Chapter 4 COMPUTER APPLICATIONS IN MEDICAL EDUCATION AND ASSESSMENT

COMPUTER APPLICATIONS IN MEDICAL EDUCATION AND ASSESSMENT

INTRODUCTION

Computers are being used throughout the continuum of medical education and assessment. The unique capabilities of computers, particularly their ability to rapidly store, process, and retrieve large volumes of data, provide assistance in and help manage education, assessment, and patient care activities. They allow a much wider range of individualized responses in tutorial sessions, a greater degree of "instructor" responsiveness to individual backgrounds and needs, and more immediate feedback in the teaching or testing setting. Also, aspects of the patient-physician encounter can be simulated for training purposes and for providing a more objective testing mechanism for assessing clinical problem-solving abilities.

The diagnosis and choice of treatments now depend largely on the subjective recall of the individual physician. Some of the deductive reasoning processes that take place in diagnosing and treating patients can be done more consistently, comprehensively, and instantaneously by using computers as "consultants." Computers are comparable to other medical aids upon which the physician has come to depend; they serve as replacements for less dependable technologies, Physicians no longer reIy on fingertip percussion of the chest wall to gauge heart size; instead, X-rays, electrocardiograms, echocardiograms, and other more sophisticated technologies are used.

Finally, computers have a unique role in the management of large volumes of data for reimbursement of health services, utilization review, and the planning, monitoring, and evaluation of medical care services.

Four different types of computer applications in medical education and assessment are discussed in this chapter. These four types are: 1) computer-supported independent study in the basic medical sciences, 2) the use of computerized simulations in training and testing for clinical competence, 3) computer-based medical consultation systems, and 4) computer uses in quality assurance/data management systems. These applications were chosen to illustrate the uses of computers across the education and assessment continuum. Each case study also was chosen to emphasize a particular capability of the computer. Each application may have multiple capabilities both in instruction and in assessment, and the concepts underlying each application are potentially applicable to other portions of the continuum.

COMPUTER-SUPPORTED INDEPENDENT STUDY PROGRAM IN THE BASIC MEDICAL SCIENCES

The Ohio State University (OSU) College of Medicine has two programs for completing the basic science portions of the Doctor of Medicine curriculum (Beran et al., 1976; Merola et al., 1978; Cramblett et al., 1979). The lecture-discussion program follows a traditional format in which students progress through a standard content sequence in a fixed amount of time—is months. In contrast, the independent study program (ISP) follows a self-study format in which students progress through the curriculum at independent rates. The principal difference between ISP and the traditional system is the former's acceptance of time as a variable in the learning process. The student progresses when he/she achieves a prescribed level of competency, not when the next lecture series begins. Individual progression rates require that students completing their basic science studies be allowed to enter their clinical clerkships in almost any month of the year.

ISP features an integrated curriculum, modular student study objectives, computerbased tutorial evaluation systems, and a means of managing a student population progressing at independent rates. In July 1970, the first class of 32 students was accepted into ISP, followed by classes of approximately 60 students each in 1971 and 1972. The latter two classes comprised a little over one-fourth of the entering medical students of each year. Currently, 40 percent of the entering classes select ISP.

The ISP curriculum is designed around two main segments: normal man and pathophysiology. Table 7 outlines the components (units) of the ISP curriculum. Each student is required not only to master the units, but also to participate in laboratory exercises, attend clinical correlation sessions and psychosocial seminars, and demonstrate satisfactory verbal communicative skills through oral exercises with basic science and clinical faculty. The student flow through an independent study learning unit is shown in figure 7.

Each student's progress is monitored by the computer through the computer-based self-evaluation or Tutorial Evaluation System (TES), a series of interactive computer-assisted instruction (CAI) "question and answer" sessions between the student and the computer terminal. When the student answers a question correctly, he/she is immediate-ly reinforced and then advanced to the next question, If a student answers a question in-correctly, the student must complete corrective feedback and/or remedial work before reanswering the question. If a student shows a total lack of understanding of a concept, he/she may be branched to a remedial review unit. If an answer is given that the course author did not anticipate, the student may be taken to a coaching sequence designed to guide him/her to the correct answer. If serious weakness in an area is demonstrated, a study prescription may be given, such as "you lack understanding in area X and should review pages Y through Z in text W" (See figure 4 in chapter 2 for a schematic summary of this process). The faculty controls generation of these study prescriptions. Figure 8 gives a sample student-computer dialog.

ISP presents logistical problems in terms of managing the flow of students through the program. With students presenting themselves individually and at different times for tutorial evaluation, lab work, and subject matter consultations, the faculty member must be assisted in monitoring each student's progress and performance. To meet this need for monitoring students, student reports are provided at various intervals. The reports comprise the computer-based management system for ISP. Sample reports are depicted in figure 9. Ch. 4-Computer Applications in Medical Education and Assessment •49

Table 7.—	The Independe	ent Study	Program	Curriculum	Outline

_

	_	
•		Associated tutorial evaluation system
Unit	Title	module(s)
	Normal man	
1 2	Cell structure and function	A B
	Exam 1	
3	The respiratory system	C
4 I 5	Kidney function, urinary system, body fluids	D E
	Evon 2	
e	EXdIII Z	G
0 7	Muscle the spinal cord and peripheral pervous system	H
8	The central nervous system	I
	, Fuer 2	
۰.	EXAM 3 The endeering metabolic system	.1
10	Nutrition	F
11	Reproduction	L
12	Hematology	0
. –		
Comp	Exam 4 prehensive exams-Sections 1.2, and 3 (anatomy, biochemistry and physiology)	of the NBME exam.
	Pathophysiology	
13	The immunologic system	Р
14	Pathologic mechanisms and genetics	Q/K
15	Mlcrobiology. ,	R
16	E pidemio logy	Т
	Exam 5	
17		s
18	The cardiovascular system	Х
	E u u O	
10		V
19	The respiratory system	Ŷ
20	The Tespiratory System	
	Exam7	
21	Hematology.	u
22	The gastrointestinal system	w
23	The renal system	Z
	Exam8	
24	The central and peripheral nervous system	1
25	The musculoskeletal system	2
26	Behaviorial science	L/3
	Exam9	
27	Introduction to pediatrics	M / 4
28	The endocrine system .	5
29	The reproductive system	

Final Comprehensive NBME Part I(all sections) SOURCE Ohio State University 1979

The student progress report provides information regarding: 1) units and subunits completed to date, 2) cumulative time spent on TES, and 3) a summary of the week's activity, including a listing of study prescriptions. The TES time is only $1 \ 1/2$ hours out of a 40-hour work week. Most student time is spent in independent study of the learning resources. The computer, however, can easily provide the reporting described herein since the student must complete a TES exercise within each subunit.

