

文献探索プロセスにおける適合性判定の実験的研究  
An Experimental Study of Relevance Judgments  
in Literature Searching Process

森 川 夏 子  
*Natsuko Morikawa*

*Résumé*

In the problems of literature searching process, the cases in which question contents cannot match with index terms are observed, and the author supposed some matching levels in this process. Since they were always related to relevance judgment by searchers, the relevance was especially examined. It was insisted that the judgment level of objective relevance was not observed in users having questions, but in searchers.

Examinations were carried out experimentally on librarians, students and researchers in medicine to examine whether judgment of relevant documents was influenced by difference of subject speciality (knowledge). Search questions on basic and clinical medicine were determined, and collected data were analyzed by variance analysis.

As the result, there was no significant difference among the judgment groups in the first test on thymus study. In the second test on cerebrovascular accidents study, however, a significant difference was observed. The results on the second test seems to be due to the problems in researcher group. However, further analysis revealed that judgment was rather correlated to each other among groups, and that judgment of student group was correlated more closely to that of researcher group, than judgment of librarian group was. This seemed to indicate the difference of the amount of subject knowledge. The examinations on the difference of the type of question in basic and clinical medicine using post-questionnaires, revealed that the judgment on clinical study was thought to be easier than basic study by the subjective groups, while judgment was better agreed among groups on basic study than on clinical study. Therefore, it was concluded that librarians show more similar judgment pattern to researchers and students when questions had more limited range.

Finally, the necessity of effort to aiming at more effective service through basic study of searching process like this was pointed out.

---

森川夏子：財団法人国際医学情報センター

Natsuko Morikawa, International Medical Information Center, Inc.

- I. 文献探索プロセスとその問題点
  - A. 文献探索プロセス
  - B. 照合レベルと適合性の判定
- II. 実験計画
  - A. 目的
  - B. 対象および方法
  - C. 予備調査
  - D. 調査表の設計
  - E. 判定者グループの選択とデータ分析の手法
- III. 調査結果の分析および考察
  - A. 探索主題の違いにおける判定の寛容度と個人差
  - B. 判定者グループにおける判定の一致度
  - C. 相関分析による判定者グループの関係
  - D. アンケートよりみたその他の判定要因
  - E. 予備調査の補遺
  - F. 結論
- IV. 結語

## I. 文献探索プロセスとその問題点

### A. 文献探索プロセス

これまで図書館や情報センターで行なわれてきた特定の主題に関する文献探索は、コンピュータの導入により、最近では機械検索が実際のサービス面でも可能となった。そして、機械化された大規模な検索システムの開発や実施が、今日の情報学における中心課題となった観がある。しかし、機械検索においても従来行なわれてきたマニュアルの探索を基盤としているものであるから、大量高速処理の利点とともに、マニュアル検索における問題点も持っている。

筆者は、先の論文で<sup>1)</sup>、マニュアルによる文献探索プロセスの問題点をとりあげ、検索結果に最も大きな影響を与えるものは、探索プロセスにおける質問分析の過程であることを強調した。これはまた、機械検索においても、人間の関与する部分に含まれるため、同様なことが言えると思われる。そして、実際に受けつけられた質問の概念数を調査し、質問の概念的複雑さを検討した。この調査を通して、文献探索プロセスは、今までに多くの研究者や図書館員によって説明されているにもかかわらず、実証的なデータの不足から、経験を通しての問題点の言及にとどまっていると思われた。また、マニュアル

検索は、技術的な熟練者に依存する面の強いサービス活動であると感じられた。ここでは、探索プロセスにおける人間の関与する部分に焦点をあてて、さらに検討し、均等なサービス・レベルを維持するための方策を考えるのに役立てたいと思う。

探索を仲介者を通して行なう場合には、情報要求をもった質問者が口頭もしくは書面で要求を言葉として表現し、それを受け取った探索者は、質問分析、質問内容の確認、検索語への変換、検索といったステップを行なうものと考えられる。A. M. Rees と T. Saracevic は、質問が出されて探索がなされる過程を、以下のようなステップを含むものと考えた。

1. 情報の要求
2. 質問の形成
3. 探索者による分析
4. 質問分析者と質問者との折衝
5. 質問の限定
6. 探索される概念の列挙
7. 探索概念の索引語への変換
8. 探索方法の選択
9. 探索の実施
10. 別の探索方法の検討<sup>2)</sup>

彼らは、質問分析を、システムの索引語へ変換できる

探索可能な概念を列挙する目的で行なうものと考えている。そして分析プロセスは、a) 単位概念の識別、b) 単位概念間の関係づけ、c) 単位概念のクラスの限定、d) 各単位概念の分析とその組合せ (Generic relationships, Specific relationships, Synonymous relationships, Collateral relationships について) の4段階を含んでいるとしている。そして、分析の結果得た関連概念について、その妥当性を質問者と協議することにより、探索の範囲の正しい決定を行ない得るとした。

探索されるべき概念を索引語に変換するステップでは、システムの索引語と探索の概念との間で、どちらかが特殊すぎるとか、包括的すぎるといふ問題がある。一方、索引化の立場からみれば、文献内容の索引語への変換と同様なプロセスであることは明白で、むしろパラレルな関係としてとらえることが多い。

探索プロセスの同様なステップは、D. W. King<sup>3)</sup> や松村<sup>4)</sup> によっても明らかにされているが、各ステップに不備のある場合は、結果がゆがめられたり、間違ったりすることになる。松村はさらに、機械化システムでは、マニュアル検索に比べて、探索の途中で利用者(質問者)との接触が不可能であり、一度検索式をたてれば、“一回チャンスの検索方式”<sup>5)</sup> をとることになると指摘し、システムと利用者との相互作用の重要性を、メドラス実績評価の結果に基づいて強調している。

実際の探索において生ずる問題点は、筆者の文献探索の経験から考えてみると、質問内容と探索に用いるツール、探索者としての図書館員の三者に関連していると思われる。質問内容の問題点は、特殊な主題で研究が開始されたばかりの概念の不明確な場合や、詳細な研究の多くなされた分野でも、諸説があって概念規定が困難な場合などである。このような時、探索者は質問分析のレベルで、内容の複雑さや難解さに困惑することになる。そして多くの場合、これが反映されて用いるツールの索引語についても的確な照合ができず、包括的な概念あるいはより特殊な概念の組合せによる探索方法をとらなければならない。したがって、ツールとしての二次資料の構成、収録範囲などが検索結果を左右することになる。

また、探索者が熟練した専門家であれば、二次資料についての探索技術を駆使して、利用者にとって有効な文献を検索できるが、それも人間の判断が関与するものであるため、個人的差違が生ずることは否めないし、また探索技術のある程度習得していても、経験の浅い探索者

では、明らかに十分満足できる結果を得られるとは保証できない。さらに機械検索と異なり、マニュアルの作業では、包括的な概念の質問の場合、関連語の広がりが大きく、数多くの索引語のもとで検索することになり、実際の探索面で時間と手間とがかかりすぎる。

一方、質問者が求める探索主題について、期待した文献数と異なる結果となれば、さらに探索方法に検討を加える必要がでてくる。R. T. Tritschler<sup>6)</sup> は、質問者が検索文献ゼロの場合も、多すぎる場合にも、検索結果に満足しないことを指摘して、確率的な適合度の表示で結果を示すことを提案している。

