HOSPITAL STAY RESPONSE ERROR ESTIMATES FOR THE HEALTH INSURANCE STUDY'S DAYTON BASELINE SURVEY

PREPARED UNDER A GRANT FROM THE U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

KENT H. MARQUIS

R-2555-HEW
MAY 1980

Rand
SANTA MONICA, CA. 90406

The research reported herein was performed pursuant to Grant No. 016B-7905 from the U.S. Department of Health, Education, and Welfare, Washington, D.C.

Library of Congress Cataloging in Publication Data

Marquis, Kent H
Hospital stay response error estimates for the health insurance study's Dayton baseline survey.

([Report] - Rand Corporation ; R-2555-HEW)
"Prepared for the Department of Health, Education, and Welfare."
Bibliography: p.
1. Hospital utilization--Ohio--Dayton--Statistical methods. 2. Health surveys--Ohio--Dayton--Statistical methods. 3. Errors, Theory of. 4. Estimation theory. I. United States. Dept. of Health, Education, and Welfare. II. Title. III. Series: Rand Corporation. Rand report ; R-2555-HEW. [DNLM: 1. Hospitalization. 2. Length of stay. 3. Costs and cost analysis. WX158 M357h]
AS36.R3 R-2555 [RA982.D3] 081s [362.1'1'072]
ISBN 0-8330-0199-X 80-11095

Published by The Rand Corporation

HOSPITAL STAY RESPONSE ERROR ESTIMATES FOR THE HEALTH INSURANCE STUDY'S DAYTON BASELINE SURVEY

PREPARED UNDER A GRANT FROM THE U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

KENT H. MARQUIS

R-2555-HEW
MAY 1980

Rand
SANTA MONICA, CA. 90406

PREFACE

This report is one of several that evaluate the quality of data obtained by the Health Insurance Study (HIS), an interdisciplinary research program conducted by The Rand Corporation under a grant from the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health, Education, and Welfare. In pursuing its goals of ascertaining the demand for health care and the effects of care on personal health, the study collects new data, using self-reporting and other methods, from people in several parts of the United States. A substantial part of the HIS research effort is devoted to evaluating these primary data for strengths and weaknesses.

The present report evaluates information about the use of inpatient hospital services, obtained in the Dayton Baseline survey. Other HIS publications, including those addressing methodology, are described in J. P. Newhouse and R. W. Archibald, *Overview of Health Insurance Study Publications*, P-6221, The Rand Corporation, November 1978.

SUMMARY

This data quality evaluation focuses on response biases and error variances in interview reports of hospital stays, their length, their cost, and when they occurred. The responses are from the Health Insurance Study's (HIS) Dayton Baseline survey conducted in 1974 to prepare for an experimental study of the demand for health services.

The hospital record check is the principal method used to evaluate response quality. Interview responses are compared with record information on a case-by-case basis, and the differences are used to infer the sizes of response biases and response error variance under varying assumptions about record quality. Record check evaluations of bias are supplemented by comparing Baseline means with similar estimates from other sources. As a further check on conclusions, our bias results are compared with conclusions reached by other record check studies of the same variables.

Very different estimates of hospital admission survey bias are produced by the two evaluation methods. Comparison of Baseline and other means suggests that the Baseline underestimated the hospital admission rate slightly. On the other hand, the record check revealed an apparently large positive survey response bias. However, the record check estimate (positive bias) was rejected on the basis that the design probably led to underestimates of omissions and that it caused random response, record, and processing errors to look like a systematic positive response bias. Therefore, we conclude that the average Baseline survey response bias is close to zero or slightly negative.

Baseline reports of the details of hospital stays (date, length, and own cost of each stay) contain small or no response biases and small to moderate amounts of response error variance, at least for the subset of survey answers that could be evaluated. Uncertainty remains concerning response effects in the survey information about family out-of-pocket expenditures for hospital stays. Bias estimates range from none to large negative; response error variance estimates range from none to 40 percent of total measured variance.

ACKNOWLEDGMENTS

I wish to express my sincere thanks to Susan Marquis, Joseph Newhouse, Seymour Sudman, and Clairice Veit for their very helpful comments on early drafts of this manuscript.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS vii

FIGURE AND TABLES xi

Section

I. INTRODUCTION 1
   - Organization of the Report 1
   - Baseline Sample and Measurement Procedures 1
   - Record Check Procedures 3
   - Overview of Error Estimation Methods 5

II. BIAS IN BASELINE INPATIENT ADMISSION REPORTS 15
   - Comparison of Average Hospitalization Rate with Estimates from Different Sources 16
   - Record Check Evaluation of Response Bias 19
   - Interpretations of the Bias Results 23

III. MEASUREMENT ERROR IN ATTRIBUTES OF HOSPITAL ADMISSIONS 32
   - Length of Stay 34
   - Own Cost of Stay 38
   - Date of Admission 44

IV. DISCUSSION AND CONCLUSIONS 47
   - Baseline Survey Bias 47
   - Baseline Response Error Variance 48
   - Implications for Use of Baseline Data 49
   - Implications for Record Check Designs 50
   - Implications for "Best-Estimation" Strategies 51

APPENDIX TABLES 53

REFERENCES 57

FIGURE

1.1. Cross-classification Outcomes for Survey Response and Hospital Record Comparisons 8

TABLES

1.1. Hospital Stays in the Baseline and Record Check Samples ... 5

2.1. Estimates of Survey Underreporting, Overreporting, and Net Bias, from Previous Record Check Studies of Hospital Stays 20

2.2. Cross-classification of Dayton Baseline and Record Reports of Hospital Stays 23

2.3. Record Check Results, by Individual Hospital 30

3.1. Upper and Lower Bound Estimators of the Percentage of Error Variance 34

3.2. Estimates of Average Length of Hospital Stay 35

3.3. Percentage of Error Variance in Length-of-Stay Reports 38

3.4. Survey and Hospital Mean Out-of-Pocket Costs 43

3.5. Percentage of Error Variance in Own Cost Reports 44

3.6. Mean Month of Admission for Survey and Hospital Reports 46

3.7. Percentage of Error Variance in Admission Month Reports 46

A.1. Age Weights Used To Adjust Admissions/Discharge Estimates 53

A.2. Annual per Capita Hospitalization Estimate, from the 1973-1974 Hospital Discharge Survey 54

A.3. Annual per Capita Hospitalization Estimate, from the 1973-1974 National Health Interview Survey 55

A.4. Annual per Capita Hospitalization Estimate, from the Dayton Baseline 56

A.5. Proportion of Hospital Bill Paid by Insurance, from Household Survey and Hospital Record Data 56

I. INTRODUCTION

This report evaluates the preenrollment, inpatient hospital use data collected by household survey[1] in Dayton, Ohio, the Health Insurance Study's (HIS) first site. The HIS uses household interview data about health and other characteristics of people to select and allocate families to treatments in the insurance experiment, to interpret experimental results, and to examine other hypotheses about individual health care behavior. Information about measurement errors in the raw data can be useful in meeting these goals effectively.

Measurement errors in survey reports of hospital stays and of hospital stay characteristics are evaluated in two ways: by comparing survey estimates with estimates derived from other data sources and by comparing individual survey answers with data in hospital records. Other evaluations of Dayton Baseline measurement error have examined the quality of outpatient expenditure reports (Marquis et al., 1976) and the quality of family income reports (Marquis and Marquis, 1976).

ORGANIZATION OF THE REPORT

Section I (Introduction) describes the data collection and record check procedures used for the Site I Baseline and the general strategy for estimating response bias. The following two sections evaluate the quality of the hospital stay data from the survey. The variables examined are reports of hospital admissions (Section II) and, for reported admissions, the length of stay, the family's out-of-pocket costs, and the admission date (Section III).

BASELINE SAMPLE AND MEASUREMENT PROCEDURES

The Baseline was a household interview survey of a stratified random sample of approximately 2000 families representative of the civilian,[2]

---

[1]After enrollment, hospital use data were collected on insurance claims.

[2]The sample also included 23 individuals on active military duty who were not living on a military base.

noninstitutionalized population in Dayton, Ohio, and surrounding geographical areas. Families with heads over 62 years of age were excluded.[3] The response rate was 80 percent of eligible families, yielding data for approximately 4400 persons. The sample used in this analysis included 4246 persons under 65 years of age.

The average interview lasted about 90 minutes and obtained information about all members of the family. All adults (18 years of age or older) responded for themselves.[4] Information about children was reported by the adult who stated he or she was most knowledgeable about the children's health care.

Interviewers asked the following questions to determine the number of times each family member had an inpatient hospital episode in the 12 months prior to the interview; parenthetical phrases were read when necessary:

4.01 Was anyone in this family unit *admitted* to a hospital at any time in the last 12 months? Please do not include emergency room visits that did not result in being admitted to a hospital.

4.02 Who was admitted to a hospital? FOR EACH PERSON MENTIONED ASK: How many times were you admitted to a hospital in the last 12 months?

4.03 Some hospital stays are hard to remember. For example, those which lasted less than a day, those which came very soon before or after other hospital stays, or admissions which were for tests, observations, minor treatment. (In addition to the stays already mentioned) was anyone in your family unit admitted to any hospital for any reason like these?

4.04 Who was that? FOR EACH PERSON MENTIONED ASK: How many (additional) times were you admitted to a hospital in the last 12 months?

---

[3]Two persons may be heads of a family (e.g., husband and wife). Use of this eligibility rule reduces, but does not eliminate, the number of observations of persons over 62 years of age.

[4]This was accomplished by repeated callbacks to families to obtain self responses from all adults. Proxy responses were accepted for the three adults who were unavailable for interview throughout the field period.

For each hospital stay identified, the following questions were asked to determine length of stay, date of admission, and the family's out-of-pocket cost.

- o What was the name of the hospital you were admitted to (most recently/just before that)?
- o IF NECESSARY: What is the address of (NAME OF HOSPITAL)? If you can't remember exactly, please describe where it is.
- o On what date were you admitted to this hospital (most recently/at that time)?
- o Counting the day you were admitted, how many days *altogether* were you there that visit? (How many days were you charged for?)
- o What did you pay for this stay in the hospital? Do not include the fees of doctors who charged or costs separately paid by insurance.

Additional questions about fees of the attending physician and the surgeon (if any) were asked; however, the quality of these answers is not evaluated in this report.

RECORD CHECK PROCEDURES

We compared Baseline survey reports about hospitalization with information in hospital records. The sampling unit is called a person-hospital pair. Three kinds of person-hospital pairs can be distinguished by using survey information:

1. *Reported stay pairs:* The particular hospital and the person's inpatient stays in it are reported in the Baseline.
2. *Outpatient use pairs:* The specific hospital is reported as (a) the usual source of the person's *outpatient* health care; (b) used for any outpatient care by the person; or (c) used for any outpatient care by anyone in the person's family.
3. *Unmentioned hospital pairs:* No outpatient use of the hospital is reported for the person or any member of his or her family, and no inpatient use of the hospital is reported for the person.

All pairs defined by categories 1 and 2 above were considered eligible for the record check. All other possible pairs (category 3) were considered ineligible.

The record check analysis is based on stays reported for persons under 65 years of age in the Baseline (Table 1.1). Not all of the Baseline stays could be included in the record check. The most frequent reason for omitting a stay was that the hospital had not completed its record check within the 10-month record check field period or had sent back a blank verification form and follow up had not produced record data.[5] Outright hospital refusals to participate were rare. Sixteen survey-reported stays were excluded because of the family's unwillingness to sign a form granting the study permission to check its medical records; 23 stays were excluded apparently because of clerical errors in the field. We examine the possible effects on response error estimates of not checking stays in category 3 and of excluding the 127 reported stays later.

