United States	1963	present	
Soviet Union	1963	present	July thru Nov 1969 omitted

was continued until no new priorities were found. Any cross-references to new accession numbers were also searched.²³

Using CAS' on-line Patent Concordance (Lockheed's "File 43"), the starting-point patent was entered and corresponding patents were printed on-line.²⁴

Access to the INPADOC service was made available by IFI Plenum Data Co., its U.S. agent. Participants were allowed to use the patent family microfiche at IFI's Washington office; alternatively, IFI volunteered to do the searching, given the list of starting-point patents. In either case, priority information was obtained from the INPADOC Numerical Data Base Index, using the starting-point patent. Using the INPADOC Numerical List, the priority dates of all priority serial numbers were verified. Each priority was then searched in the Patent Family Service microfiche for a list of equivalents. Additional priorities were sought by checking the bibliographic data on at least one other member of the patent family (particularly U.S. patents). New priorities found were then looked up in the Patent Family Service, and the process was repeated until no new priorities were found.

Once the patent family members given by each service were checked against the list of known equivalents, data for each family was tabulated and recorded on the Individual Summary Sheet (Figure 2). Data recorded included the number of equivalents known to be in the family, and for each service, the total number found, the number which should have been found (determined by referring to the country coverage charts provided by the suppliers, Tables I–III), and the number of unique equivalents found. From these data, the total number of missing patents was calculated, and an analysis was made to determine possible reasons for their absence.

The total figures for all patent familes considered were summed up on the Company Summary Sheet (Figure 3). The Company Summary Sheets were cumulated by the senior author.

ly thru Nov 1969 omitted

-- For Investigator's Use Only -
NVENTION NO. ______ Internal ID No. ______ U.S.S.N. ______

Known Issued/Published Equivalents	CAS	DERWENT	INPADOC	COMMENTS
NB: INDICA	TE THE START	ING-POINT-PAT	ENT NUMBER FIRST	
			<u> </u>	
	1		t i	
	+		 	
	+		 	
				
	+		 	
	 		L l	
	T			
Erroneous Patents	CAS	DERWENT	INPADOC	COMMENTS
			İ	
-		<u> </u>		
	 	 	 	
	 	 	1	
	+	 	├───	

Figure 1. Confidential Work Sheet.

RESULTS

The composite of the companies' findings is shown in Figure 4. Two types of recall calculations were made. One was based on the total patent family members found by each service

* [1 1 1 0 1 1	(A) TOTAL MOAN TABLET	IEHIBEKS		
		CAS	DERWENT	INPADOC
(B)	Total which should have been found:			<u> </u>
(C)	Total found by Service:			
(D)	Total of unique family members found:			<u> </u>
(E)	Total missing (A-C):			
(F)	Total not covered by service (A-B): (Includes patents not covered because (1) Country not covered at all by service, or (2) Country not covered for time period when patent published, or (3) Country not covered for non-national inventors.)			
(G)	Total missing of those which should have been found (B-C):			
Reas	ons for missing patents:			
(H)	In file, but not connected with rest of family			
(1)	In file, but connected with wrong family			
(1)	Should be in file, but is not			-
(K)	Other (explain)			
(L)	Total patents erroneously connected to family:			
(M)	Method used to search CAS (circle one): or	ı-line, n	ranual inde	xes
(N)	Method used to search Derwent (circle one):	on-lir	ne. manual	

Figure 2. Individual Summary Sheet.

indexes

NUME	ER OF PATENT FAMILIES			
(A)	TOTAL KNOWN MEMBERS OF ALL PATENT FAMILIES_			
		CAS	DERWENT	INPADO
(8)	Total which should have been found:			
(C)	Total found by Service:		1	
(۵)	Total of unique family members found:		<u> </u>	
(E)	Total missing (A-C):		t	
(F)	Total not covered by service (A-B): (Includes patents not covered because (1) Country not covered at all by service, or (2) Country not covered for time period when patent published, or (3) Country not covered for non-national inventors.)			
(G)	Total missing of those which should have been found (B-C):			
Reas	ons for missing patents:			
(H)	In file, but not connected with rest of family			
(1)	In file, but connected with wrong family			
(J)	Should be in file, but is not			
(K)	Other (explain)			
(L)	Total patents erroneously connected to family:			
(M)	Method used to search CAS (circle one): on-	-line, r	anual index	es

Figure 3. Company Summary Sheet.

divided by total members known to exist. In this version of recall, R_1 , CAS found 33.4%, Derwent found 60.8%, and INPADOC found 73%. A second recall calculation, R_2 , was made by dividing the total family members found in each service by the total family members which should have been in their service based on the countries covered. By this measure of recall, CAS found 76.7%, Derwent found 93.6%, and INPADOC found 92.5%. These results are shown in Figure 5.