## B. 照合レベルと適合性の判定

以上のような問題点は、それぞれのステップで考察すべきだが、ここでは実際の探索プロセスの流れを追って、さらに検討してみることにしたい。そこで、*Index Medicus* を例に、マニュアルによる探索を考えた場合、たとえば「妊婦における虫垂炎の統計的研究」という質問では、MeSH により、APPENDITIS の件名による one term search になると予想される。その件名のもとで、さらに副標目の occurrence と mortality が組合わされる。妊婦という限定項目は、リストされた文献の標題を参照することで検索される。しかし索引誌の限界として、標題に妊婦を明示する同義語ないしは関連語が見い出せない場合には、探索はその段階より詳細な検討をすることができずに終わる。ある場合には、さらに著者名や雑誌名によって、何らかの関連性を見つけることができれば、検索すべき文献がよりの確に選択されよう。抄録誌の場合には、さらに抄録文の検討によって、より多くの検索文献を見い出すことが可能である。ただし、MEDLARS の機械検索においては、check tag の PREGNANCY の指定により、きわめて妥当な検索を行なうことができる。

このような探索例から、文献探索プロセスにおいて、質問の概念がうまく検索語に変換できない場合についても、類推することができよう。つまり、検索は質問内容の構造いかんで、主件名のレベルで照合されるものから、主件名と副標目との組合せのレベル、さらに標題のレベル、書誌的事項を含めたレベルといった、いくつかのレベルでの照合が考えられる。これをまとめると、探索者が関与する探索プロセスにおける照合は、以下の点で行なわれるものと考えられよう。

### 1. 質問分析—質問に含まれる探索概念の識別

↓

2. 検索語の決定—探索概念に対応するトウールの索引語の選択
- ↓
3. 副標目の指定—索引語の組合せ
- ↓
4. 標題の吟味—探索概念を表わす keyword の有無
- ↓
5. 雑誌名の検討—特殊な主題を扱う雑誌の場合には有効
- ↓
6. 抄録の参照
- ↓
7. 原著論文

このような照合レベルは、すなわち質問に対する文献の適合性を決定するレベルであるともいうことができる。したがって索引側の問題点は無視しても、照合レベルの違いによって、検索文献の判定が大きく左右されるものと考えられよう。

そこで、探索技術を習得した探索者は、うまく探索概念を検索語に置き換えられる質問に対しては、コントロールされた索引システムのおかげで、信頼度の高い検索が可能であるが、標題あるいは抄録のレベルまで詳細に検討しなければ、質問との適合性を判定できない、その索引システムにうまく合致しない場合には、その分野の主題知識を背景に持つ必要があると思われる。つまり、コントロールされたシソーラスのように、関連概念の構造が明らかな場合を除いて、主題に対する知識がないと、その主題領域に関わる専門用語の関係を把握できないことになる。言いかえれば、標題や抄録のレベルで照合を行なう際にも、実際面では手がかりとなる探索概念を示す keyword の有無によって、検索すべきか否かが決定されているにすぎない。

それでは、何に基づいて文献コレクションの中から、検索すべき文献を決定しているのだろうか。その基盤となるものは、適合文献を的確に検索するための大きな要因であると考えられる。したがって適合性の判定そのものが何であるのかについて、検討する必要がある。それは、文献探索プロセスの照合レベルすべてに関わる、人間の判断力の問題であり、これについて明らかにすることによって、探索プロセスをより一層明確な形でとらえることができると思われる。

情報システムにおける根本的な目標は、利用者に対して、最も適切な情報を提供することであると言われている。文献探索プロセスにおける各ステップは、まさに常に質問との適合性についての照合の作業であると言え

る。それでは、適合性とは何についてのどのような概念を指すのであろうか。アメリカ合衆国を中心としたドキュメンテーションの分野では、‘relevance’という語がこれに対応して用いられている。ここでとりあげた適合性の問題は、したがって、この relevance の概念に合致するものと考えられる。relevance の研究は、アメリカでは数多くの論文が発表され、T. Saracevic<sup>7)</sup>によって歴史的な観点よりまとめられているが、日本においてはわずかに天野に<sup>8)</sup>よる研究があるにすぎない。最近の研究<sup>9) 10) 11) 12)</sup>では、relevance の概念があいまいに用いられている現状を批判し、より厳密な規定の中で用いられる概念として、明らかにしようとしている。しかし一般に文献レファレンスを提供するとされる reference retrieval system をとりあげず、question-answering system や実際の yes-or-no 形式の情報システムを想定して議論しており、文献レファレンスにおける relevance の明確な定義はむずかしいと考えられていることがわかる。また適合性の概念を用いて、文献検索を行なう場合の判定は、その判定の主体となる人間によって異なることが、さらに問題となるが、information need とか utility といった利用者の立場を主体とした考え方から離れて、客観的な意味合いをもつ概念として relevance を規定していると解釈できる。

一方、これまでの情報検索システムの評価のための規準として考えられていた relevance は、明らかに利用者側のものであった。たとえば F. Lancaster による MEDLARS の実績評価実験<sup>13)</sup>では、適合文献を利用者の information need にとって価値ある文献とし、適合性の評価は価値判断であるという定義に基づいて効率の評価を行なっている。一般には、したがって情報検索システムの性能評価を行なうために用いる判定規準としては、質問を出した利用者自身の評価が採用されている。その問題点としては、

“…利用者にとって、主題だけが適合性に関する要因ではない。詳しくすぎるとか簡単すぎる場合、専門的すぎるとか一般的すぎる場合、理解しやすい言語であるとか難解な言語である場合など、いずれも適合性の判断に影響を与えるはずである。つまり、適合性というのは情報の要求者の主観的評価に基準をおいているところに、評価基準としては妥当性に欠けるところがある…”<sup>14)</sup>と長沢によって指摘されている。

したがって、適合性の判定をより客観的なものとし、また安定性の高いものと規定するには、利用者自身では

ない第三者によって、質問に含まれる概念と、文献内容を示す概念との合致度を決定することが、適合性の判定を意味すると考えた方がよいだろう。

この考え方に近いものとして、天野は *relevance* の概念を定義して、“質問と記事とが、それぞれの内容に同一主題を有していて、主題を構成する要素に多少の違いはあっても、その関係に近似性があるもの”<sup>15)</sup> としている。そして *relevance* の段階的なとらえ方を、関連性とともて説明し、文献内容の要素（概念ないしトピック）を個別的にとり出し得るならば、それと質問内容との照合が、適合性の判定であると考えているので、適合性の判定はより客観的なものと思われる。

以上のことから、*relevance* の概念があいまいで、*pertinence* や *utility* と混同されて考えられていることに対し、筆者はより明確な規定として、探索者の立場から、*relevance* の概念は客観的な意味合いの強いものにすべきであると主張したい。

適合性の判定は、また質問と文献との内容に関連した人間の判断力の問題であるが、その一貫性を支持すること、すなわちどの判定者にとっても、同じような判定結果が可能であるとするのは、客観的な決定と同じことを意味する。たとえ質問を出した利用者自身にとって、実際に使える有用な文献ではないとしても、概念を言語表現した語の照合において、適合文献とされるものは存在し、探索者側ではその決定までは可能はずである。

したがって、適合性の判定を主観的で不安定なものであるという批判に対し、概念の照合における客観的なレベルが存在することを主張しておきたい。さらに *relevance* の概念は、質問とファイル中の文献との関係であると規定されるならば、質問者自身の心中の変化や価値観とは切り離された客観的な世界、言語に定着された世界における概念であると考えべきであろう。このように適合性の判定に関わる問題は、複雑な様相を呈するよう思われるので、ここでは適合性の判定における照合レベルに関連して、判定者の主題知識の有無が与える影響について、実験的な試みを行ない、検討を加えてみることにしたい。

