

## A Computerized Database and Statistical Analysis System for Radiotherapy

Sung Whan Ha, M.D., Il Han Kim, M.D.  
Wee Saing Kang, Ph.D. and Charn Il Park, M.D.

*Department of Therapeutic Radiology, College of Medicine, Seoul National University, Seoul, Korea*

A computerized system for database of radiotherapy patient and for its application was developed in 1987 and has been utilized till now. A radiotherapy planning computer (Eclipse S-140) operated under AOS (Advanced Operating System) is the main processing unit of the system which was programmed with Fortran-5. Records of 30,000 patients can be separately registered and data of 5 courses of radiotherapy delivered to one patient can be separately registered but structurally linked together. The same environment is allowed for 60 follow-up data. Our system's utility is very convenient to use and provides simple or conditional list of records or items, periodic statistics concerning many parameters and survival or complication analysis of stored database or data manually put in. Structure, operation and several retrieval formats by data processings are reported.

**Key Words:** Computerized database system

### INTRODUCTION

A patient information system is indispensable in managing a large amount of medical records and their utilization. Several computerized systems replaced traditional record system which consisted of medical charts, index or classified cards after security, easy and rapid management, and systematic retrieval of data with computers had been widely recognized<sup>1)</sup>. What is more, statistical analysis such as survival which is important in the radiotherapeutic fields could be directly done from stored data.

The PC-12 data management system, a data management system installed in PC-12 radiotherapy planning computer (Artronix, USA), was the first data management system in our department. This system, however, had inadequate data structure for our patients<sup>2)</sup> and had long operating time because data was stored in rolls of magnetic tapes. In 1986 new radiotherapy planning system (Capintec, USA) with patient registry system using Eclipse S-140 computer (Data General Corp., USA) was installed but soon we also found that above registry system was dedicated to American users only<sup>3)</sup>. So we planned to develop our own computerized database system useful for our

patients.

This paper describes the present status of system operation, details of data processing, and ways of utilities of the computerized database system we developed.

### HARDWARE AND SOFTWARE

We decide to use our planning computer (Eclipse S-140, Data General Corp., USA) as the main unit for database system. Excluding areas for radiotherapy planning, 512 kByte CPU with 16 bit and three kinds of auxiliary memory units (143 MB hard disk, floppy diskettes, and 10-inch magnetic tapes) is available for us. Our program language is the FORTRAN-5 and the computer is operated under the Advanced Operating System (AOS). The patient database system is allocated to one pathname among the six and is allowed to access from any three of five terminal consoles.

#### 1. System Operation

When a patient has finished a course of radiotherapy, a duty resident reviews the radiotherapy chart and completes a piece of registry sheet designed as input format of initial registry for the system. As a matter of fact, database system was developed to adopt the registry sheet used in our department since 1982. After a staff radiotherapist scrutinizes the registry, an operating secretary put patient information in the system through a terminal

This work was partly supported by the 1986 SNUH Special Clinical Research Fund.

console (Visual 300).

Whenever a patient visits outpatient clinic after completion of radiotherapy for periodic examination, a piece of follow-up registry sheet is written and its contents are put in the database system.

## 2. Registration by Coding System

The 37 items in initial registry consists of hospital number, radiotherapy number, demographic data, tumor-related data, and treatment-related data (Table 1). For our convenience we decide that the registration number (record number) is identical to the radiotherapy number for every patients.

Items are registered as one of four types of data (field); text, number, date and code. We defined several codes for many items. In cases of metastasis, primary organ and primary histology is registered as separate fields in addition to the treated organ and histology but code system is identical. The code system for organ and histology is quite different from ICD-O classification or WHO pathologic code system.