Figure 7.—Student Flow Through a Typical Module



- 1. Student reviews educational objectives and examines list of available instructional resources.
- II. Student pursues instructional resources in accord with his individual needs.
- Ill. Student utilizes self-evaluation system where he receives tutorial feedback as to his mastery of the educational objectives.
- IV. Student confers with faculty member for final review of instruction unit (optional).
- (Typically, from one to five submodules are included in each module, with the student proceeding through steps I-IV for each submodule.)
- V. At specified points in the curriculum he must take a written exam covering several modules, and
- VI. Demonstrate his ability to transmit information orally in prescheduled oral exams.

SOURCE: Merola et al , 1978

The student response printout can provide every response entered at the terminal sorted by student and/or by question. The student cluster report is used primarily by the curriculum directors and faculty to determine the location of the ISP students throughout the curriculum in order to schedule labs or clinical sessions. The student status report is

Ch. 4—Computer Applications in Medical Education and Assessment •51

Figure 8.—Part of a Student-Computer Dialog

Explanation	Student-Computer Dialogue
Computer poses a question.	<u>Computer</u> : Name the pathological conditions within the brain caused by blockage of the lateral and medial foramina of the inferior medullary vellum.
Student responds (unanticipated response) .	<u>Student</u> : I don't remember.
Computer poses review question.	<u>Computer</u> : Are you having difficulty placing the location of the inferior medullary velure and/or the significance of the foramina?
Student responds.	<u>Student</u> : Yes.
Computer offers coaching in- formation and poses another question.	<u>Computer</u> : The inferior medullary velure forms part of the roof of the fourth ventricle. The foramina communicate with the cisterna <i>magna</i> and/or cisterna pontis. What is contained in these cisternae?
Student responds (correctly)	Student: Cerebrospinal fluid (CSF).
Computer gives reinforcement to correct answer.	<u>Computer:</u> Correct. Now consider blockage of the foramina and answer the <i>original</i> question.
Computer re-poses the original question.	<u>Computer</u> : Name the pathological conditions within the brain caused by blockage of the lateral and medial foramina of the inferior medullary velum.
Student answers (wrong answer)	Student: Water on the brain.
Computer offers tutorial feed- back pointing to the correct answer.	<u>Computer</u> : Not acceptable. It's time to become more professional in your terminology!
Student responds (correct answer).	<u>Student</u> : Hydrocephaly.
Computer gives reinforcement to the correct answer.	<u>Computer</u> : Very good! The obstruction will cause an increased quantity of cerebrospinal fluid in the ventricular system.

SOURCE The Ohio State University College of Medicine,1978.

an abbreviated version of the student progress report and is available on demand at the CAI terminal.

Another kind of report provides item statistics information that the faculty can use to analyze, correct, and revise TES items. This report also may be used to help determine how well students responded as a group to TES questions. Most importantly, this report lists all student responses that were unanticipatedly the TES author. Responses that will improve TES sessions, as determined by faculty members, are added immediately to the TES program. Figure 10 depicts an item statistics report.

Faculty members in a specific discipline are responsible for the preparation of individual aspects of each unit and may revise and design changes in relevant modules of TES. The almost infinite variety of question types that involve corrective feedback, 52 • Computer Technology in Medical Education and Assessment

Figure 9.—Sample Computer-Based Management System Reports

ISP PROGRESS REPORT ON COURSE PILOT FOR JOHN DOE PRINTED ON MAY 5, 1976 AT 17:44 STUDENT NBR/0316 CLASS/75 PERSONAL CHART/75 FACULTY ADVISOR/AJM TYPE/LINEAR LAST USAGE /MAY 3 PROGRESS CHART USED FOR THIS REPORT/75 7/09/75 SCHED - A A B C D D F F G G H H COMP- A AB BC CD DE EF FFG G GE H H H 10/08/75 SCHED- I J K L M N O O P P COMP-HI IJK K K KL LM MNO OP P P ΟΟΡΡΡΡΡ Ρ Ρ Ρ 1/07/76 SCHED- P P R R ST T T Q Qυυ COMP-- P PQ QR RSSSSSSSJ Q *CURRENT WEEK 4/07/76 SCHED- U V V V W W W X X Y Y Z Z COMP-- u w v v-w w w w * CURRENT W NOTE: CHART SHOWS ONLY MODULES COMPLETED - STUDENT CURRENTLY AT XA

A Student Progress Report. This report gives a linear view of an individual student's progress as compared to a projected rate of progress through the entire curriculum. It also gives the faculty. adviser and the last date of usage.

STUDENT STATUS - PILOT DATE 04/07/76 TIME 4: 57 SELECTION = ALL

STUDENT NAME STUDENT NO. GROUP AREA TYREC START DATE LAST DATE TIME SEG LABEL-PROB-SEQ 030

CHARLES JONES	339	75	0	15	7/17/75	3/04 /76	22: 00	QW UAOOOT-	3-	0
PHYLLIS SMITH	340	75	0	15	7/14/75	3/20/76	32: 14	QW UAOOOT-	3-	0
DALE WITHE	341	75	0	15	7/18/75	1/ 14/76	27: 52	QW QAOOOT-	1-	0
Scott GREEN	342	75	0	15	7/14/75	4/03/76	42: 22	QW WG999Z-	1-	0
JOANN BROWN	344	75	0	15	7/22/75	3 /04 /76	15: 57	AP LAOOOT-	1-	0
ROSALIND DOE	171	73	0	15	7/19/73	8 /20/75	35: 16	QW WF100S-	1-	0
TONI BLACK	345	75	0	15	7/14/75	3/01 /76	57: 55	QW UAOOOT-	3-	0
JAMES DILL	351	75	0	15	7/14/75	4/02/76	36: 50	QW WAOOOT-	1-	0
MARY HILL	352	75	0	15	7/11/75	4/01 /76	32: 03	QW klA060S-	1-	0
DOUGLAS HALL	231	74	0	15	7/12/74	1/ 16/76	02: 47	36 5BUIOS-	1-	0
JOHN DORN	353	75	0	15	7/21/75	4/03/76	19: 03	QW QA020S-	1-	0

etc.

