HOUSEHOLD AND MEDICAL PROVIDER REPORTS ON MEDICAL CONDITIONS
National Medical Expenditure Survey, 1987

Ayah E. Johnson, Maria Elena Sanchez, AHCPR
5600 Fishers Lane, Rm. 18A-55, Rockville, MD 20857

1.0 Introduction

The source of information on a medical patient's conditions has great bearing on the quality and the reliability of the data being collected. Information on medical conditions affecting specific persons could be obtained through any of the following sources: (1) the health professionals providing medical care to the individuals; (2) the health insurance company paying for medical care can be asked for a copy of the diagnosis reports related to the bill; or, (3) the person is asked, by direct interviewing methods, to provide reasons for seeking medical care or to report on medical conditions experienced during a predetermined period of time. The last two sources are relaying information that was communicated to them by medical providers and, thus, reporting error is likely to be higher. Survey answers in particular depend on understanding, recall, and whether the household respondent feels comfortable discussing the medical condition with the interviewer.

Medical and insurance records are likely to provide data on medical diagnoses that are more reliable than that provided by the household respondent, assuming that the information can be obtained. Despite that, these sources can provide information only for subsets of the U.S. civilian noninstitutionalized population: those who seek medical care, or those covered by private or public health insurance and who use medical care. When the objective is to quantify the prevalence of disease or the cost of illness, collecting data from medical providers or insurance companies can lead to an underestimate of the statistics of interest. Prevalence or cost estimates from these sources are conditional on the fact that people sought medical care, or had health insurance coverage, which is not always the case. There are people who are ill, are not seeking medical care, but should be counted when estimating prevalence. These same persons may be experiencing work loss days which should be included in estimating cost of illness. Therefore, in order to get unbiased estimates of health characteristics, the household respondent becomes the source of interest. By using the household respondent as the source of information we alleviate the problem of population under-coverage but exacerbate the amount of response error due to different types of uncertainties such as: (1) recall of the exact nature of the medical conditions; (2) item missing data in cases when the person does not know the medical condition, or fails to report the condition due to social stigma (e.g. alcoholism); (3) inadequate questionnaire design which may elicit insufficient information to attribute an ICD-9 code to a medical condition; (4) inadequate interviewer probing resulting in incomplete information about the number of conditions; (5) issues of code structure and coding procedures; and (6) error due to proxy reporting. The quality and the amount of information elicited from household respondents must be such that the medical coder is able to translate the verbatim response to exact 4 digits ICD-9 (International Classification of Diseases, Version 9) numeric codes used in analysis of data.

1.1 Study Objectives

The main objective of this analysis is to compare household and medical provider reports on medical conditions collected by the National Medical Expenditure Survey (NMES, 1987). NMES collected medical conditions from two sources: the household (HHS survey) and medical providers who cared for HHS persons (MPS survey).

Household respondents were asked about the reasons and "condition or problem" that caused the person's medical visit or hospital stay, while medical providers were asked to supply "diagnoses" associated with a particular visit or hospitalization. The classification scheme, rules and procedures used to code condition data from households and medical provider diagnoses were identical. The scheme is the 9th Revision of the International Classification of Diseases (ICD-9), as revised for use in the National Health Interview Survey (NHIS). Although the coding scheme and procedures were identical for both sources of data, some level of disagreement is likely to remain due to the potentially different information elicited from the different questions asked in HHS and MPS, respectively.

The MPS survey dealt with medical encounters experienced by persons selected from the NMES-HHS national area probability sample of households. This special sample of persons consisted of a 25 percent national sample of households reporting care by M.D.'s and D.O.'s, and an additional sample of households with persons who were either less likely to provide medical cost information (e.g. Medicaid-covered persons) or who had experienced medical events in facilities such as hospitals and clinics. A statistical matching algorithm (CANLINK) was used to match the household and the medical provider databases. As in any statistical matching, this is subject to type I and type II errors. These errors were estimated to be at 5 percent and 13 percent, respectively (Johnson and Carlson, 1991). Although in theory levels of disagreement may be due to errors in matching, the likelihood of such a problem is not great since both type I and type II matching errors are relatively low.

This analysis is not specific to a given condition, such as breast cancer, or myocardial infarction. It is a comparison of a set of conditions from a sample for which we have both household and medical reports. This comparison can help us identify the medical conditions that household reporters can reliably describe and those that they cannot. The Assessment of the degree of agreement between reports obtained from households and medical providers is needed to validate the accuracy of survey data, and to determine whether it is feasible to attribute medical diagnoses to household reported visits for which no medical provider data were obtained.

