The Development and Evaluation of a Full-Text Drugs Database Martindale Online

James E. F. Reynolds

Thesis Submitted for the Degree
of
Doctor of Philosophy

The City University
Department of Information Science

December 1987

CONTENTS

APPENDICES

FIGURES AND TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge the support of The Pharmaceutical Society of Great Britain in this work, the guidance of Stephen Robertson during his supervision of the research and the advice of [redacted] of The City University Statistical Laboratory.

Grateful thanks are due to: [redacted] for providing test questions, search statements and relevance assessment; to [redacted] who carried out some of the clerical work; and to all my Martindale colleagues without whom the database would not have been produced.

Finally I would like to thank my family for their forbearance during this project when so much else had to be left undone.

ABSTRACT

The Development and Evaluation of a Full-text Drugs Database: Martindale Online

Martindale Online is a full-text database on drugs produced from a structured neutral database that is also used to produce a print product. Special characteristics of the database include a hierarchical record structure and a facility for linking records within the same hierarchy. The development of this database is described.

Investigation at the development stage indicated a need to index the database and this was carried out using descriptors from a specially designed thesaurus. To evaluate the effect of this indexing, three information pharmacists selected 98 queries for an assessment of retrieval effectiveness; they and the author formulated sets of search statements that were used to search the file in several different ways. It was found that searching the indexed database via descriptors and free text (when appropriate) produced significantly better results, as judged by scores that incorporated precision and recall, than searching either the indexed or the unindexed database solely in a free-text manner.

As there was evidence that searchers were slow to make use of the descriptors, highly structured search statements were created for each query using all the details from the relevant sections of the thesaurus and these statements were tested on the unindexed database. While this test produced some conflicting results, it did suggest that as far as major relevance was concerned such a method of searching might be effective with Martindale Online and is worth exploring further, especially with a view to producing a front-end system.

Detailed failure analysis was carried out on the searches performed in the recommended manner. With the information pharmacists' search statements the database was operating at a recall ratio of 60.2 for all relevant records (69.3 for records of major relevance); with the author's statements the recall ratio was 65.4 (73.2 for major relevance). Corresponding precision ratios were 63.5 (58.3 for major relevance) for the information pharmacists and 67.5 (59.6) for the author.

The largest cause of both recall and precision failure was in limitations of the search statements whether produced by the information pharmacists who had varied experience of Martindale Online, or by the author who has a detailed knowledge of the system and the contents. Limitations in the indexing also accounted for both types of failure; account has already been taken of these limitations and modifications have been made to some of the indexing guidelines.

CHAPTER 1 INTRODUCTION

The drug information world is rich in many ways. It has the benefit of a large growing primary literature base[1, 2], it has several large secondary services such as MEDLINE[3], EMBASE[4], and International Pharmaceutical Abstracts[5], and it has many well-used tertiary sources such as Martindale[6], AHFS Drug Information[7], and the British National Formulary[8] to name a few that not only survive into new editions or updates, but develop with each edition. This world also has a large and varied population with many different yet intertwined demands. It is a world where the intermediaries may be information scientists, practising pharmacists, prescribers or other clinicians; where the end-users may be pharmacists, physicians or patients; and where the producers may range from international conglomerates to individuals. Most important of all, it is a world in which it is recognised that information is essential to the effective use of drugs and medicines[9 - 14].

This study concerns the development of one drug information source, the online database version of Martindale, and an evaluation of current and potential methods of retrieving data from it.

Martindale and the database, Martindale Online, are described in detail in succeeding chapters. The book is in its 28th edition, with the 29th edition soon to be published. It consists of monographs on drugs in clinical use worldwide and these monographs are made up of facts, evaluated text, abstracts and references. Each edition is produced within an approximate 5-year cycle. The database treats each monograph as a hierarchical set of records and each record has been indexed with descriptors from a tailor-made thesaurus. It was intended to update the database every 6 months, but updating now takes place every year.

The preface to the 28th edition of Martindale defined the aim of the book in the opening paragraph;

"Ninety-nine years have passed since William Martindale produced the first edition of The Extra Pharmacopoeia - now better known as 'Martindale'. It was his aim, and it is still our aim, to provide a concise summary of the properties, actions, and uses of drugs and medicines for the practising pharmacist and medical practitioner.

However, over these years his "little book" has so grown that now it covers most of the drugs in clinical use throughout the world."
The preface also introduced the database:

"... readers.... will be able to pluck from Martindale Online sections of monographs that answer their specific questions in much the same way as some imaginary reader could answer questions if he had memorised comprehensively the whole of this edition. Martindale Online will be updated regularly with newly revised chapters while the book will continue with the current cycle of publication."

1.1 TERMINOLOGY

East[15] among others has discussed the terminology of nonbibliographic databases. Within that terminology Martindale Online can be considered to be a full-text database in that it contains the whole text of the editorial matter of each introduction and monograph. It can also be considered to be a bibliographic database, as it supplies references and summaries of journal articles. Thirdly it can also be considered to be a factual database, in that it supplies preparation data and manufacturers' names, synonyms and pharmacopoeial data. The author, however, tends to consider Martindale Online as predominantly a full-text database, even though the bibliographic records are in the majority, and it is generally treated as such throughout this thesis.

1.2 MARTINDALE IN DRUG INFORMATION

Several surveys have been carried out on the resources of drug information centres. Dombrowski and Visconti[16] in an audit of US drug information centres obtained information from 98 of 121 centres and found that 96.9% carried the 28th edition of Martindale. Rosenberg *et al.* have regularly analysed drug information centres in the US. Their latest report[17] covers 127 of the then estimated 191 US institutions with drug information centres; all 127 carried copies of the 28th edition of Martindale. Another earlier survey[18], international in scope, reported that 90% of 35 responding centres carried the 27th edition of Martindale.
The value of Martindale as a drug information resource was reported by Cardoni *et al.*[19] and by Couper[20]. However, Cassidy and Kostrewski[21] found that it was not as valuable as some other sources in providing information on household product poisoning.

A survey of varied readers, or rather purchasers, of the 27th edition of Martindale carried out by Marcus[22] identified the importance of Martindale to her sample of users. Her survey also indicated that they were most interested in adverse effects, precautions, contra-indications and interaction data. In addition she indicated areas in the book that readers would like to see improved. Briefly, improvements were requested in:

information coverage,
information arrangement,
information evaluation,
indexing.

These are areas that are under surveillance in the revision of each new edition and, in the case of the second and the last points, they are areas that might well be improved in a database version.

Miles[23] carried out a detailed study of the effectiveness of Martindale in answering a range of questions from a London drug information centre. She compared the 27th and the draft of the 28th edition of Martindale with 3 other tertiary sources (AHFS Drug Information[7], the British National Formulary[8] and the Data Sheet Compendium[24]) and with 3 secondary sources (the Iowa Drug Information Service[25], the National Abstracts Service[26] and MEDLINE[3]). She found that Martindale as a book was better than the secondary and other tertiary sources for a variety of questions, especially those dealing with availability, identification, pharmaceutics, toxicology and pharmacology. It was also the best of the tertiary sources for questions on therapeutics, administration and interactions. On overall performance, she placed Martindale in an intermediate position between quick look-up reference books and secondary information sources such as MEDLINE. She also indicated that the database version which was then being prepared would increase the value of this information source by providing information that was more up-to-date and information that was more readily accessible.

1.3 MARTINDALE ONLINE IN DRUG INFORMATION

The above proposed advantage of greater accessibility was borne out by Green[27] who in comparing Martindale Online with Drug Information Fulltext[28], which is the online version of AHFS Drug Information[7], also compared Martindale against Martindale Online. She found for her sample of questions that Martindale and the online file performed equally well,

except that Martindale Online had the advantages of inclusive searching and of being faster on searches on broad topics. This study was carried out before the database had been updated, so any benefits from updating could not be assessed.

Another evaluation of Martindale Online was carried out by Essex *et al.*[29]; they provided little detail of their experiments carried out within an information department in the pharmaceutical industry, but they did emphasise the value of Martindale Online, especially for broad searches.

Other full-text drug databases are available and being produced. Some such as the Merck Index[30] or SEDBASE[31] are available on interactive vendors' systems and can be searched using boolean operators. Some, like the database from Edinburgh on the management of poisoning[32] or the protype VADIS[33] which provides summary drug data, are available through viewdata systems. Other databases are available on compact discs and like the Micromedex group of databases[34] are menu-driven. Drug regulatory bodies[35] and the pharmaceutical industry[36] have their databases, while various hospital and general practice systems incorporate as yet small databases in their software[37]. Without even taking account of the knowledge systems that have or are being developed, it is obvious that there is a wide range of types of databases and search systems. Some attempt has been made at coordinating or standardising some of these databases[38], but more needs to be done[39].

The only comparison of Martindale Online with any other database is that carried out by Green[27] who found that Martindale Online and Drug Information Fulltext were complementary for a range of questions obtained from British and American drug information centres.

1.4 DATABASE DESIGN

In a useful review of the design and marketing of databases East'[5] intentionally did not provide details of the technology since that is rapidly changing and is market driven. While that is the case, it has to be pointed out that any database producer wishing to distribute a database through the traditional vendor system is severely restricted in any changes and enhancements that he may wish to make to the standard system. Unless the producer can demonstrate that his database is going to overshadow all the others in a vendor's portfolio, then it is

foolhardy to consider that changes will be made to the working system for any one database. Indeed the vendors stress that it is to the benefit of their searchers that all their databases conform to a recognised pattern.

1.4.1 TO INDEX OR NOT TO INDEX

A decision had to be taken in the design of Martindale Online on whether it should be indexed with descriptors or left unindexed. The factors that affected this decision are discussed in the next chapter; here it is worth considering the background to this problem.

Studies like Cranfield 2[40] and the Aberystwyth index languages test[41] as well as later work[42] had demonstrated that free-text searching could provide as good results as searching with a controlled vocabulary. Yet in the field of drug or medical information there are outstanding examples of indexed databases like MEDLINE that continue on their indexing road, and new full-text databases like Drug Information Fulltext and Sedbase, in addition to Martindale Online, have been launched with a controlled searching facility. There have been pleas that thesauruses should be viewed as search tools[43]. There have been commercial publishers adding descriptors to established databases as a user aid[44]. A review of 57 databases that offered free-text searching and were available on DIALOG showed that 40 (70%) also provided descriptor searching[45]. Studies like Henzler's[46] and Markey *et al.*'s[47] pointed to the value of a mixture of free-text and controlled searching, aiming for the high recall associated with the former and the high precision of the latter. While Hersey *et al.*[48] demonstrated improved recall and relevance with controlled searching.

Arguments for controlled indexing were put by Duckitt[49] when working on the Martindale Online project. Duckitt argued in the context of full-text databases that controlled indexing was useful in the location of information, it allowed a more natural thought approach and it allowed for syntax indexing. However, she also made the point that free-text searching of full-text databases might produce better results than such searching of bibliographic databases. Dubois[50] enlarged on these arguements some years later.

Harter[51] also pointed out that little comparative research on searching with natural language versus controlled language had been carried out with the very large databases and recommended that the best approach

should be one that is best for a particular problem with a particular database. This unfortunately means that a database producer still has to invest in the effort of providing the controlled language within the database. Harter summarised the views on the place of both types of searching. Controlled vocabularies should be used in searching if

- high precision is required without high recall,
- the subject matter is well defined,
- generic searches can be carried out.

Natural language searching should be used on the other hand if

- subtle nuances or the "specifity and expressiveness" of natural language are important,
- the terms for the concept do not exist in the control vocabulary,
- a comprehensive search is required,
- the literature is poorly defined.

There is some slight conflict in these arguments for controlled language searching. Also there is a suggestion voiced by both Duckitt and Harter that full-text databases might have their own problems because of a greater chance of false drops. The author is only aware of one study comparing the two types of searching in a full-text database which is the one carried out by Mackay on Martindale Online shortly after it became available[52]. One hundred questions of various types from a hospital drug information centre were used to test free-text or uncontrolled searching versus descriptor searching. Descriptor searching was shown to give the better performance. This was an encouraging result, but was based only on precision and the relevant documents retrieved by each search. There was no attempt to find unretrieved relevant records which limits the value of the study as 20% of the searches produced no hits.

Thus, while there is some evidence and attractive arguement in support of controlled searching, it has to be considered as slight in comparison with the evidence showing no real benefit from such indexing. More detailed studies are therefore required with full-text databases. Such studies should be carried out in the real world using facilities and systems available to the general online searcher. However, they should also take some account of the developments that might affect the way searching is done, such as front-end systems to aid the search process.

Expert systems are being devised to improve controlled searching by guiding the selection of terms from thesauruses like MESH[53]. Other systems have been designed for the computer selection of terms from dictionaries and textbooks[54] to improve search statements for free-text searching. There is already a basis for expansion of a search statement for Martindale Online - the Martindale Thesaurus. Rules could be devised whereby a front-end system extracted all terms corresponding to concepts in a question by selecting descriptors, exploded descriptors, all related terms and all lead-in terms. How would such a search statement, more detailed than one that would ever be produced manually, perform in Martindale Online? Salton and Lesk's[55] work suggests that it might not perform very well. However, they used the hierarchies in a thesaurus to expand the documents in their database as well as the search statements for their 34 queries on computer science. As in the proposed study the thesaurus would not be used to alter Martindale Online once indexed, then Salton and Lesk's findings may not be applicable here. Certainly for the future development of Martindale Online, it would be useful to know how such search statements performed against the indexed and the unindexed file.

1.5 FAILURE ANALYSIS

While much of the research mentioned above has been on optimum methods of retrieval, few studies have involved the detailed analysis of the response to each search statement. Why did this search retrieve only 12 relevant documents and at the same time retrieve 10 irrelevant ones? There is, however, one outstanding study on failure analysis in the field of medical information - that carried out by Lancaster[56] on MEDLARS. The findings of that evaluation were used to improve the service[57]. Just such a form of quality control is required for Martindale Online where the philosophy of improving the content has to be extended to improving the online service. Being a much smaller database, it might be possible with Martindale Online to carry out an analysis on all recall and precision failures rather than just a sample as was the case in Lancaster's study.

CHAPTER 2 MARTINDALE, THE EXTRA PHARMACOPOEIA

2.1 HISTORY

The Extra Pharmacopoeia was first published in 1883[58]. William Martindale expanded his personal notes to produce his 'little book' on the chemistry, pharmacy and activity of drugs that were extra to the then current British Pharmacopoeia, the latest edition of which had been published 16 years earlier in 1867[59] and reprinted with a bare 17 pages of additions in 1874. With the aid of W. W. Westcott, he incorporated selected abstracts of relevant papers that appeared in the important medical journals of the day. Six months later Martindale and Westcott produced a second edition. Altogether, they produced 10 editions before William Martindale died in 1902. His son W. H. Martindale carried on the work, initially with Westcott's assistance, and he also produced 10 editions before his death in 1933 by which time Martindale's Extra Pharmacopoeia had established itself as an indispensible reference book for pharmacy[60]. It had grown in size from 313 pages to 1975 pages in two volumes as its scope had widened to cover a range of topics such as relevant legislation and analytical data, as well as materia medica. Rather than let this reference book die, The Pharmaceutical Society of Great Britain bought the copyright and appointed one of its staff as editor[73].

Up until the 1960s The Society continued to produce editions of The Extra Pharmacopoeia according to the pattern laid down by Martindale, father and son. However, with the 25th edition edited by R. G. Todd[61] this changed as the contents of the book were altered to direct it more at the practising pharmacist and physician by concentrating on the therapeutic aspects of the drugs. The analytical data was dropped, as was much of the legal information, and a decision was made to increase the coverage of drugs available in other countries. A new pattern was set which is still being followed today. This 25th edition established Martindale, The Extra Pharmacopoeia as a standard reference work on drugs[60] that has become known simply as Martindale. Coverage of the world's drugs has increased in subsequent editions, as has the depth of the information, as the role of the pharmacist has evolved to that of being a provider of drug information to patients and physicians rather

than a compounder of medicines[1-4]. It is a legal requirement in some countries that pharmacists possess a current edition of Martindale in their practices.

2.2 DESCRIPTION

The current edition of Martindale[5], which is the edition that formed the starting point for Martindale Online, is organised into sections as shown in Table 1.

Preface	definitions of terms of reference
Part 1	chapters containing monographs on groups of major substances of medical and pharmaceutical interest
Part 2	alphabetically arranged groups of monographs on new or minor substances
Part 3	descriptions of medicines sold to the public in the UK
Indexes	directory of manufacturers index to clinical uses index to identity numbers index to drug names and preparations

Table 1 Format of Martindale

The bulk of Martindale consists of drug monographs, most of which are grouped into categories (Part 1). Monographs on old drugs, investigational compounds or toxic substances are included in alphabetical order in a section called Part 2. A third small section lists the ingredients of proprietary medicines that can be purchased 'over the counter' in the UK without a prescription. The remainder of the book is made up of a number of indexes; these include a brief index to clinical uses mentioned in the text and a detailed index of drug names.

Each monograph normally conforms to a standard pattern in which the features shown in Table 2 can be described.

nomenclature
pharmacopoeial data
physical and pharmaceutical properties
units
adverse effects and their treatment
precautions, contra-indications and interactions
pharmacokinetics
actions
uses
administration and dosage
preparations

Table 2 Contents of a Martindale Monograph

The range of information varies for many compounds. Sometimes there is little or no information for some aspects, while for others there is difficulty in trying to encompass a large amount of information. The separate sections of a monograph that describe the above features can thus vary from a few sentences to several thousand words. Sometimes, the information within a section is best conveyed by referring to another compound. For instance, salsalate has similar adverse effects to those of aspirin. Rather than repeat all these effects in the salsalate monograph, cross-reference is made to the appropriate part of the aspirin monograph and only adverse effects specific to salsalate are described wihin its own monograph. Grouping the monographs into chapters helps this cross-referencing. It also gives the writers the opportunity to describe the properties of the relevant groups within the chapter introductions. Interestingly, Miles[23] reported on the benefits of this cross-referencing system, while Marcus[22] reported that some readers did not like the feature.

Information within the monograph and an introduction takes two forms: editorial text and bibliographic material. The editorial form consists of authoritative statements written by the editorial team from a review of available published data. These statements are supplemented by

abstracts and references prepared and selected by the team from major publications and so maintaining the tradition set by Westcott in 1883.

CHAPTER 3 MARTINDALE ONLINE

There were two stimuli to the investigations into creating a databank for Martindale. The first was an awareness of the use of online services for drug information and the second was the growing use of computers in the printing industry.

The initial planning of the contents of the 28th edition took place in 1977 to 1978. That edition was to be published in 1982 and would be used by pharmacists, physicians and drug information specialists until 1987 or 1988 when the 29th edition would be published. Consideration therefore had to be given to the forms in which drug information would be required over that period.

Informal approaches from librarians in the US indicated that such a group would welcome an online version of Martindale, especially if it was to be updated more frequently than the printed book. Discussions with four of the major vendors revealed that they would be willing to undertake the distribution of such a file, but their estimates of expected use were no more than guesses. There were thus only pointers to potential use of an online file; any use would initially be by those already experienced in online searching, but ultimately it was considered that use might extend to pharmacists and doctors who were then just beginning to use computer aids in their practices. However, it was considered that setting up an online file could not be justified unless it was associated with some other cost benefit such as a saving in processing or printing costs.

At that time several printers had computer driven phototypesetters that could generate master pages or plates from tapes of coded data. For a book like Martindale of 2 to 3 million words, this generation might take 3 or 4 days. Preparing text for Martindale using this technique might produce savings over traditional setting and correction techniques and such savings might justify the setting up of an online file. Having arrived at some justification, assistance was sought from those experienced in database design and text processing. Financial assistance was also sought and obtained from the EEC and the Department of Trade and Industry; this assistance is gratefully acknowledged.

3.1 SYSTEM INVESTIGATION

In 1978 there were few organisations capable of providing the appropriate services. Database management systems such as ASSASSIN were rejected since a text processing system would not easily link with it. Printers with effective text processing systems did not provide database services that could be used in preparing an online file that was not just an electronic copy of the book. Eventually two organisations were found who could provide a double service and one of these, Peter Peregrinus Limited (PPL) of the Institution of Electrical Engineers, was selected to carry out the system investigation.
PPL's approach was to recommend the setting up of an intermediate or neutral databank that could be processed by them in two ways. In one way, a suite of typesetting programs would convert the data to a form suited to the APS 4 phototypesetter. In the other way, a second suite of programs would convert the neutral database to conform to ISO 2709 to produce an interchange tape that could be read by online vendors and then processed by them according to instructions or negotiations. It is interesting that a similar approach has been recently adopted by another publisher for the integration of their data[62].

3.1.1 NEUTRAL DATABASE

In considering a suitable unit of information, the chapter was discarded as being potentially far too long and unsuited to the management of Parts 2 and 3. The monograph was also considered unsuitable for reasons of size. PPL therefore proposed that a monograph be split into several record types as shown in figure 1. Martindale was not to be treated like a normal full-text file, or what was to be the normal treatment for a full-text file as exemplified by a prototype of a full-text file of the Journal of Medicinal Chemistry then being demonstrated by BRS. Under the BRS scheme a monograph or perhaps a chapter would be considered as a record, and searching would usually be across all sections of a record. PPL's proposal was considered to offer more flexibility for development of the database and was accepted.

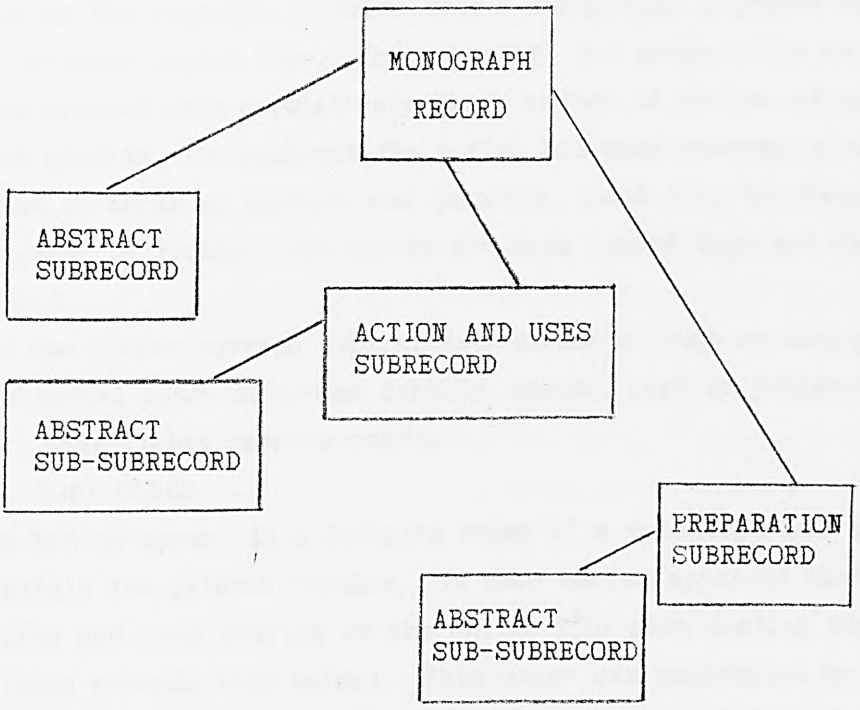


Figure 1 Hierarchy of Martindale Records

The hierarchical record structure, shown above in Figure 1, used the pharmacy section of a monograph (covering nomenclature, pharmacopoeial data, and physical and pharmaceutical properties) as the main record, and this was called the monograph record in the neutral database. A bibliographic-type record associated with this pharmacy section was treated as an abstract subrecord. All the different editorial text portions were treated as action and uses subrecords and the abstracts and references associated with each different text subrecord formed sub-subrecords. Preparation information was assigned to a preparation subrecord and associated bibliographic material was treated as an abstract sub-subrecord.

There was virtually no restriction on the number of records. Allowance was made for up to 1 million monograph records, i.e. up to 1 million drug substances. There could be up to 10 000 abstract subrecords, 130 000 action and uses subrecords each one having up to 10 000 abstract sub-subrecords, and 130 000 preparation subrecords with each having up to 10 000 abstract sub-subrecords. This meant that a monograph could have a maximum of 2.6×10^{13} records; a limit not likely to be exceeded.

Analysis of the contents of Martindale's monographs produced sets of fields for each record type. The monograph and preparation records could be divided into a relatively large number of fields, as could the abstract records. In contrast the action and uses records, although important in terms of content and quantity, could only be divided into a small number of fields. The fields for each record type are shown in Appendix A.

Most of the fields covered typical data elements, such as monograph title, chemical name, molecular formula, author, year of publication. However, some fields deserve comment.

3.1.1.1 SORT CODES

Records had to appear in a definite order if a monograph was to make sense within the printed version. It also became apparent that the order also had some bearing on the online file when dealing for instance with linked records (see below). This order was controlled by a sort code field which contained a number to denote the position of that particular record within its hierarchy. Sorting records to put them into monograph order thus involved a series of sorts for records of each record type

3.1.1.2 LINKS

One of the requirements of the system was an ability to handle linked sets of data. The linked data could be of two kinds. The first involved bibliographic material where the aim might be to provide reports of two conflicting studies that should be seen together or to list a set of references. The other kind of link involved the description of the actions of a drug that could be used in different forms. Pentazocine exists as pentazocine, pentazocine hydrochloride and pentazocine lactate. While these three forms have different physical and pharmaceutical properties, their activity in the body is that of pentazocine. This would be handled editorially by linking the three compounds so that their pharmacy would be described in individual sections, but the action and use information would be presented as applying to all three.

Within the database structure these two kinds of links affected the abstract records and the monograph records and the solution for both cases involved the use of two fields within each record type. In the abstract record, the first abstract of a linked set would contain in one field a list of the identity numbers of all the other abstract records

that followed on after it within that set; the remaining abstract records each contained in another field the identity number of the head of the set. A similar pattern applied to linked monograph records, except that the last monograph record in a series was considered to be the key record in the set since that record carried all the subrecords and was considered to be the only monograph record with a complete set of subrecords. Thus a monograph record appearing without any subrecords (other than abstract subrecords), but linked to another monograph that had some subrecords, carried in a field the identity number of that complete monograph. The monograph record of the complete set of records carried in a designated field the identity numbers of all the monograph records linked to it. This is illustrated in Figure 2.

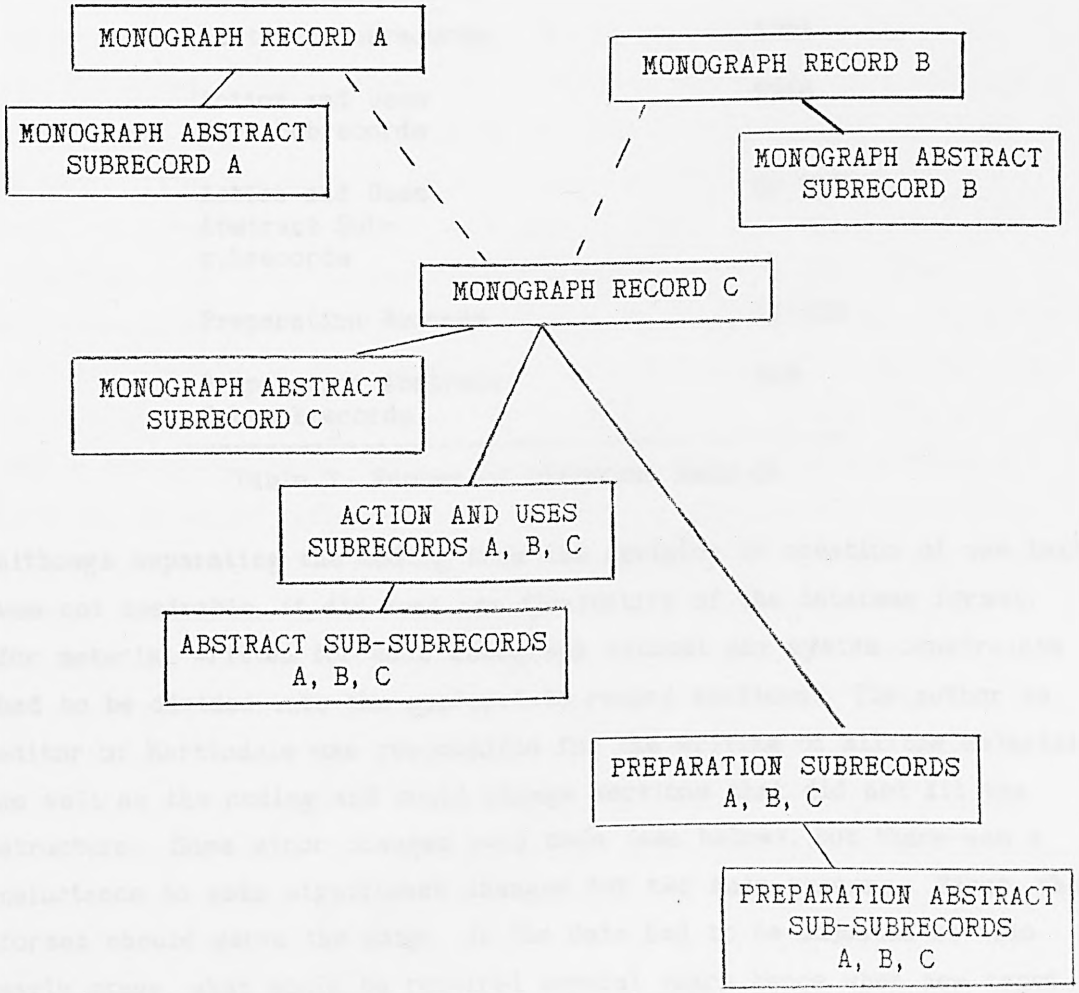


Figure 2 Linked Monograph Records

3.2 IMPLEMENTATION OF THE DATABASE DESIGN

Having accepted the design and directed modifications to handle the variety of data formats in Martindale, the next task involved coding the data into records and fields. Because of the size of Martindale and the publication schedule, revision of monographs for the 28th edition was well underway before the database design was completed. Coding was therefore treated as a separate exercise from revision. A small group of the editorial team coded 106 chapters and 5138 monographs into 82 443 records that amounted to 55.5 megabytes of data. The number of different record types is shown in Table 3.

Monograph Records (includes 106 chapter introductions)	5244
Abstract Subrecords	3061
Action and Uses Text Subrecords	8986
Action and Uses Abstract Sub-subrecords	53 515
Preparation Records	11 122
Preparation Abstract Sub-subrecords	515

Table 3 Number of Different Records

Although separating the coding from the revision or creation of new text was not desirable, it did test the flexibility of the database format, for material written for each monograph without any system constraints had to be divided into the appropriate record sections. The author as editor of Martindale was responsible for the writing of all the material as well as the coding and could change sections that did not fit the structure. Some minor changes were made (see below), but there was a reluctance to make significant changes for two main reasons. First, the format should serve the data. If the data had to be adjusted at this early stage, what would be required several years hence when new types of data might have to be handled? Secondly changes to the data required more time than was available. Changes to the text made at checking or editing were made following an analysis of the background

documents and with an awareness of the implications of any changes on the rest of the text and that could only come from detailed knowledge of the contents of that group of monographs. Making changes without knowing why every word was included has caused several errors in the past.

The database structure stood the test well. Minor changes had to be made by introducing some new fields, for instance a field to cover text appearing after the pagination of a bibliographic citation, or by expanding the coverage of some operations to a wider range of fields - the use of delimiters for allocating words to the index file for the book was extended to a wider range of fields. Further changes had to be made when it came to assigning descriptors, but that is described later.

3.2.1 CHANGES REQUIRED AT CODING

Allowable changes involved portions of the text that were not in the standard format. These anomalies usually arose because of a shortage of material and the correction seldom involved any changes to the sense of the text. Problems were to arise later because of too much data in a field, but that did not affect this coding stage. Sometimes there was not enough data to write any evaluated text but there were some references that could be included. Within this database, this meant that a subrecord would be created that did not have any parent record. The solution in such an instance was to create a dummy record with a heading but no editorial text. This happened with noxythiolin where there was no sound information for a text statement about its adverse effects, but there were references proposing and then dismissing it as a cause of adhesions. An adverse effects text subrecord was created; it contained the heading Adverse Effects but no other textual matter. The two references were coded with their abstracts as linked adverse effects abstract sub-subrecords. This would be acceptable in the printed book. The dummy record would look slightly odd in the online file; it could not be excluded since it was required for the record hierarchy, but the chance of retrieving it would be reduced by not assigning descriptors to it. The relevant fields of the three records are as follows:

```
!000        2262-a1-n
!301        Adverse Effects
!302        10
-------------------------------------------------
!000        2262-a1-1-v
!202        10
!206        Widespread filmy plastic-type adhesions
            developed in 2 boys who had had peritoneal
            lavage with noxythiolin during surgery;
            a causal relationship was suspected._
!208        Morris, M.A.
!209        letter
!211        Br.med.J.,
!212        1977
!213        1,
!215        1355
!236        2262-a1-2-s
-------------------------------------------------
!000        2262-a1-2-s
!202        20
!206        Clinical experience and a number of
            published papers did not support the
            suggestion that noxythiolin caused
            adhesions._
!207        2262-a1-1-v
!208        Rosin, R.D.
!209        letter
!211        ibid.,
!215        1644
-------------------------------------------------
```

A more common and more troublesome variation on the above problem occurred with preparations that contained ill-defined ingredients or were preparations of substances that could not be defined within the monograph structure.

Protective sprays used as skin dressings were typical examples. Details of these dressings were to be included in the chapter on Dermatological Agents, but for a number of reasons their ingredients could not be used to form Martindale monographs. Once again a dummy record was created, in this case a monograph record, so that each preparation could be coded as a preparation subrecord. The relevant fields of the dummy record and the first preparation record are displayed on the following page.

```
!000      1663-z
!101      D1M
!102      730
!103      Some Proprietary Protective Materials
!104      1
-----------------------------------------------------
!000      1663-n1-v
!502      10
!504      Nobecutane Spray
!507      Code for Astra,UK
!512      An aerosol wound dressing consisting of
          acrylic resin equivalent to total solids 5.7%,
          ethyl acetate 27.%, and propellents.  When
          applied to the skin and allowed to evaporate
          it leaves a tough elastic film impervious to
          bacteria and other contaminants
-----------------------------------------------------
```

3.2.2 CODING ACCURACY

Coding so many records was tedious, yet called for concentration in recognising the different components of a record. The coders' task was not made any easier by them having to mark each field with its number, especially when a numbering system had been adopted where each set of fields within the different record types was numbered from zero (see Appendix A). Field 12 in the monograph record, for example, was used for CAS numbers while in the abstract subrecord field 12 was used for the publication year. Although it has to be accepted that there were some advantages to this numbering system in that field 0 in each record type contained the identity number and other fields that had similar content in each record type carried a common field number, but such benefits did not outweigh the disadvantages. Later developments involved the conversion of field identifiers to the less confusing numbering scheme that is shown in the above records.

Another handicap for the coder was the automatic assignment of some data elements to several fields. Some fields always ended with standard punctuation and in such fields the punctuation could be added by program; other fields had to have their punctuation put in by hand. Coders found this confusing and would have preferred an 'all or none rule'. Adding punctuation in error caused duplicate punctuation to be presented and programs eventually had to be written to strike out any duplicate punctuation. In other cases, headings were added by program

and the coders had to remember to strike out any headings from the records, otherwise duplicate headings would be displayed.

Coding involved not only the allocation of prepared material to records and fields and the creation of identity numbers and sort codes, but the marking of linkages and the generation of various codes. Those fields that showed the links contained either one identity number or a series of identity numbers. The process of filling those fields with their numbers should have been straight-forward, but was prone to error. Numbers were found to be in error and sets were often incomplete. The field that identified the chapter to which a monograph would be allocated contained a code for that chapter; there were surprisingly few errors in the allocation of this code. Another code was used to describe the manufacturers of proprietary medicines and unfortunately this code was subject to considerable error. The idea behind the manufacturers' code was to make it easier to handle changes in manufacturers names; instead of having to correct a large number of records that contained the renamed manufacturer, only one change would be required in the look-up file that was used to translate the codes at one of the late processes before publication or release of an online tape. This facility has made corrections to manufacturers' names a simple task, but at the cost of some errors having crept in to the public database.

As the neutral database had to be capable of yielding Martindale Online and Martindale the book, data had to be included that was appropriate for one product but not the other. Cross-references for example were useful for the printed version but irrelevant to the online file. Records completely irrelevant to one version could be coded so that they were not included in that version by adding a number to a destination field. If a record was just a cross-reference then entering 2 into the destination field ensured that the record would only appear in the book. Entering 1 in the field would mark the record for the online file only. Leaving the field empty meant that the record would appear in both versions. This feature has proved useful but has to be used carefully as it caused unforeseen errors. A record might be dropped from the online file because of being a cross-reference, but it might contain a heading that affected a series of following records and those records without their heading might prove to be meaningless.

The following example shows the problem that occurred in one commercial version of the online file where the headings of abstract records are assigned to all the relevant records. Sets of abstract records in the introduction to the corticosteroids chapter were organised under various headings including:

Pancreatitis;

Polyarteritis nodosa;

Polymyalgia rheumatica;

Pregnancy and the neonate.

There were only cross-references under the headings *polyarteritis nodosa* and *polymyalgia rheumatica*, and the first record under *pregnancy and the neonate* was a cross-reference; these were all marked for exclusion from the online file. The system being used for that commercial version of the file displayed the appropriate heading for each abstract record and assigned the heading *pancreatitis* to the abstracts that had appeared under *pregnancy and the neonate*.

Coders still find the application of the destination code difficult.

3.2.3 LOADING THE DATA

Data capture was to be by optical character recognition (OCR) via a reader at PPL. Sample data was sent to a number of organisations experienced in keying more straightforward text; the returned material demonstrated that the level of accuracy in keying Martindale data was unacceptable. A compromise was reached by commissioning a printing company to key the data into their computer and supply tapes with the data in a coded form such that it simulated an OCR input to PPL's computer. This route proved reasonably economical, although some records had to go through four cycles of corrections.

The system involved as much validation as possible. Identity numbers were validated through a check-digit algorithm; manufacturers' codes and chapter codes could be checked against their respective look-up files and any codes that could not be translated were thrown up in error. Delimiters used in pairs to mark typographical style or to select data for an index could be checked to make sure that these delimiters were correctly paired. Error messages were usually provided at the beginning of the erring record that had to be proof-read by the editorial staff.

Anomalies in coding threw up strange records, sometimes with error messages. Since those who had been involved in the coding also took part in the proof-reading, they were able to see the consequences of any miscoding. Sometimes these anomalies were easily spotted as in an author's name appearing in a text field and 20 lines of text appearing in the author's field. Unfortunately some mistakes were not so easily spotted and passed by unrecognised until with luck they were picked up at a final check after typesetting.

Proof-reading, like coding called for initiative and attention to detail since the contents of the manuscript did not always correspond to that on the printout. The punctuation and headings that would be automatically added by program were not displayed at this stage and there was a confusing quantity of almost meaningless code numbers. While the editorial staff grew adept at reading the printout, outside experts had a more difficult time when they came to read and comment on the copy[53].

One unexpected error was that of the contents of a field proving to be too long. There should not have been any restrictive limitation on field length, but in practice there was a limit of 15 000 characters per field. At this stage of initial data capture only a few records contained fields that proved longer than this and luckily those records once thrown up in error could be reorganised into sets of smaller records, but this problem was to recur at a later stage.

Once the data was considered to be correctly loaded, PPL ran the initial set of programs to produce the drive tape for the phototypesetter. The output from the typesetter was proofread against the initial input copy and provided a useful check on many of the features of the system. A few errors were detected and corrected, initially on the page proofs so as to publish the book. Once that was out of the way, arrangements were made to correct the database as work concentrated on the late stages of the preparation of Martindale Online.

3.3 PREPARATION OF MARTINDALE ONLINE FROM THE NEUTRAL DATABASE

3.3.1 THE NEED FOR DESCRIPTORS

The author decided early in the project that Martindale Online should be indexed with descriptors and that these descriptors should be capable of explosion. The main reason for this decision involved the arrangement of the data; a secondary reason involved the ultimate use of the database by pharmacists and physicians.

Although various studies pointed away from the value of indexing databases[40 - 42] in any complex fashion, sections of the text of Martindale could be identified where it would be difficult to retrieve data effectively without the aid of some descriptors. The custom of describing a drug's activity by comparing it against that of a well-known compound and elaborating only on the differences saves space and provides a useful frame of reference for the educated reader. If a drug is of a certain class it can be expected to possess the general properties of that class. Indeed, it is prudent to treat new drugs as representatives of their class until sufficient evidence has been obtained to confirm any variations in their activity. It is these variations that can be dealt with in detail under each individual compound; they may perhaps involve increased activity, altered absorption, decreased toxicity, or sometimes even increased toxicity. The following example shows the text of some fields of the adverse effects text record of demeclocycline; for the purposes of clarity, the field numbers and other coding information have been omitted.

> *Adverse Effects. [of demeclocycline]*
> *As for Tetracycline Hydrochloride, (cross-reference number). Allergic reactions appear to be more common with demeclocycline and phototoxic reactions occur more frequently than with other tetracyclines. Demeclocycline probably produces the most marked tooth discoloration, turning affected teeth a strong yellow.*
> *Reversible nephrogenic diabetes insipidus has been reported in patients treated with demeclocycline.*

The adverse effects record for tetracycline describes:

gastro-intestinal effects, supra-infection anti-anabolic activity and the kidney, hepatotoxicity, blood changes, effects on teeth and bones, effects in pregnant women, allergic reactions and reaction to degradation products.

There is no mention of nephrogenic diabetes insipidus. Searching the text of demeclocycline for information on it adversely affecting the gastro-intestinal tract would not yield a response, and it is unlikely that a searcher would think to search on tetracycline in lieu of demeclocycline. These subsumed concepts could, however, be made searchable if they were indexed with appropriate descriptors.

An alternative to indexing the text would have been to have repeated the text of the major compound in each relevant record that contained a cross reference to it, but this option was soon discarded as being too costly on storage and too unwieldy in operation.

Initial use of Martindale Online would be through intermediaries experienced in searching traditional online systems. Projected use would also be by pharmacists as well as by physicians who were then beginning to be cultivated by vendors such as BRS. Drug treatment might be initiated or withdrawn as a result of a search on Martindale Online by a range of searchers. The author therefore had to accept some responsibility for the effective and safe use of the file. Areas had been identified where descriptors would or might assist retrieval and the author could not take the risk of not indexing; certainly adding descriptors should not diminish the response to any search since free-text searching would still be available. However, assigning descriptors to Martindale Online introduced what is best described as the end-user paradox; end-users, being the searchers who should make most use of the descriptors, are the least likely group to use descriptors.

Unfortunately at that time there was little that could realistically be done to resolve this paradox, but it was decided that it should not be a bar to the adding of descriptors. End-users who did not knowingly make use of them would be no worse off and might even gain some advantage from an unconscious search through the database enhanced by the addition of more terms.

3.3.2 THE CHOICE OF THESAURUS

Once the decision had been made to index Martindale Online, existing vocabularies or thesauruses had to be assessed as to their suitability for this purpose. It was hoped that at least one would be suitable, but none provided exactly the coverage required. Assessments ranged over MESH[74], SNOMED[64], MALIMET[65], IPA[5] and more limited lists like the International Classification of Diseases[66]. To confirm this assessment Eaton[67] investigated the suitability of existing reference languages. Eaton's study, based on the antibiotics chapter of Martindale, showed that there was some correlation between the nomenclature used in Martindale and that used in both MESH and MALIMET, with MALIMET having the closest correlation because of the spread of its terms. Although subjects such as pharmacokinetics were not adequately covered, it was suggested that one of these two existing thesauruses should be adopted. However, the study also showed in a few examples the advantages of exploding on descriptors for effective searching of a small sample of Martindale data and this pointed to MESH being the more suitable of the two.

