A PRELIMINARY SYSTEMS DESIGN FOR A COMMUNITY HEALTH DATA SYSTEM

by

William Weldon Moore, Jr.

A Thesis Submitted to the Faculty of the

DEPARTMENT OF SYSTEMS ENGINEERING

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

In the Graduate College

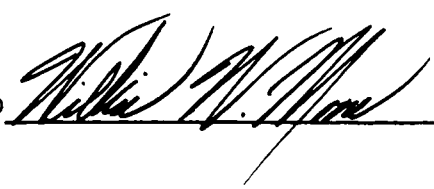
THE UNIVERSITY OF ARIZONA

1 9 7 2

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED 

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

J. L. SANDERS
Professor of Systems Engineering

May 24, 1972
Date

ACKNOWLEDGMENTS

I wish to express my gratitude to all of those persons who contributed to this research.

The time and counsel of the members of my thesis committee, who are Dr. J. C. Sanders, Dr. R. L. Baker, and Dr. W. R. Ferrell, are greatly appreciated.

A very special thanks goes to Mr. and Mrs. John R. Delair, whose collective patience, understanding and encouragement were vital to the completion of this thesis.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

ABSTRACT

Health information systems are gradually becoming popular with health agencies across the country. Little work, however, has been done to provide adequate information systems to those working in the planning area. The Health Information System for Evaluation and Planning described in this paper is a system which would provide the basic information needed for health planning in a metropolitan area. Fundamental requirements for the data base, the software configuration, and the file organization are introduced. This preliminary concept is an important first step toward a system which could satisfy almost all of the needs of health planning agencies, such as the Comprehensive Health Planning Council and the Regional Medical Program.

CHAPTER I

INTRODUCTION

1:1 Statement of The Problem

The purpose of this paper is to provide a preliminary concept for a comprehensive community health information retrieval system with an emphasis toward health planning for a well-defined metropolitan area. This Health Information System for Evaluation and Planning (HISEP) should provide a means for continually surveying the health status of individuals and specific population groups, and for evaluating the health care systems that serve these individuals and groups (13, 26).

In attempting to meet this need, HISEP should geographically pinpoint neighborhoods where significant health risks exist, and define the characteristics of the resident population. Combined with a description of the over all health care system in terms of the health services provided, facilities for providing care, manpower used in providing services, and the population served, this information will yield a means to the health planner for approaching solutions, not only for the medical problems, but also for health related social issues--such as illegitimacy and poverty (21). If HISEP can give the health planner the information he

needs, information that is complete, accurate, and timely, his ability for decision making should be greatly enhanced. It is axiomatic that if provided with complete data at the precise time it is needed, the planner will have the best possible tools with which to make an optimum decision (11, 32, 34).

Although the ultimate goal is a health information system for planning which will encompass a much wider area than a single community, it is important that the initial step be conducted in a relatively small, well-defined urban environment. The system can then be expanded in a stepwise or modular fashion. The overall system design should also provide for the future integration with other types of health information systems. Once all of the health information available is merged into one large system, the health planners will have all of the information they need for both surveillance and evaluation of the entire health delivery system. In particular, the health information systems being developed for the University of Arizona Medical Center and the Indian Health Service at the Sells Service Unit were given special attention as possible systems to be integrated with HISEP at a later date (see Sections 1:4 and 2:6).

1:2 Primary Goals, Fundamental Concepts for HISEP

HISEP is conceived in terms of the following goals (1, 33, 53):

(a) The study of the correlations among various diseases;

(b) Determination of the health status of individuals and specific population groups and their needs;

(c) Surveillance of the status of chronic diseases;

(d) Prognoses for different diseases under various treatments and circumstances;

(e) The evaluation of the extent of the load on health services by subgroups of the population.

These goals reflect a different approach for health information systems. HISEP is oriented solely toward the health planner, and as a result, many problems arise concerning the need for various types of data and how the data should be correlated. For many diseases there is no clear-cut quantifiable relationship between a health need in the pure sense and its translation into health resource requirements and health service needs (17). The health planner must be neither drowned in data, nor should he be given sketchy or poorly developed data that could lead to wrong conclusions. It is therefore imperative that the information supplied to the health planner be well organized and concise. Cost benefit analysis of data needs and resources should be used to determine what information will really be useful to the health information system (30).

1:3 Comprehensive Health Planning

Comprehensive Health Planning is the legal basis for the feasibility and future funding of a system such as HISEP.

Under Comprehensive Health Planning, the machinery to begin a realistic and concentrated effort toward planning for the needs of the nation's health system was started.

1:3:1 Definition:

Comprehensive Health Planning (CHP) was created in 1966 under Public Law 89-749, The Comprehensive Health Planning and Public Services Act, and it is now active in every state. The law sets forth as a national goal, "the highest level of health attainable for every person." It holds the individual states responsible for the requisite coordination of the planning and provides Federal funding for health planning councils which represent each planning area.

Although the health planning councils provide no medical services directly, they try to determine the type, capacity, and location of services needed in the community. They also make recommendations to the State agencies regarding needs for new policies or legislation, and encourage coordination and cost-efficiency in the delivery of health within their area (22). A general definition of comprehensive health planning includes: the assessment of health needs, resources, and development opportunities; the formulation of goals, objectives, and policies to guide long-range planning; allocation of resources available for both preventive and curative health; the preparation of

plans and programs; and the continual surveillance of the health status of individuals and population groups for the purpose of evaluating health care systems (22, 36, 41).

In a broader sense, CHP is intended to provide an institutional framework for concensus and to serve as a vehicle for relating health concerns to areas of economic and social planning (22). HISEP's design is based upon meeting these goals as effectively as possible.

1:3:2 CHP and HISEP:

The major elements in comprehensive health planning can be divided into four major areas:

(a) Establishment of goals and objectives for health planning;

(b) Generation of alternative means of achieving the established goals;

(c) Evaluation of the alternatives and the selection of the best alternatives for implementation; and

(d) Evaluation of the degree of goal achievement resulting from the implementation of chosen alternatives in preparation for refinement and improvement of the planning process.

HISEP could contribute to these objectives in the following ways:

(a) The identification of the transition of individuals from the normal population to a special category

called 'high-risk,' which is a group of people whose health is such that frequent treatment by health personnel is vital to their well being (39).

(b) The development of a realistic preventive health care program through the surveillance of the population, thus disclosing needs for preventive care among specific populations and age groups that can be identified precisely;

(c) The development of treatment monitoring and follow-up of the population;

(d) A detailed study of epidemiology and disease inter-relationships, showing the growth patterns of various diseases within a population and making correlation studies on how these various diseases might be related, (e.g., cause and effect, carriers, etc.);

(e) Disease patterns plotted versus environmental and genetic factors;

(f) The availability for study of the data on re-admissions of patients when they change health facilities;

(g) The ability to rank order health problems according to their prevalence, seriousness, length of illness, etc.;

(h) Cost-benefit analysis (e.g., how an increase in the number of physicians effects the health service, how health programs effect population subgroups differently, etc.); and

(i) The evaluation of program effectiveness through a continuous surveillance of the population and thus determining the effect each program has on the population to which it was applied.

Although HISEP will not provide all of the answers in any of these areas, it is a step toward providing the planner with the various types of information he needs. How HISEP might contribute to each of these areas will be explored in considerable detail in the following chapters.

1:4 Other Projects in Health Information Systems

There are currently many health information systems in use throughout the country, and much work has been done within the last five years to computerize the information needed in the health field. Unfortunately, each system seems to have a different objective, and they are characterized most by their diversity and fragmentation. The bulk of the work in applying computers to the health field has gone toward making the operations within large hospitals more efficient (8). These systems try to reduce the tremendous volume of paperwork and accounting procedures that are becoming commonplace in the modern hospital. The Lockheed study for the Mayo Clinic (27) and the Total Hospital Operating Medical Information System (THOMIS) developed at the New York State University Hospital (15) are both giant steps

toward making the gathering of useful data within the individual institutions both feasible and economical. These two systems provide patient information from preadmission to discharge along with facility utilization statistics in a consistent manner by using the Weed Medical Research System (54). It is vital to any health information system for planning that the data be provided from the health facilities in a timely and efficient manner (30). The aforementioned systems will contribute greatly toward making this primary goal of health planning a reality.

On the larger scale of areawide health information systems for planning, there are several systems which have taken important first steps toward providing health information for planning in a usable form. In Denver, Colorado, the neighborhood health centers have outpatient clinics that report health data on all patients to the centers, which in turn report this data to the hospitals in the city. They then try to collate this data on a master file and various high risk files. Although the concepts and organization are excellent, they have never determined how to present the data, nor have they determined how to use the information that they have (35).

Another important system for planning is the New Haven Health Information System. Its primary purpose is to pinpoint geographically those neighborhoods in New Haven

where there is a significant health risk and to define the characteristics of the resident population. This system provides information in the form of tabulations, hand-drawn maps, and analytical guidelines covering small social areas within the city. The most important contribution of this system is the manner in which it utilizes U.S. census data. It was developed as a model to make maximum use of the 1970 census (45).

In Michigan, the State Health Department has developed an excellent system whose primary purpose is to deliver statistical information on the health of the people of Michigan, in a timely way, back to the local health agencies. The Michigan Center for Health Statistics delivers technical services in guiding the collection of the information, in analyzing it, and in reporting back the results. HISEP can benefit greatly by the methods developed in Michigan to assure that the data collected are uniform and meaningful. Out of an effort similar to this, HISEP can augment its data and accrue such benefits as current population estimates for neighborhoods and other small areas, the opportunity to proceed from relatively superficial levels of inquiry to greater precision and more specific detail, and the opportunity to identify geographic areas for concern before they become extreme health hazards and to observe the rate of change for those neighborhoods in response to "treatment" (30).

In California, another planning effort is represented by the demonstration project--California Health Information for Planning Service (CHIPS). CHIPS is particularly important because it managed to achieve a cooperation among many different sources of health information. The varied organizations contributing to CHIPS include: The State Department of Public Health, the California Hospital Association, the California Medical Association, regional planning groups, and individual institutions. The cooperation of these varied agencies proved that it is possible to bring together all aspects of the medical profession into a partnership for the purpose of providing information for health planning. In addition, CHIPS provides an excellent guideline for a health information system for planning in the areas of: a post-discharge patient care abstract service for hospitals, formats for suitable areawide planning reports, development of key elements of a manpower planning subsystem, and a method for integrating the in-house hospital computer systems with the overall system required for areawide planning. CHIPS currently represents the system that most closely resembles a health information system for planning (41).

The Health Information System Study of the MIT--Harvard Joint Center for Urban Studies documents the present state of the art achieved in health information systems. It also discusses at length the motivations that are currently

present in the health field for better health information systems (35).

Other agencies with health information systems that have contributed important concepts include: the National Center for Health Statistics (30); the National Communicable Disease Center (30); the American Hospital Association (28); the National Office of Vital Statistics (30); the Metropolitan Data Center Project (52); and the Ohio Community and State Planning Project (8).

There is also much important work in health information systems being done in southern Arizona. The pilot experiment for the Indian Health Service at the Sells Service Unit is a management information system which contains three important subsystems: (a) a high risk file for delinquent reporting and a progress treatment check; (b) population surveillance; and (c) a Weed problem list (54) on each patient (4). The work that has been done on the high risk file could contribute greatly toward the high risk file proposed for HISEP. In addition, the other features of the Indian Health Service system could be integrated into HISEP at a later date to broaden HISEP's contribution to the field of health.

