(4) Buntrock, R. E. Chemical Database Evolution. *Database* **1987**, *10* (5), 143-144.
(5) Barnard, J. M. Online Graphical Searching of Markush Structures in Patents. *Database* **1987**, *10* (3), 27-34.
(6) Stobaugh, R. E. Chemical Substructure Searching. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 271-275.
(7) Simmons, E. S. Generic Structures and Chemical Patents. *Chem. Inf. Bull.* **1985**, *37* (1), 27.
(8) Lavagnino, M. R. Computer Access to Chemical Information—The Information Specialist Perspective. *Chem. Inf. Bull.* **1985**, *37* (1), 27.
(9) Kaback, S. M. Online Patent Searching: The Realities. *Online* **1983**, *7* (4), 22-31.
(10) It may be possible in some of these examples to reduce the number of structures to be iterated by the use of selected screens. However the determination of the correct screens to use is not straightforward even to those who should know (i.e., the STN Help Desk) and definitely beyond the end user.
(11) In September 1990, MARPAT contained 66 000 Markush structure records that are found in 22 000 patent citations.
(12) This has now been increased to 30 min for MARPAT.

A Comparison of Three Online Markush Databases†

KATHLEEN A. CLOUTIER

Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana 46285

Received November 15, 1990

A comparison of the content of three Markush structure databases is presented. The scope of WPIM, Pharmsearch, and MARPAT is examined by looking at the patent documents chosen for indexing for a specific publication date. The Markush structures found in the databases for five patents are also analyzed. The three organizations producing these databases are found to have different policies and approaches in both the choice of the patent documents covered and the scope of the Markush structures indexed for a given document.

The chemical information community has recently been presented with three new databases of generic (Markush) substances from the patent literature. These Markush structure files are being created by three organizations. Derwent Publications Ltd. is building its World Patents Index—Markush (WPIM) File, the French Patent and Trademark Office (INPI) is developing Pharmsearch, and Chemical Abstracts Service is producing MARPAT. WPIM and Pharmsearch are loaded on the Markush DARC software system. MARPAT is available on STN and runs under its own, registry-compatible, software. This paper looks at and compares the content of the three databases of Markush structures. First, the coverage in terms of subject matter, patent countries, and time periods will be briefly explored. The actual coverage of patent documents[1] from a specific date will then be examined. After a review of the indexing policies of the three organizations, the actual indexing of five documents covered in all of the databases will be analyzed.

COVERAGE OF THE DATABASES

Subject Areas. INPI's Pharmsearch is the most focused product, specializing in pharmaceutical patents. Derwent's WPIM covers Pharmaceutical/Veterinary, Agricultural, and General Chemical patent documents. These areas are known to subscribers as classification sections B, C, and E, respectively. The MARPAT File from Chemical Abstracts contains Markush structures of organic or organometallic molecules for patents indexed in *Chemical Abstracts*. The documentation states that "all chemical subject areas are included".[2] Both Derwent and Chemical Abstracts have announced that they are not yet able to index some classes of structures. Currently, WPIM is not handling polymers, oligosaccharides, and polypeptides. MARPAT is not currently indexing alloys, metal oxides, inorganic salts, intermetallics, and polymers. *Chemical Abstracts* Section 74 (Radiation Chemistry, Photochemistry, and Photographic and Other Reprographic Processes) is not completely covered at this time.

† Presented at the 200th National Meeting of the American Chemical Society, Washington, DC, Aug 29, 1990.

Countries. Pharmsearch covers patent documents from the European Patent Office (EP and PCT), France, and the United States. Documents that are equivalent and identical with one already indexed are simply posted as equivalents, but equivalents that are slightly different are reindexed.[3] WPIM indexes Markush structures from the "Basic" patents in Derwent's major countries. MARPAT uses the patents selected for *Chemical Abstracts* except those from the U.S.S.R.

Time Periods. Eventually, Pharmsearch will cover French patents from 1961, and United States and European documents from 1978. The database has been kept current within 6 weeks for European patents and is being extended backwards in time. Data from 1986 is being added now. It has approximately 1500 records from 1960 to 1961 as well. WPIM goes back to Derwent week 8701 which has patents from the latter part of 1986. MARPAT has indexed patent documents dated 1988 forward.

