Disease coding in a problem oriented general practice

EWEN M. CLARK, M.R.C.G.P., D.Obst.R.C.O.G.
Senior Research Fellow, Department of Community Medicine,
University of Southampton

DURING the past decade it has been recognised that disease coding in general practice has its own set of special problems. The Royal College of General Practitioners recognised this and early in the 1960s started to develop a coding system to meet the needs of general practice [1, 2]. This new system is a modification of the *International Classification of Diseases* (*I.C.D.*) [3] and follows, in general, the principles of the *ICD*. Deficits in this system have been described.[4] The *ICD* itself has its critics [5, 6] who find real difficulties in its use. It is unlikely that there is any one classification or coding system which will meet the needs of all users [7].

The University General-Practice Unit, a part of the Department of Community Medicine at the University of Southampton, recognised the many inadequacies of the National Health Service medical record for general practice. All the available options were explored in the search for a suitable medical record system which would meet the quality, content, storage and retrieval needs of a teaching general practice. Eventually the decision was made to embark on the problem orientated record system (P.O.M.R.) [8].

The major philosophy of problem orientated medical records is that of problem identification. While in no way discouraging a pathophysiological diagnostic statement where justified, it insists that the level of accuracy of the statement cannot exceed the supporting clinical data or the doctor's knowledge. Thus a problem statement of frequency and dysuria in POMR replaces the purely assumed diagnostic statement of acute cystitis in traditional medical recording. The symptomatic statement of POMR can only be upgraded to cystitis or urinary tract infection if further objective evidence is available to support the more precise diagnosis.

This situation normally leads to considerable difficulties with a classification and coding system, where one is faced with the choice of delaying encoding the problem until it has reached a level of accuracy consistent with a pathophysiological statement, or accepting an increasing number of statements in code as " other diseases of . . ." Neither of these alternatives is wholly acceptable. If the principle of recording a diagnosis as a problem statement is accepted as Weed proposes, then there is no real alternative to developing a new coding system which will handle ' problem ' and ' diagnostic ' statements equally easily.

Nine criteria

Hence the needs of the user and the necessary qualities of a disease classification system were analysed. Nine criteria emerged which should be fulfilled by the system.

(1) *Simple to use*

The person most involved in the coding activity is usually employed at the clerical level; it is therefore of vital importance that the coding process can be performed without extensive medical knowledge. This difficulty can be overcome by providing the coding clerk with an alphabetical list of diagnostic or symptomatic statements with the appropriate code number beside the statement. Obviously considerable overlap must be allowed, e.g. ischaemic heart disease, myocardial ischaemia, and coronary artery disease will appear separately in the appropriate alphabetical order yet carry the same code

number. This allows flexibility in the doctor's statement without requiring an interpretation on the part of the coding clerk.

(2) *Accuracy*
The decoded statement should mean precisely the same as the coded statement. For example ' sore throat ' should not be changed to ' acute tonsillitis ' during the encoding-decoding process. Two possibilities of error can arise under these circumstances. First, the coding clerk is required to make a judgment for which she is not trained, and, second, a serious diagnostic error can arise in the interpretation of the decoded statement. For example, some doctors may interpret acute tonsillitis to mean an infection with a bacterial agent, such as the betahaemolytic streptococcus, when, in fact, the original statement may have been made to reflect a viral infection. This can give rise to epidemiological errors.

(3) *Unambiguity*
Any system which lumps more than ten per cent of all diagnoses into an ' other ' category is undesirable. This is particularly important where the ' other ' category is very broad. For example, allowing a laceration requiring two sutures to be coded in the same category as a traumatic amputation is inadequate.

(4) *Reflect the diagnostic level*
There is a tendency in general practice to use diagnostic statements of higher precision than the objective evidence warrants. It is highly undesirable to increase this precision as a function of the coding process. For example, a diagnostic statement of ' internal derangement of the knee ' should not be coded as ' torn meniscus '. Equally the precision of this statement should not be appreciably lessened.

(5) *Use by all members of the health team*
To achieve continuity and comprehensiveness in the medical record it is highly desirable, if not essential, that all providers of health care use the same medical record document. The interdependence of social and medical diagnostic statements can be very relevant.

Thus the social diagnosis of " lives alone and cannot be bothered (or unable) to provide an adequate diet " may explain a haemoglobin level of 7·5 gm/100 ml. While treatment with iron may raise the haemoglobin level it contributes minimally to a resolution of this person's total problem.

If the doctor is to recognise the potential significance of social diagnoses, these diagnoses should be given the same prominence as the medical diagnoses. The need to code them is no less than for the medical diagnoses. Therefore, a diagnostic code must accommodate social and behavioural diagnoses as easily as it does a medical diagnosis.