## II. 実験計画

### A. 目的

文献探索サービスにおいて、経験の浅い図書館員が遭遇する困難な点は、個々の二次資料に対する探索技術の面での不備とともに、探索主題についての概念認識の不

明確さがあげられる。後者の場合、探索主題に関して、その主題専門家（研究者）に比べると、主題知識が乏しいばかりでなく、その分野の *state-of-the-art* を把握していない場合も少なくないことが、その大きな理由と思われる。しかし、ベテランの図書館員では、長年にわたる経験の積み重ねによって、ある一定のレベルに達しているものと考えられる。また図書館員であれ、主題専門家であれ、特定の主題について適合文献を探索する際には、つねに適合性判定という人間の思考過程が関与しているので、より一層むずかしい問題を含んでいる。すでに述べたように、その過程を明らかにすることは、どの探索者によっても同等なサービスが可能になるような方策を検討するために、不可欠なものと言えよう。

したがって、ここでは以上のような問題意識に基づき、専門の主題知識の有無が文献の適合性判定において、重要な要因であるかどうかを実験的方法を用いて調査することとした。すなわち図書館などで実際に行なわれている文献探索プロセスにおいて、探索主題と文献との間の適合性判定に対し、主題知識の有無がどの程度、判定に影響を与えているのかを明らかにしようとするものである。

これまでなされた実験的研究の結果や文献探索の経験的事実を通して、図書館員および研究者が適合文献を判定する際にとる行動パターンとしては、以下のことが考えられよう。図書館員は、探索経験も少なく系統立った主題知識について学習していない場合には、適合文献の重要なものを落としてしまう可能性があり、一方そのような失敗を防ぐために、探索範囲を広げ不適合文献（ノイズ）の多い探索を行なうことになる。研究者自身が探索する場合では、特に専門領域として細分化された特殊なテーマでの研究に関しては、ひじょうに厳密な判断がなされ、適合文献の判定も狭い範囲に限られてしまう。また一方でテーマに少しでも関連ある文献すべてを把握しようとする場合には、多くの文献を適合と判定するものと考えられる。このように、おのおの相反する二面があるため、行動パターンとして一義的に決定できるものではない。しかしここではそのような判定の傾向をとらえるのではなく、主題知識が判定の際に大きな要因として作用しうるものかどうかを検討してみることにしたい。

### B. 対象および方法

文献探索サービス活動を専任者において積極的に行なっている情報サービス機関は、日本においてまだあまり多くないが、財団法人国際医学情報センターでは、慶応

義塾大学医学情報センターの業務を引き継いで、医学分野における文献探索サービスを行なっている。当財団は、医学分野の探索サービスを行なう機関として人員も最も多く、処理件数も年間 700 件を越える現状であることから、財団における文献探索担当者を対象の中心とし、以下の 3 グループを設定した。第一に図書館員グループとして当財団の文献情報課で文献調査サービスを担当する者を、第二に主題の専門家である研究者グループとして慶応大学医学部内の研究者を、そして第三に医学の全般的な基礎知識を学んでいる医学生グループをその中間層として調査対象に含めた。

文献探索プロセスについて考察した時には、適合性についての照合レベルを考えたが、ここでは探索主題に対する文献の標題および書誌的事項のレベルでの判定に注目して、調査を行なうことにする。探索プロセス全体を包含する調査方法、すなわち質問の提示から探索手段（二次資料の種類など）の選択、探索語の決定、そして検索という過程をすべて含めると、各段階で探索技術面の要因が多く関わってくるため、主題知識を背景とした判定の調査にはならない。また調査対象とした各グループに対し、そのような探索実験を実施することは、実際上不可能であった。したがって最も簡単で確かなデータを得るために、あらかじめ設定した探索主題に対し、判定の材料としての文献リストを作成し、調査表の形式にまとめて調査することにした。以上の方法をさらに吟味するために、簡単な予備調査を行なって、方法決定の指針とした。

### C. 予備調査

調査の方法と調査において判定者に課せられる負担の度合や時間的制約などを検討する目的で、1973年6月から8月にかけて予備調査を行なった。探索主題は実際に受けつけられた質問「低酸素の下における心機能および血行力学的変化について」というもので、概念規定がむずかしいため、どの範囲までを検索すべきか判断しがたく、二次資料の探索においては主件名と副標目との組合せや keyword による単純な照合では、うまく探索できないものであった。判定材料として用いる文献レファレンスは、二次資料を用いて過去2年間について探索し、日本語・英語各 25 文献ずつ計 50 文献を和英別にアルファベット順に並べて文献リストとし、調査表の形式にまとめた。判定者は、図書館員9名、医学生（専門課程第4学年）9名計 18 名で、図書館員の平均在職期間は約2年1カ月となっている。

また実際の探索プロセスでは、図書館員が主題知識の不足を補うために、辞書・事典類を調べることが多いが、そうしなければ適合性の判定ができないという実際上の問題から、現場のサービスの範囲内という条件で調査することとし、参考書類の参照を制限しないものとした。

その結果、第1図に示されるように、図書館員と医学生の間では判定において大きく違ってくるとは認められない。文献ごとに多少の差はあるが、かなり相似的な判定を行なっていることがわかる。また図書館員の方が、医学生に比して適合文献と判定する割合が高く、各文献の判定の一致度をみても、適合とするものが多いことが認められた。

調査の実施に際して、図書館員の場合は、業務の合い間に協力を依頼したため、辞書類の参照を含め、短時間の面接方式をとることは困難であった。そのため実際のサービス・レベルでの判定であるという指針を与えるだけで、調査表記入方式による調査となった。一方、医学生では可能な場合には直接に面接を行なって、回答を得ることができたが、一部は調査表配布による協力依頼で回答を得た。

判定の所要時間は、医学生が平均約 24 分で、図書館員は約 68 分と 3 倍に近い時間を要した。これは、50 文献の中に日本語と英語を同数含めたため、日本語文献に比べ英語文献では辞書を必要とする場合が多く、特に図書館員にとってはそれが判定の際に多大な時間を要した原因ではないかと考えられた。また短時間の調査としては、文献数 50 は多すぎるという意見が多かった。さらに図書館員・医学生ともに、英語文献が 25 も続くことに負担を感じることを指摘している。

提示した探索主題に対しては、基礎的な実験を含めるものか、臨床の立場によるものだけであるのかといった観点のあいまいさによる疑問が多く出された。そこで医学分野においては、基礎的研究と臨床といった2つの立場の違いが、適合文献の判定にどう関連するかが問題となろう。