To obtain record data, the hospital was sent a questionnaire[6] and asked to list information about admissions and outpatient treatment for the identified patient during a 12-month reference period ending on the day before the Baseline interview. The questionnaire gave the person's name[7] and address and asked whether any inpatient or outpatient treatment or consultation was furnished to the listed person during the reference period. Inpatient information requested included number of admissions to the hospital for the sample person and, for each admission, discharge diagnosis, main condition treated, dates of admission

---

[5]Reasons for hospital noncooperation (besides refusal) were not recorded. We do not know whether the hospital could not find the requested record, whether they failed to try to locate it, or whether they refused to participate. If the hospital did not return a person's form or returned it blank, the person's survey-reported stays at that hospital were deleted from the record check analysis. They are not part of the observed survey overreports discussed in Section II.

[6]Health Insurance Study Document Number HIEI 23, "Inpatient and Outpatient (Including Emergency Room) Verification Form."

[7]Clerks were asked to list maiden names, other last names of children, and other names the person might have used.

Table 1.1

HOSPITAL STAYS IN THE BASELINE AND RECORD CHECK SAMPLES

		Number of Stays
Stays reported in the Baseline survey[a]		568
Survey-reported stays excluded from record check		127
Family refused permission for record check	16	
Hospital could not complete record check	86	
Hospital refused to check records	2	
Records not checked for other reasons	23	
Stays reported in the Baseline and included in the record check		441

[a]For persons under 65 years of age.

and discharge, total hospital bill, amount paid by insurance or other third parties, and itemized charges for room and board, ancillary services, and professional fees.

We made only one attempt to obtain hospital record verification of each survey-reported stay. Other researchers sometimes "resubmit" reported but unverified stays to the record source to minimize the effects of matching and clerical errors on the verification outcome. This procedure can introduce an unintended bias into the observations unless the probability of resubmission is independent of the verification outcome. Independence could be achieved by resubmitting verified reported stays and verified reports of no stays. The unbiased resubmission procedure, however, would be expensive and alternative uses of extra record check funds (e.g., checking the stays in category 3) might produce greater bias-reducing benefits.

OVERVIEW OF ERROR ESTIMATION METHODS

The following is a summary of some of the main issues and procedures in making the Baseline interview response bias estimates. The record check was the principal method used to estimate the interview response bias. For practical reasons, a complete record check could not be made, raising the possibility that response bias estimates for

one of the survey variables (annual per-person hospital admission rate) would be positively biased. Therefore, we used different survey bias estimation procedures, insensitive to possible record check design biases, to provide "ball park" estimates of survey bias that, under certain assumptions, bound the true survey response bias.

Response Errors

The main objective was to estimate response bias and response error variance for each hospital stay variable in the Baseline interview (number of stays, cost, date, and length). We will discuss here the underlying response model and the procedures used to estimate response bias. The discussion of response error variance estimation is deferred until Section III.

The underlying structural model of a survey response measure, M_r, is that it is determined by the true characteristic of the variable, T, by systematic reporting biases, B_r, and by unsystematic response errors, e_r. A survey measure for respondent i is written

$$M_{ri} = T_i + B_{ri} + e_{ri},$$

where the expected value of the unsystematic error, $E(e_{ri})$, is zero; that is, the value of the term would average zero if it were possible to obtain many independent measures on M for this person but, on any given trial, the term could have a nonzero value. The subscript (r) indicates the source of the error is in the response. (We will discuss sources of error in more detail below).

The mean survey response, averaging over all respondents (assuming no missing data), is

$$\bar{M}_r = \bar{T} + \bar{B}_r + \bar{e}_r.$$

Since the expected value of e_r is 0, the survey response bias can be expressed as

$$\bar{B}_r = \bar{M}_r - \bar{T}, \qquad (1)$$

the difference between the true mean and the observed mean for the people who answered the survey questions.

Record Check Estimates of Survey Response Bias

One way of estimating response bias is to compare every survey response with its corresponding hospital record value. This produces an unbiased estimate of response bias if the record values are unbiased.

A record check, in theory, can provide an estimate of T [in Eq. (1)] so that the survey response bias, $\bar{B}_r$, can be estimated.

If all hospital records for all respondents in the survey were obtained, and we assume that record measures have a structure that is similar to the survey measures, the hospital record mean, $\bar{M}_h$, may be written

$$\bar{M}_h = \bar{T} + \bar{B}_h + \bar{e}_h.$$

The difference between the survey mean and the record mean provides an estimate of the survey response bias, and the estimate is unbiased if the record mean is unbiased:

$$\bar{M}_r - \bar{M}_h = (\bar{T} - \bar{T}) + (\bar{B}_r - \bar{B}_h) + (\bar{e}_r - \bar{e}_h),$$

$$(\bar{T} - \bar{T}) = (\bar{e}_r - \bar{e}_h) = 0,$$

so

$$\bar{M}_r - \bar{M}_h = (\bar{B}_r - \bar{B}_h).$$

If we assume that $\bar{B}_h = 0$ (an unbiased record), then

$$\bar{M}_r - \bar{M}_h = \bar{B}_r; \qquad (2)$$

that is, the difference is an estimate of the survey response bias.

Some features of the record check method are the unit of analysis, the cross-classification outcome, and the interpretation of survey net bias and error variance obtained from the cross-classification outcomes.

The unit of analysis is the hospital stay reported by either the survey respondent or the hospital. Some analyses use the "matched case" as the analysis unit, which is a hospital stay reported by both sources.

The possible results of comparing survey and record values--the cross-classification outcomes--are illustrated for a binary variable in the matrix in Fig. 1.1 (assuming no missing data in either source).

SURVEY VALUE	RECORD VALUE *Yes*	RECORD VALUE *No*	
Yes	A	B	A + B
No	C	D	
	A + C		

Fig. 1.1--Cross-classification outcomes for survey response and hospital record comparisons

If the survey and record agree that a hospital admission occurred for a person, the record check result is outcome A. If they classify the event differently, the result is either outcome B or C.

A, B, and C are used to designate the obtained frequencies of each record check outcome when the variable is binary. It is not necessary (and often not possible) to include the D classification.

Net bias is estimated as $(B - C)/N$, where N is the number of people observed (not the number of hospital stays checked). This is an unbiased estimate of response bias if the record data are unbiased, no processing biases exist, and the record check design has not introduced different amounts of bias into the counts of B and C.

To see that this net bias estimate is the same as Eq. (2), note that $\bar{M}_r = (A + B)/N$ and $\bar{M}_h = (A + C)/N$; then, $\bar{M}_r - \bar{M}_h = (A + B)/N - (A + C)/N = (B - C)/N$.

It is seldom possible to check all hospital records for all survey respondents because, in the United States, there is no central health

record repository for the population. Therefore, compromise, or incomplete, record check designs are used to estimate survey response bias and strong assumptions must be employed.

Two compromise record check procedures (AB and AC designs) can introduce an estimation bias because they do not do equally well at detecting both the B and C outcomes. The pure AB record check procedure is to conduct the survey first and to check records of people who report a hospital stay. This approach will miss most of the unreported but true stays (outcome C) because reports of zero stays are not checked. The C outcomes detected are only for respondents who reported some but not all of their multiple stays. The AC design procedure is first to list hospital stays using records and then to interview the patients to see if the known stays are reported. Few, if any, truly fabricated reports of stays (outcome B) can be detected because there is no attempt to verify reported admissions that are not part of the original record sample. Sometimes B-type errors are detected for persons reporting multiple stays, one of which was part of the record sample.

The ABC record check design can detect both kinds of errors equally well. A good example is Feather's (1972) study, which consulted the central government health records for all respondents interviewed in Saskatchewan, Canada. Saskatchewan has government-run health insurance, and the government records are reasonably comprehensive for hospital stays. The records were used to check both "yes" and "no" reports obtained in the survey.

Instead of net bias, most record check studies publish overreporting and underreporting rates. Using the notation in Fig. 1.1, the survey underreporting rate, representing the proportion of recorded hospital stays that are not reported in the survey, is estimated as

$$C/(A + C).$$

The survey overreporting rate is estimated as

$$B/(A + B),$$

and it is traditionally interpreted as the proportion of reported stays that should not have been mentioned in the survey. These rates reflect a combination of the possible kinds of errors: part of the response bias, part of the response error variance, part of the record bias and record error variance, part of the processing bias, and part of the processing error variance.

Some examples of why the over- and underreporting rates reflect more than the response bias may be helpful. Let us assume that we are using an AB design to check the hospital stays reported for person i in the Baseline interview. Under AB design procedures, we will consult the hospital record only if i reports at least one stay; if i says he had no stays, we must assume that none would be found if we made a thorough search of all records in all hospitals. If at least one stay is reported, we will check the records of the hospitals that i said he stayed in. Typically, the hospitals will confirm the reported stays, but some discrepancies will occur and their pattern is predictable: There will be more reported-unrecorded (B outcome) stays than unreported-recorded (C outcome) stays.

Some of the B outcomes could arise because of a positive response bias; person i could have remembered his last hospital stay as within the survey time reference period when really it occurred earlier. He could mistake outpatient treatment at a hospital for an inpatient stay. He could make up a stay just to please the interviewer.

Other B outcomes are generated by peripheral processes that are not necessarily related to response bias. For example, a clerk could misspell i's last name so that the hospital records are checked for patient j rather than patient i. Since hospital stays are relatively rare events, the probability of discovering a stay for j in this hospital is low. Similarly, the records clerk could pull the hospital record for j instead of i by mistake. Person i could misreport the name of the hospital (a response error, of course, but not one that affects the population estimate of the per-person rate of hospital stays). The interviewer could misrecord the name of the hospital or enter the reported stay in the wrong place on the questionnaire (say, in another family member's answer column). The data entry clerk could transpose some digits in the identifying information (person's name, hospital

name, date of service) or mispunch the number of reported or recorded stays.

The peripheral errors are random errors in the sense that, for a complete record check, each error eventually generates both a B and a C outcome which cancel out when estimating the response bias. For example, if i reports the wrong hospital name, X instead of Y, a complete record check would yield an overreported stay in hospital X (outcome B) and an underreported stay in hospital Y (outcome C). The AB record check design, however, finds only the B outcome. Since response bias is estimated as (B - C)/N, the estimate from the AB design is biased positively in direct proportion to the rate of occurrence of peripheral (random) errors. The response bias estimate from an ABC design remains unbiased.

We can make response bias estimates from incomplete (AB or AC) record check designs by estimating and correcting for the design and other biases. This is one approach used in this report to estimate the reporting bias in the annual per-person hospital rate; but the conclusions are heavily dependent on the assumptions.

Meaningful confidence intervals, reflecting the potential effect of random sampling and other errors on the under- and overreport proportions, cannot be computed for one-directional record checks. Elsewhere (Marquis, 1978), I have shown that estimates of both the B and C outcomes are necessary for an unbiased estimate of the error variance, and this estimate (in the cross-classified observations) is necessary for the computation of meaningful confidence intervals.

Different bias estimation considerations arise for length of stay, date of admission, and cost responses. The response bias estimates are based on matched cases, stays that the respondent and hospital agree took place (outcome A in Fig. 1.1). Since the analysis excludes the B and C outcomes, the incomplete record check design no longer directly affects the inferences about response bias. Since the B and C outcomes have been discarded, however, the bias estimates are based on a subset of observations that may not represent the entire set adequately. Further reductions in the analyzed set are caused by missing observations on the variable of interest.

We estimate response bias for characteristics of hospital stays (length, date, cost) by the differencing method. The difference between the survey and record means is obtained on matched cases that contain nonmissing information; then, t-statistics, based on the observed means and variances, are used to evaluate statistical significance.

Comparison of Baseline and Other Estimates

We have taken some additional steps in this report to help interpret the results of the record check. The Baseline estimates of the annual per-person hospital stay rate (and other variables) are compared with estimates of the parameters made from other sources, and results of other record check studies are reviewed before discussing each Baseline bias estimate. We used these other sources to establish a reasonable range within which the present record check response bias estimates might fall. Both kinds of additional estimates raise interpretation issues. The issues in interpreting record check results have already been mentioned. Here we will discuss some of the methods issues in comparing different survey estimates.

