RECORD-KEEPING AND DATA PREPARATION PRACTICES TO FACILITATE RECORD LINKAGE

Martha Smith, Statistics Canada

Lack of adequate personal (or "entity") identifying information and appropriate documentation on what is contained in historical files can be major stumbling blocks in carrying out long-term follow-up studies. Over the past few years, considerable experience has been gained in the use of existing administrative (e.g., industrial employee, mortality, hospital, cancer, marriage, birth) survey and census data files for record linkage studies in Canada [1-3].

The purpose of this paper is to give some practical pointers for agencies and individuals involved in implementing future linkage projects, particularly those where large historical files are being used, and where no unique identity numbers are available. Specific examples will be given here which relate to occupational and environmental health studies, but many of the record linkage problems and their solutions apply also to other areas of statistical research.

Organizationally, the present paper is divided into six main sections. The first section gives the main results and conclusions. The second section outlines the kinds of data files required for occupational and environmental health studies. The third section describes the role that various broad categories of records can play in the linkage process. The fourth section gives examples of the practical problems in the preparation of existing files for linkage, along with the methods and some of the software developed to cope with these problems. The fifth section deals with the probabilistic matching technique and the art of designing an efficient linkage operation. The last section makes recommendations for future record keeping and data preparation practices to facilitate record linkage.

I. MAIN RESULTS AND CONCLUSIONS

A generalized record linkage system has been developed based on the concepts of probability and the use of 'weighted' record comparisons [4-7]. The probabilistic methods developed have several desirable features:

- records can be linked which lack unique numerical identity numbers;
- records are able to link despite discrepancies which may exist between identifying particulars;
- 'weights' can be assigned for agreement, disagreement, and partial agreement; and
- the technique discriminates between rare and common values of a given identifier.

On the basis of fairly extensive experience with computerized record linkage of a probabilistic kind, using the generalized iterative record linkage system (GIRLS), it seems unlikely that the technology and the software will be major limiting factors in the future. The major costs, which can limit the application of the approach, are often likely to be associated with the need to do data entry for additional identifiers in a standard fashion, if these have not already been captured in machine readable form. For historic data files, lack of appropriate documentation and standard data entry rules can cause problems. Some software has been developed to aid in the preprocessing of such files. It is therefore recommended that if the files are to be used for record linkage, sufficient identifying items be captured at the time of the initial data entry. Compromises whereby the amount of identifying information is restricted in order to reduce costs will be reflected in reduced accuracy of the linkages, and of the kinds of uses that can be made of the files.

Certain files may serve in the role of intermediate files that facilitate the linkage of other files.

Procedures to evaluate the quality of the linkage should be planned early. For example, it may be possible to incorporate known alive cases in a mortality search; to carry out independent manual follow-up on a sample of the file and compare with the computer results; or to carry out an alive follow-up to complement the death search.

Improvement of present data sources and the development of new sources would seem to be necessary if further demands for occupational and environmental health statistics are to be met. A checklist of data items to be collected has been described elsewhere [3-4].

Collaboration and co-operation among individuals and agencies are often required to complete studies. Suitable communication networks among investigators must be established, particularly if there is a long geographic distance between the interested groups.

II. KEY ELEMENTS IN A TYPICAL FOLLOW-UP STUDY THE KINDS OF DATA FILES REQUIRED

Certain general principles shape whatever epidemiological studies for long-term health effects are undertaken and influence the nature of the procedures for data gathering and analysis. The data gathering could include examining data systems already available which could facilitate the study. The requirements for identifying information are similar whether one is looking for changes to the exposed individual, or for inherited changes affecting the offspring from such individuals.

The key elements for data collection that should be included in any such study are described in [4]. A typical follow-up study often requires some knowledge of work histories, dose histories, health outcomes and the personal identification of the individuals involved. The software available must be capable of bringing all the various relevant files together at appropriate times.

The kinds of linkages involved may be a series of internal linkages to identify data pertaining to the same individual (e.g., to create individual work histories) as well as two-file linkages (e.g.

to match a work record against a death record). The matching techniques can use individual identification numbers (e.g., Social Insurance Numbers), probabilistic matching techniques, or a combination of the two.

III. THE FUNCTIONS OF BROAD CATEGORIES OF SOURCE RECORDS

The kinds of source records required for studies of delayed health effects may serve one or more of four possible functions in the follow-up process.