TREATMENT OF PATENT DOCUMENTS DATED MAY 16, 1989

The remainder of this paper examines the patent documents actually covered in the three databases for a given time period. A specific date was chosen as the most manageable period. Derwent has announced to their subscribers that as of indexing week 8920, all pharmaceutical and agricultural patents (sections B and C), excluding the polymers, oligosaccharides, and polypeptides, have been covered in WPIM.[4] The data used for comparison is one from Derwent week 8923 on which U.S. patents appeared. The next section of this paper will examine the coverage of patent documents published on May 16, 1989.

Both Derwent and Chemical Abstracts index only one of the first members of a patent family that they receive. The scope of this analysis must be restricted to documents which are treated as basic in these databases.

Derwent. The WPIL File has 3957 families with at least one patent dated 5/16/89. Of these, 621 are classified in the sections included in Markush DARC (B, C, and E). Three hundred thirty-eight have code M904, indicating that they have a companion structure in the WPIM File. Of these, 124 are the basic (indexed) patent document.

0095-2338/91/1631-0040$02.50/0

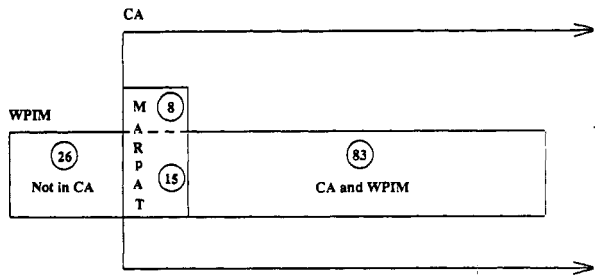


Figure 1. Overlap of coverage by WPIM and CA/MARPAT for patents dated 5/16/89.

Table I. Patent Documents Dated 5/16/89 in MARPAT but Not WPIM

Derwent section	title, patent number
B	novel complex, its production and carcinostatic agent having said complex as effective component; JP 1123803-A2
E	novel reactive flame retardant; JP 1123892-A2
	preparation of *N*-vinyl-2-oxazolidinone; U.S. 4831153
non-BCE	conductive organic polymer, JP 1123826-A2
	recovery of liquid oligomer; JP 1123807-A2
	helium refrigerator; JP 1123953-A2
	broadband antireflective coating composition and method; U.S. 4830879
	unclouded metals passivation additive; U.S. 4830730

Chemical Abstracts. As *Chemical Abstracts* does not contain family data, the only patent dates given in this file are for the indexed family member. The CA File on STN has 487 patent records dated 5/16/89. Twenty-three of these are in MARPAT.

Comparison of the Derwent and Chemical Abstracts Coverage. Figure 1 illustrates the overlap of coverage by WPIM and CA/MARPAT for this date. Of the 124 inventions in WPIM, 15 are also in MARPAT, 83 are in CA but not MARPAT, and 26 are not covered in CA at all. Eight of the 23 inventions covered in MARPAT are not in WPIM.

Each of the three regions where Chemical Abstracts and Derwent differ will now be explored. Table I gives the titles of the eight documents found in MARPAT but not in WPIM. Only one of these is in the classification sections B or C for which Derwent has attained full coverage. It claims a dextran derivative which, being a polysaccharide, is one of the classes of compounds which should be added to WPIM at a later date. Two of the eight documents are from the general chemical section E, and the remainder are not in sections B, C, or E and are thus outside of the scope of WPIM.

An inspection of the 84 patent families that are in WPIM and CA, but not MARPAT, shows that 37 of these have only specific compounds indexed in WPIM. In the Chemical Abstracts databases, they would fall within the scope of the Registry File of specific substances, and not in MARPAT. Thus, they do not pertain to this analysis of Markush patent structures. Of the 47 remaining inventions, six are in Derwent's section B and 41 in section E. The titles of the pharmaceutical inventions are given in Table II. The first three documents claim new Markush structures. The others concern the use of histamine blocker compounds in receptor-site analysis, and the novelty of two formulations.