(6) *Machine processable*
The purpose of all coding systems is to reduce a linguistic statement to a numerical or alphanumerical form. It is highly desirable from a data processing point of view to maintain a fixed number of characters or digits in the coded statement, since considerable economies in computer processing will ensue from a field of predetermined size. Naturally the number of digits in any code number must be of sufficient size to accommodate a large number of diagnostic statements. A recent study [9] of 40,000 records showed that 1,624 statements were actually used. It can be seen that a five digit code has the capacity for many times the anticipated need.

(7) *Compatibility with other systems*
In this context compatibility and comparability may mean very much the same thing. It is essential that a precise diagnostic statement can be searched for in separate systems in such a manner that comparable morbidity data can be obtained. This is a function more of the definition of a linguistic diagnostic statement than the code systems, provided the criteria described above are met. In a strict sense, compatibility means the ability of one

coding system to interact with another system through a translating interface. Provided both systems are computer processable this then becomes a relatively simple programming task.

(8) *Meet the needs of the users*

Primary medical care is different from hospital care; the two have different objectives. The latter is based almost exclusively on pathophysiological premises, the objectives being limited to current therapeutic possibilities. Primary care on the other hand fills the large gap on either side of hospital care, dealing with varying degrees of severity in the expression of the disease process on the one hand and, on the other, offering care where no known effective therapy is available.

In the example quoted above of the 'isolate' with iron deficiency anaemia, the hospital doctor may well be delighted to record a reticulocyte response to iron therapy alone, while the primary care doctor will not be content until the nutritional and social inadequacies, which caused the anaemia in the first place, are corrected as well. The needs of these two types of doctor in disease classification, while having a considerable overlap, are quite different in the way they will reflect the type of care given to the patient.

(9) *Provide for additions*

Finally, any coding system should be flexible so that new diagnostic statements can be added when necessary without disturbing the overall structure of the numerical field. This means having enough digits to allow for expansion in all of its sets or sub-sets.

Development of the constructive disease coding system (C.D.C.S.)

In 1971 Hull [10] published the basis of a new coding system, which meets many of the nine listed criteria. In summary, the method lists the physiological systems, such as respiratory, skin, or cardiovascular, and assigns to each a two digit number thus:

respiratory	01
skin	02
cardiovascular	19

Within each system an 'end organ' was identified by a further digit, thus tonsil/pharynx 01–3 or lung parenchyma 01–6. To complete the code a suitable descriptive term was selected from a list of adjectival 'type' words; for example, acute infective –01, haemorrhagic –49 or congestive –50, These two digits were inserted in the code number between the system and the 'end organ' digits. This produces a five digit code in its final form, e.g. acute tonsillitis 01013.

Experiments were carried out in a rural clinic in North Central Florida staffed by physician's assistants [11] to determine its practicality and accuracy. Hull himself performed a preliminary evaluation of the system in his own practice. Both these studies yielded encouraging results [12], [13] but highlighted some serious deficiencies.

To overcome these difficulties some modifications of the system were undertaken. (1) *Physiological system numbers.* Included in this section were code numbers for drug reactions, poisonings and those conditions which could not be diagnosed. The manner in which these had been arranged was not consistent with the rest of the system codes. Some minor changes were made in the section on poisoning and drug reactions reducing the number of system codes.

2. *Symptomatic statements.* Hull had omitted a section for frankly symptomatic statements; we felt these were essential and added a completely new section for them.

3. *Extension of end organ classifications.* We found that a number of changes were necessary in the 'end organ' lists. In some cases an 'end organ' had been completely omitted, and at other times more than one were combined in a single digit. This resulted in a degree of ambiguity when trying to code a precise condition. It was necessary to

redesign a number of the ' end organ ' lists. For example, Hull's list of ' end organs ' for system 17 (dental) consisted of:

0 Embryonic tissue
1 Tooth
2 Gingiva
3 Peridontal tissue
4 Bone

The modified version becomes:

0 Embryonic tissue
1 Tooth
2 Gingiva
3 Peridontal tissue
4 Mandible
5 Mandibular joint
6 Salivary glands
7 Salivary ducts
8 Infant teething troubles
9 Buccal mucosa (including lip).

4. *Alterations of the type code.* The ' type ' code, that is, the code which uses an adjectival word or descriptive word to indicate the disease process, gave us a number of difficulties. The original list contained about 60 terms, e.g. ' acute infective ', ' malignant ', ' hypertensive ', and ' congestive ', We found the list to be inadequate and added about a further 15 terms.