To establish tentative ranges for Baseline response bias, we compared Baseline estimates with estimates made from other (alternative) data sources, including personal interview surveys, institutional record surveys, and national accounts data.

The methodological points to note about these comparisons are that (1) the total bias rather than just the response bias is being estimated and (2) both the Baseline and the other source may contain errors.

Total bias is a general concept that includes systematic errors arising from all sources. Let us consider three sources (sampling, measurement, and processing) in a linear model:

Total bias = sampling bias + measurement bias + processing bias.

The survey mean reflects the average true score and the three sources of bias, sampling (X), measurement (M), and processing (P):

$$\bar{S} = \bar{T} + \bar{X}_s + \bar{M}_s + \bar{P}_s.$$

For the alternative data source, A, the notation for the mean is

$$\bar{A} = \bar{T} + \bar{X}_a + \bar{M}_a + \bar{P}_a.$$

Differencing the two means yields

$$\bar{S} - \bar{A} = (\bar{T} - \bar{T}) + (\bar{X}_s - \bar{X}_a) + (\bar{M}_s - \bar{M}_a) + (\bar{P}_s - \bar{P}_a).$$

If we assume the alternative source data are unbiased, then the difference of the means reflects the total survey bias:

$$\bar{S} - \bar{A} = \bar{X}_s + \bar{M}_s + \bar{P}_s.$$

This difference provides the desired "ball park" estimate of survey response bias. It is an unbiased estimate of survey response bias, $\bar{M}_s$, if there are no survey sampling or processing biases. If there are survey sampling and processing biases that have the same sign as the survey response bias, then the difference between the two observed averages provides an upper or lower bound for the survey response bias. In this case, a record check evaluation of survey response bias should yield an estimate that is somewhere between zero and the difference score.

If the alternative source data contain biases, then the difference will reflect them. Because of the subtraction, alternative source biases of the same sign as the survey biases will decrease the size of the difference, while the opposite sign biases increase the difference.

In Section II, we compare the Baseline hospitalization rate estimate to rate estimates from two other sources and find it to be somewhat below them, suggesting the possibility of a small, negative Baseline response bias. When other record check evaluations are reviewed, they indicate that small response biases (sign uncertain) may have occurred in other survey studies. The Baseline record check data

appear to show a large positive response bias, clearly not within the bounds suggested by the previous studies or the alternative rate estimates. We then examine the record check design bias and other hypotheses to reconcile the information. The conclusion we reach is that the record check design bias is probably responsible for the incongruous finding. The record check data are not useful for estimating the hospital rate response bias but can be used to estimate response errors in characteristics of hospital stays (length, cost, and admission date). The latter estimates are made in Section III and are within the bounds that other estimates and studies suggest are appropriate.

II. BIAS IN BASELINE INPATIENT ADMISSION REPORTS

In this section, we estimate response bias in the number of inpatient hospital admissions per person, per year inferred from the Dayton Baseline survey. First, we compare the Baseline estimate of the admission rate with estimates from two other sources, one a survey of hospital records and the other a household interview survey. The Baseline estimate is close to the estimate from the other interview survey but somewhat below the rate estimated from the external record survey. Next we examine other record check estimates of average survey bias in reporting hospital admissions. When the record check design is complete, little or no response bias is found. When incomplete designs are used, response biases seem to be present, but interpreting them as biases is probably incorrect. The Dayton Baseline record check results, also based on an incomplete design, seem to suggest that the survey responses contain a large positive bias, but this interpretation is also probably incorrect.

Thus, the two evaluation methods yield different results. Comparing mean rates suggests that the Baseline survey estimate is either unbiased or is a slight understatement of the population's admission rate. The record check suggests a large positive response bias. The two evaluations, then, disagree both on the sign and the size of the survey bias.

We make several attempts to reconcile the two Baseline bias estimates. First, both could be correct if there were large negative sampling biases that offset the large positive response biases. No evidence for large negative sampling biases could be found, however. Next, we examine the effect of omitting 127 survey-reported hospitalizations from the record check. But, even under the most favorable assumptions, inclusion of these cases does not affect the record check estimates importantly. Third, we investigate the effects of the record check design bias by imputing record check results for the many survey reports of no-hospitalizations that were not checked. This analysis does reconcile the two estimates by bringing the record

check outcome much closer to the result of the rate comparison findings; however, the imputation assumptions are strong and the conclusions drawn are largely a function of the assumptions made. Nevertheless, it is the only reconciliation hypothesis that is not contradicted by the data. Finally, we consider the possibility that definitional problems and hospital record errors could have affected the results. The limited tests made suggest that neither had important effects.

COMPARISON OF AVERAGE HOSPITALIZATION RATE WITH ESTIMATES FROM DIFFERENT SOURCES

The National Center for Health Statistics produces two annual estimates of the number of hospital discharges per person in the United States civilian noninstitutionalized population via the National Health Interview Survey (NHIS) and the Hospital Discharge Survey (HDS). We compare the Baseline survey rate with adjusted rates based on data from these two surveys.

The NHIS is a household survey that obtains hospitalization information for persons residing in households and classified as civilians not in institutions on the day of the survey. Hospitalization experience is limited to discharges from general, short-stay facilities and excludes stays of well newborns and stays lasting less than overnight. The HDS is a sample survey of discharge records of nonfederal, general, short-stay hospitals in the United States. As such, it includes a somewhat broader definition of hospital stays (e.g., stays lasting less than overnight), but it excludes stays of newborns and civilians in federal and military hospitals (e.g., military dependents). The population of persons included is somewhat difficult to pin down. In general, the persons represented are in the resident civilian U.S. population, but stays of military and institutionalized persons in civilian, nonfederal hospitals are included, as are stays of persons who die in hospitals or soon after discharge. This latter category of persons is not eligible for inclusion in most household interview surveys such as the NHIS or the Dayton Baseline.

To facilitate comparison with the Baseline, the NHIS and HDS estimates of annual rates of hospital discharges are adjusted in three

basic ways: (1) to include stays only of civilian, noninstitutionalized persons, (2) to include admissions for stays lasting less than overnight but, for convenience, to exclude stays of well newborn infants, and (3) to apply to medium-size, urban areas in the north central region of the United States, comparable to the HIS sample in the Dayton area. To accomplish the latter, published regional estimates are used when available. Finally, each estimate of stays (admissions or discharges)[1] is adjusted to a common age distribution for persons less than 65 years old (Appendix Table A.1).

In comparing estimates of mean stay rates, standard errors are computed only for Baseline estimates. To facilitate computation, the number of adjustments to Baseline data is kept to a minimum.

Hospital Discharge Survey Estimate (1973-1974)

We obtained a preliminary estimate from the HDS for the population aged 0-64 in the north central United States for the years 1973 and 1974 directly from published data and adjusted it for civilians discharged from military and federal government hospitals and for the age distribution.

The HDS does not detect discharges of civilians from military and federally run hospitals such as those of the Veterans Administration. In 1973, there were slightly over one million discharges[2] of civilians from VA and Department of Defense hospitals located in the United States (Croner, 1977). We increased the HDS estimate 4 percent[3] to account for these discharges, under the assumption that the number of discharges of military and institutionalized persons from enumerated hospitals is negligible (less than 1 percent of the total).

---

[1] Any differences between admissions and discharges as indicators of hospital stays is ignored.

[2] This includes 600,000 discharges from VA hospitals for persons under 65 and 420,000 discharges of persons not on military active duty and under 65 from Department of Defense hospitals (Croner, 1977, Tables 20 and 23). Regional estimates are not available.

[3] The denominator is 25.2 million HDS discharges in the United States for persons under 65 in 1973 (Lewis, 1976).

The derivation of the final HDS estimate of .16 annual hospitalizations per person under 65 years of age is shown in Appendix Table A.2.

National Health Interview Survey Estimate 1973-1974

The NHIS data come from interviews covering a representative sample of the U.S. civilian noninstitutionalized population. All types of short-term, general hospital stays for this population are included except admissions for stays less than overnight and stays of most newborn infants. We averaged the published 1973 and 1974 U.S. estimates, inflated the averages to include stays of less than a day, and adjusted them to the common age distribution and to the expected rates for small Standard Metropolitan Statistical Areas (SMSAs) in the north central region.[4]

Estimates of the ratio of less-than-overnight to all stays are based on 1973 and 1974 Hospital Discharge Survey data (Ranofsky, 1976; Lewis, 1976). The derivation of the adjusted NHIS estimate of .141 stays per person under 65 years of age appears in Appendix Table A.3.

The Dayton Baseline Estimate

The Dayton Baseline estimate of the hospital admission rate is for the population under 65 years of age[5] residing permanently in households in Dayton, Ohio, and surrounding areas corresponding to a small SMSA in the north central region of the United States. The estimate excludes families in Dayton-area households who resided there during the 12 months but not at the time of the interview. The hospital experience of the group is partly compensated for by including persons who moved into the Dayton household population during the 12-month reference period. The hospitalization experience of those who died during the year, however, is not represented.

[4]The Dayton, Ohio, area is one of these areas. The adjustment ratio (1.06) is based on regional and national estimates for the 1969-1970 surveys (Namey and Wilson, 1974) and applied to the 1973-1974 U.S. data.

[5]Persons aged 62 to 65 are slightly underrepresented in the Baseline sample, causing a possible downward bias in the estimate of mean stays. It has not been possible to estimate the size of this bias, but it is assumed to be very small.

The Baseline sample contains four persons in long-term institutions (e.g., nursing and convalescent homes) at the time of the interview, and the estimate includes their hospitalization experience. Hospitalization of newborns has been estimated and excluded: Seven persons were born during the reference period and reported hospitalized at least once during that time, so seven admissions for persons less than 17 years old have been deleted. There has been no correction for the slight underrepresentation of persons between 62 and 65 years old.

The adjusted estimate of the number of hospital admissions over 12 months per person under 65 years of age in the HIS Dayton sample is .134. The derivation is shown in Appendix Table A.4.

Admission Rate Comparisons

The Baseline estimate of .134 admissions per person under 65 years old is lower than either the HDS estimate (.16) or the NHIS estimate (.141). The standard error of the Baseline mean is .006 admissions, so the Baseline and NHIS means are within two Baseline standard errors of each other. The Baseline estimate, however, is significantly lower than the HDS estimate. A possible conclusion is that the Baseline has slightly underestimated the population hospital admission rate. We next tried a different evaluation approach.

RECORD CHECK EVALUATION OF RESPONSE BIAS

In the preceding analysis, we looked at the possibility of a total Baseline survey bias, including the effects of biases in the sample, the measurements, and the processing. There is also the possibility of a measurement bias; it is theoretically possible for offsetting measurement, sample, and processing biases to occur and be undetected by the methods just employed.

Measurement bias is estimated by comparing survey responses with a criterion source. In this research, Baseline reports of hospital admissions are compared with admission data obtained from hospital records. If the records are correct, then the differences between survey and record values are indicators of errors made in the survey reports or in processing and linking the survey and record data. The terminology and estimation methods used in record checks of survey data

were given at the end of Section I. Before examining our record check results, we consider record check findings obtained by other researchers.

Previous Record Check Studies

Record check studies of survey bias in reporting annual inpatient hospital episodes published before 1974 are reviewed in Marquis (1978). The results are summarized in Table 2.1. Estimates of gross underreporting were derived primarily from studies that sampled hospital records and interviewed the families of persons known to have been hospitalized (AC designs; see the end of Section I). Estimates of gross overreporting were obtained from research that checked survey reports of a hospital admission with hospitals, insurance companies, or records of prepaid health plans (AB designs). The full design studies were based on interviews with cross-section population samples and on checking a comprehensive set of records for survey reports of no stays, as well as for survey reports of one or more inpatient hospital episodes.

