RESULTS FROM A PILOT STUDY OF SEVERAL METHODOLOGICAL PROBLEMS ASSOCIATED WITH CASEFINDING SURVEYS[1/]

William D. Kalsbeek, Research Triangle Institute

I. INTRODUCTION

There is a growing need for nationally-scaled data on specific morbidity conditions, data which measure the magnitude of both incidence and prevalence as well as the associated economic impact of each condition upon afflicted individuals and upon society as a whole. Their usefulness to public and private health care planners and implementers is clear. Associated with the production of these data are questions as to what statistical and field methods are best suited for what is termed to be a "casefinding study" in which sources, which provide care and treatment, are surveyed for cases with the relevant condition. The initial casefinding step yields sample cases to determine incidence and prevalence while a personal interview followup of the sample cases yields estimates of the economic impact or health costs of the condition.

The survey researcher is faced with alternatives which exist as many levels of operation and which beg an answer to the question of preference. This paper pinpoints a number of important methodological issues, describes the protocol which is used to address these issues in the form of a large-scale pilot study, and presents some preliminary findings as regards the issue of source cooperation.

This pilot study has been produced in connection with a national survey to quantify morbidity levels of Head and Spinal Cord Injury (HSCI) in the United States, a survey which will be conducted under the auspices of the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS). More specifically, HSCI refers to those injuries which result in traumatic damage to the neural tissues of the brain and spinal cord. These injuries are frequently associated with automobile accidents, athletic and other sports-related injuries, combat casualties, and other gun-related accidents. Because of the variable interpretation as to what constitutes an HSCI, operational definitions were required. Several of these definitions, which will be used in this paper, are presented in appendix A.

Other investigations have been made in connection with studies which collect data from sources of medical care and treatment [1-7]. While most of these investigations deal with some issues relating to the HSCI study, the present pilot study attempts to cover specific issues, considered more generally elsewhere, as they relate to HSCI.

Four types of medical care facilities are investigated as possible sources of HSCI cases: hospitals, office-based physicians, nursing homes, and "other prevalence sources" (see appendix A). Three different "prevalence approaches" by which a casefinding study might be deployed are investigated (the following terminology is our own):

(1) Single-Year Retrospective - Past records of visits to sources during the study period (1973) are observed. A sample pool of both incidence and prevalence patients is established directly from the observed records.

(2) Five-Year Retrospective - Since hospitalization is a necessary prerequisite to HSCI eligibility, only hospital records during the five-year period 1969-1973 are observed. HSCI-eligible individuals are first observed at a point in time subsequent to their injury incident. They are then traced to determine whether they have recovered, died, or continue to suffer from their injury in 1973. The sample pool of incidence patients is thus determined directly while the pool of prevalence patients is determined indirectly.

(3) Prospective - Sources are asked to list for four weeks forward in time all possible or probable HSCI cases who seek their treatment or services. Because time constrains the length of the time period in the pilot study, a longer period (e.g., six months) would be used in a national survey. Hospitals using the prospective approach are not set up in the pilot study although they would be included if this approach were adopted.

2. PILOT STUDY OBJECTIVES

While this pilot study's objectives are many and diverse, its principal short-range objective is to provide an arena in which a statistical field test can be constructed to determine the most cost-effective means by which a casefinding study can be used to produce national morbidity statistics on HSCI. Several of these objectives are now indicated in the form of study questions to be answered.

(1) Source Cooperation - The level of effort extended by way of soliciting source cooperation is considered. Specifically, what level of "public relations" (or endorsements) should be used to solicit cooperation of sources? How do cooperation levels compare among source types, and are there differences in cooperation levels within source types (e.g., term-size of hospitals)? Is there likely to be a difference in source response rates depending on the prevalence approach or medical background of the person conducting initial contacts with sources to solicit cooperation? To what extent will cooperating sources accommodate themselves to the field requirements of alternative prevalence approaches? How does the introduction of field costs affect any conclusions?

(2) Field Costs and Productivity - Unit costs of various field activities and productivity of the field staff are needed to develop an efficient sample design for the national survey. Field staff time and expense logs are therefore maintained to provide cost data of most field activities. Using these logs, it is possible to answer the question, what are the unit

expenses and productivity associated with both data abstraction of sources and personal followup interviews of HSCI patients? How do these unit costs compare among and within source types?

(3) Medical Background of Field Staff - The medical background of this study's field staff has possible implications on both the quality of collected data as well as unit field costs in the survey. Therefore, should specifically-trained medical records personnel be retained or can regular field staff with no particular medical training or background provide data of sufficient quality in view of the study's requirements? Should medical records personnel or regular field staff be used to solicit source cooperation?

(4) Medical Data Abstraction Instrument - An integral part of a casefinding study is the ability to collect sufficient data from a patient's source record so as to be able to determine that patient's HSCI eligibility from these data. These and other descriptive data are then used to assign him to various study domains defined by such variables as nature and cause of injury, age, sex, and place of residence. Several questions arise from this. How does the availability of these data compare among and within source types? How can the abstraction instrument best be structured to accommodate the ability of the abstractor and the requirements of HSCI eligibility? How well does the system of HSCI eligibility determination work, in which the eligibility of patients is resolved by a computer algorithm on the basis of medical data that have been collected by an abstractor? What medical definition is most operational? Is misenumeration which stems from misclassification of diagnostic codes or misdiagnosis by physicians a potential problem?