2.0 Household and Medical Provider Diagnosis Reports-NMES, 1987

For the majority of the events where there are both HHS and MPS data (82.6 percent) we have condition data from both the household and the medical provider. In 12.8 percent of the cases we have only the medical provider data while in

3.2 percent of the cases we have household reported conditions but no medical data. The rate of missing data from both sources amounts only to 1.4 percent.

A hypothesis raised in the literature is that for less complex conditions the rates where both reports exist will be high, and the rate of missing data will be low. For more complex conditions, or for more stigmatized conditions such as mental health, the rates of missing data should be high. Pregnancy reports are examples of a less complex condition, and diagnoses involving mental disorders, injury and poisoning are examples where one would expect a lower rate of reporting.

A second reporting issue discussed in the literature is the number of complaints (conditions) for a given medical event. Medical providers have been shown to report a greater number of conditions than household reporters (Harlow and Linet, 1989). This difference may be partly due to the fact that respondents give the reason for the medical visit, while the provider notes the outcome of the visit, or underlying medical conditions. Underprobing on the part of survey interviewers may also serve to depress the number of mentions collected from household reporters.

Other issues related to reporting patterns include whether the household respondent is less willing or able to discuss certain medical conditions, and whether certain medical conditions are over-represented among visits that lack a provider diagnosis altogether. Reticence or inability to report on the part of the household respondent can be studied by examining the distribution of medical conditions reported by providers for cases where the household respondent did not provide any information. The medical provider information will allow us to identify the types of conditions that the household respondent cannot or will not discuss. Similarly, inspecting cases that lack provider information for a household reported condition, will allow us to examine whether there are clusters of conditions that medical providers report with less regularity than household respondents.

2.1 Household Reports: Condition Data Missing

The type of conditions that were not reported (or coded) in the household survey do not seem to cluster around stigmatized or sensitive conditions. Only 2 percent of the cases with condition data on the household side are associated with mental disorders, and this proportion is identical to that obtained from MPS data for household events lacking a condition report in HHS. Similarly, household events with a reported condition include a greater proportion of potentially sensitive injury and poisoning reports (13 percent) than the proportion obtained from MPS data for control HHS cases without any condition reported (2 percent). These results indicate that the sensitivity of the condition does not seem to affect the propensity of the household to respond. As expected, for conditions originating in the perinatal period the rate of missing data is zero. The analysis shows that, basically, except for V-codes (classification of factors influencing health status and contact with health services) and symptoms, there is no specific highly visible cluster of conditions that household respondents failed to report regularly. For both V-codes and symptoms, the differences in proportions are statistically significant and large in magnitude, unlike the observed differences in complications of pregnancy where a statistical significance is also present but the magnitude of the difference is small. There is a tendency for household events with a coded condition to overrepresent problems of the circulatory and the respiratory systems

compared to household records lacking a coded condition. Similar results for respiratory conditions are noted for several studies reviewed by Harlow and Linet, 1989 and are attributed to differences in reporting style between patients and medical providers.

2.2 Medical Provider Reports: Diagnosis Data Missing

In 4.7 percent of the events no medical diagnosis was obtained from medical providers. The distribution of conditions for household events without medical provider data, and the comparison to household events which do have the medical provider data, yield results similar to those documented above. Compared to household reports, V-codes and symptoms are over-represented, and disease of the circulatory system is under-represented, in medical provider reports.

There is no evidence that the rate of missing diagnosis data vary by the medical setting where the person was treated for either household or medical provider reports.

2.3 Number of Reported Answers

In NMES both the household respondent and the medical provider were given the opportunity to report up to 8 medical conditions as reasons for the medical visit. Approximately 99 percent of household respondents or medical providers gave no more than four diagnoses. Table 2 (available from the authors) summarizes the reporting differential between the household respondent and the medical provider in terms of the number of conditions reported.

The more pronounced differential between the two reporters is in the number of conditions that are reported for a given medical event rather than the refusal or the inability to report medical conditions associated with health events. In 32 percent of the events, after excluding cases with any missing data, both the household respondent and the medical provider reported the same number of conditions.