Finally, the 26 basic patents in WPIM but not CA are scrutinized. Examination of the Derwent records shows that 18 of these are from U.S. divisional, continuation, or continuation-in-part applications that are related to older Derwent records. Ten of these are related to families where another U.S. basic patent has been indexed by Derwent. The other eight are related to at least one family which has a non-U.S. basic patent. Derwent reindexes this type of "equivalent" invention. Chemical Abstracts does not, however, appear to cover U.S. divisional, continuation, or continuation-in-part documents which are related to a patent that they have already indexed. The next section will show that, in some cases, this can be very detrimental to the coverage of the entire invention. The eight remaining patents in WPIM but not in CA contain five from Derwent section B and three from section E. Their titles are given in Table III.

Table II. Pharmaceutical Patent Documents Dated 5/16/89 with Generic Structures in WPIM Found in CA but Not MARPAT

Derwent section	title, patent number
B	methods of preparing 4-heteropentacyclic-4-(*N*-phenylamido)piperidine derivatives and intermediate compounds; U.S. 4831192
	bengamide anthelmintics; U.S. 4831135
	β-lactam antibacterial agents; U.S. 4831130
	method and test kit for analysis of histamine receptor sites of mammalian cells; U.S. 4830961
	water-washable vehicles for topical use; U.S. 4831023
	spray-drying method for preparing liposomes and products produced thereby; U.S. 4830858

Table III. Patent Documents Dated 5/16/89 in WPIM but Not CA, Excluding U.S. Division and Continuation Cases

Derwent section	title, patent number
B	methods for the diagnosis of gastrointestinal disorders; U.S. 4830010
	catheter tip polarographic lactic acid and lactate sensor for extended use in vivo; U.S. 4830011
	method for the improvement of transplantation techniques and for the preservation of tissue; U.S. 4829984
	tamper-evident container-closure assembly; U.S. 4830208
	method and device for stimulating an erection; U.S. 4829991
E	catalytic converter; U.S. 4830833
	electrochromic material, method of making, and new electrochromic device; U.S. 4830471
	cryogenic cooling system with precooling stage; U.S. 4829785

Pharmsearch. The Pharmsearch coverage of patent documents dated May 16, 1989, was also examined. The bibliographical database has 47 records with this date. Elimination of 13 records for which only specific compounds are indexed and 10 which do not have compound numbers leaves 24 records with generic structures in the Markush database. Fourteen of these are in WPIL as equivalents to older, pre-WPIM families, four are in WPIM but not in CA, three are in WPIM and CA but not MARPAT, and three are in all of the databases. This is not an accurate comparison of the databases as Pharmsearch has indexed the equivalents of some of the patents dated 5/16/89 in WPIM and MARPAT. The Pharmsearch File does not give dates for equivalent patent numbers. Time did not allow for the necessary cross-file searching to do a thorough comparison.

Summary. This study of the selection of patents documents from a single day for Markush indexing has demonstrated several items of interest to the patent searcher. Each of the three databases was found to have indexed Markush structures for documents that did not have this type of indexing in the other two databases. This is primarily due to the differences in subject, country, and time coverage of the files as well as the present situation where each database producer has found different classes of generic structures difficult to handle and is not yet attempting all of the areas that they will eventually

cover. Chemical Abstract's omission of U.S. divisional, continuation, and continuation-in-part documents also affects the scope of the MARPAT database.

CHOICE OF COMPOUNDS FOR MARKUSH INDEXING

The selection of documents for Markush indexing in the three services has been examined. The next sections will look at the choice of Markush structures from patent documents. First, the published criteria provided by each producer will be reviewed. Then, the indexing of the five documents dated May 16, 1989, that were indexed by all three systems will be examined. Each of the databases provides other tools such as indexing terms and roles which are essential in patent searching. An examination of this additional information is, however, beyond the scope of this paper.