(5) HSCI Patient Yield - An important criterion for use in determining the sampling design for the main study is a source's patient yield, which is defined here to be the number of eligible HSCI patients in a source relative to the total number of patients. Thus, we attempt to determine how does yield compare among and within source types?

(6) Sampling of Records Within Sources - Since HSCI cases are not easily or uniformly identifiable in sources which do not generally employ diagnostic cross-referencing of records (i.e., all source types except for hospitals), sampling of records becomes a potential recourse. To this end, the pilot study attempts to answer the question, how difficult is it to sample patient records in the four source types? What is the nature of medical record systems in the source types? Are source record systems uniform enough as to develop uniform sampling techniques? Are nonstatistical field staff personnel able to implement sampling schemes independent of direct supervision?

(7) Health Costs Followup Instrument - The followup approach of patients to ascertain direct and indirect costs of HSCI is considered from three perspectives that differ in the degree to which the patient is asked to recall the type, duration, and actual costs associated with services received for HSCI. In this connection, then what is the best perspective for estimating these costs? Which perspective is most accurate and precise? Which is more feasible from the standpoint of field administration? What kinds of problems are encountered when substitute respondents are required in the cases of minor or deceased patients?

(8) Health Costs Patient Cooperation - In view of the current awareness of confidentiality and an individual's privacy, what level of cooperation may be expected in connection with requests for participation in the health costs patient followup?

(9) Multiplicity - A principal statistical problem in this type of casefinding study is multiplicity, which is defined most simply as a patient's tendency toward usage of more than one source during a study period. Since sources are surveyed initially, the patient can therefore be "casefound" by several sources. In view of multiplicity and requirements of various estimators which account for this problem [2,8], several questions will be answered. What is the extent of multiplicity among HSCI patients? How much information regarding a patient's multiplicity is available on source records? If it is not available on the records, how difficult to obtain is it?

(10) Other Statistical Measures - Quality checks involving reabstraction will be conducted in order to answer the question, what is the size of abstractor variability? Also, what is the magnitude and variation of any other variables which might be used to design the national sample?

3. PILOT STUDY DESIGN AND IMPLEMENTATION

The pilot study was conducted in eight sites which are geographically spread throughout the country. Each site consists of a single county. These counties were selected so that together they are a fair representation of settings that are likely to be encountered in the main study. A more detailed description of the eight sites is presented in table 1.

Separate design structures were established for each of the four source types. Each was constructed in such a way as to produce a direct statistical investigation of several of the above-mentioned issues and questions. Final source-specific designs for the record abstraction portion of the pilot study are presented in tables 2 through 5. The numbers of "design cells" are 40 for hospitals, 80 for office-based physicians, 24 for nursing homes, and 16 for other prevalence sources.

Each treatment combination in these tables represents a design cell in which an original and several alternate source selections were made. Up to four alternates were selected initially for each cell. Instructions as to use of alternates were the following:

(1) Approach the original selection and attempt to obtain cooperation.

(2) If unsuccessful on the basis of uniformly-imposed constraints, approach the first alternate.

(3) If unsuccessful on the basis of the same constraints, approach the next alternate.

(4) Continue in this manner until cooperation is achieved or field efforts are discontinued.

Alternates were selected for two reasons. First,

they allow for a determination of cooperation-refusal status in the event that the original selection did not make a response decision (e.g., field staff were unable to contact a physician since he was on vacation for several weeks). Second, they provide a means by which the design cell can be filled in order to measure and analyze other variables (e.g., yield of HSCI cases).

Each final design resulted from a random allocation of treatment combinations. In each design the level of Public Relations (PR) (i.e., Heavy or Light) was allocated to pairs of sites defined by Site Administrator (SA) type within Level of Urbanization. Thus, Levels of Urbanization are blocks, PR is a treatment, and SA is nested within blocks. As one site of the pair was randomly assigned one level, the other site received the other level.

In the hospital design (table 2) site within SA is a block. Each site is half of a full $\frac{1}{4} \times 2^2$ factorial in high urbanization sites and a 2^2 factorial in low urbanization sites with respect to the treatments Term-Size, Abstractor, and Prevalence Approach. A stratified sample of original and alternate hospitals was selected from a 1973 Master Facility Inventory (MFI) listing of hospitals in the eight sites. Treatments were assigned to hospitals in a predesignated order.

With nursing homes (table 3) blocking is again similar with a full 2^3 factorial within high urbanization SA and a full 2^2 within low urbanization SA. A stratified sample of nursing homes was taken from the 1973 MFI listing of nursing homes. Once again, treatments were assigned to selected nursing homes in a predesignated order.