(N) Method used to search Derwent (circle one): on-line, manual

Another measure of the value of an equivalents service is

Table II. INPADOC Data Base, 780630, Version A

ble I	ole II. INPADOC Data Base, 780630, Version A										
		country		lication of							
	AR	Argentina	730	0208-77	0531						
	AT (OE)	Austria	730	0110-78	0615						
	AU	Australia	730	0118-78	0112						
]	BE	Belgium	730	0101-78	0525						
]	BG	Bulgaria	730	0215-78	0412						
]	BR	Brazil	760	0106-78	0523						
(CA	Canada	740	0101-78	0613						
(CH	Switzerland	730	0115-78	0512						
(CS	Czechoslovikia	730	0222-78	0531						
(CU	Cuba	740213-770908								
(CY	Cypress	750	0301-77	0318						
]	DD (DL)	East Germany	730	0105-78	0531						
]	DE (DT)	West Germany	730	0101-78	0629						
]	DK	Denmark	730	0102-78	0612						
]	EG (ET)	Egypt	760	0131-78	0329						
1	E S	Spain	730	0101-78	0601						
]	FI (SF)	Finland	780	0131-78	0531						
]	FR	France	730	0105-78	0609						
(GB	Great Britain	730	0104-78	0628						
(GR	Greenland	770	0704-78	0515						
I	HK	Hong Kong	760	305-78	0505						
I	HU	Hungary	730	0129-78	0428						
I	E (EI)	Ireland	730	110-78	0104						
I	Ĺ	Israel	730	0130-78	0310						
I	N	India	750	0802-77	0226						
	T	Italy	730	0125-78	0320						
J	IP (JA)	Japan	730	0402-78	0531						
	KE	Kenia		0711-77							
]	KR (KS)	Republic of Korea		0108–78							
	ĹŬ	Luxemburg		105-78							
	MC	Monaco	751	1010-77	1118						
ľ	MN (MO)	Mongolia		l 120–77							
	MW	Malawi		0509-77							
	NL	Netherlands	730	0102-78	0615						
	NO	Norway	730	0102-78	0508						
	PH	Philippines		0703-78							
	PL (PO)	Poland		0228-78							
	PT	Portugal		0101-78							
	RO (RU)	Romania		120-78							
	SE (SW)	Sweden		0108-78							
	SU	USSR		0108-78							
	ΓR	Turkey		0101-76							
	US	United States		0102-78							
	YU	Yugoslavia		228-78							
	ZA (7D)	South Africa		730131 - 771228 730122 - 771121							
. 2	ZM (ZB)	Zambia	730)122-77	1121						
NUMBE	R OF PATENT F	AMILIES	_								
(A)	TOTAL KNOWN M	MEMBERS OF ALL PATENT FAMILIES		1484							
			CAS	DERWENT	INPADOC						
(B)	Total which s	hould have been found:	647	964	1172						
(C)	Total found b	y Service:	496	902	1084						
		ue family members found:	5	67	234						
			988	582	400						
	Total missing		988	382	400						
		ered by service (A-B): ents not covered									
		Country not covered at									
		e, or (2) Country not			}						
		ime period when patent (3) Country not covered									
		onal inventors.	837	520	312						
	Total missing	of those which should and (B-C):	151	62	. 88						

(L) Total patents erroneously connected to family:

Figure 4. Composite Summary Sheet.

have been found (B-C):

the number of unique patent family members it can provide. Unique patent family members are those equivalents that were found in one service only and not in the other two. For the 75 patent families considered, the authors found that CAS

Table III. Patent Coverage: Chemical Abstracts--Volume 87

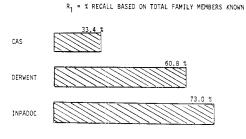
The following table summarizes CA patent coverage. For a more complete description of CA patent coverage for a particular volume or for patent coverage prior to Volume 72, the reader is referred to the introductory material at the

beginning of Issue 1 of each Volume. For details on how to obtain patent documents, see the introduction in CA Issue 1 of the current volume.