Table 2.1

ESTIMATES OF SURVEY UNDERREPORTING, OVERREPORTING, AND NET BIAS, FROM PREVIOUS RECORD CHECK STUDIES OF HOSPITAL STAYS

(In percent)

	Appropriate One-Directional Design Studies	Full Design Studies
Estimates of underreporting[a]	13-17	11-14
Estimates of overreporting	10-15	11-14
Estimates of net bias	(b)	± 3

SOURCE: Based on Marquis (1978).

[a]Excludes two lower estimates (3 percent and 4 percent) obtained for subsamples of cases and using procedures likely to cause record errors to be correlated with survey values.

[b]There are no appropriate one-directional studies in this category.

The conclusion drawn from the pattern of results is that survey reports of inpatient hospitalizations contain little or no response bias; response errors in one direction (e.g., underreports) are largely offset by compensating errors in the other direction (overreports). It is not possible to tell whether the compensating errors originate in the survey responses, the records, or in the response/record matching procedures.

Two new estimates of hospital admission response error have been published recently. Both use modified AB record check procedures that are similar in principle to the HIS Baseline record check design: Hospital records are checked for admissions reported in the survey and additional attempts are made to discover unreported admissions. Both new studies gathered information from insurance records and one study also contacted physicians listed as usual sources of health care or reported as having been visited by the sample person.

Andersen et al. (1979) estimate a net positive admission reporting bias of 3 percent.[6] The authors acknowledge the possibility that the estimate may be positively biased because the record check may not have detected enough of the apparent survey underreports. The Medical Economics Survey Methods Study (HSRDC, 1977) observed a 6 percent net survey underestimate of admissions in its urban sample and a 3 percent underestimate in a rural sample. The estimates compare survey values for cases that could be checked against records to a "best" estimate.[7] The net survey underreport finding is probably due to the logic used to construct the best estimate, since "best" appears to be defined as the highest positive value available from either the survey or the two record sources (the hospital or the insurance third party).[8] Despite

---

[6]Hospitals and other sources provided data about 1668 admissions. The observed overreport rate was 163/1558 admissions, or 10 percent. The observed underreport rate was 110/1505 admissions, or 7 percent.

[7]The study presents other best estimates that accept unchecked survey values as accurate; these estimates are not discussed here.

[8]"When information was available from one or more record sources and these sources disagreed with the household data, preference was usually given to the record sources, especially in the case of likely

the possible limitations in the record check design and the potential problems in the definition of a "best" criterion value, both new studies suggest that survey reporting biases for hospital admissions are not large. The possibility cannot be ruled out that the analysis assumptions determined the sign of the bias. The sign issue remains unresolved in the literature but does not appear to have much practical importance; most recent studies agree that the magnitude of any response bias is very small.

A small pretest of the Baseline hospital record check procedures was conducted, checking only those stays reported by responding families with records of Dayton-area hospitals. The hospitals verified only 23 of the 31 admissions reported in the pretest Baseline and reported one admission not mentioned in the survey, an implied survey overreporting rate of 26 percent. The rate of overreporting is substantially higher than those reported by other researchers using AB or ABC record check designs (Table 2.1). Further examination of the pretest data indicated one hospital was responsible for six of the eight unmatched reports, suggesting that their record-checking procedures might be biased. Provision was made to monitor the match results on a hospital-by-hospital basis during the main study so field personnel could follow up any similar problems during the data collection period.

Baseline Record Check Results

The record check in the main study shows a large overreporting rate for survey reports of hospital admissions and a small underreporting rate. The sign of the implied net bias is positive and is opposite that expected from a comparison of the Baseline mean with external

---

underreporting on the part of the respondent. When the household-reported utilization was [checked against but] not confirmed by the record source, the general rule was to include the household-reported utilization in the 'Best Data' set. . . . [T]he assumption was made that unconfirmed utilization was more likely a problem of incomplete provider records than overreporting on the part of the household respondent" (HSRDC, 1977, Appendix VIII, p. 2). The effect of the procedure is probably to incorporate the "positive" random errors into the best estimates while excluding the compensating "negative" random errors. If so, then the best estimates will be overestimates of the true population value. For additional discussion, see Marquis (1979).

estimates. The overreport estimate is 26 percent, similar to the pretest estimate and substantially larger than estimates of overreporting error produced by other record check studies of hospitalization. The underreporting rate is 7 percent. Table 2.2 gives the results.

Table 2.2

CROSS-CLASSIFICATION OF DAYTON BASELINE AND RECORD REPORTS OF HOSPITAL STAYS

(Frequencies)

Survey Response	Hospital Record		Total
	Yes	*No*	
Yes	327	114	441
No	26		
Total	353		

Overreport = 114/441 = 26 percent

Underreport = 26/353 = 7 percent

INTERPRETATIONS OF THE BIAS RESULTS

In this subsection, we discuss the two conflicting data evaluation results: The comparison to the HDS estimate suggests that the Baseline hospital stay estimate is slightly negatively biased; the record check results suggest that the Baseline Hospital stay reports contain a large positive response bias. The first hypotheses we consider are about conditions under which both results could be true. For both results to be valid, the positive response bias would have to be offset by large negative sampling or processing biases. We find evidence supporting a small negative sampling bias but not a large one. The second set of hypotheses is about problems in the record check itself. We conclude that the record check underestimated underreports, yielding the wrong conclusion about net survey response bias. Finally, we review some hypotheses about sources of the unusually large random error rate implied by the record check. A possible confusion between admissions for less than a day and outpatient treatment cannot account for the high error level, nor is there evidence that one particular hospital's poor performance produced the higher error rates (as in the pretest).

Retrospective Sampling Bias

The Dayton Baseline almost certainly contains a negative retrospective sampling bias.[9] The sample of families defined as eligible for the survey included only people alive and residing in the area at the time of the survey. Yet the hospital admission data were obtained over a retrospective 12-month reference period during which other people would have been included in the sample, principally those who had since moved or died. Those who died undoubtedly had a higher per-person hospital rate than survivors. Their hospitalization would be included in the HDS data but not in the Baseline. (For a fuller discussion, see Marquis, 1978.)

The size of the possible negative retrospective sampling bias in the Baseline estimate is probably small because the estimate is for people under 65 years old, a group that has a relatively low mortality rate. The decedent part of the bias is in two parts: people who died in the hospital during the reference period and people discharged alive during the reference period who would have been eligible for the survey except that they died before the interviewing period. In the United States, about 286,000 people under 65 died in the hospital during the 1973-1974 12-month reference period (based on Moien, 1976, and Glickman, 1977; numbers for Dayton or the north central region were not located). Using population estimates for 1973-1974 reported by Wilder (1974) and Ries (1975), this part of the decedent sampling bias is -286,000/186,075,000 = -.00154 stay per person over 12 months, too small to affect the Dayton Baseline estimate (.13 stay per person) importantly.

The second part of the decedent sampling bias cannot be estimated confidently from existing data. Based on data provided by Simmons (1967) and Sutton (1965), the rate of stays for hospitalized persons under 65 who were discharged alive but died within the survey 12-month reference period might be between 1 and 1.5 times the rate of stays of persons discharged not alive. If so, the decedent sampling bias is

[9] This statement is from the point of view of making population estimates for a past time period. The bias does not compromise one of the main Health Insurance Study Baseline objectives, which was to provide information about people eligible for enrollment in the subsequent insurance experiment.

between -.0035 and -.0044 stay per person or between -2.6 and -3.3 percent of the mean rate estimated from the Baseline. The magnitude of the negative decedent sampling bias is not large enough to offset the large positive response bias implied by the record check, nor will it account completely for the difference between the Baseline and HDS estimates.

Baseline Noncoverage and Nonresponse Sample Bias

Although unlikely in practice, the contradictory bias findings could both be true if (a) those who responded to the Baseline were healthier (had fewer hospitalizations) than sampled people who were not located or refused to respond, and (b) responders overreported their hospital admissions at a rate that would compensate for the noncoverage and nonresponse bias. This hypothesis might be entertained further if evidence supporting the first condition could be found.

No direct evidence is available, but indirect evidence fails to indicate a large negative sample bias due to noncoverage and nonresponse. Direct evidence would require knowing the true hospital stay experience of all persons in the sample and finding that the noninterviewed segment had very high admission rates. Unfortunately, true rates for sample persons are not known, so indirect indicators are necessary. The best available indirect indicator of hospitalization experience is the sample person's age; older people are more likely to be hospitalized than younger persons. The following table shows the age distribution of persons according to the 1970 census in the approximate area sampled by the Dayton Baseline and the age distribution of respondents to the Baseline:

	Percent Distribution	
Age	1970 Census	1974 Baseline
0-24	51	50
25-44	27	30
45-64	22	20

The distributions are very similar but with an apparent small (22 - 20 = 2 percent) underrepresentation of people 45-64 years old in the Baseline respondent group. The difference is probably due to definitional differences; the census includes all permanent residents of long-term institutions and the Baseline includes an accidental few. Ignoring the definition difference, the data would imply that the Baseline respondent group should have included about 85 more people in the 45-64 age range (adding 85 people would result in 22 percent of the sample being in this age range and the total respondent group size would be 4246 + 85 = 4331 people). If every other added person had been hospitalized once in the 12-month reference period, the new Baseline stay rate per person would equal the NHIS estimate. If every added person had been hospitalized one-and-a-half times, the new Baseline stay rate would equal the HDS rate. Neither of these possibilities seems likely. Thus, it is difficult to make a case for a large negative sample bias in the Dayton Baseline stay rate estimate.

Record Check Sample Biases

The record check sample did not produce record information about all of the hospital stay information reported in the Baseline; not all positive reports of stays were checked. It could be that the reported but unchecked stays contain less response error than others, leading to an overestimate of the response bias. More important, only a small proportion of the possible checks was made for Baseline reports of zero hospital stays. Had more checks been made, more C outcomes (see Fig. 1.1) might have been discovered. Since C outcomes are subtracted from B outcomes to determine survey bias, higher B outcome rates have the effect of reducing the estimate of response bias in this research. Discovery of a large number of B outcomes could change the sign of the bias estimate.

Imputation for Missing Checks of Survey-Reported Admissions. If the response errors in unchecked survey "yes" reports were different from response errors of the checked admissions, would the estimate of survey response bias change importantly? The answer is "no" and is explained next.

The survey reported 568 hospital admissions, but only 441 were checked against hospital records. (Recall from Table 1.1 that 16 could not be checked because the family refused permission to do so, that hospitals failed to furnish any data about 88 admissions, and 23 were not checked for other reasons.) A total of 127 reported stays were not checked against records.

Making the most extreme assumption, that *all* 127 unchecked reports would have been verified by records, the recalculated overreport rate would still be much larger than the recalculated underreport rate:

$$\text{Recalculated overreport rate} = \frac{114}{441 + 127} = .20,$$

$$\text{Recalculated underreport rate} = \frac{26}{353 + 127} = .05.$$

Thus, the discrepancy between the external source and record check bias estimates cannot be reconciled by imputing favorable match outcomes to the unchecked survey reports of hospital stays.

Imputation for Missing Checks of Zero Hospital Stays. The design problem judged most important in preventing a good record check estimate of response bias is the failure to check all survey reports of no hospital admissions (or to estimate the results of a complete check by checking reports from a systematic sample of respondents). Below, we derive an estimate of Baseline response bias by assuming the results of a complete check of all Dayton-area hospitals. The new response bias estimate is close to the Baseline survey bias estimate obtained using the HDS data, but the new estimate rests on strong assumptions.

The Baseline record check included some (but not all possible) checks on the validity of survey claims of no admissions. The checks made were at "mentioned" hospitals, those judged beforehand to be the most likely source of inpatient care for each checked person[10] and hence the most likely (efficient) record source to consult to detect a stay unreported in the survey.