The above code number for each item and encoded details is not displayed on CRT at the time of data input but immediately after data input for confirmation. Other codes for items of stage, pathologic grade, response of tumor...can be put in

**Table 1.** List of 37 Items Included in the Initial Registry

Group	Items
Identification	Hospital number Radiotherapy number (5 digits) ID number Name Age (YYMM) Sex
Clinical findings	Date (YYMMDD) : of diagnosis, of operation, Primary tumor : Organ (code), Organ (specify) : Histology (code), Histology (specify) Pathologic grade TNM classification Stage or Group Performance status : (code by ECOG scale) If treated tumor is : Organ (code), Organ (specify) metastatic tumor : Histology (code), Histology (specify) If multiple primary : Organ (specify), Histology (specify), tumor (+) : Date, diagnosis of another primary tumor If previous tx (+) : Types of previous treatment, due to recurred or : Date of previous treatment multiple primary
Radiotherapy factors	Date (YYMMDD) : Radiotherapy started : Radiotherapy finished : Base of follow-up Combination treatment Aim of radiotherapy Treated volume Modality of radiotherapy Applied dose Response of tumor Complication during treatment : (code), (specify)
Comment	(16 characters are allowed to be put in)

**Table 2.** List of 13 Items Included in Follow-up Registry

Group	Items
Identification	Hospital number, Radiotherapy number, ID number, Name
Date (YYMMDD)	Base of follow-up Last follow-up
Period (YYMM)	Period interval from base to last follow-up
Follow-up information	Disease state at follow-up (code), Date, diagnosis of recurrence, Treatment after recurrence, Complication after treatment : (cord), (specify) Next follow-up period

likewise.

There are two ways of data input in general, one can put data in the system either by choosing code and confirming data input at every step as above or by putting code in at once without confirmation to save time.

Important data among the 13 items in follow-up registry are base of follow-up, date of follow-up, tumor and complication status (Table 2). The method of putting code in system is not different from that for initial registry.

### 3. Structure of Database

The system's main parts are database and utility. The initial registry is classified as the first to the fifth initial registries for a single patient. The database system have capacity to store 30,000 data for the first initial registry. The second to fifth initial registry is linked to the first initial registry, and used for patients who gets the second to fifth course of radiotherapy because of reirradiation, regional relapses, distant metastasis, or other primary tumor(s). The total capacity to store the 2nd, 3rd, 4th, and 5th initial registry is 3,000, 1,000, 500 and 300 record, respectively. For the follow-up registry of single patient, maximum 60 data can be stored in total and is also linked to each initial registry.

Correction of error(s) in registry or addition of data to the blank item(s) can be easily done.

### 4. Utility

Utility part provides list of records/items, peri-

odic statistics, and survival and complication analysis. Details of each package is described in each section.

#### 1) List of Records/Items

Full items of initial registries and linked follow-up registries of a single record or records in any range, for example from registry No. 2,001 to registry No. 3,000, or the whole registered records can be retrieved with the same format as the input format by manipulation of searching scope.

As an additional function, list of registration number of blank records (records without registered data) can be obtained from the whole database.

Searching records that fulfill certain condition(s) of any item can be easily done. Maximally 10 items among the 37 items of the initial registry could be conditioned in searching. Six kinds of conditional command of EQ (equal), NE (not equal), GT (greater than), GE (greater than or equal), LT (less than), and LE (less than or equal) can be used to any items. After conditioning was completed, necessary items can be selectively listed according to a user's decision (Table 3).

#### 2) Periodic Statistics

The patterns of new records (patients) by age, sex, tumor, treatment aim, status of combined treatment, and completeness of radiotherapy can be retrieved with monthly or annual interval (Table 4). For a quick reference of follow-up status of treated patients, items of radiotherapy number, name, last follow-up date, and code of last follow-up status can be retrieved (Table 5).

#### 3) Survival and Complication Analysis

For registered patients whose characteristics are restricted with conditional commands, survival or complication analysis can be obtained in 2 formats. The first format shows a chronologic status of each patient using predetermined status code after radiotherapy, that is, a linear plot of patient's status during follow-up period (Table 6–9). This enable us rapid recognition of changing status of certain patient(s) at a glance. The code of '0' denotes 'no data input'. The second format shows tabulation of actuarial survival rates or complication rates with interval of 3 months calculated by the life table method<sup>4)</sup> (Table 10).