Aportion of the Student Status Report. (Student names are fictitious.) This report gives the location of each student in the curriculum (label), when each student started in the program, and the last date that student had any activity at the terminal.

SOURCE Pengov,1974

coaching statements, and immediate reinforcement permits individualistic approaches to the same material. The possibility of making constant changes in TES design affords each faculty member the opportunity to instantly correct errors, update content, or change the emphasis in the original system's content.

The authorship of TES changes as new faculty become involved in the program. The design of individual TES items is modified and revised by new faculty concomitantly

Ch. 4-Computer Applications in Medical Education and Assessment • 53



Figure 10.— Item Statistics Report

SOURCE Pengov 1974

with changes in objectives and source materials. Each new faculty member is encouraged during the ^{-TES} revision process to collaborate with instructional programers indesigning the module. Such collaboration can take the form of correcting obvious problem areas in TES or discussing and implementing new approaches to TES, such as greater use of problem-solving techniques.

Faculty responsibilities are somewhat different in ISP than in a traditional lecture curriculum. The emphasis in ISP is placed on integrating material into learning units prepared by an interdisciplinary team of faculty, not on designing and presenting lectures. Faculty members assess the ability of students to grasp the materials presented in the unit he/she designed through small group discussion sessions with students working on the

51-541 0 - 79 - 5

54 • Computer Technology in Medical Education and Assessment

module. Occasional clinical correlation sessions, conducted with both basic science and clinical faculty, facilitate the extension of concepts of normal function to those of pathophysiology. The emphasis placed on developing the curriculum within an interdisciplinary team and small group discussions, rather than lecture preparation, has created a curriculum that is changing constantly as faculty receive helpful suggestions from both students and other faculty.

The ISP curriculum attracts a wide range of students, and they perform comparably to lecture-discussion students on both clinical and preclinical portions of standard exams. ISP appeals to some students who wish to accelerate the completion of their medical education and to others who wish, or need, to reduce the academic pressure of medical studies. The flexibility inherent in ISP and the student's ability to control a rather intense curriculum benefit students with academic deficiencies, as well as those with strong academic credentials. Students who have had extended absences from a formal education setting, previous to their admission into the college, also benefit from ISP.

The groups that progress at the slowest rates through the program include a significant number of minority students who need special assistance to complete their medical studies. However, many students with no apparent academic weaknesses are taking more time to complete the curriculum than their lecture-discussion counterparts. Despite a delayed graduation, most of these students prefer more time than the standard 15 months the lecture-discussion program allows. The range of completion time in ISP is 10 to 29 months, the mean is 17 months.

The experience of ISP has illustrated that:

- Independent preclerkship study is a useful and viable alternative to traditional programs.
- Computer-assisted and computer-managed independent study is feasible for large numbers of preclinical medical students.
- Medical education can be tailored to the individual needs of each learner.
- Computer-assisted instructional and management techniques allow each student to progress independently of others according to his/her own educational background, learning speed, or needs. CAI/CMI allows faculty and administrators to manage large numbers of students progressing at independent rates. Such monitoring techniques are an essential element for the implementation of any largescale independent study program.
- Many faculty have become orchestraters of learning rather than disseminators of knowledge. The computer does the repetitive tutorial work, thereby providing faculty members more time for individual student interactions. Overall faculty time requirements do not change; the nature of their roles and the use of their time, however, do change.
- Student response to ISP has been very positive. They find it enjoyable to learn by discovering with a faculty member as well as from a faculty member.
- Despite the high investment costs for courseware development, overall CAI operational costs compare favorably with the costs for other instructional methods, such as labs or group discussions.

THE USE OF COMPUTER-BASED SIMULATIONS IN EDUCATION AND ASSESSMENT

Introduction

Several different models of computer simulations used for educational assessment are described below. The use of computer simulations by the Royal College of Physicians and Surgeons of Canada, the American Board of Internal Medicine (ABIM), and the National Board of Medical Examiners (NBME) in actual assessment activities is then discussed.

A wide variety of computerized, clinical encounter simulations have evolved in the last decade. Computers are used to simulate "patients;" physicians or students interact with typewriter-like computer terminals to solve the "patient's" problems. Although the setting does not simulate all aspects of a physician's interaction with a live patient, the essential feature of information flow between the two is captured well, particularly if efforts are made to make the simulated interaction more "human" and less machine-like (Senior, 1976). A more "human" system is characterized by language that is easily understood, by responses that can be corrected quickly, and by courteous and considerate instructions (Sterling, 1975). Only rudimentary typing skills, and no knowledge of computers or programing, are required of the user. Inquiries can be directed to the patient by typing in actual questions, either as words or sentences, or as code numbers referring to questions or desired studies. Successful experiments also have used a standard touch-type telephone to communicate with the computer patient (Friedman et al., 1977). Diagnoses can be entered, management decisions or actions can be initiated, and data may be reviewed. If needed, special assistance can be provided by attendants at the computer terminal site. However, brief introductory training exercises are usually conducted to familiarize the users with the terminal equipment and mode of operation.

Prior to the use of computers, simulations were primarily *paper-and-pencil* in format. The best known paper-and-pencil simulation is the patient management problem (PMP) (Hubbard et al., 1965). A traditional PMP simulates various aspects of a realistic clinical situation in which the physician or student must manage a patient. After reviewing a paragraph of information describing the patient, the user may request additional data. On the basis of the information collected, the user diagnoses, treats, or manages the patient's problem by choosing from a limited set of options or interventions that he/she thinks are necessary to reach a proper solution to the patient's complaints. Appendix I contains an example of a segment of a conventional paper-and-pencil PMP. Most often, the branching sequence offers a choice between the physical examination section, the laboratory section, and the history section of the patient management problem. Each section presents a list of items, including essential, optional, and irrelevant information. The task of the user is to choose those items that are pertinent to the resolution of the problem. Early PMPs were linear, somewhat inflexible, and offered cues to the user. "The validation of the PMP has been called into question in that attempts to show that practicing physicians behave similarly in a PMP as they do in the real-life situation have not been generally successful. This is thought to derive from the fact that in this simulated situation the physician is given a series of options and clinical investigations from which selections may be made thereby affording cues to the examinee. In the clinical situation the options, confirmatory tests, and consultations possible are legion and require an internally generated selection" (Meyer, 1978). Computer simulations can help avoid the problem of "cueing." The computerized simulations described below overcome, in varying degrees of sophistication, the limitations of paper and pencil simulations.

