Malaysian J Pathol 1990; 12(2): 101 – 106

A dBASE III SYSTEM FOR MANAGING 35 MM SLIDES IN PATHOLOGY

K T WONG, MBBS. MPath and K S CHAN, MBBS.

Department of Pathology, Faculty of Medicine, University of Malaya.

Summary

We describe the design and Management of a 35 mm slide database using a menu-driven dBASE III™ PLUS programme and a microcomputer in a large department of pathology that also caters for the individual pathologist. Existing systems described in the literature are geared towards slides of general medicine and do not address the needs of the individual pathologist. A total of 11,481 slides in fhe Department of Pathology, Faculty of Medicine, University of Malaya, were filed into a single database with each record representing one slide. Nine fields which comprised the slide accession number, reference number, slide category, SNOMED codes, and a description of the slide in natural language, seemed adequate for slide definition. The menu-driven programme had functions which included the abilities to add, delete, edit and back-up records, and to search for desired slides. Although slides may be searched for in various fields, we found that searches using natural language alone were both comprehensive and efficient, provided a standard format of description was adhered to and data entries scrutinized carefully for errors. We believe therefore, that for the pathologist working alone, coded language fields are not absolutely necessary, as manual coding and additional data entry can be time consuming. As expected, for databases larger than 10,000 slides, a 80286 microprocessor-based microcomputer was more efficient. We are of the opinion that a system such as ours is very useful for a large department of pathology or the individual pathologist to file and retrieve 35 mm slides.

Keywords: 35 mm slide, microcomputer, dBASE III™ PLUS, database, pathology, SNOMED, natural language.

INTRODUCTION

The 35 mm slide remains an important visual aid in the dissemination of medical knowledge.[1,2] This certainly applies to the field of pathology where the use of such slides is often crucial in the lecture hall, research laboratory, seminars and clinico-pathological conferences.

As the collection of such slides grows, there is a real need for an efficient system to file and retrieve them. Several authors have written about filing and retrieval systems for 35 mm slides of general medicine using microcomputers.[3,4,5] We are, however, not aware of any published report on the use of microcomputer systems to manage 35 mm slides in a large department of pathology, or which addresses the specific needs of the individual pathologist in such a task.

We, therefore, describe the design of a database for 35 mm slides of a largely pathological content, and the problems and solutions inherent in the management of such slides, with particular reference to the individual pathologist. We used easily available computer software and hardware. The minimum hardware requirements are an 8088 IBM compatible microcomputer with 640 KB RAM, a fixed (hard) disc and a floppy disc drive, a monochrome monitor and a dot-matrix printer.

MATERIALS AND METHODS

Database

A single database (SLIDE.dbf) was used to file 11,481 slides of the Department of Pathology, Faculty of Medicine, University of Malaya. Each slide was filed as a single record. New slides acquired were filed as soon as possible while the filing of backlog slides was done as an on-going process, generally but not necessarily in reverse order of their slide numbers.

Two academic staff wrote down the entries for the various fields of each record, after reviewing each slide individually. A standard format of data entry and recording was adhered to at all times. The lists were then handed to a typist to enter into the computer. All alphabetical data were entered automatically in uppercase (capital) letters using a custom screen. Entries were only considered final after at least 2 checks for mistakes by the same academic staff.

Address for reprint requests: Or K T Wong. Department of Pathology. Faculty of Medicine, University of Malaya, 59100 Kuala Lumpur, Malaysia.

Software and hardware

The commercially available software programmes used were PC DOS version 3.3 (International Business Machines Corp., Boca Raton, Fl,, USA), dBASE III PLUS version 1.1 (Ashton Tate, Torrance, CA, USA) and WordPerfect version 5.0 (WordPerfect Corp., Orem UT, USA). The hardware consisted of a 4.77 MHz, 8088 microprocessor-based IBM-compatible microcomputer with a 20 MB fixed disc, a floppy disc drive, 640 RAM, standard keyboard, CGA colour monitor, and a dot matrix printer. We also used a 16 MHz, 80826 microprocessor-based, IBM-compatible microcomputer with a 20MB fixed disc and a floppy disc drive to test our system.