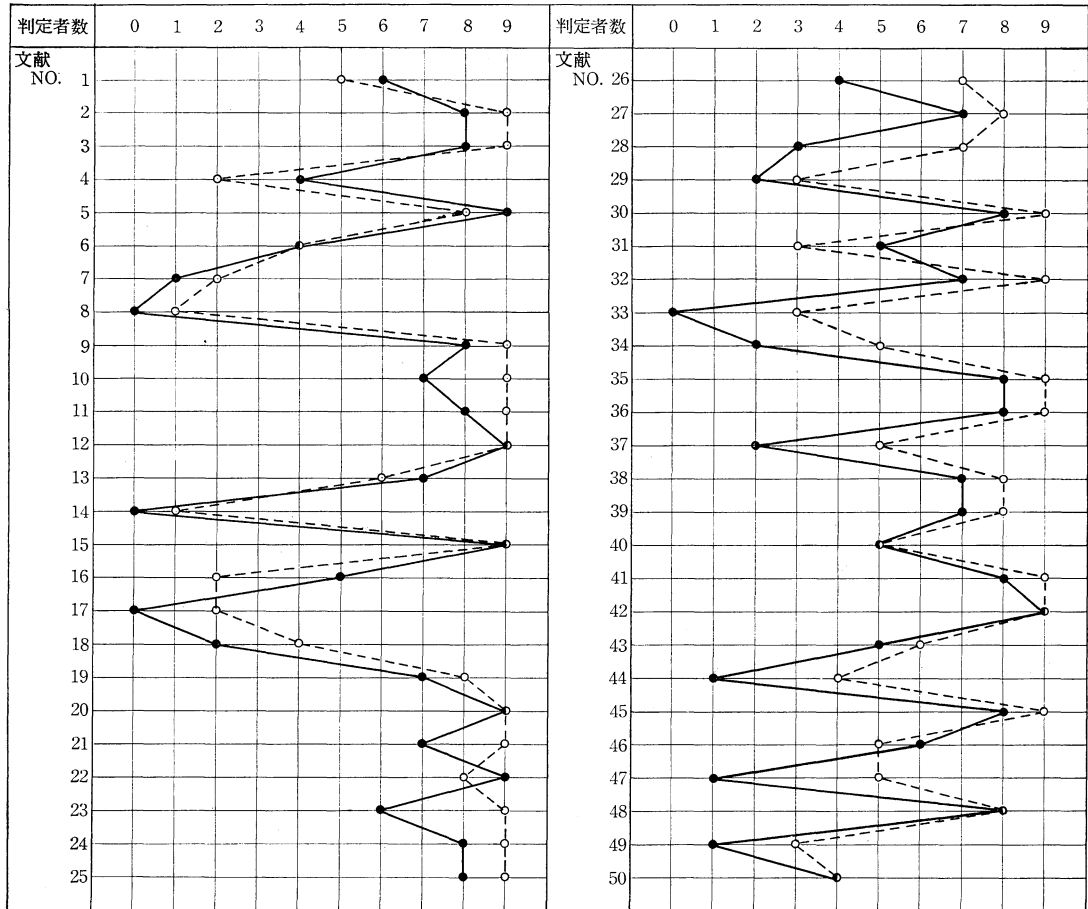
以上述べたように、この方法によって得られるデータは、かなり目的としたものをとらえることができるが、実際の調査に際しては、次のような点を考慮することにした。

1. 探索主題は、基礎的研究と臨床面の研究の2つの側面を仮定する。
2. 文献数は減らし、英語文献が判定の時に負担とならないようにする。

第1図 各文献に対する適合判定者数

●医学生

○図書館員



3. 調査の際、図書館員には調査表を配布して記入する方式をとるが、他のグループには、できるだけ面接方式で調査の意図を誤りなく伝えることに努める。
4. 判定規準としては、図書館員の場合業務として行なうサービス・レベルの範囲内ということ、医学生・研究者では主題に対する客観的判断ということを示す。
5. 判定は○×式の二者択一方式であったが、標題と書誌的事項だけの文献リストについての判定であるという制限上、あまり詳細な尺度を設定することは好ましくないとと思われるため、そのままの形を用いる。

#### D. 調査表の設計

本調査における探索主題の決定については、以下の条件を満たすものを検討した。

1. 探索主題は2つ以上の複合概念（概念の組合せ）から成ること。
2. Keyword すべてのリスト・アップが簡単にはできない包括的な概念を含むこと。
3. 探索の際に一般に用いられる件名を使う照合だけでは十分でないような特殊性を有すること。
4. 探索主題に関する実際の文献数が、極度に少ないと予想されるものではないこと。（判定材料とする文献リスト作成のため）
5. 調査対象となる研究者の協力が得やすいものであること。

ること。

このような条件に加えて、実際に国際医学情報センタに申し込まれる質問や文献探索のベテラン担当者の意見を参考に、以下の2領域における4主題を設定した。

- A 胸腺細胞の免疫能
- B 胸腺腫の病理と診断
- C 脳血管障害の生化学的研究
- D 脳卒中の治療および予後

なお A・C は概して基礎医学的な観点を中心とするもので、B・D は臨床面が中心であると仮定した。

判定材料とする文献レファレンスは、二次資料から抽出することとし、主として *Index Medicus* と「医学中央雑誌」により、胸腺研究と脳血管障害という2つの領域全体について、まず探索を行なった。調査対象とした研究者は、進展しつつある学問領域に携わる人間であるので、判定材料として用いる文献はできるだけ最新のもので、研究者がまだ目を通していないものが望ましいが、胸腺研究の領域では文献が広範な分野に出現する可能性があって、物件索引の使用が必要であるため、日本語文献の探索に用いる「医学中央雑誌」のタイム・ラグの関係上、過去2年間にさかのぼらざるを得なかった。したがって英語文献もほぼ同時期の文献を集めた。一方、脳血管障害の領域はひじょうに多くの研究がなされ、文献数も多いが、限られた分野での探索が可能であるため、できるだけ最新文献を集めることに努めた。

こうして集めた文献レファレンスのうち、以下の項目を検討して判定材料の文献を選んだ。なお文献数は予備調査における検討から、多くなると判定者の大きな負担となるため、各主題ごとに和英含めて15文献とする。

1. 標題が短かすぎるものを除く。(Keyword とされる一語のみで表現される場合などがある) e. g. 脳卒中に関する研究
2. Keyword そのものが標題であるものは除く。e. g. 脳卒中の集中治療
3. 同じような表現の標題の文献は、重複してとらない。
4. 各種の雑誌に掲載された文献をとりあげる。
5. 同一著者の文献は重複してとらない。
6. 明らかに適合しないと判定される文献も少し含める。(件名の組合せから考えて、客観的に不適合とされる文献もとりあげる。)

これらの項目は、実際問題として図書館員が文献レファレンスだけを手がかりとする時、判断に差違の生ずる可

能性のある場合と考えてよい。そうした状況を設定する目的で、この実験的調査ではこれらの項目をあげている。

したがって、ここで作成する文献リストは、各主題に関する全文献(母集団)のサンプルというのではなく、むしろ判定においてばらつきが生ずる可能性の大きい文献を集めたものといえることができる。また筆者自身が二次資料を探索して得た結果であるため、筆者の判断が bias として加わっていることは否めない。しかし、ある主題分野に関係をもつ文献全体の中から何パーセントを適合文献と判定するのかといった数値の大小について論ずるものではなく、ここでは個々の文献が、判定者の違いによってどう判定されるかについて、相対的な差を調べようとするものである。それでこのような方法で文献を選ぶことは、調査結果に大きな bias をかけるものとは考えていない。

各15、総計60の文献は、和英別にして著者のアルファベット順リストにした。それを以下の項目から成る質問表形式の調査表にまとめた。

1. 調査の目的と方法の説明
2. 文献リスト
3. 回答用紙
4. Post-questionnaire

本調査でも所要時間の記入欄を設け、さらに文献判定に関わる他の要因との関連から、アンケート用紙を最後に加えた。質問項目は次のようなものである。

#### 1. 文献内容の既知度

研究者の場合、その主題についての専門家である限り、多くの文献を通読通覧している可能性がある。それに対し、初めて文献の標題やレファレンスを知って適合性の判定を行なう場合では、判定条件が違ってくると考えられるので、どの程度既知の文献が含まれているかを調べる。