The University of Arizona Medical School computing facility will contain many of standard hospital packages such as patient billing, cost accounting, etc. This type of

information is needed for any system that serves all of the health needs of the community. In addition, the system will include some innovative packages such as a check to determine if the same disease receives fairly uniform treatment throughout the hospital. These facets of health in formation could also be integrated into HISEP at a later date (see Section 2.5).

Many research groups have contributed to the field of health information systems. However, much remains to be done, particularly in the area of health planning. The development of sophisticated input methodologies that are practical from the user's standpoint and the development of effective management techniques for correlating health data once it has been gathered are still in the future (2, 3). Although this paper does not intend to set limits on the amount of data or software for a HISEP, it does try to incorporate in the initial system all of the data base and software configurations needed for an information system which would contribute to all of the health planning components in a meaningful manner.

CHAPTER II

FACTORS INFLUENCING SYSTEM DESIGN

2:1 General Restrictions on a System for Planning

There are several considerations that must be taken into account when designing a health information system for planning that are not relevant to most information systems which are used for administration and operations. While administrative procedures and operations can be quite varied, the problems of providing information to support them are usually well-defined and the system designer knows what data are needed, what reports are to be prepared, and the manner in which the inputs must be processed to produce the required outputs. In planning, however, the planner tends to proceed heuristically with problems which can rarely be adequately stated. As a result, the requirements placed on the information system for planning are strongly dependent upon the individual planner involved. For example, the planner may use multi-variate analysis--a set of statistical procedures--to analyze the results of various computations. Only after the analysis is he able to determine what transformations, correlations, and cross-tabulations will be required (9, 52). In addition, the information required by the planners differs from that required by operations

personnel. The latter are primarily interested in the current status, while planners are interested in the past and present only to the extent in which these periods influence the future. Usually, the planner tries to predict the future by looking at the trends in the past while operations personnel need to keep current data only (52).

2:2 User Requirements

In order to accommodate the special needs of the planner, HISEP should have as a fundamental portion of its design a great amount of flexibility. It should be able to handle and cross-reference many files simultaneously. Most computer-based information systems that now exist are essentially special-purpose single-file systems. A comprehensive information system, in contrast, must handle many files at once. The planner may want to bring together many different files for purposes of comparison and analysis (38, 52). Thus, HISEP should be open-ended with a wide variety of data handling capabilities.

The system should also have a conversational and interactive mode. In other words, the user should have the capability of directly communicating with the data that is available. Critics of systems with conversational capabilities point to the difficulties in implementation, to the overhead costs, and to the greater satisfaction expressed

with batch processing by persons who are engaged in large-scale computation. Even if these criticisms were valid in general, they are largely invalid with reference to information systems. Reference-type usage and data manipulation turn out to be the kind of operations where many small steps are taken in an unplanned way, in contrast to the typical batch-type of operations. Information system users are very typical of the people whose natural work mode requires an interactive software capability, the type of model on which a conversational computer operation is based (52).

The user should be able to converse with the computer in the conventional terms of his discipline. While the batch processing mode of HISEP will be handled by experienced programmers, the user himself will seldom be a programmer. The system thus requires a relatively simple high order language of its own with relatively few instructions for the conversational mode of operation. HISEP should keep the burden of language learning for the user to an absolute minimum.

Finally, the system must also be as format free as possible. It must be able to accept and output data in whatever kind of code it may have been stored, whether alphanumeric, column-binary, unconventional bit patterns, graphic representation, or any other form. Graphic displays are particularly effective for rapid man-machine

interaction. If a system is to accept data in any format, then it must require that someone specify the format seed. HISEP must provide a means for defining and identifying all of the formats in which data appear (35, 41, 52).

2:3 System Management Requirements

Design criteria must also be influenced by the needs or system management. HISEP must have an ability to readily update and purge data which has become obsolete. The system should be able to keep track of the many locations where a particular entity might be stored, and change the item in all of the applicable files with a single command. It also should have an inbred capability for its system operating programs to be easily modified and revised (6, 30, 52).

Easily performed error correction is vital to HISEP. Past experience has shown that in areas such as survey analysis, more time is spent on detection of errors in the data and in their correction than in any other area (30). The raw data should be validated prior to entry into the system's data base (see Section 3:6).

For an ideal system from the system management viewpoint, the files should be simple, compact, and also be well-structured (38). Unfortunately, the user's requirements fundamentally conflict with these concepts. The requirement for cross-referencing and correlating data in

various files necessitates a very complex indexing system. A conversational mode requirement dictates an immense software package for implementation. Free formatting capabilities require a more lengthy file with a fairly large software system dedicated solely to storing and retrieving the data correctly (6, 38). As the system tries to satisfy the needs of the user, the time and money needed for implementation will go up drastically. The solution lies in a compromise. Systems management should have more emphasis in the earlier stages of development, with the requirements of the system user gradually having a more important role in augmenting the system (6).

2:4 Hardware Considerations

As with all large information retrieval systems, HISEP will require several levels of memory, including a large bulk memory of pseudo-random access character such as a disc. The data in the system should move along the core, disc, tape, and other media according to algorithms based on the use of the data as the computer monitors it.

To be more specific about hardware requirements is not justified at this stage of development. Rapidly changing equipment availability and the restrictions encountered in different geographic areas prohibit any meaningful approach to suggesting specific hardware (52).

2:5 Software Considerations

As a result of the user and system management specifications, the software for HISEP must meet certain minimum requirements. These can be listed:

(a) Accept and output data in a variety of formats;

(b) Handle very large volumes of data;

(c) Operate on individual elements within the data set;

(d) Manipulate and alter data set structures;

(e) Provide fast retrieval capabilities;

(f) Provide statistical analysis techniques.

In addition, HISEP must keep a record of the operations performed within it so that they are retrievable and repeatable. This allows the user to have an easy note keeping capability for all of his analysis, but more important, it is far cheaper to store a record of the operations performed and then recomputing the statistics if necessary than storing the derived statistics.

Finally, the computer should not only be faster and better than any manual system, but it should also be competitive in cost. To do this, the data structures must be efficient. HISEP should be optimized for economy of bulk data storage and ease of use, not for the speed of computations as is most common in the computer field today (8, 41, 52).

2:6 Future Integration with Other Health Information Systems

The basic design of HISEP must not hinder the possible future integration of the system with health information systems designed for other purposes. For this reason, the complete file structure should be designed with a modular concept. Some files which might be structured on an aggregate basis, such as patient medical records, should be set up on individual entities. This will enable systems which are fundamentally built around the Weed problem list (54) to be easily meshed with HISEP. The concept of HISEP being only a part of a much larger and more versatile health information system in the future will necessarily play a major role in the design criteria (24).

2:7 Legal Considerations

Without doubt, the most important legal problems in the area of a health information systems for planning are the privacy and confidentiality of information concerning specific individuals in the system (14). While it is not the concern of this paper to approach the philosophical aspects of this problem, it is most important that HISEP use the most advanced methods for assuring that the individuals concerned are protected.

There are currently several techniques available which will protect files from being accessed by unauthorized personnel. Project System MAC at MIT has been very

successful in this area and HISEP could use the procedures developed there (see Section 4.5).

Although HISEP must use some file protection devices, they should be kept at a minimum. As the codes used to access the confidential files become more complicated, the software associated with the coding becomes very large (38). Furthermore, inconvenience brought on by a lengthy file entry message may yield a propensity in the user for not using the system at all.

CHAPTER III

DATA REQUIREMENTS

3:1 General Requirements on the Data Sets

Many diverse criteria have been taken into consideration in determining the type of information that should go into the initial data base for HISEP. There seems to be a strong tendency in existing information systems for planning simply to gather great volumes of data, facilitate the retrieval of this data, and produce summary reports (8). To counteract this problem, all input data to HISEP should be carefully weighed to ascertain their usefulness to the health planner. Not only will this lower the data collection and computing cost, but it will also enable the health planner to find the pertinent information relevant to his particular problem with greater ease.

Every element in the data base should satisfy some specific present or projected need. In addition, there should be a periodic review of the information contained in HISEP to assure that each of the data entities continues to serve a useful function. To a large degree, HISEP should be a continuing development as new concepts in health planning bring new data and ideas into the system (9, 11). Presently, there is much transition in thinking among health

planners concerning what type of data is needed to adequately plan for the future. Even such standard data as the number of beds, floor space and health needs are being questioned. Consideration is gradually being given to more fundamental concepts, such as long term goals, geographic variances, and community development (25, 55).

The usefulness of a data item depends upon its validity, its ability to relate measure to important demographic and socioeconomic factors, and the availability of time trends which enable reasonable projections (8). Although determining the value of data is quite often a difficult task, some basic data seem to be essential to all health planning. For example, all health planning is fundamentally related to people and what happens to them over a period of time. Therefore, the population characteristics are of primary importance. Indicators such as demographic data, socioeconomic data, health facility utilization, and manpower allocation and requirements yield some invaluable information for present and future guidelines.

3:2 Data Required in Planning for Health Resources

The information required for health planning can be broken into two main areas: a description of the current health care system and an estimation of the current and future demand for health care services (8, 17, 30, 52).

3:2:1 A Description of the Current System of Health Care in a Planning Area

The current system of health care in a planning area can best be described in terms of health services provided, facilities for providing care, manpower used in providing services, and the population served. To quantitatively describe the health care system, the data must include:

(a) A description of the facilities by the type of facility and the services that it can provide (It is often desirable to view the different types of facilities as the strata of the system and consider different categories of population groupings for each strata if the area for planning is very large.);

(b) A classification of the various types of health services provided in the planning area;

(c) A description of the distribution of health manpower within the planning area by category;

(d) The characteristics of the persons receiving each health service (The characteristics to be measured are those which identify the person as being a member of a particular population group.);

(e) The type and length of care given for the health services provided in the planning area.

3:2:2 An Estimation of Current and Future Demand for Health Care Services

In estimating the current and future demand for health care services, the following data should be added to those listed above:

(a) The population base by demographic groups and expected growth rates (Demographic and socioeconomic data have a real bearing on such diverse areas as diagnosis, method of treatment, length of stay, compliance with medical regimens, etc.);

(b) The relative cost to the patient for each health service provided;

(c) The current excess demand for personal health services (As reflected by waiting times, expenditures on preventive care, etc.);

(d) The current supply and cost of constructing and supporting health facilities.

3:2:3 HISEP Basic Data

The foregoing description of the kind of data needed to quantitatively describe the health care system implies a vast, costly, and time-consuming program of primary data collection. It is not likely that such an extensive program can be undertaken by any CHP agency. It is possible, however, to reduce the burden of data collection through the integration and analysis of a smaller and more fundamental data set (52).

This 'basic' data can be divided into five major areas:

(a) Demographic and socioeconomic data;

(b) Medical facilities;

(c) Health services;

(d) Health manpower; and

(e) Environmental health (8, 30, 41, 52).

3:2:3:1 Demography. Demography is the study of human populations and the determinants and consequences of population change. This information can be used to identify health problems, determine possible solutions to these problems, and choose the solutions which will have the probable best effect upon the population as a whole. In addition, historical demographic data can be used to reassess the original problem to see if the action taken was in fact the optimal solution.

