

# A System of Codification of Medical Diagnoses for Application to Punch Cards, With a Plan of Operation\*

JOSEPH BERKSON, M.D.

*Division of Biometry and Medical Statistics, The Mayo Clinic, Rochester, Minn.*

THE system of coding, which is described here, is an attempt to combine in a single scheme provision for two functions, which are related but not identical: first, the cross-indexing of medical conditions, so that medical histories may be obtained from the storage files, in specified groups, for research purposes; second, the issuing of periodic statistical summaries of medical conditions seen during a definite period of time (yearly reports). While, broadly, these two functions can be thought of as parts of a single comprehensive system of keeping account of medical records, practical considerations make it desirable to plan separately for the needs of the two uses. A few examples will serve to clarify the point.

So far as providing that the record of a given case will be available when it is wanted for study of particular subjects, all that need be done is to enter this record under all the relevant categorical headings representing these subjects. If a given set of diagnostic terms is attached to a history, so far as cross-indexing is concerned, one need not attend to the relative importance or certainty of the various diagnoses,

but only to the complete entering of the record under all relevant diagnostic phrases. Thus, to give a simple example, if the diagnostic report in a case reads "carcinoma of bronchus?—or possibly thymic tumor" the case should be cross-indexed under both Carcinoma, bronchus and Tumor, thymus. Anybody who later studies either of these conditions should see this history. On the other hand, for the purpose of statistical enumeration, this would be the worst conceivable disposition, for it would lead to a count of two diagnoses, one, carcinoma of the bronchus, and the other, thymic tumor; whereas, the diagnostician did not mean that the individual was suffering from both conditions, but only that he was uncertain as to which one. For statistical purposes one would have to decide to count the case as one of carcinoma of bronchus, or one of thymic tumor, or one of undefined diagnosis. Another point is that for purposes of drawing individual groups of histories, it is desirable to keep account of certain details, which if included in a general statistical report on all cases, will only tend to clutter it up with small groups and cause confusion. For instance, for purposes of cross-index reference, one may wish to keep distinct, cases of "direct inguinal hernia" from cases of "indirect inguinal hernia," and "recurrent hernia" from "nonrecurrent

\* Read before the Vital Statistics Section of the American Public Health Association at the Sixty-fourth Annual Meeting in Milwaukee, Wis., October 9, 1935. The scheme here described is modified for more general application from one in use at The Mayo Clinic.

hernia," but these details may be undesirable for a general report. At the same time, provision must be made that a case of recurrent, direct, hernia shall not be indexed in such a way that it is not distinguished from two cases, one with a diagnosis of direct, and another with a diagnosis of recurrent inguinal hernia and so enumerated as two cases. This point assumes the greatest practical importance if diagnoses have been subdivided into refined subvarieties in the scheme of diagnoses used. For instance, if rheumatic heart disease is subdivided into rheumatic pericarditis, rheumatic mediastinopericarditis, rheumatic myocarditis, rheumatic endocarditis, and so forth, then each of these must be given an independent numerical designation.

If, now, a patient presents a combination of these rheumatic cardiac conditions which in general will be the case, there will be several numbers to assign, and the patient will be counted several times instead of once for what is actually a single condition. Finally, and perhaps most importantly, for purposes of drawing out defined groups of histories to be studied, there is no great significance to the particular order in which the records are kept. It makes no difference whether the records of cases of osteomyelitis are obtained from an alphabetical list under the letter O, or from an anatomic ordering under "bone," since all one wishes, is to get the records of cases of osteomyelitis. However, if all the diagnoses are to be enumerated in a consecutive listing, the ordering and groupings of the list have to be such as to make it intelligible. Furthermore, if the list is to be used for comparative studies, the separate rubrics contained in it should be defined on some standard basis. If one is to prepare for the report of all cases, therefore, one has to deal with a fundamental consideration of the designation of defined rubrics, and in what order

they should be arranged, that does not even arise in preparing for cross-indexing. These examples do not by any means exhaust the problems involved in adopting a code to both cross-index and summary enumeration purposes, but they will serve to focus the reasons for some of the particular details included in the coding scheme which follows.