In 65 percent of the cases the medical provider reported more conditions than the household respondents and only in 3 percent of the cases do we see the household reporting more medical conditions than the medical provider. Furthermore, only 10 percent of household reports involve more than one condition given.

This indicates that additional probing by the interviewers, a revision to the structure of the questions, or increasing the level of specificity asked about the medical condition in the household survey may be required in future surveys in order to code more than one condition on the household side. An increase in the number of household reported conditions will increase the probability that the two reporters describe similar conditions underlying the same visits.

3.0 Probability of Agreement between Household and Medical Provider

The determination of the agreement probability in reporting medical conditions is subject to the level of specificity required for the medical condition and the constraints on the type of agreement that we wish to obtain. The latter stems from the fact that we have more than one condition reported, and that there is no easy way of determining which of the reported conditions is the most salient or the most severe. Thus, in this paper, four definitions of agreement are used:

1. AGR3DIGT1 = 1 any of the household conditions match
   any of the medical provider diagnoses
   0 otherwise

2. AGR3DIGT2 = 1 the first condition reported by
   household matches the first diagnosis
   reported by the medical provider
   0 otherwise

3. AGR3DIGT3 = 1 the first condition reported by the
   household respondent matches any
   diagnosis reported by the medical
   provider
   0 otherwise

4. AGR3DIGT4 = 1 the first diagnosis reported by the
   medical provider matches any of the
   conditions reported by the household
   respondent.
   0 otherwise

The various definitions are used to capture various degrees of agreement between the data reported by the two sources. The first is the least restrictive, and should in theory yield the highest level of agreement between the two reporters. The second definition is the most restrictive and assumes that the most salient condition or diagnosis is listed first by both reporters, and should probably yield the lowest level of agreement. The third and the fourth definition assumes that the first reported condition is most salient; however it allows the other reporter up to four chances to identify the same condition. In the third definition we allow the household report to be "truth", while in the fourth definition we use the conventional definition of "truth" as that reported by the medical provider. These four definitions of agreement were implemented using the detailed 3-digits ICD-9 condition codes, and two other classifications that collapsed these 3-digits codes into major grouping of diseases. The first grouping collapsed the 3-digits codes to 20 categories of diseases by major organs or bodily system affected. The second classified the 3-digits codes into a grouping of 131 detailed diseases within the 20 bodily systems. By classifying the data into broader groups the levels of agreement ought to increase; however, the level of precision of the descriptive medical condition decreases rapidly. For example in the recode into 20 groups we are able to identify only the biological system that is being affected, such as the nervous or the respiratory system.

Table 3 summarizes the observed levels of agreement using each of the four agreement definitions, and the three medical classification groups referenced above. As expected, the highest level of agreement is achieved when we allow any of the household reported conditions to match any of the medical provider diagnoses; if the first diagnosis mentioned by the medical provider is assumed to be the most salient, then the level of agreement decreases by 11.7 percent (39.6-27.9). Since the household respondent reports one condition in the majority of the cases, agreement rates involving the first HHS mention yield results similar to companion rates based on any HHS mention.

Table 3 documents that the level of agreement between the two report sources increases when we examine the collapsed condition codes. The last column shows that physicians and household respondents are able to identify the same biological system that is affected by the illness, at a level of agreement which is 28.6 percent higher than the level of agreement observed with the detailed 3-digits code. The level of agreement with the most relaxed definition of agreement and the greatest collapsing of codes is still only 68.2 percent. After segregating medical events associated with pregnancies from those that are non-pregnancy related, the levels of

agreement change in magnitude but only for pregnancies. At the 3-digits level and for the most relaxed definition of agreement the rate is 1.7 fold the rate for non-pregnancy events.

Table 3. Rates of Agreement in Medical Condition Reports

	PERCENT OF CASES IN AGREEMENT		
Definition of	Classification of Medical Conditions		
Agreement	3-digit Codes	131-Groups	20-Groups
Overall			
ANY HHS TO ANY MPS	39.6	53.7	68.2
FIRST HHS TO FIRST MPS	26.3	38.4	52.3
FIRST HHS TO ANY MPS	37.5	51.3	65.8
FIRST MPS TO ANY HHS	27.9	40.3	54.9
Non-pregnancy			
ANY HHS TO ANY MPS	38.2	52.8	66.9
FIRST HHS TO FIRST MPS	27.2	39.9	53.4
FIRST HHS TO ANY MPS	36.0	50.4	64.5
FIRST MPS TO ANY HHS	28.9	42.2	56.0
Pregnancy			
ANY HHS TO ANY MPS	65.3	69.2	91.0
FIRST HHS TO FIRST MPS	8.9	11.9	33.0
FIRST HHS TO ANY MPS	64.1	67.7	89.4
FIRST MPS TO ANY HHS	9.7	12.6	35.3