3:2:3:2 Medical Facilities and Services. Medical facilities provide invaluable data for planning. The types, numbers, and characteristics of health facilities within the planning area are needed to determine the potential for health care for the surrounding populace. These facilities also furnish a large amount of the health service data. Hospital admission and discharge abstracts, relative frequency of purpose for seeking health care, and other important information have their origins in the medical facilities.

3:2:3:3 Manpower. Manpower statistics must take on a dynamic character. It does not suffice to know the type of services they perform, the setting in which they are provided, the equipment and technology used in providing them, the cost of providing the services, and the characteristics of the population demanding their services, as well as the demand for services, must be known. These statistics are vital to health planning.

3:2:3:4 Environmental Health. Environmental health conditions are also a valid portion of the data base for an overall system viewpoint. Environmental health conditions might include such items as water supply, waste disposal, residential environment, occupational health, injury control, etc.

3:3 Recommendations for the Initial Data Base

Since all of the available data cannot be input into HISEP, there should be some criteria to decide exactly which data are vital to health planning and which data are secondary in importance. Unfortunately, the only real criteria used to differentiate between high and low priority data are subjective. Therefore, the importance of specific data is determined almost solely by experience in the field

and personal preference among the health planners. There does, however, seem to be a basic core of data that has been proven to be of value to planning through past experience. This 'basic' data will make up the data base that is being recommended for the initial system. Although this 'basic' data is not an optimum base for HISEP; it should be at least adequate for the initial system (8, 10, 12, 14, 23, 24, 29, 30, 31, 35, 37, 40, 41, 42, 43, 44, 48, 50).

3:4 Health Information for Planning

3:4:1 Demographic and Socioeconomic Data

The data are distributed by: age, sex, race, marital status, housing conditions, place of residence, education level, employment, and income. This class of data is generally available through survey agencies, such as the U.S. Census Bureau. The approximate file style required for this data for a population of 100,000 is 5,000 words.

A. Vital Statistics

(1) Mortality data

a. Infant

b. Other

(2) Birth Statistics

(3) Marriage Statistics

(4) Divorce Rates

B. Population Characteristics

(1) Concentration of population by block

(2) Geographic distribution

(3) General characteristics of residents by block

(4) Migration characteristics

C. Area Statistics

(1) Job availability

(2) Educational facilities

(3) Transportation availability

(4) Health facility availability

a. Location

b. Service available

c. Max service potential

1. In-Patient

2. Out-patient

(5) Future development plans

D. Intervening variables

These are variables which may directly affect the demographic process.

(1) Customs

a. Marriage

b. Family size

c. Medicinal leanings

(2) Attitudes and Practices

a. Contraception

b. Sexual behavior outside marriage

3:4:2 Medical Facilities and Services

This type of data is maintained at most hospitals and clinics. The file size requirement will be approximately 150 words per facility.

A. Identity and Location

B. Type of Facility

(1) Hospital

(2) Long term care facility

(3) Health Center

(4) Rehabilitation center

(5) Mental health care facility

(6) Home care organization

C. Basic Inventory

(1) Size (number of beds)

(2) Average number of patients per day

a. In-Patients

b. Out-Patients

c. Average length of stay

D. Organizational Characteristics

(1) Ownership/Control

(2) Memberships/Affiliations

E. Service Characteristics

(1) Special Services available

(2) Utilization of services

(3) Type of area served

(4) Patient characteristics

(5) Payment policies and options

F. Manpower Distribution

G. Financial Support

3:4:3 Patient Data

For an optimum system for planning, there should be a file containing every member of the population being served with the information describing each entry listed below. This Patient Data File will be used to generate data for several other files within HISEP. If it is determined that a file for every individual is too demanding on the system, a sampling technique could be substituted. The file size required for patient data is approximately 125 words for each individual. The information should be available from hospital admissions and discharge sheets, clinic admission data, etc.

A. Identity

B. Characteristics

(1) Age

(2) Race

(3) Sex

(4) Marital Status

(5) Income/type of payment

(6) Residence

(7) Education level

C. Attending Physicians

D Medical History

(1) Immunizations

(2) Previous contagious diseases

(3) Chronic conditions

(4) Previous illness or serious injury

E. Admission Status

(1) Admitting diagnosis

(2) Relation to previous medical history

(3) Recommendations for service

(4) Anticipated length and depth of service

F. Services and Treatment Rendered

(1) In-Patient Service

(2) Out-Patient Service

(3) Extended Treatment

(4) Charges

(5) Placement on "High Risk File"

G. Referrals

3:4:4 Health Personnel and Manpower

This data is available through the health facilities. The file size requirement will be approximately 600 words for each service facility.

A. Distribution of Manpower
   (1) G.P.
   (2) Specialist
   (3) Surgeons
   (4) Dentists
   (5) Registered Nurse
   (6) Lab Technicians
   (7) X-ray Technicians
   (8) Aids/Attendants
   (9) Other

B. Current Needs

C. Future Needs

D. Other Pertinent Factors
   (1) Quality and Type Service available
   (2) Equipment and technology available
   (3) Cost factors
   (4) Demand for Service

E. Training Available

3:4:5 Health Statistics

The health statistics file will contain data which may prove to be the most difficult to collect. Although

mortality statistics are maintained by the County Recorder, morbidity statistics and some of the health facility utilization data can only be obtained through the use of special surveys, such as the Family Health Survey previously described. This file will require approximately 400 words for an area of 100,000 population.

A. Mortality Statistics
   (1) Infant
   (2) Other

B. Morbidity Statistics
   (1) Characteristics of affected population
   (2) Effects on behaviors
      a. Restricted activity
      b. Bed disability days
      c. Work loss days
      d. Hospital days
      e. Chronic disability

C. Utilization of Available Health Facilities
   (1) Percent of population using facilities
   (2) Percent of population needing facilities
   (3) Relative number of in-patients/out-patients
   (4) Average length of stay

3:4:6 High Risk Files

An important aspect of health planning is to maintain contact with those patients who are known to need

follow-up treatment. Implementation of a procedure to publish lists of patients who are delinquent in returning for needed treatment could be easily accomplished with the information already contained in HISEP. The file size requirement imposed by this file will be minimal sinze no new data is required. This "High Risk File" can be organized in the following manner:

A. Communicable Disease

A disease that is easily transmitted from one individual to another is termed a communicable disease. HISEP can watch for outbreaks by carefully keeping records of the occurrence of these diseases within the following classifications:

(1) Location defined by census block

(2) Patient characteristics

(3) Size and scope of the problem

(4) Scheduling of visits into the area

B. Non-Communicable High Risk Diseases

This classification concerns those diseases which are classified as having "high risk" potential. Persons that have contacted these diseases should be kept in constant contact with medical personnel. This can be accomplished by publishing frequent lists containing:

(1) Patient identification

(2) Patient address

(3) Last known visit to a physician

(4) Next scheduled visit

(5) Current/Overdue for treatment

C. Family Related Disease

Serious disease which can be linked to genetic causes is also of concern. Individuals who are subject to such disease should visit their physicians periodically. Therefore, the following should be published for those under this category:

(1) Patient identification

(2) Patient address

(3) Last known visit to a physician

(4) Next scheduled visit

(5) Current/Overdue for treatment

3:5 Major Sources of Information for Health Planning

The data acquisition plan must be considered as a fundamental portion of the overall design of HISEP. Almost all health information systems have a continually changing data base with a large data gathering responsibility, and HISEP will be no exception (52). In addition to the constraints put on the data in Section 3.7, consideration for choosing the information for the initial data base must be given to the ease with which the data can be gathered. The acquisition of necessary health data will probably be the largest hindrance to implementing HISEP. Many health

agencies either do not have the capability or they are not willing to sacrifice the time to furnish planners with vitally needed data. One of the primary prerequisites for the implementation of HISEP will be the formulation of a scheme to convince the various health agencies of their responsibilities in connection with providing data.

While health data is difficult to collect, the task is not impossible. A few of the agencies which can contribute invaluable data to HISEP are listed below (8, 30, 31, 35, 41, 45).

1. U.S. Department of Commerce, Bureau of the Census
2. National Center for Health Statistics
3. U.S. Department of Health, Education, and Welfare.
4. Community Profile Data Center, HEW
5. American Hospital Association
6. American Osteopathic Association
7. American Medical Association
8. State Department of Health, Division of Vital Statistics.
9. State Department of Health, Tuberculosis Division
10. State Department of Health, Environmental Division

11. State Professional Societies
12. AID to Dependent Children
13. Blue Cross
14. Local Medical Facilities
15. Local Schools' Census
16. Local Draft Boards
17. Local Election Commissions
18. Local Welfare Agencies

This list is certainly not complete, but it does mention the primary sources of useful data in the health planning area. A more specific listing of data sources with their applicability can be found in the Appendix.

3:6 Need for Additional Surveys

Some information which can contribute greatly to health planning is simply not available through any source. Surveys can be initiated to provide this data, however, and some excellent surveys have been instituted in Michigan and New Haven, Connecticut, for this purpose.

The Family Health Survey in New Haven, Connecticut, is a pilot experiment developed by the U.S. Census Bureau. This survey provides data about such critical items to health planning as: the use of health resources for the prevention of illness; sickness, disability, and activity restrictions among the population; family planning activities; and child care (48).

In Michigan, the Michigan Health Survey provides detailed information relative to the current health status or small neighborhoods in several cities. This survey was discussed earlier in Chapter II (30).

The drawback to these and other surveys, of course, is that they are a heavy financial strain on the operating budget for the health planning organization that sponsors it. It must be decided at the local level whether or not the money should be allocated to this particular function.

3:7 Validation of Input Data

A major problem for any computing system with a large and rapidly changing data base is to assess the quality of the raw data inputs. While the data analyst can and does influence how information is collected under some circumstances, quite often the procedure for data collection is inadequate. Poor sampling techniques and incomplete data can lead the planner to the wrong conclusions. In addition, the method for putting the data into the computer system can introduce errors in otherwise valid data. It is essential that the data analyst take great precaution in assuring that good data gathering techniques and a careful and thorough procedure for converting raw data into machine readable form (30).

The first step toward achieving a good and continual data gathering program is to initiate data forms which will

guarantee uniformity among the various agencies contributing similar types of information. It has been shown through experience that this simple action greatly reduces the errors in all phases of data gathering, conversion, and interpretation (52).

Prior to inputing any new type of data into the system, the data analyst must be certain that the data is accurate, complete, timely, and pertinent. To help with this evaluation of new data, the following data evaluation form is recommended and should be completed by the user needing the information and the systems analyst. Working together, they can decide if the inclusion of the new data is really worth the costs involved (35, 42).

3:8 Data Evaluation

NAME:

The complete name of the file recommended plus the mnemonic name or acronym associated with it.

CATEGORY:

The major area under which the file would be entered. (i.e., Patient Data, Medical Facilities, and Services, etc.)

TYPE OF USER:

The general type of user that would need the information and the frequency of use predicted.

LINKAGE REQUIREMENTS:

How much this new file would need to be integrated into other files already in the system.

UPDATE REQUIREMENTS:

How often the data would have to be updated.

SOURCE OF RAW DATA:

Where the data will be available and the accessibility of the data.

BRIEF DESCRIPTION:

A description of the major components and features of the file.

BRIEF HISTORY:

A brief outline of the reasons why this file is pertinent to health planning.

ACQUISITION COSTS:

A prediction of the financial effort involved to obtain the required data.