Programme design

All files, including SLIDE.dbf, were kept in the fixed disc. A menu-driven programme (written by K S Chan) provided easy access to the database. Its main programme (SLIDE.prg) and other programme files were written in dBASE language using WordPerfect instead of the in-built dBASE III word processor as the former has a superior text handling capability.

Three sub-directories were created: 2 sub-directories to keep PC DOS and dBASE III PLUS respectively, and the third contained all other programme files including SLIDE. dbf, SLIDE.prg, etc.

RESULTS

The structure of SLIDE.dbf is shown in Table 1. Nine fields were used to define each record. Theoretically, 99,999 slides can be accommodated in the database. The SLIDENO field contained the unique accession number given to each slide at the time of acquisition. SLIDE.dbf was indexed by this accession number.

In the DESCRIBE field (Table 1), natural language keywords were used to record suitable descriptions for the slide. This usually meant listing the organ or tissue, diagnosis, pathological features, and special stains, in this order. British spelling was used. As far as possible, we used nouns, e.g., a slide showing lung involved with a metastatic lesion from cervical squamous carcinoma would be entered as "LUNG, METASTASIS, SQUAMOUS CARCINOMA, CERVIX". Short forms were generally used only for special steins e.g. "PAS" for "periodic acid-schiff".

The CASENO field (Table 1) contained the case number of the specimen/patient of which the slide was taken. This could be the autopsy, surgical, cytology or other reference numbers. Apart from making it possible to obtain more information about the case than can be provided by the DESCRIBE field, this reference number provided a link between slides belonging to the same specimen/patient that were taken and filed at different times.

Five axes of the Systematised Nomenclature of Medicine (SNOMED) codes,[6] provided additional slide definition (Table 1). The codes based on procedure and occupation were not included. Since most of our slides had only one diagnosis, only one field for each of the five axes was needed. For all newly acquired slides, we entered the SNOMED codes after manually looking up the appropriate codes. In a majority of cases, it was necessary to enter only the topography and morphology codes. As most of our backlog slide collection have yet to be coded, we chosed not to code them when entering data initially. This was necessary as coding was very labour intensive. Only after entry of all backlog slide data (other than codes) into the computer was completed, did we start to enter the codes.

The field CATEGORY (Table I) recorded the type of slide and this was designated by a single letter or a combination of 2 letters e.g. "G" = Gross specimen or patient; "H" = Histology or microscopic appearance; "C" = Cytology, and "EH" = Electron microscopy and Histology, and so on.

The menudriven programme consisted of a main menu and 4 submenus (Fig. 1). From the main menu, submenus were selected by pressing a single key which would then replace the main menu on-screen. Likewise, choices made on the submenus would reveal other screens or further submenus. The important functions performed by our programme included the abilities to add, delete (or cancel), edit (or change), back-up (or duplicate into diskettes) records, and to search for required slides. Several pages of documentation provided help to the novice user.

The search function allowed for slide searches using slide numbers, case numbers, SNOMED codes, or keywords. In practice, since most of our slides were not coded, most slide searches were done using keywords. This, we have found to be entirely satisfactory. A maximum of 3 keywords may be searched for at one go. In addition, an option to exclude one word from the search was provided. In instances where the length of the word exceeded the 20 spaces provided, part of the word usually sufficed.

TABLE 1
STRUCTURE OF DATABASE **(SLIDE.dbf)**

Field	Field Name	Type	Width *	Content
1	SLIDENO	Numeric	5	Unique accession number
2	DESCRIBE	Character	60	Organ diagnosis, special features, etc.
3	CASENO	Character	10	Reference number
4	CODET	Character	5	SNOMED code: topography
5	CODEM	Character	6	SNOMED code: morphology
6	CODED	Character	5	SNOMED code: disease
7	CODEE	Character	4	SNOMED code: aetiology
8	CODEF	Character	5	SNOMED code: function
9	CATEGORY	Character	2	Type of slide

* Total width = 102

For slide searches using two or more keywords, the combination of words must all be matched before the record was selected, i.e. a conjunctional (logical and) search. We did, not provide for an alternation (logical inclusive or) search because we felt this would complicate the search procedure. Moreover, this type of search can easily be done by repezted searches using single keywords.