#### 2. 文献判定におけるレファレンスの参照

一般的な文献リストの形では、標題と共にレファレンスが付け加えられるが、判定の際著者の研究内容を知っていたり、日頃目を通す雑誌である場合に、著者名・雑誌名を参考に判定することが考えられる。

#### 3. 英語と日本語の文献の違い

一般には、英語文献の標題の方が日本語文献より詳細で内容をよく表現していると言われている。今回の調査では、そのような差を判定者自身が認めているかどうかを考慮に入れる必要がある。

#### 4. 探索質問としての基礎と臨床



基礎と臨床という2つの観点から主題を決定したが、この間に判定における差があるかどうかを調べる。

#### E. 判定者グループの選択とデータ分析の手法

財団における文献調査サービス担当の専任者および兼任者の15名（平均年齢25.1才，平均在職期間2年）から調査協力が得られたため，他のグループも各15名とした。研究者は慶応大学医学部内で協力を依頼し，胸腺研究の領域については，医学部内に「胸腺・免疫異常症研究会」が存在するところから，そのメンバーを中心として選んだ。その内訳は，病理学4名，内科6名，外科5名（平均年齢34.5才）である。また脳血管障害においては，その領域に関わる研究者と考えられた内科循環器系および脳外科を中心として，基礎の立場からも医化学の研究者を含めて，医化学5名，内科5名，脳外科5名（平均年齢37.6才）となっている。また医学生グループは専門課程第4学年（平均年齢24.5才）15名で，個別的な協力依頼による。

以上のような選択を行なっているので，判定者の抽出は統計的手法である無作為抽出法によるものではない。しかし研究者の全体数をつかむこと，すなわち各主題の専門家と言うべき研究者総数の把握は，ひじょうに困難であるし，さらに図書館員の場合では，探索活動を専門とする担当者は，日本ではまだ決して数多いわけではないために，統計的手法によるサンプリングは実際上不可能と思われた。判定者グループとしての図書館員と医学生は共に2つの領域を2回に分けて判定した。研究者の場合は，主題別に各15名計30名となっている。

集められたデータは，判定者ごとに，各主題についてリストされた15文献のうちの適合文献数をまとめたものと，各文献ごとに，判定者グループ15名のうち適合とした判定者数を示すものとに分けて集計した。前者は，判定者個人が文献を適合とみなす判定基準の違いを示すもので，探索主題に対しどの範囲までを適合するものと評価しているかに関連し，判定の寛容度として示される。後者は，個々の文献に対し判定者グループ内での程度判定が一致しているかを示すもので，判定者グループの判定の一致度として表わすことができる。さらに全体を総計して考えれば，個々の文献の適合度としても解釈することができよう。

ここで設定した判定に影響を与える要因としての主題知識の差は，図書館員，医学生，研究者の3グループにおける相違としてとらえられよう。そこで適合文献の判

定に対し，判定者グループ間で相違があるかどうかを分析し，また2回にわたって基礎と臨床という側面から調査を行なっているので，探索主題によっても判定のパターンに影響を与えることが考えられるため，探索主題における相違も分析してみる。

このような観点から，判定パターンに大きな違いがあるかどうかを分散分析法によって検定し，さらに各文献の適合度とも考えられるグループごとの判定の一致度から，図書館員，医学生，研究者の相互間の相関を考えてみることにしたい。判定がどの文献に対しても相対的に同じようになされるものと仮定すると，相関係数は高くなるが，相関によってグループ間の判定パターンの関係がよりいっそう明らかなものになると思われる。

### III. 調査結果の分析および考察

#### A. 探索主題の違いにおける判定の寛容度と個人差

調査は1973年9月から10月にわたって2回行ない，判定者グループ別にまとめると第1表のようになった。ここでは，要因Aを探索主題A・B・C・Dとし，要因Bを判定者おのおの15人として分散分析を行なった。探索主題A・BとC・Dは，2回に分けて調査を行なっていること，また研究者グループでは専門領域の違いから異なる判定者によって判定されているために，分析ではまずA・CとC・Dに分けて行ない，さらに図書館員と医学生では，4主題をまとめたものについても分析を行なった。

その結果，まず図書館員グループでは，探索主題による適合文献数の判定に有意差はないが，15人の個人差は有意であることが示された。つまり主題の相違によって，判定者の判定の寛容度が大きく変わることはないと考えてよいが，各個人の相違をみると，その寛容度には差があることになる。

医学生においては，探索主題に有意差はない。また個人差については，A・BとC・Dに分けて分析した結果では，有意差はないと判定されたが，4主題をまとめて分析した場合に，高度に有意であるという結果になったため，医学生においては，主題A・BとC・Dとの間に，判定の差があるとみなすことができよう。

研究者の場合では，主題A・Bを判定したD1グループにおいて主題による差はないが，15人の間の個人差は高度に有意であるという結果になった。主題C・Dに対するD2グループでは，主題における有意な差があり，判定者15人の個人差は有意でなかった。

文献探索プロセスにおける適合性判定の実験的研究

第1表 判定者別適合文献数

a. 医学図書館員における適合文献数

ML	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	AVER.	SD
A	10	8	10	10	15	9	9	15	8	11	7	14	9	10	11	10.40	2.39
B	11	12	12	9	12	9	12	13	9	11	8	14	9	9	13	10.87	1.82
C	10	9	7	12	14	14	11	13	13	9	10	13	10	9	12	11.07	2.05
D	9	4	11	6	10	13	10	15	13	10	8	12	6	9	12	9.87	2.90

b. 医学生における適合文献数

MS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	AVER.	SD
A	14	9	10	8	10	10	11	5	10	10	10	12	12	10	14	10.33	2.15
B	15	14	9	7	11	10	8	8	5	13	12	6	13	10	14	10.33	3.03
C	13	10	8	7	11	8	9	6	10	6	8	10	11	6	12	9.00	2.16
D	12	11	11	4	11	8	8	6	6	11	8	6	11	12	9	8.93	2.48

c. 医学研究者における適合文献数

MD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	AVER.	SD
A	10	10	12	15	7	13	11	14	11	12	8	10	9	7	7	10.40	2.44
B	15	14	15	15	5	15	15	14	8	11	10	10	9	10	6	11.47	3.38

MD	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	AVER.	SD
C	7	8	10	13	10	8	14	8	6	11	10	14	10	12	10	10.07	2.35
D	6	13	12	13	10	14	15	11	15	8	15	11	10	14	13	12.00	2.58

これまでの結果をまとめて考えると、図書館員や医学生の間では、探索主題が基礎的なものから臨床的なものへと変わることによって、判定の寛容度（15文献に対する適合文献数）が大きく変わることはないが、研究者グループにおいては、判定の基準が変わる可能性があることを示していると言えよう。

しかしD1グループでは有意差を示さないが、D2グループで差をみとめたことから考えると、その原因として次のことが言えるかもしれない。それは、第1回目の調査における胸腺の領域に関しては、胸腺を中心とした研究グループがあらかじめ把握でき、その中から判定者を選んだのに対し、第2回目の脳血管障害の分野については、明確な研究グループではなく、関連分野と考えられる研究者を対象としていることである。

探索主題ごとにまとめて、判定者グループの平均適合文献数をみると、医学生がどの主題においても最も小さい値を示している。そして標準偏差では、Dを除いて、

研究者グループが最も大きく、判定された適合文献数のちらばりという点では、図書館員や医学生に比べて大きいことが示される。また2回に分けた調査における基礎的側面と臨床面については、A・B間における図書館員の判定を除いて、概して基礎の方が臨床よりも標準偏差が小さくなっている。（第2表参照）