#### PRINCIPLE OF SCHEME

The principle on which the scheme is built is as follows: The function of summary statistical enumeration (yearly reports) is relegated to a numbering system which occupies 3 columns for what I shall call the "main number." In combination with this, I use a numbering system which occupies 2 additional columns for what I shall call the "subnumber," which performs the rest of the functions of the system. The subnumber is coded in a unique way, for certain definite reasons which will be explained later. Each diagnostic condition therefore is represented by a number combination which occupies 5 columns.

*The Main Number*—Since the main numbers will determine the rubrics and order of the general statistical report, the disposition of these numbers will depend on what order is decided on. The form of such a listing is necessarily always arbitrary. I have used the groupings and order of the *International List of Causes of Death*, not because that list recommends itself scientifically, but because it is the one which has, at present, the widest statistical application. However, it has been modified and amplified to accommodate it to hospital morbidity conditions.\*

---

\* It would be highly desirable to establish a standard form for reporting hospital morbidity similar to the *International List of Causes of Death*. The list used by The Mayo Clinic can be obtained upon request.

*The Subnumber*—The field for this occupies 2 columns. The coding applied to this subnumber does not follow the usual consecutive numbering that would permit 100 consecutive items in the 2 columns used. Instead, a number is assigned to each of the individual 12 punch positions in the 2 columns, giving a total of only 24 possibilities. I call this the "single-punch method of coding." The purpose of this method is to make possible the punching of multiple simultaneous conditions within each main number, in a single punching field. If the usual consecutive numbering system is used, multiple conditions will require multiple fields. In this method the same punching field can be used for multiple conditions within each main number. The following outline will make this clear:

Main number	Sub-number	
687	1	Inguinal hernia, unmodified
687	— 1	Inguinal hernia, direct
687	— 2	Inguinal hernia, indirect
687	— 3	Inguinal hernia, recurrent
687	—13	Inguinal hernia, direct recurrent

In the last example, direct, recurrent, inguinal hernia is made up by adding the punches of direct and recurrent in the same column, that is, the numbers 1 and 3 are both punched in the second column of the subnumber field. This is possible because only a single punch is used in the subfield for each single specified condition. Advantages of this method are emphasized in the system here described because, as will be explained later, a duplicate card is made for each number punched in a separate field, and by reducing the fields punched it is possible to reduce the number of cards necessary.

OPERATION OF THE SYSTEM

1. *Master Card*—The diagnoses in a given case appear on a summary sheet which the clerk takes up for coding. A special card has been prepared (Figure I) with the punch areas shaded, and between these shaded portions the diagnoses are written in the words given on the summary sheet. A code book alphabetically arranged is provided, and each entry has its code number entered next to it. If the

FIGURE I—A sample master card. The Roman numerals give the order of procedure for completing it.

MASTER CARD

This card is finally filed in patient case no. order. From it is made the first duplicate card.

NAME	AGE	SEX	RACE	REL.	ETHNIC	DIAGNOSIS					SURGERY													
						1	2	3	4	5	6	1	2	3										
Jones	44	M	W																					
496322																								
<div style="display: flex; justify-content: space-between;"> <div style="width: 20%;"> <p>I</p> <p>Diagnoses are written on card from history summary, coded from code book.</p> </div> <div style="width: 70%;"> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td>894</td> <td>143</td> <td>Compound fracture of tibia + fibula</td> </tr> <tr> <td>222</td> <td>1</td> <td>Diabetes mellitus</td> </tr> <tr> <td>281</td> <td>1</td> <td>Hemangioma face</td> </tr> </table> </div> <div style="width: 10%; font-size: small;"> <p>RECEIVED MAYO CLINIC LIBRARY LUTHERS BLDG. 1ST FL. S.E. CORNER</p> </div> </div>																894	143	Compound fracture of tibia + fibula	222	1	Diabetes mellitus	281	1	Hemangioma face
894	143	Compound fracture of tibia + fibula																						
222	1	Diabetes mellitus																						
281	1	Hemangioma face																						
<div style="display: flex; justify-content: space-between;"> <div style="width: 20%;"> <p>II</p> <p>Operations written, also surgeon, hospital, date; coded from code book.</p> </div> <div style="width: 70%;"> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td>8</td> <td>2</td> <td>Open reduction (1) Henderson Co? 10/3/35</td> </tr> </table> </div> </div>																8	2	Open reduction (1) Henderson Co? 10/3/35						
8	2	Open reduction (1) Henderson Co? 10/3/35																						
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>General data punched directly from history summary sheet.</p> <p>III</p> </div> <div style="width: 50%;"> <p>Diagnoses and surgery punched from coding on card visible in punch machine.</p> <p>IV</p> </div> </div>																								