[10]Mentioned hospitals are hospitals named by persons as the usual source of care and hospitals that provided outpatient care to the person or a member of the person's family during the 12-month recall period.

From these limited checks, the probability of detecting an unreported stay (Type C outcome) at a mentioned hospital is estimated as .014.[11] The probability of detecting an unreported stay by checking any unmentioned hospital can be assumed to be lower than this; for illustration purposes, it is assumed that the probability of detecting an underreported stay by checking records at a randomly selected hospital (mentioned or not) is .01.

Hypothetically, if all 3778 persons reporting zero stays were checked at all seven Dayton-area hospitals, the expected number of Type C outcomes would be 264, i.e., $3778 \times 7 \times .01 = 264$.

If the remaining six hospitals were checked for the 468 persons reporting a stay, 28 more Type C outcomes would be expected ($468 \times 6 \times .01 = 28$).

Six additional omissions, representing apparently unreported multiple stays, were obtained by the complete check of positive admission reports.

This brings the hypothetical expected number of Type C outcomes from a complete check of Dayton-area hospitals for all respondents to 298, i.e., $264 + 28 + 6 = 298$.

The number of Type B outcomes expected from a complete record check would include the 114 actually observed plus 33 that might have been detected if all survey "yes" cases had been checked.[12]

The net survey response bias is defined as $(B - C)/N$. Using the estimates above,

$$\text{Hypothetical net survey response bias} = \frac{147 - 298}{4246} = -.036.$$

---

[11]The calculation is 20/1448 = .014. Twenty underreports resulted from 1448 record checks of survey "no admission" responses at mentioned hospitals. The remaining six observed underreports were detected for people reporting at least one admission.

[12]One hundred twenty-seven admissions reported in the Baseline were not checked (Table 1.1). Assuming the unchecked admission reports are like the checked, the probability of a Type B outcome is .26 (Table 2.2), and 33 additional Type B outcomes would be expected ($127 \times .26 = 33$).

If the net bias is subtracted from the original Baseline estimate, .13 - (-.036) = .166, the corrected estimate is very close to the external HDS synthetic estimate of .16.

This apparent reconciliation should not be taken very seriously, however, because of the strong, unsupported assumptions upon which it rests.

Two points are illustrated by the preceding analysis:

1. Incomplete record checks can misestimate even the sign of the (true) response bias.
2. A systematic sampling procedure for checking "negative" survey reports of a probability sample of the universe of record sources should be included with an AB record check design. The additional design feature avoids having to invoke tenuous assumptions about the representativeness of checked cases.

Some Other Possible Sources of Measurement Error

The Baseline questionnaire included a probe question designed to elicit unreported hospital stays for which the discharge was on the same day as the admission. If respondents reported hospital outpatient treatments (e.g., clinic or emergency room visits) in response to that question, these visits would appear as survey overreports in the record check.

It is reasonably certain that this kind of confusion is not responsible for the high overreporting rates observed. The HIS record check also asked hospitals to list outpatient treatments for patients. There were 56 survey-reported admissions lasting one day or less (it is not possible to identify only admissions lasting less than one day). In two cases, the hospital did not return the verification form. Of the remaining 54 survey-reported stays, the hospital did not confirm 45, so these are apparent survey overreports. An upper limit estimate of potential respondent (or hospital) confusion between inpatient and outpatient treatment can be made by noting that, in 17 of these 45 cases, the hospital did report treating the patient on an outpatient basis during the 12-month reference period. If we assume all 17

apparent overreports are due to definition ambiguity, then subtracting them from the B outcome cell of Table 2.2 reduces the overreport rate from 26 percent to 23 percent, still substantially above rates found in other research. Although the special probe question may have increased confusion between inpatient and outpatient treatment, that confusion would not contribute importantly to the high overreport rates.

Hospitals may differ in their ability to furnish accurate verification information. If hospitals fail to confirm reported true admissions, apparent survey overreports will result. If hospitals make random matching or linking errors, the effect (within an AB record check design) is also to produce apparent survey overreports. The pretest of the hospital record check verification procedures revealed that most unmatched survey reports of admissions were to one particular hospital. An examination of the hospital-by-hospital mismatches in the present research does not show that kind of problem (Table 2.3). Hospital B

Table 2.3

RECORD CHECK RESULTS, BY INDIVIDUAL HOSPITAL
(Number of admissions)

Hospital	Record Check Agreement	Hospital *No* Survey *Yes*	Percentage of Overreport	Hospital *Yes* Survey *No*	Percentage of Underreport
A	29	10	26	0	0
B	41	32	44	11	21
C	10	0	0	1	9
D	10	2	17	2	17
E	80	19	19	7	8
F	69	31	31	1	1
G	74	20	21	1	1
All other hospitals (N = 47)	14	0	0	3	18
Total	327	114	26	26	7

NOTE: The data include 7 Dayton-area hospitals and 47 hospitals in other areas mentioned by respondents.

produced somewhat higher error rates than the others, but eliminating these data from the analysis would only reduce overreports to 22 percent and underreports to 5 percent. The unexpectedly large error rates found in the record check, therefore, cannot be attributed to the faulty performance of a single hospital.

III. MEASUREMENT ERROR IN ATTRIBUTES OF HOSPITAL ADMISSIONS

In this section, we offer estimates of response errors in reports of length of stay, out-of-pocket cost per stay, and month of admission. For matched cases, Baseline reports of length of stay contain a small amount of positive reporting bias and a small to moderate amount of reporting error variance; reports of own cost appear unbiased on the average but may contain large amounts of error variance; date of admission reports contain a small amount of net forward telescoping and a small amount of error variance. The Baseline response bias trends are similar to findings of other survey record check evaluations.

For each topic discussed, we compare the Baseline estimate of the mean to a synthetic estimate from another source (month of admission excepted); next, we evaluate evidence of survey bias from other record check research; then we use the Baseline record check results to provide estimates of the form and size of the Baseline measurement errors.

Baseline record check analyses discussed here are based only on matched cases, those hospital stays that both the survey and record agree took place for a particular person, at a particular hospital, and within the defined time reference period. We did not include stays unchecked with the record, about which there is disagreement over whether the admission occurred, or unreported by either source. The effect of these omissions on conclusions cannot be known with certainty. If the omissions occur at random, the bias estimate is within the limits of random sampling error. This would be the case, for example, if mismatches were due to random mistakes in matching the survey and record data. On the other hand, omitted cases may be due to systematic reporting, recording, or matching errors concerning the true admission. Such errors may be a sign that the attributes of the stay (length, cost, date) contain different (possibly larger) amounts of bias or random error variance than the matched stays. Furthermore, the samples are attenuated additionally by item (question) nonresponse, especially for the own cost analysis. We made no attempt to estimate or adjust for these potential sample biases in these analyses.

When measurement errors are found, it is desirable to know their source. Marquis et al. (1976) present a method for estimating survey and record error variance from record check studies. They define the measured survey value (X_S) and the measured record value (X_R) as a linear combination of the true value (X_T) and measurement error (e_S and e_R):

$$X_S = X_T + e_S,$$
$$X_R = X_T + e_R \ .$$

Assuming (1) that the errors are uncorrelated with the true values and with each other and (2) that there are no matching errors, the error variance in the survey and record measures may be estimated as

$$\text{var } e_S = \text{cov } (X_D, X_S),$$
$$\text{var } e_R = -\text{cov } (X_D, X_R),$$

where $X_D = X_S - X_R$. These estimates will overstate the survey and record error variances if there are matching errors introduced by another source. This can be demonstrated by extending the Neter et al. (1965) model of matching errors in ABC design studies. They define q as the probability that the response of person i is linked to the record of person k ($i \neq k$). By direct extension of their formulas,

$$\text{cov } (X_D, X_S) = \text{var } e_S + qN \text{ var } X_T,$$
$$\text{cov } (X_D, X_R) = - \text{ var } e_R - qN \text{ var } X_T,$$

where N is the number of person-record pairs and there are no unmatched elements. Using the covariance formulas of Marquis et al. (1976), the survey and record error variances will be overestimated by qN var X_T, the effect of matching errors. Although the size of the matching error variance is unknown, upper and lower bound estimates of the survey and record error variances can be made. The upper bound estimates assume

qN var X_T = 0; the lower bound estimates assume that one source[1] contains zero error variance, so that its discrepancy covariance equals qN var X_T, as shown in Table 3.1.

Table 3.1

UPPER AND LOWER BOUND ESTIMATORS OF THE PERCENTAGE OF ERROR VARIANCE

Bound	If var (S) > var (R)	If var (S) < var (R)
Survey upper bound	cov (S,D)/var (S)	cov (S,D)/var (S)
Survey lower bound	$\frac{\text{cov (S,D) + cov (R,D)}}{\text{var (S)}}$	zero
Record upper bound	-cov (R,D)/var (R)	-cov (R,D)/var (R)
Record lower bound	zero	$\frac{\text{cov (S,D) + cov (R,D)}}{\text{var (R)}}$

LENGTH OF STAY

The analyses that follow show that the Baseline mean length of stay is similar to the estimate derived from NHIS survey data. Both interview survey estimates are longer than the HDS estimate, which comes directly from hospital records. The difference probably reflects interview response bias, because the interview questions may not have clearly communicated to all respondents that they should count all but one of the "days" spent in the hospital when reporting length of stay.

Other record check studies consistently find survey length of stay reports to be longer than what appears in hospital records. The HIS record check, using matched stays, reveals the same type of bias at the upper ranges of the magnitudes as reported in the other research. The HIS record check analyses suggest that the survey data contain at least twice as much error variance as the record data.

[1]This is the source with the smallest measured variance.

Comparative Estimates of Average Length of Stay

Table 3.2 compares length of stay estimates from the NHIS and HDS with Baseline estimates. The definitions of eligible persons, hospitals, and stays for these surveys and sources of published data are given in Section II. We made no adjustments to the data for regional effects.

Table 3.2

ESTIMATES OF AVERAGE LENGTH OF HOSPITAL STAY
(Adjusted to common age distribution)

Source	Average Length of Stay (Days)
HDS (1973-1974 national)	6.4
NHIS (1973-1974 national plus stays of less than overnight)	6.9
Dayton Baseline (1973-1974)	7.0

The HDS estimate is lower than the estimates obtained by the household surveys (Baseline and NHIS). Because the HDS data come from written records, this estimate of length of stay may be more accurate (less subject to systematic recall and reporting bias) than the other data. HDS data exclude stays in federal hospitals, which tend to be considerably longer (Croner, 1977, Table 18). Stays by civilians in federal hospitals are reported in the household surveys but are a very small proportion of the total. It is unlikely that adding these stays to the HDS data would increase the estimate by more than 3 or 4 percent. On the basis of the information available, stays in the two household surveys are 0.5 to 0.6 day longer on the average than stays in the HDS.

Another possible reason for the difference is that household survey respondents are unclear about the definition of length of stay, specifically whether to include the day of discharge in the reported length. The HDS computes the length of stay "by counting all days from (and including) the date of admission to (but not including) the date of discharge" or by counting one day for stays of patients admitted and discharged on the same day (Lewis, 1976). The Dayton

Baseline asks the following question of respondents for each admission reported: "Counting the day you were admitted, how many days altogether were you there that visit? (How many days were you charged for?)"[2] Respondents in National Health Interview Surveys are asked: "How many nights was (person's name) in the hospital (nursing home)?" If the person correctly reports the number of nights of care or the number of days charged for, the determination of length of stay is compatible with the HDS definition for stays lasting overnight or longer. On the other hand, if the number of days spent in the hospital is reported instead (that is, both the admission and discharge days are counted), the household survey estimate will exceed the HDS estimate by one day per stay (for stays lasting overnight or longer). Perhaps the household survey respondents sometimes confuse calendar days with nights or days charged, causing these length of stay estimates to be biased upward relative to the estimates from discharge records.