Patent Coverage Summary

Type of Patent	Abbreviation	Year	1970		1971		1972		19	1973		74	1975		19	76	1977	
Type of Fatent		Vol.	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87
Australian	Austl		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Austrian	Aust		R	R	R	R	R	R	R	R	R	R	l A	A	A	A	A	A
Belgian	Belg		l R	R	R	R	R	R	R	R	R	R	A	Α	A	A	A	A
Brazilian	Br Pl		1		ŀ	į.	1			-	1					A	A	A
British	Brit		A	A	A	A	l A	A	A	A	A	A	A	À	A	Α	A	A
British Amended B	Brit B												'		1	A	A	A
Canadian	Can		R	R	R	R	l R	R	R	R	R	R	Α	Α	A	A	A	A
Czechoslovakian	Czech		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Danish	Dan		R	R	R	R	R	R	R	R	R	R	R	R	l R	R	R	R
Finnish	Finn		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
French	Fr		A	A	A	Α	A	A	A	A	A	A	A	A	A	A	A	A
French Addition	Fr (5 digit no	umber)	A	A	A	A	A	Α	A	Α	D	D	D	D	D	D	D	D
French Medicinal	Fr M		Α	Α	A	Α	A	Α	A	Α	D	D	D	D	D	D	D	D
French Addition to Medicinal	Fr CAM		A	Α	A	A	A	Α	A	Α	D	D	D	D	D.	D	D	D
German (East)	Ger E		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
German	Ger		A	Α	A	Α	A	Α	A	Α	A	Α	A	Α	A	Α	Α	A
Hungarian	Hung		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Indian	Ind		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Israeli	lsr		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Italian	Ital		R	R	R	R	R	R	R	R	R	R	R	R	R		'	1 "
Japanese	Jpn		R	R	R	R	R	R	R	R	R	R	R	R	R	Α	Α	A
Japanese Kokai	Jpn K									R	R	R	A	A	A	A	A	A
Netherlands	Neth		R	R	ĸ	R	R	R	R	ħ.	R	R	A	A	Α	A	Α	Ä
Norwegian	Norw		R	R	R	R	R	R	R	R	R	·R	R	R	R	R	R	R
Polish	Pol		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Romanian	Rom		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
South African	S Afr		A	A	Α	Α	А	Α	A	Α	Α :	Α	A	Α	A	Α	Α	A
Spanish	Span		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Swedish	Swed		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Swiss	Swiss		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
U.S.S.R.	USSR		R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
United States	US		A	Α	A	Α	A	Α	A	A	Α	Α	A	Α	A	A	A	Ä
U.S. Defensive Publications	US T		Α	Α	A	Α	Α	Α	Α	Α	Α	Α	A	Α	A	A	A	Ä
U.S. Patent Applications	US ●							-			A	A	A	A	A	A	A	Ä
U.S. Published Patent Application	US B											-	A	Α.	A	A	D	D
U.S. Reissue	USR		A	A	A	Α	A	Α	A	A	A	Α	A	A	A	A	Ā	A

- A all patents of chemical or chemical engineering interest.
- R chemical and chemical engineering patents issued only to individuals or organizations resident in the granting country.
- D documents discontinued by issuing government.



R₂ = % RECALL BASED ON FAMILY MEMBERS KNOWN IN COUNTRIES COVERED BY SERVICE



Figure 5.

provided five unique family members, while Derwent gave 67 and INPADOC gave 234.

Erroneously reported patents, i.e., those which were not true patent family members, were also recorded. CAS reported

Total analyzed

- (H) In file, but not connected with rest of family
- (I) In file, but connected with wrong family
- (J) Should be in file, but is not
- (K) Non-convention
- (K') U.S. published applications

CAS DERWENT INPADOC

137 48 58

42 5 0

15 15 1

70 23 31

8 5 26

2 0 0

Figure 6. Reasons for missing patents.

four erroneous equivalents while Derwent and INPADOC reported one each.