As a result of this study an information scientist was given the task of analysing other sections of Martindale to define the range of data. This indicated that the data could be considered as having 12 facets. Guidelines were laid down for each of these facets or sections and the assessment of descriptors for them was delegated to the assistant editor of Martindale. This further analysis indicated that MALIMET was unsuitable, mainly because of its structure. Sections of MESH appeared as if they could be used and permission was obtained from the National Library of Medicine to use their descriptors when required. However, MESH hierarchies did not always cover the range of data and Martindale staff had to create fresh hierarchies for each topic and assign descriptors to each level; it was therefore not practicable to copy sections of MESH. Unwillingly, a new thesaurus was created.

3.3.3 THE MARTINDALE THESAURUS

The Martindale thesaurus, when constructed, contained about 9500 descriptors; some typical pages are shown in Appendix B. The task of construction required considerable staff resources and had to be justified in terms of the overall production of Martindale data. As well as providing a tool for the effective indexing and searching of the

online file, the thesaurus provided a guide for the editorial staff when writing monographs.
The twelve sections of the thesaurus are as follows:

A	Drugs	G	Anatomy
B	Pharmaceutics	H	Physiology
C	Drug Administration	J	Diseases & Symptoms
D	Drug Absorption & Fate	K	Medical Procedures & Equipment
E	Pharmacological Actions & Uses	L	Environment & Technology
F	Organisms	M	Sociology

The major section of the thesaurus deals with drugs. Each compound described in Martindale had to have a descriptor which might or might not correspond to the monograph title. These drug terms were organised into hierarchies to correspond to the chapter hierarchies and any subgroups within a chapter. Only a few synonyms were included because of their common use; synonyms like heroin for diamorphine. Proprietary names, chemical names and less commonly used synonyms were excluded to keep the drugs section to a reasonable size, otherwise the index to the drug terms alone would have amounted to the equivalent of 166 pages of Martindale.
Almost every record in Martindale involved some aspect of a drug's activity or property and descriptors from other sections of the thesaurus covered these activities or properties. Martindale being a drug-based database was not being organised in a way that was suited to a pathologist but in a way that was suited to someone wanting information about the treatment of diseases. Thus an important section of the thesaurus had to deal with symptoms and diseases. Other sections deal for instance with the environment of a disease since that has some bearing on its treatment, but that bearing is limited and receives less emphasis in the thesaurus.
The relationship between descriptors, and especially between drug descriptors and the descriptors from the other sections, is established through qualifiers that form a separate section of the thesaurus. A descriptor from the diseases and symptoms section could be used to describe the adverse effects of a drug by incorporating the qualifier 'adverse effects' (abbreviated to AE) in the list of descriptors for a

particular record. The qualifier 'use' would be incorporated when the drug was being used to treat a disease. For example, pentazocine causes headache while aspirin (and sometimes even pentazocine) is used to alleviate such discomfort. The first concept is described by:

PENTAZOCINE, AE, HEADACHE.

and the second by:

ASPIRIN, USE, HEADACHE.

3.3.4 ASSIGNING DESCRIPTORS

This introduces the application of the thesaurus. As the thesaurus was being constructed, consideration was being given to the rules to be used when indexing. Consideration was also given to the way the descriptors should be organised within their fields.

Trial indexing had been carried out several years earlier when assessing the need for a thesaurus. This trial on the antibiotic group of drugs identified some of the problems that had to be addressed at this stage. Problems such as the depth of indexing, especially in records that carried several different themes.

As the database was drug orientated, so too was the indexing - but the indexing would also have to reflect the different record types. A record covering the introductory data for a group of drugs would not need to be indexed to the same depth as a record covering the adverse effects of a particular compound. Each record type or subtype was therefore analysed for its indexing requirements.

In keeping with the data, the indexing had to be accurate, thorough, and, despite the end-user paradox, organised in a way that did not require complicated search statements. Also, prospective searchers should be able to search within their frames of reference; thus a searcher interested in compounds that adversely affect the liver should be able to retrieve information by searching on the anatomical term as well as on any disease or symptom term.

When it came to investigating the way the descriptors should be organised within the field, it was decided that the format should enable the indexer to list descriptors in such a way that a record with several topics would not produce false coordinations of descriptors. The structure should also be relatively simple, both for the sake of the indexer and, more importantly, for the sake of the searcher. After testing several formats on paper, a trial was set up with Data-Star who

agreed to load groups of records indexed in four different formats. Martindale staff would then assess the efficiency and ease of use of the formats so as to select the most suitable for Martindale Online. These formats involved treating the descriptors as phrases, sentences or blocks with the linkages between the groups of descriptors being shown in a number of ways. Unfortunately the study had to be limited because of lack of time, but not before it showed that the simplest format was effective. In this format the descriptors were organised within the field into sentences. A record with several topics could thus be indexed with several sentences if there was likely to be confusion with cross-topic searching. The search software available on Data-Star's system could handle subfield searching which meant that searching could be aimed at descriptors within sentences. Any of the other formats would have called for software developments and complications in searching methods. It was considered undesirable to introduce any more searching complications. Indeed there were some doubts that searchers would accept even subfield searching. These doubts, however, were not considered to be serious when the test file with the chosen format was demonstrated at two conferences to drug information specialists and they expressed no objections to the format.

The following record shows a descriptor field divided into sentences to cover the themes in the text field. The field numbers shown in the example are those that were eventually adopted for the neutral database; they are not the numbers originally used at coding.

!000 *12405/a2/1/z*

!202 *10*

!204 *Allergy*

!206 *Either bronchospasm or a full anaphylactoid reaction had been reported in 27 patients given atracurium, of whom 4 had suffered cardiac arrest but recovered. It may be unwise to use atracurium in people with atopy or asthma.-*

!211 *Drug & Ther. Bull.,*

!212 *1985*

!213 *23,*

!215 *51*

!234 *atracurium, ae, bronchospasm, bronchi, allergy, anaphylaxis, immune response, cardiac arrest, heart.*
atracurium, precs, atopy, asthma.

When proper indexing started, the staff had guidelines that covered not only the ways that different record types should be indexed, but the ways in which the qualifiers and different sections of the thesaurus should be used. A summary of these instructions is included as Appendix C.

3.3.5 INDEXING PROBLEMS

3.3.5.1 CROSS-REFERENCES

Reference has already been made to the record that described the adverse effects of demeclocycline and to the difficulty a searcher would have in retrieving anything about demeclocycline adversely affecting the gastro-intestinal tract, since that effect was embodied within the tetracycline record to which demeclocycline referred. Having discarded as uneconomical the option of repeating the text of the adverse effects record of tetracycline under demeclocycline, there remained the problem of how to index such embodied data.

The brief demeclocycline record could be indexed with all the relevant descriptors from the invoked tetracycline record with consequent confusion for the searcher. Or the appropriate concepts in the tetracycline record could be indexed with the drug descriptors for those compounds referred to it. This had the advantage over the first option of reducing the considerable repetition of descriptors, but it had the disadvantage of providing a searcher with a record that contained descriptors that had little apparent relevance to the text. To retrieve the adverse effects of tetracycline when one had asked for demeclocycline might be accepted by those well versed in the activity of this group of drugs; it might not be acceptable to the majority of searchers. An explanatory sentence was thus added to the beginning of the text of 'cross-referred-to' records, such as the tetracycline record, by way of an additional field. This informed searchers that one or more named drugs had been referred to that compound because of similarities in use, or adverse effects, or some other property.

3.3.5.2 MINOR INFORMATION

Another problem peculiar to this database was how to handle or index drug data that was minor in nature. Diazepam's uses are described fully in the uses record of the diazepam monograph and these include its use intravenously as an anticonvulsant. Anyone wanting details of this use should retrieve this 'uses' record. However, a large number of drugs

cause convulsions as adverse effects - convulsions that may be treated with diazepam. Diazepam would be mentioned in the treatment of adverse effects record of each of those many drugs and its use would be indexed. Searching on the use of diazepam to treat convulsions would therefore lead to a large number of hits, only one of which carried the major information. These two levels of use information had to be indexed in different ways. This was achieved by creating a hierarchy for the qualifiers. Each qualifier was given a second level minor qualifier as shown below.

ACT
    UF action
    minor
*ACTM
AD
    UF administration and dosage
    minor
*ADM
.
.

Minor data could be indexed using the minor qualifier. In the diazepam example the qualifier USE would be used to index the data in the diazepam uses record while USEM would be used to index those records dealing with the treatment of adverse effects. Significant data could then be retrieved by searching on USE; should anyone want to retrieve all the records that dealt with the use of diazepam then an inclusive search could be carried out by exploding USE or the search statement could contain 'USE or USEM'.

3.3.6 CONVERSION OF NEUTRAL DATABASE TO TAPE SERVICES INTERCHANGE FORMAT

Part of PPL's remit was to produce from the neutral database tapes of Martindale data in a format that complied with ISO 2709. A tape services specification was drawn up and fields were considered for their relevance to the online service. Those fields that were selected are shown in Appendix A.

In preparing the interchange or tape services tapes, PPL stripped out all typesetting features from the selected fields.

Fields not selected included those concerned with editorial control of the data, such as those containing the checker's or revisor's initials and other similar fields of value only within the editorial office. Other fields like those for the typesize or for the book index were used solely for typesetting and were also rejected. Of course, some fields were designed solely for the online service; fields such as those for descriptors and additional authors, and as such, these fields were not selected from the neutral database before typesetting. Initially the sort code was not considered relevant to the online searcher, but it was soon realised that the vendors' programs required the sort code to ensure that linked records were displayed in the right order. On the other hand, the chapter code was included yet turned out to be of little value, at least to one vendor.

The identifying numbers of the selected fields were converted to a set of tag numbers that avoided the problems of the vendor having to deal with the same number for different fields. Rather than create a completely new set of numbers, an extra digit was added to the front of the field number. Although the neutral database only contained 4 different kinds of record format, each with their own set of fields, 2 of the record formats had more than one function. The monograph record was also used to cover the introduction data at the beginning of each chapter, while the abstract record was used for 3 different record types - the abstract subrecord of the monograph record, the abstract sub-subrecord of the action and uses subrecord, and that of the preparation subrecord. To avoid confusion, individual records were tagged with a number in the range 0 to 6 by which they could be identified, see Table 4.

Type of Record	Identity Number Format	Tag
introduction	dddddd-c	00
monograph	dddddd-c	01
monograph abstract	dddddd/dddd-c	02
action and uses text	dddddd/adddd-c	03
action and uses abstract	dddddd/adddd/dddd-c	04
preparations	dddddd/ndddd-c	05
preparations abstract	dddddd/ndddd/dddd-c	06

Table 4 Tags for Record Types

(d = a digit in the range 0-9; c = a check character
a = a letter in the range a-m; n = a letter in the range n-z)

The selected fields were to be set out using PPL's extended character set based on the standard UK ASCII character set of 8-bit codes. This character set was suitable for printing but not so good for display on terminals with relatively limited character display facilities. PPL were therefore commissioned to convert the characters to ASCII format. Conversion to EBCDIC was not considered essential and has been held as a reserve option.

3.3.7 DISTRIBUTION BY VENDOR

The decision to distribute Martindale Online through the commercial vendors was based on our desire to provide a service that could be linked or associated with the bibliographic databases that also provided drug information. It was suggested that we should act as our own vendor. Such a suggestion was quite unrealistic. The costs would have to be borne initially by one small database and those costs would encompass not only the hardware and systems, but the marketing and training, invoicing and other elements of a vendor's activities. After all that, there would still be the problems of linking Martindale Online to related bibliographic databases and persuading searchers to try yet another system for just one database.

It was also decided that the distribution rights should not be exclusive to one vendor. The straw poll at the beginning of this exercise indicated a demand among online users of DAILOG and BRS's services. It also showed that those used to searching on one online system were unwilling to change to another system just to gain access to Martindale Online. An argument in favour of an exclusive contract was that one contract could produce better financial returns than several contracts that split the market across several vendors. This arguement was not accepted as we were able to identify searchers who would not be reached if the database was limited to one vendor. However, there was some merit in the argument since spreading access over a large number of vendors might limit the use on one or more systems to a point where it was not worth those vendors holding the database.

It was interesting that negotiations differed considerably with different vendors. For instance some vendors changed their policies about new files during the negotiations and wanted the database suppliers to absorb some of the costs of development and storage. A change not welcomed by the producer. Some vendors also had marketing plans that were considered to be restrictive since they saw Martindale Online as fitting into a group of databases specially organised for a certain group of searchers. Such organisation tended to simplify searching for end-users, perhaps by limiting the facilities offered and so diminishing retrieval effectiveness.

Traditional charging for online databases at that time involved a royalty fee for the database producer based on the amount of time spent searching the database. Thus the faster a search was carried out, the better it was for the searcher. With the increasing availability of high speed data transmission, the costs to the searcher could be reduced considerably with a consequent reduction in royalties for the database producer. Associated with this was the increasing use of microcomputers to store or download the results of a search or searches, and such downloading could cover large portions of Martindale Online.

Downloading was recognised as being convenient to a searcher and was acceptable if the data so stored was for private use. To ensure that some revenue was produced when high speed terminals were in use and to counteract downloading infringements, a royalty fee was negotiated for Martindale Online whereby there was a component for the time spent searching and one for each record printed online or offline.

3.3.8 ONLINE SYSTEM DEVELOPMENT

One of the early developments in this distribution phase was to create a sample subfile of the monographs that constituted the new antibiotics chapter. This subfile was reorganised by the embryonic tape services programs and sent to one vendor to be loaded as a private test file to show how a vendor would handle the file using existing software. The aim was to identify where the system needed amending to enable the data to be retrieved effectively and in an intelligible form - bearing in mind that the subfile would not quite represent the data in its final form. Account would also have to be taken of the indexing of the subfile. The thesaurus project was then in its early stages, so a temporary set of loosely organised key words was created and these were assigned to each record. This primitive indexing proved valuable when it came to preparing the indexing guidelines for the full file.

Unfortunately that proved to be the only worthwhile product of the test. The selected vendor did load the the test material exactly as it was supplied to them by PPL and this showed various markers and codes but not the organisation of the data into the usual paragraphs or sections seen with the publicly available files. Such organisation and further development to produce an intelligible test file was not possible because of contractual problems and, much as the author wanted to proceed, the test was abandoned. All it had demonstrated was that the tape services tapes in a basic form could be read by one vendor and that that vendor would be required to carry out an unknown ammount of work to render the data fit for external use. The test project could have been tried with another vendor, but the initial experience discouraged another attempt. Instead, the features of Martindale Online were fully explained when negotiating contracts and, in general, it was agreed that both parties would together determine the format of the database on the vendor's system.

Each vendor was supplied with guidelines on how the fields in the different record should be allocated within their systems. One of these sets of guidelines is shown in Appendix D.

The philosophy behind these guidelines was to ensure that each Martindale Online record was sufficiently and meaningfully labelled with the record type as well as the name of the drug involved. In addition, any links between records should ensure that such records should be retrieved as a group yet still be individually searchable. The indexing

also demanded that there should be an operator to limit searching to sentences of descriptors.

3.3.8.1 RECORD SUMMARIES

In the early discussions it was proposed that the vendors should also produce a print summary of each record. This was a protection against searchers acting solely on the presence or absence of a hit in response to a search statement. The question 'has aspirin caused lupus erythematosus?' may be phrased in several ways and a hit may be achieved to any of those search statements. There is a risk that a searcher may go no further and not find that aspirin has been tried in the treatment of this condition. Some protection could be obtained by providing the searcher with an indication of the record type for each hit so that at least he or she would know whether the record dealt with adverse effects or uses. It was then considered that there might be some value in the searcher being told the length of the record (to provide some indication of the cost of retrieving the full record), the drug under which it was organised and the date of addition to the online service. Searchers could then decide whether they should obtain an offline rather than an online print. If the date was no later than that which coincided with the hard copy version, the searcher could always consult the copy of Martindale; in such instances the online file would act as a super index.

There was no way that the vendors could force such a record summary on their searchers, but a voluntary display could be obtained using the print command options. With Data-Star, for instance, the fields to be included in the print summary were listed under their BIBL command so that any searcher who specified ..p/bibl would retrieve the summary of relevant records.

A typical summary as shown on Data-Star is as follows:

AN	*(accession number)*	*2592/a3-s*
TI	*(monograph title)*	*Miconazole Nitrate*
RF	*(record form or type)*	*Precautions Text*
LE	*(record length)*	*525 characters*
ED	*(entry date)*	*Jan 1983.*

3.3.8.2 LINKED RECORDS

The record linkage proved a difficult task for the vendors. Data-Star managed it in such a way that an individual record within a linked set was searchable, but on printing all the linked records were displayed, each with a heading to indicate that a particular record was one of a set. Each record within the set was indexed separately, and the display of the descriptors plus the additional heading reduced the readability of the set. The file was made available on DIALOG at a later stage and so some improvement in the layout of linked records was able to be made. The descriptors had to be kept but a number of the headings were dropped from the follow-on records so that the set read more like a continuous record.

3.3.8.3 HEADINGS

Assigning the appropriate headings to each record also required some development. Each record needed to be headed with the name of the monograph or chapter title to which it belonged as well as with the type of record. The monograph name was contained in tag 103 of the type 1 record of a monograph's suite of records. This name had to be allocated to all relevant subrecords. Similarly the chapter name was contained in tag 100 and that name had to be assigned in a similar way. The record type had to be identified from the record type tag and the contents of any relevant heading tags. Linked monographs provided headings that contained the monograph titles of all the linked monographs.

3.3.8.4 THESAURUS

In addition to the data tape, the vendors were supplied with a listing of descriptors and their notation. For inclusive searching the descriptors in each record had to be matched against their notation from this listing. Facilities for inclusive searching varied with different vendors. On Data-Star, the searcher is able to flag a descriptor for inclusive searching; the system translates the descriptor into its notation and then searches on a truncated form of that notation. Under DIALOG the searcher has to identify the notation for any descriptor, then enter the truncated form of that notation as part of a search statement. Examples of the notation and some hierarchical levels can be seen in Appendix B.

3.3.9 INSTRUCTIONS ON THE USE OF MARTINDALE ONLINE

Having set up a database with an unusual record structure and one that should be searched with the aid of descriptors that had been assigned in a certain way, some guidelines were necessary for those wishing to use the file.

A user's guide was prepared following investigation of its aims[66]. A neutral guide independent of any vendor's labels and layout was printed within each thesaurus and printed separately as a free booklet. Vendor specific guides were also prepared as separate free booklets.

The guides described the organisation of Martindale Online, how it was indexed and how to retrieve information from it. The simple message contained in them was:

- search through descriptors whenever possible
- when using descriptors link them with an operator that limits searching to a co-occurrence of descriptors within sentences
- make use of the abbreviated record display.

Examples of searches included searches on drug identification, such as; *What is Elspar?* In this case natural language searching would be used.

Examples of questions that required descriptor searching included: *What drugs have been tried in the treatment of bone marrow depression?* or *What vitamins cause anaphylaxis?*

The abbreviated record display or record summary provides the accession number of the record, the title of the drug or monograph involved, the type of record, and the date the record was added to the system. In the early days of the database this summary would tell the searcher if the record was more up-to-date than the appropriate section in the book. It would also provide a warning that the information might be in a part of a monograph that they might not have expected. Asking for information on whether a drug has been used in the treatment of a disease might produce a number of hits, the summary would show what record types were involved and might show that some of the records were precautions or adverse effects records. Some of the records are several thousand characters long, so the summary should help the searcher decide whether to select an offline or an online print of the full record.

3.4 UPDATING AND MAINTENANCE

Martindale Online is updated by adding new records, deleting records and replacing records in which fields have been amended - another hurdle for vendors used to updating bibliographic databases by just adding records. It was planned that an update should incorporate rewritten monographs on groups of drugs (i.e. chapters) and changes to preparation records in all monographs to reflect the changing availability of drugs. Although the editorial team scans relevant journals to obtain material for use in revision, it was decided that it was impractical to amend the database in response to each paper; such an exercise would fully employ the staff leaving them no time to carry out the detailed review or evaluation required for revision of a monograph. However, it was realised that some reports or papers might contain information that was important enough to merit a partial update of a monograph.

A new preparation might be for a new drug, as sometimes happens, then a new monograph has to be created with its suite of subrecords. If the new preparation is for an existing drug, then a new subrecord has to be created. However, the new preparation might be for a new indication or represent a varied dose that makes a nonsense of some of the text record. In such a case there is not enough time to carry out a thorough revision of the affected record, so some extra text is appended to indicate that searchers should interpret the preceding text in the light of this new information. For example, an early change had to be made to the uses record of acyclovir, an antiviral agent, to say that since the text records of that monograph had been prepared, a preparation had been made available in the UK for use in herpes labialis.

Updating Martindale also involved covering more facets of drug information. One feature that the author wanted to add for instance was the identification of ingredients of commercial mixtures of international proprietary preparations. Such features have to be incorporated within the record structure. In this case that part of the task was relatively simple; where the problem occurred was in identifying the active ingredients and in presenting the data in an intelligible form. For example, there are very many multivitamin preparations availabe in this country, even more in the USA and other countries, so that ascorbic acid occurs as an ingredient in thousands of such products. In some of these

products, it is available as a vitamin (vitamin C), in others it is included as an antoxidant or preservative and in these cases it is not considered to be an active ingredient. However, such are editorial problems!

Deleting a record meant not only striking it out of the neutral database, but informing the vendors that they had to delete that record. A field has been allocated to each record just for this purpose. If the word delete is contained in that field, then the identity number of that record plus the delete field can be sent to the vendor so that the record can be withdrawn from the commercial version of the database. Care has to be taken to ensure that the deleted record is not the top record of a hierarchy of subrecords; if it were, then the subrecords would be considered as deleted by PPL and erased from the master file - but without a delete instruction being sent to the vendor. Care also has to be taken when deleting any records in linked sets or in records that are the subjects of cross-references.

Updating involves the addition of new data and the re-interpretation of old data, and that means that the thesaurus has to be updated. New descriptors are required. Some descriptors will need to be reclassified, as mostly happens with drug descriptors, in which case the notation needs to be altered. Only a few descriptors are deleted. The addition or the deletion of a descriptor in a database like Martindale Online means that the indexing for all records that might be affected needs to be examined and updated by re-indexing.

Not surprisingly updating Martindale Online has proved to be a problem for the vendors as well as the editorial staff. Despite the additional costs, vendors have recently opted for carrying out a complete reload of the file for each update.

CHAPTER 4 EVALUATION OF RETRIEVAL EFFECTIVENESS SEARCH SYSTEM: METHODS

A detailed study of retrieval effectiveness was carried out on Martindale Online, first of all by examining various search options including free-text searching and controlled searching, and secondly by analysing the effectiveness of searching the database as recommended. The second part of the investigation is an extension of the first, but for clarity in presenting the results, the study is presented in two parts: one dealing with what could be called a systems investigation; the other with failure analysis. This section deals with the systems investigation.

One of the most critical decisions in setting up the database was the decision to index the file with descriptors from a specially created thesaurus. Now that the database is commercially available, that decision could be assessed. Thus, the first test was set up to establish whether the records in Martindale Online had required indexing with descriptors. This test involved comparing structured searching of the indexed file with unstructured searching of the unindexed file. Initially sales of the Martindale Thesaurus were poor; this indicated that many searchers were accessing Martindale Online without knowing the descriptors to use in their search statements. The second test was intended to find out if searching the database without the benefit of the thesaurus gave inferior results to searching with statements created through the thesaurus following the guidelines recommended in the Martindale Online User's Guide. This involved a comparison of unstructured searching of the indexed file with structured searching of the indexed file.

If searchers were not searching as recommended, they were at least using the descriptors unconsciously. Were they gaining any benefit from those descriptors? This formed the basis of the third test which compared unstructured searching of the indexed file with unstructured searching of the unindexed file.

Finally in this part of the evaluation it was decided to build on the above comparisons to assess the potential value of integrating the thesaurus into the database. Such integration might, in this database,

be a step towards the automatic conversion of unstructured search statements to highly structured and detailed statements that would encompass lead-in terms, appropriate related terms, as well as explosion of hierarchies. How, in the unindexed file, would searching with such a statement, prepared manually in this instance, compare with searching by way of an unstructured statement? Also how would searching the unindexed file with these highly structured statements compare with structured searching of the indexed file which is the method recommended for optimum retrieval? These were the bases of the fourth and fifth tests. The various types of searching required for the evaluations were labelled S1 to S4 as shown in Table 5. Further details of the methods of searching are provided in Appendix K.

S1	unstructured search of the unindexed file
S2	unstructured search of the indexed file
S3	structured search of the indexed file
S4	highly structured search of the unindexed file

Table 5 Definition of Types of Searches

The comparisons made between these types of searches that were carried out to investigate the above questions are shown in Table 6.

S3 versus S1
S3 versus S2
S2 versus S1
S4 versus S1
S4 versus S3

Table 6 Summary of Search Comparisons

4.1 QUERY SELECTION

When Martindale Online was first made available searchers were contacted so that their use of the database could be analysed. Unfortunately it proved difficult to obtain enough information for any worthwhile analysis since the searchers were reluctant to provide details of their

searches. Considering that many of the users were in the pharmaceutical industry, that reluctance is understandable. Other methods of obtaining questions had to be used for this study. The Pharmaceutical Society of Great Britain runs an information department with 3 information pharmacists who answer about 14 000 queries a year. These information pharmacists were each asked to select from their logs of enquiries about 40 questions that they considered they might put to Martindale Online. They started with the most recent query and worked backwards excluding without guidance any questions that they thought might not be a test of the online file - questions for instance on the identity of foreign drugs. One of the regional drug information centres was also approached for a list of questions, but was not able to supply a list.
The author checked the selected questions and rejected those that were one-topic questions and would not involve any search statements beyond that one concept. Typical rejected questions were:

'What is silver sand?'

'What is SVI alcohol?'

Two information pharmacists supplied the same question and one of those duplicates had to be excluded. Finally one question that was baldly stated as - 'The uses of amino acids?' - was rejected as being too vague.
In all, 121 questions were submitted; 23 were rejected for the reasons stated above, leaving 98 to be used in the evaluation. The selected and rejected questions are listed in Appendix E and Appendix F respectively.

The three information pharmacists had different ranges of experience of Martindale Online. One had a good knowledge of the file structure having carried out research on the preparation of the Martindale Online User's Guide[66]. The second information pharmacist had some knowledge, but little experience of using Martindale Online, although she had carried out the early research on possible indexing languages[67]. The third member of that group had little knowledge of the database. None of them had used Martindale Online to help them answer the questions that they had submitted; the reasons for this are not clear and require investigation.

4.2 PREPARATION OF SEARCH STATEMENTS

Each of the information pharmacists was asked to draft a search statement for each of their questions without using the thesaurus or user's guide but calling on any experience that they had gained from searching other online databases. Having completed that batch of search statements, they were then asked to create statements with the aid of the thesaurus and user's guide. Finally they were requested to complete a third batch of statements using the thesaurus to provide additional terms from the information provided on related terms and scope notes. If they wanted the terms included from a hierarchy, then they could indicate that on their statement. Qualifiers such as USE or AE were converted to a standard string of terms.

It could be argued that the information pharmacists should not have constructed the statements for their own queries. However, the queries as provided were in a shorthand form (see Appendix E) and it was not possible to return to the originator to ask what was meant. Some queries thus required some interpretation which could only be done by the information pharmacist who had already answered the query, or at least tried to answer it, and who had generated the brief details in the query log.

The author also created three batches of search statements following the same guidelines to see if detailed knowledge of the file structure and database contents made any difference to the result. So as to distinguish between the author's and the other statements, the information pharmacists' statements were marked with the letter A and the author's with the letter B, as in S1A and S2B.

The comparisons listed under Table 6 should therefore be as follows:

S3A versus S1A	S3B versus S1B
S3A versus S2A	S3B versus S2B
S2A versus S1A	S2B versus S1B
S4A versus S1A	S4B versus S1B
S4A versus S3A	S4B versus S3B

Table 7 Search Comparisons

Since the two sets of three batches of statements were going to be used for two sets of four different types of searches, the author marked up the statements in the first batch of A and B so that they could be used to search the database excluding the descriptor field or paragraph and then search the database including the descriptors, i.e. for S1A and S1B and for S2A and S2B. This is discussed in more detail below under Searching, however, it is worth pointing out at this stage that this mark-up did not change the content of the search statements.
Some additions had to be made to some of the statements in the third batch where the information pharmacists had marked that they wanted some part of a hierarchy to be included. The changes that were made are also discussed below and again it should be emphasised that this did not change the statement's logic or components that were not involved in the hierarchy.
Errors in the search statements were not corrected. Statements that would obviously produce large numbers of hits were not refined, apart from two cases in which only a selection of the hits was printed for examination as the number of hits was very large.
The search statements for each query are shown in Appendix G which also shows the response to each search.

4.3 SEARCHING

There were up to 8 different searches for the 98 queries and these searches were carried out using two systems: the commercially available system from Data-Star and the internal system on the Vax 11/720 computer that was used within the Martindale editorial office. There were advantages and disadvantages to searching either system, but both had to be employed.
Searching through Data-Star involved third-party costs and difficulties in excluding the descriptor field for the S1 and the S4 searches. However, public searchers at that time only had access to Martindale Online through this vendor and the tests should involve the file as searched commercially.
Searching through the internal system incurred less cost and was more convenient in that the search software made it easy to exclude the descriptor field from any search. But there were several drawbacks to searching through this system. First, the format was different from that on the Data-Star system. The fields just had their original field

number and were not grouped into their paragraphs. There were no headings or titles, nor were the linked records organised into an easily comprehensible order. Since the searching was being carried out on the neutral database, records were retrieved that were not relevant to the online version and had to be discarded at the analysis stage. Secondly, inclusive searching was not available on the internal system. Thirdly, the internal version of the neutral database is a working file and is updated more frequently than tbe public version.

Since neither system met all the requirements of all the tests, both systems were used. Data-Star was used when any inclusive searching was required and to check on some of the searches carried out on the VAX 11/720. The internal system was used for S1, some S3 and some S4 searches. Various checks therefore had to be carried out on the output of the searches depending on which system was used.

The format of each record on Data-Star's system provides useful headings translated from the identity numbers and the hierarchical status of the record. Each record, for instance, contains the title of the monograph to which that record belongs and a heading describing the type of record, be it a Uses Abstract or Precautions Text. As mentioned above the same records on the VAX only contained their identity numbers for identification with no definition of record type or monograph title. Checks therefore had to be made to ensure that searches carried out on the VAX did not miss records that would be picked up from Data-Star through a search covering the headings and titles. All components of such headings and titles could be recognised in the search statements; any statements with such components were assigned to a search on Data-Star, either instead of through the VAX or in addition to a search through the VAX.

If Data-Star was used for S1 or S4 searches, then all records had to be discarded that had been retrieved through the descriptors or the special paragraph in the text that related to the descriptors. S1 searches were usually carried out on Data-Star so as to search the headings or titles. S4 searches were usually carried out on Data-Star for searches in which the information pharmacists required a hierarchical group of terms to be included in their search statements. In many cases it was simpler to carry out an inclusive search on the indexed file and discard those records that were retrieved as result of a hit in the descriptor

paragraph than to write out a search statement containing all the terms within the hierarchy.

The facilities of the VAX and Data-Star were used together in instances where a search on Data-Star produced a large number of hits; in such cases the records were printed out from the VAX using the identity numbers retrieved from the other system. The copies of the records retrieved in this way have been marked appropriately, otherwise the reason for its retrieval might not be apparent.

Searches on the VAX were carried out by the author with the aid of a clerical assistant. All Data-Star searches were carried out by the author. Because of the large number of searches it was not feasible for the information pharmacists to carry out the searches, whatsmore they were not familiar with the VAX systems. However, it can be argued that experimentally it was better that the information pharmacists did not carry out their own searches.

There was no feedback to produce changes in the search statements that did not produce what could be taken for reasonable results. It would have been difficult to have obtained valid alterations because of the third party status of the queries. Also the additional time commitment for the information pharmacists was too great, especially as the searching had to be completed between updates (see below). Finally on this aspect of the study, it was considered that the relevance of feedback was diminished by the aim of the study being to compare the effectiveness of different methods of retrieval.

4.4 ADDITIONAL HITS

After the searching as specified above was completed but before the analysis, the author carried out further searches using a variety of techniques to trace any additional records that might be relevant. The file was too large for a manual search of all records, but the author being the editor of the contents of the file, had written some of it, rewritten other parts, and checked all of it. The author therefore had a wider knowledge of the contents of the database than anyone else and was able to direct manual searching at likely relevant sections. Some of the search statements were widened, often by increasing inclusive searching or by altering the search logic. These endeavours might not have retrieved all the relevant records, but it is unlikely that many

were missed. Also, and more importantly, the attempts to trace the additional relevant records were consistent for all 98 queries.
The additional hits that were obtained were referred to the relevance assessor and those deemed relevant were used together with the other relevant hits to arrive at a total number of relevant hits for each query; these numbers are shown for each query in Appendix G.

4.5 TIME LIMITATION

Reference has been made above to the influence of updates on this study. Martindale Online is updated through the addition of new records, the deletion of old records, and through changes in the content of some of the retained records.
This meant that to obtain a valid set of results for all the queries, the the search statements had to be processed at a time when the file was in a static phase. Updates at Data-Star were planned for every six months, but took longer. Updates on the VAX could occur at any time after the data had been written. However, the author was able to ensure that no VAX updates occurred during the bulk of the searching. There were some occasions, usually when tracing additional records (see above) that searching on the VAX would have involved searching the updated file; on those occasions the searches were carried out on Data-Star.

4.6 RELEVANCE ASSESSMENT

Printed copies of all the hits were obtained for each search and these are available for further study. All these prints were assessed for relevance by a pharmacist on the Martindale editorial team who had some experience of assisting the information pharmacists during periods of undermanning in the information department.
This pharmacist was given up to 8 sets of prints for each query where a search had resulted in a positive response. He was also given a note of the query as reported by the information pharmacists. A few sets had the search statement written on them, but on the whole the sets were only identifiable by the query number and a code to mark the individual set. He was asked to assess each print for relevance and, if relevant, whether the relevance could be considered as major. The judgement of relevance was to be his and the only other instruction given was for him to produce a definition of what he meant by relevance and major relevance.

The notes provided by the assessor are reproduced verbatim in Appendix H. In summary, records were judged to be irrelevant if they contained no 'real useful information', even if sometimes information that others might consider of some relevance was present. Duplicate information was also judged to be irrelevant, thus if a query was answered by the same information from two sources, then one source was treated as irrelevant - although the assessor had doubts about his consistency in this. Records were judged to be of minor relevance if they led him to information that might help answer the query or if they contained useful information but did not exactly cover the query. Major relevance was used for what he termed 'bullseyes', and that included records obtained through cross-referencing, for text records even where only a portion was a bullseye, and for background information that would affect bullseyes.

While others would have made different assessments, attempts at obtaining further assessments were abandoned. One of the information pharmacists did assess some of the results from her search statements, but she was unable to complete her assessment and it was considered to be less relevant to this study than an assessment made by a third party who was able to cover all hits. The information pharmacist's assessment was therefore rejected.

The author inspected all the assessments and only referred back those hits where the same record had been scored or marked differently within the same query. The assessor was only told that there was an inconsistency; he had to decide which assessment to correct.

4.7 SCORING THE SEARCHES

A simple method of scoring the searches was devised so that the comparisons between the different types of search could be effectively carried out. The score took account of recall and precision, but did not call for any values to be set for either feature.

The formula devised with the aid of the statistics department of The City University was:

$$\theta = \frac{i + 1}{\sqrt{(i_{max} + 1) \times (r + 1)}}$$

i = the number of relevant hits retrieved

i_{max} = the total number of relevant hits in Martindale Online

r = the number of relevant and nonrelevant hits retrieved.

When $\theta = 1$ there is 100% recall and precision.

4.8 STATISTICAL ANALYSIS

As was expected the scores for each set of searches did not fall into a normal distribution and attempts at transforming the figures failed. The Wilcoxon matched-pairs test[69] was therefore selected as being suitable for the comparisons being carried out.

This test involved ranking the differences in scores obtained between the two methods of searching being compared. The smallest difference was given the rank of 1. Zero differences were ignored. When several pairs of scores showed the same differences, then a rank was given that corresponded to the midpoint of the range of differences. Whether a difference was positive or negative was ignored for the purposes of ranking, however once the ranking was complete, the plus or minus status was assigned to each rank. The sum of the positive or negative ranks, whichever was the smaller, was selected as the value to be read off the statistical tables provided for this type of test[70]. These tables provide an approximate value for P, the probability that the observed results could arise under the null hypothesis of no systematic difference between the methods being compared.

Unfortunately the statistical tables only provided data for up to 25 pairs. This proved no problem when analysing each of the 3 sets of queries from the information pharmacists as separate sets. But when it came to analysing the cumulated data for all the queries as one set, then the number of pairs exceeded those covered by the tables and the method for handling large numbers of pairs as described by Colquhoun[71] was adopted.

This method makes some approximations to provide a standard normal deviate that can be read off tables of t (with infinite degrees of freedom)[72]. Because of the large number of samples or pairs, the mean is taken as:

$$\mu = n(n + 1)/4$$

and the standard deviation as :

$$\sigma = \sqrt{\frac{n(n + 1)(2n + 1)}{24}}$$

The approximate standard normal deviate is calculated as:

$$u = \frac{|T - \mu|}{\sigma}$$

where T is the smaller sum of the signed ranks

When checked against the values obtained from tables, the above formula works reasonably well. For instance, the formula was used for a small set of pairs where the P value could be read off the tables; the tables and the formula gave the same value.

The results are displayed in a series of tables that provide the sums of θ for each set of searches being compared and the corresponding P value.

CHAPTER 5 EVALUATION OF RETRIEVAL EFFECTIVENESS SEARCH SYSTEM: RESULTS

Analysis was carried out on 8 sets of searches for the 98 queries. In some of the straightforward queries that led to simple search statements, the A and B statements turned out to be identical, or sometimes the S1 statement for either the A or the B search proved to be the same as the S4 statement. Out of the 784 statements, 116 proved to be this kind of duplicate.

These 784 statements produced altogether 2588 records for analysis once the records were discarded that had been retrieved from the VAX but would not have been retrieved from Data-Star. Sets of linked records were treated as one record. As would be expected, the different searches for the same query often retrieved some common records. Thus, of the 2588 records in the 784 sets, only 1120 were unique records. Of these 1120 unique records, 326 were considered to be relevant, with 210 of them being judged of major relevance.

Of the 326 relevant records, 168 were abstract records of which 117 were of major relevance. There were 107 action and uses records with 60 of them being of major relevance. There were 36 preparation records (24 major), 13 monograph records (7 major) and 2 chapter introductions (both major). Table 8 lists these figures.

RECORD TYPE	TOTAL N°	RELEVANT HITS MAJOR & MINOR	RELEVANT HITS MAJOR
Introduction	106	2 (1.89%)	2
Monograph	5138	13 (0.25%)	7
Action & Uses	8986	107 (1.19%)	60
Abstract	57091	168 (0.29%)	117
Preparation	11122	36 (0.32%)	24

Table 8 Relevant Hits by Record Type for All 98 Queries

The scores for each search are shown in Appendix I which is laid out in the order of the comparisons as shown in Table 7. Thus, when a comparison is made, the scores for the two types of searches are given together with their differences and ranks. The calculations to arrive at the level of significance are also provided.

5.1 STRUCTURED SEARCHING OF THE INDEXED FILE (S3) VERSUS UNSTRUCTURED SEARCHING OF THE UNINDEXED FILE (S1)

Comparisons were made at several different levels. The main comparisons involved the scores obtained for all the queries from the three information pharmacists as a result of search statements formulated by the information pharmacists or by the author (see Table 9, below).

SET	RELEVANT HITS	S3	S1	P
AT	major & minor	76.542	67.696	<0.001
BT	major & minor	77.719	70.759	<0.001
AM	major	76.672	69.682	0.001<P<0.01
BM	major	76.791	74.044	0.1<P<0.2

Table 9 Scores for S3 Versus S1 for All Queries
(A = information pharmacist generated search statements
B = author generated search statements
T = total hits
M = major hits)

When the information pharmacists' search statements were assessed the score of 76.542 for structured searching of the indexed file was significantly higher than 67.696 for unstructured searching of the unindexed file (P <0.001) when all 98 searches and all relevant hits were considered. When the author's statements were assessed S3 also provided a significantly higher score than S1; 77.719 compared with 70.759 (P <0.001).

Hits of major relevance were compared in the same way. With set AM the total score for the 98 searches was 76.672 compared with 69.682. This difference was less than that for all relevant hits but it was still

significant (0.001<P<0.01). However, no significant difference was observed between the two methods of searching when the author's search statements were used; 76.791 compared with 74.044 (0.1<P<0.2).

The 98 queries were made up of 3 groups. The group I queries (34) were those from the most experienced information pharmacist in terms of Martindale Online; the group II queries were from the information pharmacist with some knowledge of the file; the group III queries were from the information pharmacist with least experience or knowledge of the file. Comparisons were made within each of these groups.
The results obtained from group I queries are shown in Table 10.

SET	RELEVANT HITS	S3	S1	P
AT	major & minor	26.562	22.148	0.01<P<0.02
BT	major & minor	27.030	23.957	0.02<P<0.05
AM	major	25.782	21.912	0.02<P<0.05
BM	major	26.365	25.130	>0.4

Table 10 Scores for S3 Versus S1 for 34 Group I Queries

The score in the AT set was significantly higher for structured searching of the indexed file than for unstructured searching of the unindexed file; 26.562 compared with 22.148 (0.01<P<0.02). The significance was less marked when the author's search statements were assessed in the BT set; 27.030 compared with 23.957 (0.02<P<0.05). When comparing scores for hits of major relevance, the AM set showed significantly higher scores for structured searching of the indexed file (25.782 versus 21.912; 0.02<P<0.05), but there was no significant difference between the scores in the BM set (26.365 versus 25.130; P >0.4).

In the group II queries from the information pharmacist with less experience of Martindale Online the AT set provided scores for structured searching of the indexed file of 25.253 compared with 22.310 for unstructured searching of the unindexed file (0.02<P<0.05),

see Table 11. Unexpectedly the difference was somewhat greater when the BT scores were compared; respective scores were 26.052 versus 21.409 ($P < 0.001$).

SET	RELEVANT HITS	S3	S1	P
AT	major & minor	25.253	22.310	$0.02<P<0.05$
BT	major & minor	26.052	21.409	<0.001
AM	major	25.511	23.155	$0.1<P<0.2$
BM	major	26.073	22.216	$0.002<P<0.01$

Table 11 Scores for S3 Versus S1 for 32 Group II Queries

Comparisons of scores for hits of major relevance showed structured searching of the indexed file to produce significantly higher scores only with the statements generated by the author, the BM set, where the figures were 26.073 versus 22.216 ($0.002<P<0.01$). The scores in the AM set were 25.511 versus 23.155 ($0.1<P<0.2$).

The 32 queries that formed group III were from the information pharmacist with least experience of Martindale Online. None of the comparisons showed any significant difference between the scores obtained from searching the indexed file in a structured way and searching the unindexed file in an unstructured way, see Table 11 below.

SET	RELEVANT HITS	S3	S1	P
AT	major & minor	24.727	23.238	0.2
BT	major & minor	24.637	25.393	>0.4
AM	major	25.379	24.615	>0.4
BM	major	24.353	26.698	$0.2<P<0.4$

Table 12 Scores for S3 Versus S1 for 32 Group III Queries

5.2 STRUCTURED SEARCHING OF THE INDEXED FILE (S3) VERSUS UNSTRUCTURED SEARCHING OF THE INDEXED FILE (S2)

The scores for the total number of queries for the four sets are shown below in Table 13.