Other Record Check Studies

Nine record check studies present some information about survey and record agreement on length of hospital stay. When the analysis is confined to the cases for which both sources acknowledge the admission (matched cases), survey data appear to contain a net positive reporting bias of small magnitude.

Belloc (1954), Barlow et al. (1960), and Andersen et al. (1979) found an average overstatement of length of stay of 0.6 day.[3] Cannell et al. (1965) found a 0.3 day average overstatement and Balamuth et al. (1961) reported that their respondents made a positive 0.2 day average error.

Other studies have published estimates in a somewhat different form. Cannell and Fowler (1963) found that a higher percentage of respondents overestimate length of stay than underestimate it. Kulley

[2]The phrase in parentheses is used at the interviewer's discretion.

[3]It is not entirely clear whether Andersen et al. limited the analysis to matched cases or also included Type B outcomes. If the latter, then the average bias for matched cases would be larger.

(1974) states there is substantial agreement between survey and record values[4] but does not analyze the direction or magnitude of the differences. Richardson and Freeman (1972) present survey and record distributions for a sample of aged respondents that indicate little or no net survey bias in stay length reports.[5] The only study that suggests length of stay is underestimated in a survey is the Medical Economics Survey Methods Study (HSRDC, 1977); survey reports range from an average of 2 to 5 percent *below* a "best" estimate of length of stay. The rules for constructing the best estimate, however, may have forced this result.[6] Where studies have computed the difference between survey and record length of stay values on a case-by-case basis and limited the analysis to stays for which both sources agree took place, the average survey response bias is an overstatement that ranges from 0.2 to 0.6 day per stay.

Baseline Record Check Results

The HIS record check also indicates that Baseline survey reports of length of stay are biased upward compared with information in hospital records. The average length of stay[7] in matched case Baseline reports is 7.51 days compared with a 6.85-day mean record report. The difference of .67 day is statistically significant (the standard error of the difference is .21 day).

---

[4]The intercept term in a regression using a dichotomous dependent accuracy variable is .72 (that is, 72 percent accuracy), and the sign of most of the explanatory variable coefficients is positive.

[5]Means are not presented and cannot be recalculated with sufficient accuracy due to the broad categories used for the length of stay classifications.

[6]Procedures for constructing the best estimate of length of stay are not specified in the report. Their other best estimates are constructed mainly by accepting the single highest estimate available from the survey and two record sources. Comparing any single source with this kind of best estimate will result in inferring that the source contains a negative bias.

[7]All analyses reported here assign a value of one day to inpatient stays lasting less than a day.

The estimate is in the same direction and slightly above the +0.6-day bias found in other record check studies employing cross-section household survey samples. The response bias estimate is also consistent with the total survey bias estimate obtained by comparing the Baseline mean with the adjusted HDS mean length of stay.[8]

The error variance estimates (Table 3.3) suggest that the Baseline survey data contain more error variance than the record data. The

Table 3.3

PERCENTAGE OF ERROR VARIANCE IN LENGTH-OF-STAY REPORTS
(Matched cases)

Source	Variance	Covariance with Difference	Estimate of Percentage of Error Variance	
			Upper Bound[a]	Lower Bound[b]
Survey (S)	43.35	10.10	23	13
Record (R)	37.72	-4.47	12	00
Difference (D)	14.56			

NOTE: N = 320 matched cases.

[a]Upper bound = $\frac{\text{cov (S,D)}}{\text{var (S)}}$ or $-\frac{\text{cov (R,D)}}{\text{var (R)}}$.

[b]Lower bound = $\frac{\text{cov (S,D)} + \text{cov (R,D)}}{\text{var (S)}}$, or zero.

amount of total measured survey variance that is simple response error variance is in the range of 13 to 23 percent for the matched cases. The random error variance in record data or introduced by other sources is in the range of zero to 12 percent of total measured record variance.

OWN COST OF STAY

The next analysis shows that the average family out-of-pocket cost for hospital stays reported in the Baseline is considerably lower than

[8]The record check sample means are higher than the Baseline mean, which indicates that short stays reported in the Baseline have a lower probability of inclusion in the record check sample than longer stays.

an estimate derived from national accounts data compiled by the Social Security Administration. Yet our record check finds no response bias. Only one other record check study has evaluated the family's own cost estimates, and its results are equivocal. Record check evaluations of *total* bill reporting generally agree that the average bias is close to zero. The HIS record check finds that the average survey/record own-cost discrepancy is not significantly different from zero for matched cases, suggesting the Baseline data either are unbiased on the average[9] or the stays with large, true out-of-pocket expenditures are not in the matched case or Baseline samples. The Baseline survey net cost data appear to contain a moderate to large amount of error variance, but this may reflect errors introduced in the record check matching and transcription operations.

Comparative Estimates of Average Own Cost of Hospital Care

We compared Baseline estimates of the family's out-of-pocket costs for hospital care with estimates derived from data published by the Office of Research Statistics of the Social Security Administration (SSA). The Baseline estimates are substantially smaller than the SSA estimates.

The "external" estimate desired is the amount paid for hospital inpatient care by the living, Dayton, noninstitutionalized, civilian population under 65 years of age in FY1974 and not paid or reimbursed by third parties. National estimates of out-of-pocket hospital expenditures for FY1974 for all age groups are available from the SSA and can be adjusted tentatively to reflect spending of the younger age group. It is not possible to adjust either for regional effects or the out-of-pocket expenses of decedents and the institutionalized. Worthington (1975) estimates total out-of-pocket payments for hospital care in the United States in FY1974 to be $4274 million. Mueller and Gibson (1976) estimate total private sector payments of the U.S. population over 64 years old at $1577 million in FY1974. These consist of family

[9] It also suggests the estimate based on SSA data contains an upward bias when applied to the Dayton sample.

out-of-pocket payments, private insurance company payments, and payments by industry and philanthropic organizations. Mueller and Piro (1976) suggest that 50 to 60 percent of civilians over 64 years of age have some private insurance covering costs of hospital care (e.g., supplements to Medicare coverage). On this basis, we assume here that 50 percent of the private sector payments for hospital care of patients over 65 are paid by third parties and the remainder ($788 million) are out-of-pocket. Using these assumptions, FY1974 out-of-pocket spending for hospital care of patients under 65 is $4274 million - $788 million = $3486 million.

Dividing the SSA estimate by the size of the 1974 civilian U.S. population (193 million) suggests that per capita own cost for hospital care for patients under 65 was $18.06 in FY1974. The own cost per hospital discharge for this population is estimated to be $124.50 based on 28 million discharges. The SSA hospital care data include expenditures for outpatient care in hospitals. This is estimated to be $1.80 per person, using data from the 1970 Health Care Survey (Andersen et al., 1973) and inflated to FY1974 dollars. The per capita estimate based on SSA and health care survey data is then $16.26.

The estimates of out-of-pocket costs for inpatient hospital care for the Dayton Baseline are $9.26 per capita and $69.08 per admission[10] for patients under 65. (Own costs were unknown for 50 reported stays and imputed at the patient age-group mean for these estimates. The Dayton data are not adjusted to the U.S. age distribution since this adjustment would not affect the comparison with SSA estimates.) The Baseline estimates are considerably lower than the SSA-based estimates.

The SSA-based estimate reported here may be wrong for several reasons. First, the basic estimate of out-of-pocket payments is the small residual left over after estimating total payments to hospitals and all government, philanthropic, industry, and private health insurance payments. Small errors in any of these components (or the total) can have large effects on the out-of-pocket remainder estimate. Second, the

---

[10] Excluding one outlier, the Baseline per capita mean is $7.84 and the per stay estimate is $58.64.

adjustments we used to arrive at the own cost estimate for persons under 65 are very inexact. Unfortunately, although the estimates are error prone, it is not possible to state the directional effects the errors might have. More direct estimates of own costs will be available for future research and policy planning.[11]

If the estimate based on SSA data is correct, then either the Baseline respondents underreport their out-of-pocket hospital expenses by a substantial amount on the average or there is a sample/nonresponse bias that excludes high out-of-pocket cost stays at a higher rate than less expensive stays. We found no evidence of a negative *response* bias in subsequent analyses (reported below); therefore, we might tentatively conclude that the Baseline estimate is subject to a sample bias as it excludes the higher expenses of decedents and the institutionalized (which are part of the SSA-based estimate).

Other Record Check Studies

Interest in household survey reports of out-of-pocket expenditures for inpatient care is recent, so there is not yet much accumulated record check evidence of response effects in face-to-face interviews. Three estimates of bias in reporting the total hospital bill are available and only one or two for the own cost reporting bias.

Barlow et al. (1960) found an average 5-percent positive bias in reporting the total hospital bill for matched cases and suggest that the difference might be due to a different treatment of personal expenses and blood transfusions by respondents and records. Including the checked cases that did not agree about an admission with the matched cases, Andersen et al. (1979) found no average bias in survey reports of the total hospital bill.[12] The Medical Economics Survey Methods

---

[11]These will be provided by the Health Insurance Study's health care demand experiment and the National Medical Care Expenditures Study.

[12]An earlier analysis (Kulley, 1974) of a subset of the Andersen et al. data, using regression and ignoring direction of errors, suggests that out-of-pocket costs are reported more accurately than total hospital costs. To be scored "accurate" on a dichotomous dependent variable, both the respondent and record had to report the cost

Study (HSRDC, 1977) suggests an average negative survey bias in face-to-face interviews of 8 percent of the total bill when the analysis is confined to verified cases and a 5 percent underreporting average using their full best-estimate data set. The three studies disagree on the direction of a potential response bias in reporting the total hospital bill but are consistent in suggesting that the size of any bias may be small (e.g., less than 10 percent).

Concerning own cost, an unpublished study (Bryant, 1962) found very similar distributions of the fraction of the bill paid by insurance obtained from a survey and from hospital records, suggesting the absence of a major net response bias in the interview data (Appendix Table A.5). The Medical Economics Study (HSRDC, 1977) provides the only other estimates of average survey response bias in reporting out-of-pocket payments for inpatient care. It is not clear whether these estimates were obtained using the full best-estimate data set or only data for the verified cases. The Baltimore sample that was interviewed face-to-face showed an apparent underestimation bias of 32 percent. The face-to-face rural Washington County sample average response bias estimate is *plus* 22 percent.

Baseline Record Check Results

The Baseline mean survey and record estimates of family out-of-pocket spending per hospital episode are similar in the matched case sample.[13]

The mean own cost estimates for matched cases that contain data from both sources are in Table 3.4. The hospital record estimate is the difference between the recorded total cost of the stay and the amount paid by insurance and other third parties directly to the

---

value within $50 of each other. The average accuracy score for total hospital cost appears to be about .5. For own cost, agreement accuracy appears close to perfect if insurance paid the whole hospital bill and closer to .7 if the own cost was nonzero.

[13] They are below the estimate using all stays reported in the Baseline. A Baseline outlier case was not included in the record check sample, and this appears to account for much of the mean difference.

Table 3.4

SURVEY AND HOSPITAL MEAN OUT-OF-POCKET COSTS

(Matched cases)

Source	Mean	Number of Stays
Survey	$55.40	247
Hospital	54.70	247
Difference	0.69	247

NOTE: The standard error of difference is $13.19.

hospital.[14] In a small number of cases, the difference is negative and has been truncated to zero. The Baseline estimate comes from answers to the direct question about how much the family paid out-of-pocket for each stay. Unpaid obligations may not have been included or they might appear as missing data. For the latter, we excluded the case from the analysis. The difference between the Baseline and hospital record means is small and not statistically significant, although this masks quite a bit of apparent disagreement between the sources on a case-by-case basis.