First, such records may identify an individual as belonging to an "at risk" population (or to a "control" population with which the other is to be compared). In this case they are referred to as "starting point" records which initiate the follow-up process.

Alternatively they may identify an end effect, such as cancer or death in an individual who is a member of a study population, in which case they are referred to as "endpoint" records. One example of an endpoint file is the Canadian Mortality Data Base consisting of the records of all deaths in that country dating back to 1950. Follow-up thus will consist of using a file of starting-point records to search a file for potential end-point records, and of linking those records from the two files which relate to the same individuals.

The third possible function of a record file is that of an intermediate file which facilitates the searching and the linkage process. For example, where a starting-point record carries the maiden surname of a woman who later married, and the endpoint record contains her married surname, the search of the endpoint file may be more productive and accurate where reference can be made to another file, such as a marriage file or the Social Insurance Number Index which contains both of these names.

The fourth function of record files is as a source of the detailed statistical variables required for the analysis. For example, linkage may be required to bring together individual work histories, dose histories and smoking histories.

In considering the possible uses of various available files, all four functions must be kept in mind.

IV. PREPARING THE INPUT FILES

Prior to linkage of any kind, the records being used need to be brought into the formats that are required for making the necessary comparisons, and into the sequences that are appropriate for the linkages. The quality of the identifiers needed for linkage may also be tested by looking for blank fields and for values of the identifiers that are not permitted (such as day of birth = 32). If data collection and data entry have not been done with record linkage initially in mind, this phase can be quite time consuming and costly.

We have found the Statistical Analysis System (SAS) very helpful at this stage, and as a routine we systematically scrutinize the values of fields in files to be used in linkage. These are compared with any available documentation regarding coded values and their meanings. One can check how many fields have non-missing values, valid values, ranges, codesets, or invalid characters or values. Whereas blank fields can only be filled from other sources, fields which have unacceptable values may sometimes be corrected.

One may wish to create a new field for each record to indicate the "availability" and validity of fields on the same record. For example the value "120112001" could indicate "present and with the valid code range" (1), "present, but with an invalid code" (2), or "absent" (0). A SAS distribution of this word facilitates one's assessment of the likelihood that one will be able to link the files.

It is necessary to obtain copies of the forms of the original source documents, the record layouts and any file documentation, along with detailed information regarding how the administrative system works.

Some problems one may expect to encounter have to do with the quality of the records, and some methods which have been used to deal with the problems are as follows:

(1) Lack of a standard format - particularly for the name and address fields

If name fields have been entered in string format and if a variety of delimiters have been used to separate surnames from forenames, it may be necessary to put the values of the fields into a standard fixed format. It is particularly difficult to separate the components in a name field if blanks have been used as the delimiter. A simple NAMESCAN routine has been developed, which changes all alphabetic characters to "A" and leaves all other characters intact. A SAS distribution can then be made to look at the various patterns on the file.

When standardizing name fields, titles should be put in a separate field e.g., Mrs, Jr, Sr. Two-part surnames can be concatenated (SMITH-JONES to SMITHJONES) and retained along with alternate entries for SMITH and for JONES, special characters may be eliminated (O'CONNOR to OCONNOR) and prefixes concatenated (VAN DYK to VANDYK). A prefix list is shown in Table 1. Geographic and disease codes will usually have changed over time. It may be necessary to recode fields so that all records share a common system of codes, or to use ranges of codes that are comparable.

Table 1. --List of Surname Prefixes

BON	DI	LE	O
D	DO	LES	ST
DA	DU	LI	STE
DE	EL	LU	VAN
DEL	FITZ	LOS	VANDEN
DEN	L	M	VANDER
DER	LA	MAC	VON
DES	LAS	MC	VONDER

(2) Spelling errors

To get around spelling errors in surnames, a phonetic encoding scheme can be used. We currently use the modified New York State Identification and Intelligence System (NYSIIS) surname code [8]. In the 1950-79 Mortality Data Base file, there were about 200,000 unique surnames which mapped into about 40,000 NYSIIS codes. Based on evaluation studies of earlier linkage projects, we are currently considering making modifications to this coding scheme based on some of the phonetics involved with Canadian names (particularly French names).