With other prevalence sources (table 4) high and low urbanization are combined with urbanization defining the blocks with a full 2^3 factorial within blocks. A stratified sample of other prevalence sources was selected from a listing of all pilot site sources which provide any of a number of predetermined relevant services to HSCI patients. Treatments were assigned to select sources in a predesignated order.

Finally, office-based physicians (table 5) blocking is similar with a full 3×2^2 factorial within high urbanization SA and a full 2^3 within low urbanization SA. A stratified sample of office-based physicians was selected from a current listing of AMA physicians who indicate that they have office-based practices and that they are one of several predetermined specialty groups. Type strata were formed on the basis of a previous assignment of specialty groups to one of those subjectively-determined categories depending on the relative proportion of HSCI cases that are likely to be seen. Treatments were assigned to selected physicians in a predesignated order.

Field implementation associated with these activities began in January 1975 with attempts to solicit national endorsements of the HSCI project by relevant professional associations. Endorsements were received from some 21 organizations. Since allocation of PR was made by site, attempts were made to solicit endorsements from relevant state and local association in Heavy PR sites in addition to the national endorsements. Each endorsing organization was asked to provide RTI with a formal letter to which RTI field staff might refer when cooperation of sources was requested. Endorsement solicitation at this level was very successful. In Light PR sites no such local endorsement solicitations were made.

Two types of Site Administrators were used in the eight sites. Medical Record Accountants (MRA's) were recruited for abstraction on the basis that they have had formal training and/or experience in working with medical records. The top MRA applicants in four sites were retained as SA's. In three of the other four sites, Field Supervisors (FS's) from the off-site field staff of the RTI national general purpose sample were used as SA's. In Wake County, North Carolina an on-site Survey Specialist with no previous exposure to the project was used as an SA. Neither of the latter two types of field staff had any previous experience with casefinding surveys.

The SA's were trained in a single three-day training session at RTI in which all field activities were thoroughly discussed. No attemps were made to preferentially train any SA. All were given well-documented training manuals which included procedure documentation, a medical glossary, and a number of appendices which describe record systems and the contents and vernacular of medical records.

Immediately prior to the training session, mailouts were sent to all original and alternate sources. Mailouts differed according to level of PR. In Heavy PR sites two mailouts were made. The first mailout contained a brief description of the study and listed endorsing national organizations. In a few days a second mailout was sent which contained a colorful brochure clearly describing the study. Relevant state and local endorsement letters were included, and the letter indicated that an SA would contact them. In Light PR sites a single mailout consisted of a brief letter describing the study, listing the endorsing national organizations, and indicating that an SA would contact them. The mailouts to alternates were worded somewhat differently since contact for cooperation was not always made.

In all sites SA's made initial telephone followups approximately two weeks after the mailouts were completed. Attempts were made to set up an appointment for an introductory meeting at which time a detailed discussion of the source's role in the study was discussed. This discussion included a description of the level of endorsements received, the prevalence approach, the type of abstractor to be used (if requested) and the procedures involved in the health costs followup.

A good deal of difficulty was encountered with telephone followups in getting through to a source staff member who could make a participation decision. In several instances arrangements could not be made to speak or meet with this individual even after numerous attempts over a period of several days. Substitution of an alternate was then made by central RTI staff. Whenever possible, telephone followups were continued on successive alternates until a cooperating source was encountered.

4. RESPONSE ANALYSIS METHODOLOGY AND RESULTS

Two measurement units are used in the preliminary analyses of source cooperation which are presented in this section. First, note that for any cell-i of the designs of tables 2 through 5, n_i original and alternate sources were selected and designated for use in a predetermined ordered sequence. Let m_i be the number of selections ($m_i \leq n_i$) used before encountering a cooperating source or until time constraints prevented further solicitation for cooperation. Also define R_{ij} as the response of source-j in cell-i. Let R_{ij} assume dummy values according to the following:

$$R_{ij} = \begin{cases} 1 \text{ if source cooperates} \\ 2 \text{ if source refuses} \\ 3 \text{ otherwise (e.g., unable to meet, unable to contact, ineligible).} \end{cases}$$

If it is assumed that at least a single "1" or "2" is contained in each cell-i sequence R_{ij} ($j = 1,2,\ldots,m_i$), then one cell-i measure is

$$I_i = \begin{cases} 1 \text{ if a "1" appears first in the sequence } R_{ij} \ (j = 1,2,\ldots,m_i) \text{ [not counting "3"];} \\ 0 \text{ if a "2" appears first in the sequence } R_{ij} \ (j = 1,2,\ldots,m_i) \text{ [not counting "3"].} \end{cases} \quad (4.1)$$

Next, let C_{ij} be defined as SA costs in source-j of cell-i associated with arranging the introductory meeting by telephone and then conducting the meeting at the source. These costs reflect only the time required for these activities and do not indicate travel or other miscellaneous expenses. A second cell-i measure is defined by

$$C_i = \begin{cases} \sum_{j=1}^{m_i} C_{ij} & \text{if a "1" appears last in the sequence } R_{ij} \ (j=1,2,\ldots,m_i); \\ \text{missing data} & \text{if otherwise.} \end{cases} \quad (4.2)$$

The measure of (4.1) thus reflects the response of the first source in a sequence that makes a cooperation-refusal decision. The measure of (4.2) reflects the cost-effectiveness of cooperation solicitation in a cell since it is determined by sources' inclination to cooperate and the associated costs of solicitation.

