Ildikó ZSIGÓ - János MARTON

Computerised Identification of Disciplines and Core Journals in the Medical Sciences

Scientific journals are the most important sources of scientific information. Their quantity exceeds ten thousands and their prices are measured in millions of dollars. Well, it is absolutely not indifferent what is the effectiveness of subscriptions and it is very important to know whether the most significantjournals are available in a library or not.

Our aim was the computerized thematical classifying of medical science, and automatic identification of the core journals of its disciplines.

In our investigation we used 306 journals categorized as important ones in the journal list of Excerpta Medica from every field of life sciences together with the 25 most frequently citing journals of each. The computer program we had developed thematically grouped these journals into 56 life science disciplines with the help of citation connections. The citing journals of the members of the thematical groups formed the core journals. The core journals were ranked firstly according to the number of their citations to the members of the disciplines and secondly according to their impact factor.

Now we would like to show the 12 steps of our computerized procedure.

In the first step we fed the 306 base journals and their 25 most frequently citing journals into the computer database from the 1995 Journal Citation Reports volume of Science Citation Index.

During the second step we introduced the concept of the FELLOW JOURNALS in order to establish connection between the base journals. Two base journals are fellow journals if at least 5 of their citing journals are identical.
Figure 1.

In the third step in order to express connections only with the help of base journals, we introduced the term of PARTNER JOURNALS. Thus, two journals are called partner journals if at least 80 % identity exists between their two sets of fellow journals. So we reached that the citation connections between the base journals were determined only by the base journals.
Figure 2.

In the course of the fourth step, we could form the primary disciplinary clusters. First we ranked the journals according to the number of their partners. Starting with the journal having the least partners, the first primary discipline cluster were created from this journal and its partners. Taking the journals one by one, all those journals became a disciplinary cluster leader which have not yet been participated in a previously generated disciplinary cluster as a partner. One journal could become the member in several disciplinary clusters, depending on how many partners of the journal became disciplinary cluster leader.
Figure 3.

In the fifth step, we wanted to narrow the big similarity among some of the primary disciplinary clusters, so we selected the two most characteristic journals of each disciplinary cluster, which had the most common citing journals among each other. In the following, let's consider these pairs ofjournals as determining pairs of the disciplines.
Figure 4.

In the sixth step the common citing journals of the determining pairs were selected as members of the disciplinary clusters.
Figure 5.

In the seventh step we omitted from the clusters those journals which were not among the base journals.

To increase the differentiation among the disciplinary clusters, during the eighths step journals being determining pair members were omitted from the clusters where they were not in the determining pair.

In the ninth steps we picked out those citing journals of the disciplinary cluster members, which cited more than half of the member journals.

Next, in the tenth step, we repeated the seventh and eighth steps. Journals which were not among the base journals were omitted, and we left the earlier selected determining pairs of journals only in their own disciplinary cluster. As a consequence of this procedure more differentiated disciplinary clusters were created.

In the eleventh step, we eliminated journals included in more than one disciplinary cluster. They remained in the cluster which attracted them strongest.