56 • Computer Technology in Medical Education and Assessment

Sophisticated *computerized versions* of PMPs contain multiple branches and require the physician to make choices that lead to any number of strategic routes, several of which could result in an acceptable solution to the patient's problem. These computerized versions allow much more complex branching; provide immediate feedback to the physician; permit the review of previously selected items; utilize audio/visual capabilities; and provide mechanisms for timing, control, or monitoring of the examination. Additionally, user performance recordings are more detailed, retracing by the user is minimized, and scoring of the problem is instantaneous. For each problem the user receives both introductory information about the patient, the setting of the problem, and an explanation of his/her role and tasks.

Sim One is a lifesize manikin that has several functions, including respiratory activity, pulse rate, blood pressure, skin color, and pupillary size, that are controlled by the computer. Each function responds to drug administration and other interventions used by an anesthesiologist to manage patients in surgery. This model simulates a variety of operating room situations. The clinician can repeatedly experience these situations in credentialing or educational settings without risk to real patients and with immediate feedback regarding the effects of judgments and actions (Abrahamson et al., 1969).

INDEX has a glossary or index list of over 1,000 history questions, physical findings, laboratory procedures, and diagnoses that physicians use to obtain information from the patient. A sample report summarizing an interaction between a physician and a patient is illustrated in figure 11. The physician, identified as "No. 136," has examined a patient who has blood loss anemia that is caused by fibroid tumors of the uterus and von Willebrand's disease. All inquiries or diagnoses made are listed in order as four-digit numbers. The final diagnosis and management plan, the time the examinee took to complete the problem, the dollar costs of studies ordered, the time consumed, and an estimate of the aggregate discomfort and risk to the patient from the workup are listed in the figure. Such a record is available for analysis, evaluation, and scoring in printed and computer-readable form.

MATRIX is a modification of a teaching program designed to illustrate that a diagnosis is based on statistical probabilities of the condition in question. These probabilities are based on factors such as the patient's age, sex, symptom complex, physical findings, and results of laboratory tests. The teaching program was adapted to a testing program on the diagnosis of acute abdominal pain. As in INDEX, the candidate is given an index of inquiry possibilities and then enters the inquiry by code (Senior, 1976).

Mac Puf is a computerized cardiovascular-pulmonary model in which the computer program calculates the effects of changing one variable on all the other physiological variables (Dickinson, 1977). Values of the variables are interrelated by equations, and the computer solves all the equations simultaneously. The program can be made to simulate diseases such as acute myocardial infarction, arterial or venous bleeding at any rate, pulmonary emboli, and obstructive lung disease. It can be designed to respond to treatments such as digoxin, intravenous fluids, oxygen, respirators, rotating tourniquets, vasopressor drugs, and bicarbonate. If untreated, the patient may deteriorate and "die" or stabilize at some lower level of function. Appropriate and timely therapy may reverse the downward trend of simulated illness, and vital signs will improve. Complications and new medical crises can be programed to occur at any time. Thus, this model of a changing and responsive patient represents a potentially useful tool for assessing management skills or the ability of a clinician both to treat crisis problems and to avoid complications.

Ch. 4—Computer Applications in Medical Education and Assessment •57

#136 CASE: 4 TIME: 14: 50 DATE: 10/31/72 HISTORY PHYSICAL STUDIES DIAGNOSES PROBLEM BEGAN 1010 MAIN PROBLEM 1000 FAMILY BLEEDING 2716 HOSPITALIZATIONS 2340 TMM ALCOHOL 2330 PROLONGED BLEEDING 1963 ANTICOAGULANTS 1564 HEPATITIS 1760 ALLERGIES 2550 PALLOR 3024 **MENSTRUATION 1920 MENORRHAGIA 1926** SPLEEN PALPABLE 3748 FECAL BLOOD 3834 LIVER ENLARGED 3747 SMEAR AND DIFFERENTIAL: 5008 **ROUTINE STUDIES: 5000** PLATELET COUNT: 5058 **PROTHROMBIN TIME: 5066** CAPILLARY FRAGILITY: 3506 BLEEDING TIME: 5060 PETECHIAE 3054 APLASTIC ANEMIA IRON DEFICIENCY ANEMIA THROMBOCYTOPEMIC PURPURA VON WILLEBRAND'S DISEASE (IMPRESSION LIST) SPECIAL COAGULATION TESTS: 5086 (FINAL LIST) **IRON DEFICIENCY ANEMIA** VON WILLEBRAND'S DISEASE PATIENT MANAGEMENT PLAN FERROUS GLUCONATE 300 MG PO TID ADVISE PATIENT TO REMAIN UNDER SUPERVISION FOR BASIC DISEASE DOLLAR COST OF STUDIES: \$25 TOTAL TIME TO COMPLETE ALL LAB STUDIES: 1 DAYS 12 HRS O MIN PATIENT TIME CONSUMED: O DAYS 1 1-HRS 25 MIN PATIENT RISK: 7 PATIENT DISCOMFORT: 35 ELAPSED TIME FOR PROBLEM SOLUTION: 28 MINUTES.

ELAPSED TIME FOR PROBLEM SOLUTION: 28 IN SOURCE Senior,1976 *CASE*, a computer-assisted simulation of the clinical patient encounter, was developed in the late-1960's at the University of Illinois (Harless et al., 1971; 1973). CASE uses an unprompted and undirected process. The interaction occurs at a computer terminal and the physician elicits whatever history, physical examination, and laboratory data that are needed for the diagnosis and management of the patient's problem by typing unstructured, English language phrases or sentences into the computer. The computer provides brief introductory materials to the physician, including a description of the clinical setting, the circumstances around the patient's visit, and some physical characteristics of the patient. After the introduction, the physician receives no further clues. Candidates must determine what information they think is important in order to define and manage the patient's problem. Appendix I contains a typical undirected interaction between a physician and a CASE patient.

A wide and representative variety of internists' patients can be simulated by this computer model, and each simulated, computerized patient (CASE) can be refined to respond appropriately to approximately 95 percent of the physician's inquiries. Over 50 CASES have been created to date; most are housed at the OSU College of Medicine.