SET	RELEVANT HITS	S3	S2	P
AT	major & minor	76.542	67.606	<0.001
BT	major & minor	77.719	71.370	0.001<P<0.01
AM	major	76.672	69.456	0.001<P<0.01
BM	major	76.791	73.875	0.1<P<0.2

Table 13 Scores for S3 Versus S2 for All Queries

The total score for the structured searches using the information pharmacists' statements was 76.542 which was significantly higher than the score of 67.606 obtained from unstructured searching even of the indexed file (P <0.001). When statements generated by the author were used for comparison the difference was still significant, although at a lower level; 77.719 compared with 71.370 (0.001<P<0.01). Consideration only of the hits of major relevance showed a significantly higher score for structured searches of the indexed file in the AM set (0.001P<0.01), but not in the BM set (0.1<P<0.2).

When the group of queries selected by the information pharmacist with the most experience of Martindale Online (group I) were considered, structured searching yielded higher scores in the AT set, 26.562 compared with 21.653 (0.01<P<0.02) and in the BT set, though at a lower level of significance, 27.030 versus 24.262 (0.02<P<0.05), see Table 14. S3 also produced significantly higher scores in the AM set; 25.782 compared with 21.720 (0.02<P<0.05).
No significant difference was detected between the scores in the BM set.

SET	RELEVANT HITS	S3	S2	P
AT	major & minor	26.562	21.653	0.01<P<0.02
BT	major & minor	27.030	24.262	0.02<P<0.05
AM	major	25.782	21.720	0.02<P<0.05
BM	major	26.365	25.376	>0.4

Table 14 Scores for S3 Versus S2 for 34 Group I Queries

Analysis of the results for the group II set of queries (the information pharmacist with intermediate knowledge of the database) showed that structured searching of the indexed database provided higher scores in the AT set; 25.253 versus 22.447 (0.02<P<0.05), see Table 15. It did not show a significant difference when only the major hits were assessed; 25.511 versus 23.237 (0.1<P<0.2).

SET	RELEVANT HITS	S3	S2	P
AT	major & minor	25.253	22.447	0.02<P<0.05
BT	major & minor	26.052	21.680	<0.001
AM	major	25.511	23.237	0.1<P<0.2
BM	major	26.073	21.995	0.002<P<0.01

Table 15 Scores for S3 Versus S2 for 32 Group II Queries

Oddly results obtained with sets of these queries where the author had generated the search statements showed higher scores for S3 over S2 in both the BT and the BM sets; P <0.001 and 0.002<P<0.01 respectively.

The group III queries from the information pharmacist with little experience or knowledge of Martindale Online provided results in all sets that showed no significant difference between structured and unstructured searching, see Table 16.

SET	RELEVANT HITS	S3	S2	P
AT	major & minor	24.727	23.505	0.2<P<0.4
BT	major & minor	24.637	25.428	>0.4
AM	major	25.379	24.499	>0.4
BM	major	24.353	26.504	0.1<P<0.2

Table 16 Scores for S3 Versus S2 for 32 Group III Queries

5.3 UNSTRUCTURED SEARCHING OF THE INDEXED FILE (S2) VERSUS UNSTRUCTURED SEARCHING OF THE UNINDEXED FILE (S1)

There were no significant differences between any of the scores when all queries were considered or when groups I, II or III were considered (Tables 17, 18, 19 and 20)

SET	RELEVANT HITS	S2	S1	P
AT	major & minor	67.606	67.696	>0.4
BT	major & minor	71.370	70.759	>0.4
AM	major	69.456	69.682	0.4
BM	major	73.875	74.044	>0.4

Table 17 Scores for S2 Versus S1 for All Queries

SET	RELEVANT HITS	S2	S1	P
AT	major & minor	21.653	22.148	>0.4
BT	major & minor	24.262	23.957	>0.4
AM	major	21.720	21.912	>0.4
BM	major	25.376	25.130	>0.4

Table 18 Scores for S2 Versus S1 for 34 Group I Queries

SET	RELEVANT HITS	S2	S1	P
AT	major & minor	22.447	22.310	>0.4*
BT	major & minor	21.680	21.409	>0.4
AM	major	23.237	23.155	>0.4*
BM	major	21.995	22.216	>0.4

Table 19 Scores for S2 Versus S1 for 32 Group II Queries

SET	RELEVANT HITS	S2	S1	P
AT	major & minor	23.505	23.238	>0.4*
BT	major & minor	25.428	25.393	>0.4*
AM	major	24.499	24.615	>0.4*
BM	major	26.504	26.698	>0.4*

Table 20 Scores for S2 Versus S1 for 32 Group III Queries

The * against the P value in Tables 19 and 20 indicates that the differences in the rankings were too few for any value to be read off the tables; the largest P value was therefore selected.

5.4 HIGHLY STRUCTURED SEARCHING OF THE UNINDEXED FILE (S4) VERSUS UNSTRUCTURED SEARCHING OF THE UNINDEXED FILE (S1)

Once again none of the comparisons yielded a significant difference as can be seen from Tables 21 to 24.

SET	RELEVANT HITS	S4	S1	P
AT	major & minor	68.237	67.696	0.3<P<0.4
BT	major & minor	71.929	70.759	0.2<P<0.3
AM	major	72.558	69.682	0.05<P<0.1
BM	major	74.398	74.044	>0.4

Table 21 Scores for S4 Versus S1 for All Queries

SET	RELEVANT HITS	S4	S1	P
AT	major & minor	23.150	22.148	0.4
BT	major & minor	23.619	23.957	>0.4
AM	major	23.038	21.912	0.2<P<0.4
BM	major	24.067	25.130	>0.4

Table 22 Scores for S4 Versus S1 for 34 Group I Queries

SET	RELEVANT HITS	S4	S1	P
AT	major & minor	22.100	22.310	>0.4
BT	major & minor	23.205	21.409	0.05<P<0.1
AM	major	24.776	23.155	0.1<P<0.2
BM	major	23.800	22.216	0.2

Table 23 Scores for S4 Versus S1 for 32 Group II Queries

SET	RELEVANT HITS	S4	S1	P
AT	major & minor	22.987	23.238	>0.4
BT	major & minor	25.105	25.393	>0.4
AM	major	24.613	24.615	>0.4
BM	major	26.531	26.698	>0.4

Table 24 Scores for S4 Versus S1 for 32 Group III Queries

5.5 STRUCTURED SEARCHING OF THE INDEXED FILE (S3) VERSUS HIGHLY STRUCTURED SEARCHING OF THE UNINDEXED FILE (S4)

When all relevant hits were considered, searching the indexed file in a structured way yielded significantly higher scores than did highly structured search statements used against the unindexed file (Table 25).

SET	RELEVANT HITS	S3	S4	P
AT	major & minor	76.542	68.237	0.001<P<0.01
BT	major & minor	77.719	71.929	0.01<P<0.02
AM	major	76.672	72.558	0.05<P<0.1
BM	major	76.791	74.398	0.1<P<0.2

Table 25 Scores for S3 Versus S4 for All 98 Queries

However the difference was not significant when only hits of major relevance were considered.

When the three groups of queries were considered, group I provided significantly higher scores with S3 than S4 in all sets, see Table 26. In group II, only the results for BT showed S3 providing significantly higher scores than S4; none of the other sets showed any significant difference, see Table 27.

None of the sets in group III showed any significant difference between the two types of searching.

SET	RELEVANT HITS	S3	S4	P
AT	major & minor	26.562	23.150	0.02<P<0.05
BT	major & minor	27.030	23.619	0.02<P<0.05
AM	major	25.782	23.038	0.02<P<0.05
BM	major	26.365	24.067	0.05

Table 26 Scores for S3 Versus S4 for 34 Group I Queries

SET	RELEVANT HITS	S3	S4	P
AT	major & minor	25.253	22.100	0.1
BT	major & minor	26.052	23.205	0.02<P<0.05
AM	major	25.511	24.776	>0.4
BM	major	26.073	23.800	0.05<P<0.1

Table 27 Scores for S3 Versus S4 for 32 Group II Queries

SET	RELEVANT HITS	S3	S4	P
AT	major & minor	24.727	22.987	0.1<P<0.2
BT	major & minor	24.637	25.105	>0.4
AM	major	25.379	24.613	>0.4
BM	major	24.353	26.531	0.2<P<0.4

Table 28 Scores for S3 Versus S4 for 32 Group III Queries

CHAPTER 6 EVALUATION OF RETRIEVAL EFFECTIVENESS
FAILURE ANALYSIS

Failure analysis was carried out on each query following procedures similar to those used by Lancaster[56].

Analysis was only carried out, however, on the results of the searches executed according to the guidelines laid down in the User's Guide, i.e. the results of the S3 searches.

The comparisons of retrieval effectiveness with different types of searches carried out in the first part of this evaluation used a single score that incorporated both recall and precision. This made for ease of comparison within that part of the evaluation; it does not make for easy comparison with Lancaster's analysis which employed recall and precision ratios. These ratios were therefore calculated for the results of the S3A and S3B searches for all 98 queries (see Appendix J).

The average recall and precision ratios are given below in the introductions to the relevant sections.

The following pages contain detailed reports on the reasons for the various failures. Record numbers have been cited throughout these reports. Anyone wishing to follow up this failure analysis should be aware that the contents of the records used in this study are likely to have been altered at any updates carried out subsequent to this research; copies of the records that were used in the analysis are available from the author.

6.1 RECALL FAILURE

Of the 98 searches carried out for each A and B statement, 49 suffered some recall failure; 48 of these failures occurred in the A group and 42 in the B group. A total of 215 relevant records were not retrieved in the A group due to 241 factors; corresponding figures in the B group are 190 relevant records not retrieved due to 198 different factors. This is summarised in Table 29.

	A	B
Number of searches	98	98
Number with recall failure	48	42
Records not retrieved	215	190
Number of factors	241	198

Table 29 Summary of Recall Failures

The average recall ratios for A and B searches are given below together with the corresponding average from Lancaster's analysis.

	A	B	Lancaster
All records	60.2	65.4	57.7
Major records	69.3	73.2	65.2

Table 30 Average Recall Ratio

Table 31 overleaf summarises the reasons for the recall failure and the extent of that failure within each group.

		N° missed records		% missed records		N° searches involved		% searches involved	
		A	B	A	B	A	B	A	B
SEARCHING									
Statement	total	159	122	73.95	64.21	35	29	35.71	29.59
too limited	major	111	79	81.02	64.23	26	21	26.53	21.43
INDEXING									
Indexing	total	58	52	26.98	27.66	19	19	19.39	19.39
rules	major	40	33	30.77	27.50	12	12	12.24	12.24
Indexing	total	13	14	6.05	7.45	10	10	10.20	10.20
deficiency	major	8	8	6.15	6.66	6	7	6.12	7.14
Indexing	total	8	8	3.72	4.26	4	4	4.08	4.08
sentence	major	7	7	5.38	5.83	3	3	3.06	3.06
Thesaurus	total	1	1	0.47	0.53	1	1	1.02	1.02
limitation	major	0	0	-	-	-	-	-	-
PROCESSING									
Systems	total	7	7	3.26	3.72	5	5	5.10	5.10
failure	major	6	6	4.62	5.0	4	4	4.08	4.08

Table 31 Reasons for Recall Failure

6.1.1 RECALL FAILURE DUE TO SEARCH STATEMENT BEING TOO LIMITED

In almost all cases the recall failure in this category was due to the search statement being too narrow or specific. In query 68, for example, the request was for information on carbamazepine in pregnancy. The A search statement limited the search to adverse effects in pregnancy ignoring precautions and absorption; also it did not incorporate the PREG qualifier. As a result 7 records were missed. Sometimes the search statement was made deliberately specific or general knowing that this might lead to recall failure as in query 1 which asked for alternative CNS stimulants to Ritalin. This meant retrieving all the records that dealt with CNS stimulants, Ritalin was not excepted. To limit the large number of records, only compounds classified as CNS stimulants in the thesaurus were specified. This excluded 6 records on compounds with stimulant activity but which were otherwise classified in the thesaurus. Also to limit the numbers further, the author printed only the Uses text records; this excluded 5 relevant records that were other than Uses text records.

In other cases the search statements matched the query but because of the relevance assessment had to be treated as being too specific or too general. Query 4 asked for drugs that could be used orally to reduce facial hair. The search statements incorporated oral administration as a concept and failed to retrieve 10 records that were assessed as relevant although they made no mention of oral administration. Another example is typified by query 19 which was a request for suitable analgesics for a patient with cirrhosis. The 6 records not retrieved but considered relevant did not deal with cirrhosis but more generally with liver damage or jaundice. In one query, 57, information was requested on nitrazepam causing depression of mood. The search statement asked for mental depression; a document dealing with CNS depression was deemed relevant, but that relevance has to be doubted. Similar cases occur in recall failure attributed to indexing problems.

Only 2 questions suffered because of problems with the logical operator. In the A search statement of query 97 a nondescriptor and a descriptor were linked by the operator WITH which demands that the two terms be in the same sentence. That demand could not be met and 11 records were not retrieved. The other query was 55 which was a request for substances that turn pink when wet. This was a difficult test for Martindale Online and the B search only retrieved a relevant record

through a contrived free-text search based on detailed knowledge of the file. The A statement used WITH, limiting the search too severely.

6.1.1.1 DETAILS FOR EACH QUERY

Query 1 What alternative CNS stimulants are there to Ritalin?

The A statement failed to retrieve 4 records, 2 of them classed as of major relevance; the B statement failed to retrieve 10 records, 3 of which were of major relevance. The identity numbers of these records are as follows (an asterisk signifies major relevance):

A	B
622-a5-10-b*	622-a5-10-b*
2067-a1-r	2067-a1-r
1419-a6-1-d	1419-a6-1-d
1410-k*	1410-k*
	1487-a5-y
	1478-a5-l
	1472-a5-g
	622-a5-t*
	2067-a3-e
	2058-a5-d

Such a general question meant retrieving all the records that dealt with CNS stimulants. To reduce the number of records that this yielded (127 for the A search), only compounds classified as CNS stimulants in the thesaurus were selected and the selection was further limited by printing only the uses text records. The discrepancy between the A and B sets is due to the A statement including the term CENTRAL STIMULANT. This is a descriptor from the action facet and it was assigned to records describing the action of such compounds as caffeine and appetite suppressants that were not classified as CNS stimulants within the drugs section of the thesaurus.

Thus for several reasons, these recall failures are attributed to the limitations of the search statements.

Query 3 Are there any new drugs for phenylketonuria?
Record 2156-n93-3-r was not retrieved by either the A or B search; it dealt with the diagnosis of this condition, not its treatment. Diagnosis was not considered in either search statement and could be considered as a limitation.

Query 4 Can anything be given orally to reduce facial hair?
Both the A and B statements specified the oral concept and as a result both excluded 10 records considered to be of major relevance although they made no reference to the oral route.

Query 8 Is there an interaction between prednisolone and oral contraceptives?
Record 1060-a5-34-m provided information on corticosteroids interacting with hexoestrol and is of some relevance to the query. Neither the A nor B search retrieved this record; a more general search statement would have done.

Query 10 What gelatin plasma expanders are available?
The A and B searches failed to retrieve 9 relevant records; 4 of these records were considered to be of major relevance. None of the 9 provided any information on availability or on preparations and both search statements specified these features. The records judged to be of major relevance dealt with the uses of gelatin, oxypolygelatin and polygeline (which are used as plasma expanders); the records of minor relevance dealt with the adverse effects and precautions of gelatin.

Query 17 Stability of phenobarbitone in aqueous solution
Five records, all of major relevance, were not retrieved by the A or the B search since both search statements specified the term SOLUTION. None of the 5 records contained this descriptor; they dealt with mixtures or did not deal with solutions in sufficient detail to be indexed. A more general search statement would have retrieved these records

Query 19 Which analgesic should be given to a patient with cirrhosis?

Cirrhosis was specified in both search statements. The 5 relevant records not retrieved, 4 of which were assessed as being of major relevance, dealt with liver damage or jaundice. The statements were too specific.

Query 23 Why do you need to take plenty of water with lithium?

Two records (1 major) were not retrieved by the A search being too specific. One was a record dealing with adverse effects and not precautions as specified, the other dealt with the use of intravenous fluids and not water as specified. A third relevant record (5057-a3-w) was not retrieved by either the A or B search statements; in this case relevant information was contained within the text which discussed dehydration without specifically mentioning water or fluids which were both specified in the search statements.

Query 29 Ampicillin in pregnancy

This query was interpreted as calling for precautions information and the PRECS qualifier was used in both A and B search statements. Neither statement retrieved 5 records considered to be of major relevance as they dealt with use or absorption; understandably they were not indexed with PRECS.

Query 30 How to collect large spillages of mercury?

Record 5306-a1-25-e* mentions problems of mercury spillage but did not deal with methods of disposal. The search statements were too specific to retrieve this record which was assessed as being of major relevance.

Query 37 Toxicity of dyestuffs used in food

The A search failed to retrieve 31 relevant records, 26 of which were considered to be of major relevance. The B search failed to retrieve 17 records of which 14 were considered to be of major relevance. In all cases, the search was too specific; either the qualifier AE or the descriptor FOODS was the cause of the failure. Often the information in the records considered to be relevant was present in an implied form that could only have been retrieved by searching on a drug name and AE. There is also an overlap with recall failure due to indexing; 8 of the A

records, one of which is also a B record, displayed some indexing peculiarities and are also considered in that section of the failure analysis.

Query 39 Treatment of tuberous sclerosis

Record 6602-a1-64-s was not retrieved by the B search statement which specified use? or treat?. The record which was of minor relevance dealt with adverse effects; the statement was therefore too specific.

Query 40 Local anaesthetic injection reactions

The A search failed to retrieve 10 records of which 9 were assessed as of major relevance. The B search did not retrieve 6 records, all of them major. The records are listed below with comments against each on the specificity of the search statements. The A search statement was:

LOCAL-ANAESTHETICS£ WITH AE WITH INJECTION

The B search statement was:

LOCAL-ANAESTHETICS£ WITH AE WITH (INJECTION OR PARENTERAL-ROUTE)

A	B
7609-a4-54-n*	7609-a4-54-n*

This would have been retrieved if AEM had been used in the search statement instead of AE (and if the route had been indexed, therefore also under index failure).

A	B
7600-a1-x*	7600-a1-x*
7601-a1-1-h*	7601-a1-1-h*
7620-a1-n*	7620-a1-n*

These records were indexed with INTRAVENOUS ROUTE. B should have used PARENTERAL-ROUTE£. A was also too specific in limiting the search to INJECTION and not a route.

A	B
7601-a1-46-q*	7601-a1-46-q*
7601-a1-9-e*	7601-a1-9-e*

These records dealt with the adverse effects of epidural block and details of the block were not included, thus not indexed other than with EPIDURAL BLOCK. Specifying INJECTION in A was too specific; specifying PARENTERAL ROUTE in B was too general.

A

7630-a1-m*

7644-a1-4-h*

7644-a1-5-g*

141-a1-f

The first 3 of these records were indexed with PARENTERAL ROUTE, thus INJECTION in the A statement was too specific. 141-a1-f also had an indexing limitation, but allowing for that, INJECTION was too specific and contributed to the recall failure.

Query 41 Hydroquinone cream - use and formulation

Neither the A nor the B searches retrieved 1612-a1-1-m* which is an adverse effects record not indexed with USE. Although the relevance is doubtful here, this failure has to be attributed to limitation of the search statements which called for USE.

The B search failed to retrieve 1612-a2-1-r since the record did not specify the form of the preparation and OINTMENT OR TOPICAL APPLICATION in the search statement was thus too specific. B also failed to retrieve 1612-n1-k* which described a lotion of hydroquinone; this record was indexed with the descriptor LOTION, a term distinct from that used for cream (OINTMENT). The failure has been attributed to the search statement being too specific, although again there is some doubt about the relevance.

Query 42 Details of a honey dressing

Two records of major relevance not retrieved by either the A or B search did not specify a dose form such as dressing; both statements asked for DRESSINGS£ and the B statement also asked for TOPICAL DOSAGE FORM as an alternative. The statements were thus too specific.

Query 43 Which vitamin is used to treat fragile capillaries?

Three records considered to be of major relevance were not retrieved by either the A or the B search since both statements asked for VITAMINS£. The 3 records dealt with biflavanoids which are not classified as vitamins in the thesaurus, but there is a related-term link betwen them and the vitamins. It could therefore be considered that the statements should have taken account of the link and included biflavanoids.

Query 44 Effects of mepacrine on the liver

Neither A nor B retrieved one record considered to be of major relevance since both statements asked for AE; the relevant record stated that mepacrine is concentrated in the liver. The statements were too specific.

Query 45 Carcinogenicity of Sudan IV

The A statement asked for CARCINOGEN and failed to retrieve one record (2274-a1-v) of dubious relevance. This record did not deal with carcinogenicity; it contained the following:

ointments containing more than 5% of scarlet red [Sudan IV] *are irritant and may cause systemic effects.*

Failure has however, been attributed to the statement being too specific.

Query 49 Reports of progesterone causing cancer

One record not retrieved by either A or B dealt with neoplasms and not malignant neoplasms as requested; this record also dealt with chlormadinone and not progesterone. Another record not retrieved, but considered as relevant, dealt with hyperplasia associated with progesterone but not malignant neoplasms. These failures have been attributed to the statements being too specific.

Query 53 Moduretic causing hyponatraemia

Moduretic is the proprietary name for a combination of two drugs, amiloride and hydrochlorothiazide. The 5 records considered to be relevant (1 of which was of major relevance) that were not retrieved by either the A or the B search dealt only with hydrochlorothiazide; the search statements specified both drugs and were thus considered to be too specific.

Query 54 Is clindamycin absorbed topically?

Record 54-a7-2-n* was not retrieved by the A search statement which was

CLINDAMYCIN WITH SKIN WITH DRUG-ABSORPTION.

It was not clear in the text of the record if the absorption was through the skin, therefore it was not indexed. The statement was too specific.

Query 55 Substance that turns pink when wet

The A search failed to retrieve the one relevant record through being too specific in using the WITH qualifier instead of AND. This was a difficult query for Martindale Online and it is unlikely that anyone without a good knowledge of its contents would be able to retrieve the relevant record.

Query 57 Nitrazepam causing depression of mood

Record 4054-a2-1-p was not retrieved by the A or the B search; it dealt with CNS and not mental depression which was what was asked for in the search statements. The statements were considered as being too specific, although the relevance assessment is doubtful in this instance.

Query 60 Iodine content of kelp.

The B search did not retrieve 4571-a1-24-c which dealt with seaweed and did not mention kelp. Kelp is not a descriptor and was used in the B search as a free-text term, hence the recall failure. The A search asked for ALGAE£ and successfully retrieved this record.

Query 63 Use of ephedrine nose drops in neonates.

Record 2058-a5-d* which was not retrieved by either A or B dealt with children and not neonates; the statements called for neonates and hence were too specific.

Query 67 Possible adverse effects of topically applied menthol.

The A statement was too specific in not retrieving 2 records of major relevance. One was not retrieved because TOPICAL DOSAGE FORM was requested instead of TOPICAL APPLICATION. The other record dealt with an ointment which would have been retrieved if TOPICAL DOSAGE FORM£ had been specified.

Query 68 Carbamezepine in pregnancy.

The A statement was too specific in restricting the search with the AE qualifier; as a result it failed to retrieve 7 records (6 major) that dealt with the absorption of carbamazepine in pregnancy or the precautions to be taken when using the drug in pregnancy.

Query 70 Calcium and vitamin D supplement for a four-year-old unable to take milk (what doses and source)?
The A search did not retrieve 15 records (11 major) and the B search 13 records (9 major). In A's case 9 records (6 major) were not retrieved because they dealt with calcium salts and not calcium as requested in the descriptor field; the remaining 6 records (5 major) were not retrieved by A because VITAMIN D SUBSTANCES was specified and not VITAMIN D SUBSTANCES&.
The B search did not retrieve 4 records (2 of major relevance) because of the request for an inclusive search on CALCIUM SALTS; the records contained descriptors with the word calcium, but calcium was not the important part of the molecule and the substances had not been classified under the CALCIUM SALTS hierarchy. The remaining 9 records were not retrieved because B specified CALCIUM SALTS WITH VITAMIN D SUBSTANCES& and they either dealt solely with one or other of the groups of compounds.
The search statements were too specific.

Query 73 Increasing aqueous solubility of salicylic acid - which polyethylene glycol to be used (i.e. which one is best and at what percentage)?
Both the A and the B search asked for SOLUBILITY and as a result failed to retrieve the one record considered to be relevant that dealt not with solubility but with incompatibility!

Query 78 Strength of saltpetre to use as a preservative.
Saltpetre is potassium nitrate and that was specified in both the A and B search statements. Record 13254-a2-x* dealt with sodium nitrate which has similar properties to the potassium salt. The statements were obviously too specific.

Query 85 Use of water-soluble vitamin K in malabsorption.
Two relevant records (1 major) were not retrieved by either A or B. Neither record dealt with the soluble aspect of vitamin K and, as this was asked for in different ways in both statements, they have to be considered as too specific.

Query 87 Parenteral formulation of thiazide diuretic.
Both statements asked for proprietary or nonproprietary preparation information. The 4 records considered to be of minor relevance that were not retrieved by either A or B did not deal with any preparation information but provided some data about the route of administration. The statements could be considered to be too specific.

Query 91 Is citric soda the same as potassium citrate?
The A search failed to retrieve 1180-t because it included COMP [comparison] in the search statement making it too specific; it also contained some odd logic.

Query 92 In which infusion fluids is ampicillin unstable?
Seven relevant records (2 major) were not retrieved by the A search statement because it included the term INFUSION as a free-text word or descriptor. The B search statement was also too specific in that it failed to retrieve 8 records (3 major) because it included PARENTERAL ROUTE OR PARENTERAL DOSAGE FORM and, in the case of 1 record, did not include INFUSION as a free-text search term.

Query 96 Can you preserve hydrogen peroxide solution, if so with what?
Neither the A nor B search statements contained DISINFECTANTS£, but limited the search to concepts of stability or preservation. As a result record 2206-r* was not retrieved. Also A did not retrieve 5904-h by being too specific in asking for STABILITY and not STABILITY OR STABILISATION.

Query 97 Are there any official preparations containing rose or extract of rose?
There were 11 recall failures (9 major) in the A search due to the search statement linking a nondescriptor and a descriptor with the operator WITH. The statement is:

ROSE WITH NPROP

Using AND as the operator would have retrieved the records. However, because of indexing policy and indexing failure, using the descriptor for rose (ROSE-OIL) would not have retrieved them; despite that, the search statement was considered to be too specific.

6.1.2 RECALL FAILURE DUE TO INDEXING RULES

Table 31 below summarises the 12 reasons for recall failure due to some aspect of the indexing guidelines. Each of these reasons is then discussed in more detail on the following pages.

		Missed records		% missed records		Searches involved		% searches involved	
		A	B	A	B	A	B	A	B
Dose not generally indexed for AE data	total	17	18	7.91	9.57	2	3	2.04	3.06
	major	13	14	10.0	11.67	2	3	2.04	3.06
Minor ingredients of preparation not indexed	total	7	11	3.26	5.85	3	4	3.06	4.08
	major	3	5	2.31	4.17	1	2	1.02	2.04
Water not indexed outside its monograph	total	9	1	4.19	0.53	2	1	2.04	1.02
	major	8	1	4.26	0.83	2	1	2.04	1.02
Records not indexed to avoid duplication	total	5	5	2.33	3.85	4	4	4.08	4.08
	major	1	1	0.77	0.83	1	1	1.02	1.02
PRECS not routinely indexed for AE data	total	7	7	3.26	3.72	3	3	3.06	3.06
	major	6	6	4.62	5.0	2	2	2.04	2.04
Incidental data not indexed	total	4	4	1.86	2.13	4	4	4.08	4.08
	major	2	2	1.54	1.67	2	2	2.04	2.04
AE and USE data indexed as sep. sentences	total	3	3	1.40	1.60	1	1	1.02	1.02
	major	3	3	2.31	2.50	1	1	1.02	1.02
PREG and PREGNANCY not indexed in same sentence	total	2	-	0.93	-	1	-	1.02	-
	major	2	-	1.54	-	1	-	1.02	-

(cont'd)

		Missed records		% missed records		Searches involved		% searches involved	
		A	B	A	B	A	B	A	B
Incompatibility data not indexed in detail	total	1	1	0.47	0.53	1	1	1.02	1.02
	major	-	-	-	-	-	-	-	-
Indications of preparations not indexed	total	1	1	0.47	0.53	1	1	1.02	1.02
	major	-	-	-	-	-	-	-	-
SOU qualifier not used for general information	total	1	1	0.47	0.53	1	1	1.02	1.02
	major	1	1	0.77	0.83	1	1	1.02	1.02
Minor information indexed with a minor qualifier	total	1	-	0.77	-	1	-	1.02	-
	major	1	-	0.77	-	1	-	1.02	-

Table 32 Recall Failures due to Indexing Rules

6.1.2.1 DOSAGE INFORMATION NOT ROUTINELY INDEXED FOR ADVERSE EFFECTS DATA

Missed Record A Search	Missed Record B Search	Query
	5057-a1-147-x*	33
2042-a1-1	2042-a1-1	40
7609-a4-54-n*	7609-a4-54-n*	40
141-a1-5-t	141-a1-5-t	40
860-a1-2-z	860-a1-2-z	40
7600-a1-7-h*	7600-a1-7-h*	40
7600-a1-17-w*	7600-a1-17-w*	40
7600-a1-34-j*	7600-a1-34-j*	40
7600-a1-4-j*	7600-a1-4-j*	40
7601-a1-23-m*	7601-a1-23-m*	40
7601-a1-28-d*	7601-a1-28-d*	40
7609-a1-2-j*	7609-a1-2-j*	40
7609-a1-3-a*	7609-a1-3-a*	40
7617-a3-t*	7617-a3-t*	40
7633-a1-6-q*	7633-a1-6-q*	40
7644-a1-5-g*	7644-a1-5-g*	40
7751-a2-q*	7751-a2-q*	79
9205-a3-z	9205-a3-z	79

Table 33 Records not Retrieved due to Dosage Indexing Rule

Dosage information is conveyed by the AD qualifier. In considering the use of this qualifier, it was envisaged that searchers wanting to obtain data on recommended doses would best obtain it through AD. As a consequence, information about side-effects would not be relevant and this led to the indexing rule that AD should not be used routinely when indexing side-effects. It would be used when the dose played an important part in the adverse effect; appropriate descriptors would also be used to cover overdosage and poisoning.

As the first update was being prepared it was realised that this was not a sound rule and the guidelines were changed so that AD would be

indexed for any dosage information. Searchers wanting recommended doses could obtain them through AD WITH USE.

6.1.2.2 MINOR INGREDIENTS OF PREPARATIONS NOT INDEXED

Missed Record A Search	Missed Record B Search	Query
6402-n11-j	6402-n11-j	2
1306-n1-g	1306-n1-g	2
2246-n5-f	2246-n5-f	2
5602-n11-j*	5602-n11-j*	27
5602-n13-b*	5602-n13-b*	27
5602-n22-f*	5602-n22-f*	27
7890-n2-y	7890-n2-y	42
	699-n1-s	97
	1901-n10-r	97
	5279-n1-g*	97
	6408-n5-w*	97

Table 34 Records not Retrieved due to Method of Indexing Preparations

Preparations of medicines usually contain several ingredients. Some preparations are mixtures of active ingredients, others contain one major active ingredient and a number of other substances for flavouring, preserving, stabilising, or other reasons to do with formulation. These ingredients may be inactive or possess activities irrelevant to the indications for which the preparation is used, e.g. peppermint water in many mixtures. It was therefore decided to index only significant preparation information, otherwise there would be thousands of hits for descriptors such as SODIUM CHLORIDE. Searchers wishing to retrieve minor ingredients could do so by way of free-text searching. Unfortunately searchers would have to know that the compounds in which they were interested were minor ingredients, or at least guess as much after a fruitless search using drug descriptors. And this leads to the problem of a searcher knowing that 'zero hits' is a true response to a

search or whether it is an indication of the need for further refining of the search statement.

In considering the records that were not retrieved because minor ingredients of preparations were not indexed, free-text searching in query 2 retrieved the 3 records not retrieved by the S3 searches. Free-text searching did not retrieve anything for query 27, but it did retrieve the missed relevant records for queries 42 and 97.

The relevance assessment for the 3 missed records in query 2 was idiosyncratic. This query called for new uses of lavender oil; one of the missed relevant records described a preparation of aromatic vinegar from a 1934 formulary. Another missed record described Formaldehyde Solution with Soap, also from a 1934 source. The third record described a massage lotion used at Hadassah University Hospital. Lavender oil was a very minor ingredient in all 3 preparations and its use as a perfume in all of them can hardly be described as new.

6.1.2.3 WATER NOT BEING INDEXED OUTSIDE ITS MONOGRAPH

Missed Record A Search	Missed Record B Search	Query
673-n7-z*		17
4061-y*		17
4061-n1-m*		17
4061-n3-c*		17
4061-n3-1-x*		17
4061-11-r*		17
4061-13-e*		17
2317-a3-y		23
5057-a1-147-b*	5057-a1-147-b*	23

Table 35 Records not Retrieved due to Water not being Indexed outside its Monograph

WATER is a descriptor in Martindale Online and was designed for use with water as the subject of a Martindale monograph. Water mentioned as a solvent or vehicle for other drugs was not considered worth indexing

since it occurred several thousand times (4626) in the database and would have made retrieval of records from the Water monograph exceedingly difficult. The author does not envisage a change in this indexing rule. The fact that the B search only failed to retrieve 1 record indicated that the author's knowledge of this rule was an advantage. Making searchers aware of the special nature of the WATER descriptor might help those who took notice of such hints and remembered to act on them.

6.1.2.4 AVOIDANCE OF DUPLICATE INDEXING

Missed Record A Search	Missed Record B Search	Query
884-a3-14-g	884-a3-14-g	15
884-a3-16-j	884-a3-16-j	15
1590-j	1590-j	41
12821-a1-s*	12821-a1-s*	43
263-a5-1	263-a5-1	65

Table 36 Records not Retrieved due to Avoidance of Duplicate Information

The 2 records not retrieved in query 15, which was for information on the use of minoxidil in baldness, were both abstract records that dealt with the uses of minoxidil and mentioned hypertrichosis as an adverse effect. This effect was also mentioned in the appropriate adverse effects abstracts where it was indexed. A similar case occurred with query 65 which was on the hazards of camphorated oil. The record listed as a recall failure covered the uses of camphor and stated that the 'use of camphorated oil is discouraged because of its toxicity'. This point was also covered in the adverse effects section which was retrieved. The indexing rule was not to index adverse effects data mentioned in uses records, if that point was adequately covered under the appropriate adverse effects section. In other words, the indexing was to be guided by the editorial emphasis on the data. This decision calls for vigilance on the part of the indexer as it would be easy to

ignore the rule, and in recent indexing there appears to be a tendency in that direction.

The record considered to be a recall failure in query 41 was an introduction record that provided information covered in more detail elsewhere; this introduction record was intentionally not indexed. Similarly the record considered to have been missed in query 43 was a cross-reference to another record, so once again there was no point in indexing a lead-in to another record that was already adequately indexed.

6.1.2.5 PRECAUTIONS NOT ROUTINELY INDEXED WITH ADVERSE EFFECTS

Missed Record A Search	Missed Record B Search	Query
6214-a2-9-w*	6214-a2-9-w*	22
2381-a1-4-n*	2381-a1-4-n*	37
2386-a1-1-k*	2386-a1-1-k*	37
2411-a2-1-x*	2411-a2-1-x*	37
2415-a1-1-g*	2415-a1-1-g*	37
7600-a3-w	7600-a3-w	40
7620-a6-f*	7620-a6-f*	40

Table 37 Records not Retrieved due to Precautions not being Indexed with Adverse Effects

Once again the indexing was guided by the editorial selection. Data considered for the precautions section of a monograph is intentionally separated from data for the adverse effects section.

The record not retrieved in query 22 was an adverse effects record that was not indexed with PRECS; the search statement included PRECS. The first record listed against query 40 in the above table showed the converse: a precautions record not indexed with AE; in this case AE was specified in the search statement. There could be some arguement for adding PRECS to an adverse effects record, but there is none for the converse. Conditions that form precautions should not be confused with conditions generated as adverse effects; precautions are not adverse

effects and AE cannot be added routinely to precautions records. The second record was a uses record that contained a warning that cocaine should not be given by injection for local anaesthesia; this was indexed with PRECS but not with AE.
The first 3 recall failures listed above against query 37 (toxicity of dyestuffs used in food) were general precautions records on recommended intakes, they contained incidental adverse effects data not considered worth indexing - data such as *toxicological studies were required.*

6.1.2.6 INCIDENTAL DATA NOT INDEXED

Missed Record A Search	Missed Record B Search	Query
4904-a8-193-p*	4904-a8-193-p*	29
2411-a1-12-j*	2411-a1-12-j*	37
7601-a4-81-v	7601-a4-81-v	40
12887-a1-15-s	12887-a1-15-s	79

Table 38 Records not Retrieved due to Incidental Data not being Indexed

Record 4904-a8-193-p* dealt with the treatment of infections in a variety of patients, some of whom were pregnant. No further point was made of the pregnant state which was not deemed worth indexing, hence this record was not retrieved for the query on ampicillin in pregnancy. Also this record happened to be one where USE and AE were indexed in the same sentence since cross-over with the other descriptors was considered to be of little significance.
2411-a1-12-j* included a mention of food that was considered to be incidental as regards indexing. However, it did deal with reactions to tartrazine, so FOOD ADDITIVES could have been worth adding as a descriptor, in which case it would have been retrieved by the B search.
7601-a4-81-v only provided a passing mention of the adverse effects of epidural injection. It was intentionally not indexed and therefore not retrieved for the query on local anaesthetic injection reactions.

Query 79 asked for data on amyl nitrite in cyanide poisoning, the record listed above that was not retrieved dealt with laetrile poisoning and the delay in treatment because it was not realised that laetrile released cyanide. The treatment measures were mentioned but considered to be standard and incidental to the main point on the delay.

6.1.2.7 ADVERSE EFFECTS AND USES DATA INDEXED IN SEPARATE SENTENCES

Missed Record A Search	Missed Record B Search	Query
9105-a4-20-f*	9105-a4-20-f*	98
9081-a3-23-n*	9081-a3-23-n*	98
9105-a1-r*	9105-a1-r*	98

Table 39 Records not Retrieved due to Adverse Effects and Uses Data being Indexed in Separate Sentences

Query 98 asked for information on the adverse effects of the 'morning-after pill'. The search statement for both the A and the B search was POSTCOITAL-CONTRACEPTION WITH AE. The first two recall failures listed above dealt with stilboestrol and levonorgestrel respectively for postcoital contraception and contained minimal information on adverse effects. The third recall failure dealt with the adverse effects of stilboestrol and stated that *nausea is common with the high doses employed for postcoital contraception.*

In the first two records AE was not indexed partly because of the guideline that indexing should follow the editorial emphasis which in these cases was on uses. The third record was not indexed with POSTCOITAL CONTRACEPTION since the emphasis was on the adverse effects of stilboestrol. Here the indexers were influenced by the concern that indexing AE and POSTCOITAL CONTRACEPTION would be confusing a use with an adverse effect.

The 3 recall failures if completely indexed would have had as part of their indexing:

1. STILBOESTROL, USE, POSTCOITAL CONTRACEPTION....
   STILBOESTROL, AE....

2. LEVONORGESTREL, USE, POSTCOITAL CONTRACEPTION....
   LEVONORGESTREL, AE....
3. STILBOESTROL AE, NAUSEA....
   STILBOESTROL, USE, AD, POSTCOITAL CONTRACEPTION...

Such indexing was not considered worthwhile. Even if it had been carried out, the search statements which employed - 'POSTCOITAL-CONTRACEPTION WITH AE' would still have failed.

6.1.2.8 PREG AND PREGNANCY NOT INDEXED IN THE SAME SENTENCE

Missed Record A Search	Missed Record B Search	Query
6214-a2-9-w*		22
6214-a1-4-a*		22

Table 40 Records not Retrieved due to PREG and PREGNANCY not being Indexed in the Same Sentence

PREG was envisaged as a qualifier that would encompass pregnancy and the neonate. Early indexing rules therefore instructed that PREG was sufficient to cover the state of pregnancy; the PREGNANCY descriptor would be used for details of pregnancy. This was found to be an unwise decision and latterly PREGNANCY has been assigned in a more general fashion. The above 2 records should therefore have received PREGNANCY as a descriptor.

6.1.2.9 INCOMPATIBILITY DATA NOT INDEXED IN DETAIL

Only 1 record was not retrieved for this reason by either the A or B search: record 9-2-y in query 92. This record dealt with the incompatibility of ampicillin in vitro in an infusion solution. Indexing incompatibility data could get as complicated as indexing interaction data, with separate sentences for each incompatibility or for each component of a mixture if incompatible with another substance. It was therefore decided only to index the substance that demonstrated the incompatibility and the fact of the incompatibility; the substances with

which it was incompatible would not be indexed.. It is unlikely that this rule will change.

6.1.2.10 INDICATIONS OF PREPARATIONS NOT INDEXED

Record 7911-n16-a, which was not retrieved by either the A or the B search in query 3, described a multivitamin preparation for use in phenylketonuria.

Indexing instructions were not to index such preparations in detail as it was best to have the detailed indexing under records describing the individual active ingredients. This rule has subsequently been changed and any indications provided for mixed preparations are now being indexed.

6.1.2.11 SOU QUALIFIER NOT USED FOR GENERAL INFORMATION

Record 5425-w* was not retrieved by the A or B searches for query 32. SOU is a qualifier used for the source of a drug and is usually indexed with the relevant drug and the family of the species that acts as its source. This missed record stated that gelatin *...is a protein obtained by partial hydrolysis of animal collagenous tissues such as skin.....* The indexer considered this statement too general for indexing with SOU. Current indexing does not interpret that guideline so strictly and SOU would now be indexed with GELATIN.

6.1.2.12 MINOR INFORMATION INDEXED WITH THE MINOR QUALIFIER

Record 9237-12-a* was not retrieved by the A search in query 77 because it was indexed with the minor qualifier PHARMM and not the ordinary qualifier PHARM; the search statement specified PHARM. The record dealt with glyceryl trinitrate and its adsorption onto plastics. The major information was about glyceryl trinitrate and it was considered that a searcher wanting to obtain information about the pharmacy of plastics would not want this record but the monograph record from the Plastics monograph. By indexing this record with PHARMM and PLASTICS the searcher was given the facility to retrieve it and other secondary pharmacy data or to exclude it and request only the key pharmacy information. An inclusive search on PHARM& would retrieve PHARM and PHARMM; a search on PHARM would retrieve the significant information. The information pharmacist formulating the A statement did not make use of this facility, the author did.

A similar policy was adopted with interaction information. The sample queries did not test whether this policy worked effectively.

6.1.3 RECALL FAILURE DUE TO INDEXING DEFICIENCY

		Missed records		Searches involved	
		A	B	A	B
Inexcusable omission of descriptors	total	7	7	6	6
	major	5	6	4	5
Indexing not sufficiently detailed	total	5	6	5	6
	major	2	2	2	2

Table 41 Recall Failures due to Indexing Deficiency

6.1.3.1 INEXCUSABLE OMISSION OF DESCRIPTORS

Missed Record A Search	Missed Record B Search	Query
2317-a3-y	2317-a3-y	23
592-a1-1-g*	592-a1-1-g*	42
7910-c*	7910-c*	43
15304-a1-1*	15304-a1-1*	43
9021-f*	9021-f*	59
5408-a1-b	-	60
-	9237-12-a*	77
12887-a1-b*	12887-a1-b*	88

Table 42 Records not Retrieved due to Inexcusable Omission of Descriptors

Omissions due to human or system failure are included here so that there is some overlap with failure due to system faults (see below). The first two records listed in Table 42 against queries 23 and 42 contained

no descriptors. While this might be a systems fault (and is listed there) it could be an indexing omission.