Using the model and assumptions in Marquis et al. (1976), about 40 percent of the measured variance in the survey net cost reports is random error, and 44 percent of the variance in hospital estimates of patient out-of-pocket cost is random error (Table 3.5). This indicates a substantial amount of disagreement between the sources on a case-by-case basis, but the disagreements tend to offset each other in the sample examined, so that the net bias is close to zero. The existing data cannot be used to determine whether the error variance originates in the survey/hospital reports themselves or is introduced later by, for example, transcription errors, linking errors, or processing mistakes. Because the estimates of error variance are similar

---

[14] Third-party reimbursement directly to the patient is not reflected in the records. It is rare but could cause an upward bias in the response bias estimate.

Table 3.5

PERCENTAGE OF ERROR VARIANCE IN OWN COST REPORTS
(Matched cases)

Source	Variance	Covariance with Difference	Estimate of Percentage of Error Variance: Upper Bound	Estimate of Percentage of Error Variance: Lower Bound
Survey	49,504	19,695	40	00
Hospital	53,072	-23,263	44	07
Difference	42,958			

NOTE: N = 247 matched stays. This excludes 80 matched cases with missing own cost data in either the survey or the record.

for both sources, the hypothesis that the error is introduced after the data acquisition stage should be seriously entertained. Indeed, the lower bound estimate of survey error variance is zero (Table 3.5).

DATE OF ADMISSION

The Baseline record check, as well as other published record checks, suggest that reported admission dates are telescoped forward on the average. The net telescoping effect is not large in any of the studies and is about one-fourth of a month in the matched case Baseline reports. This should not affect most analyses, including those that impute room and board charges based on the reported admission date. Estimates of random measurement error variance in the admission date data indicate that it is only a small percentage of the total measured variance in the matched case sample.

Other Record Check Studies

Other record check studies suggest that interview reports of admission or discharge dates are slightly telescoped forward in time for matched cases.

Using matched cases, Cannell et al. (1965) found that 11 percent of their respondents reported an admission month closer to the present

and 7 percent reported it farther back in time. Cannell and Fowler (1963, Procedure A) found the same bias trend for discharge date in a second record check study: 14 percent reported the month as more recent and 9 percent as less recent than the record entry. They also used a self-administered reporting procedure that encouraged the respondent to check his or her own records and found that the bias was removed but misclassification errors still remained (8 percent earlier, 8 percent later).

Three record check studies (Belloc, 1954; Cannell et al., 1965; and McNerney et al., 1962) suggest that date-reporting errors are minimal if the analyst needs to assign the admission to a broadly defined time period such as a calendar quarter. All three studies found that the survey report of admission/discharge date is within ± one month of the hospital report for 95 percent of the matched cases for which admission date information was available.

Baseline Record Check Results

We recoded the month of admission so that January 1973 had a value of 1 and December 1974 a value of 24. The difference between the means of the recoded survey and hospital distributions for matched cases was small but statistically significant. The data are given in Table 3.6.

On the average, respondent reports of month of admission are about one-fourth of a month closer to the date of the interview than hospital reports for the matched cases. Reported stays telescoped into the reference period are excluded from the matched case analysis, so the full extent of forward telescoping may not be reflected by the +.25-month estimate. On the other hand, stays unreported because they were telescoped out of the reference period are also excluded from the matched case analysis. Estimates of the percentage of measured survey and record variance that is random measurement error (Table 3.7) are low. The survey responses contain more random error than the record data.

Table 3.6

MEAN MONTH OF ADMISSION FOR SURVEY AND HOSPITAL REPORTS

(Matched cases)

Source	Average Month of Admission[a]	Number of Stays[b]
Survey	13.52	322
Hospital	13.27	322
Difference	.25	322

NOTE: Standard error of the difference is .09.

[a]Months are coded on an interval scale with 1 = January 1973 and 13 = January 1974. Most interviews were conducted in June and July 1974 (codes 18 and 19).

[b]Five matched stays are excluded because of missing information about month of admission.

Table 3.7

PERCENTAGE OF ERROR VARIANCE FOR ADMISSION MONTH REPORTS

(Matched cases)

Source	Variance	Covariance with Difference	Estimate of Percentage of Error Variance: Upper Bound	Estimate of Percentage of Error Variance: Lower Bound
Survey	13.15	2.03	15	09
Record	11.92	-0.80	7	00
Difference	2.83			

NOTE: N = 322 matched stays.

IV. DISCUSSION AND CONCLUSIONS

BASELINE SURVEY BIAS

Two kinds of estimates of Baseline survey bias have been given for each of four hospital stay variables. The bias estimates are summarized here along with my opinion about how to interpret them.

Annual Per-person Admission Rate

Comparison with other surveys: The Baseline survey rate appears to be a slight underestimate of the population rate.

Record check: The Baseline survey respondents appear to overreport stays much more often than they underreport stays. In the absence of other information, this would imply that the annual admission rate would be overestimated.

Conclusion: The true response bias is probably small. The record check was not designed to estimate the net response bias accurately; random response and processing errors are mistaken for positive response errors. The Baseline survey annual per-person admission rate slightly underestimates the population rate because the sample excluded persons who died during the retrospective reference period and who were likely to have been hospitalized.

Length of Hospital Stay

Comparison with other surveys: The Baseline survey (and the Health Interview Survey) responses appear to overestimate the average length of stay by about 0.5 to 0.6 of a day compared with Hospital Discharge Survey data.

Record check: The Baseline survey responses overestimate the duration of the average hospital stay by 0.7 day (S.E. = 0.2 day). This is similar to findings from other record check studies.

Conclusions: The two kinds of estimates are similar, and the net positive Baseline response bias may be due, in part, to differences between the respondents' and the records' definition of length of stay.

Own Cost of Stay

Comparison with independent estimate: The Baseline estimate is substantially below an estimate derived from data compiled by the Social Security Administration.

Record check: The net Baseline response bias is very small and not significantly different from zero.

Conclusion: Uncertainty about the magnitude of the Baseline bias remains because of problems with both kinds of estimates. The SSA-based estimate may contain very large errors; the record check estimate excludes many stays for which family out-of-pocket cost information was unavailable. Response biases for these stays may be different from those included in the record check. Nevertheless, when it was possible to check response validity, no net response bias was found. A limited conclusion is that out-of-pocket-cost survey data were unbiased if both the survey and record agreed about the existence of the stay and were able to furnish own cost estimates.

Admission Month

Baseline record check data suggest that there is a slight forward telescoping of the admission date; this result is similar to findings of other record check studies. No external survey estimate of month of admission distribution was available to compare with the Baseline survey estimate.

BASELINE RESPONSE ERROR VARIANCE

The record check data were used to compute the following bounds on response error variance as a proportion of measured variance for the three attributes of reported hospital stays:

	Lower Bound	Upper Bound
Length of stay	.13	.23
Own cost	.00	.40
Month of admission	.09	.15

The estimates were made using only the matched stays. They suggest that a moderate amount of "noise" exists in the reported length and admission date information. Uncertainty remains about how much response error variance is in the own cost reports; it could be a little or a lot depending on the unknown quantity of error variance introduced in transcribing, matching, and other processing of the matched case data.

IMPLICATIONS FOR USE OF BASELINE DATA

The effects of the response biases and response error variances inferred for Baseline hospital stay data will depend on how the data are used. In general, response biases affect estimates of the first moment of the population distribution (e.g., the mean or the constant term in a regression model), while the response error variance influences statistics that depend upon unbiased estimates of the true score variance (e.g., the correlation coefficient).[1]

The Health Insurance Study will use health care price and consumption data from the subsequent demand experiment to estimate a price regression coefficient in a prediction equation that employs Baseline hospitalization and other variables as "covariates." The covariates will be included to increase the precision of the price coefficient estimate. The Finite Selection Model (Morris, 1975), used to allocate families to price treatments, assures orthogonality between the covariates and the price treatments, so any measurement bias in hospitalization covariates will not bias the price coefficient estimate. The moderate amounts of response error variance in Baseline hospitalization variables will bias the coefficients on Baseline hospitalization stay predictor variables toward zero. The covariates, however, can still be expected to explain some of the consumption variance and, hence, help to increase the precision of the estimated price coefficient.

---

[1] For more detailed discussions of the link between response effects and statistics, see Cochran (1968), Marquis and Marquis (1977b, Section III), Marquis (1977, pp. 11-18), or Marquis et al. (1979, pp. 11-18).

Other analysts may use Baseline hospital data for dependent variable measures in demand estimation regression equations. For this use, main effect response biases and response error variance will not affect estimates of (unstandardized) regression coefficients; precision of the estimates, however, will be affected. If the Baseline data are used to construct price predictors, to estimate "intercepts," for stratified sampling, or to estimate site means for comparison with national data, the response biases and error variances become more relevant.

IMPLICATIONS FOR RECORD CHECK DESIGNS

This research demonstrates the pitfalls of relying on incomplete record check designs to estimate the size and sign of the survey response bias. An uncritical acceptance of the record check results would lead to the conclusion that Baseline respondents overreported their hospital stays by as much as 26 percent, yet other evidence suggests that the response bias was small and possibly negative. Future record check designs should include features to obtain unbiased estimates of both underreporting and overreporting so that more realistic assessments of net survey bias are possible.

A second problem concerns the inability to partition observed survey/record discrepancy variance into its sources. In this research, it was not possible to infer the extent to which response, record, and processing mistakes were responsible for the disagreements between survey and record information. The problem results from underidentified error models; in this case, the models allowed for processing errors and the possibility of imperfect record information. This creates too many parameters for estimation with available kinds of observations. Future bias estimation designs should include additional kinds of observations, such as retest estimates of response, record, and processing error variance, so that estimation assumptions and restrictions can be relaxed in favor of empirical parameter estimates. There is a limit, of course, on how much can be estimated empirically. As with any question about truth, it will always be necessary to assume, without direct empirical support, the characteristics of the true score distribution for one source of data.

IMPLICATIONS FOR "BEST-ESTIMATION" STRATEGIES

An alternative use of record check data is to substitute information furnished by the hospital for that reported by the family. Several "best-estimate" strategies can be considered for the HIS record check data, some better than others.

An extreme approach is to use only hospital record data in an analysis. The effect of this procedure is to reduce the number of stays in the analyzed data set, thus reducing precision and biasing the mean of the annual per-person distribution of stays. The hospital data will be more accurate with respect to nonzero length of stay and month of admission values. However, if the response error effects removed are main effects or "constant" biases, the accuracy of a regression coefficient estimate may not be improved. A direct substitution of hospital data for interview data, then, will probably reduce estimation precision and bias estimates of central tendency, but leave unstandardized regression coefficients relatively unbiased.

A second best-estimate strategy is to select the "best" available data when the interview and records disagree. There are at least two kinds of risks in this approach. The first is that the wrong choice is made and erroneous data are included instead of accurate data. If the choice of best data is really random and both sources of potential data are equally erroneous, one might not lose precision or increase bias by the procedure. Exercising systematic selection rules could result in poorer quality data if the selection rules are wrong. A second kind of risk enters when part of an individual's survey data is replaced by hospital data (e.g., length of stay, given that the admission is reported in the survey). The substantial number of disagreements between survey and record reports of admissions and of characteristics of the stay suggests the possibility of large amounts of clerical error (transcription mistakes, mixing up identification labels, reading from the wrong record, etc.). Hospital record data may be very accurate at one level but sometimes may be inaccurately transcribed or linked to the incorrect person or stay in record check studies. Substituting these kinds of data for survey reports may introduce large amounts of random error. Even substituting hospital reports for items the family

could not report (survey missing data) may be less desirable than more conventional imputation strategies.

A third best-estimate strategy is to combine both survey and hospital information into a single minimum variance estimate. Although this strategy has been recommended in the past (Marquis et al., 1976), we should await better information about the true source of the errors found in the Dayton Baseline record check data before adopting it. If the error is introduced in the record check matching and transcription procedures (that is, it is really not in the survey or record data), a combination best-estimate strategy may create noisier data than that from the survey alone.