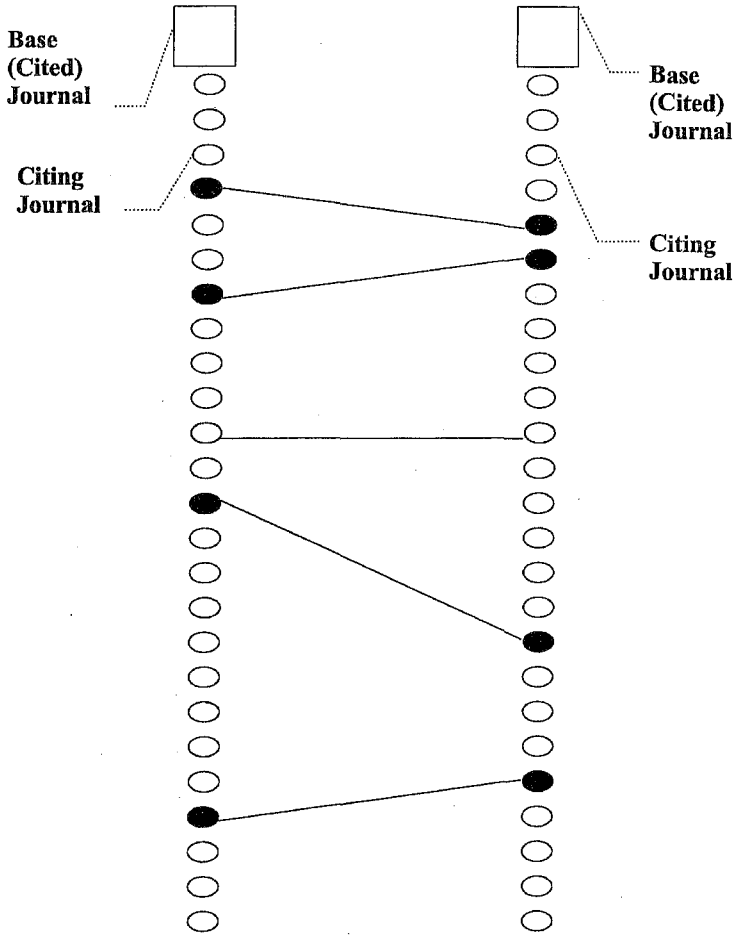
In the twelfth step, with the help of a special algorithm we eliminated the citing journals with the less occurrence number. Following this, we united clusters having 80% identity in their core journal lists. We repeated the previous two steps while there were no 80% identity among the core journal lists.

Now we could create the ranked lists of disciplines from the citing journals of the disciplinary clusters. Ranking principle was the number of occurrence as a citing journal.
Figure 6.

The method is proper to evaluate any journal collection, for example a library subscription list. If we use the members of the collection as the starting base journals of the clustering procedure, the theme groups of the library will be identified automatically and the presence and absence of the most important, highest ranked core journals helps in carrying out the strengthening of the completeness of the best equipped fields and the liquidation of the poorly represented ones.

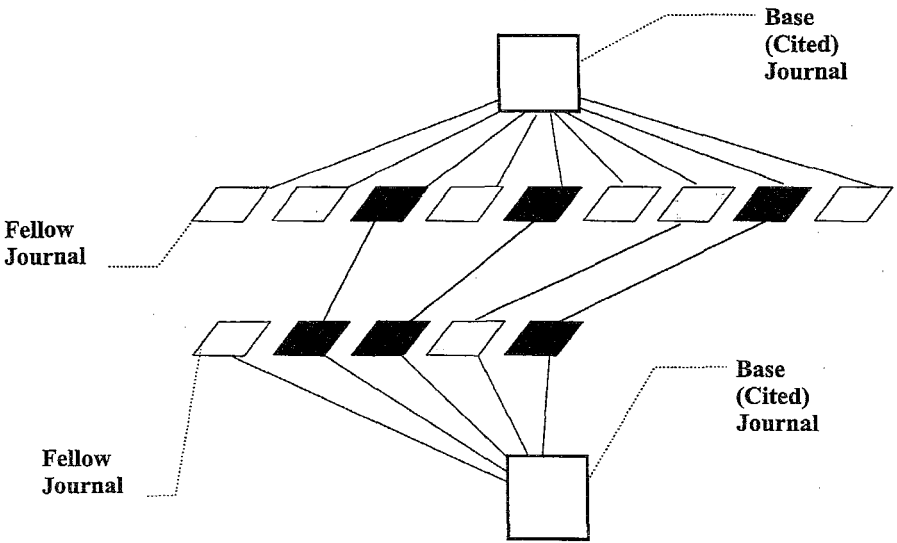
Above the evaluation of library stocks we can create a science map with the help of the citation connections of disciplines. We selected the two most similar clusters of each cluster and by drawing these connections a net of disciplines was created.
Figure 7.

Figure 1.