PROOF OF VALUE:

If available, a description of how and where the data has already been proven to be useful to health planning.

STORAGE FORMAT:

Are there any special requirements on the format in which the data is to be stored?

STORAGE CAPACITY REQUIREMENTS:

Estimated file size for a population of 100,000.

HARDWARE REQUIREMENT:

Are there any special requirements on the storage device that is used to store the data?

REPORT GENERATION REQUIREMENTS:

Will the system be required to generate a special format for the file output? How much flexibility must there be for the report format?

ADAPTION TO GEOCODING:

Will the data be appropriate for geocoding, and if so, will it be in a form that can easily be correlated with geography?

DOCUMENTATION:

Has the file been carefully documented?

PROCESSING ACCESS METHOD:

Describe the levels of accessibility required (i.e. sequential processing, random access).

EVALUATION OF USER ORIENTATION:

Will the users be able to utilize the data without any orientation?

REMARKS:

Levels of analysis, etc.

SUMMARY:

Of advantages, disadvantages, specific features particularly useful to HISEP, etc.

CHAPTER IV

BASIC SYSTEM CONCEPTS

4:1 Data Base Organization

The cost of data processing associated with a large scale computer system such as HISEP is directly dependent upon the data base organization. The costs of storing and processing data, combined with the indirect cost of the response time of the system are directly involved with the manner in which the files are arranged (38). The overall file organization should aid in decreasing the total operating cost. Dependent upon the volume, variety, and use of the data, there are several techniques for structuring the data base from which to choose. Some file structures that have been considered are, duplicate files, index files, correspondence tables, inverted files, and list processing (19, 20, 52). The more sophisticated file organizational techniques have been rejected for use with HISEP due to the highly volatile nature of the data.

As the organization becomes more complicated in a rapidly growing data base, the file maintenance costs go up exponentially. Therefore, a fairly simple file system, each file containing information about a predefined set of entities, has been chosen for HISEP. In data processing

terminology, each attribute is represented by a data item and each record corresponds to a particular entity and contains the values, for that entity, of the set of data items included in the file (52).

In particular, a form of serial file organization has been chosen in which entities are blocks of data stored in a two dimensional form which is accessible in any order. The serial ordering of a file is determined by an 'organizing key'. For example, if the key is an employee's number, then the numerical sequence can be used as the 'organizing key' (19, 52).

A file can be organized on any attribute, or set of attributes, and a variety of derivative files may be produced. In HISEP, the primary file will consist of individual entities, with each entity described by the characteristics listed for the patient data section of the data base recommended in Chapter III. From this primary file, much of the information required in other files can be derived (e.g., demographic data). When an individual moves or is born into the planning area, his patient data file and the corresponding derived attributes will be immediately entered into HISEP. The 'organizing key' for this file and the other HISEP files can be chosen at the time of system development to best serve the specifically proposed use of the system (19, 41, 52).

Since this primary file will ideally consist of every individual served by the health system within the planning area, an adequate identifier, or record header is of concern. There is currently much discussion as to how an individual can be uniquely identified (4, 10, 12). The minimum length identifier should consist of the following:

(a) Name at birth

(b) Birth date

(c) Social security number (if applicable)

(d) Sex

The remaining basic files mentioned in Chapter III should be structured in the same serial format as the primary file. Once these files have been compiled, all of the remaining files that are needed can be derived from them, utilizing the structuring techniques mentioned earlier. (e.g., inverted files).

To stipulate the details of the data base structure beyond this general approach is not practical. A primary difficulty in designing any data base is that it is not known a priori how to stratify the various data files to best meet unknown demands by the system users. The more specific design of the data base should come when a more specific plan for utilizing the data is agreed upon among the health planners of the particular planning area. However, something more can be said about the software needs

to maintain the maximum amount of utility and flexibility for HISEP.

4:2 Data Linkage and Integration

To a large degree, it has been found that great volumes of segregated data do not satisfy the needs of the health planner (8, 11). The data must have more than just the current available two-dimensional aspect. The planners need complete and uniform information in many dimensions to explain the underlying circumstances which have given rise to a particular health problem (52). For example, the characteristics of the population should be compared simultaneously to chronic conditions, utilization of facilities, and severity of conditions. Through the proper integration of data, valid cause and effect relationships can be established (30). To assure this proper integration of data, the original basic files for HISEP were designed to provide sufficient disaggregation of data to permit the study and modification of variables affecting the supply and demand of health resources (52).

Quite often the information needed for planning is not available at the unit record level. Several methods are currently being developed to integrate data at the aggregate level. Analytical techniques, such as least squares regression analysis, make it possible to estimate the values of dependent variables on the basis of information that is

available. Computer mapping is another good technique for obtaining estimates on the aggregate level through the integration of data that is available (30, 35).

The planning process obviously requires a wide range of data. However, the requirements are for aggregate data for specific small geographic areas. They are not for information on identifiable individuals, but rather for information on all individuals with certain specified characteristics (30). (The Patient Data file will be the main source of data for this type of derived files.)

There are several factors that should be considered when integrating data from various sources into a single file. Data required for planning must necessarily come from many sources and in several forms, and problems of compatibility and comparability may be involved. These problems must be thoroughly considered in the system design (8). Also, resources may be most effective when they are pooled. Several planning agencies and other groups may possibly have similar information needs. Every effort should be made to implement a common system (52). Finally, attention should be given to structure and relationships to assure that the data integration will enhance dialogue between health planners and statisticians (30).

Some examples of output data which utilizes the integration of various basic data files include (30):

(a) How services relate to each other within an institution and among other institutions;

(b) How personnel are utilized, rather than merely a head count;

(c) Why physicians admit patients to various hospitals;

(d) Use of demographic data to produce population projections and population distribution variables by area and by charactistics;

(e) The rates of hospitalization from sub-groups within the population yield the relationships of facilities with each other and comparison rates of hospitalization for various diagnoses.

Certainly many other examples are possible. The principle used in these examples is that of merging the basic files to create new files for outputting whatever type of information the planner may desire.

4:3 Requirements on the System Software

In order for HISEP to satisfy all of the requirements set forth previously, there are several criteria that must be met by the software package. Chapter II listed the more vital requirements. Within each requirement are some very important specific operating characteristics which are described below in terms of their technical nature (35, 38, 41, 52).

4:3:1 Input/Output

HISEP must be able to accept and output data in a variety of formats. The key to effective handling of external data is the separation of record description from actual program code throughout the routines in the system. Since the data may be arranged in the record in a variety of ways, it is the function of the format description to communicate information about the data format. As described earlier, files may be either very simple in form or a more complicated form created by merging the simple files. These simple files and intermixed files may have a hierarchial structure of many levels. HISEP must distinguish these hierarchial details for effective processing of the input and output data. Another problem area with input/output concerns data item representation. The conversion of data to the proper coding for internal storage rather than storing actual text yields great savings in storage requirements. HISEP should utilize the 8th Revision of the International Classification of Diseases, and other well-known coding schemes to store the data for maximum storage and retrieval efficiency. Use of these standard coding schemes will also aid in the implementation of data provided from other agencies such as the U.S. Census Bureau.

4:3:2 Large Volume Data Handling

The capability to handle large volumes of data is another requirement that must be met. Whenever insufficient

capacity exists to handle a given amount of data, a solution must be sought that involves partitioning and recombining the data through some process that requires larger processing time than that would normally be required. HISEP should have the means available for circumventing process and secondary storage device limitations through imbedded capabilities for handling partitioned and multidevice data sets. Substantial savings can be obtained if HISEP has the capability of accepting more than one data set, simultaneously or in sequence, as input to a process. An equally desirable feature is the ability to produce several outputs concurrently, each output having a different content and structure from the others.

4:3:3 Handling Individual Data Items

HISEP must be able to perform a wide range of processes that concern both individual and sets of data items. Specific capabilities needed to handle individual data values include the following functions:

4:3:3:1 Data Transformations. Data transformations provide the capability to combine, convert, replace, and create new item values as a function of input data. While most languages provide for the more basic data transformations such as logarithm, trigonometric functions, etc., HISEP will require more advanced forms of manipulation. These

advanced capabilities should include complete freedom of algorithmic expression, reference to data items by name or by position, reference to constituents by data items (e.g., bytes in an alphanumeric string) and to external, preprogrammed functions. This data transformation capability may be implemented by a procedural transformation language or a data specification language, which includes the means for defining data items as a function of other data.

4:3:3:2 Data Validation. One of the most serious problems of processing large volumes of survey data is the validation of input. Some error checking is implicit in data item definition and in the contingency checking process it concerns. For example, any value not allowed in the field is manifestly erroneous. The system must provide means for specifying the type of exception action to be taken in the case of input error detection. Methods by which users can perform selective examination and replacement of erroneous values should be within the system's processing capability.

4:3:3:3 Text Processing. The requirement for text processing will arise frequently (e.g., handling records containing uncontrolled address information). The capabilities to handle natural text are provided through special file formats and text processing programs (e.g., permutation indexing).

4:3:3:4 Entity Classification. A special kind of transformation on the input data involves classification of entity records according to user-specified criteria. A requirement for entity classification arises in a variety of contexts, ranging from stratification of data for file summaries, to multivariate classification problems. The membership of an entity in a class is denoted by assigning the value of the class name to a data item reserved for that purpose in the entity record.

4:3:3:5 Dictionary Lookup. Dictionary lookup is often employed for associating one set of data with another when prior reordering of the two sets into a compatible sequence is inconvenient. A similar requirement arises in serial file processing where a set of records is to be matched, on a many-to-one basis, with another set of records in the same file. A good example of the latter arises in converting address information to corresponding geographic codes.

4:3:3:6 Entry Value Modification. Quite often a small subset of a particular file will need updating or modifying. Some means for specifying the correction of individual values in a data set should be provided by the system, in addition to the capability for the regenerating an entire file.

4:3:4 Handling Data Sets

In order to handle sets of data, HISEP's capabilities should include the seven following functions. It should be noted that these capabilities are in many ways dependent on the capability requirements for individual data items.

4:3:4:1 Entity Selection. HISEP should have facilities for the selection and retrieval of appropriate subsets of data items by name and for the selection of entries on the basis of enumeration or some other logical criteris.

4:3:4:2 Organizing Data Sets. The system should provide facilities that permit data sets to be sorted on specified data fields in ascending or descending order, treating key fields as either numeric or alphabetic. Data set index tables may be generated to correspond to file organizing keys, relating specific areas of files to specific storage media.

4:3:4:3 Combining Data Sets. The system should have capabilities to combine data from one or more files according to a variety of criteria. The result of a data file merge process would consist of all entities that occur in the source files, with values of only those properties that these entities have in common. A match-merge process

results in a file consisting only of those entities that are represented in each file.

4:3:4:4 Data Set Reduction. Date set reduction is another important software function needed by HISEP. A file summary capability represents the simplest type of data set reduction. In this process, the set of values associated with a particular property of entities that possess the same value for a key attribute is reduced to a single value. The result may be a mean, median, etc. A cross tabulation occurs when the source values are being summarized into several result values. An accumulation of geographic areas into separate totals, each corresponding to a certain demographic code, is an example of a cross tabulated summary.

4:3:4:5 Processing of Hierarchically Organized Data. Entities in HISEP will probably be linked through a variety of hierarchical relationships. It may be necessary to associate several types of data records, varying in format and composition, to make the data more compact and less redundant. Thus, a process to handle hierarchically organized data is necessary.