Using the 8088 microcomputer, and allowing the entire database to be searched, about 9, 10, and 13 minutes were needed to search 1, 2, or 3 words respectively (the time interval being measured from the completion of keyword/s entry to the completion of search). The time needed to search for 3 words and exclude I word was about 20 minutes. In practice, however, less time was often needed, as our database contained many slides of the same subject which were of a similar quality so that once the user felt he had an adequate number of desired slides, the search could be terminated by pressing the escape key without having to search through the entire database.

In terms of speed, excluding one word was equivalent to searching for an additional word, while the length of the word did not alter the speed appreciably. There was no difference in speed even if words were entered in different orders, e.g. in searching for "OVARY", "CYSTADENOMA" and "BORDERLINE", entries made in this order were the same as the reverse order. We tested the search function with the 80286 microcomputer and the speeds were only about 2 minutes in all instances.

The search output list took 2 optional forms (the BRIEF or the COMPLETE report formats). In most instances, we used the BRIEF format (fields SLIDENO, CASENO CATEGORY, DESCRIBE only) as the, information in these fields was usually adequate for slide selection (Fig. 2). There was the added advantage that the on-screen display

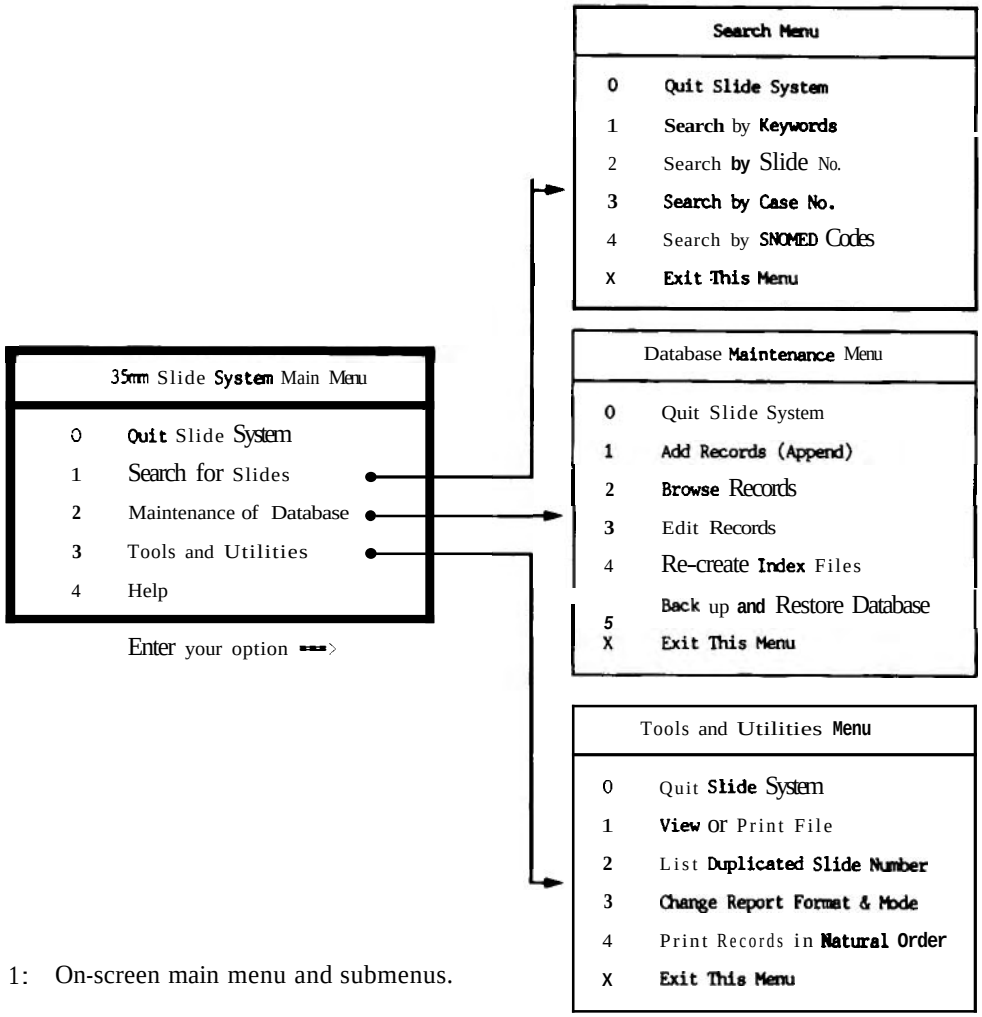

Fig. 1: On-screen main menu and submenus.

of search output lists was simple and uncluttered. For the purpose of checking data entries. the COMPLETE format (all fields displayed) was used.