Independent using of this software is also allowed. The format calculated by the life table method can be retrieved by manual input of analyzed survival or non-registered patients or both. In this case the first format isn't displayed.

**Table 3.** Example Format of List of 8 Items (radiotherapy number, name, age (YYMM), base of follow-up (date), histology code, specification of histology, code for host performance status, and code for tumor response at the time of completion of radiotherapy) of the Patients Who were Selected with 4 Conditions ; organ is cervix, registered (thus treatment started) in 1980, planned dose was fully applied, radiotherapy aim was radical.

LIST OF ITEM				Ther. Radiol. SNUH			
CONDITIONS							
1	:	8	GE 800101				
2	:	8	LT 810101				
3	:	11	EQ 61				
4	:	29	EQ 1				
5	:	32	EQ 4				
ITEM NUMBER (S) FOR LISTING							
			ITEM ( 1) = 2				
			ITEM ( 2) = 3				
			ITEM ( 3) = 4				
			ITEM ( 4) = 10				
			ITEM ( 5) = 13				
			ITEM ( 6) = 14				
			ITEM ( 7) = 22				
			ITEM ( 8) = 33				
1	:	416	OH C----- I-	4900	80 116	6	sq not spe 2 2
2	:	454	PARK D--- B---	5000	80 214	1	sq L N K 2 2
3	:	485	LEE S---- N---	5600	80 229	2	sq L K 2 1
4	:	489	HEO P--- S---	5200	80 313	1	sq L N K 3 2
5	:	518	CHOI C----- B--	5600	80 327	0	0 0
6	:	533	KIM O-- P---	4800	80 4 4	1	sq L N K 0 1
7	:	550	JIN O--- B---	5000	80 415	1	sq L N K 0 1
8	:	560	KIM C----- K---	4300	80 423	6	sq not spe 0 2
9	:	561	YOO T--- S---	5400	80 423	6	sq not spe 0 0
10	:	564	BAE M--- J-	5700	80 423	2	sq L K 0 1
11	:	567	YOON D--- S----	5400	80 3 0	6	sq not spe 0 0
12	:	572	LIM Y---- K-----	5000	80 430	1	sq L N K 0 1
13	:	573	SONG Y---- S----	4100	80 430	6	sq not spe 0 1
14	:	579	KIM S---- A-	6500	80 5 7	2	sq L K 0 1
15	:	602	BYEON C----- H---	5900	80 516	6	sq not spe 0 1

**Table 4.** Monthly Distribution of all the Cervix Cancer Patients who was Registered in 1980 by Applied Dose per Planned Dose Retrieved by Utility Programme of Periodic Statistics

APPLD DOSE	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TTL
< Half	1	1	--	--	--	2	2	2	2	--	--	--	10
< 3/4	--	--	1	--	--	1	--	--	--	--	1	--	3
< Full	--	1	--	--	1	--	--	--	--	--	--	2	4
Full Dose	6	5	14	12	9	6	9	9	5	8	3	10	96
No inform	1	--	--	--	--	--	--	1	--	--	--	1	3
<b>Total</b>	<b>8</b>	<b>7</b>	<b>15</b>	<b>12</b>	<b>10</b>	<b>9</b>	<b>11</b>	<b>12</b>	<b>7</b>	<b>8</b>	<b>4</b>	<b>13</b>	<b>116</b>

**Table 5.** An Example List of Radiotherapy Number, Name, Last Follow-up Date, and Code for the Disease Status at that Date of the Same Cervix Cancer Patients as Those in Table 4.