4:3:4:6 File Inversion and Dictionary Construction. Inverted files and dictionaries are used for efficient retrieval of records from a data base in which a complete serial search through the file would be time consuming and

expensive. HISEP should have the means of constructing such indexes for arbitrary properties defined as keys to the remaining record contents.

4:3:4:7 Matrix Operations on Serial Data Sets. Calculation of regression and other types of estimates useful in demographic projection can be represented more efficiently as a matrix operation on a file containing initial predictor values. This type of file is usually treated as a matrix and is then multiplied by the estimating coefficients. Therefore, HISEP needs matrix manipulation capability.

4:3:5 Report Formatting

The ability to output data as desired also puts requirements on the HISEP software. The system must be able to generate reports with both standardized formats and user-specified formats. Formatting capabilities should provide for page headings, arbitrary arrangement of information, column and row summaries, intermixing and insertion of data into text, and variations in page format as a function of data content. The display of graphic data is also important to HISEP. The software for this function will quite often be provided by the manufacturer of the display hardware.

4:3:6 Query Processing

System capabilities to optimize efficient data retrieval and output the information may be collectively

referred to as query processing. Query processing is one of the major functions of an information system. To minimize record-research time, queried data bases are stored on direct-access devices for on-line query.

4:3:7 Data Analysis

For data analysis capabilities, HISEP should provide for production of percentaged frequency tables; statistical tests such as chi-square and gamma; dispersion and linear correlation; and multivariate analysis. It is expected that more sophisticated analysis techniques would become part of the system as it matures.

4:3:8 Executive

The executive component of HISEP functions is the internal "supervisor." It performs such functions as linking system components, controlling internal operations of the hardware and software, recording internal processes, etc. Six major activities of process management are:

(a) User/system communication;

(b) job specification retention;

(c) linkage to user-supplied special purpose programs;

(d) procedural control of computing tasks;

(e) process monitoring and recording; and

(f) data set assess and release protection, control, and auditing.

4:3:9 Updating

The problem of updating and purging data in a large scale system such as HISEP is of a different character than the other software problems. The problem arises when one permits individual disparate pieces of data to enter, leave, and change the data base. The system must be aware constantly of the different files to which a single datum may belong or be relevant. At the extreme, a data change may become appropriate to files other than the ones which it originally affected, or it may lose its appropriateness. Thus, every fact must be tagged to the files to which it belongs, and algorithms need to specify what other files should be amended. Several estimates for each basic file must be made. These include: the appropriate rate of changes, the maximum to which a live file will grow, when back-up data can be discarded, the man-power needed to transform data to machine-usable form, and what economics in clerical operations can be achieved by designing appropriate consoles and other input devices such as computer readable record forms.

4:4 General Purpose Software Available

There have been developed over the past few years several software packages for use in health information systems. A few of these packages which may be useful with HISEP are briefly described below.

4:4:1 New Haven Health Information System

The U.S. Census Bureau has developed some software for use in the New Haven Health Information System. The requirement for linking diverse data sources was solved by an address matching system called ADMATCH. ADMATCH used only the street address on each record, rather than a coded geographic identifier, to permit the matching and merging of local data with Census data. DIME is a geographic base file which yields precise geographic location of a data entity. GRIDS is a program package which provides three mapping options within a grid pattern for computer mapping. In addition to the software packages, some methodology was developed and implemented, thereby making their analysis techniques available. In general, one takes quantitative statistical methods and applies them to a data set. With the empirical results, the neighborhoods may be classed into a minimum number of types and then they are ranked within each type. Correlational and factor analysis are then used to form typological indicators of the population. Cluster analysis and discriminatorial analysis examine the correlation on a less aggregated scale to determine if the relationships are genuine (45, 49, 50).

4:4:2 MIT Projects

The Urban Information Systems Task Force of the M.I.T. Urban Systems Summer Study, The Health Information

Systems Project of the M.I.T. Harvard Joint Center for Urban studies, and the ADMINS project in the Center for International Studies, M.I.T., have all produced software packages which might contribute to the goals of HISEP. ADMINS is of special interest. ADMINS is a "tool kit" of computer programs for the designer and developer of information systems. ADMINS Mark III is a particular information system which is currently functioning at M.I.T. for the particular purpose of social science survey analysis. ADMINS Mark V is now under development and it should solve some of the problems of how to build software systems for information systems for any one of a variety of conceivable purposes (35).

4:4:3 FFS

FFS (Formatted File System) was developed for the Defense Intelligence Agency. Files may be on tape or disc and may be defined, maintained, and queried. Cross-index and table-of-contents capabilities are provided via a definition process and maintenance on tape or disc. A unique key identifier field is required in each record for identification purposes. Some geographical and statistical capabilities are also provided (52).

4:4:4 GPSP

GPSP (General Purpose Software Programs) is a package of ten programs for the IBM 1401 or 1460. The programs

provide for: (a) stratified record tabulation into a standard tabular format, (b) file summary, (c) record selection and recording, (d) record reformatting, and (e) file updating. This package provides many of the fundamental file processing requirements that we-e described earlier in this chapter (52).

4:4:5 SPAN

The SPAN (Statistical Processing and Analysis) system was developed to deal with standard file manipulation and data reduction, matrix operations, multivariate analysis processes, geographic data functions, and report generation. The system includes the MIDAS (Mixed Data Structures) Processor and Report Generator. MIDAS permits the user to accept inputs of records that are not uniform, and from them, construct files arrayed hierarchically according to an organizing scheme expressed at execution time. The same set of inputs therefore may be used, with different organization and solution keys, for different purposes. Report formats can be defined by means of heading and print operators, or through a picture method in which the user inputs a format scheme that employs the actual print positions to be used (30, 52).

4:5 Privacy Safeguards

The protective features and safeguards available for protecting the sensitive data in HISEP are much more reliable

than the current trend of thinking indicates (18). The problem of unauthorized access has generally been solved by using a subset of lock-out procedures characterized as file protection devices. Elaborate encoding and decoding has also been employed and can be viewed as a different subset of lock-outs. However, this method is of secondary importance and its discussion is not included here (35).

File protection is a set of procedures assuming that access to a file will be granted only to those with explicit authority to receive or record information. This authority can be made subject to a variety of factors and/or conditions and can be modified whenever circumstances dictate. Because of the discrete and segmented organization of computers, access can be granted to subsets of data in a file or any combination of subsets. Access to the record can be limited to those categories of data which are relevant, and the others, like individual identifying information, can be blocked. The file protection system used on the Project MAC system at M.I.T. is probably the best example of file protection currently available. The system has been in existence for over four years, has hundreds of users, all of whom maintain separate, private files within the computing facility. The permission to use a particular file can be restricted in a variety of ways and evoked by the grantor at any time. A detailed description of this and

and other useful protection procedures can be found in Computers and Privacy, A Survey, by Lance Hoffman.

The difficulty with privacy in large scale systems is not its technical feasibility. The dilemma is to specify the standards and guidelines for how protection shall be implemented, who shall be protected and under what conditions. The major issues concerning privacy in HISEP are the exactment of such standards and guidelines in the form of legislation rather than the feasibility of the technical implementation of protective measures (35).

4:6 Documentation

In a large scale system such as HISEP, the enormous volume of data and its continually changing nature impose a major requirement upon the system to locate and identify elements of relatively permanent data and to keep track of transient data. Data documentation is necessary to provide this ability. Basically, data documentation is concerned with "data about data" that must be processed continuously so that system users will be abreast of the growing and changing contents of the data base as well as where and how they may be accessed. When a data base is limited in size and scope and changes slowly, documentation can be accommodated readily on a few typewritten pages. However, with a large scale system, the problems of systematic data documentation became critical (52).

The planning agency must implement and operate the data base documentation system before data collection starts. The documentation methodology involves: (a) organizing the data base, including classifying, naming, and relating its components; (b) maintaining descriptions of data base contents; and (c) showing the data base contents to the user systematically and currently (52).

The basic goal of documentation, therefore, is to account for all significant objects in the data base. The several classes of data should be distinguished: data items, as the smallest meaningful units of information within entity records; data lists (record types); files and matrices; tape reels and other types of physical storage media; computer programs; and maps. All data must be documented and the contents of all tapes, discs, and card drawers must be described. Reports should be generated periodically to keep the user updated. These reports should include a "data dictionary," in which data items are indexed by subject-matter and permuted descriptor sets. Through the dictionary, the user can locate particular data items in the various records and files. A "tape file dictionary" would be an example of a comprehensive guide to the contents of some specific physical storage device. Obviously, some of the reports need to be produced more frequently than others (6, 20, 52).

CHAPTER V

SYSTEM OUTPUT AND REPORTS

5:1 General System Output

Earlier chapters have discussed the need for a wide variety of output. Not only must HISEP produce periodic preformatted reports, but it must also have a large degree of flexibility to produce output in formats that are established by user requirements. An overall viewpoint of the periodic reports that should be furnished by HISEP is given by the following six general areas (4, 8, 41).

(a) Routine and special status reports are needed to keep the planners informed of the current status of the various facilities serving the field of health. Such information as the level of occupancy for hospital beds, admissions for various diseases, births, and deaths would be included in these reports.

(b) Effectiveness status monitoring would inform the planner of the effect a particular health program is having on the population being served, and also the effect the program is having in other health areas.

(c) Community status monitoring would maintain a vigil on the general health of the community and inform

the planner of a breakdown in health standards in a particular area.

(d) Descriptive profiles would contain information concerning the characteristics of various segments of the population.

(e) Analysis of future projects are necessary to keep the planner current with how expansion plans and new programs might directly or indirectly effect the health status of the community.

(f) Delinquency and follow-up reporting would serve to keep track of those patients who are delinquent in important visits to their physician, and assess the value a particular action had toward solving a health problem. Although this type of report is not usually included in other health information systems for planning, it is necessary to meet the goals set forth for HISEP. The original health problems must be reassessed to determine if the action taken was indeed the proper solution, and, if not, to determine what the proper solution might be.

In general, the purpose of all of the reporting done by HISEP should reflect the primary goal of determining trends affecting the health system, thus enabling the planner to predict the future with some degree of success and reliability.

5:2 Routine Periodic Reporting

Many agencies have spent much time in determining

exactly what information should be included in the reports generated for health planning (2, 4, 7, 12, 17, 21, 41, 47). Although the opinions vary to some degree and each planning agency will have its own personal preferences, a general outline can be stated that will include all of the vital information needed by the various agencies. Whenever special types of reports are needed by the health planners, these can be easily generated, but they should not be considered as a part of the normal system reporting. As far as specific formatting of reports is concerned, these should be set up to satisfy the tastes of the individual users. The following is an outline for the information that probably will be required in the periodic reports of HISEP.

5:2:1 Summary Information

A. Bed status summary

Status listed by institution class and institution

(1) Current capacity

(2) Current occupancy

(3) Licensed beds

(4) Expansion plans

B. All hospital summary

(1) Description of population being served

(2) Patient days by disease classification

(3) Patient characteristics

(4) Future Projections

C. Travel time analysis

(1) Time from each census tract to the nearest hospital

(2) Analysis by service offered from the census tract to the service

D. High-risk neighborhood identification

(1) List by disease

(2) Population characteristics

(3) Nearby services available

E. Search for large variances for a particular service among the different institutions offering the service.

F. Use of health indicators

Information given in terms of means, ratios, medians growth rates, and proportions.