Obviously the ideal size of the list of these descriptive terms would be very large; unfortunately, as the list of terms increases so does the difficulty in selecting an appropriate descriptive term from a number of terms with similar meanings. To keep the list of descriptive terms to a reasonable size we were faced with two options, either we could add a new term to the list, or we could use a term already on the list and in a certain context give it a special meaning.

This can best be illustrated by two examples. First, attempting to code diabetic retinopathy. Retinopathy is identified by the system code 18 and the end organ seven; however the original descriptive ' type ' list contained no term which would indicate the diabetic aetiology of the retinal problem. Since diabetes causes a number of similar conditions it is obvious that a new term should be added to the descriptive ' type ' list, thus we added ' diabetogenic.' This term can also be used in skin ulceration, and neuropathy, when the condition is secondary to the diabetes.

Let us illustrate the second option by coding some of the cardiac conduction problems, taking for example Wolff-Parkinson-White syndrome and the Stokes-Adams syndrome. The system code of 19 refers to the cardiovascular system and the ' end organ ' code nine to the conduction mechanism. Again, the difficulty lies in the selection of the appropriate descriptive ' type ' term. The solution to the difficulty lies in selecting two descriptive ' type ' terms which have not already been assigned in combination with 19—9, and in the context of 19—9 giving the term a diagnostic significance which is only very loosely related to the meaning of the actual word used. Thus: 19459 Wolff-Parkinson White syndrome, the ' 45 ' number is the code given to the descriptive ' type ' term epileptiform; in Stokes-Adams syndrome 19399, the ' 39 ' term is ' retentive ' in the descriptive ' type ' list. A number of conventions such as the two illustrated were employed.

5. *Evaluation.* The final step in the development was to code a large number of diagnoses using the new system. Each diagnosis with its code number was entered on punch cards.

Both an alphabetical and numerical listing were produced. Checks were made for the use of the same code number for more than one diagnosis and the appropriate corrections made. However, in a number of cases the same diagnosis and code number could be generated by diagnostic statements with virtually identical meanings; for example the term " hardening of the arteries " and " arteriosclerosis " both have the same code number.

Evaluation of the constructive disease coding system (C.D.C.S.)

A trial of the constructive disease coding system was conducted in the University Practice at Southampton University. Every doctor-patient contact was recorded on a problem oriented basis. A pilot study was performed using 167 records. The problem statement was coded first by the author using the new system, and then by the practice manager using the Royal College of General Practitioners version of which she had considerable previous experience. To prevent bias in decoding, I called out the Royal College of General Practitioners' number which was translated by the practice manager. This was reversed when decoding the new system.

The decoded problem statement was then compared to the original encoded statement. This comparison resulted in three possibilities; it was either correct, wrong, or doubtful. Another doctor who did not know which system was involved was asked to pronounce judgment on the ' doubtful ' statements. In a number of instances these were adjudged to be correct, a very few wrong, and the remainder were felt to be ambiguous and so remained in the doubtful group. Table 1 shows the result of this trial.

TABLE 1

COMPARISON OF CONSTRUCTIVE DISEASE CODING SYSTEM WITH R.C.G.P. SYSTEM (PILOT STUDY)

n=167	*Correct*		*Doubt*		*Wrong*	
		%		%		%
Royal College of General Practitioners	116	*69·5*	37	*22·2*	14	*8·4*
Constructive disease coding system	154	*92·2*	12	*7·2*	1	*0·6*

Some weeks later the complete process was repeated on 500 consecutive records. On this second occasion, however, I coded both systems, but the same process for decoding was used as for the pilot trial. The results of this comparison are shown in table 2.

TABLE 2

COMPARISON OF CONSTRUCTIVE DISEASE CODING SYSTEMS WITH R.C.G.P. SYSTEM (FINAL STUDY)

n=500	*Correct*		*Doubt*		*Wrong*	
		%		%		%
Royal College of General Practitioners	338	*67·6*	151	*30·2*	11	*2·2*
Constructive disease coding system	465	*93*	33	*6·6*	2	*0·4*

It can be seen that the constructive disease coding system was able to cope with the problem statements considerably better than the Royal College of General Practitioners version. It is certainly reasonable to wonder if this difference exists because of the problem statements imposed by the problem oriented medical records, or is it an inherent difference in the two coding systems. There is no question that the range of precise diagnostic statements in the constructive disease coding system is as extensive as the

Royal College of General Practitioners' code and it also includes problems which are non-pathophysiological but which present common management problems in day-to-day patient care. For example, a patient with mouth ulcers resulting from ill-fitting dentures should have two problem statements: (a) ill-fitting dentures, (b) mouth ulcers, both of which can be readily coded by the constructive disease coding system but cannot be adequately coded and retrieved from the Royal College of General Practitioners' version.

General application of the C.D.C.S.