The third record, 7910-c* might well have been subject to a systems fault in that the internal database carried descriptors for it, but no descriptors were seen on Data-Star's version. By some coincidence the descriptors that were included in the internal version were very general and more detailed indexing should have been carried out, so if there had not been some fault with the location of the descriptors, this record would still have been listed as a recall failure in the next group.

Record 15304-a1-1* dealt with the uses of rutin; it made no reference to hesperidin. However, hesperidin was reported in another record to have similar uses to rutin and this point should have been indexed in the rutin record. This indexing omission meant that the rutin record was not retrieved in response to a request for information on vitamins for fragile capillaries.

Record 9021-f* should have been indexed with the qualifier DES since it contained a description of the different types of oral contraceptives. This error might be explained by the indexers associating DES just with the description of individual drugs in the monograph record.

Record 5408-a1-b described kelp as a source of iodine; iodine was not indexed.

In 9237-12-a* sorbtion should have been indexed.

Finally 12887-a1-b* was a record about laetrile which should have been indexed with the descriptor BENZALDEHYDE.

6.1.3.2 INDEXING NOT SUFFICIENTLY DETAILED

Missed Record A Search	Missed Record B Search	Query
2224-a3-p	2224-a3-p	16
-	5057-a3-64-v	23
1612-a1-2-q*	1612-a1-2-q*	41
7890-n2-y	7890-n2-y	42
827-a4-c*	827-a4-c*	76
6630-a3-p	6630-a3-p	78

Table 43 Records not Retrieved due Insufficiently Detailed Indexing

Record 2224-a3-p dealt with disinfection of water supplies and was indexed with:

WATER SUPPLY, DISINFECTANT, CONTAMINATION, DISINFECTION.

The term purification was requested in the search statement; as it is a broader term than DISINFECTION it could have been added. On the other hand it could also be argued that the recall failure was due to the search statement being too limited, or to a consequence of the indexing guidelines that instruct indexers always to use the most specific descriptors.

Record 5057-a3-64-v could have been indexed with the additional descriptors AD and FLUID AND ELECTROLYTE REPLACEMENT as well as descriptors for the relevant routes of administration. The text of the record and its indexing were as follows:

A warning that if patients with lithium-induced polyuria must not take fluids by mouth before an operation, they should be given intravenous fluids the night before the operation.

DESCRIPTORS: LITHIUM, PRECS, SURGICAL PROCEDURES, POLYURIA.

The query was 'why do you need to take plenty of water with lithium', the fact that this record mentioned polyuria has some relevance, albeit minor.

Record 1612-a1-2-q* dealt with the toxic concentrations of hydroquinone topical preparations; concentrations over 3% were not recommended. This was an adverse effects record and it could be argued that it contained

some 'use' information for indexing. USE was not indexed, so the record was not retrieved for the request for information on' hydroquinone cream, its use and formuulations'.
Query 42 asked for details of honey dressings and record 78901-n2-y was not retrieved in response to either the A or the B search statement because honey was not treated as a significant ingredient of a preparation and so was not indexed.
Record 837-a4-c* was a record on the uses of potassium perchlorate; sodium perchlorate had been referred to it. Sodium perchlorate had been added as a descriptor, but only to the first sentence of descriptors - it should have been used more extensively.
Record 6630-a3-p dealt with preservatives in food and listed a number of preservatives including ...*sodium nitrate (or potassium nitrate)*...Sodium nirate was indexed, potassium nitrate was not.

6.1.4 RECALL FAILURE DUE TO ERRORS IN THE INDEXING SENTENCE

Missed Record A Search	Missed Record B Search	Query
9034-a4-q*	9034-a4-q*	4
2400-a1-1-f*	2400-a1-1-f*	37
2370-a6-8-1*	2370-a6-8-1*	37
2373-a1-1-1*	2372-a1-1-1*	37
2375-a1-1-j*	2375-a1-1-j*	37
2378-a1-1-k*	2378-a1-1-k*	37
141-a1-f	141-a1-f	40
2058-a5-d*	2058-a5-d*	63

Table 44 Records not Retrieved due to Errors in the Indexing Sentence

These eight records were indexed with the appropriate descriptors but they were split incorrectly into sentences in such a way that descriptors that should have been linked were not and search statements constructed using the WITH qualifier could not retrieve the relevant records. There is however, some artificiality in ascribing some of

these record failures to this indexing fault in that the A search statement was too broad for queries 37, 40 and 63 and that the B statement was too broad for query 63; therefore if the indexing had been correct the records listed against those queries would still not have been retrieved.

6.1.5 RECALL FAILURE DUE TO THESAURUS LIMITATIONS

Record 2317-n2-y* was not retrieved by either the A or the B search in query 87 which requested information on parenteral formulations of thiazide diuretics. The record described Chlorothiazide for Injection (U.S.P.) and it was indexed with the descriptor INJECTION. Unfortunately INJECTION, which has here been used to describe a dose form, is classified as a method of administration and not as a parenteral dose form; thus the inclusive search on PARENTERAL DOSAGE FORM& as requested in both statements could not retrieve this record. Admittedly the indexers should have been aware of this anomally and added the descriptor PARENTERAL DOSAGE FORM. This section of the thesaurus needs revising; until it is, the indexers have been instructed to index with PARENTERAL DOSAGE FORM whenever they use INJECTION to describe a dose form.

Incidentally 2 record failures from 2 queries were initially considered for inclusion in this section of failure due to thesaurus limitation. One involved the treatment of tuberous sclerosis and the other the iodine content of kelp; neither tuberous sclerosis nor kelp are descriptors, but the words appeared as such in the records marked as relevant and a more general search would have retrieved them both. These two records are included in the section on failure due to limitation in the search statement.

6.1.6 RECALL FAILURE DUE TO SYSTEM FAULTS

Missed Record A Search	Missed Record B Search	Query
2601-a4-35-w*	2601-a4-35-w*	19
2686-a4-16-t*	2686-a4-16-t*	19
2686-a5-4-r*	2686-a5-4-r*	19
6243-a4-z*	6243-a4-z*	22
2317-a3-y	2317-a3-y	23
592-a1-1-q*	592-a1-1-q*	42
7910-c*	7910-c*	43

Table 45 Records not Retrieved due to System Faults

The first four records contained descriptors that matched the search statements, yet they were not retrieved, even on repeated searching. In these two queries the searching had to be carried out on Data-Star's system. The fifth and sixth records contained no descriptors; this could have been a failure on the part of the indexer (and is included in that section of the failure analysis) or the descriptor field could have been 'lost' during processing. The final record contained descriptors in the internal database but not on the database as held by Data-Star. No reasons have been found for these system faults.

6.2 PRECISION FAILURE

Of the 98 searches carried out for each A and B statement, 14 in the A group and 18 in the B group suffered some precision failure; 88 irrelevant records were retrieved in the A group and 76 in the B group. Some irrelevant records were retrieved as a result of more than one factor.

	A	B
Number of searches	98	98
Number with precision failure	14	18
Irrelevant records retrieved	85	72
Number of factors	93	80

Table 46 Summary of Precision Failures

The average precision ratios for A and B searches are given below together with the corresponding average from Lancaster's analysis.

	A	B	Lancaster
All records	63.5	67.5	50.4
Major records	58.3	59.6	25.7

Table 47 Average Precision Ratio

	Number of irrelevant records		% irrelevant records		Number of searches involved		% searches involved	
	A	B	A	B	A	B	A	B
SEARCHING								
Statement too limited	62	50	72.94	69.44	10	12	10.20	12.24
INDEXING								
Indexing rules	12	3	14.12	4.17	1	2	1.02	2.04
Indexing deficiency	6	6	7.06	8.33	2	2	2.04	2.04
Indexing sentence	1	1	1.18	1.39	1	1	1.02	1.02
Thesaurus limitation	12	19	14.12	26.39	3	4	3.06	4.08
SYSTEM								
Duplicate reference	-	1	-	1.39	-	1	-	1.02

Table 48 Reasons for Precision Failure

6.2.1 PRECISION FAILURE DUE TO SEARCH STATEMENT BEING TOO LIMITED

The largest group of irrelevant records was assessed as having been retrieved because the search statements could have been formulated more accurately; in most cases the statement was too general.
In some queries such as query 1 which asked for alternative CNS stimulants to Ritalin, the statement was made deliberately wide to retrieve a range of compounds with CNS activity. Many of these compounds were judged subjectively as unsuitable alternatives, though they possessed CNS stimulant activity. Interestingly, this query also caused problems in recall where, depending on the viewpoint, the search

statements could be considered to be too specific or too general; it is the kind of question least easily handled by Martindale Online.
In another example, query 15, narrowing the search statement to exclude retrieval of the irrelevant record would have meant not retrieving the relevant records; the irrelevant record was a necessary price.
The subjective nature of the relevance assessment was obvious in several questions. Query 37 for instance, requested toxicity information on dyestuffs used in foods. The record judged to be irrelevant described dyes in foods being associated with hyperkinetic states and could therefore have been judged as relevant.

6.2.1.1 DETAILS FOR EACH QUERY

Query 1 What alternative CNS stimulants are there to Ritalin?

A	B	
12342-a1-k	12342-a1-k	Dealt with almitrine, a respiratory stimulant; this type of stimulant was specified in both statements, but was not considered relevant.
1433-a5-f		The uses record of methylphenidate, which is Ritalin. It was not excluded to to prevent problems with the NOT operator.
1431-a5-x	1431-a5-x	Methylamphetamine's uses record. It could be considered relevant; was probably rejected due to restricted use.
1432-a1-a	1432-a1-a	Methylenedioxyamphetamine, a drug of abuse, but still a CNS stimulant; probably rejected for the same reason as 1431-a5-x.
1425-a1-t	1425-a1-t	Levamphetamine, a CNS stimulant probably rejected for the same reason as 1431-a5-x.
632-a3-p	-	Coffee, a stimulant of limited clinical use.
628-a5-h	-	Aminophylline; this drug has a wide range of activities, including respiratory and slight CNS stimulant activity.
546-a4-v	-	Strychnine; it has CNS activity but offers no clinical benefit.

(cont'd)

2540-a5-w	-	Viloxazine, an antidepressant with some CNS stimulant action.
2534-a3-f	-	Protriptyline, an antidepressant similar to viloxazine.
2075-a5-n	-	Isoprenaline, a sympathomimetic with some CNS stimulant action.
15303-a3-n	-	Nicotine; has CNS activity, but no clinical uses.
13321-a1-x	-	Theodrenaline; has CNS activity, but is used in hypotension.
13116-a1-2	13116-a1-z	Pimeclone, a respiratory stimulant as requested.
12933-a1-k	12933-a1-k	Metamivan, a respiratory stimulant as requested.
12931-a1-z	12931-a1-z	Mepixantone, a respiratory stimulant as requested.
12784-a1-x	-	Ginseng; possesses CNS and other actions, but has doubtful clinical value.
12738-a1-j	12738-a1-j	Fenethylline, a CNS stimulant; presumably rejected because of minor use.
12664-a1-l	12664-a1-l	Dimorpholamine, a respiratory stimulant as requested.
12537-a1-v	-	Catha, a CNS stimulant of doubtful value.
12364-a1-d	-	4-Aminopyridine, a CNS stimulant but used to reverse muscle relaxants and in myasthenia.
7620-a6-f	-	Cocaine, a CNS stimulant, but use was not considered to be relevant.
7618-a1-q	-	Coca, a CNS stimulant but no longer used clinically.
7045-a4-k	-	Flupenthixol, a CNS stimulant but used as an antipsychotic.
6617-a3-e	-	Pheneturide, a CNS stimulant but used as an anticonvulsant.
6243-a6-v	-	Morphine, a CNS stimulant and depressant used as an analgesic.
6153-a2-d	-	Phenindamine, a CNS stimulant but used as an antihistamine.

Query 3 Are there any new drugs for phenylketonuria?
Both the A and the B search statements retrieved 10 irrelevant records from the section on Proprietary Amino-acid Preparations (668-1). These all dealt with dietary preparations.
Although there is no descriptor to cover the concept 'new', these precision failures have not been considered under the section on thesaurus limitations. It would not be realistic to expect such a descriptor in a database that is not updated frequently; searchers wishing to obtain new information can do so using the update commands.

Query 4 Can anything be given orally to reduce facial hair?
Three records retrieved as a result of both the A and the B search dealt with griseofulvin being used to treat fungal hair infections; the statements were too general. Both statements also retrieved a record that described the dose of clomiphene in the treatment of polycystic ovary disease; no mention was made of facial hair, although hirsutism is one of the symptoms of this condition and that was why the condition was included in the statements.

Query 10 What gelatin plasma expanders are available?
Both search statements retrieved 2 irrelevant records that contained data about preparations that could be given parenterally and contained gelatin; unfortunately the preparations were not for plasma expansion. If the statements had included terms to cover volume expansion, these records would have been excluded.

Query 15 Use of minoxidil in baldness
Record 884-a1-10-d was retrieved by both search statements and dealt not with baldness but with a change of hair colour. The statements were too general. All the relevant records contained the descriptors HAIR and AE; refining the search by excluding one of those descriptors would have meant not retrieving the relevant records.

Query 18 Nonpharmaceutical uses of gum arabic
Both the A and the B search retrieved the same 8 irrelevant records as a result of the statements being too general. Seven of these records contained some pharmacy information, but according to the indexing guidelines, 5 of them were not indexed with PHARM (see also below).

Refining the search statement by including 'not PHARM?' would have meant that 6 and not 8 irrelevant records would have been retrieved. The eighth record dealt with gum arabic or acacia in food; although this was not a pharmaceutical use it was considered irrelevant.

Query 28 Does a nicotine spray exist?

15303-a3-n was retrieved by both the A and the B search and provided some information on nicotine being used in horticulture as a vapour or spray. While this could be considered by some to be relevant, it does not provide information on availability and a more restrictive search on preparations of nicotine would not have retrieved it - or anything else.

Query 37 Toxicity of dyestuffs used in food

The B search retrieved 2370-a6-10-e which was about dyes in foods being associated with hyperkinetic states and could be considered as highly relevant, instead of irrelevant as judged; one can only treat this as a possible failure of the search statement.

Query 41 Hydroquinone cream - use and formulation

The A statement did not limit the search to a cream or similar preparation and as a result retrieved 901-2-a which described the compound's use as an antoxidant in tablets.

Query 45 Carcinogenicity of Sudan IV

The A statement asked for any colouring agent instead of limiting the search to SCARLET RED which is the descriptor for Sudan IV. As a result, the A search but not the B search retrieved 6 irrelevant records.

Query 49 Reports of progesterone causing cancer

The A statement did not specify AE and retrieved 2 irrelevant records that dealt with the precautions or use of progesterone in breast cancer.

Query 55 Substance that turns pink when wet

The B search statement was: SUPPLEMENTARY-DIAGNOSTIC-AGENTS£ WITH WATER, and this was only arrived at through knowledge of the file. It produced 4 hits, one of which was relevant; any refining of this odd search statement would not have retrieved the only relevant record.

Query 61 Use of anisaldehyde

The B search statement included the term ANISE which resulted in 4 irrelevant records. This is also discussed under the section on the limitations of the thesaurus which does not contain a descriptor for anisaldehyde.

Query 70 Calcium and vitamin D supplement for a four-year-old unable to take milk (what doses and source)?

The A statement produced no hits, the B statement produced 8, of which 2 were relevant. The 6 irrelevant records were the result of intentionally asking in the search statement for preparations of calcium and vitamin D without specifying children in case some of the preparations might have been suitable; the statement also asked for information on the administration of the compounds to children.

Query 90 Alupent syrup, is it sugar-free?

The B search statement did not specify the dose form and as a result retrieved a record that dealt with Alupent injection.

6.2.2 PRECISION FAILURE DUE TO INDEXING RULES

Reason	Irrelevant record A search	Irrelevant Record B search	Query
Indexing an absent effect	1478-a5-l	1478-a5-l	1
	5736-a5-d	-	1
	2112-a2-v	-	1
	2096-a1-m	-	1
	2094-a5-b	-	1
	2080-a4-m	-	1
	2067-a3-e	2067-a3-e	1
Indexing a secondary effect	2099-a3-t	-	1
	2095-a3-z	-	1
	2087-a5-v	-	1
	2078-a4-s	-	1
	2050-a1-l	-	1
Use of PREG qualifier	-	2601-a5-29-y	22

Table 49 Irrelevant Records Retrieved due to Indexing Rules

In query 1, seven irrelevant records were retrieved because an absent effect was indexed as instructed in the guidelines. If a substance was said not to possess CNS stimulant activity, then it was considered that that activity should be indexed. In this instance, these records which were so indexed were considered to be irrelevant. In other cases it might be valuable to know that the compounds had no activity. It is unlikely that this rule will change.

In the same query, 5 irrelevant records were retrieved because secondary CNS activity was indexed; to counterbalance this, some recall failures were due to incidental or minor information not being indexed.

The irrelevant record retrieved by the B search for query 22 was a result of the PREG qualifier being used in a broad sense to cover the postpartum state as well as pregnancy and the neonate.

6.2.3 PRECISION FAILURE DUE TO INDEXING OMISSION OR ERROR

Irrelevant record A search	Irrelevant Record B search	Query
529-a2-j	529-a2-j	18
5400-n	5400-n	18
5402-n1-c	5402-n1-c	18
5463-a2-g	5463-a2-g	18
5465-a2-2-c	5465-a2-2-c	18
7023-a1-m	7023-a1-m	35

Table 49 Irrelevant Records Retrieved due to Indexing Omissions or Error

The last record shown in Table 49 dealt with the adverse effects of chlorpromazine and was retrieved in response to a search for information on the normal dose of clopenthixol depot injection. Clopenthixol was cross-referenced to chlorpromazine and the indexer had indexed chlorpromazine's problems on injection as being applicable to clopenthixol. Indexing cross-referenced data is likely to lead to problems in identifying how much of that data is applicable to the secondary compounds.

The other irrelevant records have been included in this section out of completeness for they could only be considered to be failures due to indexing omissions at a remove! Query 18 was a request for information on the nonpharmaceutical uses of gum arabic. The 5 records listed above against the A and the B search were retrieved with the search statement 'ACACIA WITH USE' and this statement has already been assessed as being too general (see above). However, none of the records contained PHARM in the same index sentence as USE, so if a more precise statement had been formulated employing USE NOT PHARM?', these irrelevant records

would still have been retrieved - but then the record judged as being relevant would not have been retrieved.

6.2.4 PRECISION FAILURE DUE TO ERRORS IN THE INDEXING SENTENCE

Both the A and the B search retrieved 33-a5-r in query 40 which was a request for information on local anaesthetic injection reactions. This record stated that the pain of cefoxitin might be reduced if it were given with lignocaine. The indexing should have been such that these 2 topics were split into two sentences along the following lines:

CEFOXITIN, AE, PAIN, INJECTION, PARENTERAL ROUTE,....

LIGNOCAINE, USE, PAIN, INJECTION, ...

Unfortunately the indexing was not so split and 33-a5-r was retrieved as a record describing the adverse effects of lignocaine.

6.2.5 PRECISION FAILURE DUE TO THESAURUS LIMITATIONS

Irrelevant record A search	Irrelevant Record B search	Query
1446-a2-t	1446-a2-t	1
1432-a1-a	1432-a1-a	1
4609-w	4609-w	61
2042-a2-2-z	2402-a2-2-z	71
2042-a2-1-e	2402-a2-1-e	71
9237-n17-e	9237-n17-e	71
9237-a5-47-b	9237-a5-47-b	71
9237-a5-44-z	9237-a5-44-z	71
9237-a5-43-e	9237-a5-43-e	71
9237-a5-37-t	9237-a5-37-t	71
9237-a4-5-z	9237-a4-5-z	71
9237-a4-3-a	9237-a4-3-a	71
	1340-t	84
	6301-a1-36-m	84
	1350-a4-11-y	84
	1346-a2-4-p	84
	1341-a4-d	84
	3715-a1-1-m	84
	9004-a4-n	84

Table 51 Irrelevant Records Retrieved due to Thesaurus Limitations

Query 1 was that problem query on CNS stimulants as alternatives to Ritalin. Record 1446-a2-t described propylhexedrine which is classified as a CNS stimulant in the thesaurus although its main use is as a decongestant - the classification in this instance was not suitable. Record 1432-a1-a described methylenedioxyamphetamine, a drug of abuse which is classified as a CNS stimulant although it is not used clinically; this could be considered as a marginal limitation of the thesaurus.

The retrieval of 4609-w could also be considered to be due to a marginal limitation of the thesaurus. The query was on the use of anisaldehyde. There is no descriptor for this substance, so the search statements had to involve free-text searching on anisaldehyde linked to the descriptor USE by the operator 'and'. If there had been a descriptor for anisaldehyde, then it could have been linked to USE by the 'with' operator, in which case 4609-w would not have been retrieved. However, it is unlikely that anisaldehyde will become a descriptor.

Query 71 was for information on the sublingual dose of glyceryl trinitrate. The descriptor that covers the sublingual route is BUCCAL ROUTE which in this instance was too general.

Query 84 was a request for information on normal blood cholesterol levels. There are no suitable descriptors for such a request as physiological descriptors are very general and other descriptors that are connected with cholesterol take a disease viewpoint. The B search was carried out using free-text searching and so retrieved the 7 irrelevant records.

6.2.6 PRECISION FAILURE DUE TO DUPLICATE INFORMATION

Record 4061-a5-18-v was retrieved by the B search for query 68 which was a request for information on carbamazepine in pregnancy. This record was a shortened version of 6602-a4-55-y which appeared within the phenytoin set of records and which dealt with phenytoin, phenobarbitone and carbamazepine. Because of slight differences in the content of the abstracts it was decided to load both versions. Normally duplicates are excluded or not indexed.

CHAPTER 7 DISCUSSION

7.1 SEARCH SYSTEM

The first group of comparisons (S3 versus S1) was set up to see whether it had been worthwhile indexing the database. Did searching Martindale Online in the way the author envisaged it should be done produce better results than searching it as a free-text file that had no descriptors assigned to it? When all relevant hits were being considered, the difference in favour of searching making use of descriptors was highly significant ($P<0.001$) for search statements formulated by third parties. As would be expected, the author who is familiar with both the system and the content of the database achieved scores that showed a slightly less marked difference, although still significant.

Inspection of the search statements for the S3 set shows that descriptor searching was used solely in 83 of the A searches (searches using statements formulated by the information pharmacists) and 80 of the B searches (in these, the statements were formulated by the author). In the A set, 9 searches employed a mixture of descriptor and free-text searching, 5 searches were purely free-text searches, and 1 search was a mixture of descriptor and record format searching. In the B set the respective figures were 9, 8, and 1. As indicated in the User's Guide, there is a place for both controlled and free-text searching in Martindale Online, especially when dealing with drug nomenclature. Free-text searching was used effectively in this study, but not to a very great extent, which can probably be explained by drug nomenclature queries being excluded from the test sample.

It could be argued for a full-text file such as Martindale Online that a comparison of hits of major relevance would provide the most 'relevant' results and certainly this is reinforced by Lancaster's comments on recall analysis[56]. Hits considered to be only of major relevance gave a slightly different pattern. The difference in scores was still significant for the A set, but to a lesser extent. However, there was no significant difference between the two types of searching for the B set. While this supports the author's view that someone who knows the database well is able to obtain satisfactory results from free-text

searching, it is impractical to consider that everyday users would have or wish to have similar knowledge.

The results for the individual information pharmacists are interesting in their variability. The queries selected by the most experienced of the three provided results that had a pattern similar to that for the total. Results from searches on the queries selected by the information pharmacist with intermediate knowledge showed that the difference between S3 and S1 was significant for all relevant hits; with hits of major relevance, there was no significant difference with the A set. Finally, there were no significant differences in any of the sets involving the queries (group III queries) selected by the information pharmacist with least experience of the database. It would appear that S1 searches (and S2 for that matter) produced better results with this third information pharmacist than with the others. It would also appear that the lack of difference between S3 and S1 was not due to poor search statements for S3, for those prepared by the author produced similar results. It is also worth commenting that no striking features emerged from the failure analysis of the group III queries. The selection of the queries might have played some part in the result, although all three information pharmacists provided similar numbers of searches for which there were no relevant records in the database; 11 in group I, 12 in group II, and 12 in group III.

While the first set of comparisons does show that indexing the file improved the effectiveness of searching, it does not reflect what is being done by searchers. What it does compare is what should be done with what might have been done. A more relevant question, considering that the database is already indexed, is: how does what is being done compare with what should be done?

Sales of the Martindale Thesaurus have been slow (information supplied by the Pharmaceutical Press); slow enough to indicate that many users for some time were searching Marindale Online without its aid. The S2 searches, which used the same unstructured search statements as S1, were carried out on the indexed database in an attempt to represent what a user without a thesaurus might do. As discussed elsewhere, attempts to obtain information from outside searchers proved difficult, which is unfortunate since it would have been preferable to have studied some of their searches. However, the three information pharmacists did at least represent three different types of searchers. Comparisons for all

relevant hits showed similar differences in scores in favour of S3 as in the comparison of S3 and S1. The AT set showed a highly significant difference; the BT and AM sets also showed S3 performing better than S2, though the significance was less than with the AT set; The scores in the BM set showed no significant difference.

Once again it is difficult to arrive at a conclusion from consideration of the comparisons within each of the three group. The most striking feature was the recurring lack of difference between results in group III.

Having demonstrated that searching as it should be done produces better scores than searching as it is probably done is really a pyrrhic victory if the users are not searching in the most effective way. The key comparison therefore is: has the additional work on indexing the database helped the searcher in spite of the searcher not making the best use of this facility (S2 versus S1)? All comparisons provided insignificant differences between the scores. Sadly, searchers who have not consciously been using the descriptors have therefore not been helped by the indexing.

The author was aware in designing the online database that end-users or inexperienced users such as pharmacists, prescribers and perhaps patients, would be the least likely searchers to use search aids such as thesauruses, and this has been supported by others[75, 76]. Yet it could be argued that these are the searchers most in need of such aids, at least with this database - the end user paradox. What was not expected was the finding that searchers already using online facilities when Martindale Online was made available, experienced online searchers and intermediaries in the main, would not make use of cost saving aids. The end user paradox might be partly resolved by building the thesaurus into the database, and this might even be of benefit to the experienced searcher. If the thesaurus contained enough lead-in terms and if rules could be formulated for explosion and the selection of related terms, then complex search statements might be constructed automatically from a searcher's relatively unstructured search statement or question. Such a complex statement would include therefore a spread of descriptors and nondescriptors that might provide useful retrieval from the unindexed database, provided that arrangements could be made for the retrieval of data from cross-referenced material. Cutting back on the indexing would

be a useful cost saving. This in summary was the rational behind the final two sets of comparisons.

The information pharmacists and the author constructed highly structured search statements with the aid of the thesaurus and compared these with the unstructured search statements in searches of the unindexed file (S4 versus S1). The difference in scores for all the sets searched was insignificant - not a result that was expected. Nor were the results of the last set of comparisons, for while the AT and the BT sets showed that S3 searches carried out in the recommended manner on the indexed database produced higher scores than the S4 searches of the unindexed database, the AM and BM sets showed no such difference. These confusing results indicate that further study is required before either dismissing the S4 type of searching or considering it as an alternative to the recommended methods.

7.2 FAILURE ANALYSIS

Having assessed the different methods of searching Martindale Online and determined that the optimum method involved a mix of descriptor and free-text searching, it was decided to carry out a failure analysis on the searches made using this optimum method with a view to quality control of the database.

Lancaster's study[56] served as a guide for this analysis. However, there were some differences between the two studies. First of all this study is much smaller than Lancaster's; 98 searches were analysed compared with 302. This smaller size did mean though that it was practicable to carry out the precision analysis on all the records and not on a representative sample. Also the recall figures, which did not need to be derived from a separate set, were arrived at following the assessor's scanning of all records retrieved in response to the search statement plus those retrieved by the author's additional manual amd computer searches for all queries. Secondly, this study did not involve any 'user system interaction' since the queries being used had been extracted by the information pharmacists' from their logs of past queries; poor 'user system interaction' accounted for a considerable proportion of Lancaster's failures. Thirdly, the two databases and search systems possessed a number of differences, for instance in types of record, the approach to indexing, and methods of searching.

For all these differences, some general comparisons can be made between the two studies.

The average (macroaverage) recall ratio for all records was 60.2 for the searches formulated by the information pharmacists (A) compared with 65.4 for the searches formulated by the author (B) and 57.7 from Lancaster's study. A similar relationship is evident when records of major relevance are considered; the average recall ratios are 69.3 for A searches, 73.2 for B searches, and 65.2 for Lancaster.

The differences are more evident when the precision ratios are considered. Precision failure analysis was carried out by assessing the retrieved records that were of no relevance and the average ratios resulting from this analysis are 63.5 for A and 67.5 for B; Lancaster's figure was 50.4. However, in discussing overall performance Lancaster also calculated the average precision ratio of 25.7 based solely on records of major relevance; corresponding figures for A and B are 58.3 and 59.6. The difference in precision ratios is reflected in this study producing a smaller number of unwanted records and a smaller number of affected searches when precision failures due to searching, indexing and the thesaureus were compared (Tables 52 and 53) - which is only to be expected in a small database with a specially constructed thesaurus.

Reason for Irrelevance	A	B	Lancaster
Searching	62	51	983
Indexing	31	29	393
Thesaurus Limitations	12	19	1094

Table 52 Number of Irrelevant Records Involved in the Different Sets of Searches

Reason for Irrelevance	A	B	Lancaster
Searching	10	12	67
Indexing	4	5	60
Thesaurus			
Limitations	3	4	91

Table 53 Percentage of Searches Involved in Precision Failure

7.2.1 SEARCHING

The largest cause of both recall and precision failure in this study lay in the search statements' limitations. This held true for both the A and the B searches, thus detailed knowledge of the file and the search system did not provide the benefits that might be expected in this area. Lancaster examined 302 searches and found that 35% of total recall failures were due to problems in searching; equivalent figures in this study are 74% for A and 64% for B.

Thirty six different queries were involved in some recall failure due to limitations of the search statements, 35 in the A set and 29 in the B set. In 16 of these queries the search statements reflected the content of the query, while the relevance assessment marked records that satisfied part of the query or an oblique aspect of the subject as relevant. Query 8, for example, was a request for information on any interaction between prednisolone and an oral contraceptive and the search statement concentrated on that specific corticosteroid, on oral contraceptives, and on INT. Not surprisingly, such a specific search did not retrieve a record on an interaction between hexoestrol and an unspecified corticosteroid. In another example, query 30 called for information on the collection of mercury spillage; the recall failure dealt with the problem of mercury spillage but not on the problem of its collection. Such records can be seen to be of some relevance and might well have been retrieved if there had been some interaction with the user. In the context of this study there was no such interaction, nor was there a second attempt at an alternative search strategy if the first one failed to produce a satisfactory response. This perhaps

explains why the limitations of the search statements accounted for such a large proportion of the failures.

In 2 queries, the statement could be considered to be based on too narrow an interpretation of the query. The statement for query 23 specified PRECS and WATER as part of the statement for a request for information on the reason for taking large amounts of water with lithium; by so specifying, the search did not retrieve relevant records that dealt with this topic from the adverse effects angle or dealt with the use of infusion fluids in conjunction with lithium. Query 24 was a request for information on ampicillin in pregnancy and was interpreted as being a request for information on the precautions to be taken in using this antibiotic in pregnant patients; such a limited interpretation excluded relevant records that dealt with uses or absorption and fate.

Limitations due to faulty logic occurred in 3 queries, 55, 91, and 97. In 91 there was also the inclusion of an unnecessary qualifier. In 55 and 97 the statements called for a coordination of a descriptor and a nondescriptor in the same sentence, which is an impossibility in this database.

Queries 1 and 55 proved to be difficult for this database. Both these questions could only be satisfied by a contrived search, and in the case of 55, one of the statements introduced a logic error as discussed above. In the case of query 1, a request for alternative CNS stimulants to Ritalin, the statement had to be very general to retrieve a representative group of alternatives.

In 12 queries the statements were too limited in that they failed to cover the content of the question as provided or failed to make use of the available system facilities. Query 40 was a request for information on reactions to local anaesthetic injections; missed relevant records could have been retrieved if the statements had included AEM (as well as AE), INTRAVENOUS ROUTE, and EPIDURAL BLOCK. In formulating search statements for Martindale Online, it is difficult to know how far to go in following up all related terms. Query 49 called for information on progesterone causing cancer and the missed relevant records dealt with neoplasms (not malignant neoplasms), with hyperplasia, and with chlormadinone; it is unlikely that many searchers would include hyperplasia or substances like chlormadinone in a search statement.

Precision failure due to limitations of the search statement affected 14 searches, 10 in the A set and 12 in the B set. In 3 of the queries any refinement of the statements would have meant that the relevant records would not have been retrieved; thus in queries 15, 28, and 70 the irrelevant records were a necessary evil. In most cases the statements turned out to be obviously too specific or too general; for example, in query 45 the A statement called for any colouring agent instead of limiting the search to the descriptor for sudan IV. Query 4 contained two examples of statement limitations. The request was for information on oral drugs for facial hair. The statement contained the descriptor POLYCYSTIC OVARY DISEASE since this is associated with hirsuitism; as requested the search retrieved records on this condition but which made no mention of hirsuitism or its treatment. The search also retrieved data on the oral treatment of fungal hair infections and this could only have been excluded by incorporating 'NOT FUNGAL INFECTIONS' into the search statement. It is likely that a searcher might think that it is not worth including POLYCYSTIC OVARY DISEASE, but it is not very likely that a searcher would think to exclude fungal infections.

The balance between recall and precision is nicely demonstrated by the response to query 41, a request for information on hydroquinone cream, its use and formulation. Specifying the dose form in one statement meant that 2 relevant records were not retrieved; not specifying the dose form in the other statement meant that an irrelevant record on hydroquinone tablets was retrieved. A similar conflict occurred in queries 60 and 61. The first of these called for information on the iodine content of kelp while the second for information on the use of anisaldehyde. Neither kelp nor anisaldehyde are descriptors. A search on kelp failed to retrieve a relevant record on seaweed, but a search on ALGAE& retrieved it. There was no relevant information in the database on the use of anisaldehyde, but a search on the descriptor ANISE, under which the aldehyde might have been indexed, produced some precision failures.

7.2.2 INDEXING

The next largest cause of both recall and precision failure involved the indexing and particularly the indexing rules. In Lancaster's analysis about 37% of recall failures were due to indexing; in this study 37% in the A set and 40% in the B set were due to some indexing problem.

The indexing rule that caused the greatest number of recall failures was the instruction not to use the AD qualifier routinely when indexing adverse effects data. It had been considered that the AD qualifier should be used for information about the administration of a drug in the treatment of a disorder and only be used in adverse effects records where there was considerable dosage information. This was soon recognised to be a mistake and the indexing guidelines have now been changed so that any dosage information in any record is indexed with the qualifier AD. It is too early to know what effect this change will have on precision; certainly the former rule did not lead to any precision failures.

A related cause of recall failure was the instruction not to index precautions information with the adverse effects qualifier unless there was some relevant statement about adverse effects. Seven recall failures occurred in 4 searches because of this rule. It is unlikely that this rule will change, for adding AE to every precautions indexing sentence would greatly increase the chances of false coordination when searching on adverse effects; conditions that are the subject of precautions in using a drug are not necessarily conditions that can be caused by the drug. Before leaving the problems of qualifiers, it is interesting that the SOU and the PREG qualifiers also caused recall problems by being used too rigidly. Their use has since been widened, even though 1 precision failure occurred because PREG had already been used widely to cover the postpartum state as well as the pregnant state.

Not indexing minor ingredients of preparations was the second largest cause of recall failure in this category. It was envisaged that these minor ingredients should be retrieved using free-text searching. Controlled searching did not work in the queries involved, but in 3 of the 4 searches free-text searching would have been successful. There is no intention to change this rule. Searchers should however be reminded of the occasions when free-text searching should be employed. Indeed in the DIALOG system, which has not been considered in this study, searchers of Martindale Online have had to be instructed in great detail about when and when not to use controlled or free-text searching because of the way the data has had to be organised for that system.

While it is a policy in this database to index in some detail, certain measures had to be taken to restrict the size of the descriptor field.

One of these measures involved not routinely indexing vehicles or solvents outside their monographs. This meant that there were not thousands of entries for SODIUM CHLORIDE or DEXTROSE or WATER and it was this lack of indexing of water that led to a number of recall failures. The B searches produced 1 recall failure because of this, while the A searches produced 9. More information about the best ways to search the database might improve performance in this type of search.

Another attempt to keep the size down as well as to improve precision, was the instruction not to index duplicate information. The relevance assessment marked as recall failures 4 records that contained information duplicated elsewhere. In contradiction, one record that was a shortened version of another and was included because it was not an exact duplicate was marked as a precision failure. The assessment marking the duplicates as recall failures was rather surprising and, although the assessor did indicate that he might not have been consistent, it merits further investigation to discover whether other searchers find this service of no value and are still expecting full-text files to be treated similarly to bibliographic databases. Similarily, general incidental information was not indexed and this lead to 4 recall failures, although in all 4 cases other records provided relevant and more detailed information. It is noteworthy that there were 5 records that did have incidental information indexed and were considered precision failures in query 1. There were also special types of incidental information that were not indexed, such as some of the incompatibility data and the indications of preparations. Each type resulted in only one recall failure. While this might appear to be an acceptable result, problems can be foreseen with further use of the database. Detailed indexing of incompatibilities is a large task and is not being undertaken; however, it has already been decided to index the indications of preparations.

Indexing negative information led to 7 precision failures in query 1. The irrelevant records described some drugs as having no CNS activity; as the query was for information on compounds with this activity it is understandable that these records should be considered irrelevant. However, another question might well be answered by this indexing and there are no plans to change this rule.

The descriptor field was structured to take sentences of descriptors with the aim of providing precise searching by reducing the chances of false coordinations of terms; indeed, there was only 1 false coordination in this study (record 33-a5-r in query 40). Describing the different topics of a record in separate sentences of descriptors meant that searching across the topics could not be carried out if the searcher wished to prevent any false coordinations. Nausea occurring with a drug being used to treat diarrhoea would therefore be indexed with 2 sentences: one containing the drug descriptor and NAUSEA and AE; the other containing the drug and DIARRHOEA and USE. Following the search guidelines and employing the 'with' operator would mean that a hit would not ensue from this record if the statement consisted of USE with NAUSEA with DIARRHOEA. In query 98, three recall failures occurred indirectly because of this approach; indirectly for the indexer had not carried out the instruction to the letter but instead had not indexed what could be considered the minor of the two concepts. Had the indexing been as instructed, the recall failures would still have occurred. Information can be retrieved from such records by carrying out two-stage searches. In the above example, the drugs used in the treatment of diarrhoea could be identified and this set of drugs could be searched to see which of them caused nausea. This was a point of concern in the design of the database and it has to be admitted that the author is surprised that it only affected the one query in this study.

One final aspect of the indexing rules not properly tested by this study was the application of the minor qualifiers. Only 1 recall failure occurred in the A search because the minor qualifier PHARMM had been used according to the rules; the record was retrieved by the B search. There are quite complex indexing rules for the handling of interactions and the use of minor qualifiers and it is a pity that these were not put to the test.

Errors in the indexing, whether due to omission, inaccurate assigning of descriptors, or poor sentence structure in the descriptor field accounted for a relatively small proportion of recall and precision failures. This is encouraging considering that the content of the indexing was not checked for the first release of Martindale Online because the indexing at that stage was carried out after the revision and editing. At

subsequent updates, records have been indexed at the same time as revision so that the indexing has been subjected to checks for coverage at the time of normal text checking and editing; this should reduce further the incidence of this sort of error.

Two recall failures and one precision failure occurred because of an indexing error in cross-referenced material. However, this feature of the database was not sufficiently tested for it is often difficult for the indexer to decide which descriptor in the major record should apply to the drug being cross-referenced to it.

7.2.3 THESAURUS LIMITATIONS

Thesaurus limitations caused more of a problem for precision than for recall. There was only one recall failure and that was due to the descriptor INJECTION being used by the indexer to describe a dose form although the descriptor was classified as a method of administration. Indexers have already been informed that they may use INJECTION in that sense, but that if doing so they should also index with PARENTERAL DOSAGE FORM; this is not a very tidy solution, but it has prevented a restructuring of the dosage and administration section of the thesaurus.

Some of the precision failures were due to inappropriate drug classification as happened in query 1. Drugs can be classified in a variety of ways, so there are bound to be occasions when a question would be better answered if a drug had been classified differently. There may be an improvement in this type of failure as the drug classifications are being reconsidered and amended as chapters are revised. Other errors of this type are due to a conscious decision that some descriptors should be used to cover several detailed concepts, e.g. BUCCAL ROUTE, or that a descriptor is not required for minor compounds mentioned within monographs and that could be covered by the descriptor for that encompassing monograph substance.

7.2.4 SYSTEM

A few errors were attributed to system faults; it is difficult at this stage to see what measures can be taken to prevent such errors.

7.2.5 ZERO HITS

Of the 98 queries, 34 produced no relevant hits, even accounting for the additional searches; 11 of the 34 were in group I, 11 in group II and 12 in group III.

Such a response was either because the database had failed to cover the relevant information or because there was no information to cover, in which latter case a negative response was an acceptable one. This study was not aimed at studying the coverage of Martindale Online, so that both types of result were scored similarly if the system failed to produce any hits; zero hits in any search for which there were no known relevant records produced a score of 1. However, any data on coverage is useful in the planning of revision of the database.

In 17 of the 34 queries with 0 relevant hits the zero response was not considered to be a failure. Five of the 17 queries dealt with interactions (queries 11, 34, 69, 75 and 93). In such cases a zero response could be considered to be a useful response and indicated that the editorial staff had not been able to identify any such interactions.

Seven of the 17 queries dealt with other topics that would have been researched by the editorial staff. If there had been any relevant information at the time of revision, it would have been included. The seven queries were 36, 47, 51, 58, 74 and 94. In the last of these (94), the query was 'Is lactulose safe in pregnancy?' Information would have been included if lactulose had been shown not to be safe in pregnancy.

Four of the 17 dealt with topics not normally covered by Martindale;

- query 21 Any drugs from the tropical rain forrest?
- query 31 How to remove the smell of spilt methyl salicylate?
- query 56 Crystal size of adrenaline.
- query 62 Detection of benodiazepines on the breath?

The last of the 17 queries in this group (query 28 on a nicotine spray) did produce some potentially useful information, but the record was not assessed as being relevant.

The other 17 of the 34 queries could be considered as failures of the database since they were on topics that generally should have been researched and, in the author's opinion, should have yielded some information. Some of the topics might be outside the scope of Martindale, but since there is some doubt they have been included in this group. Such topics include:

- query 9 Dilution, storage and extending of semen.

- query 20 Strontium acetate in toothpaste.
- query 21 Normal blood cholesterol values.