MARPAT. The most informative statement in the MARPAT documentation relevant to their indexing criteria is that "the records contain the Markush structures found in the Claims and sometimes the Disclosure of the patent".[2] In the Markush structure, EX is used to identify specific groups from the disclosure, and SC denotes specifically claimed structures. Stereochemistry displays if the patent has used it.

Pharmsearch. The Markush structures in the Pharmsearch File follow the indexing policy of "indexing all single compounds mentioned anywhere in the disclosure, examples, and claims, and widening the specific values resulting from such indexation, using the generic terms of the Markush structure described in the widest claim".[5] The generic terms chosen from the widest claim must be "chemically related" to the specific terms. Chemical compounds are indexed in the structure file if they are[6]

(1) New compounds claimed as active components of pharmaceutical compositions.
(2) New compounds claimed as intermediates in preparation processes (synthesis and purification processes) of pharmaceutically active compounds.
(3) Known compounds which are active ingredients of new pharmaceutical compositions.
(4) Known pharmaceutically active compounds with new preparation or purification processes or with new pharmaceutical activities.

WPIM. "Rules governing coverage in the WPIM database are equivalent to the rules for coverage in the Derwent fragmentation coding systems."[7] "Coding is, in general, applied to the widest disclosure of a patent".[8] Compounds are indexed if they are[7]

(1) Described as new.
(2) Products of new processes, including materials purified in new ways.
(3) Removed, when this is important to the novelty of the invention.
(4) Used to effect removal when this is important.
(5) Analyzed or detected, when this is the novelty of the invention.
(6) Used in analysis or detection, when this is important to the invention.
(7) New catalysts.
(8) Important ingredients of compositions.

COMPARISON OF THE FIVE PHARMACEUTICAL PATENTS INDEXED BY ALL THREE DATABASES

The examination of the Pharmsearch coverage of patent documents dated May 16, 1989, given above stated that only three of the documents were also indexed in MARPAT and WPIM. A closer look at all the family members of the 15 inventions covered in both WPIM and MARPAT revealed that the equivalents of two additional patents are indexed in the Pharmsearch database. The Markush indexing of each of these five inventions will now be analyzed.

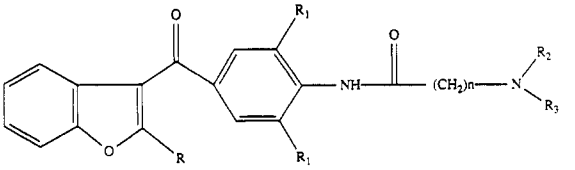


Figure 2. Common substructure of Markush structure in U.S. 4,831,054 assigned to Taro Pharmaceuticals, Ltd.

1. U.S. 4,831,054. "2-Alkyl-2-Benzoylbenzofurans Useful for Treating Cardiac Arrhythmia" assigned to Taro Pharmaceuticals, Ltd. This patent contains one Markush structure (Figure 2) which is found in the main claim. The generic variables are all unsubstituted lower alkyl. In this very straightforward situation, all three databases have selected the same compounds. The only difference is in the WPIM File where Derwent has indexed all possible alkyls and does not additionally provide the CHK[9] superatom. The other two databases have given the specific alkyl groups which are used in the document as well as the generic variable. They have not generated any additional possibilities.

2. U.S. 4,831,017. "Increased Litter Size in Monogastric Domestic Animals After Treatment in Mid-gestation of Gravid Female with Avermectin or Milbemycin Compound" assigned to Merck and Co., Inc. This invention concerns a new use for known compounds. Two Markush structures are claimed. They are the avermectins and milbemycins previously described in several patents. Both the claims and disclosure provide identical descriptions of these compounds. The avermectins are described in both specific and generic terms at each of four positions. Additionally, one bond in the structure may be either single or double. The milbemycins are characterized in a table of 11 specific compounds and three oxime derivatives of the form $=N\text{-}OR_6$ where R_6 is either a hydrogen or lower alkyl. All of the biological data in the patent is for ivermectin, a mixture of two avermectins.