HHS -- household reported medical condition
MPS -- medical provider reported conditions

Again when collapsing the condition codes into the 20 major groups the rate of agreement increases; for pregnancies it hits and expectedly high rate of 91 percent. It should be noted that for non-pregnancy visits the level of agreement does not exceed 66.9 percent, even when using the least restrictive definition of agreement, and the most extensive recode of the medical condition. This difference in reporting suggests that the household respondent has in many cases a different perception of the reason for the visit than the medical provider. It is difficult without a structured study to attribute a specific reason for the low level of agreement. It might be due to: (1) recall; (2) lack of communication between doctor and patient; (3) interviewer or medical coders' error; and potentially (4) errors due to the statistical matching. In statistical comparisons (tables are available from the authors) the level of agreement for pregnancy events is significantly different than for non-pregnancy events.

The higher rates of agreement are found when allowing all conditions reported by the medical provider to factor into the definition of agreement. The lower levels of agreement are found when we restrict the definition of agreement to only the first condition reported by the medical provider. This reinforces the hypotheses that the medical provider reports a set of related outcomes describing the medical evaluation rather than a reason for the visit. Moreover, by examining various medical conditions listed for a medical event we found that many combinations of ICD-9 codes could be reported by the medical provider, and it is the vector of outcomes that describes the visit rather than the order in which the medical condition is reported. Given these observations we have decided to conduct further analysis using only the first definition of agreement which allows any household reported condition to match to any condition reported by the medical provider.

3.1 Rates of Agreement in Condition Reporting by Type of Event

NMES classified health visits to medical providers by "type of event". Household event types included doctor visits, emergency room and outpatient visits, and hospital stays. For

the Medical Provider Survey, NMES distinguished between medical care provided in an ambulatory setting, in a hospital outpatient setting and during hospital stays. In both surveys pregnancies were handled separately. The comparisons of the level of agreement is done within types of events as defined in the Household Survey, and the objective is to assess whether recall or accuracy of the reported condition is associated with the setting or the severity of the condition. A surrogate measure of the severity of the condition in this case is whether the patient received hospital care or just sought ambulatory care.

The total number of medical encounters reported, for which we have both household and provider questionnaires is 33,514. In 64 percent of these cases the visits are doctor visits, 12 percent are emergency room visits, 14 percent are hospital outpatient visits and 10 percent are hospital stays. Table 4 summarizes the rates of agreement allowing any household condition to match to any medical provider reported conditions. In this tabulation we have broken down the events to pregnancy and non-pregnancy events.

Table 4 Rates of Agreement by Type of Event

Type of Event	PERCENT OF CASES IN AGREEMENT		
	Total	Non-Pregnancy	Pregnancy
3-digit codes			
Total	39.6	38.2	65.3
Emergency Room Visits	43.2	43.0	61.4
Ambulatory Visits	37.0	37.8	86.0
Outpatient Visits	40.2	34.5	77.4
Inpatient Stay	36.6	40.2	13.4
131-groups			
Total	53.7	52.7	69.2
Emergency Room Visits	55.4	55.3	65.0
Ambulatory Visits	54.1	52.4	87.4
Outpatient Visits	49.5	47.6	81.6
Inpatient Stay	54.9	59.9	22.4
20-groups			
Total	68.2	53.4	91.0
Emergency Room Visits	73.4	61.5	73.7
Ambulatory Visits	67.2	51.7	94.2
Outpatient Visits	61.2	48.1	89.1
Inpatient Stay	78.0	61.5	87.3

The pattern and the magnitude of the rates of agreement do not change radically when controlling for the place where medical care was given. There is hardly any difference between the overall rates of agreement and the rate of agreement for non-pregnancy events; this basically reflects the fact that the number of pregnancy cases is relatively low. For emergency room visits and inpatient stays involving non-pregnancy events, the levels of agreement are slightly higher than for ambulatory and outpatient visits-- 43 and 40.2 percent compared to 37.8 and 34.5 percent respectively. Although slight in magnitude, these differences are statistically significant.