Topics that should have been covered without doubt in Martindale are:

- query 12 Lithium succinate preparations.
- query 13 Hydrogen peroxide for contact lens cleaning.
- query 25 Feverfew in migraine.
- query 33 What is best for ringworm in the pregnant patient?
- query 38 Use of salicylsulphonic acid and its formulations.
- query 46 Candida extract for intradermal use.
- query 48 Use of chloroform water in infants.
- query 64 Half life of potassium perchlorate.
- query 80 Papain for mouth ulcers.
- query 83 Is Lomotil liquid sugar-free?
- query 86 Castor oil as an abortifacient
- query 89 Dose of Auralgen ear drops.
- query 90 Is Alupent syrup sugar-free?
- query 95 Dose of aspirin and papaveretum tablets.

7.3 INFLUENCE OF SPECIAL FEATURES OF MARTINDALE ONLINE ON RETRIEVAL

Special Martindale Online features include linked records, the indexing of cross-referenced data and the different types of records. How did these features affect the results of the searches for the 98 queries?

7.3.1 LINKED RECORDS

A growing feature of Martindale is the bringing together of abstracts and references that support or contradict each other; they may be brought into a review or into a set of linked records.

The various methods of searching, including the additional searches, produced 326 relevant records for the 98 queries. In 18 of these queries there were 44 sets of linked relevant records (13.5%), 18 were produced by 14 of the S3 searches. The 26 relevant linked sets produced by other types of searches included 5 picked up by the additional searches.

Out of the 44 relevant sets of abstract records,

27 had as the linked data further references, with, on 3 occasions, just a mention of criticism or an indication that there were contrary

views to those contained in other records of the set (queries 1, 29 and 50).

14 provided useful additional information as in query 29 where record 10-a6-24-a provides data on concentrations of ampicillin in maternal and cord sera and in amniotic fluid. Linked to it is 10-a6-25-f which provides information on the time required for transplacental diffusion.

3 provided contradictory information. In query 17 the linked set under 4061-11-r brings together conflicting data on the stability of phenobarbitone sodium solutions. In query 54, the first record reported no absorption of clindamycin from its topical administration, the second record provided evidence of absorption. In query 72, the first record deals with hexamine being effective as an insect repellent, the linked data refers to 2 studies that had not shown this effect.

It is still editorial policy to increase the linking of references and abstracts or to include a review or commentary. It would be interesting to analyse this feature after another update using an outside relevance assessor. No comment was made at the assessment on any sets of linked records.

7.3.2 CROSS-REFERENCED DATA

The indexing of cross-referenced text was examined in the context of the S3 searches and only 4 queries were involved.

In the first (query 57), the search on nitrazepam causing mental depression retrieved the diazepam record to which nitrazepam had been referred. The nitrazepam record states that this compound has similar adverse effects to diazepam. Therefore the relevant adverse effects in diazepam were indexed with nitrazepam and a paragraph of text was added to the diazepam record to explain the connection. In the second query (43), relevant information was not retrieved because the cross-referenced data had not been indexed and this is referred to under the section on indexing omissions. In the third query (35), there was a precision failure due to the characteristics of the cross-referred to drug being too generously assigned to the compound considered to have similar properties. In the fourth query (76) there was a recall failure due to an opposite effect where the cross-referred to drug was not indexed in sufficient detail for the other compound.

The retrieval of cross-referenced data played an important part in the decision to index the database, so it is pity that the sample questions

did not test this feature more fully, although it is disturbing that 3 of the 4 queries involved should display an error in this feature.

7.3.3 RECORD TYPES

Mackay[52] has indicated that the headings assigned to the different record types might form useful search features. These headings contain words such as *Uses Text* or *Adverse Effects Abstract*.

Within this study, records were retrieved in some of the searches for 13 of the 98 queries partly because of a hit in the heading paragraph (the RF paragraph on Data-Star). Of the 326 unique relevant records retrieved for the 98 queries, 46 (14.1%) were retrieved in this manner from all the searches; the A searches produced 29 and the B searches 40 of the relevant records. In addition to the 46 relevant records, there were 40 irrelevant records, 25 in the A sets and 33 in the B sets.

Of the 46 relevant records, 26 were also retrieved by S3 searches that did not invoke the heading paragraph (13 in the A set and 24 in the B set).

Two of the S3 search statements specified that part of the search should be carried out in the RF paragraph. In the first, query 1, 64 records were retrieved and 26 were relevant; in the second, query 14, 2 records were retrieved and both were relevant. The author invoked the RF paragraph in query 1 to keep the number of hits to a reasonable number by requesting that only the Uses text records be searched. The author also invoked the RF paragraph in query 14 to limit the search to preparation records and retrieved the same 2 relevant records as S3A which did not so limit the search.

While this study was not designed to investigate the RF paragraph, it does provide a further indication that it can be a useful search feature. In practice there are queries that can only be answered by using this feature, questions such as the availability of drugs in certain countries and such queries were intentionally not selected by the information pharmacists.

ENDPIECE

Searching Martindale Online through descriptors supplemented as appropriate by free-text searching is more effective than complete free-text searching of the unindexed database or uncontrolled searching of the indexed database.

Highly structured searching of the unindexed database produced conflicting results, but indicated that further study on a front-end system for free-text searching might be worthwhile.

Failure analysis showed that poor construction of the search statements within this study accounted for the largest number of recall and precision failures. This should be less of a problem in searches that involve some interaction.

Failure analysis also highlighted some problems in the indexing policy and some of these are being corrected.

The special features of Marindale Online can be of value in searching.

REFERENCES

1. Annual Lists of Journals Indexed in Index Medicus. Bethesda, Md, National library of Medicine.
2. AD HOC WORKING GROUP FOR CRITICAL APPRAISAL OF THE MEDICAL LITERATURE. *Annals of Internal Medicine*, Vol.106, 1987, pp.598-604.
3. *MEDLINE*. Bethesda, Md, National Library of Medicine.
4. *EMBASE*. Amsterdam, Excerpta Medica.
5. *International Pharmaceutical Abstracts*. Washington, D.C., American Society of Hospital Pharmacists.
6. REYNOLDS, J.E.F.(Ed.). *Martindale, The Extra Pharmacopoeia. 28th Edition*. London, The Pharmaceutical Press, 1982.
7. *AHFS Drug Information*., Washington, D.C., American Society of Hospital Pharmacists.
8. *British National Formulary*. London, The British Medical Association and The Pharmaceutical Society of Great Britain.
9. *USP DI Drug Information for the Health Care Provider*. Rockville, Md, United States Pharmacopeial Convention Inc.
10. HAYNES, R.B., McKIBBON, K.A., FITZGERALD, D., GUYATT, G.H., WALKER, C.J., SACKETT, D.L. How to Keep up with the Medical Literature. *Annals of Internal Medicine*, Vol.105, 1986, pp. 149-53 [this is the first of a series of papers on this topic].
11. ANON. Topics in drug therapy. *Prescribers Journal*, Vol. 24, 1984, p.89.
12. HERXHEIMER, A., LIONEL, N.D.W. Minimum Information Needed by Prescribers. *British Medical Journal*, Vol. 2, 1978, p.1129-32.
13. Report of the Working Party on Drug Information Services. *Pharmaceutical Journal*, Vol. 213, 1974, pp.297-300.
14. THE NUFFIELD FOUNDATION. *Pharmacy: The Report of a Committee of Inquiry Appointed by the Nuffield Foundation*, London, The Nuffield Foundation, 1986.
15. EAST, H. *Designing and Marketing Databases. British Library Research Paper 7*. London, The British Library, 1986.
16. DOMBROWSKI, S.R., VISCONTI, J.A. National Audit of Drug Information Centres. *American Journal of Hospital Pharmacy*, Vol.42, 1985, pp.819-26

17. ROSENBERG, J.M., MARTINO, F.P., KIRSCHENBAUM, H.L., ROBBINS,J. Pharmacist-operated Drug Information Centers in the United States - 1986. *American Journal of Hospital Pharmacy*, Vol.44, 1987, pp.337-44.

18. GALLO, G.R., WERTHEIMER, A.I. An International Survey of Drug Information Centers. *Drug Information Journal*, Vol.19, 1985, pp.57-61.

19. CARDONI, A.A., PALMER, M.A., GROVER, R. Drug Information and the Community Pharmacist. *Journal of the American Pharmaceutical* Association, Vol.17, 1977, pp.680-84.

20. COUPER, I.A. Hospital-based Drug Information Centres. *Journal of Hospital Pharmacy*, Vol.30, 1972, pp.125-9.

21. CASSIDY, S.L., KOSTREWSKI, B.J. An Evaluation of Information Sources in Household Product Poisoning. *Journal of Information Science*, Vol. 12, 1986, pp.143-51.

22. MARCUS, C. *The Use of Martindale: the Reference Book and the Databank*. M.Sc. Thesis. London, The City University, 1981.

23. MILES, C.S. *Martindale: a Comparative Study*. M.Sc. Thesis. London, The City University, 1981.

24. *A.B.P.I. Data Sheet Compendium*. London, Datapharm Publications Limited.

25. IOWA DRUG INFORMATION SERVICE. Iowa City, The University of Iowa.

26. NATIONAL ABSTRACTS SERVICE. London, Drug Information Pharmacists Group.

27. GREEN, E. *A Comparison of Two Fulltext Databanks of Drug Information: Martindale Online and Drug Information Fulltext*. M.Sc. Thesis. London, the City University, 1984.

28. *Drug Information Fulltext*. Washington, D.C., American Society of Hospital Pharmacists.

29. ESSEX, A., KOLODIZIEJ, B., SHAYLOR, R. Martindale Online - an Evaluation. *AIOPI Newsletter*. No.1, 1987, pp.3-5.

30. WINDHOLTZ, M.(Ed.). *The Merck Index. 10th Edition*. Rahway, N.J., Merck Co. Inc., 1983.

31. *Sedbase*. Amsterdam. Excerpta Medica.

32. SCOTTISH POISONS INFORMATION SERVICE. Edinburgh, The Royal Infirmary.

33. *VADIS*. Edinburgh, The Royal Infirmary.

34. *POISINDEX, DRUGDEX, EMERGINDEX, IDENTIDEX, CCIS.* Denver, Co, Micromedex Inc.

35. *SWEDIS.* Uppsala, Socialstyrelsens lakemedeksavdelsing.

36. *Philex.* Exeter, Exeter Database Systems Ltd.

37. STOCKLEY, I.H. Drug Interaction Service. University of Nottingham.

38. NATIONAL BOARD OF HEALTH AND WELFARE,DEPT OF DRUGS. *Preliminary Report on: Questionaire on National Drug Databases in Europe*, Uppsala, 1986.

39. REYNOLDS, J.E.F. Databases and Patient Information. *Pharmaceutical Journal*, Vol. 238, 1987, p.446.

40. CLEVERDON, C.W., MILLS, J., KEEN, E.M. *Factors Determining the Performance of Indexing Systems, Aslib Cranfield Research Project*, Cranfield, College of Aeronautics, 1966.

41. KEEN, E.M. The Aberystwyth Index Languages Test. *Journal of Documentation*, Vol.29, 1973, pp.1-35.

42 SPARCK JONES, K. Retrieval Systems Tests 1958 - 1978, in SPARCK JONES, K. (Ed.). *Information Retrieval Experiment.* London, Butterworth, 1981, pp.213-55.

43. PITERNICK, A. B. Searching Vocabularies: a Developing Category of Online Search Tools. *Online Review*, Vol. 8, 1984, pp.441-9.

44. ANON. Indexing Terms for Our Abstracts. *Annals of Internal Medicine*, Vol.105, 1986, pp.11-12.

45. FIDEL, R. Writing Abstracts for Free-text Searching. *Journal of Documentation*, Vol. 42, 1986, pp.11-21.

46. HENZLER, R.G. Free or Controlled Vocabularies. *International Classification*, Vol. 5, 1978, pp.21-26.

47. MARKEY, K., ATHERTON, P., NEWTON, C. An Analysis of Controlled Vocabulary and Free Search Statements in On-line Searches. *Online Review*, Vol. 4, 1980, pp.225-36.

48. HERSEY, D.F., FOSTER, W.R., STALDER, E.W., CARLSON, W.T. Free Text Word Retrieval and Scientist Indexing: Performance Profiles and Costs. *Journal of Documentation*, Vol.27, 1971, pp.167-83.

49. DUCKITT, P. The Value of Controlled Indexing Systems in Online Full Text Databases. *Proceedings of the 5th International Online Information Meeting*, Oxford, Learned Information, 1981, pp.447-53.

50. DUBOIS, C.P.R. Free Text Versus Controlled Vocabulary: a Reassessment. *Online Review*, Vol. 11, 1987, pp.243-53.

51. HARTER, S.P. *Online Information Retrieval. Concepts, Principles, and Techniques*. Orlando, Fa, Academic Press, 1986, pp.56-8.

52. MACKAY, E.M. *Comparison of Controlled and Natural Language Searching in Martindale Online*. M.Sc. Thesis. London, The City Unversity, 1984.

53. VICKERY, A., BROOKS, H. Expert Systems and Their Applications in LIS. *Online Review*, Vol.11, 1987, pp.149-65.

54. O'CONNOR, J. Computer Use of a Medical Dictionary to Select Search Words. *Information Processing and Management*, Vol. 22, 1986, pp.477-86.

55. SALTON, G., LESK, M.E. Computer Evaluation of Indexing and Text Processing. in SALTON, G. (Ed.). *The SMART Retrieval System, Experiments in Automatic Document Processing*, Englewood Cliffs, NJ, Prentice Hall, 1971, p.173.

56. LANCASTER, F.W. *Evaluation of the MEDLARS Demand Search Service*, Bethesda, Md, National Library of Medicine, 1968.

57. LANCASTER, F.W. Aftermath of an Evaluation. *Journal of Documentation*, Vol. 27, 1971, pp.1-10.

58. MARTINDALE, W. *The Extra Pharmacopoeia*, London, H.K. Lewis, 1883.

59. *British Pharmacopoeia*, London, Spottiswoode, 1867.

60. ANON. *Pharmaceutical Journal*, Vol. 231, 183, 493.

61. TODD, R.G. (Ed.). *Martindale, The Extra Pharmacopoeia. 25th Edition*, London, The Pharmaceutical Press.

62. KIRCZ, J.G., BLEEKER, J. The Use of Relational Databases for Electronic and Conventional Scientific Publishing. *Journal of Information Science*, Vol. 13, 1987, pp.75-89.

63. GILMOUR, M.J. Martindale: A Good Read. *Pharmaceutical Journal*, Vol.229, 1982, pp.761-3.

64. CÔTÉ, R.A.(Ed.). *Systematized Nomenclature of Medicine. 2nd Edition*. Skokie, Il, College of American Patholigists, 1979.

65. *MALIMET*. Amsterdam, Excerpta Medica.

66. WORLD HEALTH ORGANIZATION. *International Classification of Diseases. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death*. Geneva, World Health Organization, 1977.

67. EATON, J.K. *Reference Language for a New Drug Database*. M.Sc. Thesis. London, The City University, 1979.

68. MAYNE J. *The Compilation of a User Guide for Martindale Online*. M.Sc. Thesis. London, The City University, 1984.

69. WILLCOXON, F. *Biometrics Bulletin*, Vol. 1, 1945, p.40.

70. LENTNER, C. *Geigy Scientific Tables. Volume 2*. Basle, Ciba Geigy, 1982.

71. COLQUOHOUN, D. *Lectures in Biostatistics*, Oxford, Clarendon Press, 1971.

72. FISHER, R.A., YATES, F. *Statistical Tables for Biological, Agricultural and Medical Research*, Edinburgh, Oliver and Boyd, 1936.

73. REYNOLDS, J.E.F. Martindale's Development and Future. *Pharmaceutical Journal*, Vol. 231, 1983, pp.507-8.

74. *Medical Subject Headings*. Bethesda, Md, National Library of Medicine.

75. FENICHEL, C.H. An Examination of the Relationship between Searching Behavior and Searcher Background. *Online Review*, Vol. 4, 1980, 341-7.

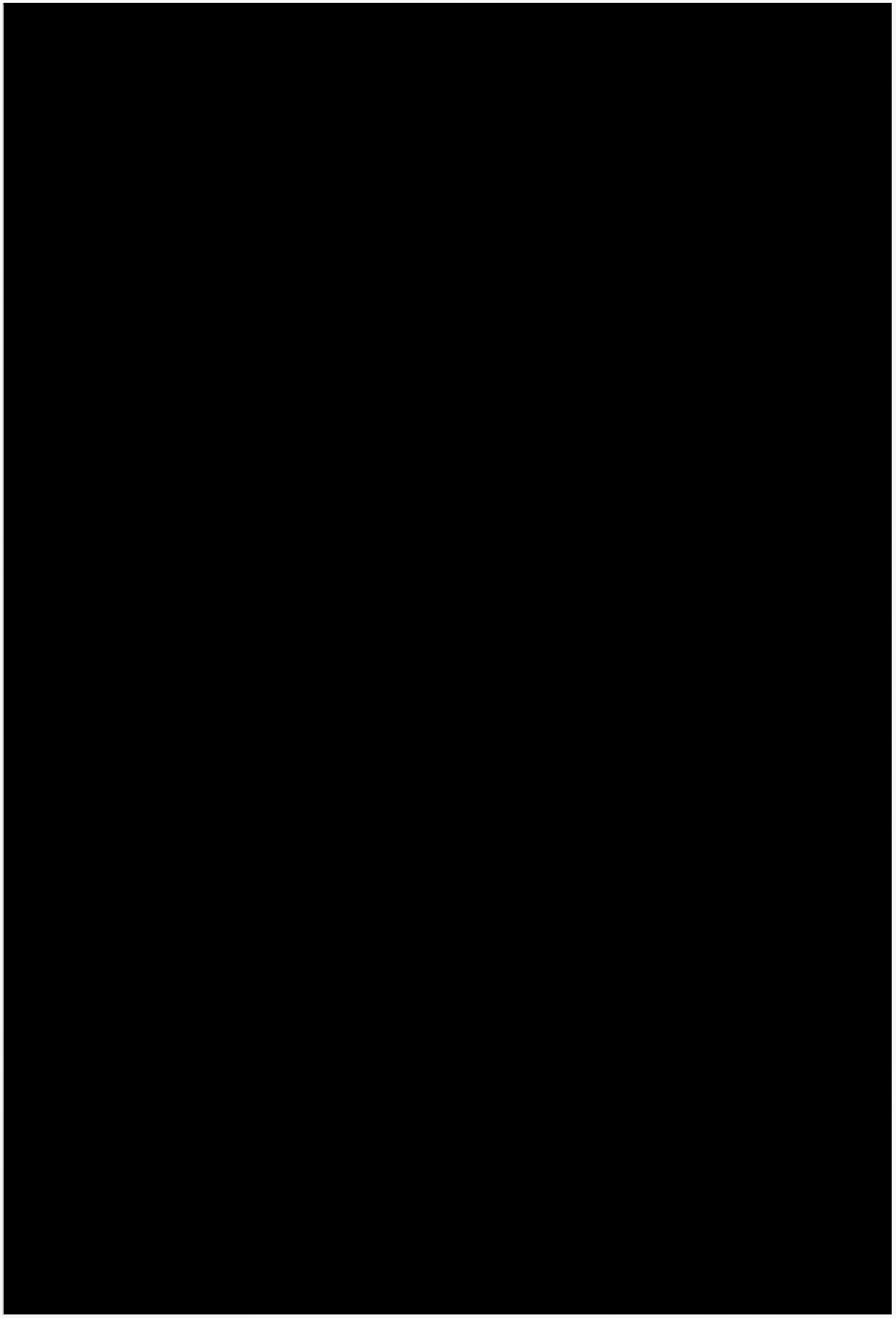
76. NORTON, R.A., WESTWATER,J. Starting End-users. *Aslib Proceedings*, Vol.38, 1986, pp.381-8.

Appendix A
Record structures
(p.127-131)
has been removed for
copyright reasons

Appendix B
Sample thesaurus pages
(p.132-133)
has been removed for
copyright reasons

Appendix C
Summary of indexing guidelines
(p.134-145)
has been removed for
copyright reasons

Appendix D
Guidelines for vendors
(p.146-155)
has been removed for
copyright reasons



APPENDIX E

QUERIES USED FOR ASSESSMENT OF RETRIEVAL EFFECTIVENESS

Group I

1 What alternative CNS stimulants are there to Ritalin?
2 Are there any new uses for lavender oil?
3 Are there any new drugs for phenylketonuria?
4 Can anything be given orally to reduce facial hair?
5 What are the side-effects of vitamin B6?
6 Are ivermectin and avermectin the same?
7 Use of taurine in cardiovascular disease?
8 Is there an interaction between prednisolone and oral contraceptives?
9 What is used to dilute, extend, or store semen?
10 What gelatin plasma expanders are available?
11 Can Bisto be taken with MAOI inhibitors?
12 Information on the availability and/or formulations of lithium succinate.
13 Use of hydrogen peroxide for cleaning contact lenses.
14 Are any tetracycline sprays available?
15 Use of minoxidil in baldness.
16 Grade of chlorinated lime used for purifying drinking water?
17 Stability of phenobarbitone in aqueous solution?
18 Nonpharmaceutical uses of gum arabic?
19 Which analgesic should be given to a patient with cirrhosis?
20 Is strontium acetate used in toothpaste formulations?
21 Are any drugs obtained from plants in the tropical rain forest?
22 Is codeine safe in pregnancy?
23 Why do you need to take plenty of water with lithium?
24 Which diuretic is safe when breast feeding?
25 Use of feverfew in migraine?
26 What is used for lice infestation of eyelashes?
27 Which cough preparation to give to a diabetic?
28 Does a nicotine spray exist?
29 Ampicillin in pregnancy?

30 How to collect large spillages of mercury?
31 How to remove the smell of spilt methyl salicylate?
32 Is gelatin in gelatin capsules of plant or animal origin?
33 What is best to treat ringworm in a pregnant patient?
34 Is there a glibenclamide/aspirin interaction?

Group II

35 Normal dose of clopenthixol depot injection?
36 Microbial contamination of isopropyl alcohol?
37 Toxicity of dyestuffs used in food?
38 Use of salicylsulphonic acid and its formulations?
39 Treatment of tuberous sclerosis?
40 Local anaesthetic injection, reactions?
41 Hydroquinone cream - use and formulations?
42 Details of a honey dressing?
43 Which vitamin is used to treat fragile capillaries?
44 Effects of mepacrine on the liver?
45 Carcinogenicity of Sudan IV?
46 Candida extract for intradermal use?
47 Use of topical cefotaxime?
48 Use of chloroform water in infants?
49 Reports of progesterone causing cancer?
50 Dangers of chewing betel nut?
51 Adverse effects on the blood of disulfiram?
52 Use of AMP in pain?
53 Moduretic causing hyponatraemia?
54 Is clindamycin absorbed topically?
55 Substance that turns pink when wet?
56 Crystal size of adrenaline?
57 Nitrazepam causing depression (of mood)?
58 Diluent for polio oral vaccine?
59 Details of types of oral contraceptives?
60 Iodine content of kelp?
61 Use of anisaldehyde?
62 Detection of benzodiazepines on the breath?
63 Use of ephedrine nose drops in neonates?
64 Half-life of potassium perchlorate?

65 Hazards of camphorated oil?

66 Use of iodine to sterilise drinking water?

Group III

67 Possible adverse effects of topically applied menthol.

68 Carbamazepine in pregnancy?

69 Interaction between colchicine and warfarin?

70 Calcium and vitamin D supplement for a four-year-old unable to take milk (what doses and source)?

71 Maximum sublingual dose of glyceryl trinitrate per 24 hours?

72 Brewers yeast/vitamin B as an insect repellant?

73 Increasing aqueous solubility of salicylic acid - which polyethylene glycol to be used (i.e. which one is best and at what percentage)?

74 Dosage of piroxicam in the elderly?

75 Interaction between Bronztan tablets and warfarin?

76 Desirability of giving high doses (100 mg or greater) of sodium perchlorate injection to block the thyroid before technetium brain scan

77 Glyceryl trinitrate adsorption onto polypropylene syringes - does it occur?

78 Strength of saltpetre to use as a preservative?

79 Amyl nitrite in cyanide poisoning, is it injected or inhaled?

80 Use of papaine in mouth ulcers?

81 Use of hydrogen peroxide ear drops?

82 Dosing of phenazopyridine - before or after meals?

83 Lomotil liquid - is it sugar free?

84 What are the normal blood cholesterol levels?

85 Use of water soluble vitamin K in malabsorption?

86 Potential of castor oil as an abortifacient?

87 Parenteral formulation of thiazide diuretic?

88 Use of benzaldehyde solution in cancer?

89 Auralgen drops - what quantity used in ear?

90 Alupent syrup, is it sugar-free?

91 Is citric soda the same as potassium citrate?

92 In which infusion fluids is ampicillin unstable?

93 Is there an interaction between pyridostigmine and phenylamine?

94 Is lactulose safe in pregnancy?

95 What is the dose for aspirin and papaveretum tablets?

96 Can you preserve hydrogen peroxide solution, if so with what?

97 Are there any official preparations containing rose or extract of rose?

98 Adverse effects of the 'morning-after pill'?

APPENDIX F

REJECTED QUERIES

Group I

1 What is the truth drug?
2 How much chloral hydrate is in chloral syrup BPC?
3 What is acidophilus?
4 Information on a new alkylating agent?
5 Constituents of aqua regis?
6 Constituency of tallow?
7 Constituency of Baby Bio?

Group II

8 Information on clebopride?
9 Details of Condy's fluid.
10 Uses of amino acids (too general).
11 Formula of calcium gluconate gel.
12 What is silver sand?
13 Details of vitamin K4.
14 Citric acid mouthwash?
15 Castor oil to promote abortion (duplicate query)

Group III

16 What is denatured alcohol?
17 What is SVI alcohol?
18 Formula of clindamycin lotion?
19 Formula of ointment of cocoa butter?
20 What is chlorodyne mixture?
21 What is a scoville unit?
22 Formula of Gabriel's lotion?
23 Formula of wool alcohol's ointment?

APPENDIX G

SEARCH STATEMENTS AND RESPONSES

Group I

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
1	**What alternative CNS stimulants are there to Ritalin?**		[30	20]
S1A	(central nervous system stimulant?) or (cns stimulant)	10	6	3
S1B	(stimulant? and cns and use?) not (ritalin or methylphenidate)	2	2	1
S2A	as for S1A	10	6	3
S2B	as for S1B	30	16	11
S3A	(central-stimulants£ or cns-stimulants£ or supplementary-central-and-respiratory-stimulants£) with (use or act)	127		
	Because this response was so large, only uses text records were printed. Of the 64 text records, 26 and 19 were relevant as shown	64	26	19
S3B	(A117.8£ or A532.197£) and (text and uses).RF.	30	21	18
S4A	as for S3B	30	21	18
S4B	as for S3B	30	21	18

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
2	**Are there any new uses for lavender oil?**		[7	0]
S1A	(lavender oil) or lavender or (oleum lavendulae) or (lavender intermedia) or (lavender angustifolia)	26	5	0
S1B	(lavender oil) and use?	9	5	0
S2A	as for S1A	26	5	0
S2B	as for S1B	9	5	0
S3A	(lavender or labiatae) with use	4	4	0
S3B	(lavender or labiatae) with use	4	4	0
S4A	lavender or (lavendin oil) or (lavender oil) or (spike lavender oil) or (lavendula adj (latifolia or officinalis or spica or vera))	36	7	0
S4B	[as S4A] and use?	10	6	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
3	**Are there any new drugs for phenylketonuria?**		[6	0]
S1A	phenylketonuria	20	6	0
S1B	phenylketonuria and (use? or treat?)	14	5	0
S2A	as for S1A	22	6	0
S2B	as for S2B	20	6	0
S3A	use with phenylketonuria	14	4	0
S3B	(use or usem) with phenylketonuria	14	4	0
S4A	as for S1A	20	6	0
S4B	as for S1B	14	5	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
4	**Can anything be given orally to reduce facial hair?**		[15	15]
S1A	(hirsut? or (facial hair)) and (oral or mouth)	5	2	2
S1B	((hair and (face or facial)) or beard) and (use? or treat? or reduc?)	12	1	1
S2A	as for S1A	6	2	2
S2B	as for S1B	14	1	1
S3A	(hirsutism or hair or polycystic-ovary-syndrome) with (oral-dosage-form or mouth or oral-route) with use	8	4	4
S3B	as for S3A	8	4	4
S4A	(hirsutism or hair or (polycystic adj ovary adj syndrome) and (use? or treat?)	158	[modified for reduction]	
	hirsutism and (use? or treat?)	29	13	13
S4B	(hirsutism or hair or (polycystic adj ovary adj syndrome) and (use? or treat?) and (face or facial)	7	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
5	**What are the side-effects of vitamin B6?**		[1	0]
S1A	((vitamin b6) or pyridoxine) and (side or (adverse effect?))	29	0	0
S1B	(adverse or side?) and vitamin b6	0	0	0
S2A	as for S1A	29	0	0
S2B	as for S1B	0	0	0
S3A	pyridoxine with ae	1	1	0
S3B	pyridoxine with ae	1	1	0
S4A	(pyridoxine or pyridoxal) and ((adverse effect) or (adverse reaction) or (side effect?))	29	0	0
S4B	(pyridoxal or pyridoxine) and (adverse or toxic or side?)	30	0	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
6	**Are ivermectin and avermectin the same?**		[1	1]
S1A	ivermectin and avermectin	1	1	1
S1B	as for S1A	1	1	1
S2A	as for S1A	1	1	1
S2B	as for S1A	1	1	1
S3A	ivermectin with des	1	1	1
S3B	as for S3A	1	1	1
S4A	ivermectin or avermectin	2	1	1
S4B	as for S4A	2	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
7	**Use of taurine in cardiovascular disease**		[1	1]
S1A	((taurine or (aminoethanesulphonic acid)) and (cardiovascular? or heart?)	1	1	1
S1B	taurine and (use? or treat?) and (cardiovascular or heart or cardiac or vascular)	1	1	1
S2A	as for S1A	1	1	1
S2B	as for S1B	1	1	1
S3A	taurine with use with (cardiovascular-disorders£ or actions-and-uses-relating-to-the-cardiovascular-system£ or cardiovascular-system-physiology£)	1	1	1
S3B	taurine with use with cardiovascular-disorders£	1	1	1
S4A	taurine and [cardiovascular disorders]* or (cardiovascular system)	1	1	1
S4B	as for S1B	1	1	1

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
8	**Is there an interaction between prednisolone and oral contraceptives?**		[1	0]
S1A	(prednisolone or corticosteroid?) and ((oral contracepti?) or (sex hormone?))	5	1	0
S1B	prednisolone and interact? and (oral contraceptive?)	0		
S2A	as for S1A	7	1	0
S2B	as for S2B	0		
S3A	prednisolone with int with oral-contraceptives£	0		
S3B	1 = sex-hormones£ not (anabolics£ or androgens£) 1 with prednisolone with int	0		
S4A	[the equivalent of 1 from S3B] and [corticosteroids] and interact?	2	1	0
S4B	prednisolone and interact? and [oral contraceptives]	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
9	**What is used to dilute, extend or store semen?**		[0	0]
S1A	semen and (stabil? or dilut? or exten? or stor?)	6	0	0
S1B	(sperm or semen) and (diluent or dilut? or shelf or storage or store or extender)	9	0	0
S2A	as for S1A	7	0	0
S2B	as for S1B	9	0	0
S3A	semen with (storage or stability)	0		
S3B	semen with (storage or stability or diluting)	0		
S4A	(semen or ejaculate or seminal fluid)) and (dilut? or storage or stability)	1	0	0
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
10	**What gelatin plasma expanders are available?**		[10	5]
S1A	gelatin and (plasma or blood or infusion)	22	9	5
S1B	gelatin and preparation? and (plasma expand?)	1	1	1
S2A	as for S1A	23	10	5
S2B	as for S1B	1	1	1
S3A	gelatin with (parenteral-dosage-form or infusion or injection) with (prop or nprop)	3	1	1
S3B	as for S3A	3	1	1
S4a	gelatin and ((plasma products) or plasma or blood)) and ((infusion or injection or (intravenous injection) or (intravenous route))	5	2	1
S4B	as for S4A	5	2	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
11	**Can Bisto be taken with MAOI inhibitors?**		[0	0]
S1A	((maoi? or (monoamine oxidase inhibit?)) and (bisto or gravy)	0		
S1B	((maoi or (monoamine oxidase) and (gravy or bisto)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	((maoi-antidepressants£ or maoi-antihypertensives£ or maoi-inhibitor) with int) and bisto	0		
S3B	((maoi-antidepressants£ or maoi-antihypertensives£ or maoi-inhibitor) with int) and (bisto or gravy)	0		
S4A	([maoi-antidepressants] or [maoi-antihypertensives] or (maoi inhibitor) and (bisto or gravy)	0		
S4B	as for S4A	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
12	**Information on the availability and /or formulations of lithium succinate.**		[0	0]
S1A	lithium succinate	0		
S1B	preparation and (lithium succinate)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	as for S1A	0		
S3B	as for S1B	0		
S4A	as for S1A	0		
S4B	as for S1B	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
13	**Use of hydrogen peroxide for cleaning contact lenses.**		[0	0]
S1A	(hydrogen peroxide) and (contact lens?)	0		
S1B	(hydrogen peroxide) and (contact lens?) and clean?	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	hydrogen-peroxide with contact-lenses&	0		
S3B	as for S3A	0		
S4A	(hydrogen peroxide) and ((contact lens?) or (eye care) or (contact lens solution?))	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
14	**Are any tetracycline sprays available?**		[2	2]
S1A	(tetracycline? or chlortetracycline or demeclocycline or doxycycline or lymecycline or meclocycline or minocycline or oxytetracycline or rolitetracycline) and (spray? or aerosol? or nebuliser?)	5	2	2
S1B	(tetracycline? or chlortetracycline or demeclocycline or doxycycline or lymecycline or meclocycline or minocycline or oxytetracycline or rolitetracycline or clomocycline) and (spray? or aerosol? or nebuliser?) and preparation?	3	2	2
S2A	as for S1A	5	2	2
S2B	as for S1B	3	2	2
S3A	tetracycline& with (spray or aerosol or topical-medicated-spray)	2	2	2
S3B	(tetracycline& with (spray or aerosol or topical-medicated-spray)) and (preparation?).rf.	2	2	2
S4A	[tetracycline] and (spray? or aerosol? or nebuliser?)	4	2	2
S4B	[tetracycline] and (spray or aerosol or (topical medicated spray)) and (preparation?).rf.	3	2	2

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
15	**Use of minoxidil in baldness**		[6	0]
S1A	minoxidil and (bald? or alopecia?)	0		
S1B	as for S1A	0		
S2A	as for S1A	0		
S2B	as for S1A	0		
S3A	minoxidil with (alopecia or hair)	5	4	0
S3B	as for S3A	5	4	0
S4A	minoxidil and (baldness or alopecia? or (hair loss) or hair)	1	0	0
S4B	as for S4A	1	0	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
16	**Grade of chlorinated lime used for purifying drinking water**		[1	0]
S1A	((chlorinated lime) or (bleaching powder) or chlorine) and water and (purif? or sterilis?)	6	0	0
S1B	(chlorinated lime) and water and purif?	0		
S2A	as for S1A	7	0	0
S2B	as for S1B	0		
S3A	chlorinated-lime with purification with (water or water-pollution or water-supply)	0		
S3B	as for S3A	0		
S4A	(chlorinated lime) and (water or (water pollution) or purification or (water purification) or (water supply))	5	0	0
S4B	as for S4A	5	0	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
17	**Stability of phenobarbitone in aqueous solution**		[7	7]
S1A	phenobarbit? and stability and (aqueous or water)	1	1	1
S1B	phenobarbitone and stability and solution and aqueous	0		
S2A	as for S1A	16	6	6
S2B	as for S1B	1	1	1
S3A	phenobarbitone with (stability or deterioration or storage) with pharm with solution with water	0		
S3B	phenobarbitone with (stability or deterioration or storage) with pharm with solution	3	3	3
S4A	phenobarbitone and stability and (water or solution)	1	1	1
S4B	phenobarbitone and (stability or deterioration) and (water or solution)	1	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
18	**Nonpharmaceutical uses of gum arabic**		[3	2]
S1A	((gum arabic) or acacia) and use?	15	3	2
S1B	as for S1A	15	3	2
S2A	as for S1A	21	3	2
S2B	as for S1A	21	3	2
S3A	acacia with use	11	3	2
S3B	as for S3A	11	3	2
S4A	as for S1A	15	3	2
S4B	as for S1A	15	3	2

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
19	**Which analgesic should be given to a patient with cirrhosis?**		[13	11]
S1A	(cirrhosis or liver) and analgesi?	13	3	3
S1B	analgesic? and cirrhosis and (use? or treat? or choice)	0		
S2A	as for S1A	22	7	6
S2B	as for S1B	1	0	0
S3A	(analgesics£ or supplementary-analgesics£ or supplementary-narcotic-analgesics£) with (cirrhosis£ or alcoholism or chronic-hepatic-failure)	3	3	3
S3B	as for S3A	3	3	3
S4A	(cirrhosis or alcoholism or (chronic hepatic failure) or (primary biliary cirrhosis)) and (analgesi? or pain)	9	1	1
S4B	(cirrhosis or alcoholism or (chronic hepatic failure) or (primary biliary cirrhosis)) and ([analgesics] or [narcotic analgesics])	7	6	6

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
20	**Is strontium acetate used in toothpaste formulations?**		[0	0]
S1A	(strontium acetate) and (toothpaste? or toothpowder? or dentifrice?)	0		
S1B	(strontium acetate) and toothpaste	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	(strontium acetate) and (dentifrices or teeth or dental-care£)	0		
S3B	as for S3A	0		
S4A	(strontium acetate) and (dentifrice? or toothpaste? or teeth or formulation)	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
21	**Are any drugs obtained from plants in the tropical rain forest?**		[0	0]
S1A	(plant? or (medicinal plant?) or herb) and (tropic? or forest?)	1	0	0
S1B	(tropical rain forest) and (deriv? or source)	0		
S2A	as for S1A	1	0	0
S2B	as for S1B	0		
S3A	tropical-climate with (plants& or plant-anatomy&)	0		
S3B	tropical and sou	0		
S4A	plant? and ((rain forest) or (tropical climate) or tropics)	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
22	**Is codeine safe in pregnancy?**		[4	4]
S1A	codeine? and pregnan?	4	2	2
S1B	as for S1A	4	2	2
S2A	as for S1A	5	3	3
S2B	as for S1A	5	3	3
S3A	codeine with pregnancy£ with precs	1	1	1
S3B	codeine with preg with (precs£ or use£)	3	2	2
S4A	codeine and (pregnancy or foetus or teratogen? or [congenital disorders]	3	2	2
S4B	as for S4A	3	2	2

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
23	**Why do you need to take plenty of water with lithium?**		[4	1]
S1A	lithium? and (water or fluid?)	19	3	1
S1B	lithium and water and (administ? or dose or dosage or take?)	4	1	1
S2A	as for S1A	19	3	1
S2B	as for S1B	4	1	1
S3A	lithium with water with precs	0		
S3B	(lithium with ad) and (water or fluid?)	0		
S4A	lithium and water and administration	3	0	0
S4B	as for S1B	4	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
24	**Which diuretic is safe when breast feeding?**		[12	4]
S1A	(diuretic? or diuresis or thiazide?) and (lactat? or breastfeed? or milk)	9	0	0
S1B	diuretic? and ((breast feeding) or lactat?)	4	0	0
S2A	as for S1A	10	0	0
S2B	as for S1B	6	0	0
S3A	(breast-feeding or lactation or breast-milk) with diuretics&	12	12	4
S3B	as for S3A	12	12	4
S4A	[diuretics] and ((breast feeding) or lacatat?)	2	0	0
S4B	as for S4A	2	0	0

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
25	**Use of feverfew in migraine**		[0	0]
S1A	(migraine or headache?) and (feverfew or (chrysanthemum parthenum))	0		
S1B	feverfew and migraine	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	headache&.de. and (feverfew or (chrysanthemum parthenum))	0		
S3B	as for S3A	0		
S4A	(feverfew or (chrysanthemum parthenum)) and (migraine or headache or (cluster headache))	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
26	**What is used for lice infestation of eyelashes?**		[1	1]
S1A	(lice or louse or pediculosis) and (eyelash? or lash?)	1	1	1
S1B	(lash or pediculosis) and eyelash?	1	1	1
S2A	as for S1A	1	1	1
S2B	as for S1B	1	1	1
S3A	pediculosis with use with eye&	1	1	1
S3B	as for S3A	1	1	1
S4A	as for S1A	1	1	1
S4B	as for S1B	1	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
27	**Which cough preparation to give to a diabetic?**		[3	3]
S1A	(cough or anti-tussive or expectorant) and (diabetes mellitus)	0		
S1B	cough and (diabetes or diabetic) and (use? or treat?)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	diabetes-mellitus£ with (cough-suppressants£ or supplementary-cough-suppressants£ or expectorants£ or coughing or cough) with use	0		
S3B	as for S3A	0		
S4A	(diabetes mellitus) and (cough or [cough suppressants] or antitussive? or [expectorants])	0		
S4B	(diabetes mellitus) and ([cough suppressants] or [supplementary cough suppressants] or [expectorants] or cough or coughing) and use	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
28	**Does a nicotine spray exist?**		[0	0]
S1A	nicotine and (spray? or aerosol? or nebulis?)	1	0	0
S1B	as for S1A	1	0	0
S2A	as for S1A	1	0	0
S2B	as for S1B	1	0	0
S3A	(nicotine or tobacco-products) with (spray or aerosol or inhalation-spray or nasal-spray or nebuliser or throat-spray)	1	0	0
S3B	as for S3A	1	0	0
S4A	(nicotine or (tobacco product?)) and (spray? or aerosol? or nebuliser?)	1	0	0
S4B	as for S4A	1	0	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
29	**Ampicillin in pregnancy?**		[7	7]
S1A	ampicillin and (pregnan? or teratogen or foetus)	7	6	6
S1B	as for S1A	7	6	6
S2A	as for S1A	10	7	7
S2B	as for S1B	10	7	7
S3A	ampicillin with pregnancy£ with precs	2	2	2
S3B	ampicillin with preg with precs	2	2	2
S4A	ampicillin and (pregnancy or foetus or teratogen? or (congenital disorder?))	4	3	3
S4B	as for S4A	4	3	3

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
30	**How to collect large spillages of mercury?**		[1	1]
S1A	mercury and (disposal or spill? or waste)	1	1	1
S1B	mercury and (spill? or collect?)	1	1	1
S2A	as for S1A	1	1	1
S2B	as for S1B	1	1	1
S3A	mercury with waste-disposal	0		
S3B	as for S3A	0		
S4A	mercury and (spillage or (waste disposal))	1	1	1
S4B	as for S4A	1	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
31	**How to remove the smell of spilt methyl salicylate?**		[0	0]
S1A	((methyl salicylate) or wintergreen?) and (spill? or disposal or waste)	0		
S1B	smell and remov? and (methyl salicylate)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	methyl-salicylate with waste-disposal	0		
S3B	as for S3A	0		
S4A	(methyl salicylate) and (spillage or (waste disposal)	0		
S4B	as for S1B	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
32	**Is gelatin in gelatin capsules of plant or animal origin?**		[1	1]
S1A	gelatin and capsule?	31	1	1
S1B	gelatin and capsule? and (plant or animal)	2	1	1
S2A	as for S1A	31	1	1
S2B	as for S1B	2	1	1
S3A	gelatin with sou	0		
S3B	as for S3A	0		
S4A	gelatin and capsule? and (source or origin)	0		
S4B	gelatin and capsule? and (source or origin) and (plant or animal)	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
33	**What is used to treat ringworm in a pregnant patient?**		[0	0]
S1A	pregnan? and (ringworm or tinea? or antifungal)	0		
S1B	(ringworm or tinea) and pregnan?	0		
S2A	as for S1A	8	0	0
S2B	as for S1B	0		
S3A	tinea& with pregnancy with use	0		
S3B	tinea with preg with use	0		
S4A	as for S1B	0		
S4B	as for S1B	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
34	**Is there a glibenclamide/aspirin interaction?**		[0	0]
S1A	glibenclamide and aspirin	0		
S1B	glibenclamide and aspirin and interact?	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	glibenclamide with int with aspirin	0		
S3B	as for S3A	0		
S4A	glibenclamide and interaction and (aspirin or salicylate? or (acetylsalicylic acid))	0		
S4B	as for S4A	0		