Several analyses of variance were performed using the measure C_i and the Statistical Analysis System (SAS) subroutine package [9]. The general linear model subroutine of SAS assumes

$$E(\underline{C}) = \underline{\beta}X$$

where $\underline{C}$ is an observation array of the C_i, $\underline{\beta}$ is a coefficient array, and X is a matrix of dummy variables that SAS applies to perform analysis of variance. The model assumes that the C_i are normally distributed and of equal variance and were variably assigned to analysis sets defined by source type, level of urbanization, and treatment type. SAS treats a missing observation (i.e., no final cooperation obtained) by deleting it from the analyses since the program calculates only one sums-of-squares and sums-of-cross-products array, regardless of the model indicated.

Analysis of PR and other treatments (i.e., Abs, PA, TS, TC, Type) were done separately. This was done since PR is only of relevance to response analyses while the others are of interest here as well as in subsequent investigations of HSCI patient yield, abstraction costs, etc. Furthermore, unlike the other treatments, PR is administered in sites but applied to individual sources as a treatment. Results from separate analyses of PR by source type are presented in table 6.

Analyses for the other treatments by source type were done by level of urbanization except with other prevalence sources. This was done to accommodate the fact that TS, TC, and Type strata can be produced more specifically in urban sites due to larger numbers of available sources than in rural sites. Results of these analyses are presented in table 7. Means of treatment levels associated with tables 6 and 7 are presented in table 8.

Normality assumptions preclude use of the analysis of variance for the dichotomous I_i. For present purposes only marginal means are computed for treatments. These are computed by source type and level of urbanization and presented in table 9.

5. DISCUSSION

Tables 6 through 9 give rise to a number of observations:

(1) With $\alpha = 0.1$ SA type is at or near significance with costs in tables 6 and 7 except for low urbanization hospitals. Costs are appreciably greater with FS sources in most instances across low and high urbanization. Furthermore, the rates of FS to MRA costs from table 8 are generally greater than the corresponding ratios of hourly wages (1.58) indicating that MRA's tend to require less time to complete the activities. Table 9 generally supports the above observations particularly in high urbanization areas.

(2) PR is generally not significant with costs except for nursing homes. Table 9 indicates that cooperation may be somewhat greater with heavy PR hospitals although 2 x 2 chi-square tests of PR yield no significance here or with other source types.

(3) A number of interactions are significant particularly among PR analyses of nursing homes and office-based physicians in table 6. These interactions involve PR in both cases.

(4) Observing the means of I_i associated with treatment levels in table 9 moderate differences appear among PA levels in nonhospital sources, Term-Size levels in high urbanization hospitals, Type levels in high urbanization office-based physician practices, and Type of Service levels in high urbanization other prevalence sources. Less-noticeable level differences are generally consistent with other observations. Differences among levels of Abs in low urbanization sources appears to be an artifact attributable to small sample sizes.

While the first two observations appear to be clearcut, the following limitations of the above analyses are noted:

(1) Sample sizes and accompanying degrees of freedom are such that the power of many tests is low particularly in low urbanization analyses.

(2) Costs from which the C_i are computed exclude travel and out-of-pocket field expenses; however, if one can assume that this exclusion effects only location changes in the C_i, then this effect on the results is minimal.

(3) Missing costs data, particularly in office-based physicians (22 of 80 empty cells), dispels much of the orthogonality of comparisons which was originally created by the source type designs.

(4) Significant interactions in PR analyses of office-based physicians and nursing homes indicate that separate analyses of covariance, involving levels of PR and/or urbanization, might be appropriate in subsequent investigations of the response data.

APPENDIX A

HSCI Eligibility - Eligibility was determined separately for those patients with diagnoses corresponding to HSCI "included" and "case-finding" ICDA-8 codes that were indicated on their medical records. The criteria are as follows:

Included: (1) Hospitalization is required as a direct result of the HSCI;
(2) One or more sequelae may be indicated (e.g., dementia; persistent vegetative state; motor dysfunction; speech, hearing or visual dysfunction; epilepsy; post-concussion symptoms);

Casefinding: (1) Hospitalization is required as with "included";
(2) One or more of the following symptoms was noted within five days of the HSCI incident: loss of contact, seizures, headaches, vomiting cerebrospinal rhinorrhea, weakness of extremities;
(3) One or more sequalae may be indicated as with "included";

Pilot Study Period - The periods of time for which casefinding was applied were calendar year 1973 for retrospective-approach sources and a four-week period during June and July of 1975 for prospective-approach sources.

HSCI Recovery - During the study period (1) no treatment and/or services for HSCI has been sought and (2) no residual physiological defects (e.g., speech deficiencies, paralysis) are suffered.

Incidence Patient - An individual who has experienced an HSCI eligible incident during the pilot study period.

Prevalence Patient - An individual who has experienced an HSCI eligible incident during or prior to the pilot study period and who has not recovered before the start of the study period.