For pregnancy-related visits, the levels of agreement for all events except inpatient stays is, as expected, significantly higher than for non-pregnancy related visits. It is between 1.4 to 2.4 folds higher. The level of agreement (3- digits codes) for inpatient stays associated with pregnancies, 13.4 percent, is unusually low; this rate remains low, at 22.4 percent even when recoding the conditions into 131 major groups. It is only when the conditions are grouped into 20 major categories that the agreement level for pregnancy-related inpatient stays rises to what is expected for such a medical condition. We plan to investigate the matter more carefully.

When comparing the overall rates of agreement between every two medical settings, the differences in rates, although small in magnitude, are almost always statistically significant except for the following paired comparisons: at the 3-digit coding level, ambulatory setting vs inpatient setting (Z=.45); at the 131 grouping, emergency room visits vs ambulatory visits (Z=1.52), and hospital stay vs emergency room (Z=-0.43) and ambulatory visits (Z=0.87); at the 20-groups classification, the only nonsignificant difference is noted when comparing levels of agreement between emergency events and inpatient stays.

3.2 Agreement in Reporting Medical Conditions within Broad Levels of Diagnoses

One can assume that when medical conditions are visible, recognizable, or can be easily communicated by providers, the household respondent should be able to report them with higher reliability. By looking at broad levels of conditions we will be able to test whether this hypotheses holds. This in turn should enable us to identify classes of conditions that could be predicted using household reported data. It will also guide future work in questionnaire design that might help obtain more accurate information from household respondents.

Two groups of conditions are analyzed. The first focuses on biological organs that are affected by the condition. These are basically the 20-groups of diseases. The second group of conditions discussed under 3.2.2 was constructed based on a list of medical conditions or symptoms that both NMES and the Health Interview Survey identify by name in the interview, thus assuming that they are common and understandable. The levels of agreement between household and medical provider data are being examined.

3.2.1 Probability of Agreement by Type of Conditions

There are twenty classes of conditions that are summarized in the table below. The probability of agreement is computed under three scenarios. The first focuses on the comparison of household and medical provider reports at the detailed three digits level; the second examine the probability of agreement after recoding the data into 131 major groups of diseases and the third deals with the recodes of the 3-digits into the 20 major groups. The first puzzling finding that this table exhibits is that there are groups of diseases where the levels of agreement of the three digit level is relatively very low, but when collapsing to the 20 major groups of conditions, the probability of agreement almost doubles. An example of such a pattern is the first listed in Table 5. For infectious and parasitic diseases the level of agreement at the 3-digit is 36 percent, compared to 71 percent at their appropriate recode into 20 major disease groupings.

When ordering the medical diagnosis using the probability of agreement for the 20-group recode, the probability of agreement is high for groups that encompass recurrent and recognizable conditions. The top five are: (1) complication of pregnancy, childbirth and puerperium (0.91); (2) injury and poisoning (0.86); (3) diseases of the circulatory system (0.82); (4) diseases of the nervous system and sensory organs (0.81); and (5) diseases of the respiratory system (0.79). The lowest probability of agreement, 0.58, is found for the least specific set of conditions: classification of factors influencing health status and contact with health services.

It is important to note that the rank order of these diseases by agreement rate varies with the level of recode of the medical condition. If the sort order is based on the level of

agreement at the 3-digit level, the top five groups of conditions are: (1) the circulatory system (0.43); (2) endocrine, nutritional and metabolic disorders, and immunity disorder (0.40); (3) disease of the respiratory system (0.36); (4) classification of factors influencing health status and contact with health services (0.36); (5) Infectious and parasitic diseases (0.36); (6) neoplasms (0.34). This ordering does not replicate for the other two recode levels.