(1) Social Organization

a. Marital

b. Children age

c. Children under 18 at home

d. Racial, ethnic, and age backgrounds

(2) Socioeconomic Strata

a. Income (Median family income)

b. Education

c. Occupation

d. Property ownership

e. Housing adequacy

(3) Fertility rate

(4) Infant Mortality and Injury

(5) Migration characteristics

(6) Health status overall

5:2:2 High Risk File

A. Check for proper number of revisits

B. Check for concentrated areas of high risk disease

C. Check for concentrated areas of each high risk disease

D. Denote priority level for critical nature of each patient in the data file

5:2:3 Institutional Output:

A. Summary Information

(1) Admissions

(2) Births

(3) Daily census/percentage of bed use

(4) Discharges

(5) Deaths

B. Service Utilization

Information matches bed licensing categories against service categories, thus yielding the extent to which each hospital service is utilized.

C. Summary Information on Discharges

(1) Discharge Status

a. Dead/alive

b. Diseased/healthy

c. Disabled/fit

d. Dissatisfied/satisfied

e. Uncomfortable/comfortable

(2) Average length of stay

(3) Consultations

(4) Autopsy

(5) Continuing care

D. Summary Information on Newborn

(1) Discharge status

(2) Continuing care scheduling

E. Operation and treatment summary

List by type of operation

(1) Breakout into length of stay

(2) Average length of stay

(3) Breakout into age and sex

F. Diagnoses Summary

(1) List by diagnosis

(2) Age and sex breakout

G. Diagnoses Index

List by coded disease for all patients under each category.

(1) Secondary diseases

(2) Surgical procedures

(3) Attending physician

(4) Consulting physicians

(5) Discharge date

(6) Discharge status

(7) Sex

(8) Age

H. Physician Index

List by physician

(1) Break-up by service rendered

a. Attending

b. Surgeon

c. Admitting

d. Consulting

(2) Patient information under each service

a. Discharge date and status

b. Sex and age

5:2:4 Facility Profile

A. Facility status

B. Inspection status

C. Occupancy status

D. Full time personnel

E. Nursing personnel

5:3 Computer Mapping

One method for distilling the data in HISEP and presenting it simply and forcefully is computer mapping. When the data has demographic or geographic aspects, the information involved can be more easily grasped if it is in the form of a map. Computer mapping serves a number of vital functions. Where a spatial dimension is involved, like geographic distribution, it visually displays the allocation of the data over the area and immediately highlights the extremes in the data picture. More important, it gives the health planner a powerful tool to assist him in swiftly communicating complex situations to non-technical associates (5, 16).

Given the proper instructions, the computer and automatic plotter can draw any map that can be drawn by hand, and maps that no human would attempt to daaw. Multiple colors, varying thicknesses, and varying paper widths are all available to make the features of the map more recognizable to the viewer (5).

Four major elements are involved in the computer mapping process: (a) selecting and specifying the data to be mapped; (b) linking the data file to a geographic base file; (c) manipulating and organizing the data to fit available programs and equipment; and (d) deciding on the cartographic features of the map to emphasize the significance of an important set of data (46). Most of the human

effort spent on computer mapping is consumed by coding the data geographically and manipulating the data to fit the available equipment and programs (35).

Whenever the geographic source of data is reported along with the data, the data is geocoded. The geocode may refer to a point or an area. A point might be a street address, where an area might refer to a block, a precinct, a city, etc. The code is then transformed into the coordinates of the map where it is associated with the data to be mapped. This transformation of points and areas into map coordinates can be accomplished through the use of an address matching program and a geographic base file describing the relationship between addresses and the geocodes (35). The transformation problem was met by the U.S. Census Use Study and the development of ADMATCH and DIME resulted. The DIME file and the ADMATCH routines which were described earlier are the basis for SYMAP, the best known, most comprehensive, and most widely used computer mapping program currently available (46).

SYMAP (Synographic Computer Mapping) enables three basic types of maps to be produced through three primary options: the conformal option, where the areas shaded are approximate representations of the polygonal geographic areas that they represent; the contour option, where the data values are assigned to particular points and the program shades contours by establishing equal-width intervals

representing the range of values between each pair of points; and the proximal option, where the data values are again assigned to particular points and the program shades the respective values from each point to an imaginary line equidistant between each pair of adjacent points. In addition, SYMAP has a large number of statistical-support-options which permit calculations of means, standard directions, histograms, and percentile groups (46). The maps in the appendix illustrate the contour option and the proximal option.

SYMAP utilizes a standard high-speed computer printer which prints typewriter-like characters, on standard computer printout paper. Where the desired width of the map exceeds the width of the paper, the printing is done in strips. Shading that represents aggregated data values such as the number of housing units per block can be shown by an overprint of characters.

Line-drawn maps can be generated by a number of plotting devices, namely pen plotters and cathode ray tubes. The pen plotters actually draw lines to represent spatial arrangements described by the data. The cathode ray tubes project images onto photosensitive paper or films. Use of these devices affords considerable precision but usually at a considerably higher cost than character printing (46).

The extra time and effort involved in massaging the data to produce maps rather than tabular listings should be

weighted against the need to communicate the data. Material that is to be viewed by many people with various backgrounds will certainly justify the expenditure on graphic presentation. Computer-assisted maps for HISEP will enable the planners to assimilate date quickly and easily, and, most importantly, allow them to convey to others the reasons behind their decisions (5).

5:4 Recommended Displays

For the display of all types of information generated by HISEP, there are three basic choices. The three are: (a) cathode ray tubes (crt's); (b) impact printers; and (c) inkline plots. While all three are classified as display devices, the use of each one is optimized for entirely different purposes (35).

The crt permits the rapid, visual display of material and the easy editing of the information, whether the information is a command, data being entered into the system, or data being received.

Impact printers (listers) are the most conventional means of providing hard copy output. These include listers and typewriters. A lister permits typing an entire line of characters with each stroke, while the typewriter types only a single character with each stroke. Impact printers have significance for HISEP, not only in the conventional means of employment, listing on hard copy, but also as a

means of economically satisfying a variety of mapping needs of the health planner (35, 46).

The third category of displays is inkline plots. It permits the construction of a hard copy display with greater detailed accuracy and flexibility than can be accomplished with a lister, and it is therefore preferable where a high degree of graphic accuracy is necessary. Plotters, however, are relatively expensive, complex, and are controlled by elaborate programming instructions. The construction of maps is less constricted, but it is considerably more lengthy and difficult than the SYMAP process (46).

A display of some kind is essential in order to communicate HISEP's output. There is no question but that at the minimum there must be a lister. The real question is whether or not to provide alphanumeric crt and/or inkline plotters. There is not much question of the cost benefit of the crt. Not only is it vital for user/crt interaction, but it also may be used in lieu of the typewriter or the printer if no hard copy is required. Or, with an added investment, a crt copier may be provided. The inkline plotter is of questionable value. For the time being, SYMAP provides an adequate mapping capability considering the high cost of the plotter. The degree of mapping precision and flexibility which the inkline plotter provides, over and above that of SYMAP, does not justify its use, given the added cost (35).

CHAPTER VI

BACK-UP PROCEDURES

6:1 Necessity of a System Back-up Configuration

The overall system design for HISEP must provide for the possibility of a system failure that would purge all or part of the system's software and/or data base. Back-up requirements are dictated by the large size and the continual changes of the data base files in HISEP. There are many ways this file back-up procedure can be accomplished. The procedure chosen for HISEP, should allow for a maximum amount of ease in revision of the software packages, and also a maximum amount of protection, since changes are often not documented concurrently with a programming or data base change.

6:2 Recommendations

If a new or revised program is to be entered into this system, it must first be proven that it will mesh properly with all of the other operating routines. After it has been approved, it should be read onto a magnetic tape. Then the new or revised program can be read into the system on-line software. The tape can then be converted into an object deck and placed in storage at a location removed

from the central computing facility. In the case of a program revision, it is usually recommended that the program being replaced is retained in storage for 120 days.

For the data base, changes will occur so often that the back-up file should remain on tape. However, there should be no updating of files during the normal day-to-day operation. Instead, the new files should be constructed off-line on tape and then read into the system at night, or other times when the system is not operating. The new data files that are now on tape can then be stored at a remote location. The old tapes should be retained for a period covering at least ten file updates.

APPENDIX A

DEMOGRAPHIC DATA

The following list of health and health related information was compiled so that health planners and community health researchers might select the particular types of data they see as "ideally" needed in a system such as HISEP. Accompanying facts on possible sources of these types of information was gathered to illustrate the extent to which data identified as needed are actually available in usable form. No claim is made to the completeness that would be expected from a compendium of sources material; rather the list simply reflects the possible avenues the search for particular types of information might take.

The types of health and health related data are grouped into the following categories:

1. Demographic data
2. Health statistics data
3. Health manpower and resources data
4. Patient data

TABLE A-1. DEMOGRAPHIC DATA

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Age and Sex:		
Under 1 5 year intervals 1-85 Over 85	U.S. Census	Decennial
School age census	City and County School Censuses and Enrollment data	Annual
Age 18	Local draft boards	Continuing
Adults 18 and over	Local election commissions	Annual
Migration:		
Intra-State	State Department of Commerce State Employment Agency	Continuing Continuing
Inter-State	U.S. Census	
Population Projections:		
Annual estimated change by central and suburban Counties	U.S. Census	Annual
Annual estimated change by city block	Metropolitan Area Planning Council*	Annual
Fertility ratios, trends	U.S. Census	Decennial

* Almost all large communities have a local planning council which takes limited annual surveys.

TABLE A-1. DEMOGRAPHIC DATA (Continued)

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Ethnicity:		
Foreign born; foreign born parentage; Irish stock; Italian stock; non-white	U.S. Census	Decennial
Race of school children 1-12 grades	School racial census	Annual
Social Characteristics:		
Economic status/ family income	U.S. Census	Decennial
Household size and composition	U.S. Census Metropolitan Area Planning Council	Decennial Annual
Employment status and occupation	U.S. Census State Employment Agency U.S. Department of Health Education and Welfare (HEW)	Decennial Continuing Continuing
Educational attainment	U.S. Census	Decennial

TABLE A-1. DEMOGRAPHIC DATA (Continued)

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Analyses of social and economic population measures (home ownership, working mothers, non-family population, etc.)	U.S. Census	Decennial
Social and economic groupings within local population	U.S. Census United Community Fund	Decennial Annual
Housing by construction date and quality	U.S. Census of Housing Metropolitan Area Planning Council	Decennial Annual
Reported violations of Sanitary Code	Local health departments	Continuing
Economically troubled households by income level, occupation, and family size	HEW AID to Dependent Children Other welfare agencies	Continuing
Urban Travel:		
Intra-city	City transit studies	Continuing
Travel to work	City transit studies	Continuing
Method of travel to work	U.S. Census	Decennial

TABLE A-1. DEMOGRAPHIC DATA (Continued)

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Social Pathology:		
Juvenile Court appearances by community of residence	State Department of Probation	Continuing
Committed drug offenders	State Department of Correction	Continuing
Neglected Children	State Department of Public Welfare	Continuing
	HEW	Continuing
	Society for the prevention of cruelty to children	Continuing
Economic Activity:		
Economic base data	State Department of Commerce	Continuing
Employment and labor turnover by major industry	State Employment Agency	Continuing
Employment trends	State Employment Agency	Continuing
	U.S. Bureau of Labor Statistics	Continuing
Unemployment levels	U.S. Bureau of Labor Statistics	Continuing
	State Employment Agency	
Unemployment levels	U.S. Bureau of Labor Statistics	Continuing
	State Employment Agency	Continuing
	HEW	Continuing

TABLE A-1. DEMOGRAPHIC DATA (Continued)

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Vital Statistics:		
Births by sex, race, weight, residence, birth injury	County Recorder's Office	Continuing
	State Department of Health, Division of Vital Statistics	Continuing
Deaths by age, sex, diagnosis and residence	County Recorder's Office	Continuing
	State Department of Health, Division of Vital Statistics	
Marriages, divorces	U.S. Department of Commerce	Continuing
	U.S. Census	Decennial
	City Licensing Agency	Continuing

There are other agencies which may be able to contribute to demographic data. These include: The State Department of Mental Health will take various types of surveys in the lower income level areas of the cities. The local voluntary health agencies keep data in connection with high-risk areas. City, county, and regional planning commissions keep various types of sample data on housing, transportation, and employment levels.