(3) Incomplete files

Due to the rules regarding cutoff dates for preparation of statistics from certain files, one may find that records are missing due to late registrations. If the files are assigned numbers in an orderly fashion, a sequence and continuity check of the numerically sorted file can be carried out, missing gaps listed, as well as the first and last record numbers of the files. We have done this for the Mortality Data Base file. Where exposure data files have been maintained separately from the Master Identification file, some utilites can be used to match files for "orphan" records i.e. an exposure record with no corresponding record on the master identification file or vice versa.

(4) Missing identifiers

These can be assessed from SAS output of individual fields, as well as using the availability word for a number of variables. It is advisable to split a field into its component parts - for example, for birth date use year, month and day. Sometimes sex code has been found missing from files.A list of all forenames appearing on the Mortality Data Base has been created. This has been used to impute a sex code e.g., 1=male only, 2=female only, 3=either male or female forename. Sex code is required so that appropriate weights can be assigned for forenames in the frequency weighting.

(5) Lack of documentation of old historical files

Here we have found SAS output very helpful, and created documentation regarding the contents of each field.

(6) Possible correlated data items

Certain data fields may be correlated, therefore caution has to be taken when assigning weights to these items e.g., birth place of father, mother, and a child. In certain instances the information relates to identical items (e.g., an address and postal code); in other cases it may reflect multiple wrong guesses (e.g., a birthdate being incorrectly reported).

(7) Duplicate records not properly identified

It is important that for a two-file linkage, all records that are known in advance to relate to the same individual be properly identified. This is to ensure that any groups to which either record of such a pair may belong can be combined by the linkage system. Typical examples are records relating to women who have both a maiden name and a married surname. One is unlikely to want to discard one record and keep the other, because there may be records on the other file that relate to either surname. A field can be added to the record to contain a value of 1,2,3 etc. to indicate whether this is the first, second or third "duplicate" entry for this record. If no duplicate exists, the value of the field can be set to zero. Such duplicate records must all be assigned the same unique number (in the GIRLS system this is referred to as the SEQUENCE number).

If an intermediary file is used, alternative entries can be put in with different versions of the identifying information. These may be either entries from both files separately or in hybrid form (i.e. certain items from one file and other items from the other file).

(8) An internal linkage should have been done first

Any file that is going to be used for a two-file linkage, should first be examined to determine whether an internal linkage is required to bring together all records which refer to the same individual (or entity). If one is uncertain about whether there are duplicates, sometimes a fairly inexpensive first check may be to sort the file by surname, first forename, and birth date and to create a microfiche copy of the file for visual examination. A great deal of work in a two-file linkage can be saved by first unduplicating in this fashion the two files that are to be linked.

(9) Length of data fields

If two fields are to be compared, the lengths of the data fields need to be compatible. For example, as a standard, we encode ten letters of the surname into the NYSIIS code. If the number of letters in one file is less than ten characters, problems can arise when the codes are compared. It is therefore advisable to use a surname field that is ten characters or greater. If special characters were originally used, the data entry of the field should be large enough to allow for the elimination of these special characters in the preprocessing.

(10) Separating out values where the same field was used for more than one purpose

As an example, the same field on some files may be used for maiden as for alias surname. One may wish to try to separate out the two types of surnames that have been entered, so that during the linkage step appropriate rules can be used.

(11) Several unique numbering systems used over time

In certain files, several numbers may have been used over time to refer to the same individual. In administrative sytems, there may be a rather different problem; one often needs to clarify whether such numbers have ever been reassigned to other individuals.

In certain cases, one may wish to chain all the various numbers that were used by the same person over time and use this as a pocket identifier within which a probabilistic match could be made.

V. PROBABILISTIC RECORD LINKAGE TECHNIQUES

The Basic Principle

There are three major difficulties to be overcome in order to achieve efficient record linkage. The personal identifying items are often inadequate to discriminate between the person to whom a record truly refers, and other persons in the population who have similar names. A second difficulty arises because when people report personal identifiers they frequently make mistakes. The third difficulty arises because of the large volume of records involved in record linkage. Some related difficulties include the setting of appropriate threshold values for acceptance and rejection of linkages, deciding how most efficiently to carry out a multi-step operation, deciding on the number of partial agreements to use and the selection of pocket identifiers.