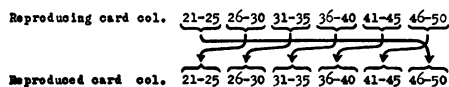
condition is one to be included in the statistical summary, an x is punched in the third column of the main number field, in addition to the number shown in the code book. For instance, to take the example previously cited, if the diagnostic sheet says "carcinoma of the bronchus?—or possibly thymic tumor," the numbers for both diagnoses are punched, but only the carcinoma number has the x punched. There is room for 6 such diagnostic fields. When more are necessary, an auxiliary master card is prepared. There is also finally punched, in a column assigned for the purpose, the number of diagnostic fields which have been used; this is to facilitate the preparation of duplicates, which will be described.

2. *Duplicates*—The completion of the punch cards as so far outlined does not render them ready for practical use. It would be impracticable to sort the entire series of accumulated cards each time the records of cases corresponding to a particular diagnosis were wanted. Moreover, since the diagnosis in question might have been punched in any one of the 6 successive fields, all the cards would have to be sorted in that number of fields, a task which would be prohibitive if a large accumulation of cards was involved. Therefore, some system of duplication of cards is necessary.

There are now prepared from each master card, as many duplicates as there have been fields punched, which will be the same in this system of coding, as the number of diagnoses different in respect of the main number. All the master cards prepared as described, are assembled periodically (each week) and from these, duplicates are made on a special colored card. The reproducing machine of the International Machines Co. permits the easy duplication of cards with a transfer of columns. The duplicate is made in such a way that it has the diagnosis

for which it is made, in the first diagnostic field, no matter in which field it was punched on the master card (Figure II). To accomplish this, the reproducer is wired so that fields 1, 2, 3, 4, 5, 6 of the reproducing card are punched in the order of 2, 3, 4, 5, 6, 1 on the reproduced card (Figure III).

FIGURE III—The scheme of wiring the reproducing punch machine for punching the duplicates in order to make the pertinent diagnosis appear always in the same columns.



Thus, diagnostic field 2 is shifted to the position of field 1. If 3 or more fields have been punched, the second card is now used as the reproducing card, resulting in the third duplicate having the fields in the order of 3, 4, 5, 6, 1, 2; thus the third field is shifted to the position of field number 1. This is continued progressively for as many fields as occupied.\* This procedure is actually extremely easy, because the number of diagnostic fields occupied on the master card has been punched in a special column, and the master cards are first sorted on this column to determine how many duplicates are to be made. I have found that about 1,500 master cards can be duplicated, sorted, and filed in less than 2 hours.

3. *Filing of Cards*—When the duplicating process is completed, we have the master cards and the duplicates to file. The master cards are ordered on the patient case number (not diagnoses), and are placed in a file in this order. Here are kept, in patient number order, consecutively, all the master cards with all the diagnoses written on them. The duplicate cards are ordered

\*Where more than 6 diagnoses have appeared on the summary sheet, an auxiliary master card has been made, so that the operation of making duplicates is the same.

on the 3 columns of the main number in the first diagnostic field (all the relevant diagnoses are now in the first field on the duplicates) and are placed in a file of duplicates, each set of cards corresponding to a particular main number behind a tab card bearing this number.\*

4. *Utilization of Cards After Filing*—Cross-index. If one is concerned with obtaining, for research reference, the case number of histories corresponding to, say, indirect inguinal hernias, which were recurrent, one determines the number for this, which is 687—13. All the duplicates under tab number 687 are removed from the file, and by using the selecting mechanism of the sorting machine, the cards with punches 1 and 3 in column 5 are selected. The selected cards are now ordered on patient case number, and by using the "controlling" device of the tabulator, a list is obtained of case numbers in order, each case appearing only once no matter how many duplicates (visits) this patient has made. At the same time, a count of patients (control count) and visits (card count) is obtained if desired.

Report. For a periodic summary enumeration, one goes to each set of duplicates consecutively from main number 001 forward, and in each group, first sorts on column 3 of diagnostic field 1 to obtain the x cards (those to be reported). The selected cards are then ordered on patient case number, and are run through the tabulator for a "controlled" count (to obtain a count of each patient only once). The main numbers which represent the rubrics used for summary report purposes and totals are printed on the tabulator print sheet consecutively, so that this sheet

need merely be copied to give the final summary report.