MARPAT and WPIM thoroughly index both of the Markush structures. Pharmsearch, however, only covers the 11 milbemycins specifically claimed, ignoring the oxime derivatives as well as all of the avermectins and ivermectin. This omission seems to be a result of the policy to only index generic terms which chemically correspond to at least one specific compound. This approach has had a detrimental effect on the indexing of this use patent which cites several other patents as sources for the scope of its generic structures.

Aside from this policy difference, MARPAT and WPIM each commit indexing errors. MARPAT uses nPr for an "isopropyl" group, and misses the monoalkylamino option when amino, monoalkylamino, and dialkylamino are all specified options for a variable group. In WPIM, the -O- linkage of a -O-(CO)-CH(Me)-nBu group is omitted. The same error is reproduced in the Derwent Documentation Abstract for this invention.

3. U.S. 4,831,031. "Aryl Piperazinyl—(C_2 or C_4)—Alkylene Heterocyclic Compounds Having Neuroleptic Activity" assigned to Pfizer, Inc. This is the first U.S. patent to issue in its family. Another U.S. patent has, however, resulted from a divisional application. Derwent has assigned an abstract number in week 9006 for this second patent, dated 11/28/89, but has not yet indexed it. Chemical Abstracts does not have the second patent in their CA or CAPreviews Files and thus will not be indexing it. Figure 3 shows the Markush structure which has three variable parts. The indexing of each variable will now be examined in turn.

(a) Ar. This variable is only claimed as benzoisothiazole. The disclosure is much wider—covering 11 types of aromatic ring systems each with multiple isomers and attachment points.

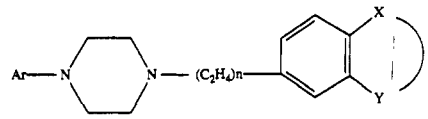


Figure 3. Common substructure of Markush structure in U.S. 4,831,031 assigned to Pfizer, Inc.

Only two of the 17 example compounds are the claimed benzoisothiazolyl.

MARPAT has the narrowest indexing, only covering the claimed benzoisothiazolyl as the sulfoxide or sulfone, optionally substituted in the benzo ring by the claimed list of substituents.

Pharmsearch indexes all of the eight ring systems used in the examples. The attachment points, and types and positions of substituents, are limited to the examples, ignoring both the broader claimed list for benzoisothiazole and the scope of the wider disclosure.

WPIM indexes the same eight ring systems from the examples, but broadens the substitution to the scope of the disclosure. The additional ring systems which are disclosed are covered by the superatoms HEF and ARY.[10]

(b) $(C_2H_4)_n$ Where $n = 1, 2$. Pharmsearch and MARPAT index this correctly, but WPIM erroneously gives C_1 or C_3 as the options rather than C_2 or C_4.

(c) Fused Ring System. Ten ring systems are claimed. The oxo form of five of these are given as examples. The wider disclosure gives 11 more ring systems, including a spiro system which is used in an example.

MARPAT covers all of the claimed ring systems with their specified substitution. The wider disclosure compounds (including the one used as an example) are not indexed for the Markush structure.

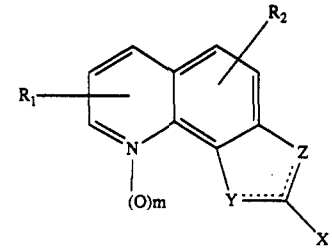
Pharmsearch indexes the example rings, including the nonclaimed spiro system. Only the optional oxo form of the examples is illustrated. The claimed optional substituents which are not demonstrated in the examples are not covered. Those that are indexed are extremely specific. The only option given for a "phenyl optionally substituted by one Cl or F" is "mBeCl".

WPIM includes everything mentioned in the patent specification. The total scope of the claims, examples, and wider disclosure is covered in the Markush structure.