The patents which were missed were also tabulated, and four of the participants investigated the reasons for these missed patents. The categories of reasons and numbers in each category are shown in Figure 6.

DISCUSSION

 $R_{\rm I}$ is simply a measure of which equivalents service retrieves the most equivalents. Based on the results of this study, INPADOC should be the service of choice when the most total equivalents are desired. INPADOC found three-quarters of all equivalents, while Derwent finished a close second at three-fifths. CAS was only able to reveal one-third of known equivalents. The reasons for these figures are directly at-

tributable to country coverage of these services. INPADOC covers 46 countries, Derwent covers 24, and CAS covers 26 but only 15 of them fully.

R₂ takes into account these differences in country coverage by measuring the efficiency with which they disclose the patent family members of countries they claim to cover. This measurement does not hold the services responsible for revealing equivalents outside the scope of their service. Derwent and INPADOC display essentially the same degree of efficiency, covering approximately 93%. CAS reports little more than three-quarters of the patents they claim to cover.

It is remarkable that Derwent is on a par with INPADOC, since INPADOC has the advantage of a special arrangement with national patent offices which enables it to receive the information in machine-readable form. Derwent, on the other hand, does in-house data processing from original documents. That all three services found unique equivalents indicates that if every possible family member must be found then all three must be consulted. The erroneous patents found can be attributed to errors, and indicate that Derwent and INPADOC are almost perfect in assigning patents to patent families.

There are a number of reasons why patents were missed in spite of the fact that the countries were within the scope of the service. The four participants who analyzed the missing hits found that CAS was often remiss in connecting equivalents to the other members of their patent family. Some contributing factors to CAS' poor performance are:

- (1) Until the beginning of 1978, CAS had the policy of entering divisions, additions, reissues, continuations, and continuations-in-part as separate patent families in their file.
- (2) Multiple priority patent families cannot be thoroughly searched in the CA Patent Concordance file because of the absence of priority information.¹³

Both CAS and Derwent were occasionally guilty of connecting equivalents to the wrong patent family. A possible explanation for Derwent's misconnected patents is a delay in processing corrections to its on-line file which persisted to the end of 1978.14

A large percentage of the missed patents were those that were simply omitted from the files altogether. CAS was particularly remiss on this point, while Derwent is the most efficient. A possible explanation for CAS' poor showing is their slowness in processing patents into their file as documented by Oppenheim and Sutherland.²² Nonconvention patents accounted for a number of the missed equivalents, particularly in the INPADOC file. Derwent is the only one of the three services which attempts to locate and link together nonconvention equivalents. Two CAS misses were accounted for by the unusual circumstance of these equivalents being part of the U.S. Patent and Trademark Office's experimental program of publishing applications prior to grant.

CONCLUSIONS

From the above, the authors came to the following conclusions for pharmaceutical patents. First, for disclosure of the maximum number of equivalents, the service of choice is INPADOC. For reliability in covering the patents which they admit responsibility for, Derwent and INPADOC are the most efficient. To ensure that all known equivalents are found, all three equivalents services (Derwent, INPADOC, and CAS) must be consulted.

ACKNOWLEDGMENT

The authors wish to acknowledge the full cooperation of all three equivalents services suppliers in the performance of this study, particularly IFI Plenum for making the INPADOC files freely available.

REFERENCES AND NOTES

- (1) Two notes on terminology: The word "equivalent" should not be taken literally when applied to patents. Although equivalent patents may protect the same invention, they seldom are identical, word-for-word documents. Because of differences in local law and examining procedure, equivalent patents, when published, may have noticeable differences in the text. In this study the word "patent" will be used generically to mean any published patent document, e.g., unexamined patents. (Strictly speaking, a patent document must be sealed or granted before it is considered a
- legally enforceable patent.)
 (2) P. Newman and E. I. Hoegberg, "What the Patent Attorney Needs from a Patent Information Point of View", J. Chem. Inf. Comput. Sci., 18,
- (3) Examining countries are those countries whose patent offices subject their patent applications to a rigorous appraisal as to the patentability of the invention, on such grounds as novelty of subject matter
- (4) The claims of the patent are the summarization of the invention in legal terminology which appear at the end of the patent document. In examining countries, the claims are often amended during examination to comply with local law and to overcome rejections put forth by the patent examiner.