B. 判定者グループにおける判定の一致度

各文献ごとに与えられた適合とする判定者数をまとめたものは、主題別に第3表のようになった。さらに判定者グループ全体の総計は、文献の適合度とみることができるので、判定者数の多い順にランクづけ、その中での判定者グループにおける判定結果をグラフ表示した。（第2～5図）このグラフで明らかのように、総数によるランクづけを行なった場合、3つの判定者グループの間では、主題Aにおいてかなりパラレルな判定パターンになっていることがわかる。しかしB・C・Dでは、判定は相対的なパターンを描かず、グループ間での適合

第2表 適合文献数の判定者グループによる相違

ML: 図書館員 MS: 医学生 D1, D2: 研究者

## [A] 胸腺細胞の免疫能

	AVERAGE	SD
ML	10.40	2.39
MS	10.33	2.15
D1	10.40	2.44

## [C] 脳血管障害の生化学的研究

	AVERAGE	SD
ML	11.07	2.05
MS	9.00	2.16
D2	10.07	2.35

## [B] 胸腺腫の病理と診断

	AVERAGE	SD
ML	10.87	1.82
MS	10.33	3.03
D1	11.47	3.38

## [D] 脳卒中の治療および予後

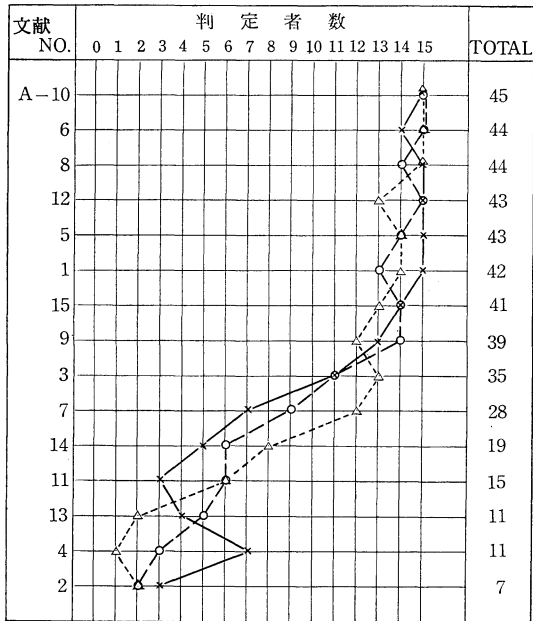
	AVERAGE	SD
ML	9.87	2.90
MS	8.93	2.48
D2	12.00	2.35

第3表 探索主題別による適合文献判定者数

	[A] 胸腺細胞の免疫能														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ML	15	3	11	7	15	14	7	15	13	15	3	15	4	5	14
MS	14	2	13	1	14	15	12	15	12	15	6	13	2	8	13
D1	13	2	11	3	14	15	9	14	14	15	6	15	5	6	14
TOTAL	42	7	35	11	43	44	28	44	39	45	15	43	11	19	41
	[B] 胸腺腫の病理と診断														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ML	15	14	9	11	15	8	12	15	6	5	5	14	6	14	14
MS	13	11	11	15	8	7	6	14	8	5	9	14	8	13	13
D1	13	11	10	15	7	9	9	11	12	10	11	15	13	14	12
TOTAL	41	36	30	41	30	24	27	40	26	20	25	43	27	41	39
	[C] 脳血管障害の生化学的研究														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ML	6	15	14	8	14	6	15	9	9	8	12	15	10	12	13
MS	2	15	6	9	14	6	13	9	4	8	12	10	4	15	8
D2	4	15	11	11	13	7	14	9	6	6	10	13	7	14	11
TOTAL	12	45	31	28	41	19	42	27	19	22	34	38	21	41	32
	[D] 脳卒中の治療および予後														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ML	13	15	12	10	8	6	13	8	9	8	11	2	14	10	9
MS	13	13	9	8	11	6	8	6	8	12	11	1	12	12	4
D2	13	12	13	13	13	12	11	10	13	14	14	4	14	13	11
TOTAL	39	40	34	31	32	24	32	24	30	34	36	7	40	35	24

文献探索プロセスにおける適合性判定の実験的研究

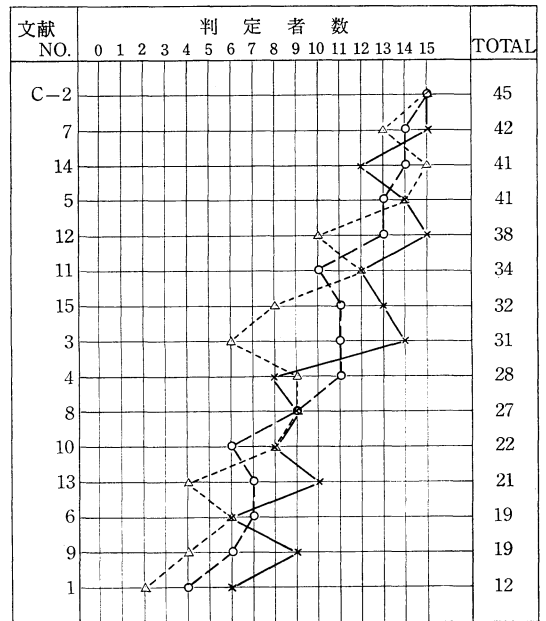
第2図 探索主題 A における適合度ランク



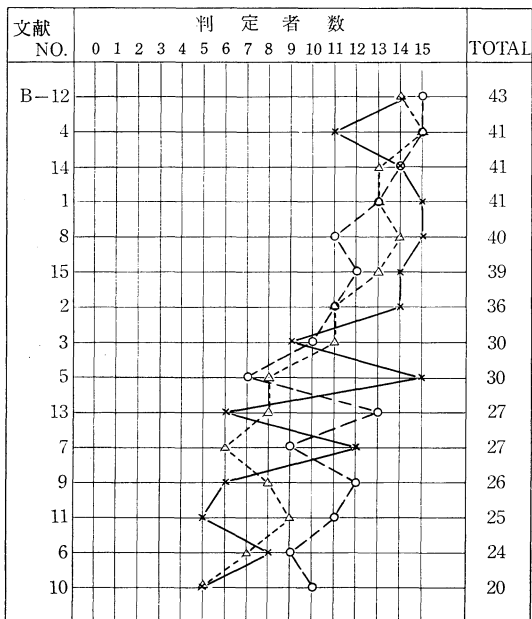
×図書館員 △医 学 生 ○研 究 者

(注) TOTAL が同数の場合は、研究者グループにおける数値の高いものを上位にランクづけた

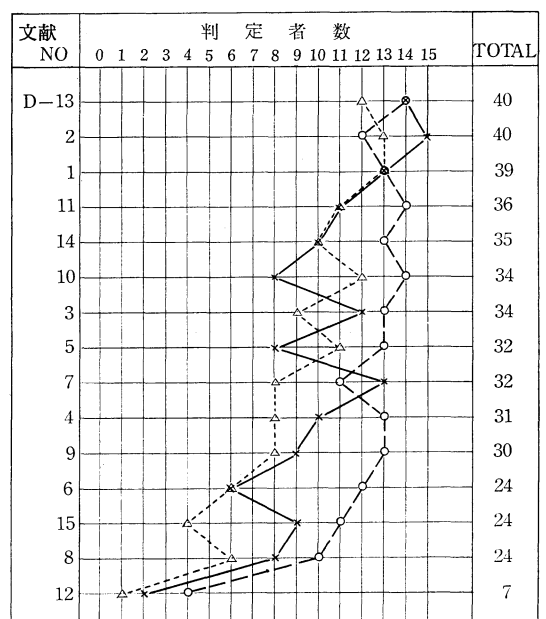
第4図 探索主題 C における適合度ランク



第3図 探索主題 B における適合度ランク



第5図 探索主題 D における適合度ランク



度の差（ここでは判定者数の相違で示される）がある。

ここではグループごとに文献単位で、15人の判定者の一致度が示されるので、各主題における判定者グループ間に相違があるかどうかを分析した。要因Aを、判定者グループ ML, MS, MD とし、要因Bを15文献のおのおのとして、分散分析した結果、主題AとBにおいて判定者グループ間に有意差はなかった。一方、Cでは5%の有意水準で有意であるとされ、Dでは1%の有意水準で高度に有意であるという結果になった。