With the search list (Fig. 2) as a guide, suitable slides were retrieved from a storage unit consisting of slide boxes, where slides were arranged according to the accession numbers. In addition to a print-out, our programme offered an option to save the search list into an ASCII file on a floppy disc for future reference.

DISCUSSION

One of the most important function any slide filing system should have is the ability to retrieve desired slides expeditiously and effectively because valuable teaching slides are as good as lost if they cannot be retrieved.[1,3,7] To achieve this, accurate slide description or definition is a prerequisite.

In the context of anatomical pathology databases, it has been suggested that the use of both coded language (e.g. SNOMED codes) and natural language is probably ideal for recording diagnoses.[8,9] Consequently, we used both coded and natural language for slide definition in our database.

If the SNOMED coding system is used, S axes (topography, morphology, aetiology, function and disease) are adequate for most diagnostic details.'[0] Although our experience with the SNOMED code for slides is limited, as most of our slides have yet to be coded, we believe that the multiaxial SNOMED code is at least as good as, if not better than, the other coding system which has been used by others for slide definition, viz., the WHO International Classification of Diseases.[3,7]

In recent years, controversy has arisen regarding the relative advantages of using coded language and natural language in computerized medical records. Some authors[11] were of the

Word(s) Searched : PYELONEPHRITIS

Page No. 1

Slide#	Case#		Category and Description of Slides
723		H	KIDNEY, DIABETES MELLITUS, CHRONIC PYELONEPHRITIS
725		G	KIDNEY, CHRONIC PYELONEPHRITIS
726		H	KIDNEY, ACUTE PYELONEPHRITIS
727		G	KIDNEY, ACUTE PYELONEPHRITIS
728		H	KIDNEY, ACUTE PYELONEPHRITIS
729		H	KIDNEY, ACUTE PYELONEPHRITIS
730		G	KIDNEY, ACUTE PYELONEPHRITIS
731		G	KIDNEY, PYELONEPHRITIS WITH ABSCESS
732		H	KIDNEY, PYELONEPHRITIS, ATROPHY
1310		G	PERICARDIUM, SEPTIC PERICARDITIS, PYELONEPHRITIS
1448		G	KIDNEY, ACUTE PYELONEPHRITIS, HYDRONEPHROSIS
1457		G	KIDNEY, ACUTE PYELONEPHRITIS
1458		G	KIDNEY, ACUTE PYELONEPHRITIS
9855	S4064/72	G	KIDNEY, CHRONIC PYELONEPHRITIS
9856	S4064/72	G	KIDNEY, CHRONIC PYELONEPHRITIS
9857	S3982/72	H	KIDNEY, PYELONEPHRITIS, GRAM -VE SEPTICAEMIA
9989	S3207/72	G	KIDNEY, CHRONIC PYELONEPHRITIS, HYDRONEPHROSIS
9990	S3207/72	G	KIDNEY, CHRONIC PYELQNEPHRITIS, HYDRONEPHROSIS
9991	S3207/72	G	KIDNEY, CHRONIC PYELONEPHRITIS, HYDRONEPHROSIS
9992	S3207/72	G	KIDNEY, CHRONIC PYELONEPHRITIS, HYDRONEPHROSIS

Fig. 2: An example of a search output list.

opinion that the latter will assume a greater role, while Friedman[8] suggested that natural language searches could sometimes be more efficient than SNOMED searches.

In situations like ours, where coding lagged behind natural language slide description, we found the use of natural language alone in slide searches to be entirely satisfactory, provided a standard format of description was strictly adhered to. Moreover, the DESCRIBE field. where natural language slide definition was entered, was flexible enough to accommodate such diverse slide descriptions as "vascular invasion" or "hyperchromatic nuclei" that would otherwise be impossible with established coded language. Likewise, multiple diagnoses and cross-indexing were easily accommodated.

We believe, therefore, that for the individual pathologist operating alone, or who has scarce clerical resources at his disposal, it might be advantageous not to include a coded language field or fields in the database. This is because manual coding and additional data entry involved can be very time consuming.