Hospital - Those sources with six or more impatient beds that are recognized by the Master Facility Index (MFI) of the National Center for Health Statistics (NCHS) as hospitals.

Nursing Home - Those sources recognized by the MFI of NCHS which provide care ranging from that comparable to a hospital to that which offers little more than room and board of limited quality.

Office-Based Physician - Any licensed, actively practicing Medical Doctor or Doctor of Osteopathy who is a member of the American Medical Association or American Osteopathic Association, respectively.

Other Prevalence Source - Those sources, other than a hospital, nursing home, or office-based physician, in which an HSCI patient received care which satisfies a need that is attributable to his HSCI.

REFERENCES

[1] Kramm, Elizabeth R., Crane, Marian M., Sirken, Monroe G., Brown, Morton L., "A Cystic Fibrosis Pilot Survey in Three New England States," American Journal of Public Health, Vol. 52, No. 12 (December 1962), pp. 2041-2057.

[2] National Center for Health Statistics: Design of Sample Surveys to Estimate the Prevalence of Rare Diseases. Vital and Health Statistics. PHS Pub. No. 1000-Series 2-No. 11. Public Health Service. Washington, U. S. Government Printing Office, October 1965.

[3] National Center for Health Statistics: Participation of Hospitals in the Pilot Study of the Hospital Discharge Survey. Vital and Health Statistics. PHS Pub. No. 1000-Series 2-No. 19. Public Health Service. Washington, U. S. Government Printing Office, October 1966.

[4] National Center for Health Statistics: Pilot Study on Patient Charge Statistics. Vital and Health Statistics. PHS Pub. No. 1000-Series 2-No. 28. Public Health Service. Washington, U. S. Government Printing Office, May 1968.

[5] National Center for Health Statistics: Development of the Design of the NCHS Hospital Discharge Survey. Vital and Health Statistics. PHS Pub. No. 1000-Series 2-No. 39. Public Health Service. Washington, U. S. Government Printing Office, September 1970.

[6] National Center for Health Statistics: National Ambulatory Medical Care Survey. Vital and Health Statistics. PHS Pub. No. 1000-Series 2-No. 61. Public Health Service. Washington, U. S. Government Printing Office, April 1974.

[7] National Center for Health Statistics: Design and Methodology for a National Survey of Nursing Homes. Vital and Health Statistics. PHS Pub. No. 1000-Series 1-No. 7. Public Health Service. Washington, U. S. Government Printing Office, September 1968.

[8] Sirken, Monroe G., "Stratified Sample Surveys with Multiplicity." Journal of the American Statistical Association. Vol. 67, No. 337 (March 1972), pp. 224-227.

[9] Barr, Anthony J., and Goodnight, James H., A Users Guide to the Statistical Analysis System, Department of Statistics, North Carolina State University, August 1972.

Table 1. Description of eight pilot sites

Region	State	County 1970 population	Principal City, 1970 population	Population: 1970 Percent change 1960-1970	Percent urban
Northeast	New York	Westchester, 894,406	White Plains, 50,220	10.6	93.8
Northeast	Pennsylvania	Allegheny, 1,605,016	Pittsburgh, 520,117	-1.4	94.8
Southeast	North Carolina	Wake, 228,453	Raleigh, 121,577	35.4	69.4
Southeast	Texas	Harris, 1,741,912	Houston, 1,232,802	40.1	95.5
Midwest	Missouri	Jefferson, 105,248	Festus, 7,530	37.6	16.8
Midwest	Wisconsin	Milwaukee, 1,054,063	Milwaukee, 717,099	-10.1	100.0
West	California	Placer, 77,306	Roseville, 17,895	26.3	40.5
West	California	Los Angeles 7,032,075	Los Angeles, 2,816,061	4.2	98.7

Table 2. Final hospital abstraction design for pilot study*

Level of Urbanization(U)	Site Administrator (SA)											
	MRA						FS					
High: Public Relations(PR)	Allegheny, Pa. H			Milwaukee, Wisc. L			Harris, Texas L			Los Angeles, Calif. H		
Term-Size (TS)	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.
(1) S1	M	S	+	M	F	+	M	F	+	M	S	+
(2) S1	I	F	+	I	S	+	I	S	-	I	F	+
(3) S2	M	F	+	M	S	+	M	S	-	M	F	+
(4) S2	I	S	+	I	F	+	I	F	-	I	S	+
(5) S3	M	S	+	M	F	+	M	F	+	M	S	+
(6) S3	I	F	+	I	S	+	I	S	+	I	F	+
(7) L	M	F	+	M	S	+	M	S	+	M	F	+
(8) L	I	S	+	I	F	+	I	F	-	I	S	+
Low: Public Relations(PR)	Jefferson, Mo. L			Placer, Calif. H			Westchester, N.Y. L			Wake, N.C. H		
Term-Size (TS)	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.
All	M	S	+	M	F	+	M	S	+	M	F	+
All	I	F	+	I	S	+	I	F	+	I	S	+

*The body of the table gives the levels of the various treatments (PR, TS, Abs., and PA) to be applied with each site.