### C. 相関分析による判定者グループの関係

判定者グループ間における判定の一致度に関しての相関関係を明らかにする目的で、図書館員 vs. 研究者、図書館員 vs. 医学生、研究者 vs. 医学生の3つの組合せについて、4主題それぞれの相関係数を求めると第4表の

第4表 判定者グループ間の相関係数表

主題	ML vs. MS	ML vs. MD	MC vs. MS
A	0.864	0.940	0.943
B	0.626	0.109	0.688
C	0.658	0.848	0.871
D	0.724	0.628	0.793

ようになった。相関係数の検定では、B主題における図書館員 vs. 研究者の場合に、5%の有意水準で無相関となった。その他には、相関関係があるとみなされた。基礎と臨床という観点よりみると、C・D間の図書館員 vs. 医学生を除いて、すべてが基礎における相関係数がより高い値を示した。それで各グループ間での判定は、かなり相対的なものであり、臨床の主題より基礎の方が判定における一致度の差が少ないと言える。さらに、図書館員と研究者の相関より、つねに研究者と医学生の相関が高いことが明らかにされた。

分散分析による有意差検定では、大きな差をみとめなかったけれど、相関分析においてグループ間の関係をみた場合には、判定のしかた（一致度）は、専門の研究者に対し、図書館員よりも医学生の方が近い判定を行なうのではないかと考えられる。これは、実験計画の当初に考えたように、より専門的知識の量的近似性を反映したものといえよう。

### D. アンケートよりみたその他の判定要因

判定実験のあとに付け加えたアンケートについては、以下のような結果が得られた。

#### 1. 文献リスト中の既知文献

第1回目の胸腺領域において、図書館員に1名だけ、業務上最新文献の現物チェックを行なっているために内容を知っている文献があると答えた者があるが、他は初めて知ったと答えている。医学生では、胸腺に関心をもつ学生が4名、いくつかの文献を知っていたと答え、一方研究者グループ（D1）では、そのほとんどがいくつかは知っていると答え、半分ぐらいと回答した者も1名あった。

第2回目の脳血管障害の領域では、図書館員は全員が初めて知ったと答え、医学生では1名だけ、いくつか知っていると答えた者がある。研究者グループ（D2）では、大部分を知っていたとする者1名、半分と答えた者2名、いくつかは知っていたとした者5名となったが、D1グループに比べると初めて知ったとする者が多くなっている。これは、第2回目の領域で基礎領域の研究者として、医化学の専門家を選んだので、日頃脳血管障害に直接関与しているわけではないことによるものと思われる。

#### 2. 著者名・雑誌名の参照

2回の調査の両方で、図書館員の半数以上が雑誌名を判定の際に参照したと答えている。医学生では、1～2名が著者および雑誌名を参照したと答えたにすぎない。研究者では、3分の2程度が、著者名や雑誌名を参照したと答えている。研究者の場合、同じ専門領域の著者に対する知識があるため、論文の標題だけでなく、著者名によっても論文の内容を推測することができ、それが判定の際にも行なわれたものと考えられる。

#### 3. 日本語文献と英語文献の判定の容易さ

どのグループにおいても、どちらともいえないとする者が大半を占めたが、どちらかという図書館員は日本語文献の方を判定しやすいとし、研究者では英語文献とする者が多い。これは、図書館員にとって、英語という外国語の判定が、多少とも負担になることを示すものといえる。

#### 4. 基礎と臨床の判定のむずかしさ

基礎的研究の色彩が強いと仮定して設定した主題A・Dは、臨床的なB・Dよりも判定がしにくいと、どのグループにおいても考えられているようである。その理由としては、探索主題の表示が不明確で、範囲がはっきりしていないとする者が多く、また医学生や研究者における回答では、関心の持ち方や専門の分野との関連において、判定のしやすさが決まると考えられている。

ついで、判定実験における平均所要時間は、第5表に

第5表 判定実験の平均所要時間

	ML	MS	MD
第1回	27.8分	15.7分	9.0分
第2回	28.9分	17.9分	10.9分

示す通りで、図書館員は研究者の3倍近い時間をかけたことになる。

#### E. 予備調査の補遺

予備調査では、実際に受けつけた探索質問を用いて調査したので、質問者自身の判定との相違を検討できる。そこで「低酸素の下における心機能および血行力学的変化について」という質問を出した研究者に対し、予備調査で用いた文献について同様な判定の調査を行なった。質問者は地方の病院に所属している者であったが、調査表を郵送して判定を依頼し、回答を得ることができた。

その結果、質問者は50文献中23文献を適合するものと判定した。個々の文献についての医学生と図書館員9名ずつの判定者数を比較すると、必ずしも判定者全員が一致して適合であるとした文献を、質問者も適合するものと考えていないことがわかった。また反対に、質問者が適合文献としたものでも、判定者の数が少ない例もある。

そこで、さらに質問者の判定を基準として、recall ratio と precision ratio を算出し、第6表を得た。ここでは、質問者の判定と一致した文献数の平均について、図書館員の方が医学生より多くなっており、recall ratio が高いが、50文献に対する適合文献数の平均も多いため、precision ratio は低くなっている。医学生を医学分野における専門家と解釈するには多少無理がある

第6表 予備調査における判定の一致と効率

判定者	適合文献数	一致数*	precision	recall
A	35	21	60.0	91.3
B	38	23	60.5	100.0
C	22	14	63.6	60.9
D	28	15	53.6	65.2
E	24	17	70.8	73.9
F	30	20	66.7	87.0
G	27	16	59.3	69.6
H	28	18	64.3	78.3
I	43	23	53.5	100.0
平均	30.6	18.6	61.4%	80.7%

判定者	適合文献数	一致数	precision	recall
J	31	18	58.1	78.3
K	38	21	55.3	91.3
L	31	19	61.3	82.6
M	44	22	50.0	95.7
N	31	19	61.3	82.6
O	49	23	46.9	100.0
P	35	19	54.3	82.6
Q	25	17	68.0	73.9
R	37	21	56.8	91.3
平均	35.7	19.9	58.0%	86.5%

\* 一致数とは、質問者が適合と判定した23文献に対する判定の一致を示す。

かもしれないが、この予備調査の条件のものでは、これまでの経験的事実とされるように、図書館員は recall ratio を重視し、有効な文献をおとす危険を避ける行動をとることが、この結果から裏付けられる。

#### F. 結論

これまでの分析結果をまとめると、分散分析による主題間の相違については、研究者D2グループにおけるC・D間に有意差がみとめられ、一方判定者グループ間の相違については、CとDにおいて有意差があった。そこで、さらにこれらの場合を検討してみると、研究者D2グループに関わる場合に限定されていることがわかる。D2グループは他の判定者グループと異なる判定パターンを持っていると推測されるが、その原因としては、前にもふれたように研究者グループとして、専門性のよさに欠ける面のあることが指摘できよう。