A fourth approach would be to retain the highest value available from either the survey or the hospital record whenever the two disagreed. Unfortunately, this procedure does more than eliminate any survey underreporting bias; it also retains any survey overreporting bias and, more important, converts up to half of the random error variance (regardless of source) into a positive measurement bias.

A final best-estimate strategy would be to use only the survey data, ignoring the record observations completely. If my earlier conclusions are correct, this is the better strategy because the survey response biases are negligible, and combining the survey and record observations will introduce both additional error variance due to match errors and bias due to the one-directional record check procedure. As a reviewer pointed out: "It may well be the case that for hospital stays the most cost-effective procedure is to use only survey data."

APPENDIX TABLES

Table A.1

AGE WEIGHTS USED TO ADJUST ADMISSIONS/DISCHARGE ESTIMATES

Age Category (years)	Weight	Age Category (years)	Weight
0-14 15-44 45-64	.29 .48 .23	0-16 17-44 45-64	.34 .43 .23
Sum	1.00	Sum	1.00

SOURCE: U.S. Bureau of the Census, *Current Population Reports*, Series P-25, No. 539, January 1975. The data are consistent with estimates published in Ries (1975) and Ranofsky (1976).

NOTE: Means or proportions for each age category are estimated from a data source, multiplied by the appropriate age weight from the table, and the products summed to produce an estimate for the entire sample or population.

Table A.2

ANNUAL PER CAPITA HOSPITALIZATION ESTIMATE, FROM THE 1973-1974 HOSPITAL DISCHARGE SURVEY

Criteria	Patient Age 0-14	15-44	45-64	Total under Age 65
Number of discharges,[a] North Central U.S.				
1973[b]	1,318	4,110	2,438	7,866
1974[c]	1,285	4,352	2,542	8,179
Average	1,302	4,231	2,490	8,023
Civilian, noninstitutionalized population North Central U.S.[a]				
1973[b]	15,538	24,681	11,517	51,736
1974[c]	15,076	24,452	11,472	51,000
Average	15,307	24,566	11,495	51,368
Preliminary annual discharge rate (discharges ÷ population)	.085	.172	.217	.156
Age adjustment weights	.29	.48	.23	--

Age adjusted estimate (annual discharges per person)	.157
Plus 4 percent for discharges of civilians from federal military hospitals[d]	+.006
Final adjusted HDS estimate	.163

[a]In thousands.

[b]Lewis (1976).

[c]Ranofsky (1976).

[d]From Croner (1977).

Table A.3

ANNUAL PER CAPITA HOSPITALIZATION ESTIMATE, FROM THE 1973-1974 NATIONAL HEALTH INTERVIEW SURVEY

Criteria	Patient Age			Total under Age 65
	0-16	17-44	45-64	
Number of overnight discharges, U.S.[a]				
1973[b]	4,503	12,312	7,073	23,888
1974[c]	4,437	12,133	7,484	24,054
Estimated number of discharges for stays of less than overnight[a]				
1973[d]	185	406	141	731
1974[e]	182	376	127	685
Adjusted average discharges[a]	4,654	12,614	7,412	24,679
1973[b]-1974[c] average U.S. civilian, noninstitutionalized population[a]	63,477	79,899	42,699	186,075
Discharges per person	.0733	.1579	.1736	.1326
Age weight	.34	.43	.23	--
Preliminary estimate				.133
Plus adjustment for small SMSA, north central region[f]				+.008
Final adjusted estimate				.141

[a]In thousands.

[b]Wilder (1974).

[c]Ries (1975).

[d]Lewis (1976).

[e]Ranofsky (1976).

[f]Based on Namey and Wilson (1974).

Table A.4

ANNUAL PER CAPITA HOSPITALIZATION ESTIMATE, FROM THE DAYTON BASELINE

Criteria	Patient Age 0-16	17-44	45-64	Total under Age 65
Number of admissions reported	107	307	154	568[a]
Number of admissions for newborns	-7	--	--	-7
Adjusted number of admissions	100	307	154	561
Number of persons	1502	1907	837	4246
Annual admission rate per person	.0666	.1610	.1840	.1321
Age adjustment weights	.34	.43	.23	--
Adjusted baseline estimate of annual admissions per person under age 65				.134

[a]Potentially includes stays of four persons in the institutionalized population.

Table A.5

PROPORTION OF HOSPITAL BILL PAID BY INSURANCE, FROM HOUSEHOLD SURVEY AND HOSPITAL RECORD DATA

Proportion of Bill Paid by Insurance	Percent Distribution of Hospital Stays: Household Survey Data	Hospital Record Data
None	18	20
Less than 1/2	5	4
1/2 to 3/4	12	11
More than 3/4	63	64
Other	3	0
Total	100	100

SOURCE: Bryant (1962).

NOTE: Based on a weighted analysis of 1590 recorded episodes, excluding records with missing insurance data but including recorded stays not mentioned in the survey. Distributions do not sum to 100 percent because of rounding.

REFERENCES

Andersen, Ronald, Joanna Kravitz, Odin W. Anderson, and Joan Daley, *Expenditures for Personal Health Services, National Trends and Variations (1953-1970)*, DHEW Publication No. (HRA) 74-3105, 1973.

Andersen, Ronald, et al., *Total Survey Error*, Jossey-Bass, San Francisco, 1979.

Balamuth, Eve, Sam Shapiro, and Paul M. Densen, "Health Interview Responses Compared with Medical Records," *Vital and Health Statistics*, Series D, No. 5, June 1961.

Barlow, Robin, James Morgan, and Grover Wirick, "A Study of the Validity in Reporting Medical Care in Michigan," *Proceedings of the Social Statistics Section*, American Statistical Association, August 1960, pp. 54-65.

Belloc, Nedra B., "Validation of Morbidity Survey Data by Comparison with Hospital Records," *Journal of the American Statistical Association*, Vol. 49, 1954, pp. 832-846.

Bryant, E. Earl, "Proportion of Hospital Bill Paid by Insurance--A Comparison of Household Interview Results with Hospital Records," unpublished manuscript, National Center for Health Statistics, July 20, 1962. Available from the Survey Methodology Information System, U.S. Bureau of the Census, Document 72 257 024.

Cannell, Charles F., and Floyd J. Fowler, "Comparison of Hospitalization Reporting in Three Survey Procedures," *Health Statistics*, Series D, No. 8, January 1963; reprinted in *Vital and Health Statistics*, Series 2, No. 8, July 1965.

Cannell, Charles F., Gordon Fisher, and Thomas Bakker, "Reporting of Hospitalization in the Health Interview Survey," *Vital and Health Statistics*, Series 2, No. 6, July 1965.

Cochran, W. G., "Errors of Measurement in Statistics," *Technometrics*, Vol. 10, 1968, pp. 637-666.

Croner, Charles M., *The Nation's Use of Health Resources, 1976 Edition*, DHEW Publication No. (HRA) 77-1240, 1977.

Feather, Joan, *A Response/Record Discrepancy Study*, University of Saskatchewan, Saskatoon, November 1972. Available from the National Technical Information Service, Springfield, Virginia.

Glickman, Linda S., "Inpatient Utilization of Short-Stay Hospitals by Diagnosis, United States--1974," *Vital and Health Statistics*, Series 13, No. 30, July 1977.

HSRDC (Health Services Research and Development Center), *Medical Economics Survey-Methods Study, Final Report*, Contract HRA 106-74-150, The Johns Hopkins Medical Institutions, Baltimore, Md., August 1977.

Kulley, Andrew Mark, *The Validity of Survey Measurement of Health Services Utilization: A Verification Study of Respondent Reports of Hospitalization*, unpublished Ph.D. dissertation, Purdue University, July 1974.

Lewis, W. Frank, "Utilization of Short-Stay Hospitals; Summary of Nonmedical Statistics--United States 1973," *Vital and Health Statistics*, Series 13, No. 23, July 1976.

Marquis, Kent H., *Record Check Validity of Survey Responses; A Reassessment of Bias in Reports of Hospitalizations*, The Rand Corporation, R-2319-HEW, May 1978.

Marquis, Kent H., *Some Implicit Assumptions in Survey Measurement Error Estimation*, The Rand Corporation, P-6341, June 1979.

Marquis, Kent H., *The Methodology Used To Measure Health Care Consumption During the First Year of the Health Insurance Experiment*, The Rand Corporation, R-2126-HEW, August 1977.

Marquis, K. H., M. S. Marquis, and Joseph P. Newhouse, *The Measurement of Expenditures for Outpatient Physician and Dental Services: Methodological Findings from the Health Insurance Study*, The Rand Corporation, R-1883-HEW, April 1976.

Marquis, K. H., J. E. Ware, Jr., R. H. Johnston Jr., M. S. Marquis, M. E. Michnich, B. J. Rose, J. A. Stein, and C. T. Veit, *Appendix B to Summary Report: An Evaluation of Published Measures of Diabetic Self-Care Variables*, The Rand Corporation, N-1152-HEW, June 1979.

Marquis, M. Susan, and Kent H. Marquis, *Measurement Evaluation of the Health Insurance Study Screening Interview*, The Rand Corporation, R-2107-HEW, November 1977a.

Marquis, M. Susan, and Kent H. Marquis, *Survey Measurement Design and Evaluation Using Reliability Theory*, The Rand Corporation, R-2088-HEW, June 1977b.

McNerney, Walter J., et al., *Medical and Hospital Economics*, Hospital Research and Educational Trust, Chicago, 1962.

Moien, Mary, "Inpatient Utilization of Short-Stay Hospitals by Diagnosis, United States--1973," *Vital and Health Statistics*, Series 13, No. 25, October 1976.

Morris, Carl, "A Finite Selection Model for Experimental Design of the Health Insurance Study," *Proceedings of the Social Statistics Section*, American Statistical Association, 1975, pp. 78-85.

Mueller, Marjorie Smith, and Robert M. Gibson, "Age Differences in Health Care Spending, Fiscal Year 1975," *Social Security Bulletin*, Vol. 39, June 1976, pp. 18-30.

Mueller, Marjorie Smith, and Paula A. Piro, "Private Health Insurance in 1974: A Review of Coverage, Enrollment, and Financial Experience," *Social Security Bulletin*, March 1976, pp. 3-20.

Namey, Christy W., and Ronald W. Wilson, "Health Characteristics by Geographic Region, Large Metropolitan Area, and Other Places of Residence," *Vital and Health Statistics*, Series 10, No. 86, January 1974.

Neter, John E., Scott Maynes, and R. Ramanathan, "The Effect of Mismatching on the Measurement of Response Errors," *Journal of the American Statistical Association*, Vol. 60, 1965, pp. 1005-1027.

Newhouse, Joseph P., and Rae W. Archibald, *Overview of Health Insurance Study Publications*, The Rand Corporation, P-6221, November 1978.

Ranofsky, Abraham L., "Utilization of Short-Stay Hospitals: Annual Summary for the United States, 1974," *Vital and Health Statistics*, Series 13, No. 26, September 1976.

Richardson, Arthur H., and Howard E. Freeman, "Evaluation of Medical Care Utilization by Interview Survey," *Medical Care*, Vol. 10, 1972, pp. 357-362.

Ries, Peter W., "Current Estimates from the Health Interview Survey, United States--1974," *Vital and Health Statistics*, Series 10, No. 100, September 1975.

Simmons, Walt R., "Analysis of Bias in Interview-Reported Hospital Experience," *Bulletin of the International Statistical Association*, 1967, pp. 1-39.

Sutton, Gordon F., "Hospitalization in the Last Year of Life--United States 1961," *Vital and Health Statistics*, Series 22, No. 1, September 1965.

Wilder, Mary H., "Current Estimates from the Health Interview Survey, United States--1973," *Vital and Health Statistics*, Series 10, No. 95, October 1974.

Worthington, Nancy L., "National Health Expenditures, 1929-1974," *Social Security Bulletin*, Vol. 38, February 1975, pp. 3-20.