The objective of the Generalized Iterative Record Linkage System was to make it possible for computer procedures to efficiently carry out the data processing involved in the probabilistic matching of data files, and to do so easily for a wide variety of diverse data requests. The GIRLS system has involved optimizing four major tasks: (1) the search operation, (2) the decision-making step, (3) the grouping of records, and (4) the retrieval of information.

In the searching step, the sequencing information is used as a means of avoiding the many unprofitable pairings that would have to be examined if every record initiating a search were compared with every other record in the file being searched. Generally for searches of the Mortality Data Base, comparison pairs are created only where both the sex and the phonetically coded form of the surname agree.

For other applications, the sequencing may be by one of several systems of numerical identifier or by phonetically coded surname. Regardless of the means by which the record pairs are brought together, the next step will be a detailed comparison of the remaining identifiers. This is necessary even where the numeric identifiers agree, because such identifiers are occasionally used improperly by persons to whom they do not belong, and sometimes even by a relative of the rightful owner who has the same surname.

At the present time, a test is being made to provide a measure of the usefulness of employing personal identifiers from the Social Insurance Number (SIN) index file to supplement those from the work records, for the purposes of carrying out automated death searches. Not only are the names, birth dates and such more likely to be recorded on the SIN record, they are also more likely to be complete, and as well they will frequently include the mother's maiden surname, which carries considerable discriminating power and is quite unlikely to be available from any work record.

In the decision-making step, each of the remaining identifiers is compared in turn, wherever it is represented on both members of the comparison pair of records.

The odds associated with any specified outcome from the comparison of any identifier are:

$$\text{odds} = \frac{\text{freq of specific outcome in linked pairs}}{\text{freq of specific outcome in unlinked pairs}}$$

This applies equally to agreements, disagreements and to any degree of similarity or dissimilarity no matter how it is defined (as long as both definitions are identical above and below the line).

When pairs are sorted in descending order of total weight, a point is reached at which the record pairs should be judged unlinkable or borderline. To calculate where this threshold should be, two further values are required to be weighted for a two-file linkage. These are:

(1) the likelihood that the individual is represented in the file being searched, so that there is a potential for linkage, and (2) the size of the file being searched, since the opportunity for fortuitous agreement increases in proportion to the file size.

The logarithms of both of these values will be negative. When added in with the weight from the identifier comparisons, the resultant sum is known as the "absolute total weight".

$$W^* = W + \log_2 \frac{Na(L)}{Na} + \log_2 \frac{1}{Nb}$$

where,

W^* = $\log_2$ of the absolute odds in favour of a correct linkage;

W = $\log_2$ of the relative odds in favour of a correct linkage = w1 + w2 + w3 .... where these are each logs to the base 2 of the odds ratios for the successive identifier comparison outcomes;

Na and Na(L) are respectively, the total number of records in the file initiating the searches and the number out of these that will be linked with matching records in the file being searched (or a reasonably close estimate of Na(L)/Na may be used initially); and

Nb = the size of the file being searched.

To calculate w1, w2...., for reasons of convenience it is desirable to treat separately the data derived from linked pairs and that which applies to unlinked pairs. If w is the net weight for the particular identifier comparison outcome, $\log_2$ (frequency in linked pairs) is the negative component of this net weight, and $\log_2$ (1/ frequency in unlinked pairs) is the positive component of the net weight.

Because the negative components of weight vary with the quality of the file initiating the searches i.e. with the reliability of the identifiers as recorded on that file, these negative components need to be recalculated for each new linkage before the final weighting is done. The data may be obtained initially from preliminary machine linkage, numerical linkages where available, or from manual linkages. Examples of how the weights are obtained are discussed in reference [9].

The positive components tend to be stable where the files being searched are the same on successive occasions (e.g., the death file) and can usually be calculated from the frequencies of the identifier values in that file.

The Art of Record Linkage

The art of designing an efficient computerized linkage operation depends less upon theory than an intuitive perception of how best to carry out the comparisons and what outcomes from these are most likely to be revealing, so that they ought to be recognized by the computer.

Some of the intuitively obvious refinements that have actually been put to use in Statistics Canada's death searches have to do with:

(1) Recognition of partial agreement outcomes, e.g., of
- surnames (three levels of agreement/disagreement);
- given names (eight levels of agreement/disagreement, including agreement truncation where the initials agree);
- birth year (up to 6 levels of agreement/disagreement);
- birth month (3 levels);
- birth day (4 levels).