TABLE A-2. HEALTH STATISTICS

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Disability:		
Number and percent persons with chronic conditions, restricted activity days, bad disability days, work loss days	Health Interview Survey National Center for Health Statistics	Five Years
Morbidity:		
Kinds of disease in children from low income areas	Local OEO Agencies State Department of Health	Continuing Continuing
Incidence of acute conditions by age, residence	State Department of Health Tuberculosis Society United Community Service Local Health Agencies American Hospital Association Social Security Administration (Medicare)	Continuing Continuing Continuing Continuing Continuing Continuing
Disability in the Workmen's Compensation by age, nature and length of injury	State Industrial Accident Board	Continuing

TABLE A-2. HEALTH STATISTICS (Continued)

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Courses of rejection from military service	Selective Service	Continuing
Incidence of reportable disease (T.B., V.D., etc.)	State Department of Health HEW	Continuing Continuing
Disabling highway Accidents	Registry of Motor Vehicles, Division of Accident Records	Continuing
Handicapped and emotionally disturbed children	State Department of Education Local and State aid societies	Continuing
Sickness among school children	Local Schools	Continuing
Number, residence, characteristics and discharge diagnosis of short-term hospital patients	Local Hospitals Hospital planning agencies Blue Cross	Continuing Continuing Continuing
In-patients receiving (Medicare (as above)	Social Security Administration	Continuing

TABLE A-2. HEALTH STATISTICS (Continued)

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Utilization of hospitals; number of admissions, discharges, births, deaths, census	State Department of Health American Hospital Association Local health facilities	Continuing Continuing Continuing
Services paid for under public assistance and Medicaid programs	State Department of Public Welfare	Continuing
Blue Shield patients; characteristics, type and volume of service rendered	Blue Shield	Continuing
Number and characteristics of mental hospital patients; age, sex, diagnostic group, and utilization data	State Department of Health	Continuing
Physician Care:		
General out-patient care; volume, age, sex, residence	Individual hospitals and clinics	Continuing

TABLE A-2. HEALTH STATISTICS (Continued)

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Other Health and Health-related Services:		
Number and kinds of Services provided		
Home Health Services	State Department of Health	Semi-annual
Laboratory Services	Blue Cross	Continuing
	State Department of Health	Continuing
	Social Security Administration	Continuing
Immunizations	State Department of Health	Continuing
Physical and psychiatric rehabilitation services	State Department of Vocational Rehabilitation	Continuing
	State Center for the Blind	Continuing
Classes for handicapped children	State Department of Education	Continuing

Other sources of information on health statistics in general are:

HEW, Division of Vital Statistics, Registration Unit
National Center for Health Statistics, HEW
State Division of Communicable Diseases

TABLE A-3. HEALTH MANPOWER AND RESOURCES

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Health Facilities and Agencies:		
Hospitals for each institution: control, type of service, beds, certification status, etc.	Joint Commission on Accreditation of Hospitals State Department of Health	Continuing
Extended care facilities for each institution: control, beds, certification status, affiliations	State Department of Health	Annual
Nursing Homes - (Same as Above)	State Department of Health	Annual
Visiting nurse services: location, staffing, functions	State Department of Health	Annual
Manpower Resources:		
Physicians and osteopaths: number in active practice, by current office location, age, specialty, hospital affiliation	American Hospital Association	Continuing

TABLE A-3. HEALTH MANPOWER AND RESOURCES (Continued)

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Dentists: number in active practice, age, office location	Board of Registration	Annual
Nurses: registered and licensed practical		
Number employed, by age and marital status, in hospitals, health departments, schools, industry, other	Board of Registration	Annual
Public Health	State Department of Health	Biennial
Physical therapists	Board of Registration	Biennial
Technicians - X-ray and Laboratory	U.S. Bureau of Labor Statistics	Continuing

Some other sources of Health resources and manpower data are:

American Osteopathic Association
HEW
Lutheran Hospital Association of America

TABLE A-3. HEALTH MANPOWER AND RESOURCES (Continued)

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Some other sources of Health resources and manpower data (Continued)		
	American Association of Homes for the Aging Board of Hospitals and Homes of the Methodist Church Mental Health Federation, Inc. Regional and Local Hospital Planning Agencies American Medical Association	

TABLE A-4. PATIENT DATA

TYPE OF DATA	SOURCE OF DATA	DATA COLLECTION INTERVAL
Identity: by age, sex, race, residence, marital status	Individual hospitals	Continuing
Type of Payment:		
Analysis of utilization data according to source of payment:		
Hospitals	Blue Cross	Continuing
	Individual Hospitals	Continuing
	American Hospital Association	Continuing
No charge service utilization	State Department of Health	Continuing
	Social Security Administration	Continuing
	State Department of Public Welfare	Continuing
Attitudes Toward Health Services:		
Knowledge of how to get needed health care, according to family characteristics	State Department of Education	Annual
	State Department of Health	Continuing
	State Department of Public Welfare	Continuing

APPENDIX B

SYMAP COMPUTER MAPS WITH A STANDARD HIGHSPEED PRINTER

The accompanying two maps illustrate the ability for SYMAP to produce computer maps with a standard highspeed printer. MAP 1 is configured with the contour option and MAP 2 with the grid option (a variation of the proximal option).

