

# Summarizing and coding case records: can the task be delegated?

A.M.W. PORTER, MD, MRCGP  
General Practitioner, Surrey

C. TIBBOTT  
State Registered Nurse, Surrey

**SUMMARY.** After several preliminary joint sessions a doctor and a nurse independently constructed summary cards for 100 alphabetically consecutive case records and coded them according to the classification of the Royal College of General Practitioners. The searches and codings were repeated after an interval of at least one month. The error rates of the doctor and the nurse were similar. The inter-observer concordance for the ratings was acceptable for broad diagnostic categories, but was less good when strict criteria were applied. It is concluded that the task may be delegated.

## Introduction

THE teaching and research capabilities of a practice are greatly enhanced by rapid and easy access to the epidemiological data contained in the records of patients. Hitherto this task has been so daunting and the computerization so expensive that it has only been undertaken by the enthusiast. The appreciation that records should have a summary card, the publication of the new College classification<sup>1</sup> and the availability of inexpensive microcomputers place an obligation on all teaching practices to reconsider this matter. Summarizing and coding notes remains, however, a time-consuming task<sup>2</sup> which few doctors would willingly undertake. Is this an appropriate task to delegate to a nurse? This study attempts to answer this question in respect of one practice and to measure the inter- and intra-observer variation between a doctor and a nurse and their relative error rates.

## Method

The records of 117 alphabetically consecutive patients were searched by a general practitioner (A.P.) and a nurse (C.T.). It was agreed independently that 17 patients were no longer in the practice. These records were discarded and this left the case notes of 100 patients from 54 families for study; 79 were those of adults over the age of 18 years.

Most of the records already had a completed summary sheet, but as these proved inadequate for the purposes of the study we each independently completed a new summary on a card designed for the study. This card included an opportunity to record the demographic details of the patient and a structured smoking history. The diagnostic entries were coded according to the classification of the Royal College of General Practitioners.<sup>1</sup> Current drug therapy was recorded using headings from the *British national formulary*.

These summaries were then put aside. After an interval of at least one month each set of notes was again independently summarized without seeing the original results. Only when this was completed were the four sets of summaries compared. The key

codings were taken to be the initial coding of the doctor. Comparison of these with the initial coding of the nurse gave an indication of the feasibility of delegation by measuring the interobserver variation and by comparing error rates. The two comparisons between the first and second codings for the nurse and for the doctor gave a measure of intraobserver variation.

The subsequent comparisons took four forms. The first comparison simply compared the number of entries on each card. The second compared broad diagnostic categories in order to establish the extent of concordance. Minor variations in diagnostic descriptions or mistakes over dates were ignored. If, for example, two diagnostic categories were common to each summary the concordance would have been entered as two. The third comparison was much stricter. The diagnostic category and code, the category — first (F), new (N) or other (O) — and the year all had to coincide exactly if any two entries were to be declared concordant. The only concession was to make the F and N categories interchangeable as, in practice, it is rarely possible to discriminate between them. The fourth comparison involved only the initial comparisons of the nurse and doctor. Each card was scrutinized for mistakes and oversights as judged by the best possible interpretation of a careful joint perusal of the notes. Mistakes were categorized as 'minor' or 'serious'. Most of the minor errors were incorrect entries and most of the serious errors were omissions.

Examples of minor errors were: entry of the wrong year (this was the most common minor error); misreading tonsillitis for bronchitis; confusion of pityriasis rosea and rosacea; coding pneumonia as bronchitis; entering two barium enemas instead of three; and omitting to enter the removal of an intrauterine device. Examples of serious errors were: omission of hysterectomy; omission of intrauterine device; omission of reactive depression; coding a barium meal as a barium enema; tonsils — misreading 'removal' as 'reassurance'; omission of a past history of steroids and omission of tuberculosis.

We each recorded with a stopwatch the time spent searching notes and constructing a summary card and the time spent looking up the code numbers.

## Results

The doctor spent a mean of 8.5 minutes on each record. Only 18% of his time was spent looking up code numbers, the rest was spent searching the notes and writing the summaries. The same figures for the nurse were 12.0 minutes and 9%.

One hundred consecutive code numbers for both the doctor and the nurse were checked for search and transcribing errors and none were found.

The number of entries made from 100 summary cards and the concordance rates are shown in Table 1 and the number of correct entries and errors made are shown in Table 2. Inspection shows little difference between the doctor and the nurse. The nurse was more consistent than the doctor and achieved an intraobserver concordance rate of 83% for diagnostic categories, that is 83% of the entries on her first set of cards appeared on the second set. The concordance of diagnostic categories for the nurse and doctor for their first coding was 70% but with the application of strict criteria this figure fell to 47%.