4. U.S. 4,831,040. "Method of Prevention and Treatment of Peptic Ulcers" assigned to Rorer Pharmaceutical Corp. Pharmsearch indexed the European patent from this family. This invention is a new use for known compounds disclosed in "U.S. Patent No. 4,698,346 (which is incorporated herein by reference)". Figure 4 shows the Markush structure of heterocyclo-fused quinolines. The U.S. patent specifically claims three compounds, two of which are not quinoline derivatives. A wider disclosure of the corresponding benzodiazine compounds which contain an additional N in the "left" ring is given in the specification which states that "The preferred compounds include those in which the quinoline ring structure is present. The present invention also embraces the corresponding benzodiazines, e.g., 1,4-benzodiazines of the formula: ... as well as the 1,2- and 1,3-benzodiazines. The invention is illustrated by way of the preferred quinoline compounds but it is within the skill of the art to extend the illustration to benzodiazines." The example compounds are all quinoline derivatives—four ring systems and the N-oxides of two of them.

(a) Ring System (Variable Groups *m*, Y, Z). MARPAT gives all five quinoline analogues and their N-oxides as well as the NR_3 generic group. None of the benzodiazines are included. A typographical error giving NO_2 for an amino group is found in this part of the record.

Pharmsearch also covers all five quinoline analogues, but includes the N-oxides for only the two which have examples



Z = O, S, N
Y = O, S, N, NR_3

Figure 4. Common substructure of heterocyclo-fused quinoline Markush structure in U.S. 4,831,040 assigned to Rorer Pharmaceutical Corp.

of it. The variable R_3 is not included. This is most likely due to the fact that none of the examples contain a substituted nitrogen. They do not index any benzodiazine derivatives either.

WPIM indexes all possible ring systems claimed plus those in the wider disclosure. The N-oxides and all possible substitutions are covered. This gives 36 tricyclic ring systems—including all isomers of the benzodiazine analogues.

(b) X Option. Ten types of substitution are generically claimed. The examples cover all but one of them. MARPAT includes all options for all of the ring systems indexed. Only one of the specific examples is included.

Pharmsearch indexers matched the substitution options with the rings specifically containing them. Thus, not all possibilities are covered, and the variables are different for each ring system. Some specific atoms are given as well as generic superatoms.

WPIM gives the same options for all ring systems, and indexes the additional specific cases given in the examples and disclosure as well as those in the claims.

(c) R_1 and R_2. MARPAT covers the claimed generic options, but does not pick up the specifics in the examples.

Pharmsearch indexes these variable groups as unattached to avoid generating all of the possibilities. (The Pharmsearch manual indicates that this approach is taken whenever necessary.) No generic options are given that are not supported by an example. Different substitution patterns are again shown for each ring system.

WPIM's coverage is thorough. Seven specific examples are given as well as all of the generic options. All ring systems have the same data.

5. U.S. 4,831,185. "Asymmetric Synthesis of Natural Vitamin E" assigned to Hoffman-La Roche, Inc. The Derwent WPIL database has seven family records for this invention. All have a U.S. patent as the basic document. Only two of the families are old enough to have Markush indexing at this time. Pharmsearch indexes the European patent for this invention. MARPAT has covered only this U.S. family member. The others are not in the CA or CAPreviews Files, indicating that they will not be indexed. This U.S. patent claims the process for resolving a racemate. The additional U.S. patent indexed by Derwent and the Pharmsearch indexing for the European Patent give other intermediates from this synthetic process. This U.S. patent document does not clearly indicate which of these compounds might be claimed in further divisions. By not indexing any of the additional U.S. patents in this family, Chemical Abstracts will not cover any of these additional structures. The Markush structure important to this patent's claim is shown in Figure 5.

(a) OR. The OR group is claimed as a "hydrolyzable ester protecting group". The disclosure provides some very general examples of preferred groups.

MARPAT shows its variable group "R"[11] and provides the textual description "hydrolyzable ester protecting group". The

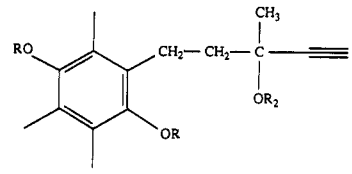


Figure 5. Common substructure of Markush structure in U.S. 4,831,185 assigned to Hoffman-La Roche, Inc.

generic definition from the disclosure is indexed as well as the specifically claimed acetyl.

Pharmsearch indexes acetyl and the superatom PRT[12] only.