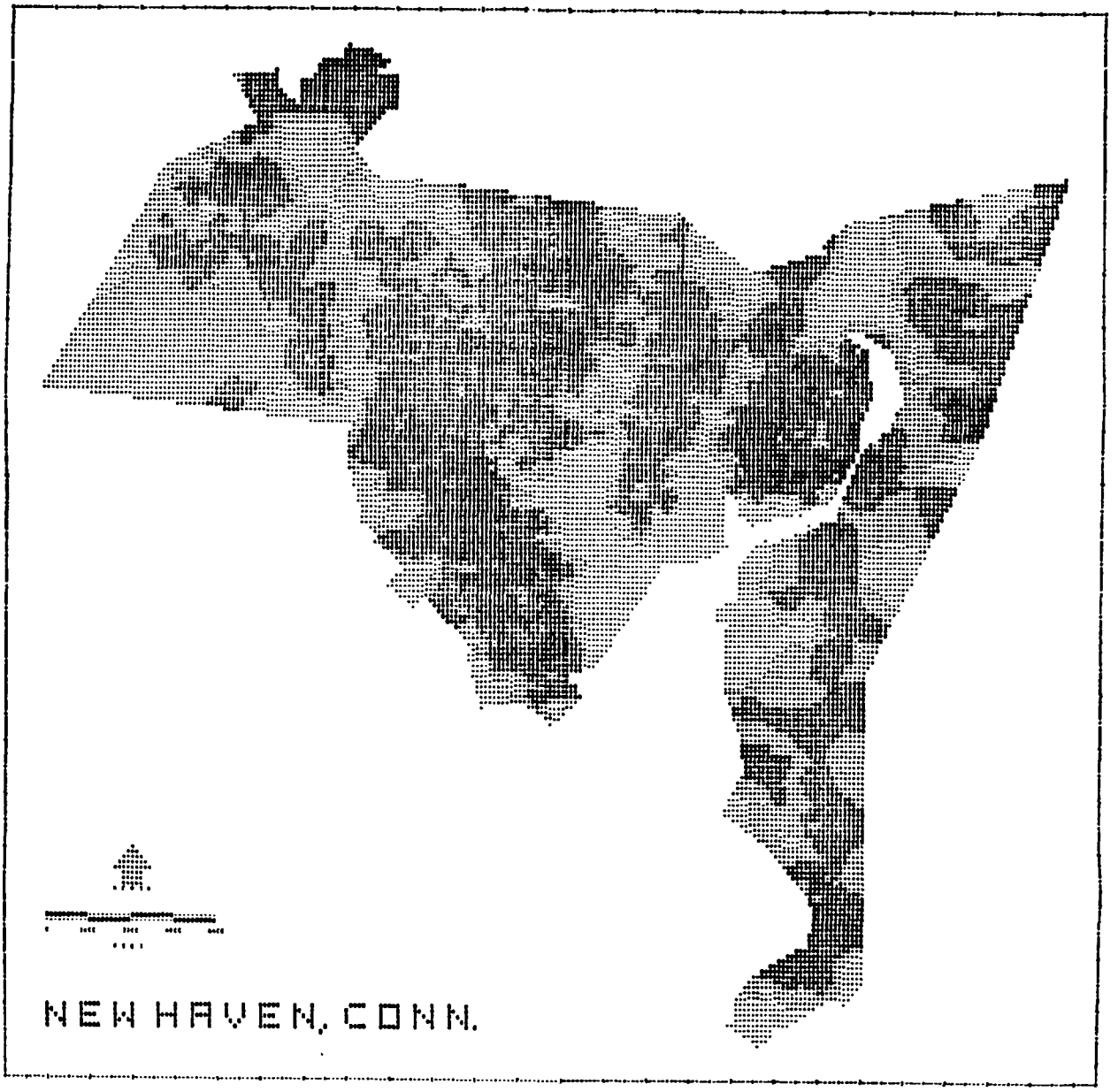


Figure 1. Contour Option Map

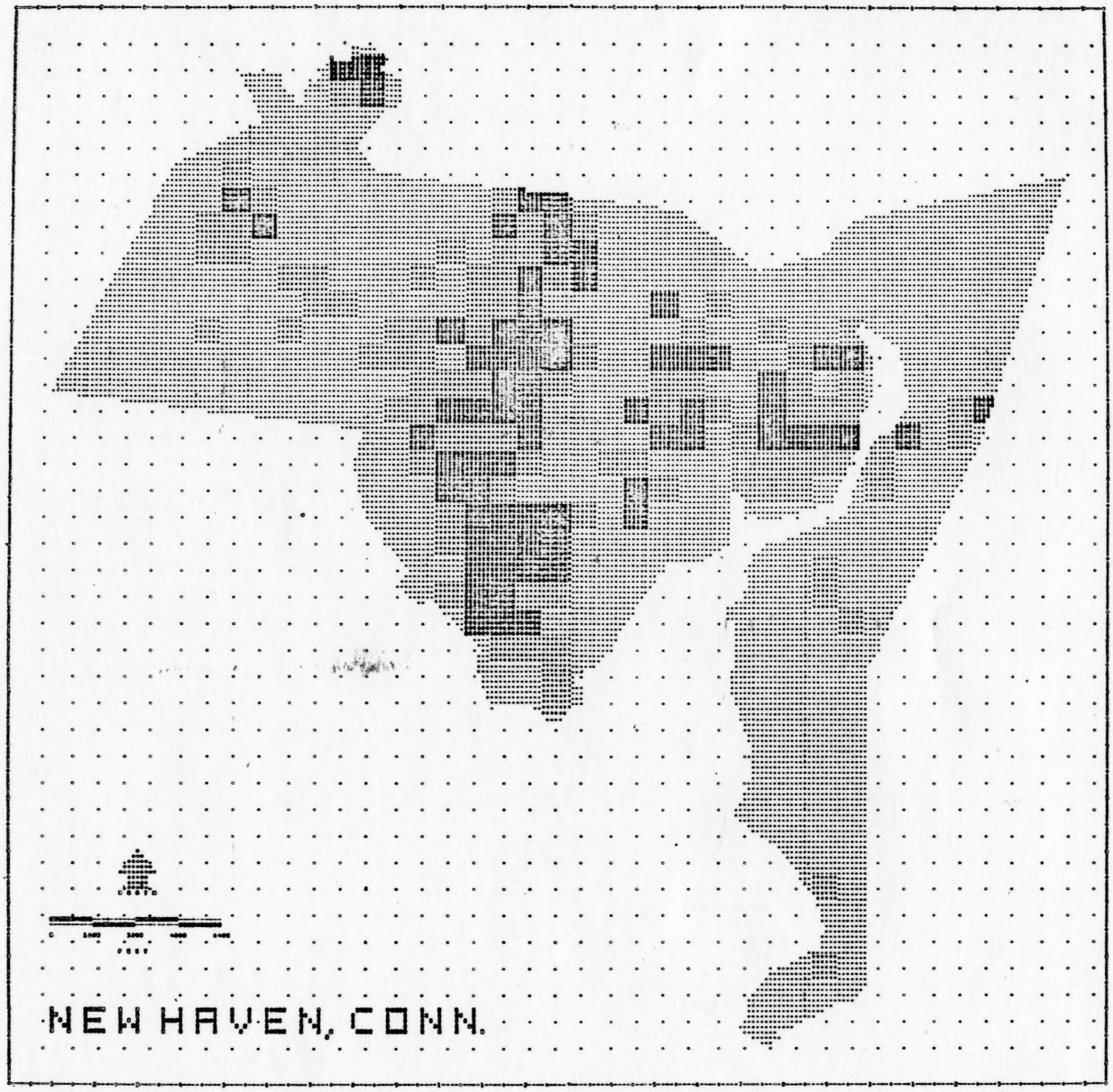


Figure 2. Proximal Option Map

REFERENCES

1. Ackoff, R., Scientific Method: Optimizing Applied Research Decisions, New York: John Wiley and Sons, Inc., 1962.

2. Barry, M.C., and C. G. Sheps, "A New Model for Community Health Planning," American Journal of Public Health, 59, No. 2, Feb. 1969, pp. 226-236.

3. Battistella, R. M., and T. P. Weil, "Comprehensive Health Care Planning--New Effort or Redirected Energy," New York State Journal of Medicine, 6, No. 18, Sept. 1969.

4. Bell Aerosystems Company, "Technical Proposal, Management Information System for the Sells Service Unit, Division of Indian Health, U.S. Public Health Service," submitted to Bureau of Health Services, Public Health Service, Department of Health, Education, and Welfare, Arlington, Tucson, May 11, 1968.

5. Berger, S., and A. Paller, "A Map is Worth a Thousand Printpoints," Computer Decisions, 2, No. 11, Nov. 1970, pp. 38-41.

6. Brandon, D. H., Management Planning for Data Processing, Princeton, New York: Brandon/Systems, 1970.

7. Chiang, C. L., "An Index of Health: Mathematical Models," Public Health Service Publication, No. 1000, Series 2, No. 5, May 1965.

8. Consad Research Corporation, Urban Information Systems and Requirements for Health Information, U.S. Public Health Service, June 1967.

9. Engel, A., Perspectives in Health Planning, London: The Athelone Press, 1968.

10. Ertel, P., R. C. Chase, and P. J. Ambuel, "A Multi-Purpose Outpatient Data System," Systems Development Project Comment Series, No. 9-6(19).

11. Evans, J. A., and R. V. Campbell, "The structuring of Health Information System Options: Summary and Assessment of Methodology," Socio-Economic Planning Sciences, 4, No. 2, June 1970.

12. Gabrielson, I. W., E. Siker, K. B. Sohler, and E. G. Stockwell, "Relating Health and Census Information for Health Planning," American Journal of Public Health, 59, No. 7, July 1969, pp. 1169-1176.

13. Garfield, S. R., "The Delivery of Medical Care," Scientific American, 222, No. 4, April 1970, pp. 15-23.

14. Geiger, H. J., "Tufts University Comprehensive Community Health Action Program," Office of Economic Opportunity Project, June 1970.

15. Geisler, R., "The Thomas Medical Information System," Datamation, 16, No. 6, June 1970, pp. 133-136.

16. Greene, R. A., and V. W. Sidel, The Use of Computer Mapping in Health Research, Boston: M.I.T. Press, 1968.

17. Hillboe, H. E., "Health Planning on a Community Basis," Medical Care, 2, No. 3, May-June 1968, pp. 203-214.

18. Hoffman, L., Computers and Privacy, Stanford University Accelerator Document, 51AC-Pub-479, Aug. 1968.

19. Hsiao, D. K., A File System for a Problem Solving Facility, Dissertation in Electrical Engineering, University of Pennsylvania, 1968.

20. IBM, IBM Systems/360 Principles of Operation, A22-6821-2, IBM Corporation, Poughkeepsie, New York, 1966.

21. Ichniowski, F. C. J., "Operations Research in Regional Medical Programs," First National Conference, Hospital and Health Services Division, American Institute of Industrial Engineers, Department of Health, Education, and Welfare, Rockville, Maryland, Feb. 1970.

22. Jacobs, A. H., and R. B. Froh, "Significance of Public Law 89-749," The New England Journal of Medicine, 279, No. 24, Dec. 1968, 1314-1318.

23. Kamrany, N. M., and A. N. Christakis, "Social Indicators in Perspective," Socio-Economic Planning Sciences, 4, No. 2, June 1970.

24. Kansas State Board of Health and Kansas Regional Medical Program," Requirements Definition Study for a Health Information System in Kansas," April 1969.

25. Kissick, W. L., "Health Policy Directions for the 1970's," The New England Journal of Medicine, 282, No. 24, June 11, 1970, pp. 1343-1354.

26. Lindberg, D. A. B., The Computer and Medical Care, Springfield, Illinois: Charles C. Thomas, 1968.

27. Lockheed Missles and Space Company, "Proposal for A Mayo Clinic Medical Information System," LMSC-894666.

28. McLachlan, G., and R. A. Shegog, Computers in the Service of Medicine, Essays on Current Research and Applications, London, New York, Toronto: Oxford University Press, 1968.

29. Murnaghan, J. H., and K. L. White, "Report of the Conference on Hospital Discharge Abstracts Systems," Medical Care, 8, No. 4, July-Aug. 1970.

30. NCHS, "Health Statistics for Comprehensive Health Planning and Evaluation," Working Draft Report by the Study Committee, The Public Health Conference on Records and Statistics, Health Services and Mental Health Administration, Department of Health, Education, and Welfare, Doc. No. 618.11 Rev., Feb. 1, 1970.

31. NCHS, "Medical Care, Health Status, and Family Income," Vital and Health Statistics, Series 10, No. 9, Jan. 1970, pp. 54-67.

32. Navano, V., "A Systems Approach to Health Planning," Health Services Research, 4, No. 2, Summer 1969, pp. 96-111.

33. Naylor, J. H., J. L. Balintfy, D. S. Burdick, and K. Chu, Computer Simulation Techniques, New York: John Wiley and Sons, Inc., 1968.

34. Packer, A. H., "Applying Cost Effectiveness Concepts to the Community Health System," Operations Research, 16, No. 2, May-June 1968, pp. 227-253.

35. Pool, I. S., S. McIntosh, and D. Griffel, "On the Design of Computer-Based Information Systems," Health Information System Report for Joint Center for Urban Studies, MIT Center for International Studies, Cambridge, Massachusetts, Sept. 1968.

36. Public Law 89-749, Comprehensive Health Planning and Public Health Services Amendments of 1966, 80 Stat. 1180, Nov. 3, 1966.

37. Roemer, M. I., A. T. Moustafa, and G. E. Hopkins, "A Proposed Hospital Quality Index: Hospital Death Rates Adjusted for Case Severity," Health Services Research, Summer 1968, pp. 96-118.

38. Rosove, P. E., Developing Computer-Based Information Systems, New York: John Wiley and Sons, 1967.

39. Rothenberg, R. E., The New American Medical Dictionary and Health Manual, New York: The New American Library, Inc., 1968.

40. Sanders, B. S., "Measuring Community Health Levels," American Journal of Public Health, 54, No. 7, July 1944, pp. 1063-1070.

41. Saylor, L. F., California Health Information for Planning Service, Summary Report for The Division of Patient Care, Facilities and Services, State of California, Department of Public Health, PHS Grant hm-00446, 1965-1968.

42. Smith, W. F., "Cost Effectiveness and Cost Benefit Analysis for Public Health Programs," Public Health Reports, 83, No. 11, Nov. 1968, pp. 899-906.

43. Sparling, F., "Measuring Medical Care Quality: A Comparative Study," Hospitals, 36, March 16, 1966, 1966, pp. 63-68.

44. Taylor, K. O., "Trends in Hospital Use," Public Health Reports, 84, No. 12, Dec. 1969, pp. 1037-1041.

45. U.S. Bureau of The Census, Census Use Study: General Description, Report No. 1, Washington, March 1970.

46. U.S. Bureau of The Census, Census Use Study: Computer Mapping, Report No. 2, Washington, August 1967.

47. U.S. Bureau of The Census, Census Use Study: Data Tabulation Activities, Report No. 3, Washington, March 1970.

48. U.S. Bureau of The Census, Census Use Study: Family Health Survey, Report No. 6, Washington, Sept. 1969.

49. U.S. Bureau of The Census, Census Use Study: Health Information System, Report No. 7, Washington, October 1969.

50. U.S. Bureau of the Census, Census Use Study: Data Uses In Health Planning, Report No. 8, Washington, June 1970.

51. U.S. Department of Health, Education, and Welfare, Cycle 1 of The Health Examination Survey: Sample and Response, Series 11, No. 1, Washington, 1964

52. U.S. Department of Housing and Urban Development, Urban and Regional Information Systems: Support for Planning in Metropolitan Areas, Washington, D.C.: U. S. Government Printing Office, July 1968.

53. Wagner, C. J., "The Health System and Its Components. Theoretical Relationships," Washington: Unpublished paper, 1967.

54. Weed, L. L., Medical Records, Medical Education, and Patient Care, Cleveland: The Press of Case Western Reserve University, 1969.

55. Williams, J. D., "Comprehensive Health Planning: An Organizational Means for Transition," American Journal of Public Health, 59, No. 1, Jan. 1969.