 (5) R. H. Hope, "Patent Families" in "Proceedings of the International Patents
- Conference 1978" (Stratford-upon-Avon, England, April 12-14, 1978), Derwent Publications Ltd., London, pp 236-238.
- A. K. Valicenti, "The Information Chemist's View of the Patent Information Needs of Research Workers and Patent Attorneys", J. Chem. Inf. Comput. Sci., 18, 85-87 (1978).
- (7) The word "priority application" here refers to the original application on which all other equivalents are based. Under the terms of the Paris Convention⁸ to which nearly all countries of the world belong, an applicant is granted a year after filing an original application in which to file applications in other countries and be able to claim the date of the priority application. This gives all members of the patent family the same effective filing date as the priority application.
- (8) The Paris Convention of March 20, 1883, resulted in the foundation of the "International Union for the Protection of Industrial Property which came into force on July 7, 1884. The Convention was revised at Brussels on Dec 4, 1900; at Washington on June 2, 1911; at The Hague on Nov 6, 1925; at London on June 2, 1934 (with effect as from Aug 1, 1938); at Lisbon on Oct 31, 1958 (with effect as from Jan 4, 1962); and at Stockholm on July 14, 1967.
- (9) Non-convention filings occur when the applicant files in additional countries within the convention year but inadvertently omits the claim of priority, or when the applicant files in additional countries after the
- convention year has expired.

 (10) "Basic" is the term coined by Derwent to designate the first member of a patent family which it picks up in its service.
- (11) Derwent did not commence full coverage of chemistry until 1970. Prior to that, coverage of various areas of chemistry began in step-wise fashion. Pharmaceuticals began in 1963, agricultural chemicals in 1965, and polymers in 1966.
- (12) INPADOC's service is marketed in the U.S. by the IFI Plenum Data Corp.
- (13) R. J. Rowlett, Chemical Abstracts Service, private communication.
- (14)M. D. Dixon, Derwent, private communication
- (15) H. Allcock, IFI Plenum, private communication.
 (16) W. Pilch and W. Wratschko, "INPADOC: A Computerized Patent
- Documentation System", J. Chem. Inf. Comput. Sci., 18, 69-75 (1978).

 (17) P. J. Pollick, "Patent Processing at Chemical Abstracts Service—A (17) P. J. Politick, "Patent Processing at Chemical Abstracts Service—A Foundation for the Future", paper presented at the Patent Information Committee of the Pharmaceutical Manufacturer's Association Science Information Subsection Meeting, Washington, D.C., March 23, 1978.
 (18) M. M. Duffey, "Searching Foreign Patents", J. Chem. Inf. Comput. Sci., 17, 126-130 (1977).
 (19) C. W. Cleverdon, "Evaluation of Operational Information Retrieval Systems Part 1. Identification of Criteria" Cranfield England: College
- Systems. Part 1. Identification of Criteria", Cranfield, England: College of Aeronautics, 1964.
 (20) R. G. Smith, L. P. Anderson, and S. K. Jackson, "On-line Retrieval of
- Chemical Patent Information. An Overview and a Brief Comparison of Three Major Files", J. Chem. Inf. Comput. Sci., 17, 148-157 (1977).
- (21) S. M. Kaback, "Retrieving Patent Information Online", ONLINE, 2 (1), 16-25 (1978)
- (22) C. Oppenheim and E. A. Sutherland, "Studies on the Metallurgical Patent Literature. I. The Coverage of Patents by Abstracts Journals in Metallurgy", J. Chem. Inf. Comput. Sci., 18, 122-126 (1978).
 (23) If a manual search were done instead, the starting-point patent would first be searched in the Patent Number Index. Priorities would then
- be determined by referring to the WPI Gazette, Alerting Bulletins, or Basic Abstracts Journals. These priorities would then be looked up in the Priority Concordance (cumulations and all recent supplements) and if any earlier or later priorities were found, they would then be looked
- up in turn. This process would be repeated until no new priorities were found. All family members based on these priorities would be noted.

 (24) If a manual search were done instead, a search of the starting-point patent in the CAS Patent Concordance would give an associated CAS abstract number with the family members listed next to it, or another patent number would be given. That number would then be consulted to find the list of family members.