WPIL does not index the very general PRT superatom but provides all of the disclosure definitions and acetyl.

(b) R_7. MARPAT and WPIM cover all of the options given in the specification for this variable.

Pharmsearch limits its choices to the narrow scope of the examples. A substituted benzyl group is covered only as its exemplified ortho isomer.

CONCLUSIONS

Even though this analysis is performed over a very small set of five patent families, several trends in the handling of Markush structures in the three databases are revealed. INPI's focus in Pharmsearch on specific examples is found to be at the expense of the coverage of entire aspects of an invention. This was seen quite vividly in the Merck use patent where the entire series of avermectin compounds was not indexed, even though the only compound for which biological data was given was an avermectin. The reliance on examples has also led to optional substituents being presented as required if they are illustrated in all of the specific compounds. Chemical Abstracts, despite their stated scope of Markush structures in claims or disclosure, neglected the wider disclosure information in the Rorer and Pfizer patents. The file sample studied is too small to determine if this occurs with any consistency throughout the entire database. MARPAT was weaker than the other databases in its coverage of specific examples from the Rorer patent. The omission of the data from U.S. patents resulting from divisional, continuation, or continuation-in-part applications was shown to impair Chemical Abstracts' coverage of inventions which fall into this category. Derwent's long experience in patent indexing carries over to their Markush file. In both of the cases where the disclosure was wider than the claims, WPIM had a greater depth of indexing than the other two databases.

In addition to looking at how the indexed documents are handled, it cannot be forgotten that each of the databases covers patent families that one or both of the others do not include. This is the result of differences in subject scope, country coverage, time periods covered, policy for handling equivalent documents, and the current inability to handle some structure types. Despite their thoroughness, all producers of this information were also shown to be subject to human-generated errors in the Markush structures which appear in the files. All of this needs to be taken into consideration as the patent searcher chooses which combination of these databases needs to be checked for a given search and, importantly, how the search query is phrased for each chosen database. A query used to search the Pharmsearch database may need to be more generic than one used in the other files in order to retrieve the more specifically described variable groups. The additional indexing tools provided with each Markush database will also influence the final approaches taken in a patent search.

REFERENCES AND NOTES

(1) The term "patent" is used in this paper as a synonym for "invention". It is used loosely to refer to any type of patent document (both granted patents and published patent applications).
(2) *MARPAT Database Description Sheet*, revision April 1990. STN International, 1990.
(3) Information on treatment of equivalents obtained during phone call to Questel, Inc.
(4) *Chemical Patents Index and Electrical Patents Index Annual Meetings—1990.* Derwent Publications Ltd.: Quebec City, Quebec, Canada, 1990; p 19.
(5) *Pharmsearch Database.* Institut National de la Propriete Industrielle, undated, pp 4-5.
(6) *Markush DARC Workshop Manual.* Jan 1989 ed., 1989, p C-2-6.
(7) Ibid. p C-1-6.
(8) *Central Patents Index Chemical Retrieval Manual.* Sept 1981 ed., Derwent Publications, Ltd.: London, England, 1981; p I.M.3A-p.0.2.
(9) The CHK superatom in Markush DARC is defined as "alkyl, alkylene". "All members of this group are carbon chains, in which each carbon atom is bonded to other carbon atoms or hydrogens." (*Markush DARC Workshop Manual.* Ref 6; p C-3-2).
(10) The HEF superatom in Markush DARC is defined as "fused heterocycle, including all ring sizes and degrees of unsaturation". ARY is "monocyclic or fused carbocyclic system containing at least one benzene ring". (*Markush DARC Workshop Manual.* Ref 6; pp C-3-2,3).
(11) The R node is used in MARPAT File structures for "unstructurable groups described in the patent, e.g., "organic group," "electron withdrawing group." Whatever phase the patent used to describe the R group appears in the TX (TeXt) field after the R". (*MARPAT User Guide.* April 1990; Chemical Abstracts Service: Columbus, OH, 1990; p 49.)
(12) The PRT superatom in Markush DARC is defined as "protecting group" (*Markush DARC Workshop Manual.* Ref 6; p C-3-3).