Legend:

Public Relations = PR: H-Heavy; L-Light

Term Size = TS: S1-Short-Less than 175 Beds; S2-Short-175 to 400 Beds; S3-Short-Greater than 400 Beds; L-Long-All Sizes

Abstractor = Abs: M-Medical Records Accountant; I-Experienced Field Interviewer

Prevalence Approach = PA: S-Single Year Retrospective; F-Five Year Retrospective

Site Admn. = SA: MRA-Medical Records Accountant; FS-Field Supervisor

Status=Stat.: "+" - Cell has cooperating source; "-" - Cell has no cooperating source

Table 3. Final nursing home abstraction design for pilot study

Level of Urbanization(U)	Site Administrator (SA)											
	MRA						FS					
High:	Allegheny, Pa.			Milwaukee, Wisc.			Harris, Tex.,			Los Angeles, Calif.		
Public Relations(PR)	H			L			L			H		
Type of Care (TC)	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.
(1) Nursing	M	P	+	M	R	+	M	R	+	M	P	+
(2) Nursing	I	R	+	I	P	+	I	P	+	I	R	+
(3) Personal	M	R	+	M	P	+	M	P	+	M	R	+
(4) Personal	I	P	+	I	R	+	I	R	+	I	P	+
Low:	Jefferson, Mo.			Placer, Calif.			Westchester, N.Y.			Wake, N. C.		
Public Relations(PR)	L			H			L			H		
Type of Care (TC)	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.
(1) All	M	R	+	M	P	+	M	P	+	M	R	+
(2) All	I	P	+	I	R	+	I	R	+	I	P	+

*The body of the table gives the levels of the various treatments (PR, TC, Abs., and PA) to be applied within each site.

Public Relations=PR: H-Heavy; L-Light

Abstractor=Abs.: M-Medical Records Accountant; I-Experienced Field Interviewer

Prevalence Approach=PA: P-Prospective (Follow-up); R-Retrospective (Follow-back)

Site Admn.=SA: MRA-Medical Records Accountant; FS-Field Supervisor

Status=Stat.: "+" - Cell has cooperating source; "-" - Cell has no cooperating source

Table 4. Final other prevalence sources abstraction design for pilot study

Level of Urbanization(U)	Site Administrator (SA)											
	MRA						FS					
High:	Allegheny, Pa.			Milwaukee, Wisc.			Harris, Tex.,			Los Angeles, Calif.		
Public Relations(PR)	H			L			L			H		
Type of Service (TS)	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.
(1) M	M	P	+	I	P	+	I	R	-	M	R	-
(2) N	M	R	+	I	P	+	I	R	-	M	P	+
Low:	Jefferson, Mo.			Placer, Calif.			Westchester, N.Y.			Wake, N. C.		
Public Relations(PR)	L			H			L			H		
Type of Service (TS)	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.	Abs.	PA	Stat.
(1) M	I	R	+	M	P	+	I	P	+	M	R	+
(2) N	M	R	+	M	P	+	I	R	+	I	P	+

*The body of the table gives the levels of the various treatments (PR, TS, Abs., and PA) to be applied within each site.

Legend:

Public Relations=PR: H-Heavy; L-Light

Type of Service=TS: M-Medical; N-Nonmedical

Prevalence Approach=PA: P-Prospective (Follow-up); R-Retrospective (Follow-back)

Abstractor=Abs.: M-Medical Records Accountant; I-Experienced Field Interviewer

Site Administrator=SA: MRA-Medical Records Accountant; FS-Field Supervisor

Status=Stat.: "+" - Cell has cooperating source; "-" - Cell has no cooperating source