Group II

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
35	**Normal dose of clopenthixol depot injection**		[1	1]
S1A	clopenthixol and dose and (depot injection)	0		
S1B	clopenthixol and (dose or dosage) and injection	1	1	1
S2A	as for S1A	0		
S2B	as for S1B	3	1	1
S3A	clopenthixol with ad with injection	2	1	1
S3B	as for S3A	2	1	1
S4A	clopenthixol and (injection or (parenteral dosage form) or (intramuscular route) or (subcutaneous route)) and (dose or dosage)	1	1	1
S4B	as for S4A	1	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
36	**Microbial contamination of isopropyl alcohol**		[0	0]
S1A	(isopropyl alcohol) and (contamination, microbial)	0		
S1B	(isopropyl alcohol) and contamin?	6	0	0
S2A	as for S1A	0		
S2B	as for S1B	6	0	0
S3A	isopropyl-alcohol with microbial-contaminants	0		
S3B	as for S3A	0		
S4A	(isopropyl alcohol) and ((microbial contaminants) or impurities)	6	0	0
S4B	(isopropyl alcohol) and ((microbial contaminants) or impurities or contamin?)	6	0	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
37	**Toxicity of dyestuffs used in food**		[32	27]
S1A	(colouring, food) and toxicity	0		
S1B	(dye? or colour? or color?) and food and (adverse or toxic?)	8	5	3
S2A	as for S1A	0		
S2B	as for S1B	12	7	5
S3A	colouring-agents& with foods with ae	0		
S3B	colouring-agents& with foods& with ae	6	5	3
S4A	(colouring or amaranth or caramel) and (adverse or toxic or side) and (foods or (food additives))	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
38	**Use of salicylsulphonic acid and its formulations**		[0	0]
S1A	(salicylsulphonic acid) and use	0		
	(salicylsulphonic acid and formulation	0		
S1B	(salicylsulphonic acid) and (use? or preparation?)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	(salicylsulphonic acid).pr,ti, sy,ab,tx.	0		
S3B	salicylsulphonic acid	0		
S4A	((salicylsulphonic acid) or (salicylsulfonic acid) or (salicyl sulfonic acid)) and ((pharmaceutical manufacturing) and [pharmaceutical form] or use)	0		
S4B	(salicylsulphonic acid)	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
39	**Treatment of tuberous sclerosis**		[1	0]
S1A	(tuberous sclerosis) and treatment	0		
S1B	(tuberous sclerosis) and (use? or treat?)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	(tuberous sclerosis).tx,ab.	1	1	0
S3B	as for S1B	0		
S4A	(tuberous sclerosis) and use	0		
S4B	as for S1B	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
40	**Local anaesthetic injection reactions**		[31	23]
S1A	(anaesthetic, local) and (adverse effect)	0		
S1B	(local anaesthetic) and injection and (adverse or toxic or side)	9	2	0
S2A	as for S1A	0		
S2B	as for S1B	13	4	2
S3A	local-anaesthetics& with ae with injection	6	5	4
S3B	local-anaesthetics& with ae with (injection or parenteral-route)	8	7	6
S4A	([local anaesthetics] or (local anaesthesia)) and injection and (adverse or toxic or side?)	47	26	20
S4B	[local anaesthetics] and (injection or (parenteral route)) and (adverse or toxic or side?)	41	26	20

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
41	**Hydroquinone cream - use and formulations**		[11	7]
S1A	(hydroquinone and use and skin) or (hydroquinone and formulation)	4	3	1
S1B	hydroquinone and (cream or applied) and (use? or preparation?)	3	2	2
S2A	as for S1A	6	5	2
S2B	as for S1B	4	3	3
S3A	(hydroquinone with use) or (hydroquinone with nprop)	9	8	6
S3B	(hydroquinone with use with (ointment or topical-application)) or (hydroquinone with ointment with (nprop or prop))	6	6	5
S4A	(hydroquinone and ointment and use?) or (hydroquinone and ointment and (pharmaceutical manufacturing))	3	3	3
S4B	(hydroquinone and (ointment or cream) and (use or uses)) or (hydroquinone and (ointment or cream) and preparation?)	6	6	5

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
42	**Details of a honey dressing**		[3	2]
S1A	dressing and honey	1	1	0
S1B	honey and (dressing or applied)	2	1	0
S2A	as for S1A	1	1	0
S2B	as for S1B	2	1	0
S3A	honey with dressings£	0		
S3B	honey with (dressings£ or topical-dosage-form£)	0		
S4A	(dressings or (medicated dressings) and honey	0		
S4B	honey and (dressing? or topical)	1	1	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
43	**Which vitamin is used to treat fragile capillaries?**		[4	4]
S1A	vitamin and use and (capillary fragility)	0		
S1B	vitamin and capillar? and fragil?	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	vitamins& with use with (microvasculature or fragility)	0		
S3B	as for S3A	0		
S4A	[vitamins] and use? and ((capillary fragility) or capillar? or (peripheral circulation)	2	2	2
S4B	as for S4A	2	2	2

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
44	**Effects of mepacrine on the liver**		[2	2]
S1A	mepacrine and (adverse effect) and liver	0		
S1B	mepacrine and liver	2	1	1
S2A	as for S1A	0		
S2B	as for S1B	3	2	2
S3A	mepacrine with ae with liver	1	1	1
S3B	as for S3A	1	1	1
S4A	mepacrine and (liver or [liver disorders] and adverse	1	1	1
S4B	as for S4A	1	1	1

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
45	**Carcinogenicity of Sudan IV**		[1	0]
S1A	(sudan iv) and (adverse effect) and cancer	0		
S1B	((sudan iv) or (scarlet red)) and (cancer or carcin? or neoplasm?) and adverse	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	colouring-agents& with carcinogen	6	0	0
S3B	scarlet-red with ae	1	1	0
S4A	(sudan? or (sudan iv) or [diagnostic dyes] or [colouring agents] or [disinfectant dyes] or colorant) and (carcinogens or [malignant neoplasms])	0		
S4B	as for S1B	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
46	**Candida extract for intradermal use?**		[0	0]
S1A	candida and (injection, intradermal)	0		
S1B	candida and (use? or treat?) and intradermal	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	candida with intradermal-route	0		
S3B	as for S3A	0		
S4A	(candida or c) and extract and intradermal	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
47	**Use of topical cefotaxime**		[0	0]
S1A	cefotaxime and use and skin	0		
S1B	cefotaxime and (use? or treat?) and topical	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	cefotaxime with use with topical-dosage-form£	0		
S3B	cefotaxime with use with (topical-application or topical-dosage-form£)	0		
S4A	cefotaxime and ((topical application) or [topical dosage form] or (transdermal route))	0		
S4B	as for S4A	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
48	**Use of chloroform water in infants**		[0	0]
S1A	(chloroform water) and infant	0		
S1B	(chloroform water) and (infant? or child? or neonate?) and (use? or treat?)	1	0	0
S2A	as for S1A	0		
S2B	as for S1B	3	0	0
S3A	chloroform with children with oral-route	0		
S3B	chloroform with (children or neonate) with oral-route	0		
S4A	(chloroform or (chloroform water)) and use and (neonates or children)	0		
S4B	chloroform and (neonate? or child? or infant?) and (use? or treat?) and water	1	0	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
49	**Reports of progesterone causing cancer**		[3	1]
S1A	progesterone and (adverse effect) and cancer	0		
S1B	progesterone and (cancer or neoplasm?) and adverse	1	0	0
S2A	as for S1A	0		
S2B	as for S1B	2	1	0
S3A	progesterone with (carcinogen or malignant-neoplasms£)	3	1	1
S3B	progesterone with ae with malignant-neoplasms£	1	1	1
S4A	progesterone and (adverse or toxic or side?) and (carcinogens or [malignant neoplasms])	1	1	1
S4B	progesterone and (adverse or toxic or side?) and (carcinogens or [malignant neoplasms] or cancer)	1	1	1

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
50	**Dangers of chewing betel nut**		[2	2]
S1A	(betel nut) and toxicity	0		
S1B	(betel nut) and chew? and adverse	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	areca with ae	2	2	2
S3B	as for S3A	2	2	2
S4A	(areca or (betel nut?) and (adverse or toxic or side-effect?)	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
51	**Adverse effects of disulfiram on the blood**		[0	0]
S1A	disulfiram and (adverse effect) and blood	0		
S1B	disulfiram and adverse and (blood or haemat?)	1	0	0
S2A	as for S1A	0		
S2B	as for S1B	1	0	0
S3A	disulfiram with ae with blood£	0		
S3B	disulfiram with ae with (blood£ or blood-disorders£)	0		
S4A	disulfiram and (adverse or toxic or side-effect?) and blood	1	0	0
S4B	as for S4A	1	0	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
52	**Use of AMP in pain**		[0	0]
S1A	amp and use and pain	0		
S1B	((adenosine phosphate) or amp) and (pain or analges?) and (use? or treat?)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	adenosine-phosphate with pain	0		
S3B	as for S3A	0		
S4A	((adenosine phosphate) or amp) and (use or treat?) and (pain or algesic or analgesic)	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
53	**Moduretic causing hyponatraemia**		[5	1]
S1A	amiloride and hydrochlorothiazide and (adverse effect) and sodium	0		
S1B	(moduretic or amiloride) and (hyponatraemia or sodium) and adverse	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	amiloride with hydrochlorothiazide with ae with hyponatraemia	0		
S3B	as for S3A	0		
S4A	amiloride and hydrochlorothiazide and (adverse or toxic or side-effect?) and (hyponatraemia or hyponatremia or (sodium depletion) or hyponatremic)	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
54	**Is clindamycin absorbed topically?**		[1	1]
S1A	clindamycin and (absorption, skin)	0		
S1B	clindamycin and absorb? and (skin or topical)	1	1	1
S2A	as for S1A	0		
S2B	as for S1B	1	1	1
S3A	clindamycin with skin with drug-absorption	0		
S3B	clindamycin with drug-absorption with (skin or topical-application or topical-route)	1	1	1
S4A	clindamycin and (drug absorption) and (skin or (percutaneous route) or (transdermal route)	0		
S4B	clindamycin and absorption and (skin or topical)	2	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
55	**Substance that turns pink when wet?**		[1	0]
S1A	(test or indicator) and water	48	0	0
S1B	pink and (moisture or water)	29	0	0
S2A	as for S1A	50	0	0
S2B	as for S1B	29	0	0
S3A	supplementary-diagnostic-agents& with water	0		
S3B	supplementary-diagnostic-agents& and water	4	1	0
S4A	water and ([diagnostic dyes] or [colouring agents] or (diagnosis and testing))	149	1	0
S4B	as for S4A	149	1	0

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
56	**Crystal size of adrenaline**		[0	0]
S1A	adrenaline and crystal	0		
S1B	adrenaline and size and crystal	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	as for S1A	0		
S3B	as for S1A	0		
S4A	(adrenaline or epinephrine) and (crystal or (physiochemical characteristics) or (particle size))	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
57	**Nitrazepam causing depression (of mood)**		[2	0]
S1A	nitrazepam and (adverse effect) and depression	0		
S1B	nitrazepam and depression and adverse	3	1	0
S2A	as for S1A	0		
S2B	as for S1B	4	2	0
S3A	nitrazepam with ae with mental-depression	1	1	0
S3B	as for S3A	1	1	0
S4A	nitrazepam and (adverse or toxic or side-effect?) and (mental depression) or (mental depressant)	0		
S4B	nitrazepam and depression and (adverse or toxic or side-effect?)	3	1	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
58	**Diluent for polio oral vaccine**		[0	0]
S1A	(poliovaccine, oral) and diluent	0		
S1B	polio and vaccine and diluent	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	oral-poliomyelitis-vaccine with diluent	0		
S3B	as for S3B	0		
S4A	(oral poliomyelitis vaccine) and (diluent or diluting)	0		
S4B	(poliomyelitis or polio) and vaccine and dilu?	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
59	**Details of types of oral contraceptives**		[1	1]
S1A	(contraceptives, oral) and classification	0		
S1B	(oral contraceptives) and classification	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	oral-contraceptives with des	0		
S3B	as for S3A	0		
S4A	(oral contraceptives) and (type? or class?)	5	1	1
S4B	as for S4A	5	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
60	**Iodine content of kelp**		[2	0]
S1A	kelp and iodine	1	1	0
S1B	as for S1A	1	1	0
S2A	as for S1A	1	1	0
S2B	as for S1B	1	1	0
S3A	algae& with iodine	1	1	0
S3B	as for S1A	1	1	0
S4A	iodine and algae	1	0	0
S4B	as for S1A	1	1	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
61	**Use of anisaldehyde**		[0	0]
S1A	anisaldehyde and use	1	0	0
S1B	(anisaldehyde or anise) and use	4	0	0
S2A	as for S1A	1	0	0
S2B	as for S1B	5	0	0
S3A	as for S1A	1	0	0
S3B	as for S1B	5	0	0
S4A	as for S1A	1	0	0
S4B	as for S1B	4	0	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
62	**Detection of benzodiazepines on the breath**		[0	0]
S1A	benzodiazepine and concentration and breath and test	0		
S1B	benzodiazepine? and breath	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	(benzodiazepine-hypnotics£ or benzodiazepine-tranquillisers£) with breath	0		
S3B	as for S3A	0		
S4A	[benzodiazepine hypnotics] and breath and ((diagnosis and testing) or [diagnostic agents])	0		
S4B	([benzodiazepine hypnotics] or [benzodiazepine tranquillisers]) and breath and ((diagnosis and testing) or [diagnostic agents])	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
63	**Use of ephedrine nose drops in neonates**		[1	1]
S1A	ephedrine and (nose drop) and neonate	0		
S1B	ephedrine and (neonate? or infant) and (nose or nasal)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	ephedrine with nasal-drops with neonate with use	0		
S3B	ephedrine with (nasal-drops or nasal-route) with neonate with use	0		
S4A	ephedrine and (nose or (nasal route) or [nasal dosage form]) and neonate and use	0		
S4B	as for S4A	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
64	**Half-life of potassium perchlorate**		[0	0]
S1A	(potassium perchlorate) and half-life	1	0	0
S1B	(potassium perchlorate) and (half life)	0		
S2A	as for S1A	1	0	0
S2B	as for S1B	0		
S3A	potassium-perchlorate with half-life	0		
S3B	as for S3A	0		
S4A	(potassium perchlorate) and (half life)	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
65	**Hazards of camphorated oil**		[7	3]
S1A	(camphorated oil) and toxicity	1	1	0
S1B	(camphorated oil) and adverse	2	2	2
S2A	as for S1A	1	1	0
S2B	as for S1B	2	2	2
S3A	camphor with ae	6	6	3
S3B	as for S3a	6	6	3
S4A	(camphor or (camphorated oil) or (camphor liniment)) and (adverse or toxic or side-effect?)	7	5	3
S4B	as for S4A	7	5	3

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
66	**Use of iodine to sterilise drinking water**		[0	0]
S1A	iodine and water and sterilisation	2	0	0
S1B	iodine and water and sterilis?	24	0	0
S2A	as for S1A	2	0	0
S2B	as for S1B	24	0	0
S3A	chemical-sterilisation with water with iodine	0		
S3B	as for S3A	0		
S4A	iodine and use and water and ((chemical sterilisation) or (chemical steriliser) or disinfection	3	0	0
S4B	as for S4B	3	0	0

Group III

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
67	**Possible adverse effects of topically applied menthol**		[2	2]
S1A	menthol and skin and (adverse reactions)	0		
S1B	menthol and (topic? or applic?) and (adverse or side? or toxic)	1	1	1
S2A	as for S1A	0		
S2B	as for S1B	1	1	1
S3A	menthol with ae with topical-dosage-form	0		
S3B	menthol with ae& with (topical-dosage-form& or topical-application)	2	2	2
S4A	menthol and (topical or skin or transdermal) and (adverse or toxic or side-effect?)	0		
S4B	menthol and (adverse or toxic or side-effect?) and ((topical application) or [topical dosage form])	1	1	1

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
68	**Carbamazepine in pregnancy**		[10	9]
S1A	carbamazepine and teratogenicity	0		
S1B	carbamazepine and (pregnan? or teratogen?)	14	8	7
S2A	as for S1A	0		
S2B	as for S1B	16	9	8
S3A	carbamazepine with ae with (teratogen or pregnancy)	3	3	3
S3B	carbamazepine with (preg or pregm)	11	10	9
S4A	carbamazepine and (pregnancy or congenital or foetus)	14	8	7
S4B	as for S1B	14	8	7

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
69	**Interaction between colchicine and warfarin**		[0	0]
S1A	colchicine and warfarin and (drug interactions)	0		
S1B	colchicine and warfarin and interact?	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	colchicine with int with warfarin	0		
S3B	as for S3A	0		
S4A	([anticoagulants] or warfarin) and interaction and ([antigout agents] or colchicine)	6	0	0
S4B	as for S1B	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
70	**Calcium and vitamin D supplement for a four-year-old unable to take milk; what doses and source?**		[15	10]
S1A	calcium and (vit d) and (food supplement) (not milk)	0		
S1B	calcium and (vitamin d) and childh? and (dose or dosage)	2	1	0
S2A	as for S1A	0		
S2B	as for S1B	2	1	0
S3A	(calcium or vitamin-d-substances) with lactose-intolerance with (ad or prop)	0		
S3B	vitamin-d-substances& with calcium-salts& with ((children with ad) or prop)	8	2	1
S4A	(calcium or [calcium salts] or [vitamin d substances]) and [ad] and (malabsorption or (carbohydrate malabsorption syndrome)	9	0	0
S4B	[vitamin d] and (dose or dosage or admin?) and childh? and ([calcium salts] or calcium) and (preparation? or proprietary)	5	5	3

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
71	**maximum sublingual dose of glyceryl trinitrate per 24 hours**		[1	0]
S1A	(glyceryl trinitrate) and dosage and (sublingual administration)	1	0	0
S1B	(glyceryl trinitrate) and dose and sublingual and maximum and (daily or (24 hours))	0		
S2A	as for S1A	1	0	0
S2B	as for S1B	0		
S3A	glyceryl-trinitrate with ad with buccal-route	10	1	0
S3B	as for S3A	10	1	0
S4A	(glyceryl trinitrate) and ((buccal route) or (buccal tablet)) and (dose or dosage or admin?)	3	0	0
S4B	(glyceryl trinitrate) and (dose or dosage or admin?) and (day or daily) and (buccal or sublingual)	1	0	0

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
72	**Brewers yeast/vitamin B as an insect repellant**		[1	1]
S1A	((brewers yeast) or (vitamin b)) and (insect repellants)	0		
S1B	((brewers yeast) or (vitamin b)) and insect and repellant	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	(dried-yeast or vitamin-b-substances£) with use with insect-repellant)	1	1	1
S3B	(dried-yeast or vitamin-b-substances£) with (insect-repellant or insecticide)	1	1	1
S4A	[vitamin b substances] or (dried yeast) and use and insecticide	0		
S4B	([vitamin b substances] or yeast) and insect?	1	1	1

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
73	**Increasing aqueous solubility of salicylic acid - which polyethylene glycol to be used (i.e. which one is best and at what percentage)?**		[1	0]
S1A	(salicylic acid) and solubility and (aqueous or water) and macrogol?	0		
S1B	as for S1A	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	(salicylic-acid with pharm with solubility) and macrogols	0		
S3B	macrogols with use with solubility with salicylic-acid	0		
S4A	(salicylic acid) and (macrogols or (macrogol ethers)) and (solubility or solvent)	1	1	0
S4B	macrogols and (salicylic acid) and solubility	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
74	**Dosage of piroxicam in the elderly**		[0	0]
S1A	piroxicam and geriatrics and dosage	0		
S1B	piroxicam and (dose or dosage) and (elderly or old or geriatric?)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	piroxicam with ad with old-people	0		
S3B	as for S3A	0		
S4A	piroxicam and (dose or dosage or admin?) and ((old people) or age)	0		
S4B	piroxicam and (old or elderly) and (dose or dosage or admin?)	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
75	**Interaction between Bronztan tablets and warfarin**		[0	0]
S1A	warfarin and bronztan and (drug interaction)	0		
S1B	bronztan and warfarin and interaction	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	warfarin with int with sunscreen-agents£	0		
S3B	(warfarin with int) and bronztan	0		
S4A	[anticoagulants] and interaction and (bronztan or [sunscreen agents])	0		
S4B	warfarin and interaction? and bronztan	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
76	**Desirability of giving high doses (100 mg or greater) of sodium perchlorate injection to block the thyroid before technetium brain scan.**		[1	1]
S1A	(sodiium perchlorate) and thyroid and technecium and (brain scan)	0		
S1B	(sodium perchlorate) and technetium and brain and scan	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	sodium-perchlorate with ad with thyroid-gland	0		
S3B	sodium-perchlorate with ad with parenteral-route£ with (encephalography£ or radionuclide-imaging£)	0		
S4A	(sodium perchlorate) and (dose or dosage or admin?) and [antithyroid agents]	0		
S4B	as for S4A	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
77	**Glyceryl trinitrate adsorption onto polypropylene syringes - does it occur?**		[1	1]
S1A	(glyceryl trinitrate) and polypropylene and adsorption	0		
S1B	(glyceryl trinitrate) and plastic and adsorb?	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	glyceryl-trinitrate with plastics with pharm	0		
S3B	glyceryl-trinitrate with plastics with sorption	0		
S4A	(glyceryl trinitrate) and sorption and (plastics or polypropylene)	0		
S4B	(glyceryl trinitrate) and (sorption or adsorption or absorption) and (plastic? or polypropylene) and syringe?	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
78	**Strength of saltpetre to use as a preservative?**		[4	1]
S1A	saltpetre and preservatives and food	0		
S1B	saltpetre and preservative	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	potassium-nitrate with use with food-preservation	2	2	0
S3B	potassium-nitrate with use with preservative	2	2	0
S4A	(potassium nitrate) and use and (food preservation)	0		
S4B	(potassium nitrate) and preservative	3	3	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
79	**Amyl nitrite in cyanide poisoning, is it injected or inhaled?**		[3	1]
S1A	(amyl nitrite) and cyanide and poisoning and (inhalation or injection) not abuse	3	3	1
S1B	(amyl nitrite) and cyanide and (inject? or inhal?)	3	3	1
S2A	as for S1A	3	3	1
S2B	as for S1B	3	3	1
S3A	amyl-nitrite with ad with cyanide-poisoning	0		
S3B	amyl-nitrite with ad with tae with cyanide-poisoning	0		
S4A	(amyl nitrite) and use and ((cyanide poisoning) or cyanide)	0		
S4B	(amyl nitrite) and cyanide and treatment and adverse	1	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
80	**Use of papaine in mouth ulcers**		[0	0]
S1A	papain and ((mouth ulcers) or gums or throat)	0		
S1B	papain and (mouth or oral) and ulcer?	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	papain with use with aphthous-stomatitis	0		
S3B	as for S3A	0		
S4A	papain and use and ((oral ulceration) or ulcer)	0		
S4B	papain and ((mouth ulcer) or (aphthous stomatitis))	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
81	**Use of hydrogen peroxide ear drops**		[2	1]
S1A	(hydrogen peroxide) and ears	0		
S1B	(hydrogen peroxide) and ((ear drops) or aural) and (use? or treat?)	1	1	1
S2A	as for S1A	0		
S2B	as for S1B	2	2	1
S3A	hydrogen-peroxide with use with ear-drops	2	2	1
S3B	as for S3A	2	2	1
S4A	(hydrogen peroxide) and use and (ear-drops or instillation)	1	1	1
S4B	(hydrogen peroxide) and (use? or treat?) and ear-drops	1	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
82	**Dosing of phenazopyridine - before or after meals?**		[1	1]
S1A	phenazopyridine and (drug administration) and food	0		
S1B	phenazopyridine and (dosage or dose or administrat?) and (food or meals)	1	1	1
S2A	as for S1A	0		
S2B	as for S1B	1	1	1
S3A	phenazopyridine with ad with (after-food or before-food)	1	1	1
S3B	as for S3A	1	1	1
S4A	phenazopyridine and ((after food) or (before food)) and (dose or dosage or administrat?)	1	1	1
S4B	phenazopyridine and (dose or dosage or administrat?) and (food or meal? or (absorptive state) or (fasting state))	1	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
83	**Lomotil liquid - is it sugar free?**		[0	0]
S1A	(lomotil or (diphenoxylate hydrochloride)) and (sugar or (dental caries))	0		
S1B	lomotil and (syrup or liquid) and (sugar or glucose or dextrose)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	lomotil and (prop with pharm)	0		
S3B	lomotil and (sugar or dextrose or glucose) and prop	0		
S4A	diphenoxylate and ((pharmaceutical manufacturing) or diluting)	0		
S4B	as for S4A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
84	**What are the normal blood cholesterol levels?**		[0	0]
S1A	cholesterol and ((blood levels) or (body fluids))	0		
S1B	cholesterol and (blood or plasma or serum) and normal	3	0	0
S2A	as for S1A	0		
S2B	as for S1B	4	0	0
S3A	blood-and-lymphatic-system-physiology with cholesterol	0		
S3B	cholesterol and concentration and blood	7	0	0
S4A	cholesterol and ([blood and lymphatic system physiology] or blood)	4	0	0
S4B	cholesterol and concentration and (blood or plasma or serum)	15	0	0

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
85	**Use of water soluble vitamin k in malabsorption**		[2	1]
S1A	(vitamin k) and (water solubility) and malabsorption	0		
S1B	(vitamin k) and soluble and water and malabsorption	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	(vitamin-k-substances£ and water-soluble-vitamins£) with use with malabsorption	0		
S3B	(vitamin-k-substances£ and soluble and water) with use with malabsorption	0		
S4A	((vitamin k) or acetomenaphthone or menadiol or menadione or menaquinone or menatetrenone or phytamenadione or (vitamin k deficiency)) and (dose or dosage or administr?) and malabsorption	0		
S4B	([vitamin k] or (vitamin k deficiency)) and (use or treat?) and malabsorption	2	2	1

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
86	**Potential of castor oil as an abortifacient**		[0	0]
S1A	(castor oil) and (abortion or miscarriage)	0		
S1B	(castor oil) and abort?	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	castor-oil with use with abortion	0		
S3B	as for S3A	0		
S4A	((castor oil) or ricin) and ((abortion induction) or (uterine stimulant))	0		
S4B	((castor oil) or ricin) and (abortion or emmenagogue or uterine)	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
87	**Parenteral formulation of thiazide diuretic**		[5	1]
S1A	(thiazide diuretics) and (injection or (intravenous injection))	3	1	0
S1B	thiazide and (parenteral or injection) and (formula or preparation)	0		
S2A	as for S1A	4	1	0
S2B	as for S1B	0		
S3A	thiazide-diuretics£ with (prop or nprop) with parenteral-dosage -form£	0		
S3B	thiazide-diuretics£ with (parenteral-dosage-form£ or parenteral-route) with (prop or nprop)	0		
S4A	[thiazide diuretics] and ([parenteral dosage form] or [parenteral route])	0		
S4B	[thiazide diuretics] and ([parenteral dosage form] or [parenteral route] or infusion or injection)	18	1	1

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
88	**Use of benzaldehyde solution in cancer**		[1	1]
S1A	(cancer or antineoplastics) and benzaldehyde	1	1	1
S1B	benzaldehyde and (cancer or neoplasm? or malignan?)	1	1	1
S2A	as for S1A	1	1	1
S2B	as for S1B	1	1	1
S3A	benzaldehyde with use with (antineoplastic or malignant-neoplasm£)	0		
S3B	as for S3A	0		
S4A	benzaldehyde and ([antineoplastics] or antineoplastic)	0		
S4A	benzaldehyde and [malignant neoplasms] and (use? or treat?)	1	1	1

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
89	**Auralgen drops - what quantity used in ear?**		[0	0]
S1A	auralgen and dosage and eardrops	0		
S1B	auralgin and drops and ear and (drops or dosage)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	auralgen and (ad with eardrops)	0		
S3B	as for S3A	0		
S4A	auralgen and (dose or dosage or admin?) and (ear-drops or instillation)	0		
S4B	as for S3A	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
90	**Alupent syrup, is it sugar-free?**		[0	0]
S1A	(orciprenaline or alupent) and (sugar or (dental caries))	0		
S1B	alupent and (syrup or liquid) and (sugar or glucose or dextrose)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	alupent and syrup and (prop with pharm)	0		
S3B	alupent and (sugar or dextrose or glucose) and prop	1	0	0
S4A	orciprenaline and ((pharmaceutical manufacturing) or diluting)	0		
S4B	as for S1B	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
91	**Is citric soda the same as potassium citrate?**		[1	0]
S1A	(citric soda) and (potassium citrate)	0		
S1B	as for S1A	0		
S2A	as for S1A	0		
S2B	as for S1A	0		
S3A	(citric soda) and comp with (potassium citrate)	0		
S3B	(citric soda) or (potassium-citrate with des)	1	1	0
S4A	as for S1A	0		
S4B	((citric soda) or (potassium citrate)) and definition	3	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
92	**In which infusion fluids is ampicillin unstable?**		[8	3]
S1A	ampicillin and stability and (intravenous infusions)	0		
S1B	ampicillin and stability and (infusion or injection or solution)	6	3	1
S2A	as for S1A	0		
S2B	as for S1B	27	5	3
S3A	(ampicillin with pharm with stability) and infusion	1	1	1
S3B	ampicillin with (stability or incompatibility) with (parenteral-dosage-form or parenteral-route)	0		
S4A	ampicillin and (stability or deterioration or incompatibility) and infusion	1	1	1
S4B	as for S4A	1	1	1

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
93	**Is there an interaction between pyridostigmine and phenylamine?**		[0	0]
S1A	pyridostigmine and phenylamine and (drug interactions)	0		
S1B	pyridostigmine and aniline and interact?	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	pyridostigmine with int with phenylamine	0		
S3B	pyridostigmine with int with aniline	0		
S4A	(pyridostigmine or [anticholinesterase parasympathomimetics]) and phenylaniline	0		
S4B	as for S1B	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
94	**Is lactulose safe in pregnancy?**		[0	0]
S1A	lactulose and (pregnancy or teratogenicity)	0		
S1B	lactulose and pregnan? and (use or precautions or safe?)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	lactulose with ae with (teratogen or pregnancy)	0		
S3B	lactulose with ae with (preg or teratogen)	0		
S4A	lactulose and (pregnancy or congenital or foetus)	0		
S4B	lactulose and (pregnan? or teratogen? or congenital)	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
95	**What is the dose for aspirin and papaveretum tablets?**		[0	0]
S1A	aspirin and papaveretum and tablets and dosage	0		
S1B	aspirin and papaveretum and tablet? and (dose or dosage)	0		
S2A	as for S1A	0		
S2B	as for S1B	0		
S3A	aspirin with papaveretum with ad	0		
S3B	as for S3A	0		
S4A	aspirin and papaveretum and tablets and (dose or dosage or admin?)	0		
S4B	apirin and papaveretum and (dose or dosage or admin?)	0		

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
96	**Can you preserve hydrogen peroxide solution, if so with what?**		[5	2]
S1A	(hydrogen peroxide) and stability	2	0	0
S1B	(hydrogen peroxide) and solution and preserv?	2	1	1
S2A	as for S1A	13	4	1
S2B	as for S1B	2	1	1
S3A	hydrogen-peroxide with pharm with stability	3	3	1
S3B	hydrogen-peroxide with pharm with (preservation or antoxidation or stabilisation)	4	4	1
S4A	(hydrogen peroxide) and (stability or stabilisation) and [preservatives]	0		
S4B	as for S4A	0		

* Search terms enclosed within [ ] are descriptors for which the equivalent of inclusive searching was requested.

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
97	**Are there any official preparations containing rose or extract of rose?**		[11	9]
S1A	rose and (extract of rose) (or rose extract) and formulation	0		
S1B	(rose or (rose and extract)) and preparation?	30	11	9
S2A	as for S1A	0		
S2B	as for S1B	30	11	9
S3A	rose with nprop	0		
S3B	(rose-oil or red-rose-petal) with nprop	7	7	7
S4A	(rose or (rose fruit)) and (pharmaceutical manufacturing)	0		
S4B	rose and preparation	30	11	9

S	QUERY AND STATEMENTS	HITS	TOTAL RELEVANT	MAJOR RELEVANT
98	**Adverse effects of the morning-after pill**		[3	3]
S1A	(postcoital contraception) and (adverse drug reaction)	0		
S1B	contraceptive and (postcoital or (morning after)) and (adverse or side? or toxic)	3	3	3
S2A	as for S1A	0		
S2B	as for S1B	3	3	3
S3A	postcoital-contraception with ae	0		
S3B	as for S3A	0		
S4A	((postcoital contraception) or (female contraception)) and (adverse or toxic or side-effect?)	3	3	3
S4B	(postcoital contraception?) and (adverse or toxic or side-effect?)	3	3	3

APPENDIX H

ROUGH NOTES BY RELEVANCE ASSESSOR

NOTES ON X, MAJ, MIN, ONE OF WHICH WAS ASSIGNED TO EACH RECORD

X- Used when there was a conjunction of the ideas contained in the question but the match provided no useful information. I may differ from [redacted] assessment in this. [[redacted] was the information pharmacist who attempted some relevance assessment on the responses to [redacted] queries.]

Where the same information appeared twice, sometimes in 2 different ways, as in chapter introduction and then in monograph, one record should have been marked with an X. Do you agree with this? I hope I was consistent.

MIN- Was often used where the record retrieved might lead to solid information. For example, a record stating that a WHO Committee had reported on a topic, but omitting detail. If you read the WHO Report, it might contain the precise data wanted.
Criterion applied similarly (may not be apparently consistently) to 'Further references.....'
Also used for near-misses if the enquirer might rather be told those facts than remain ignorant of them.

MAJ- Used, as well as for bullseyes, for
(a) a few pertinent words in a long text field;
(b) useful background information as in the case of absorption information underpinning a precaution statement;
(c) where !309 was invoked in a major drug text record (but this needs qualifying). [!309 = the field that contains the statement that a number of drugs have been cross-referred to this record as its contents can be considered to apply to them.]

I erred, knowingly (and being me) on the side of condemnation rather than on the side of ready approval.

APPENDIX I

SCORES FOR EACH SEARCH AND CALCULATIONS FOR EACH COMPARISON

1AT Structured Searching of the Indexed File Versus Unstructured Searching of the Unindexed File (S3AT v S1AT)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AT	S1AT		GROUPI	TOTAL
1	0.601	0.379	222	8	12
2	0.791	0.408	383	15	30
3	0.488	0.577	-89	-4	-5
4	0.417	0.306	111	5	7
5	1	0.129	871	20	43
6	1	1	0	-	-
7	1	1	0	-	-
8	0.707	0.577	130	6	8
9	1	0.378	622	19	42
10	0.302	0.629	-327	-14	-28
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.707	293	12	21.5
15	0.771	0.378	393	16	31
16	0.707	0.267	440	17	36.5
17	0.354	0.500	-146	-7	-10
18	0.577	0.500	77	3	4
19	0.535	0.286	249	9	16
20	1	1	0	-	-
21	1	0.707	293	12	21.5
22	0.632	0.600	32	1	1
23	0.477	0.400	47	2	2
24	1	0.088	912	21	44
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-

28	0.707	0.707	0	-	-
29	0.612	0.875	-263	-10	-17
30	0.707	1	-293	-12	-21.5
31	1	1	0	-	-
32	0.707	0.250	457	18	38
33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AT	S1AT		GROUPII	TOTAL
35	0.816	0.707	109	1	6
36	1	1	0	-	-
37	0.174	0.174	0	-	-
38	1	1	0	-	-
39	1	0.707	293	6.5	21.5
40	0.401	0.177	224	3	13
41	0.822	0.516	306	8	27
42	0.500	0.707	-207	-2	-11
43	0.447	0.447	0	-	-
44	0.816	0.577	239	4.5	14.5
45	0.267	0.707	-440	-12	-36.5
46	1	1	0	-	-
47	1	1	0	-	-
48	1	1	0	-	-
49	0.500	0.500	0	-	-
50	1	0.577	423	9.5	33
51	1	1	0	-	-
52	1	1	0	-	-
53	0.408	0.408	0	-	-
54	0.707	0.707	0	-	-
55	0.707	0.101	606	13	41
56	1	1	0	-	-
57	0.816	0.577	239	4.5	14.5
58	1	1	0	-	-

59	0.707	0.707	0	-	-
60	0.816	0.816	0	-	-
61	0.707	0.707	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	0.707	293	6.5	21.5
65	0.935	0.500	435	11	35
66	1	0.577	423	9.5	33

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AT	S1AT		GROUPIII	TOTAL
67	0.577	0.577	0	-	-
68	0.603	0.302	301	6	26
69	1	1	0	-	-
70	0.250	0.250	0	-	-
71	0.426	0.500	-74	-1	-3
72	1	0.707	293	4	21.5
73	0.707	0.707	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.775	0.447	328	7	29
79	0.500	1	-500	-9	-39
80	1	1	0	-	-
81	1	0.577	423	8	33
82	1	0.707	293	4	21.5
83	1	1	0	-	-
84	1	1	0	-	-
85	0.577	0.577	0	-	-
86	1	1	0	-	-
87	0.408	0.408	0	-	-
88	0.707	1	-293	-4	-21.5
89	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AT	S1AT		GROUPIII	TOTAL
90	1	1	0	-	-
91	0.707	0.707	0	-	-
92	0.471	0.333	138	2	9
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.816	0.236	580	10	40
97	0.289	0.289	0	-	-
98	0.500	0.500	0	0	-

CALCULATIONS FOR S3AT v S1AT

For All Searches

$n = 44$

$n(n+1)/2 = 990$

$\Sigma ranks = 990$

$T = \Sigma$-ve ranks

$= 192.5$

$\mu = n(n+1)/4$

$= 495$

$$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

$= 85.688$

$$u = \frac{T-\mu}{\sigma}$$

$= 3.530$

$P < 0.001$

For Each of the Three Groups

Group I	Group II	Group III
$n = 21$	$n = 13$	$n = 10$
$\Sigma ranks = 231$	$\Sigma ranks = 91$	$\Sigma ranks = 55$
Σ-ve ranks = 47	Σ-ve ranks = 14	Σ-ve ranks = 14
$0.01<P<0.02$	$0.02<P<0.05$	$P = 0.2$

1BT Structured Searching of the Indexed File Versus Unstructured Searching of the Unindexed File (S3BT v S1BT)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BT	S1BT		GROUPI	TOTAL
1	0.677	0.311	366	15	41
2	0.791	0.671	120	6	10
3	0.488	0.586	-98	-4	-8
4	0.417	0.139	278	11	24
5	1	0.707	293	12.5	30
6	1	1	0	-	-
7	1	1	0	-	-
8	0.707	0.707	0	-	-
9	1	0.316	684	17	49
10	0.302	0.426	-124	-7	-11
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.816	134	8	12
15	0.771	0.378	393	16	42
16	0.707	0.707	0	-	-
17	0.707	0.354	353	14	39.5
18	0.577	0.500	77	3	6.5
19	0.535	0.267	268	10	23
20	1	1	0	-	-
21	1	1	0	-	-
22	0.671	0.600	71	2	5
23	0.447	0.400	47	1	3.5
24	1	0.124	876	18	51
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.875	-263	-9	-22
30	0.707	1	-293	-12.5	-30
31	1	1	0	-	-
32	0.707	0.816	-109	-5	-9

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BT	S1BT		GROUPII	TOTAL
35	0.816	1	-184	-5	-17
36	1	0.408	592	16	48
37	0.395	0.348	47	2	3.5
38	1	1	0	-	-
39	0.707	0.707	0	-	-
40	0.471	0.168	303	10	35
41	0.764	0.433	331	12	38
42	0.500	0.577	-77	-3	-6.5
43	0.447	0.477	0	-	-
44	0.816	0.666	150	4	14
45	1	0.707	293	8	30
46	1	1	0	-	-
47	1	1	0	-	-
48	1	0.707	293	8	30
49	0.707	0.354	353	13	39.5
50	1	0.577	423	14	43
51	1	0.707	293	8	30
52	1	1	0	-	-
53	0.408	0.408	0	-	-
54	1	1	0	-	-
55	0.632	0.129	503	15	47
56	1	1	0	-	-
57	0.816	0.577	239	6	20
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.816	0.816	0	-	-
61	0.408	0.447	-39	-1	-1.5
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.935	0.612	323	11	36
66	1	0.200	800	17	50

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BT	S1BT		GROUPIII	TOTAL
67	1	0.816	184	4.5	17
68	0.957	0.701	256	7	21
69	1	1	0	-	-
70	0.250	0.289	-39	-1	-1.5
71	0.426	0.707	-281	-8	-25
72	1	0.707	293	10.5	30
73	0.707	0.707	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.775	0.447	328	13	37
79	0.500	1	-500	-15.5	-45.5
80	1	1	0	-	-
81	1	0.816	184	4.5	17
82	1	1	0	-	-
83	1	1	0	-	-
84	0.354	0.500	-146	-2	-13
85	0.577	0.577	0	-	-
86	1	1	0	-	-
87	0.408	0.408	0	-	-
88	0.707	1	-293	-10.5	-30
89	1	1	0	-	-
90	0.707	1	-293	-10.5	-30
91	1	0.707	293	10.5	30
92	0.333	0.504	-171	-3	-15
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.913	0.471	442	14	44
97	0.816	0.622	194	6	19
98	0.500	1	-500	-15.5	-45.5

CALCULATIONS FOR S3BT v S1BT

For All Searches

$n = 51$

$n(n+1)/2 = 1326$

$\Sigma ranks = 1326$

$T = \Sigma$-ve ranks

$= 310.5$

$\mu = n(n+1)/4$

$= 663$

$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

$= 106.684$

$u = \frac{T-\mu}{\sigma}$

$= 3.304$

$P < 0.001$

For Each of the Three Groups

Group I	Group II	Group III
$n = 18$	$n = 17$	$n = 16$
$\Sigma ranks = 171$	$\Sigma ranks = 153$	$\Sigma ranks = 136$
Σ-ve ranks = 37.5	Σ-ve ranks = 9	Σ-ve ranks = 66
$0.02 < P < 0.05$	$P < 0.001$	$P > 0.4$

1AM Structured Searching of the Indexed File Versus Unstructured Searching of the Unindexed File (S3AM v S1AM)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AM	S1AM		GROUPI	TOTAL
1	0.514	0.263	251	7	14
2	0.447	0.192	255	8	15
3	0.258	0.218	40	2	2
4	0.417	0.306	111	5	7
5	0.707	0.183	524	17	38
6	1	1	0	-	-
7	1	1	0	-	-
8	1	0.408	592	18.5	39.5
9	1	0.378	622	20.5	42
10	0.408	0.511	-103	-4	-4
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.707	293	12	22
15	0.408	1	-592	-18.5	-39.5
16	1	0.378	622	20.5	42
17	0.354	0.500	-146	-6	-8
18	0.500	0.433	67	3	3
19	0.577	0.309	268	10	17
20	1	1	0	-	-
21	1	0.707	293	12	22
22	0.632	0.600	32	1	1
23	0.707	0.316	391	14	30
24	0.620	0.141	479	16	37
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.875	-263	-9	-16
30	0.707	1	-293	-12	-22
31	1	1	0	-	-
32	0.707	0.250	457	15	35