(2) Recognition of cross-agreement, e.g., of
- initials (where there is no straight agreement);
- month and day of birth - as for initials.

(3) Recognition of degrees of compatibility/incompatibility e.g., in
- last known alive year versus death date (up to 4 levels);
- marital status (up to 4 levels for each status on a search record).

(4) Comparison of place of work versus place of death.

(5) Calculation of age at the time of the matching death to determine the likelihood of death in a particular year using life-table data.
(6) Use of death file size for that same year as influencing the odds for a fortuitous similarity of the identifying particulars.

A potential refinement may be judged worth retaining as a part of the linkage procedure where it is used often enough in doubtful matches, and makes a large enough difference in the final decision to link or not to link, to justify the possible added complexity in the programming. The GIRLS system makes it possible to gather such data after a preliminary linkage and again after a final production run.

The best tactic when designing a linkage procedure for a specific operation is to gather such empirical data after a preliminary linkage so that the procedure can be revised before the final weighting. The information needed earliest has to do with the frequencies among linked pairs of the different comparison outcomes recognized by the preliminary linkage procedures. The tabulations ("info outcomes") should recognize all the comparison outcomes likely to be useful in the decision process.

We often find that what one learns by looking at some manual linkages first can be very helpful in planning a study. This aids in working out the appropriate methods to use and in preparing cost estimates. One may have to decide whether there is enough identifying information available to do the linkage. To get an overall estimate of this, one can first imagine how strongly unfavourable the odds would be if one did not know whether any of the items agreed or disagreed, and were linking to a file of a given size. Then, as one compares each item, in turn, and assumes they agree, this will demonstrate the possible extent of the increase in likelihood favouring correct linkage. One can use a global overall weight for the items employed in this exercise, and hence get a ballpark impression of whether or not there are enough items available to make it work (see Tables 2 and 3 for an example).

After the linkage status decision has been made, the system can identify groups of records which potentially refer to the same entity and it can indicate where conflicts exist. A conflict exists where groups do not fit your requirement e.g., one record relating to more than one death record. In the GIRLS system there are two ways of resolving these conflicts - automatic resolution by the system based on the 'best' linkage, or by manual resolution. A combination of the two often works best.

The retrieval of information operation of the system is designed to quickly and concisely aid the user in making decisions regarding the future direction of the linkage process. The GIRLS system can produce reports at the detailed level on weight sets, linked pairs, group reports, information about the linked pairs, and it can also produce estimates for updating the weights. One may wish to produce reports based only on links for which a given condition is true (e.g., all links above a given weight) or for which a condition using variables on the source records may be true (e.g., all known dead cases as known earlier on the worker's nominal roll file).

VI. FUTURE DIRECTIONS

There are three main directions for our future endeavours:
(1) The improvement and expansion of existing search and linkage facilities which could include further development and enrichment of our current files (e.g., addition of occupation and industry on the death file). The NYSIIS code routine needs to be evaluated more fully taking into account the kinds of names found in Canada. A dictionary of accredited comparison procedures needs to be developed from past linkage studies that could serve as a guide for future studies. Results from earlier studies need to be carefully evaluated,

Table 2. —Example of a Possible Census-to-Death Linkage -- Likelihood of Fortuitously Selecting the Correct Death Record, Using no Identifiers Other than Sex (Assumes enumeration in 1971 at age 42, death in 1979 at age 50, and male sex)

COMPARISON ITEMS	ODDS	CUMULATIVE ODDS	WEIGHT (10 X $\log_2$)	CUMULATIVE WEIGHT (10 X $\log_2$)	NOTES
Random chance of finding death in 1979 male death file, assuming it is there	1/96,532	1/96,532	-166	-166	1
Likelihood of dying in that year, if alive at the beginning of it	1/131	1/12,645,692	- 70	-236	2
Likelihood of being alive at the beginning of 1979 if enumerated in 1971	1/1.04	1/13,151,520	- 1	-237	3

Note: (1) From death file size, for males dying in 1979.
(2) From life tables for likelihood of death in a 12 month period, for a male of age 50.
(3) From life tables, for the likelihood of survival to age 50 among a cohort of males still alive at age 42.