The costs of creating, refining, and field testing each CASE are estimated to range from \$5,000 to \$15,000, depending on the degree and type of review provided. The relatively low development cost for a simulation as complex as CASE is made possible by the existence of a series of computer programs that guide the creation or generation of new CASES. This generating system (GENESYS) allows the creation of a large number of CASES in a relatively short time.

GENESYS essentially involves a three-phase process: 1) author preparation, 2) technical creation, and 3) refinement. During the first phase, the author compiles data describing his/her patients for subsequent use in a CASE simulation. The technical creation phase entails entering the author-supplied information into the computer according to a standardized format. The final phase, refinement, is designed not only to ensure that the CASE simulation can respond to the widest possible variety of anticipated questions, but also to "humanize" the computerized patient.

Computerized PMPs, INDEX, and CASE have been used and are being tested for actual physician assessment activities by the Royal College of Physicians and Surgeons in Canada, by NBME, and by ABIM. The use of each simulation is described in more detail below.

Pediatric Fellowship Examination of the Royal College of Physicians and Surgeons of Canada.

The Royal College of Physicians and Surgeons of Canada uses Computerized Patient Management Problems (CPMP) regularly in their Pediatric Fellowship Examination process to test the ability of pediatric candidates to manage patients (see appendix I for an example of a CPMP). The College also uses conventional multiple-choice questions and oral examinations in the Fellowship exam. The specific objectives of the CPMP section are to assess the candidate's skill and ability in data gathering, formulating hypotheses regarding the patient's problem, and resolving the problem. Candidates have a choice of taking an examination in either English or French.

The computer portion of the examination has been administered annually since 1974 to all candidates eligible for certification in pediatrics. The development of the CPMP examination was precipitated by several factors (Skakun, 1978): 1) concerns about the inability of multiple choice and essay examinations to measure various clinical aspects of

medicine; 2) concerns about using a single examination, such as mutiple choice or essay, to certify specialists; and 3) concerns about the degree of reliability, validity, and standardization of oral examinations.

The types of CPMPs used in the Royal College exam are the linear and the sequential. The linear uses two response strategies. In the simple choice, the candidate selects six items or less from a list of possibilities; in the ordered choice, the candidate chooses a fixed number of options and then rank orders them. In the sequential type of CPMP, the candidate scales each option on a scale of 1 to 20; the most pertinent information is given a weight of 20 and the least pertinent a weight of 1. In this type of problem each candidate is presented with the same amount of information and is led, in an identical step-by-step manner, to the next options in the problem.

Officials of the Royal College note that the examination could have been administered using conventional paper-and-pencil PMPs. Use of the computer, however, has advantages in that it prevents the retracing of physician questioning and allows for more complex branching problems, incorporation of visual and audio components (i. e., Xrays, heart sounds, etc.), and immediate scoring of the problem (Skakun, 1978).

Most of the CPMPs used are developed and tested on the IBM 1500 computing system at the Division of Education Research Services in the Faculty of Education at the University of Alberta. Each member of a five-person committee is asked to develop two problems in areas relevant to pediatric practice. Each problem is then reviewed by the entire committee. Once the committee reaches agreement on the content, the problems are programed for presentation by a computer. A group of pediatricians not involved in the preparation or initial review of the problems takes an examination in which the problems are presented. On completion of the examination, their comments and suggestions are used to further refine CPMPs. These updated problems are then used in the Royal College examination. Problems are reviewed annually for accuracy and relevance by practitioners not involved in their initial construction.

CPMPs also have been used since 1974 as part of the examination process for approximately 110 final year medical undergraduates at the University of Alberta. Additionally, CPMPs have been incorporated into a self-assessment package that is available to fellows of the Royal College during the time of the College's annual meeting.

The Royal College exam is administered at eight different centers across Canada, using either available computing facilities in each center, or a network such as the CHARGEX system of the Canadian Imperial Bank of Commerce.

Each year, approximately 150 candidates take the Royal College's CPMP examination in addition to taking a multiple choice test, an oral examination, and an in-training survey. The cost per candidate in 1974 for administering four CPMP problems was approximately \$225. These costs have been decreasing gradually since 1974; in 1977, the cost per candidate was approximately \$150. For the self-assessment tests held at the time of the College's annual meeting, approximatel, 200 fellows participate at a cost of significantly less than \$100 per fellow.

Evaluation studies of the Royal College Fellowship exam indicate (Fincham et al., 1976; Berner et al., 1977; Skakun et al., 1977):

• No positive correlations exist between CPMP, multiple choice, and oral examination portions of the Fellowship Examination. This finding suggests that different aspects of competence are being measured by the various types of tests.

- . There are low interproblem correlations between CPMPs. This suggests that candidate performance might be case- or problem-specific, or that the skills required to manage the patient, such as problem solving and decisionmaking, differ from case to case.
- Candidates had no difficulty in using the computer terminal and viewed the problems as realistic and relevant to pediatrics. They suggested that CPMPs were better tests of pediatric skills than the multiple choice questions.
- Nationwide computer-based testing is feasible and computerized patient management problems can be translated to multiple computer languages and hardware systems and still operate successfully to support such testing.

The Computer-Based Examination (CBX) Project of the American Board of Internal Medicine and the National Board of Medical Examiners

The CBX project, initiated in 1968, is based on a computer simulation of a patient with a given disease or diseases. Although various computer simulations have been explored, all involve physician interactions with a computerized patient that is or is not hospitalized. In the current CBX simulation model (which evolved from INDEX), laboratory tests, medical procedures or consultations, and drug therapies interact with the patient's disease in a time sequence similar to one occurring in real life. By simulating the effect of physician action on the patient, and by adjusting the patient's status accordingly, the computer model provides an almost life-like, dynamic simulation of the patient/physician encounter.

The computer model, through a sampling of the physician's interaction with a patient, permits the assessment of a number of factors, including cost of workup, risk (pain, complications, and mortality) to the patient, time taken to initiate corrective therapy, and length of a patient's hospital stay. The model also permits a step-by-step evaluation of the physician's action in ordering tests and prescribing therapy. Appendix I contains a sample interaction between a physician and a simulated patient.