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AM	S1AM		GROUPII	TOTAL
35	0.816	0.707	109	1	5.5
36	1	1	0	-	-
37	0.189	0.189	0	-	-
38	1	1	0	-	-
39	0.707	1	-293	-6	-22
40	0.386	0.204	182	3	10
41	0.783	0.316	467	11	36
42	0.577	0.408	169	2	9
43	0.447	0.447	0	-	-
44	0.816	0.577	239	4	12
45	0.378	1	-622	-12	-42
46	1	1	0	-	-
47	1	1	0	-	-
48	1	1	0	-	-
49	0.707	0.707	0	-	-
50	1	0.577	423	9.5	33.5
51	1	1	0	-	-
52	1	1	0	-	-
53	0.707	0.707	0	-	-
54	0.707	0.707	0	-	-
55	1	0.143	857	13	44
56	1	1	0	-	-
57	0.707	1	-293	-6	-22
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.707	0.707	0	-	-
61	0.707	0.707	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	0.707	293	6	22
65	0.756	0.354	402	8	31
66	1	0.577	423	9.5	33.5

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AM	S1AM		GROUPIII	TOTAL
67	0.577	0.577	0	-	-
68	0.632	0.316	316	9	29
69	1	1	0	-	-
70	0.302	0.302	0	-	-
71	0.302	0.707	-405	-10	-32
72	1	0.707	293	5	22
73	1	1	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.408	0.707	-299	-7.5	-27.5
79	0.707	0.707	0	-	-
80	1	1	0	-	-
81	0.816	0.707	109	1	5.5
82	1	0.707	293	5	22
83	1	1	0	-	-
84	1	1	0	-	-
85	0.707	0.707	0	-	-
86	1	1	0	-	-
87	0.707	0.408	299	7.5	27.5
88	0.707	1	-293	-5	-22
89	1	1	0	-	-
90	1	1	0	-	-
91	1	1	0	-	-
92	0.707	0.500	207	2	11
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.577	0.333	244	3	13
97	0.316	0.316	0	-	-
98	0.500	0.500	0	-	-

CALCULATIONS FOR S3AM v S1AM

For All Searches

$n = 44$

$n(n+1)/2 = 990$

$\Sigma ranks = 990$

$T = \Sigma$-ve ranks

$= 257$

$\mu = n(n+1)/4$

$= 495$

$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

$= 85.688$

$u = \frac{T-\mu}{\sigma}$

$= 2.778$

$0.001 < P < 0.01$

For Each of the Three Groups

Group I	Group II	Group III
$n = 21$	$n = 13$	$n = 10$
$\Sigma ranks = 231$	$\Sigma ranks = 91$	$\Sigma ranks = 55$
Σ-ve ranks = 49.5	Σ-ve ranks = 24	Σ-ve ranks = 22.5
$0.02 < P < 0.05$	$0.1 < P < 0.2$	$P > 0.4$

1BM Structured Searching of the Indexed File Versus Unstructured Searching of the Unindexed File (S3BM v S1BM)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BM	S1BM		GROUPI	TOTAL
1	0.705	0.252	453	15	42
2	0.447	0.316	131	5	10
3	0.258	0.258	0	-	-
4	0.417	0.139	278	9	25
5	0.707	1	-293	-11.5	-31
6	1	1	0	-	-
7	1	1	0	-	-
8	1	1	0	-	-
9	1	0.316	684	17	47
10	0.408	0.577	-169	-7	-15
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.866	134	6	11
15	0.408	1	-592	-16	-45.5
16	1	1	0	-	-
17	0.707	0.354	343	13	38
18	0.500	0.433	67	1	4
19	0.577	0.289	288	10	26
20	1	1	0	-	-
21	1	1	0	-	-
22	0.671	0.600	71	2	5
23	0.707	0.632	75	3	6
24	0.620	0.200	420	14	40
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.875	-263	-8	-22
30	0.707	1	-293	-11.5	-31
31	1	1	0	-	-
32	0.707	0.816	-109	-4	-7

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BM	S1BM		GROUPII	TOTAL
35	0.816	1	-184	-5	-17
36	1	0.408	592	16	45.5
37	0.286	0.252	34	1	2
38	1	1	0	-	-
39	1	1	0	-	-
40	0.476	0.065	411	13	39
41	0.802	0.530	272	9	24
42	0.577	0.333	244	7	20
44	0.447	0.447	0	-	-
44	0.816	0.666	150	4	13.5
45	0.707	1	-293	-11	-31
46	1	1	0	-	-
47	1	1	0	-	-
48	1	0.707	293	11	31
49	1	0.500	500	15	43.5
50	1	0.577	423	14	41
51	1	0.707	293	11	31
52	1	1	0	-	-
53	0.707	0.707	0	-	-
54	1	1	0	-	-
55	0.447	0.183	264	8	23
56	1	1	0	-	-
57	0.707	0.500	207	6	19
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.707	0.707	0	-	-
61	0.408	0.447	-39	-2	-3
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.756	0.866	-110	-3	-8
66	1	0.200	800	17	49

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BM	S1BM		GROUPIII	TOTAL
67	1	0.816	184	5.5	17
68	0.913	0.653	260	7	21
69	1	1	0	-	-
70	0.201	0.174	27	1	1
71	0.302	1	-698	-15	-48
72	1	0.707	293	9.5	31
73	1	1	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.408	0.707	-299	-12	-36
79	0.707	0.707	0	-	-
80	1	1	0	-	-
81	0.816	1	-184	-5.5	-17
82	1	1	0	-	-
83	1	1	0	-	-
84	0.354	0.500	-146	-3	-12
85	0.707	0.707	0	-	-
86	1	1	0	-	-
87	0.707	0.707	0	-	-
88	0.707	1	-293	-9.5	-31
89	1	1	0	-	-
90	0.707	1	-293	-9.5	-31
91	0.707	1	-293	-9.5	-31
92	0.500	0.378	122	2	9
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.516	0.666	-150	-4	-13.5
97	0.894	0.568	326	13	37
98	0.500	1	-500	-14	-43.5

CALCULATIONS FOR S3BM v S1BM

For All Searches

$n = 49$

$n(n+1)/2 = 1225$

$T = \Sigma$-ve ranks

$= 473.5$

$\mu = n(n+1)/4$

$= 612.5$

$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

$= 100.530$

$u = \frac{T-\mu}{\sigma}$

$= 1.383$

$0.1 < P < 0.2$

For Each of the Three Groups

Group I	Group II	Group III
n = 17	n = 17	n = 15
Σranks = 153	Σranks = 153	Σranks = 120
Σ-ve ranks = 58	Σ-ve ranks = 21	Σ+ve ranks = 38
$P > 0.4$	$0.002 < P < 0.01$	$0.2 < P < 0.4$

2AT Structured Searching of the Indexed File Versus Unstructured Searching of the Indexed File (S3AT v S2AT)

QUERY	SCORES θ		DIFF	RANK GROUPI	RANK TOTAL
	S3AT	S2AT			
1	0.601	0.379	222	8	14
2	0.791	0.408	383	15	32
3	0.488	0.552	-64	-2	-3
4	0.417	0.283	124	5	8
5	1	0.129	871	21	45
6	1	1	0	-	-
7	1	1	0	-	-
8	0.707	0.500	207	7	12.5
9	1	0.354	646	19	43
10	0.302	0.677	-375	-14	-31
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.707	293	12	24.5
15	0.771	0.378	393	16	33
16	0.707	0.250	457	17.5	39.5
17	0.354	0.600	-246	-10	-19
18	0.577	0.426	151	6	10
19	0.535	0.446	89	3	5
20	1	1	0	-	-
21	1	0.707	293	12	24.5
22	0.632	0.730	-98	-4	-6
23	0.477	0.400	47	1	2
24	1	0.084	916	22	46
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.853	-241	-9	-18
30	0.707	1	-293	-12	-24.5
31	1	1	0	-	-
32	0.707	0.250	457	17.5	39.5

33	1	0.333	666	20	44
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AT	S2AT		GROUPII	TOTAL
35	0.816	0.707	109	1	7
36	1	1	0	-	-
37	0.174	0.174	0	-	-
38	1	1	0	-	-
39	1	0.707	293	7.5	24.5
40	0.401	0.177	224	4	15
41	0.822	0.655	167	2	11
42	0.500	0.707	-207	-3	-12.5
43	0.447	0.447	0	-	-
44	0.816	0.577	239	5.5	16.5
45	0.267	0.707	-440	-12	-38
46	1	1	0	-	-
47	1	1	0	-	-
48	1	1	0	-	-
49	0.500	0.500	0	-	-
50	1	0.577	423	9.5	35
51	1	1	0	-	-
52	1	1	0	-	-
53	0.408	0.408	0	-	-
54	0.707	0.707	0	-	-
55	0.707	0.099	608	13	42
56	1	1	0	-	-
57	0.816	0.577	239	5.5	16.5
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.816	0.816	0	-	-
61	0.707	0.707	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	0.707	293	7.5	24.5
65	0.935	0.500	435	11	37
66	1	0.577	423	9.5	35

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AT	S2AT		GROUPIII	TOTAL
67	0.577	0.577	0	-	-
68	0.603	0.302	301	8	29
69	1	1	0	-	-
70	0.250	0.250	0	-	-
71	0.426	0.500	-74	-2	-4
72	1	0.707	293	6	24.5
73	0.707	0.707	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.775	0.447	328	9	30
79	0.500	1	-500	-11	-41
80	1	1	0	-	-
81	1	0.577	423	10	35
82	1	0.707	293	6	24.5
83	1	1	0	-	-
84	1	1	0	-	-
85	0.577	0.577	0	-	-
86	1	1	0	-	-
87	0.408	0.365	43	1	1
88	0.707	1	-293	-6	-24.5
89	1	1	0	-	-
90	1	1	0	-	-
91	0.707	0.707	0	-	-
92	0.471	0.333	138	3	9
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.816	0.546	270	4	20
97	0.289	0.289	0	-	-
98	0.500	0.500	0	0	-

CALCULATIONS FOR S3AT v S2AT

For All Searches

$n = 46$

$n(n+1)/2 = 1081$

$\Sigma ranks = 1081$

$T = \Sigma$-ve ranks

$= 221.5$

$\mu = n(n+1)/4$

$= 540.5$

$$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

$= 91.530$

$$u = \frac{T-\mu}{\sigma}$$

$= 3.485$

$P < 0.001$

For Each of the Three Groups

Group I	Group II	Group III
$n = 22$	$n = 13$	$n = 11$
$\Sigma ranks = 253$	$\Sigma ranks = 91$	$\Sigma ranks = 66$
Σ-ve ranks = 51	Σ-ve ranks = 15	Σ-ve ranks = 19
$0.01<P<0.02$	$0.02<P<0.05$	$0.2<P<0.4$

2BT Structured Searching of the Indexed File Versus Unstructured Searching of the Indexed File (S3BT v S2BT)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BT	S2BT		GROUPI	TOTAL
1	0.677	0.548	129	7	15
2	0.791	0.671	120	5	13
3	0.488	0.577	-89	-3	-9
4	0.417	0.133	284	12	27
5	1	0.707	293	13.5	31.5
6	1	1	0	-	-
7	1	1	0	-	-
8	0.707	0.707	0	-	-
9	1	0.316	684	17	47
10	0.302	0.426	-124	-6	-14
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.816	134	8	17
15	0.771	0.378	393	16	39
16	0.707	0.707	0	-	-
17	0.707	0.500	207	10	21
18	0.577	0.426	151	9	18
19	0.535	0.189	346	15	38
20	1	1	0	-	-
21	1	1	0	-	-
22	0.671	0.730	-59	-2	-7
23	0.447	0.400	47	1	4
24	1	0.105	895	18	49
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.853	-241	-11	-24
30	0.707	1	-293	-13.5	-31.5
31	1	1	0	-	-
32	0.707	0.816	-109	-4	-11.5

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BT	S2BT		GROUPII	TOTAL
35	0.816	0.707	109	5	11.5
36	1	0.408	592	15	46
37	0.395	0.386	9	1	1
38	1	1	0	-	-
39	0.707	0.707	0	-	-
40	0.471	0.236	235	7	23
41	0.764	0.516	248	8	25
42	0.500	0.577	-77	-4	-8
43	0.447	0.477	0	-	-
44	0.816	0.866	-50	-3	-6
45	1	0.707	293	9.5	31.5
46	1	1	0	-	-
47	1	1	0	-	-
48	1	0.500	500	13	43
49	0.707	0.577	130	6	16
50	1	0.577	423	12	40
51	1	0.707	293	9.5	31.5
52	1	1	0	-	-
53	0.408	0.408	0	-	-
54	1	1	0	-	-
55	0.632	0.129	503	14	45
56	1	1	0	-	-
57	0.816	0.775	41	2	3
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.816	0.816	0	-	-
61	0.408	0.408	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.935	0.612	323	11	36
66	1	0.200	800	16	48

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BT	S2BT		GROUPIII	TOTAL
67	1	0.816	184	4	19
68	0.957	0.731	226	6	22
69	1	1	0	-	-
70	0.250	0.289	-39	-1	-2
71	0.426	0.707	-281	-7	-26
72	1	0.707	293	9.5	31.5
73	0.707	0.707	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.775	0.447	328	12	37
79	0.500	1	-500	-14.5	-43
80	1	1	0	-	-
81	1	1	0	-	-
82	1	1	0	-	-
83	1	1	0	-	-
84	0.354	0.447	-93	-3	-10
85	0.577	0.577	0	-	-
86	1	1	0	-	-
87	0.408	0.408	0	-	-
88	0.707	1	-293	-9.5	-31.5
89	1	1	0	-	-
90	0.707	1	-293	-9.5	-31.5
91	1	0.707	293	9.5	31.5
92	0.333	0.378	-45	-2	-4
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.913	0.471	442	13	41
97	0.816	0.622	194	5	20
98	0.500	1	-500	-14.5	-43

CALCULATIONS FOR S3BT v S2BT

For All Searches

$n = 49$

$n(n+1)/2 = 1225$

$\Sigma ranks = 1225$

$T = \Sigma$-ve ranks

$= 302$

$\mu = n(n+1)/4$

$= 612.5$

$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

$= 100.530$

$u = \frac{T-\mu}{\sigma}$

$= 3.089$

$0.001 < P < 0.01$

For Each of the Three Groups

Group I	Group II	Group III
$n = 18$	$n = 16$	$n = 15$
$\Sigma ranks = 171$	$\Sigma ranks = 136$	$\Sigma ranks = 120$
Σ-ve ranks = 39.5	Σ-ve ranks = 7	Σ-ve ranks = 61
$0.02 < P < 0.05$	$0.0002 < P < 0.001$	$P > 0.4$

2AM Structured Searching of the Indexed File Versus Unstructured Searching of the Indexed File (S3AM v S2AM)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AM	S2AM		GROUPI	TOTAL
1	0.514	0.251	278	9	15
2	0.447	0.192	255	10	16
3	0.258	0.209	49	1	1
4	0.417	0.283	134	5	7
5	0.707	0.183	524	17	38
6	1	1	0	-	-
7	1	1	0	-	-
8	1	0.354	646	20	42
9	1	0.354	646	20	42
10	0.408	0.500	-92	-2	-2
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.707	293	12	22
15	0.408	1	-592	-18	-39
16	1	0.354	646	20	42
17	0.354	0.600	-246	-8	-14
18	0.500	0.369	131	4	6
19	0.577	0.421	156	6	8
20	1	1	0	-	-
21	1	0.707	293	12	22
22	0.632	0.730	-98	-3	-3
23	0.707	0.316	391	14	30.5
24	0.620	0.135	485	16	37
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.853	-241	-7	-13
30	0.707	1	-293	-12	-22
31	1	1	0	-	-
32	0.707	0.250	457	15	36

33	1	0.333	666	22	44
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AM	S2AM		GROUPII	TOTAL
35	0.816	0.707	109	1	4.5
36	1	1	0	-	-
37	0.189	0.189	0	-	-
38	1	1	0	-	-
39	0.707	1	-293	-6	-22
40	0.386	0.204	182	3	10
41	0.783	0.401	382	8	29
42	0.577	0.408	169	2	9
43	0.447	0.447	0	-	-
44	0.816	0.577	239	4	12
45	0.378	1	-622	-12	-40
46	1	1	0	-	-
47	1	1	0	-	-
48	1	1	0	-	-
49	0.707	0.707	0	-	-
50	1	0.577	423	10.5	34.5
51	1	1	0	-	-
52	1	1	0	-	-
53	0.707	0.707	0	-	-
54	0.707	0.707	0	-	-
55	1	0.140	860	13	45
56	1	1	0	-	-
57	0.707	1	-293	-6	-22
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.707	0.707	0	-	-
61	0.707	0.707	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	0.707	293	6	22
65	0.756	0.354	402	9	32
66	1	0.577	423	10.5	34.5

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AM	S2AM		GROUPIII	TOTAL
67	0.577	0.577	0	-	-
68	0.632	0.316	316	8	28
69	1	1	0	-	-
70	0.302	0.302	0	-	-
71	0.302	0.707	-405	-10	-33
72	1	0.707	293	5	22
73	1	1	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.408	0.707	-299	-7	-27
79	0.707	0.707	0	-	-
80	1	1	0	-	-
81	0.816	0.707	109	1	4.5
82	1	0.707	293	5	22
83	1	1	0	-	-
84	1	1	0	-	-
85	0.707	0.707	0	-	-
86	1	1	0	-	-
87	0.707	0.316	391	9	30.5
88	0.707	1	-293	-5	-22
89	1	1	0	-	-
90	1	1	0	-	-
91	1	1	0	-	-
92	0.707	0.500	207	2	11
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.577	0.309	268	3	17
97	0.316	0.316	0	-	-
98	0.500	0.500	0	-	-

CALCULATIONS FOR S3AM v S2AM

For All Searches

$n = 45$

$n(n+1)/2 = 1035$

$T = \Sigma$-ve ranks

$= 259$

$\mu = n(n+1)/4$

$= 517.5$

$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

$= 88.593$

$u = \frac{T-\mu}{\sigma}$

$= 2.918$

$0.001 < P < 0.01$

For Each of the Three Groups

Group I	Group II	Group III
n = 22	n = 13	n = 10
Σranks = 253	Σranks = 91	Σranks = 55
Σ-ve ranks = 50	Σ-ve ranks = 24	Σ-ve ranks = 22
$0.02 < P < 0.05$	$0.1 < P < 0.2$	$P > 0.4$

2BM Structured Searching of the Indexed File Versus Unstructured Searching of the Indexed File (S3BM v S2BM)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BM	S2BM		GROUPI	TOTAL
1	0.705	0.470	235	10	21
2	0.447	0.316	131	5.5	12.5
3	0.258	0.218	40	1	3
4	0.417	0.133	284	12	26
5	0.707	1	-293	-13.5	-30.5
6	1	1	0	-	-
7	1	1	0	-	-
8	1	1	0	-	-
9	1	0.316	684	18	46
10	0.408	0.577	-169	-8	-16
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.866	134	7	14
15	0.408	1	-592	-17	-44
16	1	1	0	-	-
17	0.707	0.500	207	9	19
18	0.500	0.369	131	5.5	12.5
19	0.577	0.204	373	15	38
20	1	1	0	-	-
21	1	1	0	-	-
22	0.671	0.730	-59	-2	-5
23	0.707	0.632	75	3	6
24	0.620	0.169	451	16	40
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.853	-241	-11	-22
30	0.707	1	-293	-13.5	-30.5
31	1	1	0	-	-
32	0.707	0.816	-109	-4	-8.5

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BM	S2BM		GROUPII	TOTAL
35	0.816	0.707	109	3	8.5
36	1	0.408	592	14.5	44
37	0.286	0.314	-28	-1	-2
38	1	1	0	-	-
39	1	1	0	-	-
40	0.476	0.164	312	11	36
41	0.802	0.632	170	5	17
42	0.577	0.333	244	6	23
44	0.447	0.447	0	-	-
44	0.816	0.866	-50	-2	-4
45	0.707	1	-293	-9.5	-30.5
46	1	1	0	-	-
47	1	1	0	-	-
48	1	0.500	500	13	41.5
49	1	0.408	592	14.5	44
50	1	0.577	423	12	39
51	1	0.707	293	9.5	30.5
52	1	1	0	-	-
53	0.707	0.707	0	-	-
54	1	1	0	-	-
55	0.447	0.183	264	8	25
56	1	1	0	-	-
57	0.707	0.447	260	7	24
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.707	0.707	0	-	-
61	0.408	0.408	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.756	0.866	-110	-4	-10
66	1	0.200	800	16	48

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BM	S2BM		GROUPIII	TOTAL
67	1	0.816	184	5	18
68	0.913	0.690	223	6	20
69	1	1	0	-	-
70	0.201	0.174	27	1	1
71	0.302	1	-698	-14	-47
72	1	0.707	293	8.5	30.5
73	1	1	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.408	0.707	-299	-11	-35
79	0.707	0.707	0	-	-
80	1	1	0	-	-
81	0.816	0.816	0	-	-
82	1	1	0	-	-
83	1	1	0	-	-
84	0.354	0.447	-93	-2	-7
85	0.707	0.707	0	-	-
86	1	1	0	-	-
87	0.707	0.707	0	-	-
88	0.707	1	-293	-8.5	-30.5
89	1	1	0	-	-
90	0.707	1	-293	-8.5	-30.5
91	0.707	1	-293	-8.5	-30.5
92	0.500	0.378	122	3	11
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.516	0.666	-150	-4	-15
97	0.894	0.568	326	12	37
98	0.500	1	-500	-13	-41.5

CALCULATIONS FOR S3BM v S2BM

For All Searches

$n = 48$

$n(n+1)/2 = 1176$

$\Sigma ranks = 1176$

$T = \Sigma$-ve ranks

$= 440$

$\mu = n(n+1)/4$

$= 588$

$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

$= 97.499$

$u = \frac{T-\mu}{\sigma}$

$= 1.518$

$0.1 < P < 0.2$

For Each of the Three Groups

Group I	Group II	Group III
$n = 18$	$n = 16$	$n = 14$
$\Sigma ranks = 171$	$\Sigma ranks = 136$	$\Sigma ranks = 105$
Σ-ve ranks = 69	Σ-ve ranks = 16.5	Σ+ve ranks = 27
$P > 0.4$	$0.002 < P < 0.01$	$0.1 < P < 0.2$

3AT Unstructured Searching of the Indexed File Versus Unstructured Searching of the Unindexed File (S2AT v S1AT)

QUERY	SCORES θ		DIFF	RANK	RANK
	S2AT	S1AT		GROUPI	TOTAL
1	0.379	0.379	0	-	-
2	0.408	0.408	0	-	-
3	0.552	0.577	-25	-6	-7
4	0.283	0.306	-23	-4	-5
5	0.129	0.129	0	-	-
6	1	1	0	-	-
7	1	1	0	-	-
8	0.500	0.577	-77	-9	-11
9	0.354	0.378	-24	-5	-6
10	0.677	0.629	48	7	9
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	0.707	0.707	0	-	-
15	0.378	0.378	0	-	-
16	0.250	0.267	-17	-2	-3
17	0.600	0.500	100	10	12
18	0.426	0.500	-74	-8	-10
19	0.446	0.286	160	12	15
20	1	1	0	-	-
21	0.707	0.707	0	-	-
22	0.730	0.600	130	11	13
23	0.400	0.400	0	-	-
24	0.084	0.088	-4	-1	-2
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.853	0.875	-22	-3	-4
30	1	1	0	-	-
31	1	1	0	-	-
32	0.250	0.250	0	-	-

33	0.333	1	-666	-13	-17
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S2AT	S1AT		GROUPII	TOTAL
35	0.707	0.707	0	-	-
36	1	1	0	-	-
37	0.174	0.174	0	-	-
38	1	1	0	-	-
39	0.707	0.707	0	-	-
40	0.177	0.177	0	-	-
41	0.655	0.516	139	2	14
42	0.707	0.707	0	-	-
43	0.447	0.447	0	-	-
44	0.577	0.577	0	-	-
45	0.707	0.707	0	-	-
46	1	1	0	-	-
47	1	1	0	-	-
48	1	1	0	-	-
49	0.500	0.500	0	-	-
50	0.577	0.577	0	-	-
51	1	1	0	-	-
52	1	1	0	-	-
53	0.408	0.408	0	-	-
54	0.707	0.707	0	-	-
55	0.099	0.101	-2	-1	-1
56	1	1	0	-	-
57	0.577	0.577	0	-	-
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.816	0.816	0	-	-
61	0.707	0.707	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	0.707	0.707	0	-	-
65	0.500	0.500	0	-	-
66	0.577	0.577	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S2AT	S1AT		GROUPIII	TOTAL
67	0.577	0.577	0	-	-
68	0.302	0.302	0	-	-
69	1	1	0	-	-
70	0.250	0.250	0	-	-
71	0.500	0.500	0	-	-
72	0.707	0.707	0	-	-
73	0.707	0.707	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.447	0.447	0	-	-
79	1	1	0	-	-
80	1	1	0	-	-
81	0.577	0.577	0	-	-
82	0.707	0.707	0	-	-
83	1	1	0	-	-
84	1	1	0	-	-
85	0.577	0.577	0	-	-
86	1	1	0	-	-
87	0.365	0.408	-43	-1	-8
88	1	1	0	-	-
89	1	1	0	-	-
90	1	1	0	-	-
91	0.707	0.707	0	-	-
92	0.333	0.333	0	-	-
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.546	0.236	310	2	16
97	0.289	0.289	0	-	-
98	0.500	0.500	0	-	-

CALCULATIONS FOR S2AT v S1AT

For All Searches

n = 17

Σranks = 153

Σ-ve ranks = 74

P >0.4

For Each of the Three Groups

Group I	Group II	Group III
n = 13	too few diffs	too few diffs
Σranks = 91	to read off	to read off
Σ+ve ranks = 40	the tables	the tables
P >0.4	P >0.4*	P >0.4*

3BT Unstructured Searching of the Indexed File Versus Unstructured Searching of the Unindexed File (S2BT v S1BT)

QUERY	SCORES θ		DIFF	RANK	RANK
	S2BT	S1BT		GROUPI	TOTAL
1	0.548	0.311	237	9	21
2	0.671	0.671	0	-	-
3	0.577	0.586	-9	-2	-2
4	0.133	0.139	-6	-1	-1
5	0.707	0.707	0	-	-
6	1	1	0	-	-
7	1	1	0	-	-
8	0.707	0.707	0	-	-
9	0.316	0.316	0	-	-
10	0.426	0.426	0	-	-
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	0.866	0.866	0	-	-
15	0.378	0.378	0	-	-
16	0.707	0.707	0	-	-
17	0.500	0.354	146	8	15
18	0.426	0.500	-74	-5	-10
19	0.189	0.267	-78	-6	-11
20	1	1	0	-	-
21	1	1	0	-	-
22	0.730	0.600	130	7	14
23	0.400	0.400	0	-	-
24	0.105	0.124	-19	-3	-3
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.853	0.875	-22	-4	-4
30	1	1	0	-	-
31	1	1	0	-	-
32	0.816	0.816	0	-	-

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S2BT	S1BT		GROUPII	TOTAL
35	0.707	1	-293	-9	-22
36	0.408	0.408	0	-	-
37	0.386	0.348	38	1	6
38	1	1	0	-	-
39	0.707	0.707	0	-	-
40	0.236	0.168	68	3	9
41	0.516	0.433	83	4	12
42	0.577	0.577	0	-	-
43	0.447	0.447	0	-	-
44	0.866	0.666	200	6	18
45	0.707	0.707	0	-	-
46	1	1	0	-	-
47	1	1	0	-	-
48	0.500	0.707	-207	-7	-19
49	0.577	0.354	223	8	20
50	0.577	0.577	0	-	-
51	0.707	0.707	0	-	-
52	1	1	0	-	-
53	0.408	0.408	0	-	-
54	1	1	0	-	-
55	0.129	0.129	0	-	-
56	1	1	0	-	-
57	0.775	0.577	198	5	17
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.816	0.816	0	-	-
61	0.408	0.447	-39	-2	-7
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.612	0.612	0	-	-
66	0.200	0.200	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S2BT	S1BT		GROUPIII	TOTAL
67	0.816	0.816	0	-	-
68	0.731	0.701	30	1	5
69	1	1	0	-	-
70	0.289	0.289	0	-	-
71	0.707	0.707	0	-	-
72	0.707	0.707	0	-	-
73	0.707	0.707	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.447	0.447	0	-	-
79	1	1	0	-	-
80	1	1	0	-	-
81	1	0.816	184	4	16
82	1	1	0	-	-
83	1	1	0	-	-
84	0.447	0.500	-53	-2	-8
85	0.577	0.577	0	-	-
86	1	1	0	-	-
87	0.408	0.408	0	-	-
88	1	1	0	-	-
89	1	1	0	-	-
90	1	1	0	-	-
91	0.707	0.707	0	-	-
92	0.378	0.504	-126	-3	-13
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.471	0.471	0	-	-
97	0.622	0.622	0	-	-
98	1	1	0	-	-

CALCULATIONS FOR S2BT v S1BT

For All Searches

n = 22

Σranks = 253

Σ-ve ranks = 100

P >0.4

For Each of the Three Groups

Group I	Group II	Group III
n = 9	n = 9	too few diffs
Σranks = 45	Σranks = 45	to read off
Σ-ve ranks = 21	Σ-ve ranks = 18	the tables
P >0.4	P >0.4	P >0.4*

3AM Unstructured Searching of the Indexed File Versus Unstructured Searching of the Unindexed File (S2AM v S1AM)

QUERY	SCORES θ		DIFF	RANK	RANK
	S2AM	S1AM		GROUPI	TOTAL
1	0.263	0.263	0	-	-
2	0.192	0.192	0	-	-
3	0.209	0.218	-9	-2	-3
4	0.283	0.306	-23	-5	-6
5	0.183	0.183	0	-	-
6	1	1	0	-	-
7	1	1	0	-	-
8	0.354	0.408	-54	-8	-10
9	0.354	0.378	-24	-6.5	-8
10	0.500	0.511	-11	-3	-4
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	0.707	0.707	0	-	-
15	1	1	0	-	-
16	0.354	0.378	-24	-6.5	-8
17	0.600	0.500	100	10	14
18	0.369	0.433	-64	-9	-11
19	0.421	0.309	112	11	15
20	1	1	0	-	-
21	0.707	0.707	0	-	-
22	0.730	0.600	130	12	16
23	0.316	0.316	0	-	-
24	0.135	0.141	-6	-1	-2
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.853	0.875	-22	-4	-5
30	1	1	0	-	-
31	1	1	0	-	-
32	0.250	0.250	0	-	-

33	0.333	1	-667	-13	-17
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S2AM	S1AM		GROUPII	TOTAL
35	0.707	0.707	0	-	-
36	1	1	0	-	-
37	0.189	0.189	0	-	-
38	1	1	0	-	-
39	1	1	0	-	-
40	0.204	0.204	0	-	-
41	0.401	0.316	85	2	12
42	0.408	0.408	0	-	-
43	0.447	0.447	0	-	-
44	0.577	0.577	0	-	-
45	1	1	0	-	-
46	1	1	0	-	-
47	1	1	0	-	-
48	1	1	0	-	-
49	0.707	0.707	0	-	-
50	0.577	0.577	0	-	-
51	1	1	0	-	-
52	1	1	0	-	-
53	0.707	0.707	0	-	-
54	0.707	0.707	0	-	-
55	0.140	0.143	-3	-1	-1
56	1	1	0	-	-
57	1	1	0	-	-
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.707	0.707	0	-	-
61	0.707	0.707	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	0.707	0.707	0	-	-
65	0.354	0.354	0	-	-
66	0.577	0.577	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S2AM	S1AM		GROUPIII	TOTAL
67	0.577	0.577	0	-	-
68	0.316	0.316	0	-	-
69	1	1	0	-	-
70	0.302	0.302	0	-	-
71	0.707	0.707	0	-	-
72	0.707	0.707	0	-	-
73	1	1	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.707	0.707	0	-	-
79	0.707	0.707	0	-	-
80	1	1	0	-	-
81	0.707	0.707	0	-	-
82	0.707	0.707	0	-	-
83	1	1	0	-	-
84	1	1	0	-	-
85	0.707	0.707	0	-	-
86	1	1	0	-	-
87	0.316	0.408	-92	-2	-13
88	1	1	0	-	-
89	1	1	0	-	-
90	1	1	0	-	-
91	1	1	0	-	-
92	0.500	0.500	0	-	-
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.309	0.333	-24	-1	-8
97	0.316	0.316	0	-	-
98	0.500	0.500	0	-	-

CALCULATIONS FOR S2AM v S1AM

For All Searches

n = 17

Σranks = 153

Σ+ve ranks = 57

P = 0.4

For Each of the Three Groups

Group I	Group II	Group III
n = 13	too few diffs	too few diffs
Σranks = 91	to read off	to read off
Σ+ve ranks = 33	the tables	the tables
P >0.4	P >0.4*	P >0.4*

3BM Unstructured Searching of the Indexed File Versus Unstructured Searching of the Unindexed File (S2BM v S1BM)

QUERY	SCORES θ		DIFF	RANK	RANK
	S2BM	S1BM		GROUPI	TOTAL
1	0.470	0.252	218	9	20
2	0.316	0.316	0	-	-
3	0.218	0.258	-40	-4	-7
4	0.133	0.139	-6	-1	-2
5	1	1	0	-	-
6	1	1	0	-	-
7	1	1	0	-	-
8	1	1	0	-	-
9	0.316	0.316	0	-	-
10	0.577	0.577	0	-	-
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	0.866	0.866	0	-	-
15	1	1	0	-	-
16	1	1	0	-	-
17	0.500	0.354	146	8	16
18	0.369	0.433	-64	-5	-10
19	0.204	0.289	-85	-6	-11
20	1	1	0	-	-
21	1	1	0	-	-
22	0.730	0.600	130	7	15
23	0.632	0.632	0	-	-
24	0.169	0.200	-31	-3	-4
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.853	0.875	-22	-2	-3
30	1	1	0	-	-
31	1	1	0	-	-
32	0.816	0.816	0	-	-

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S2BM	S1BM		GROUPII	TOTAL
35	0.707	1	-293	-9	-21
36	0.408	0.408	0	-	-
37	0.314	0.252	62	3	9
38	1	1	0	-	-
39	1	1	0	-	-
40	0.164	0.065	99	5	13
41	0.632	0.530	102	6	14
42	0.333	0.333	0	-	-
43	0.447	0.447	0	-	-
44	0.866	0.666	200	7	18
45	1	1	0	-	-
46	1	1	0	-	-
47	1	1	0	-	-
48	0.500	0.707	-207	-8	-19
49	0.408	0.500	-92	-4	-12
50	0.577	0.577	0	-	-
51	0.707	0.707	0	-	-
52	1	1	0	-	-
53	0.707	0.707	0	-	-
54	1	1	0	-	-
55	0.183	0.183	0	-	-
56	1	1	0	-	-
57	0.447	0.500	-3	-1	-1
58	1	1	0	-	-
59	0.707	0.707	0	-	-
60	0.707	0.707	0	-	-
61	0.408	0.447	-39	-2	-6
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.866	0.866	0	-	-
66	0.200	0.200	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S2BM	S1BM		GROUPIII	TOTAL
67	0.816	0.816	0	-	-
68	0.690	0.653	37	1	5
69	1	1	0	-	-
70	0.174	0.174	0	-	-
71	1	1	0	-	-
72	0.707	0.707	0	-	-
73	1	1	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.707	0.707	0	-	-
79	0.707	0.707	0	-	-
80	1	1	0	-	-
81	0.816	1	-184	-3	-17
82	1	1	0	-	-
83	1	1	0	-	-
84	0.447	0.500	-53	-2	-8
85	0.707	0.707	0	-	-
86	1	1	0	-	-
87	0.707	0.707	0	-	-
88	1	1	0	-	-
89	1	1	0	-	-
90	1	1	0	-	-
91	1	1	0	-	-
92	0.378	0.378	0	-	-
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.666	0.666	0	-	-
97	0.568	0.568	0	-	-
98	1	1	0	-	-

CALCULATIONS FOR S2BM v S1BM

For All Searches

n = 21

Σranks = 231

Σ+ve ranks = 110

P >0.4

For Each of the Three Groups

Group I	Group II	Group III
n = 9	n = 9	too few diffs
Σranks = 45	Σranks = 45	to read off
Σ-ve ranks = 21	Σ+ve ranks = 21	the tables
P >0.4	P >0.4	P >0.4*

4AT Highly Structured Searching of the Unindexed File Versus Unstructured Searching of the Unindexed File (S4AT v S1AT)

QUERY	SCORES θ		DIFF	RANK	RANK
	S4AT	S1AT		GROUPI	TOTAL
1	0.677	0.379	298	15	34
2	0.465	0.408	57	2	3
3	0.577	0.577	0	-	-
4	0.639	0.306	333	17	37
5	0.129	0.129	0	-	-
6	0.816	1	-184	-10	-18
7	1	1	0	-	-
8	0.816	0,577	239	11	22
9	0.707	0.378	329	16	36
10	0.369	0.629	-260	-13	-26
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	0.775	0.707	68	3	5
15	0.267	0.378	-111	-6	-9
16	0.289	0.267	22	1	2
17	0.354	0.500	-146	-8	-13.5
18	0.500	0.500	0	-	-
19	0.169	0.286	-117	-7	-10
20	1	1	0	-	-
21	1	0.707	293	14	30
22	0.671	0.600	71	4	6
23	0.224	0.400	-176	-9	-17
24	0.160	0.088	72	5	7
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.632	0.875	-243	-12	-24
30	1	1	0	-	-
31	1	1	0	-	-
32	0.707	0.250	457	18	40

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S4AT	S1AT		GROUPII	TOTAL
35	1	0.707	293	10	30
36	0.408	1	-592	-15	-45
37	0.174	0.174	0	-	-
38	1	1	0	-	-
39	0.707	0.707	0	-	-
40	0.689	0.177	512	14	43
41	0.577	0.516	61	2	4
42	0.500	0.707	-207	-5.5	-19.5
43	0.775	0.447	328	12	35
44	0.816	0.577	239	7	22
45	0.707	0.707	0	-	-
46	1	1	0	-	-
47	1	1	0	-	-
48	1	1	0	-	-
49	0.707	0.500	207	5.5	19.5
50	0.577	0.577	0	-	-
51	0.707	1	-293	-10	-30
52	1	1	0	-	-
53	0.408	0.408	0	-	-
54	0.707	0.707	0	-	-
55	0.115	0.101	14	1	1
56	1	1	0	-	-
57	0.577	0.577	0	-	-
58	1	1	0	-	-
59	0.577	0.707	-130	-4	-11
60	0.408	0.816	-408	-13	-39
61	0.707	0.707	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	0.707	293	10	30
65	0.750	0.500	250	8	25
66	0.500	0.577	-77	-3	-8

QUERY	SCORES θ		DIFF	RANK	RANK
	S4AT	S1AT		GROUPIII	TOTAL
67	0.577	0.577	0	-	-
68	0.701	0.302	399	9	38
69	0.378	1	-622	-13	-46
70	0.079	0.250	-171	-3	-15
71	0.354	0.500	-146	-2	-13.5
72	0.707	0.707	0	-	-
73	1	0.707	293	7	30
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.447	0.447	0	-	-
79	0.500	1	-500	-10.5	-41.5
80	1	1	0	-	-
81	0.816	0.577	239	5	22
82	1	0.707	293	7	30
83	1	1	0	-	-
84	0.447	1	-553	-12	-44
85	0.577	0.577	0	-	-
86	1	1	0	-	-
87	0.408	0.408	0	-	-
88	0.707	1	-293	-7	-30
89	1	1	0	-	-
90	1	1	0	-	-
91	0.707	0.707	0	-	-
92	0.471	0.333	138	1	12
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.408	0.236	172	4	16
97	0.289	0.289	0	-	-
98	1	0.500	500	10.5	41.5

CALCULATIONS FOR S4AT v S1AT

For All Searches

$n = 46$

$n(n+1)/2 = 1081$

$\Sigma ranks = 1081$

$T = \Sigma$-ve ranks

$= 460$

$\mu = n(n+1)/4$

$= 540.5$

$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

$= 91.530$

$u = \frac{T-\mu}{\sigma}$

$= 0.879$

$0.3 < P < 0.4$

For Each of the Three Groups

Group I	Group II	Group III
$n = 18$	$n = 15$	$n = 13$
$\Sigma ranks = 171$	$\Sigma ranks = 120$	$\Sigma ranks = 91$
Σ-ve ranks = 65	Σ-ve ranks = 50.5	Σ-ve ranks = 37
$P = 0.4$	$P > 0.4$	$P > 0.4$

4BT Highly Structured Searching of the Unindexed File Versus Unstructured Searching of the Unindexed File (S4BT v S1BT)

QUERY	SCORES θ		DIFF	RANK GROUPI	RANK TOTAL
	S4BT	S1BT			
1	0.677	0.311	366	10	29
2	0.746	0.671	75	5	8
3	0.586	0.586	0	-	-
4	0.177	0.139	38	2	4
5	0.127	0.707	-580	-14	-36
6	0.816	1	-184	-8	-16.5
7	1	1	0	-	-
8	0.707	0.707	0	-	-
9	0.707	0.316	391	11	30
10	0.369	0.426	-57	-3	-5
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	0.866	0.866	0	-	-
15	0.267	0.378	-111	-7	-10
16	0.289	0.707	-418	-13	-32
17	0.354	0.354	0	-	-
18	0.500	0.500	0	-	-
19	0.661	0.267	394	12	31
20	1	1	0	-	-
21	1	1	0	-	-
22	0.671	0.600	71	4	7
23	0.400	0.400	0	-	-
24	0.160	0.124	36	1	3
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.632	0.875	-243	-9	-20
30	1	1	0	-	-
31	1	1	0	-	-
32	0.707	0.816	-109	-6	-9

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S4BT	S1BT		GROUPII	TOTAL
35	1	1	0	-	-
36	0.408	0.408	0	-	-
37	0.174	0.348	-174	-6	-15
38	1	1	0	-	-
39	0.707	0.707	0	-	-
40	0.736	0.168	568	12	35
41	0.764	0.433	331	10	27
42	0.707	0.577	130	2.5	11.5
43	0.775	0.447	328	9	26
44	0.816	0.666	150	5	14
45	0.707	0.707	0	-	-
46	1	1	0	-	-
47	1	1	0	-	-
48	0.707	0.707	0	-	-
49	0.707	0.354	353	11	28
50	0.577	0.577	0	-	-
51	0.707	0.707	0	-	-
52	1	1	0	-	-
53	0.408	0.408	0	-	-
54	0.816	1	-184	-7	-16.5
55	0.115	0.129	-14	-1	-1
56	1	1	0	-	-
57	0.577	0.577	0	-	-
58	1	1	0	-	-
59	0.577	0.707	-130	-2.5	-11.5
60	0.816	0.816	0	-	-
61	0.447	0.447	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.750	0.612	138	4	13
66	0.500	0.200	300	8	24

QUERY	SCORES θ		DIFF	RANK	RANK
	S4BT	S1BT		GROUPIII	TOTAL
67	0.816	0.816	0	-	-
68	0.701	0.701	0	-	-
69	1	1	0	-	-
70	0.612	0.289	323	8	25
71	0.500	0.707	-207	-3	-18
72	1	0.707	293	6.5	22.5
73	0.707	0.707	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.894	0.447	447	10	34
79	0.707	1	-293	-6.5	-22.5
80	1	1	0	-	-
81	0.816	0.816	0	-	-
82	1	1	0	-	-
83	1	1	0	-	-
84	0.250	0.500	-250	-5	-21
85	1	0.577	423	9	33
86	1	1	0	-	-
87	0.187	0.408	-221	-4	-19
88	1	1	0	-	-
89	1	1	0	-	-
90	1	1	0	-	-
91	0.707	0.707	0	-	-
92	0.471	0.504	-33	-1	-2
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.408	0.471	-63	-2	-6
97	0.622	0.622	0	-	-
98	1	1	0	-	-