相関分析による結果からは、判定者グループの相互の間に、相関関係があり、相関係数から判断すると、図書館員よりも医学生の方が研究者との相関が高かった。

さらにアンケートの結果と考え合わせてみると、判定者にとって判定しやすい主題が、必ずしも高い判定の一致をみるとは限らないことがわかる。判定者の感覚としては、基礎的な主題より臨床的なものの方が概念としてもとらえやすく、判定しやすいと考えがちであるので、判定も一致度が高いと予想されたが、結果はその反対で、基礎的なものの方が高くなっている。

予備調査の補足として付け加えた質問者による判定との比較では、図書館員は recall ratio を高める努力をし、適合文献の検索ミスを防ごうとする傾向にあることが示されている。これに関連して、実験調査でも、医学

生の平均適合文献数は一番低いことが指摘される。予想では、研究者も図書館員より低いと考えられたが、B・Dにおいては、むしろ研究者グループが最も多くの適合文献をあげている。それゆえ、臨床的な主題はわかりやすく判定しやすいと考えられていても、主題の提示が広く漠然としている場合には、判定のばらつきも大きく、研究者は多くの文献を適合とみなす傾向にあるという結果になった。

ここでは、限られた2回の調査、4主題を検討しただけなので、医学分野における様々なタイプの探索質問に対して、一般化して述べることはできないが、基礎的な主題や範囲の限定された主題においては、判定者としての図書館員が、医学生や研究者とかなり近い判定パターンを示すと言ってよい。

## VI. 結 語

ここで行なった実験的調査は、マニュアルによる文献探索プロセスの中で、質問内容がうまくシステムに合致しない場合の問題点として、適合性判定のための照合レベルを考察し、大きく関係すると思われる判定者の主題知識の違いから、比較分析した試みであった。

文献探索プロセス全体を含めて、適合性の判定を考えた場合には、さらに質問者との相互作用が一番大きな要因であると考えられる。それは、経験的な観点からも、文献探索サービスを初めて行なう時、利用者との面談が情報入手の手段として有効にはたらないことから説明できる。つまり図書館員自身が、利用者に対処すべく技術的な二次資料についての知識と、探索質問の性格に対する識別力を有していなければ、面談による相互の理解と、より有効な検索とを期待することはできないのである。

この相互作用の問題は、明らかに重要であると言えるが、人間対人間の接触の上に成り立つものであるため、その効果についての計量的な把握は、方法論的に困難な点が多い。しかし事例研究などの積み重ねによって、その問題点をより明らかにすることは可能であろう。

ここでは、適合性判定の問題をシステム内に取り入れた形で質問との関連から考えて、利用者自身の判定を考慮に入れなかった。それは、前にも指摘したように利用者の判定がきわめて主観的なもので、価値判断を含んでいること、利用者自身の目的に合致するか、自分の読めるレベル、言語、スタイルかといった点で決まるものであることによる。

適合性の判定について、利用者（質問者）の側に立てば、当然これらのものを含めなければ意味がないが、適合性の概念そのものに混乱をきたすことも否めない。そこで、ここでは図書館員の側に立ち、文献探索サービスに提出される質問に対しての適合性を考え、客観的なレベルで、どの専門家によっても概念的に合致するものを適合するものと考えた。

これは、探索技術だけでなく、専門分野の関連知識を経験的に身につけた図書館員によって成り立つ文献探索サービスにおいて、未経験の探索者が、より早く均質なレベルまで到達するには、客観的な適合性の存在を考えて、それが第一の目標となるべきであるとする考え方に基づいている。そのレベルをどこに設定して、図書館員の教育やレベル・アップに結びつけるかは、今後の課題であるが、そのためには、これまで考察してきた限られた条件の判定パターンの実験を踏み台に、さらに多くの要因を含めた調査実験を行なう必要があろう。

機械検索は、詳細な索引化と論理式を用いて、迅速に文献を検索することができるが、ひとたびシステムにうまく合致しない探索質問が受け付けられると、システムとの間に判定者の介入を必要とするし、また大量情報の蓄積が可能であるといえども、限られたデータ・ベースをもとにしていることが指摘できる。その点、マニュアルにおける文献探索は、様々なタイプの二次資料を用いることによって、機械検索ではむずかしい質問に対し、きめ細かい検索を行なうことができる。今日、わが国においても、医学分野ではMEDLARSテープの機械検索サービスが実施されるようになったが、日本語文献に対する要求も多く、まだマニュアルの検索は重要な役割を担っていると言える。したがって、機械検索とともに、さらにより効率のよい結果をもたらすための努力として、検討を加える余地は多く残っており、今後ともこのような基礎研究をしてゆく必要があると思われる。

- 1) 森川夏子. "文献探索における概念の数の問題. 慶応義塾大学医学情報センターにおける文献調査サービスを中心として," *Library system*, vol. 10, July 1971, p. 65-94.
- 2) Rees, Alan M. and Saracevic, Tefko. "Conceptual analysis of questions in information retrieval systems," *Proceedings of the American Documentation Institute*, Part 2, 1963, p. 175.

- 3) King, Donald W. "Design and evaluation of information systems," *Annual review of information science and technology*, vol. 3, 1968, p. 83.
- 4) 松村多美子. "ランカスターのメドラーズ実績評価を通してみた情報システムと利用者の相互作用," *Library and information science*, no. 10, 1972, p. 171.
- 5) *Ibid.*, p. 171.
- 6) Tritschler, R. J. "Effective information-searching strategies without "perfect" indexing," *American documentation*, vol. 15, July 1964, p. 179-84.
- 7) Saracevic, Tefko. The concept of 'relevance' in information science: A historical review. <Saracevic, T. ed. *Introduction to information science*, London, Bower, 1970> p. 111-51.
- 8) 天野善雄. "文献探索における relevance の諸問題," *Library system*, vol. 6, Oct. 1967, p. 217-33.
- 9) Königová, Marie. "Mathematical and statistical methods of noise evaluation in a retrieval system," *Information storage and retrieval*, vol. 6, May 1971, p. 437-44.
- 10) Cooper, W. S. "A definition of relevance for information retrieval," *Information storage and retrieval*, vol. 7, June 1971, p. 19-37.
- 11) Foskett, D. J. "A note on the concept of 'relevance'," *Information storage and retrieval*, vol. 8, Aug. 1972, p. 77-8.
- 12) Wilson, Patrick. "Situational relevance," *Information storage and retrieval*, vol. 9, Aug. 1973, p. 457-71.
- 13) Lancaster, F. W. *Evaluation of the operating efficiency of MEDLARS*. Final Report. Bethesda, Md., National Library of Medicine, Jan. 1968. p. 20.
- 14) 長沢雅男. 情報検索入門. 東京, 森北出版, 1971. p. 182.
- 15) 天野, *op. cit.*, p. 219.

#### 参 考 文 献

1. Cuadra, C. A. and Katter, R. V. *Experimental studies of relevance judgments*. TM-3520/001, 002, 003/00. 3 vols. System Development Corporation, Santa Monica, Calif., 1967.
2. Rees, Alan M. *et al.* *A field experimental approach to the study of relevance assessments in relation to document searching*. 2 vols. Cleveland, Ohio. Center for Documentation and Communication Research, School of Library Science, Case Western Reserve University. Oct. 1967.