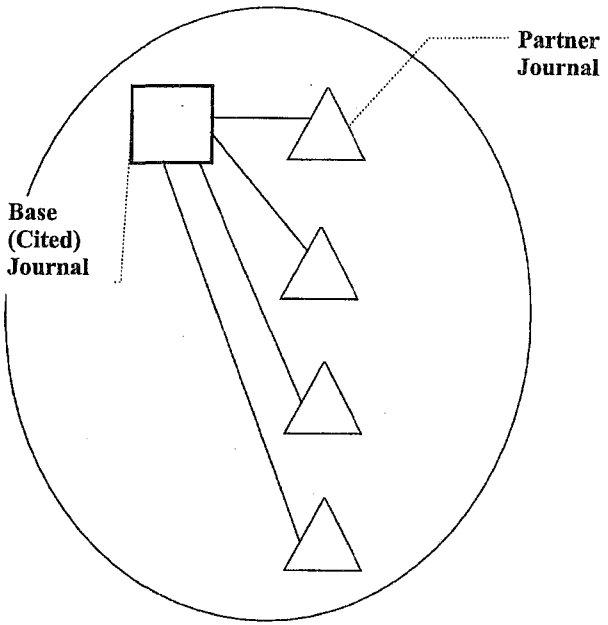
Fellow Journals => at least 5 common citing journals

Figure 2.



Partner Journals => at least 80 % common fellow journals

Figure 3.



Primary disciplinary cluster => base (cited) journal + its partners

Figure 4.

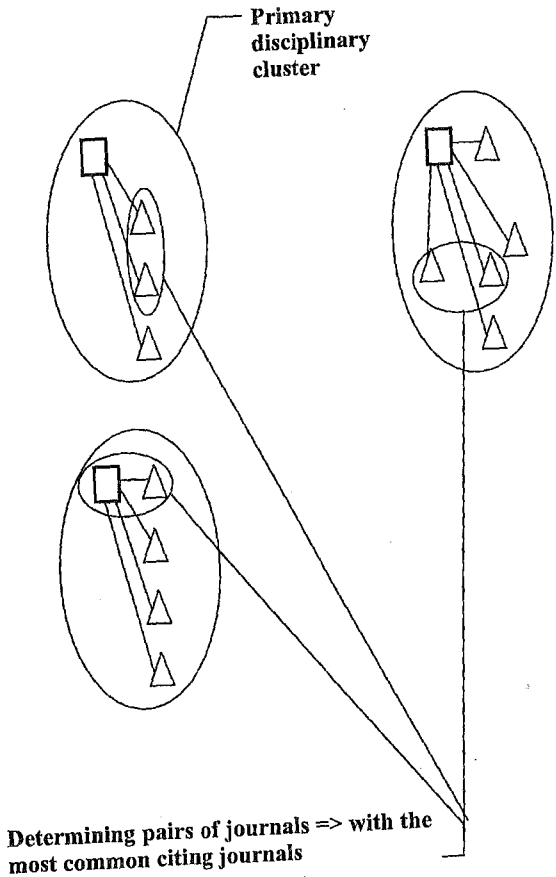

Primary
disciplinary
cluster
Determining pairs of journals => with the
most common citing journals

Figure 5.