The major reservation to implementing this system centers around the ability of the project: 1) to develop a scoring strategy and to determine whether the competencies assessed by this model are appropriate to measure, and 2) to develop a method for the cost-effective production of new simulations. Future plans involve further testing and refinement of the model as well as development of additional patients or scripts. Studies have been initiated to evaluate the model and determine its economic feasibility for large-scale testing and use by NBME and ABIM as part of their certification and recertification processes (Friedman, 1978).

A Model for Evaluation and Recertification through Individualized Testing (MERIT) of the American Board of Internal Medicine.

MERIT was begun in 1973 to investigate: 1) an approach to recertification that uses the specific patient problems of each internist as the basis for the evaluation of his/her skills, knowledge, and clinical judgment; 2) the use of an advanced computer-assisted simulation of the clinical patient encounter (CASE) as the examination instrument for the individualized evaluation; 3) the development of a scoring system for CASE to assess various dimensions of clinical behavior; 4) the involvement of each participating internist in an evaluation process that will help him/her to identify deficiencies and thus to plan for more meaningful, individualized continuing education; and, 5) the attitudes of practicing internists toward a recertification process that embraces the preceding components (Harless et al., 1978).

The MERIT process has undergone extensive evaluation during the past 5 years with 450 Board-certified internists in 3 field studies. The first study was conducted in San Francisco with 28 physicians in *1973-74;* the second study, held in 1976, included *90* physicians located in 9 health care and training institutions throughout the State of Ohio; and the third study was conducted in January 1978 at 15 centers nationwide and involved 332 physicians.

The MERIT process was designed to incorporate individualized evaluations of a candidate's performance in patient care as a major part of the recertification procedure. It is meant to supplement the multiple choice questions and paper-and-pencil PMPs used in current examinations. MERIT incorporates a number of steps, the first of which is the gathering of information from each candidate about the types of patients he/she sees in practice. This patient information is summarized into an individual practice profile that is used to select the three most representative computerized patient simulations for that physician's exam. Table 8 depicts a typical practice profile.

The MERIT process uses the CASE simulation and generates six separate scores (a seventh, history and physical exam minimum data base score, is under consideration):

- 1. *Diagnosis Score*. Diagnoses for each CASE that a candidate is required to list have been identified. Each diagnosis is then assigned some proportion of a total 100 percent. This weight is based on the degree to which the disease can be treated and the degree to which it threatens life. An individual's diagnosis score is the total of the weighted value for the correct diagnoses listed; the possible range is *O* to 100 percent.
- 2. *Management Score*. The management steps (medications, procedures or consultations, and patient instructions) that a candidate is required to list have been identified. Each management step listed as a scoring criterion is weighted differently, totaling 100 percent across all three therapy areas. An individual's score is the sum of the weights of each correct therapeutic step he/she lists in the interaction.
- 3. Critical Concept Score. Information from the history, physical examination, and laboratory workup for each CASE that is considered critical in determining both the patient's problems and an appropriate treatment plan has been identified. Each piece of information ("critical concept") is then weighted nine, four, or one to reflect its importance in the patient's workup. An individual's critical concept score is either the total weighted value of critical concepts he/she elicited, divided by the weighted total of critical concepts in the CASE, or the percentage of the total possible critical concept value elicited.
- 4. *Efficiency Score*. This score is applied in the laboratory section only. In addition to identifying critical laboratory information, procedures or tests that are neither critical nor excessive (0-weight) have been identified. An individual's efficiency score is the total number of critical laboratory results elicited, divided by the sum of critical and inefficient procedures (or divided by total procedures minus 0-weighted procedures). An individual's efficiency score is, then, the percentage of total laboratory results elicited that are critical to the CASE.
- 5. Danger/Discomfort (D/D) Score. Laboratory procedures that entail some risk and/or discomfort are weighted (-9, -4, and 1) to reflect the degree of risk or discomfort. An individual's D/D score is a total of the negative weight received. This is not a percentage score; zero is perfect.

ú
ASE
riate (
prop
ect Ap
to Sel
rofile
tice P
e Prac
ample
8. – S
Table

Dietrih	ution of							
diagno	ses in	Referi	als	Distribution o	Circumstances o	1		
Diagnostic areas prac	tice	드	Out	complaints	patients visits	De	mographic data	
Infectious disease 1	1%	%—	%-	General symptoms 20 ⁴	% Outpatient 6	5% Fem	ale65'	%
Neoplasms 5		2	40	Vervous system 2	Inpatient 3	3% Male	35	
Diabetes 2	~	ł	1	Skin, nails, hair –	Home	2		
Other endocrine 1		ł	1	Cardiovascular and		Age (0-19	%
Nutritional and metabolic . 4	**			lymphatic 7	New complaint 3	5% Age:	20-4425	
Hematology 2	~	I	0	Respiratory 30	Followup	0 Age	45-6435	
Organic brain syndrome		I	Ι	Musculoskeletal · · · · 5	General checkup 1	5 Age	greater than 64 30	
Psychiatric disorder 2	~	ł	20	Digestive 15				
Neurology	~	1	1	Jrinary 7	Patients hospitalized	Low	socioeconomic 25	%
Hypertension 14	*	20	1	Male reproductive 1	after visit	0% Midd	lle 50	
Cardiac disease	**	75	Ι	Female reproductive. 2		High		
Cerebrovascular disease 2	2	1	ł	Eyes and ears 3				
Other vascular disease 1	-	1	Ι	Mental health 5		Blac	k 10'	%
Upper respiratory disease . 3	~	I	I	Vonsymptomatic 3		Chic	ano2	
Pulmonary disease	2	I	I		1	Oriei	ntal7	
Gastrointestinal5	•	ო	t	Total 100	%	Whit	e 78	
Genitourinary 2	~					Othe	er 3	
Gynecology 1	-	1	6					
Obstetrics		I	ļ					
Dermatology	~		6					
Rheumatology 6	m	1	5					
Congenital disease		I						
Trauma and injury	-	I	I					
Adverse chemical and								
drug effects 1	-	I	Ι					
Symptoms	e 1	[[
Well patients 6	۳ ۳	1	1					
Total H 00	0% 1	00% 1	%00					
Source: MERIT project.								

6. Cost Score. Every laboratory procedure or test has a dollar cost value associated with it. An individual's cost score is the sum of the value of every procedure and/or test ordered.