Table 5. Final office-based physician abstraction design for pilot study*

Level of Urbanization	Site Administrator (SA)															
	MRA								FS							
High: Public (PR) Relations	Allegheny, Pa. H					Milwaukee, Wisc. L					Harris, Tex. L					Los Angeles, Calif. H				
		Type	Abs.	PA	Stat.		Type	Abs.	PA	Stat.		Type	Abs.	PA	Stat.		Type	Abs.	PA	Stat.
	(1)	L	M	P	+	(1)	L	M	P	+	(1)	L	M	P	+	(1)	L	M	P	-
	(2)	L	M	R	+	(2)	L	M	R	+	(2)	L	M	R	-	(2)	L	M	R	-
	(3)	L	I	P	+	(3)	L	I	P	+	(3)	L	I	P	+	(3)	L	I	P	+
	(4)	L	I	R	-	(4)	L	I	R	-	(4)	L	I	R	-	(4)	L	I	R	-
	(5)	M	M	P	+	(5)	M	M	P	+	(5)	M	M	P	+	(5)	M	M	P	+
	(6)	M	M	R	+	(6)	M	M	R	-	(6)	M	M	R	+	(6)	M	M	R	-
	(7)	M	I	P	+	(7)	M	I	P	-	(7)	M	I	P	+	(7)	M	I	P	+
	(8)	M	I	R	+	(8)	M	I	R	-	(8)	M	I	R	-	(8)	M	I	R	+
	(9)	H	M	P	+	(9)	H	M	P	+	(9)	H	M	P	+	(9)	H	M	P	+
	(10)	H	M	R	+	(10)	H	M	R	+	(10)	H	M	R	-	(10)	H	M	R	+
	(11)	H	I	P	+	(11)	H	I	P	+	(11)	H	I	P	+	(11)	H	I	P	+
	(12)	H	I	R	+	(12)	H	I	R	+	(12)	H	I	R	-	(12)	H	I	R	-
Low: Public (PR) Relations	Jefferson, Mo. H					Placer, Calif. L					Westchester, N.Y. L					Wake, N. C. H				
		Type	Abs.	PA	Stat.		Type	Abs.	PA	Stat.		Type	Abs.	PA	Stat.		Type	Abs.	PA	Stat.
	(1)	L	M	P	+	(1)	L	M	P	+	(1)	L	M	P	+	(1)	L	M	P	+
	(2)	L	M	R	-	(2)	L	M	R	+	(2)	L	M	R	+	(2)	L	M	R	+
	(3)	L	I	P	+	(3)	L	I	P	+	(3)	L	I	P	+	(3)	L	I	P	+
	(4)	L	I	R	+	(4)	L	I	R	-	(4)	L	I	R	-	(4)	L	I	R	+
	(5)	MH	M	P	+	(5)	MH	M	P	+	(5)	MH	M	P	+	(5)	MH	M	P	+
	(6)	MH	M	R	+	(6)	MH	M	R	-	(6)	MH	M	R	-	(6)	MH	M	R	+
	(7)	MH	I	P	+	(7)	MH	I	P	+	(7)	MH	I	P	+	(7)	MH	I	P	+
	(8)	MH	I	R	+	(8)	MH	I	R	+	(8)	MH	I	R	-	(8)	MH	I	R	+

*The body of the table gives the levels of the various treatments (PR, Type, Abs., and PA) to be applied within each site.

Legend:

Public = PR Relations: H-Heavy; L-Light

Type: Proportion HSCI Cases Seen: L-Low; M-Medium; H-High (in rural sites M and H combined=MH)

Abstractor=Abs.: M-Medical Records Accountant; I-Experienced Field Interviewer

Prevalence Approach=PA: P-Prospective (Follow-up); R-Retrospective (Follow-back)

Site Administrator=SA: MRA-Medical Records Accountant; FS-Field Supervisor

Status=Stat.: "+" - Cell has cooperating source; "-" - Cell has no cooperating source

Table 6. Analyses of public relations for costs in dollars (C_i)

Hospitals			Office-Based Physicians		
Source	df	Significance Level	Source	df	Significance Level
PR	1	0.28	PR	1	0.14
U	1	0.49	U	1	0.44
SA(U)	2	0.00+	SA(U)	2	0.00+
PRxU	1	0.13	PRxU	1	0.01
PRxSA(U)	2	0.60	PRxSA(U)	2	0.00+
Error	28		Error	50	
Total	35		Total	57	
Nursing Homes			**Other Prevalence Sources***		
Source	df	Significance Level	Source	df	Significance Level
PR	1	0.07	PR	1	0.40
U	1	0.04	U	1	0.80
SA(U)	2	0.00+	SA(U)	2	0.27
PRxU	1	0.01	PRxU	1	0.32
PRxSA(U)	2	0.07	PRxSA(U)	1	0.21
Error	16		Error	6	
Total	23		Total	12	

*The $X'X$ matrix is singular making resulting sums of squares questionable.

Table 7. Analyses of other treatments for costs in dollars (C_i)

Level of Urbanization(U)	Hospitals			Office-Based Physicians			Nursing Homes			Other Prevalence Sources*		
	Source	df	Sig. Level	Source	df	Sig. Level	Source	df	Sig. Level	Source	df	Sig Level
High	SA	1	0.00+	SA	1	0.00+	SA	1	0.11			
	Site(SA)	2	0.83	Site(SA)	2	0.22	Site(SA)	2	0.50			
	TS	3	0.30	Type	2	0.52	TC	1	0.30			
	Abs.	1	0.93	Abs.	1	0.82	Abs.	1	0.20			
	PA	1	0.44	PA	1	0.53	PA	1	0.40			
	TS x Abs.	3	0.98	Type x Abs.	2	0.28	TC x Abs.	1	0.49			
	TS x PA	3	0.74	Type x PA	2	0.05	TC x PA	1	0.50			
	Abs. x PA	1	0.20	Abs. x PA	1	0.99	Abs. x PA	1	0.50	SA	1	0.13
	Error	12		Error	19		Error	6		Site(SA	5	0.40
	Total	27		Total	31		Total	15		TS	1	0.23
										Abs.	1	0.68
										PA	1	0.30
										Error	3	
										Total	12	
Low	SA	1	0.83	SA	1	0.00+	SA	1	0.07			
	Site(SA)	2	0.18	Site(SA)	2	0.01	Site(SA)	2	0.17			
	Abs.	1	0.34	Type	1	0.54	Abs.	1	0.39			
	PA	1	0.79	Abs.	1	0.97	PA	1	0.38			
	Error	2		PA	1	0.34	Error	2				
	Total	7		Type x Abs.	1	0.25	Total	7				
				Type x PA	1	0.38						
				Abs. x PA	1	0.52						
				Error	16							
				Total	25							

*High and low urbanization other prevalence sources were analyzed together.