#### SUMMARY OF PROCEDURES

1. On registration of a patient, a blank master card is filled in with name and patient number and placed in a file, awaiting the dismissal of the patient. This is merely a check to prevent the missing of histories.

2. After dismissal of patient, the history of the case is sent to the indexing department, where the previously prepared master card (see 1) is removed from the waiting file. On this card, a clerk copies in appropriate spaces the diagnoses and operations; and then codes these on the same card from an alphabetically arranged code book (Figure I).

3. History and master card go to a second clerk who independently checks the coding.

4. The master card is taken up by a punch clerk who punches from the summary sheet of the history, general data (columns 1 through 20). The coding of the diagnoses and operation (see 2) on the card in the punch machine is now visible to the punch clerk, and from these the punching is completed (columns 21 through 65). The number of diagnostic fields occupied is also punched (column 19).

5. The history and master card pass to a checking clerk who independently punches the first duplicate from the notes on history and the coding on the master card. The card is light-checked by this clerk with the original master card. This step is included as a check on punching. These two cards later are checked again for punching, by running them through the checking device of the reproducing machine.

6. The accumulated first duplicates are sorted on the previously punched number of duplicates (column 19) into groups requiring 2, 3, 4, 5, 6 duplicates. Each of these is run through the

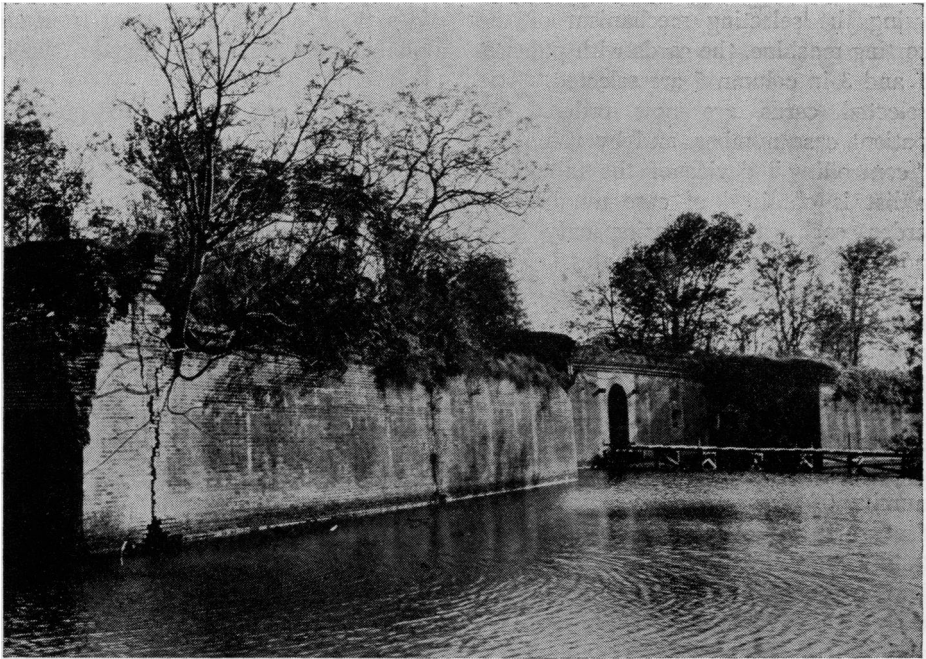
\* In sorting the duplicates, I have arranged to sort with them a specially prepared tab punch card with the number printed on the tab so that the duplicates after ordering are conveniently separated by these marked tabs and easily filed.

duplicating machine the required number of times. For the cards having six diagnoses, for instance, the second duplicate makes the third; the third, the fourth; and so forth, until the sixth is complete.

7. The master cards are sorted on

patient case number and then filed.

8. The duplicates are sorted with inserted tab punch cards on the main number of the first diagnostic field (column 20 through 22). This groups them in main number order and they are so filed.



*Old Fort Macomb, formerly Fort Woods—Located on the Old Spanish Trail 22 miles east of New Orleans at Chef Menteur. Built in 1842 on land which was formerly the property of the King of France. This great mass of embattled brick with all the typical accoutrements of an ancient stronghold of war, with turrets and moats and drawbridges, is one of a considerable number of old forts to be found in the vicinity of New Orleans.*