The CASE model, when used in an examination, requires that the examinee use his/her own intellectual resources to resolve the problems of the simulated patient. No direction is provided, such as in a multiple choice list, to resolve the problem. Differences between CASE and a multiple choice list are analogous, though more pronounced, to the difference between a multiple choice and an essay exam; in both CASE and an essay exam, the uncued behavior and deeper thinking of the examinee are being assessed.

Patterns of behavior can be determined for each physician being examined because each is acting without cues. Behavior patterns that are dysfunctional or potentially harmful to a patient can be corrected. For example, a physician may demonstrate a tendency to use some drugs excessively or some procedures unnecessarily. Alternatively, patterns that are creative and effective can be rewarded and shared with the physician community.

Dramatic errors can be uncovered since the examinee is not cued or restricted in what he/she does to the simulated patient. An example of this occurred during MERIT field trials. A phyician dealing with a simulated patient determined that exploratory surgery should be performed on his patient. As it turned out, the patient was in congestive heart failure, but the physician had not discovered this during his workup. A cued or multiple choice exam would never uncover this situation, since the treatment plan suggested by the physician was so radical it probably would not have been a multiple-choice option.

In the course of the MERIT project a variety of research questions have been addressed with the following results:

- It is feasible to individualize an examination based on the kinds of patients an internist sees in practice.
- It is technically and logistically feasible to reliably examine physicians simultaneously at different sites nationwide.
- Internists gather information and initiate patient management steps similarly with CASE as with real patients.
- Internists generally found the MERIT process to be an acceptable and desirable form of evaluation.
- The need for reliability suggests that any implementation of MERIT include the use of five to seven, as opposed to three, CASES.
- Comparison of MERIT scores with other component scores in the recertification examination confirmed that MERIT measures substantially different aspects of competence.
- The validity of MERIT CASE scores compares favorably with other examination methods employed in the study.

ABIM is improving the current scoring system and examining the logistics and costs of full-scale implementation of this methodology (Meskauskas and Norcini, 1978).

COMPUTER-BASED MEDICAL CONSULTATION SYSTEMS

Computer-based consultation systems range from those used in specific disease or problem areas to those used for more general purposes. A number of these systems are

briefly summarized below. More general purpose computer systems also are used to help guide the physician in diagnosing and managing patient problems. Some of these will be described in the next section on quality assurance/data management systems.

Most of the consultation systems accumulate data on patient characteristics, histories, physical findings, laboratory data, treatments, and outcomes. These data are then analyzed for correlations that might be found between certain patient characteristics and treatments with specified outcomes. This information is then used to help guide the physician in managing the patient and in discovering new knowledge regarding patient care.

Duke University Cardiovascular information System. This system provides physicians with a large data base regarding clinical experiences with coronary artery disease. The data describe outcomes of patients with various sets of attributes (Rosati et al., 1975). The patient attributes, laboratory and physical findings, history, and outcome of a large number of patients are stored in a computer, which then classifies the information. When attributes of a new patient are entered, the computer selects the most closely matched subgroup. The disease courses and outcomes of all patients previously categorized in the same subgroup are displayed. The computer's memory is accurate, unbiased by recent or dramatic events, and enhanced by the great number of entries derived from the entire institution, rather than from one physician's experience. Therefore, the physician's management decision can be based on far more accurate and unbiased information than could be possible without the computer.

Electrolyte and Acid-Base Consultation System. Beth Israel Hospital in Boston, Mass., has constructed a computer consultation program to help physicians manage patients with electrolyte and acid-base disorders (Bleich, 1971). The program directs a dialog in which the physician, or other user, enters clinical and laboratory information. On the basis of the abnormalities detected, the program asks traditional questions to further characterize the electrolyte and acid-base disturbance. During or after the completion of the dialog, an evaluation note is produced that contains a list of diagnostic possibilities, an explanation of underlying pathophysiology, therapeutic recommendations, precautionary measures, suggestions for further studies, and references to the medical literature.

HELP. The HELP system is the core of a complex group of computerized subsystems developed at the Latter Day Saints Hospital in Salt Lake City, Utah (Giebink and Hurst, 1975). A variety of findings on symptoms data and patient status information is incorporated into the system after being acquired through laboratory tests, blood gas analysis, intensive care monitoring units, multiphasic health testing stations, and medical record abstracts. Decision logic that uses a variety of statistical techniques is employed to develop probabilities for certain diagnoses or treatment selections based on the historical data base. In addition, the system provides warnings of patient conditions that may require intervention in several areas of the hospital, including the emergency room. The rapidly growing data base and the decision logic are also being used in evaluation and research.

INTERNIST. The INTERNIST was developed at the University of Pittsburgh as a computer-based diagnostic consultation system for problems in internal medicine (Lawrence, 1978). It is based on assigning rough estimates of the likelihood of the association of a disease, given a finding, and a similar estimate of the likelihood of the finding, given a disease. It currently covers about two-thirds of internal medicine. Internist represents an attempt to model, within a computer program, the thinking processes that the author uses to evaluate a case and make a diagnosis. It is not a model based on probabilities in the statistical sense and, therefore, is subject to the individual perspective brought to the model. The system mimics the diagnostic behavior of the "excellent clinician" and is an

example of a number of decision analysis types of systems currently being developed and tests.

Indiana University Medical Center Computer Reminders System. The computer reminders system is based on the assumption that the physician must apply a few simple rules to a few items of information many times (McDonald, 1976 and 1978). Using a very simplified computer language, developed especially for this use, the physician writes the rules he/she wishes applied to the data and then lets the computer repeatedly apply them.

The system provides reminders for a large percentage of the simple clinical decisions; it assures that baseline screening tests have been done, checks that abnormal test findings are examined further, and assures that treatments are followed with appropriate measures. The system occasionally suggests a diagnosis when a particular pattern of abnormalities is evident, but more often suggests treatments that might correct a pattern of abnormalities. Studies of the system's impact on patient care are currently underway.

COMPUTER USES IN QUALITY ASSURANCE/DATA MANAGEMENT SYSTEMS

As described in the previous chapter, the practicing physician is subject to the requirements of various "quality assurance" programs designed to measure, evaluate, monitor, and/or improve medical services. Computers provide the basis for the data processing components of these programs. Some of these activities have been described earlier using examples from both the private and governmental sectors. These examples included activities sponsored by the Joint Commission on Accreditation of Hospitals, private health insurance companies (e.g., Blue Shield) in payment of claims and quality and utilization review, hospital discharge abstract services, the Medicaid Management Information System (MMIS), and activities related to Professional Standards Review Organizations (PSROs). The latter two programs are briefly summarized below, followed by short descriptions of two multipurpose systems.