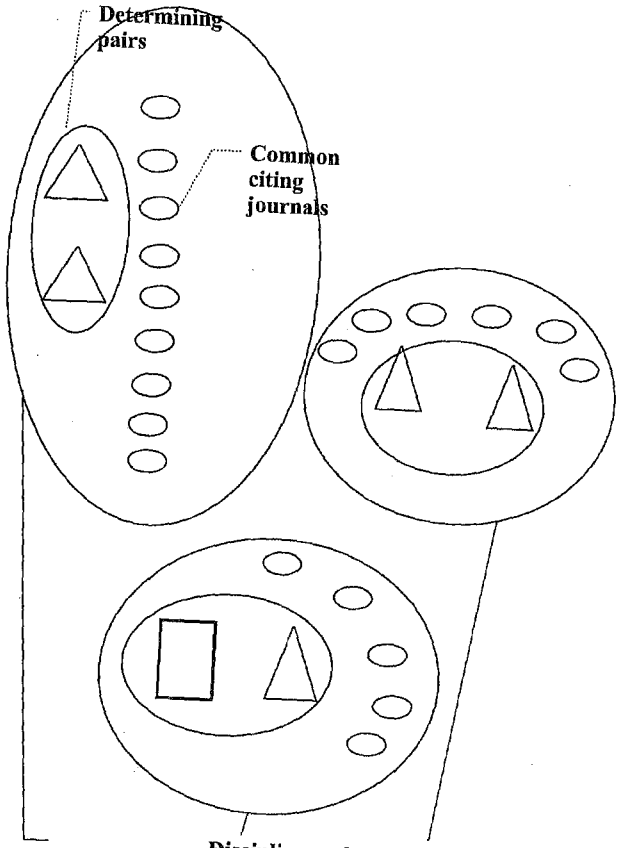

Determining pairs
Common citing journals
Disciplinary cluster =>
determining pairs + common citing journals

Figure 6.

Example for the ranked core journal list of a discipline

GYNECOLOGY

Cluster members

AM J OBSTET GYNECOL
BRIT J OBSTET GYNAECOL
EUR J OBSTET GYN R B
FERTIL STERIL
GEBURTSTH FREAUENHEILK
OBSTET GYNEACOL

Ranked core journal list	*Occurence No.*
AM J OBSTET GYNECOL	6
FERTIL STERIL	6
OBSTET GYNEACOL	6
ACTA OBSTET GYN SCAN	6
J REPROD MED	6
EUR J OBSTET GYN R B	6
BRIT J OBSTET GYNAEC	5
GYNECOL ONCOL	5
HUM REPROD	5
GEBURTSTH FREAUENHEILK	5
INT J GYNECOL OBSTET	5
LANCET	4
J CLIN ENDOCR METAB	4
PRENATAL DIAG	4
CLIN OBSTET GYNECOL	4
J PERINAT MED	4
GYNECOL OBSTET INVES	4
BAILIERE CLIN OB GY	4
Z GEBURTSH PERINATOL	4
J ULTRAS MED	3
SEMIN PERINATOL	3
HYPERTENS PREGNANCY	2

MAP OF MEDICAL SCIENCES

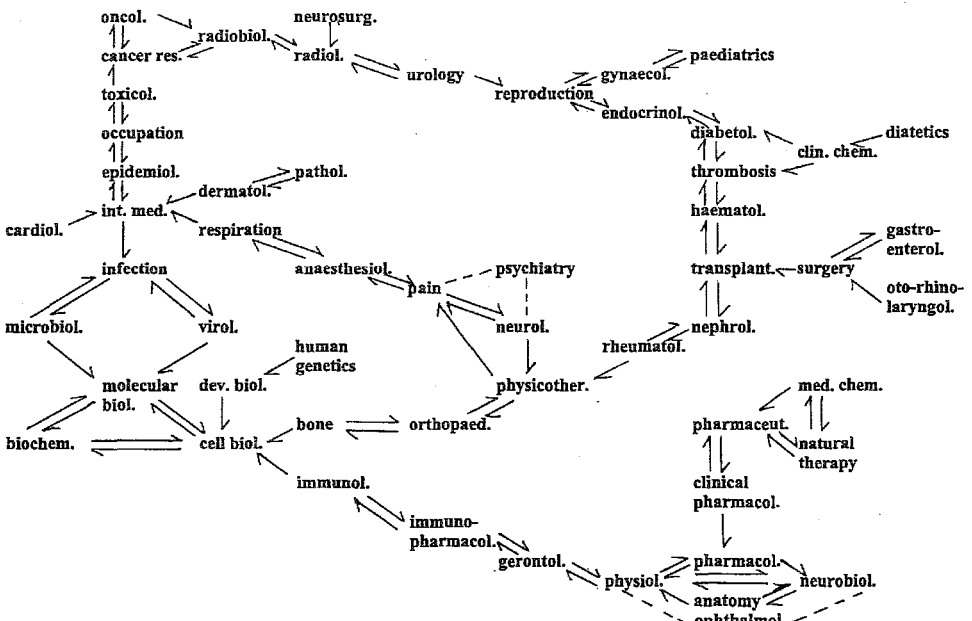


Figure 7.