CALCULATIONS FOR S4BT v S1BT

For All Searches

$n = 36$

$n(n+1)/2 = 666$

$\Sigma ranks = 666$

$T = \Sigma$-ve ranks

$= 261$

$\mu = n(n+1)/4$

$= 333$

$$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

$= 63.651$

$$u = \frac{T-\mu}{\sigma}$$

$= 1.131$

$0.2 < P < 0.3$

For Each of the Three Groups

Group I	Group II	Group III
$n = 14$	$n = 12$	$n = 10$
$\Sigma ranks = 105$	$\Sigma ranks = 78$	$\Sigma ranks = 55$
Σ+ve ranks = 45	Σ-ve ranks = 16.5	Σ-ve ranks = 21.5
$P > 0.4$	$0.05 < P < 0.1$	$P > 0.4$

4AM Highly Structured Searching of the Unindexed File Versus Unstructured Searching of the Unindexed File (S4AM v S1AM)

QUERY	SCORES θ		DIFF	RANK	RANK
	S4AM	S1AM		GROUPI	TOTAL
1	0.705	0.263	442	17	40
2	0.164	0.192	-28	-1	-1
3	0.218	0.218	0	-	-
4	0.639	0.306	333	16	35
5	0.183	0.183	0	-	-
6	0.816	1	-184	-11	-15
7	1	1	0	-	-
8	0.577	0.408	169	9	12.5
9	0.707	0.378	329	15	34
10	0.333	0.511	-178	-10	-14
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	0.775	0.707	68	4	5
15	0.707	1	-293	-13.5	-26.5
16	0.408	0.378	30	2	2
17	0.354	0.500	-146	-8	-11
18	0.433	0.433	0	-	-
19	0.183	0.309	-126	-7	-9
20	1	1	0	-	-
21	1	0.707	293	13.5	26.5
22	0.671	0.600	71	5	6
23	0.354	0.316	38	3	3
24	0.258	0.141	117	6	8
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.632	0.875	-243	-12	-20
30	1	1	0	-	-
31	1	1	0	-	-
32	0.707	0.250	457	18	41

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S4AM	S1AM		GROUPII	TOTAL
35	1	0.707	293	7.5	26.5
36	0.408	1	-592	-14	-44
37	0.189	0.189	0	-	-
38	1	1	0	-	-
39	1	1	0	-	-
40	0.169	0.204	415	13	39
41	0.707	0.316	391	12	38
42	0.577	0.408	169	4	12.5
43	0.775	0.447	328	10	33
44	0.816	0.577	239	5	19
45	1	1	0	-	-
46	1	1	0	-	-
47	1	1	0	-	-
48	1	1	0	-	-
49	1	0.707	293	7.5	26.5
50	0.577	0.577	0	-	-
51	0.707	1	-293	-7.5	-26.5
52	1	1	0	-	-
53	0.707	0.707	0	-	-
54	0.707	0.707	0	-	-
55	0.082	0.143	-61	-1	-4
56	1	1	0	-	-
57	1	1	0	-	-
58	1	1	0	-	-
59	0.577	0.707	-130	-3	-10
60	0.707	0.707	0	-	-
61	0.707	0.707	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	0.707	293	7.5	26.5
65	0.707	0.354	353	11	37
66	0.500	0.577	-77	-2	-7

QUERY	SCORES θ		DIFF	RANK	RANK
	S4AM	S1AM		GROUPIII	TOTAL
67	0.577	0.577	0	-	-
68	0.653	0.316	337	10	36
69	0.378	1	-622	-13	-45
70	0.095	0.302	-207	-2	-17
71	0.500	0.707	-207	-2	-17
72	0.707	0.707	0	-	-
73	0.707	1	-293	-6.5	-26.5
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.707	0.707	0	-	-
79	0.707	0.707	0	-	-
80	1	1	0	-	-
81	1	0.707	293	6.5	26.5
82	1	0.707	293	6.5	26.5
83	1	1	0	-	-
84	0.447	1	-553	-12	-43
85	0.707	0.707	0	-	-
86	1	1	0	-	-
87	0.707	0.408	299	9	32
88	0.707	1	-293	-6.5	-26.5
89	1	1	0	-	-
90	1	1	0	-	-
91	1	1	0	-	-
92	0.707	0.500	207	2	17
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.577	0.333	244	4	21
97	0.316	0.316	0	-	-
98	1	0.500	500	11	42

CALCULATIONS FOR S4AM v S1AM

For All Searches

$n = 45$

$n(n+1)/2 = 1035$

$\Sigma ranks = 1035$

$T = \Sigma$-ve ranks

$= 363$

$\mu = n(n+1)/4$

$= 517.5$

$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

$= 88.593$

$u = \frac{T-\mu}{\sigma}$

$= 1.744$

$0.05 < P < 0.1$

For Each of the Three Groups

Group I	Group II	Group III
$n = 18$	$n = 14$	$n = 13$
$\Sigma ranks = 171$	$\Sigma ranks = 105$	$\Sigma ranks = 91$
Σ-ve ranks = 62.5	Σ-ve ranks = 27.5	Σ-ve ranks = 42
$0.2 < P < 0.4$	$0.1 < P < 0.2$	$P > 0.4$

4BM Highly Structured Searching of the Unindexed File Versus Unstructured Searching of the Unindexed File (S4BM v S1BM)

QUERY	SCORES θ		DIFF	RANK GROUPI	RANK TOTAL
	S4BM	S1BM			
1	0.705	0.252	453	12	31
2	0.302	0.316	-14	-1	-1
3	0.258	0.258	0	-	-
4	0.177	0.139	38	2	2
5	0.180	1	-820	-14	-36
6	0.816	1	-184	-6	-14.5
7	1	1	0	-	-
8	1	1	0	-	-
9	0.707	0.316	391	10	29
10	0.333	0.577	-244	-8	-17
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	0.866	0.866	0	-	-
15	0.707	1	-293	-9	-21.5
16	0.408	1	-592	-13	-34
17	0.354	0.354	0	-	-
18	0.433	0.433	0	-	-
19	0.714	0.289	425	11	30
20	1	1	0	-	-
21	1	1	0	-	-
22	0.671	0.600	71	4	5
23	0.632	0.632	0	-	-
24	0.258	0.200	58	3	3
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.632	0.875	-243	-7	-16
30	1	1	0	-	-
31	1	1	0	-	-
32	0.707	0.816	-109	-5	-9.5

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S4BM	S1BM		GROUPII	TOTAL
35	1	1	0	-	-
36	0.408	0.408	0	-	-
37	0.189	0.252	-63	-1	-4
38	1	1	0	-	-
39	1	1	0	-	-
40	0.661	0.065	596	12	35
41	0.802	0.530	272	8	19
42	0.408	0.333	75	2	6
43	0.775	0.447	328	10	26
44	0.816	0.666	150	5	12
45	1	1	0	-	-
46	1	1	0	-	-
47	1	1	0	-	-
48	0.707	0.707	0	-	-
49	1	0.500	500	11	32.5
50	0.577	0.577	0	-	-
51	0.707	0.707	0	-	-
52	1	1	0	-	-
53	0.707	0.707	0	-	-
54	0.816	1	-184	-7	-14.5
55	0.082	0.183	-101	-3	-8
56	1	1	0	-	-
57	0.500	0.500	0	-	-
58	1	1	0	-	-
59	0.577	0.707	-130	-4	-11
60	0.707	0.707	0	-	-
61	0.447	0.447	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.707	0.866	-159	-6	-13
66	0.500	0.200	300	9	24

QUERY	SCORES θ		DIFF	RANK	RANK
	S4BM	S1BM		GROUPIII	TOTAL
67	0.816	0.816	0	-	-
68	0.653	0.653	0	-	-
69	1	1	0	-	-
70	0.492	0.174	318	7	25
71	0.707	1	-293	-5	-21.5
72	1	0.707	293	5	21.5
73	1	1	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.707	0.707	0	-	-
79	1	0.707	293	5	21.5
80	1	1	0	-	-
81	1	1	0	-	-
82	1	1	0	-	-
83	1	1	0	-	-
84	0.250	0.500	-250	-3	-18
85	0.816	0.707	109	2	9.5
86	1	1	0	-	-
87	0.324	0.707	-383	-9	-28
88	1	1	0	-	-
89	1	1	0	-	-
90	1	1	0	-	-
91	0.500	1	-500	-10	-32.5
92	0.707	0.378	329	8	27
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.577	0.666	-89	-1	-7
97	0.568	0.568	0	-	-
98	1	1	0	-	-

CALCULATIONS FOR S4BM v S1BM

For All Searches

$n = 36$

$n(n+1)/2 = 666$

$\Sigma ranks = 666$

$T = \Sigma$-ve ranks

$= 307$

$\mu = n(n+1)/4$

$= 333$

$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

$= 63.651$

$u = \frac{T-\mu}{\sigma}$

$= 0.408$

$P > 0.4$

For Each of the Three Groups

Group I	Group II	Group III
$n = 14$	$n = 12$	$n = 10$
$\Sigma ranks = 105$	$\Sigma ranks = 78$	$\Sigma ranks = 55$
Σ+ve ranks = 41	Σ-ve ranks = 21	Σ+ve ranks = 27
$P > 0.4$	$P = 0.2$	$P > 0.4$

5AT Structured Searching of the Indexed File Versus
Highly Structured Searching of the Unindexed File (S3AT v S4AT)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AT	S4AT		GROUPI	TOTAL
1	0.601	0.677	-76	-4	-8.5
2	0.791	0.465	326	14	29
3	0.488	0.577	-89	-6	-7
4	0.417	0.639	-222	-9	-17
5	1	0.129	871	19	46
6	1	0.816	184	8	13
7	1	1	0	-	-
8	0.707	0.816	-109	-7	-9
9	1	0.707	293	12.5	25.5
10	0.302	0.369	-67	-3	-3
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.775	225	11	19
15	0.771	0.267	504	17	40
16	0.707	0.289	418	16	35
17	0.354	0.354	0	-	-
18	0.577	0.500	77	5	6
19	0.535	0.169	366	15	32
20	1	1	0	-	-
21	1	1	0	-	-
22	0.632	0.671	-39	-2	-2
23	0.447	0.224	223	10	18
24	1	0.160	840	18	45
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.632	-20	-1	-1
30	0.707	1	-293	-12.5	-25.5
31	1	1	0	-	-
32	0.707	0.707	0	-	-

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AT	S4AT		GROUPII	TOTAL
35	0.816	1	-184	-2	-13
36	1	0.408	592	15.5	42.5
37	0.174	0.174	0	-	-
38	1	1	0	-	-
39	1	0.707	293	8.5	25.5
40	0.401	0.689	-288	-7	-22
41	0.822	0.577	245	6	21
42	0.500	0.500	0	-	-
43	0.447	0.775	-328	-10	-30.5
44	0.816	0.816	0	-	-
45	0.267	0.707	-440	-13	-37
46	1	1	0	-	-
47	1	1	0	-	-
48	1	1	0	-	-
49	0.500	0.707	-207	-4	-16
50	1	0.577	423	12	36
51	1	0.707	293	8.5	25.5
52	1	1	0	-	-
53	0.408	0.408	0	-	-
54	0.707	0.707	0	-	-
55	0.707	0.115	592	15.5	42.5
56	1	1	0	-	-
57	0.816	0.577	239	5	20
58	1	1	0	-	-
59	0.707	0.577	130	1	10
60	0.816	0.408	408	11	33.5
61	0.707	0.707	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.935	0.750	185	3	15
66	1	0.500	500	14	38.5

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AT	S4AT		GROUPIII	TOTAL
67	0.577	0.577	0	-	-
68	0.603	0.701	-98	-2	-8
69	1	0.378	622	11	44
70	0.250	0.079	171	3	11
71	0.426	0.354	72	1	4
72	1	0.707	293	5.5	25.5
73	0.707	1	-293	-5.5	-25.5
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.775	0.447	328	7	30.5
79	0.500	0.500	0	-	-
80	1	1	0	-	-
81	1	0.816	184	4	13
82	1	1	0	-	-
83	1	1	0	-	-
84	1	0.447	553	10	41
85	0.577	0.577	0	-	-
86	1	1	0	-	-
87	0.408	0.408	0	-	-
88	0.707	0.707	0	-	-
89	1	1	0	-	-
90	1	1	0	-	-
91	0.707	0.707	0	-	-
92	0.471	0.471	0	-	-
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.816	0.408	408	8	33.5
97	0.289	0.289	0	-	-
98	0.500	1	-500	-9	-38.5

CALCULATIONS FOR S3AT v S4AT

For All Searches

$n = 46$

$n(n+1)/2 = 1081$

$\Sigma ranks = 1081$

$T = \Sigma$-ve ranks

$= 260$

$\mu = n(n+1)/4$

$= 540.5$

$$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

$= 91.530$

$$u = \frac{T-\mu}{\sigma}$$

$= 3.065$

$0.001 < P < 0.01$

For Each of the Three Groups

Group I	Group II	Group III
$n = 19$	$n = 16$	$n = 11$
$\Sigma ranks = 190$	$\Sigma ranks = 136$	$\Sigma ranks = 66$
Σ-ve ranks = 44.5	Σ-ve ranks = 35	Σ-ve ranks = 16.5
$0.02 < P < 0.05$	$P = 0.1$	$0.1 < P < 0.2$

5BT Structured Searching of the Indexed File Versus
Highly Structured Searching of the Unindexed File (S3BT v S4BT)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BT	S4BT		GROUPI	TOTAL
1	0.677	0.677	0	-	-
2	0.791	0.746	45	2	3
3	0.488	0.586	-98	-6	-8
4	0.417	0.177	240	10	27
5	1	0.127	873	17	51
6	1	0.816	184	9	17
7	1	1	0	-	-
8	0.707	0.707	0	-	-
9	1	0.707	293	11.5	33.5
10	0.302	0.369	-67	-4	-5
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.866	134	8	13
15	0.771	0.267	504	15	46
16	0.707	0.289	418	14	41
17	0.707	0.354	353	13	39
18	0.577	0.500	77	5	7
19	0.535	0.661	-126	-7	-11
20	1	1	0	-	-
21	1	1	0	-	-
22	0.671	0.671	0	-	-
23	0.447	0.400	47	3	4
24	1	0.160	840	16	50
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.632	-20	-1	-1
30	0.707	1	-293	-11.5	-33.5
31	1	1	0	-	-
32	0.707	0.707	0	-	-

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BT	S4BT		GROUPII	TOTAL
35	0.816	1	-184	-3.5	-17
36	1	0.408	592	17	49
37	0.395	0.174	221	7	24.5
38	1	1	0	-	-
39	0.707	0.707	0	-	-
40	0.471	0.736	-265	-9	-29
41	0.764	0.764	0	-	-
42	0.500	0.707	-207	-6	-22.5
43	0.447	0.775	-328	-13	-38
44	0.816	0.816	0	-	-
45	1	0.707	293	11	33.5
46	1	1	0	-	-
47	1	1	0	-	-
48	1	0.707	293	11	33.5
49	0.707	0.707	0	-	-
50	1	0.577	423	14	42.5
51	1	0.707	293	11	33.5
52	1	1	0	-	-
53	0.408	0.408	0	-	-
54	1	0.816	184	3.5	17
55	0.632	0.115	517	16	48
56	1	1	0	-	-
57	0.816	0.577	239	8	26
58	1	1	0	-	-
59	0.707	0.577	130	2	12
60	0.816	0.816	0	-	-
61	0.408	0.447	-39	-1	-2
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.935	0.750	185	5	20
66	1	0.500	500	15	44.5

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BT	S4BT		GROUPIII	TOTAL
67	1	0.816	184	5.5	17
68	0.957	0.701	256	10	28
69	1	1	0	-	-
70	0.250	0.612	-362	-14	-40
71	0.426	0.500	-74	-1	-6
72	1	1	0	-	-
73	0.707	0.707	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.775	0.894	-119	-3	-10
79	0.500	0.707	-207	-8	-22.5
80	1	1	0	-	-
81	1	0.816	184	5.5	17
82	1	1	0	-	-
83	1	1	0	-	-
84	0.354	0.250	104	2	9
85	0.577	1	-423	-15	-42.5
86	1	1	0	-	-
87	0.408	0.187	221	9	24.5
88	0.707	1	-293	-12	-33.5
89	1	1	0	-	-
90	0.707	1	-293	-12	-33.5
91	1	0.707	293	12	33.5
92	0.333	0.471	-138	-4	-14
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.913	0.408	505	17	47
97	0.816	0.622	194	7	21
98	0.500	1	-500	-16	-44.5

CALCULATIONS FOR S3BT v S4BT

For All Searches

$n = 51$

$n(n+1)/2 = 1326$

$\Sigma ranks = 1326$

$T = \Sigma$-ve ranks

$= 412.5$

$\mu = n(n+1)/4$

$= 663$

$$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

$= 106.684$

$$u = \frac{T-\mu}{\sigma}$$

$= 2.348$

$0.01 < P < 0.02$

For Each of the Three Groups

Group I	Group II	Group III
$n = 17$	$n = 17$	$n = 17$
$\Sigma ranks = 153$	$\Sigma ranks = 153$	$\Sigma ranks = 153$
Σ-ve ranks = 29.5	Σ-ve ranks = 32.5	Σ+ve ranks = 68
$0.02 < P < 0.05$	$0.02 < P < 0.05$	$P > 0.4$

5AM Structured Searching of the Indexed File Versus
Highly Structured Searching of the Unindexed File (S3AM v S4AM)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AM	S4AM		GROUPI	TOTAL
1	0.514	0.705	-191	-7	-13
2	0.447	0.164	283	10	19
3	0.258	0.218	40	3	4
4	0.417	0.639	-222	-8	-16
5	0.707	0.183	524	18	38
6	1	0.816	184	6	11
7	1	1	0	-	-
8	1	0.577	423	17	34.5
9	1	0.707	293	11.5	23.5
10	0.408	0.333	75	5	7
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.775	225	9	17
15	0.408	0.707	-299	-13	-28.5
16	1	0.408	592	19	40.5
17	0.354	0.354	0	-	-
18	0.500	0.433	67	4	6
19	0.577	0.183	394	16	33
20	1	1	0	-	-
21	1	1	0	-	-
22	0.632	0.671	-39	-2	-3
23	0.707	0.354	343	14	31
24	0.620	0.258	362	15	32
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.632	-20	-1	-1
30	0.707	1	-293	-11.5	-23.5
31	1	1	0	-	-
32	0.707	0.707	0	-	-

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AM	S4AM		GROUPII	TOTAL
35	0.816	1	-184	-4	-11
36	1	0.408	592	13	40.5
37	0.189	0.189	0	-	-
38	1	1	0	-	-
39	0.707	1	-293	-7.5	-23.5
40	0.386	0.619	-233	-5	-18
41	0.783	0.707	76	2	8
42	0.577	0.577	0	-	-
43	0.447	0.775	-328	-10	-30
44	0.816	0.816	0	-	-
45	0.378	1	-622	-14	-42.5
46	1	1	0	-	-
47	1	1	0	-	-
48	1	1	0	-	-
49	0.707	1	-293	-7.5	-23.5
50	1	0.577	423	11	34.5
51	1	0.707	293	7.5	23.5
52	1	1	0	-	-
53	0.707	0.707	0	-	-
54	0.707	0.707	0	-	-
55	1	0.082	918	15	44
56	1	1	0	-	-
57	0.707	1	-293	-7.5	-23.5
58	1	1	0	-	-
59	0.707	0.577	130	3	9
60	0.707	0.707	0	-	-
61	0.707	0.707	0	-	-
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.756	0.707	49	1	5
66	1	0.500	500	12	36.5

QUERY	SCORES θ		DIFF	RANK	RANK
	S3AM	S4AM		GROUPIII	TOTAL
67	0.577	0.577	0	-	-
68	0.632	0.653	-21	-1	-2
69	1	0.378	622	10	42.5
70	0.302	0.095	207	4	15
71	0.302	0.500	-198	-3	-14
72	1	0.707	293	5.5	23.5
73	1	0.707	293	5.5	23.5
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.408	0.707	-299	-7	-28.5
79	0.707	0.707	0	-	-
80	1	1	0	-	-
81	0.816	1	-184	-2	-11
82	1	1	0	-	-
83	1	1	0	-	-
84	1	0.447	553	9	39
85	0.707	0.707	0	-	-
86	1	1	0	-	-
87	0.707	0.707	0	-	-
88	0.707	0.707	0	-	-
89	1	1	0	-	-
90	1	1	0	-	-
91	1	1	0	-	-
92	0.707	0.707	0	-	-
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.577	0.577	0	-	-
97	0.316	0.316	0	-	-
98	0.500	1	-500	-8	-36.5

CALCULATIONS FOR S3AM v S4AM

For All Searches

$n = 44$

$n(n+1)/2 = 990$

$\Sigma ranks = 990$

$T = \Sigma$-ve ranks

$= 349$

$\mu = n(n+1)/4$

$= 495$

$$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

$= 85.688$

$$u = \frac{T-\mu}{\sigma}$$

$= 1.704$

$0.05 < P < 0.1$

For Each of the Three Groups

Group I	Group II	Group III
$n = 19$	$n = 15$	$n = 10$
$\Sigma ranks = 190$	$\Sigma ranks = 120$	$\Sigma ranks = 55$
Σ-ve ranks = 42.5	Σ-ve ranks = 55.5	Σ-ve ranks = 21
$0.02 < P < 0.05$	$P > 0.4$	$P > 0.4$

5BM Structured Searching of the Indexed File Versus
Highly Structured Searching of the Unindexed File (S3BM v S4BM)

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BM	S4BM		GROUPI	TOTAL
1	0.705	0.705	0	-	-
2	0.447	0.302	145	7	14
3	0.258	0.258	0	-	-
4	0.417	0.177	240	9	25
5	0.707	0.180	527	15	48
6	1	0.816	184	8	18
7	1	1	0	-	-
8	1	1	0	-	-
9	1	0.707	293	10.5	31.5
10	0.408	0.333	75	3.5	6.5
11	1	1	0	-	-
12	1	1	0	-	-
13	1	1	0	-	-
14	1	0.866	134	5	12
15	0.408	0.707	-299	-12	-36.5
16	1	0.408	592	16	49.5
17	0.707	0.354	353	13	40
18	0.500	0.433	67	2	5
19	0.577	0.714	-137	-6	-13
20	1	1	0	-	-
21	1	1	0	-	-
22	0.671	0.671	0	-	-
23	0.707	0.632	75	3.5	6.5
24	0.620	0.258	362	14	41
25	1	1	0	-	-
26	1	1	0	-	-
27	0.500	0.500	0	-	-
28	0.707	0.707	0	-	-
29	0.612	0.632	-20	-1	-1
30	0.707	1	-293	-10.5	-31.5
31	1	1	0	-	-
32	0.707	0.707	0	-	-

33	1	1	0	-	-
34	1	1	0	-	-

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BM	S4BM		GROUPII	TOTAL
35	0.816	1	-184	-6.5	-18
36	1	0.408	592	17	49.5
37	0.286	0.189	97	3	8
38	1	1	0	-	-
39	1	1	0	-	-
40	0.476	0.661	-185	-8	-21
41	0.802	0.802	0	-	-
42	0.577	0.408	169	5	15
43	0.447	0.775	-328	-13	-39
44	0.816	0.816	0	-	-
45	0.707	1	-293	-11	-31.5
46	1	1	0	-	-
47	1	1	0	-	-
48	1	0.707	293	11	31.5
49	1	1	0	-	-
50	1	0.577	423	15	45
51	1	0.707	293	11	31.5
52	1	1	0	-	-
53	0.707	0.707	0	-	-
54	1	0.816	184	6.5	18
55	0.447	0.082	365	14	42
56	1	1	0	-	-
57	0.707	0.500	207	9	23
58	1	1	0	-	-
59	0.707	0.577	130	4	11
60	0.707	0.707	0	-	-
61	0.408	0.447	-39	-1	-2
62	1	1	0	-	-
63	0.707	0.707	0	-	-
64	1	1	0	-	-
65	0.756	0.707	49	2	3
66	1	0.500	500	16	46.5

QUERY	SCORES θ		DIFF	RANK	RANK
	S3BM	S4BM		GROUPIII	TOTAL
67	1	0.816	184	4.5	18
68	0.913	0.653	260	8	26
69	1	1	0	-	-
70	0.201	0.492	-291	-9	-27
71	0.302	0.707	-405	-16	-44
72	1	1	0	-	-
73	1	1	0	-	-
74	1	1	0	-	-
75	1	1	0	-	-
76	0.707	0.707	0	-	-
77	0.707	0.707	0	-	-
78	0.408	0.707	-299	-13	-36.5
79	0.707	1	-293	-11	-31.5
80	1	1	0	-	-
81	0.816	1	-184	-4.5	-18
82	1	1	0	-	-
83	1	1	0	-	-
84	0.354	0.250	104	2	9
85	0.707	0.816	-109	-3	-10
86	1	1	0	-	-
87	0.707	0.324	383	15	43
88	0.707	1	-293	-11	-31.5
89	1	1	0	-	-
90	0.707	1	-293	-11	-31.5
91	0.707	0.500	207	6.5	23
92	0.500	0.707	-207	-6.5	-23
93	1	1	0	-	-
94	1	1	0	-	-
95	1	1	0	-	-
96	0.516	0.577	-61	-1	-4
97	0.894	0.568	326	14	38
98	0.500	1	-500	-17	-46.5

CALCULATIONS FOR S3BM v S4BM

For All Searches

$n = 50$

$n(n+1)/2 = 1275$

$\Sigma ranks = 1275$

$T = \Sigma$-ve ranks

$= 497$

$\mu = n(n+1)/4$

$= 637.5$

$\sigma = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

$= 103.592$

$u = \frac{T-\mu}{\sigma}$

$= 1.356$

$0.1 < P < 0.2$

For Each of the Three Groups

Group I	Group II	Group III
$n = 16$	$n = 17$	$n = 17$
$\Sigma ranks = 136$	$\Sigma ranks = 153$	$\Sigma ranks = 153$
Σ-ve ranks = 29.5	Σ-ve ranks = 39.5	Σ+ve ranks = 50
$P = 0.05$	$0.05 < P < 0.1$	$0.2 < P < 0.4$

APPENDIX J

RECALL AND PRECISION RATIOS

RECALL RATIOS

Query	Relevant Hits Retrieved/ All Relevant Hits		Recall Ratios		Relevant Major Hits Retrieved/ All Relevant Hits		Major Recall Ratios	
	A	B	A	B	A	B	A	B
1	26/30	20/30	86.7	66.7	18/20	17/20	90.0	85.0
2	4/7	4/7	57.1	57.1	0/0	0/0	100	100
3	4/6	4/6	66.7	66.7	0/0	0/0	100	100
4	4/15	4/15	26.7	26.7	4/15	4/15	26.7	26.7
5	1/1	1/1	100	100	0/0	0/0	100	100
6	1/1	1/1	100	100	1/1	1/1	100	100
7	1/1	1/1	100	100	1/1	1/1	100	100
8	0/1	0/1	0	0	0/0	0/0	100	100
9	0/0	0/0	100	100	0/0	0/0	100	100
10	1/10	1/10	10.0	10.0	1/5	1/5	20.0	20.0
11	0/0	0/0	100	100	0/0	0/0	100	100
12	0/0	0/0	100	100	0/0	0/0	100	100
13	0/0	0/0	100	100	0/0	0/0	100	100
14	2/2	2/2	100	100	2/2	2/2	100	100
15	4/6	4/6	66.7	66.7	0/0	0/0	100	100
16	0/1	0/1	0	0	0/0	0/0	100	100
17	0/7	3/7	0	42.9	0/7	3/7	0	42.9
18	3/3	3/3	100	100	2/2	2/2	100	100
19	3/13	3/13	23.1	23.1	3/11	3/11	27.3	27.3
20	0/0	0/0	100	100	0/0	0/0	100	100
21	0/0	0/0	100	100	0/0	0/0	100	100
22	1/4	2/4	25.0	50.0	1/4	2/4	25.0	50.0
23	0/4	0/4	0	0	0/1	0/1	0	0
24	12/12	12/12	100	100	4/4	4/4	100	100
25	0/0	0/0	100	100	0/0	0/0	100	100

	Relevant Hits Retrieved/ All Relevant Hits		Recall Ratios		Relevant Major Hits Retrieved/ All Relevant Hits		Major Recall Ratios	
Query	A	B	A	B	A	B	A	B
26	1/1	1/1	100	100	1/1	1/1	100	100
27	0/3	0/3	0	0	0/3	0/3	0	0
28	0/0	0/0	100	100	0/0	0/0	100	100
29	2/7	2/7	28.6	28.6	2/7	2/7	28.6	28.6
30	0/1	0/1	0	0	0/1	0/1	0	0
31	0/0	0/0	100	100	0/0	0/0	100	100
32	0/1	0/1	0	0	0/1	0/1	0	0
33	0/0	0/0	100	100	0/0	0/0	100	100
34	0/0	0/0	100	100	0/0	0/0	100	100
35	1/1	1/1	100	100	1/1	1/1	100	100
36	0/0	0/0	100	100	0/0	0/0	100	100
37	0/32	5/32	0	15.6	0/27	3/27	0	11.1
38	0/0	0/0	100	100	0/0	0/0	100	100
39	1/1	0/1	100	0	0/0	0/0	100	100
40	5/31	7/31	16.1	22.6	4/23	6/23	17.4	26.1
41	8/11	6/11	72.7	54.5	6/7	5/7	85.7	71.4
42	0/3	0/3	0	0	0/2	0/2	0	0
43	0/4	0/4	0	0	0/4	0/4	0	0
44	1/2	1/2	50.0	50.0	1/2	1/2	50.0	50.0
45	0/1	1/1	0	100	0/0	0/0	100	100
46	0/0	0/0	100	100	0/0	0/0	100	100
47	0/0	0/0	100	100	0/0	0/0	100	100
48	0/0	0/0	100	100	0/0	0/0	100	100
49	1/3	1/3	33.3	33.3	1/1	1/1	100	100
50	2/2	2/2	100	100	2/2	2/2	100	100
51	0/0	0/0	100	100	0/0	0/0	100	100
52	0/0	0/0	100	100	0/0	0/0	100	100
53	0/5	0/5	0	0	0/1	0/1	0	0
54	0/1	1/1	0	100	0/1	1/1	0	100
55	0/1	1/1	0	100	0/0	0/0	100	100
56	0/0	0/0	100	100	0/0	0/0	100	100

Query	Relevant Hits Retrieved/ All Relevant Hits		Recall Ratios		Relevant Major Hits Retrieved/ All Relevant Hits		Major Recall Ratios	
	A	B	A	B	A	B	A	B
57	1/2	1/2	50.0	50.0	0/0	0/0	100	100
58	0/0	0/0	100	100	0/0	0/0	100	100
59	0/1	0/1	0	0	0/1	0/1	0	0
60	1/2	1/2	50.0	50.0	0/0	0/0	100	100
61	0/0	0/0	100	100	0/0	0/0	100	100
62	0/0	0/0	100	100	0/0	0/0	100	100
63	0/1	0/1	0	0	0/1	0/1	0	0
64	0/0	0/0	100	100	0/0	0/0	100	100
65	6/7	6/7	85.7	85.7	3/3	3/3	100	100
66	0/0	0/0	100	100	0/0	0/0	100	100
67	0/2	2/2	0	100	0/2	2/2	0	100
68	3/10	10/10	30.0	100	3/9	9/9	33.3	100
69	0/0	0/0	100	100	0/0	0/0	100	100
70	0/15	2/15	0	13.3	0/10	1/10	0	10
71	1/1	1/1	100	100	0/0	0/0	100	100
72	1/1	1/1	100	100	1/1	1/1	100	100
73	0/1	0/1	0	0	0/0	0/0	100	100
74	0/0	0/0	100	100	0/0	0/0	100	100
75	0/0	0/0	100	100	0/0	0/0	100	100
76	0/1	0/1	0	0	0/1	0/1	0	0
77	0/1	0/1	0	0	0/1	0/1	0	0
78	2/4	2/4	50.0	50.0	0/1	0/1	0	0
79	0/3	0/3	0	0	0/1	0/1	0	0
80	0/0	0/0	100	100	0/0	0/0	100	100
81	2/2	2/2	100	100	1/1	1/1	100	100
82	1/1	1/1	100	100	1/1	1/1	100	100
83	0/0	0/0	100	100	0/0	0/0	100	100
84	0/0	0/0	100	100	0/0	0/0	100	100
85	0/2	0/2	0	0	0/1	0/1	0	0
86	0/0	0/0	100	100	0/0	0/0	100	100
87	0/5	0/5	0	0	0/1	0/1	0	0

Query	Relevant Hits Retrieved/ All Relevant Hits		Recall Ratios		Relevant Major Hits Retrieved/ All Relevant Hits		Major Recall Ratios	
	A	B	A	B	A	B	A	B
88	0/1	0/1	0	0	0/1	0/1	0	0
89	0/0	0/0	100	100	0/0	0/0	100	100
90	0/0	0/0	100	100	0/0	0/0	100	100
91	0/1	1/1	0	100	0/0	0/0	100	100
92	1/8	0/8	12.5	0	1/3	0/3	33.3	0
93	0/0	0/0	100	100	0/0	0/0	100	100
94	0/0	0/0	100	100	0/0	0/0	100	100
95	0/0	0/0	100	100	0/0	0/0	100	100
96	3/5	4/5	60.0	80.0	1/2	1/2	50.0	50.0
97	0/11	7/11	0	63.6	0/9	7/9	0	77.8
98	0/3	0/3	0	0	0/9	0/3	0	0

PRECISION RATIOS

Query	Relevant Hits/ Retrieved Hits A	Relevant Hits/ Retrieved Hits B	Precision Ratios A	Precision Ratios B	Relevant Major Hits/Retrieved Hits A	Relevant Major Hits/Retrieved Hits B	Major Precision Ratios A	Major Precision Ratios B
1	26/64	20/30	40.6	66.7	18/64	17/30	28.1	56.7
2	4/4	4/4	100	100	0/4	0/4	0	0
3	4/14	4/14	28.6	28.6	0/14	0/14	0	0
4	4/8	4/8	50.0	50.0	4/8	4/8	50.0	50.0
5	1/1	1/1	100	100	0/1	0/1	0	0
6	1/1	1/1	100	100	1/1	1/1	100	100
7	1/1	1/1	100	100	1/1	1/1	100	100
8	0/1	0/0	0	0	0/0	0/0	100	100
9	0/0	0/0	100	100	0/0	0/0	100	100
10	1/3	1/3	33.3	33.3	1/3	1/3	33.3	33.3
11	0/0	0/0	100	100	0/0	0/0	100	100
12	0/0	0/0	100	100	0/0	0/0	100	100
13	0/0	0/0	100	100	0/0	0/0	100	100
14	2/2	2/2	100	100	2/2	2/2	100	100
15	4/5	4/5	80.0	80.0	0/5	0/5	0	0
16	0/0	0/0	0	0	0/0	0/0	100	100
17	0/0	3/3	0	100	0/0	3/3	0	100
18	3/11	3/11	27.3	27.3	2/11	2/11	18.2	18.2
19	3/3	3/3	100	100	3/3	3/3	100	100
20	0/0	0/0	100	100	0/0	0/0	100	100
21	0/0	0/0	100	100	0/0	0/0	100	100
22	1/1	2/3	100	66.7	1/1	2/3	100	66.7
23	0/0	0/0	0	0	0/0	0/0	0	0
24	12/12	12/12	100	100	4/12	4/12	33.3	33.3
25	0/0	0/0	100	100	0/0	0/0	100	100
26	1/1	1/1	100	100	1/1	1/1	100	100
27	0/0	0/0	0	0	0/0	0/0	0	0
28	0/1	0/1	0	0	0/1	0/1	0	0
29	2/2	2/2	100	100	2/2	2/2	100	100
30	0/0	0/0	0	0	0/0	0/0	0	0

Query	Relevant Hits/ Retrieved Hits A	B	Precision Ratios A	B	Relevant Major Hits/Retrieved Hits A	B	Major Precision Ratios A	B
31	0/0	0/0	100	100	0/0	0/0	100	100
32	0/0	0/0	0	0	0/0	0/0	0	0
33	0/0	0/0	100	100	0/0	0/0	100	100
34	0/0	0/0	100	100	0/0	0/0	100	100
35	1/2	1/2	50.0	50.0	1/2	1/2	50.0	50.0
36	0/0	0/0	100	100	0/0	0/0	100	100
37	0/0	5/6	0	83.3	0/0	3/6	0	50.0
38	0/0	0/0	100	100	0/0	0/0	100	100
39	1/1	0/0	100	0	0/1	0/0	0	100
40	5/6	7/8	83.3	87.5	4/6	6/8	66.7	75.0
41	8/9	6/6	88.9	100	6/9	5/6	66.7	83.3
42	0/0	0/0	0	0	0/0	0/0	0	0
43	0/0	0/0	0	0	0/0	0/0	0	0
44	1/1	1/1	100	100	1/1	1/1	100	100
45	0/6	1/1	0	100	0/6	0/1	0	0
46	0/0	0/0	100	100	0/0	0/0	100	100
47	0/0	0/0	100	100	0/0	0/0	100	100
48	0/0	0/0	100	100	0/0	0/0	100	100
49	1/3	1/1	33.3	100	1/3	1/1	33.3	100
50	2/2	2/2	100	100	2/2	2/2	100	100
51	0/0	0/0	100	100	0/0	0/0	100	100
52	0/0	0/0	100	100	0/0	0/0	100	100
53	0/0	0/0	0	0	0/0	0/0	0	0
54	0/0	1/1	0	100	0/0	1/1	0	100
55	0/0	1/4	0	25.0	0/0	0/4	100	0
56	0/0	0/0	100	100	0/0	0/0	100	100
57	1/1	1/1	100	100	0/1	0/1	0	0
58	0/0	0/0	100	100	0/0	0/0	100	100
59	0/0	0/0	0	0	0/0	0/0	0	0
60	1/1	1/1	100	100	0/1	0/1	0	0
61	0/1	0/5	0	0	0/1	0/5	0	0
62	0/0	0/0	100	100	0/0	0/0	100	100

Query	Relevant Hits/ Retrieved Hits A	B	Precision Ratios A	B	Relevant Major Hits/Retrieved Hits A	B	Major Precision Ratios A	B
63	0/0	0/0	0	0	0/0	0/0	0	0
64	0/0	0/0	100	100	0/0	0/0	100	100
65	6/6	6/6	100	100	3/6	3/6	50.0	50.0
66	0/0	0/0	100	100	0/0	0/0	100	100
67	0/0	2/2	0	100	0/0	2/2	0	100
68	3/3	10/11	100	90.9	3/3	9/11	100	81.8
69	0/0	0/0	100	100	0/0	0/0	100	100
70	0/0	2/8	0	25.0	0/0	1/8	0	12.5
71	1/10	1/10	10	10.0	0/10	0/10	0	0
72	1/1	1/1	100	100	1/1	1/1	100	100
73	0/0	0/0	0	0	0/0	0/0	100	100
74	0/0	0/0	100	100	0/0	0/0	100	100
75	0/0	0/0	100	100	0/0	0/0	100	100
76	0/0	0/0	0	0	0/0	0/0	0	0
77	0/0	0/0	0	0	0/0	0/0	0	0
78	2/2	2/2	100	100	0/2	0/2	0	0
79	0/0	0/0	0	0	0/0	0/0	0	0
80	0/0	0/0	100	100	0/0	0/0	100	100
81	2/2	2/2	100	100	1/2	1/2	50.0	50.0
82	1/1	1/1	100	100	1/1	1/1	100	100
83	0/0	0/0	100	100	0/0	0/0	100	100
84	0/0	0/7	100	0	0/0	0/7	100	0
85	0/0	0/0	0	0	0/0	0/0	0	0
86	0/0	0/0	100	100	0/0	0/0	100	100
87	0/0	0/0	0	0	0/0	0/0	0	0
88	0/0	0/0	0	0	0/0	0/0	0	0
89	0/0	0/0	100	100	0/0	0/0	100	100
90	0/0	0/1	100	0	0/0	0/1	100	0
91	0/0	1/1	0	100	0/0	0/1	100	0
92	1/1	0/0	100	0	1/1	0/0	100	0
93	0/0	0/0	100	100	0/0	0/0	100	100
94	0/0	0/0	100	100	0/0	0/0	100	100

Query	Relevant Hits/ Retrieved Hits A	B	Precision Ratios A	B	Relevant Major Hits/Retrieved Hits A	B	Major Precision Ratios A	B
95	0/0	0/0	100	100	0/0	0/0	100	100
96	3/3	4/4	100	100	1/3	1/4	33.3	25.0
97	0/0	7/7	0	100	0/0	7/7	0	100
98	0/0	0/0	0	0	0/0	0/0	0	0

A score of 0/0 may have a precision ratio of 0 or 100 depending on whether there are known relevant documents. If there are no relevant documents, then it has been considered acceptable to assign a ratio of 100; where there are unretrieved relevant documents, then a score of 0/0 is assigned a ratio of 0.

APPENDIX K

DETAILS OF SEARCHING

Searching was carried out using standard boolean logic employing operators such as AND, OR, NOT and WITH.

The unindexed file in the context of this study means the database without its descriptors; the inverted file can still be searched.

In order to search the unindexed file the search statement had to be terminated with ..DE. or its equivalent. This ensured that searching did not take place in the descriptor field and was employed for S1 and S4 searches.

Multi-word descriptors for Martindale Online are hyphenated on Data-Star to prevent problems with double posting. Such descriptors were not hyphenated on the internal database and allowance had to be made for this at analysis.

Explosion was carried out by adding the character obtained by keying "SHIFT" and "3" (usually this is the £ key) at the end of the descriptor. This facility was not available on searches carried out internally.

Right-hand truncation was available.

Further details are provided in the User Guide to Martindale Online.

S1 Searches

These searches are described as unstructured searches of the unindexed file. Search statements were constructed in a similar manner to any other search on Data-Star employing logical operators and truncation, if required. The Martindale Online Thesaurus was not consulted. Each search statement was terminated with ..DE. or its equivalent.

S2 Searches

These searches are described as unstructured searches of the indexed file. The S1 search statement was used, but without ..DE.

S3 Searches

These searches are described as structured searches of the indexed file. The search statement was constructed following the guidelines in the User Guide. Briefly, the Thesaurus was consulted for appropriate descriptors for any concept in the query. If more than one descriptor was used, they were linked with the operator WITH. Searching might also be directed at some of the other paragraphs to which data had been assigned by Data-Star; this feature was sometimes used in the less structured searches to limit a large number of hits as in query 1. Searching could be directed at the descriptor field by specifying .DE. at the end of the statement or descriptor. This was not usually necessary since the WITH operator called for the the descriptors linked by it to be in the same sentence and the chance of such a co-occurrence being matched in the text is quite rare.

S4 Searches

These searches are described as highly structured searches of the unindexed file. The Thesaurus was consulted for each search to provide as detailed a search statement as possible. Descriptors were selected as terms that might appear in the text and could be used to search the database minus its descriptors. Subsidiary data such as that in scope notes, node labels and related terms were also used as sources of search terms. Hierarchies provided more specific terms (or more general terms) for incorporation into the statement which was finally constructed using boolean operators. As in S1 searches, the statements were terminated with ..DE. to ensure that searching was not carried out in the descriptor field.