General application of this system would now be premature. I believe we have a working model of a system which responds well to the ' problem statement ' philosophy of the problem oriented medical record system and is also suitable for ' pathophysiological diagnostic statements.' I believe these needs correspond closely to the needs of general practice in the accurate recording of morbidity. The present library of ' diagnostic ' statements contains over 920 different titles, and the list continues to grow, but recently the rate of new additions to the list has dropped dramatically.

The next developmental step should be for another group to use the existing library for coding purposes and to co-operate in the coding of new statements. Tentative arrangements have been made to establish this.

The final objective is to provide a list of statements sufficiently extensive and comprehensive to be generally applicable in general practice for use by clerical staff with only occasional reference to a doctor.

Discussion

The continuing emergence of general practice as a discipline in its own right supported by a scientific methodology cannot proceed without a method of classifying and comparing data. The uniqueness of general practice lies, in part, in the manner in which ' soft ' data are used by the doctor to form a specific treatment plan. The clinical history and physical signs of an illness process are often of such a nature that a true ' diagnosis ' in the classical sense cannot be made with accuracy or honesty. There has been an understable reluctance on the part of the profession to admit this overtly, yet this is the very nature of general practice.

The constructive disease coding system provides a technique of coding non-specific ' diagnostic ' statements and thus can, to some extent, standardise ' soft ' morbidity data, and provide a basis comparison and quantification. In conjunction with the problem oriented medical record, for investigation and treatment plans for similar problem statements can be related to their outcome.

In addition to these benefits, the understanding of the natural history of a disease process can be studied as the problem statements change from reflecting a presenting symptom to a confirmed pathophysiological entity. The problem list number will provide the link between the changing problem statement codes as the ' diagnostic ' evolution unfolds. The ability to compare this evolutionary process between doctors provides the potential for an exciting new dimension in the study of primary care.

As the problem oriented medical record provides the solid foundation of data, the constructive disease coding system may provide a more flexible tool for measurement and comparison of these data than has been so far available in the study of primary care.

Acknowledgements

This study has been partly supported by Nuffield Foundation Funds.

REFERENCES

1 College of General Practitioners Research Committee (1959). *Journal of the College of General Practitioners*, **2**, 140–159.

2 College of General Practitioners Research Committee (1963). *Journal of the College of General Practitioners*, **6**, 207–216.

3 World Health Organisation (1967). Geneva: W.H.O.
4 Westbury, R. C. & Tarrant, M. (1969). *Canadian Medical Association Journal*, **101**, 82–87.
5 Adelstein, A. M. (1971). *Lancet*, **1**, 655.
6 Clemmesen, J. (1969). *Lancet*, **2**, 162.
7 Hurtado, A. V. & Greenlick, M. R. (1971). *Health Services Research*, **6**, 235–250.
8 Weed, L. L. (1969). *Medical Records, Medical Education and Patient Care*. Ohio: Press of Case Western University, Cleveland.
9 Greenwood, R. M. (1972). *International Journal of Biomedical Computing*, **3**, 43–57.
10 Hull, F. M. (1971). *Journal of the Royal College of General Practitioners*, **21**, 581–585.
11 Henry, R. A. (1972). *Health Services Report*, **87**, 687–692.
12 Clark, E. M. (1971). Diagnostic Coding for Physician Assistant Clinics. Preliminary Report, Department of Community Health, University of Florida.
13 Hull, F. M. (1973). Personal communication.

SIXTH WORLD CONFERENCE ON GENERAL PRACTICE

In association with VIP Travel the College has arranged a tour to MEXICO for this Conference, and it is still not too late for those who wish to go to apply.

The tour departs on Saturday 26 October 1974, from London (Heathrow) airport and returns to Heathrow mid-day on Tuesday 12 November 1974. We have arranged an all inclusive rate of £269·00 per person, which includes air fare and a twin-bedded room with private bath at the first-class Hotel Chateau Royal on a bed-and-breakfast basis, and transfers between the airport and hotel. (There is a supplement for a single bedroom of £40·00.) Alternatively, we have arranged the same facilities at the recommended good second-class Hotel Luma for £253·00. (Single room supplement £33·00.)

Local tours within Mexico have been arranged and full details can be obtained from Mr Endres.

The tour is open to all members of the College and their families, and also their friends, who may wish to take this opportunity to visit Mexico. For those attending the Conference there is an additional registration fee of US dollars 50.00 for members, and US dollars 15.00 for wives.

Please send your deposit of £20·00 per person as soon as possible to Mr K. Endres, Tours Manager, VIP Travel Ltd., 42 North Audley Street, London, W.1 (Phone: 01-629 2243).

STUART CARNE.