Medicaid Management Information System

MMISs are used to assist States in the management of their Medicaid programs. The Surveillance and Utilization Review (S/UR) module of MMIS: 1) produces a pre-screen set of profiles that have been compared to the average pattern of care as defined by the State, 2) limits production of profiles to those providers demonstrating aberrant behavior, 3) performs postperformance utilization review, and 4) offers options in the individual State's approach to utilization review. *

Professional Standards Review Organizations

PSROs are federally mandated programs organized primarily to determine the medical necessity and appropriateness of medical services provided to patients in federally financed programs. Major functions of the review process include admissions certification, continued stay review, retrospective review, and medical care evaluation studies.

Computers are used principally to schedule both admissions certification and concurrent review activities and to evaluate patterns of care. Computer applications that

^{*}See ch. 3 for a more detailed discussion.

66 . Computer Technology in Medical Education and Assessment

provide length of stay indicators and norms and standards are also used to properly certify the appropriate length of stay for a particular case. Review function computer applications include profiling and, to some extent, evaluating medical care. Since profiling is an aggregation of data elements, the computer is particularly useful because it handles a wide variety and a large number of cases. Statistical analyses are performed to display the information in a variety of ways. The computer is used in medical care evaluation studies not only to process data that focus on the individual problems being studied, but also to monitor those problems after solutions have been implemented to determine whether or not the action taken has solved the problem. *

Computer-Stored Ambulatory Record System

The Computer-Stored Ambulatory Record System (COSTAR) developed at Massachusetts General Hospital (Barnett, 1976) is used in the Harvard Community Health Plan to provide automated medical records and business office support. The data base is also available for retrospective quality assurance review and to preprogram reminders or alerts within the system for patient followup and selection of preferred therapies.

Problem Oriented Medical Information System

The Problem Oriented Medical Information System (PROMIS), developed at the University of Vermont (USDHEW, 1977) is unique in two respects. It not only radically restructures the medical record, but also directs the process of clinical care. The PROMIS laboratory staff developed these capabilities in order to address problems hindering the provision of medical care: dependence on the physician's memory, ineffective organization for massive amounts of medical data, and lack of meaningful feedback about the appropriateness of care.

In PROMIS, data is organized by patient problems. The computer record is structured around four phases of medical action: an initial *data base* on each patient, including medical history and physical examination; a *list* of the patient's problems; diagnostic and treatment *plans* for each problem; and *progress notes* on each problem indicating how the patient is progressing during therapy. Except for the initial data base, every entry into the computer record is associated with a particular problem of the patient. Thus, when a technician enters the result of a laboratory test, the data is entered under the problem for which the test was initially ordered. By structuring the record in this way, all information pertinent to a problem is organized logically for review by the physician and other medical care professionals.

Personnel enter data about patients through video terminals. The videoscreen of the terminal displays an array of choices, and the provider makes a selection by touching the screen. Data is entered by the medical care professionals who originate them. For example, physicians and nurses enter notes about the patient's progress, radiologists enter notes as they read films, and technicians in the clinical laboratory enter results of tests. In addition, patients enter their own medical histories. Each staff member has a unique identification code that allows entry and access only to those parts of the computer record necessary for the provision of care.

^{*}See ch. 3 for a more detailed discussion, including the application to ambulatory care in New Mexico and Utah.

SUMMARY

Medical education and assessment and patient care could be integrated, interactive subsystems of the medical care system. Computer technology not only has shown promise in helping achieve this integration, but also has aided advances in this direction. The capabilities of computer-assisted instruction, testing, and management to provide individualized assistance in the medical school curriculum are being extended toward individualized examination on entry to and during practice. In the future, some of these examinations will be based on data extracted from the actual practice experience of examinees, thereby allowing both for better measurement of practice strengths and weaknesses and for better definition of continuing education needs.

Computer capabilities to maintain and analyze institutional, regional, and national data bases on actual patient care will allow better correlation between the condition of the patient (outcome) and patient care techniques (process). These capabilities also will help provide better indicators both of patient care techniques that are successful or unsuccessful for specific diseases, individuals and situations, and of the costs, for example, of alternative methods of diagnosis and treatment. This type of information will enhance an outcomes approach to quality of care measurement and reduce the present emphasis on process measures. Better definition and correlation also will restructure the content and emphasis of undergraduate and graduate medical education.

The creation and use of computer-based simulations enable better assessments of aspects of clinical competency. These simulations may hasten consensus formation regarding the definition of certain dimensions of competent care because of the requirement for specificity in therapeutic responses to simulated illnesses. Since patterns of performance behavior can be recorded in detail by the computer, cumulatively, they can be used to create diagnostic and therapeutic profiles for particular diseases. These cumulative profiles of care can be used in PSRO and other quality review situations as more detailed standards against which to measure appropriate care. For example, the extent to which a repertoire of tests or a particular test is indicated or required can be used in addition to the rougher measures of quality of care, such as length of hospital stay.

Inversely, using computerized recordkeeping profiles, appropriate care for disease entities can be defined and verified in great detail within norms and ranges. These profiles of appropriate care can be added to the medical knowledge base and create computer simulations for teaching and testing. Such simulations could expose students to patients and diseases not experienced in the clinical setting. Through individualization of instruction and the use of computers for patient scheduling, students in clinical rotations can be matched to patients to provide them with the proper, rather than random, exposure to different diseases. These experiences can be recorded so that simulated patients may be used appropriately to fill voids in experience.

The computer's information storage and retrieval capabilities can increase the accessibility of the medical knowledge base. Textbook as well as clinical data can be quickly updated and restructured to provide the same data organized in a variety of ways. This ability to reorganize, restructure, and analyze data also facilitates creation of new medical knowledge.

The computer can be a viable and cost-effective resource if used in a coordinated manner. Having one computerized data base for education, another for assessment,

68 • Computer Technology in Medical Education and Assessment

another for patient care, and another for administration, not only is inefficient in most cases, but also prohibits reaping of combined benefits. Thus, one of the challenges facing the increased use of computer technology in medicine is the design of a system that facilitates and enhances data and information flow among the subsystems described in this chapter.