Table 3. --Example of a Possible Census-to-Death Linkage -- Cumulative Effect of Successive Agreements on the Odds in Favour of a Correct Match, when all Identifiers are Present and Agree

IDENTIFIER AGREEING	ODDS	CUMULATIVE ODDS	WEIGHT	CUMULATIVE WEIGHT
			(10 x $\log_2$)	
(Random chance)	-	1/13,151,520	-237	-237
Surname	2,287/1	1/5,745	+112	-125
First initial	14/1	1/410	+ 38	- 87
Second initial	14/1	1/29	+ 38	- 49
Rest of first name	87/1	3/1	+ 64	+ 15
Marital status	26/1	8/1	+ 14	+ 29
Year of birth	56/1	437/1	+ 58	+ 87
Month of birth	12/1	5,242/1	+ 36	+123
Birth prov/country	8.6/1	45.078/1	+ 31	+154
Ethnicity	3.5/1	157,773/1	+ 18	+172
Parental birthplaces	1.2/1	189,328/1	+ 2	+174
Industry, major	6/1	1,135,968/1	+ 41	+215
Occupation, major	11/1	12,495,648/1	+ 31	+246
Residence province	4.4/1	54,980,851/1	+ 21	+267
Residence city	72/1	3,958,621,272/1	+ 62	+329

particularly with respect to the use of intermediate files and the use of alive follow-up procedures as were used in the Ontario miners study [10]. Further refinements are needed in developing a file of non-links to get weight estimates, particularly where the comparisons are fairly complex (e.g., weighting of forenames).

(2) **The development of new and much needed data bases** which would identify, in a more systematic fashion than heretofore, the occupational and environmental circumstances of people, and which could be used as starting point files, to initiate the searches for subsequent health histories. Here data collection rules and forms need to be more clearly developed which could be used by industry. Use of new files such as census of agriculture, farm registers, and census of population files can be exploited. The use of existing files for alive and morbidity follow-up need to be explored.

(3) **The exploration with other agencies of any collaborations** that would be productive for generation of the required statistics, and for setting up the necessary communication network and financial support to implement such recommendations.

ACKNOWLEDGMENTS

The author would like to thank Dr. Newcombe for his contribution to this paper. The opinions expressed in this paper are those of the author and do not necessarily represent the views of Statistics Canada.

REFERENCES

[1] Smith, M.E. and Newcombe, H.B., "Automated Follow-up Facilities in Canada." AJPH Vol. 70, No. 12, pp. 1261-1268, 1980.

[2] Smith, M.E., "Long-term Medical Follow-up in Canada." In: Peto, R., Schneidermen, M. eds. Quantification of Occupational Cancer. Banbury Report 9, Cold Spring Harbor Laboratory, pp. 675-688, 1981.

[3] Smith, M.E., "Development of a National Record Linkage Program in Canada." American Statistical Association - 1982 Proceedings of the Section on Survey Research Methods, pp. 303-308, 1982.

[4] World Health Organization, "Guidelines for the Study of Genetic Effects in Human Populations." Environmental Health Criteria 46 (in press).

[5] Howe, G.R. and Lindsay, J., "A Generalized Iterative Record Linkage Computer System for Use in Medical Follow-up Studies." Computers and Biomedical Research, Vol. 14, pp. 327-340, 1981.

[6] Smith, M.E. and Silins, J., "Generalized Iterative Record Linkage System." American Statistical Association - 1981 Proceedings of the Social Statistics Section, pp. 128-137, 1981.

[7] Hill, T., "Generalized Iterative Record Linkage System: GIRLS." Research and General Systems, Informatics Services and Development Division, Statistics Canada, Ottawa, 1985.

[8] Lynch, B.T. and Arends, W.L., "Selection of a Surname Coding Scheme for the SRS Record Linkage System." Statistical Reporting Services, U.S. Department of Agriculture, Washington, D.C., 1977.

[9] Newcombe, H.B. and Abbatt, J.D., "Probabilistic Record Linkage in Epidemiology." Red Book Series Report No. 5. Eldorado Resources Limited, Suite 400, 255 Albert Street, Ottawa, Ontario., K1P 6A9, November 1983.

[10] Muller, J., Wheeler, W.C., Gentleman, J.F., Suranyi, G., Kusiak, R., "Study of Mortality of Ontario Miners." International Conference on Occupational Radiation Safety in Mining, Vol. I, pp. 335-343, 1984.