A minor or serious error was made by the doctor in 27% of

© *Journal of the Royal College of General Practitioners*, 1986, 36, 67-68.

**Table 1.** Concordance between the doctor and the nurse for coding 100 summary cards.

	Total No. of entries	Comparison by diagnostic categories		Comparison applying strict criteria	
		No. of concordant entries	Concordance rates (%)	No. of concordant entries	Concordance rates (%)
<i>Doctor</i>					
First coding	456}	353	77.4 <sup>a</sup>	245	53.7 <sup>a</sup>
Second coding	446}				
<i>Nurse</i>					
First coding	546}	454	83.2 <sup>a</sup>	381	69.8 <sup>a</sup>
Second coding	571}				
<i>Doctor versus nurse</i>					
Doctor's first coding	456}	321	70.4 <sup>b</sup>	216	47.4 <sup>b</sup>
Nurse's first coding	546}				

<sup>a</sup>Relative to first coding. <sup>b</sup>Relative to doctor's first coding.

**Table 2.** The number of correct entries and errors made by the doctor and the nurse for their first codings. The percentage of the total number of correct entries and errors which were made by the doctor is given in parentheses.

	Number of correct entries	Number of minor errors	Number of serious errors	Total
Doctor	332 (43.6)	52 (58.4)	72 (47.7)	456 (45.5)
Nurse	430	37	79	546
Total	762	89	151	1002

the diagnoses, as opposed to 21% for the nurse (Table 2). The nurse made only a few more serious errors than the doctor — 68% of all the errors made by the nurse were serious versus 58% for the doctor.

## Discussion

Few doctors would wish to include all problems on the summary card. In four preliminary sessions before coding began we agreed which problems to include and which to exclude. Despite this misunderstandings still occurred for it proved impossible to identify and anticipate all the potential areas of confusion. For example, we had agreed to code all instances where a woman had received a hormone in the past; nonetheless, the nurse failed to code women who had been taking oral contraceptives. We therefore recommend as many preliminary sessions as possible. The nurse compensated for possible difficulties by entering and coding more diagnoses than the doctor (546 compared with 456 for the first codings).

The intraobserver concordance rates gave cause for concern. The concordance rate of the nurse was better than that of the doctor (Table 1). Interpretation of handwriting proved to be a difficulty for both the doctor and the nurse. There was evidence to suggest that the doctor concentrated on entries in his own hand and sometimes overlooked entries in other handwriting.

The low concordance figure between the doctor and the nurse when strict criteria were applied was disconcerting but of relatively little importance. What matters is that an episode of an illness should be recorded and not if the year entered is incorrect.

The doctor's excess of minor errors may be more apparent than real because the nurse's extra codings were not, in retrospect, assessed with the same care as those held in common. The doctor's serious errors were often due to carelessness, while the

nurse's were mostly due to lack of familiarity with medical terms or to a misunderstanding. The error rates of 72 serious errors out of 456 entries (15.8%) for the doctor and 79 out of 546 (14.5%) for the nurse are probably acceptable. It is doubtful if anyone else would do much better given the limitations of constructing morbidity codes and the tedium of the task. It is, however, a limitation which must always be considered when interpreting retrospective morbidity codes from general practice.

There are many examples in the literature of the delegation of tasks in general practice to nurses but none are directly relevant to this study.

It has been suggested that this study may not be particularly valuable since the first few letters of a diagnosis can now be entered into a computer and the appropriate coding will then be automatically selected and stored. This criticism can be challenged in two respects. First, the exercise greatly improves the quality and succinctness of summaries. The College classification<sup>1</sup> should be used for the construction of all summaries, even if it is not intended to computerize records. Secondly, every scheme of morbidity coding should have a retrospective element. There is little point in starting a coding prospectively on 1 January 1985 and omitting to code that a woman had a panhysterectomy for carcinoma of the cervix on 31 December 1984. If there is to be a retrospective element then it should be as accurate as possible. This means searching the notes and producing a summary card. The construction of these cards took almost all of the time in this study. Even if code numbers were not determined *Occasional paper 26* would have to be referred to frequently in order to discover the classification options for a given condition. Thus, there is little extra work involved in entering the code numbers. This leaves the practice the option of purchasing an inexpensive computer and entering the codes directly at a later stage.

This study demonstrates that for one practice the task of summarizing and coding entries may be delegated to a nurse.

## References

1. Royal College of General Practitioners. *Classification of diseases, problems and procedures 1984. Occasional paper 26.* London: RCGP, 1984.
2. Porter AMW. New RCGP classification. *J R Coll Gen Pract* 1985; 35: 43.

## Acknowledgement

We thank Dr Richard Morris for advice.

## Address for correspondence

Dr A.M.W. Porter, 37 Upper Gordon Road, Camberley, Surrey GU15 2HJ.