Table 8. Marginal treatment means for costs in dollars (C_i)*

Level of Urbanization(U)	Treatments: Levels	Hospitals	Office-Based Physicians	Nursing Homes	Other Prevalence Sources†
Combined	PR: Light	16.53 (16)	10.69 (27)	10.96 (12)	14.05 (6)
	Heavy	17.81 (20)	9.21 (31)	9.12 (12)	10.26 (7)
	U: Low	16.29 (08)	10.52 (26)	12.68 (08)	13.74 (8)
	High	17.51 (28)	9.40 (32)	8.72 (16)	9.23 (5)
High	SA: MRA	11.29 (16)	5.80 (19)	6.98 (08)	8.96 (8)
	FS	25.80 (12)	14.66 (13)	10.46 (08)	16.89 (5)
	TS,TC, or Type:	S1 16.77 (08)	L 9.90 (09)	Nur. 9.78 (08)	M 8.38 (6)
		S2 12.36 (08)	M 9.80 (10)	Per. 7.65 (08)	N 15.12 (7)
		S3 16.78 (04)	H 8.73 (13)		
		L 23.50 (08)			
	Abs: MRA	18.25 (15)	9.19 (18)	10.06 (08)	10.70 (7)
	EFI	16.66 (13)	9.66 (14)	7.38 (08)	13.54 (6)
	PA:	S 16.48 (14)	R 8.32 (10)	R 9.56 (08)	R 14.96 (5)
		F 18.54 (14)	P 9.89 (22)	P 7.87 (08)	P 10.16 (8)
Low	SA: MRA	16.69 (14)	6.88 (13)	6.29 (04)	
	FS	15.88 (04)	14.15 (13)	19.06 (04)	
	TS,TC, or	(04)	L 11.55 (13)	(04)	
	Type:		MH 9.48 (13)	10.75 (04)	
	Abs: MRA	18.32 (04)	10.99 (13)	14.61 (04)	
	EFI	14.25 (04)	10.05 (13)	14.61 (04)	
	PA:	S 16.80 (04)	R 10.92 (10)	R 14.65 (04)	
		F 15.78 (04)	P 10.27 (16)	P 10.70 (04)	

†No urbanization distinction is made.

*Sample sizes are indicated parenthetically.

Table 9. Marginal treatment means for response (I_i)*

Level of Urbanization(U)	Treatments: Levels	Source Types: Hospitals	Office-Based Physicians	Nursing Homes	Other Prevalence Sources≠
Combined	PR: Light	0.60 (20)	0.45 (40)	0.83 (12)	0.63 (8)
	Heavy	0.80 (20)	0.55 (40)	0.92 (12)	0.75 (8)
	U: Low	0.75 (08)	0.62 (32)	0.75 (08)	0.88 (8)
	High	0.69 (32)	0.42 (48)	0.94 (16)	0.50 (8)
High	SA: MRA	1.00 (16)	0.63 (24)	1.00 (08)	0.88 (8)
	FS	0.38 (16)	0.21 (24)	0.88 (08)	0.50 (8)
	TS,TC, or Type:	S1 0.63 (08)	L 0.19 (16)	Nur. 0.88 (08)	M 0.75 (8)
		S2 0.50 (08)	M 0.38 (16)	Per. 1.00 (08)	N 0.62 (8)
		S3 1.00 (08)	H 0.69 (16)		
		L 0.63 (08)			
	Abs: MRA	0.69 (16)	0.46 (24)	0.88 (08)	0.75 (8)
	EFI	0.69 (16)	0.38 (24)	1.00 (08)	0.63 (8)
	PA:	S 0.75 (16)	R 0.25 (24)	R 0.88 (08)	R 0.38 (8)
		F 0.63 (16)	P 0.58 (24)	P 1.00 (08)	P 1.00 (8)
Low	SA: MRA	0.75 (04)	0.75 (16)	0.75 (04)	
	FS	0.75 (04)	0.50 (16)	0.75 (04)	
	TS,TC, or Type:		L 0.56 (16)		
			M 0.69 (16)		
	Abs: MRA	0.50 (04)	0.56 (16)	1.00 (04)	
	EFI	1.00 (04)	0.69 (16)	0.50 (04)	
	PA:	S 0.50 (04)	R 0.38 (16)	R 0.50 (04)	
		F 1.00 (04)	P 0.88 (16)	P 1.00 (04)	

≠No urbanization distinction is made.

*Sample sizes are indicated parenthetically.

Footnote

1/The work upon which this publication is based was performed persuant to a contract (No. N01-NS-4-2334) with the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) of the National Institutes of Health (NIH). However, the opinions expressed herein do not necessarily reflect the position or policy of NINCDS or NIH, and no official endorsement of the contents by NINCDS or NIH should be inferred.