NO	NAME	FUP	DATE	ST	NO	NAME	FUP	DATE	ST
408	CHOO S--- J-	80	9 17	2	715	KIM K----- S---	89	11 17	1
416	OH C----- I-	82	4 26	1	721	KIM H--- B---	87	5 14	1
454	PARK D--- B---	89	1 19	1	730	KIM B--- S---	No F-up		
455	SEONG S- J---	No F-up			732	KANG T--- S---	No F-up		
459	CHOI Y----- J-	80	9 11	1	741	SOO H----- J-	81	11 15	8
485	LEE S--- N--	80	12 30	1	743	SEO O--- S---	88	11 15	1
489	HEO P--- S---	81	5 23	8	746	IM I- R-----	81	7 27	6
518	CHOI C----- B--	80	9 11	2	748	LEE S--- H-	82	1 15	8
533	KIM O- R--	88	12 23	9	749	SEOK K--- S---	80	11 24	1
550	JIN O--- B---	85	10 17	1	753	CHANG J--- S---	84	2 25	8

**Table 6.** Linear Plotting of Codes for the Disease Status of the Same Patients as Those were Conditioned in Table 3.

REG.	1	6	12	18	24	30	36	42	48	54	60	7Y	9Y
	---	V---	V--V	---V---	V	---	V--V	---	V--V	---	V---	V	---
416 :	000000000000	000000000000	001										
454 :	000000000000	000000000000	000011111111	111111111111	111111111111	111111111111	11111111						
485 :	011111111111												
489 :	022222222222	28											
518 :	00002												
533 :	000000000000	000000000000	000011111111	111111111111	111111111111	111111111111	11111111						
550 :	000000000000	000000000000	000011111111	111111111111	111111111111	111111111111	1						
560 :	002												
561 :	000000000000	000000000000	011111111111	111111111111	111111111111	111111111111	11111111						
564 :	000000000000	000000000000	00000000111	111111111111	111111111111	111111111111	1111						
567 :	000000005555	555555558											
572 :	000000000000	000011111111	111111111111	111111111111	111111111111	111111111111	1118						
573 :	000000000000	006000000000	00000000111	111111111111	111111111111	111111111111	1111111111						
579 :	000000000000	000000000000	000001111111	111111111111	111111111111	111111111111	1111111111						
602 :	000000000000	000000000000	0000000001										

**Table 7.** Code for Disease State at Follow-up

Code	State at follow-up
1	No evidence of tumor
2	Residual tumor
3	Recurrence, local
4	Recurrence, regional
5	Recurrence, locoregional
6	Distant metastasis
7	Combination of above
8	Dead
9	F/U at other department or hospital

## DISCUSSION

Although the system has runned very well and its recent function is very convenient to use for our patients, easily manageable, and quite reliable in keeping security of data we registered, we now think that the system need modification in some function through our 3 years of experience of the system operation.

To put patients' name in the system, we can use English only. It has been sometimes not easy to search record by name because of differences in

**Table 8.** Linear Plotting of Codes for Complication Status of the Same Patients as Those were Conditioned in Table 3.

REG.	1	6	12	18	24	30	36	42	48	54	60	7Y	9Y
	V	V	V	V	V	V	V	V	V	V	V	V	V
416 :	000000000000			000000000000		000							
454 :	000000000000			000000000000		000011111111		111111111111		000000111111		00000011	
485 :	000000000000												
489 :	000000000000		00										
518 :	00000												
533 :	000000000000			000000000000		000011111111		111111111111		100000011111		10000000	
550 :	000000000000			000000000000		000011111111		111111111111		000000000000		0	
560 :	000												
561 :	000000000000			000000000000		033333333333		333333111111		000000000000		00000111	
564 :	000000000000			000000000000		000000000000		000000000000		000000000000		0000	
567 :	000000000000			0000000000									
572 :	000000000000			000000000000		000000000000		000000000000		000000000000		0000	
573 :	000000000000			000000000000		000000000000		000000000000		000000000000		0000000000	
579 :	000000000000			000000000000		000001111111		111111111111		111000000000		0000011111	
602 :	000000000000			000000000000		000000000001							

**Table 9.** Code for Complication State

Code	State of Complication
1	None
2	Early
3	Late
4	Early and late

usage of English letters for name. This problem, however, might be solved by establishment of standard way of writing Korean names in English. To overcome the problem, we made a separate database system in dbase III plus using 3 fields, name in Korean letter, radiotherapy number, and birthday using a personal computer (Pro-3000). Using this PC database, we can easily and correctly find radiotherapy number of patients by full name or part of name.