As the CATEGORY field seems to be very useful for further slide classification in a slide database with a large pathological content we recommend its inclusion in the slide database design. Thus, we believe the minimal number of fields in any pathological slide database should include – in the context of our database – the fields SLIDENO, CASENO, CATEGORY and DESCRIBE. New fields may be added whenever the need arises.

One problem that has to be addressed in using natural language description is that preferred choices of keywords must be communicated to other users in a large department. A list of commonly used keywords made available to all users would partially overcome this problem. For the single-user database, this problem would not arise if preferred keywords are used consistently all the time.

As one's database grows, the time needed to search for slides naturally increases. For a database of about 10,000 records, using a 4.77 MHz, 8088 microcomputer is probably too slow to be practical. An 80286 microcomputer, on the other hand, would be more suitable since slide searches by natural language keywords can be achieved in only about 2 minutes. Thus, in choosing computer hardware,

the projected size of the database should be borne in mind.

Typographical errors during data entry would undoubtedly hamper slide search using any field and may even render the slide inaccessible. This is particularly true for codes.[9] Thus, careful checking of data entries cannot be overemphasized.

Although the use of a menu-driven programme generally simplifies access into the database, it may impose certain limitations to the experienced user well versed with dBASE language. In searching for slides e.g., one may be able to execute a more specific slide search "manually" than may be possible with a menu-driven programme. Nevertlieless, for a collection of slides which caters for many users, some of whom may be computer illiterate, a menu-driven programme offers a practical solution.

Automated coding if incorporated into the system, will enable coding to be done speedily and remove the tedium of manual coding. Significant advances in optical disc technology such as WORM (Write Once. Read Mainly) has enabled mass storage of information and if available, can be used to store slide images[12] so that viewing of slides for selection becomes much easier as the image can be brought on-screen.

We wish to stress that our system is a low-cost, stand alone system, operable on a relatively unsophisticated microcomputer. We find our computerised system efficient in the management of 35 mm slides in a large department of pathology. It offers the speed and comprehensiveness needed to cut down the time -pent searching for desired slides, thereby reeing the pathologist to concentrate on more productive work. The individual pathologist. too, should have no problem using it to build up a 35 mm slide database, manage and retrieve slides needed for teaching and demonstrations.

ACKNOWLEDGEMENT

We are thankful to Dr Jayaranee for her assistance in listing and checking data entries and to Ms Sarimah for entering data into the computer. We also would like to thank Prof. L M Looi, Head, Department of Pathology, University of Malaya for permission to publish this article.

REFERENCES

1. Barker VF, Harden RM. The storage and retrieval of 35 mm slides. Med Educ 1980; 14 (1): 59–71.
2. Hedley AJ, Morton R. The clinical slide library: A valuable learning resource in continuing medical education. Med Biol Illust 1976; 26: 203–7.
3. Gilson CC, Collins JM. Use of the microcomputer in a department of medical illustration for retrieval of clinical teaching slides. J Audiovisual Media Med 1982; 5: 130–4.
4. Brown SE. Ashton Tate's dBASE II as a patient-index manager. J Audiovisual Media Med 1985; 8: 5–8.
5. Robertson SJ. 'Cardbox' – an ideal program for clinical filing. J Audiovisual Media Med 1985; 8: 19–21.
6. Systematized Nomenclature of Medicine (SNOMED), ed 2. Skokie, Ill, College of American Pathologists, 1979.
7. Nayler J. The CPHA's ICD-9.CM as a coding source for niedical slide filing. J Audiovisual Media Med 1990; 13: 55–60.
8. Friedman BA. The impact of new features of laboratory information systems on quality assurance in anatomic pathology. Arch Pathol Lab Med 1988; 112: 1189–91.
9. Ulirsch RC. Status of anatornic pathology data management systems. Arch Pathol Lab Med 1984; 108: 884–7.
10. Cote RA, Robboy S. Progress in medical information management. Systematised Nomenclature of Medicine (SNOMED). JAMA 1980; 243 (8): 756–62.
11. Moore GW, Hutchins GM, Miller RE. Strategies for searching medical natural language text. Am J Pathol 1984; 115: 36–41.
12. Aller RA. Surgical pathology computer systems have come of age. Arch Pathol Lab Med 1988; 112: 1186–7.