Data of some patients who have had more than 5 courses of radiotherapy or whose follow-up period exceed 10 years, small the number is, elicits a problem of data storage or retrieval. Thus we feel the necessity of expansion of memory allocated to the programme or modification of output format.

Graphic presentation of survival data<sup>5,6)</sup> or comparison between survival by parameters, though we can do them with other personal computer in our

**Table 10.** Tabular Format of Overall Survival Rate Calculated by the Life Table Method of the Same Patients as Those were Conditioned in Table 3.

SNUH Dept. of Therapeutic Radiology  
**OVERALL SURVIVAL RATE by LIFE TABLE METHOD**

MONTH	NUMBER	DEAD	DROP	PROB
0 - 3 :	43	0	2	1.000
3 - 6 :	41	0	3	1.000
6 - 9 :	38	0	1	1.000
9 - 12 :	37	1	2	.972
12 - 15 :	34	2	0	.915
15 - 18 :	32	1	0	.886
18 - 21 :	31	0	1	.886
21 - 24 :	30	2	1	.826
24 - 27 :	27	0	0	.826
27 - 30 :	27	1	1	.795
30 - 33 :	25	0	1	.795
33 - 36 :	24	0	1	.795

department, is not possible with this system. Thus we are trying to make above functions possible with this system.

To develop a custom-tailored database system

and its utility programme is not easy, but its successful accomplishment and systematic operation promise a department consistent frame with which analysis of patients by whatever a department want to assess will results in enormous gains in productivity as we experienced.

## REFERENCES

1. Sanders DH: Computers Today. 2nd ed, McGraw-Hill book Co, New York, 1987
2. Artronix: PC-12 Data Management System. User's Manual. Artronix, St. Louis, 1975
3. Capintec Systems: Tumor Registry Manual. Capintec Systems Inc., Ramsey 1983
4. American Joint Committee on Cancer: Manual for Staging of Cancer. 2nd ed, JB Lippincott Co, Philadelphia, 11-21, 1983
5. Sladen JG: Life table using lotus 1-2-3 software on a personal comuter. Am J Surg 151:339-342, 1986
6. Ito S, Irie G: A patient information system for radiotherapy. Strahlentherapie 155:369-375, 1979

＝ 국문초록 ＝

## 방사선 치료 환자 자료처리 및 통계의 전산화에 관한 연구

서울대학교 의과대학 치료방사선과학교실

하성환 · 김일한 · 강위생 · 박찬일

우리나라 실정에 적합한 방사선 치료 환자들의 자료처리와 그 응용을 위한 전산화 시스템이 본과에서 1987년 개발된 이래 현재까지 매우 유용하게 운영되고 있다. 방사선 치료계획용 Eclipse 전산기를 주처리기기로 한 본 시스템은 Fortran-5로 프로그램되어 있고 AOS로 프로그램이 처리되고 있다. 방사선 치료를 받은 환자 30,000명에 관한 자료를 수록할 수 있는 용량으로 개발되었다. 본 시스템은 동일 환자가 재발 또는 전이 등의 이유로 재차 치료를 받는 경우가 있을지라도 5회까지는 각각 독립적인 자료로 입출력 가능하지만 그 자료들은 서로 연계적으로 수록되는 구조로 되어 있다. 동일한 방식으로 방사선 치료후 60회까지의 정기검진 자료가 수록 및 입출력 가능하다. 본 시스템의 응용 소프트웨어는 사용이 매우 간편하며 이를 이용하여 특정 환자에 대하여 수록된 자료 전체를 검색하거나 특정 조건에 맞는 환자들의 자료중 필요한 부분만 선택하여 검색할 수 있고, 특정 기간중에 치료받은 환자들을 대상으로 하여 분석하고자 하는 인자들에 관한 통계적 분석이 가능하며, 수록된 자료중 특정 조건에 맞는 환자들의 생명표적 생존율 또는 합병율 산출이 가능하다. 상기의 생명표적 분석은 미수록된 자료인 경우에도 자판을 통한 입력으로도 가능하다. 본 시스템의 구조, 운용 및 자료 처리에 의한 응용예 등을 보고한다.