Table 5 Probability of Agreement by Types of Conditions

Medical Diagnosis	Probability of Agreement		
	3-digits	131-groups	20-groups
1. Infectious and parasitic diseases	0.36	0.59	0.71
2. Neoplasms	0.34	0.55	0.79
3. Endocrine, nutritional and metabolic diseases, and immunity disorders	0.40	0.67	0.77
4. Diseases of the blood and the blood forming organs	0.26	0.61	0.72
5. Mental disorders	0.25	0.60	0.70
6. Diseases of the nervous system and sensory organs	0.25	0.75	0.81
7. Diseases of the circulatory system	0.43	0.63	0.82
8. Diseases of the respiratory system	0.36	0.56	0.79
9. Diseases of the digestive system	0.26	0.51	0.66
10. Diseases of the genitourinary system	0.20	0.59	0.71
11. Complications of pregnancy, childbirth, and puerperium	0.25	0.83	0.91
12. Disease of the skin & subcutaneous tissue	0.05	0.50	0.62
13. Diseases of the musculo-skeletal system and connective tissue	0.27	0.52	0.69
14. Congenital anomalies	0.22	0.48	0.62
15. Certain conditions originating in the perinatal period	0.26	0.35	0.68
16. Symptoms, signs, and ill defined condition	0.17	0.50	0.65
17. Injury and poisoning	0.24	0.61	0.86
18. Classification of factors influencing health status and contact with health services (V codes)	0.36	0.45	0.58
19. Classification of impairment	0.17	0.62	0.75
20. All others	0.15	0.28	0.38

The observed levels of agreement for all groupings would deter any derivation of medical conditions based only on household data when medical provider data are missing.

3.2.2 Probability of Agreement by Chronic and Acute Conditions

As part of the NMES study a list of most common chronic conditions, and a checklist of symptoms were identified by names. This list was patterned after similar lists in the Health Interview Survey. This list did not include ICD-9 codes, but rather names of diseases or descriptions of symptoms. In this analysis, the hypothesis is that at least for the chronic conditions the probability of agreement between the household and the medical provider should be high. Table 6 provides the probability of agreement at the three levels of recode for that list of diseases. The list is sorted in descending rank order based on the probability of agreement at the 3-digits condition code level.

The top five conditions with the highest level of agreement at the 3-digits codes are common, recognizable, and knowable: Diabetes, high blood pressure, hemorrhoids, more than two ear infections and hay fever. Also, for these conditions people generally take medication or visit the doctor on a frequent or regular basis. It is of interest to note that even with these types of conditions the levels of agreement are

below 80 percent whether at the 3-digits condition coding or after recoding into 131 major groups. One possibility which we already mentioned is the fact that in most instances the household respondent did not provide more than one medical condition. Thus the likelihood of agreement between the medical provider and the household respondent is decreased.

The conditions with the lowest level of agreement in this instance are: abdominal pains, stroke, enuresis, rheumatism, and arteriosclerosis. These conditions could be misreported easily or could be conditions associated with social stigma. Stroke could be reported as heart problems, rheumatism and arthritis are frequently confused, and enuresis is associated with social stigma: people may not want to report bed wetting to an interviewer. If the same probabilities (Table 6) are sorted in rank order of the probability of agreement based on the recoding of the 3-digit ICD-9 codes into the 20- major medical groupings, the top five conditions change to include more of the chronic conditions: myocardial infarction, gallbladder disease, hay fever, cancer of any kind, and high blood pressure. With the lowest agreement level at the same 20-groups collapse we find abdominal pains, acne, enuresis, rheumatism, and indigestion/runny nose/sore throat. Two of these conditions, abdominal pains and enuresis appear in the lowest agreement grouping regardless of the recoding pattern used.

Table 6 Probability of Agreement for Chronic Conditions, and Known Symptoms

Condition	Probability of Agreement		
	3-digit Codes	131-groups	20-groups
Diabetes	0.66	0.73	0.82
High blood pressure	0.65	0.72	0.83
Hemorrhoids (piles)	0.61	0.64	0.79
More than two ear Infections	0.56	0.77	0.83
Hay Fever	0.51	0.65	0.83
Cancer of any kind	0.50	0.59	0.83
Gallbladder disease	0.48	0.72	0.85
Skin Rash	0.47	0.59	0.69
Frequent Headaches	0.44	0.54	0.69
Emphysema	0.42	0.58	0.75
Varicose veins	0.41	0.52	0.70
Weakness or fatigue	0.41	0.54	0.68
Frequent sinus problem	0.41	0.57	0.77
Weight loss or upset stomach	0.40	0.56	0.69
Myocardial infarction	0.39	0.54	0.87
Anemia	0.39	0.61	0.72
Indigestion/SoreThroat/Runny nose	0.38	0.52	0.68
Heart Disease	0.38	0.57	0.83
Arthritis	0.36	0.60	0.73
Acne	0.35	0.47	0.56
Shortness of breath	0.35	0.50	0.72
Pain or swelling in joint	0.34	0.56	0.71
Repeated Backache	0.34	0.52	0.69
Arteriosclerosis	0.31	0.51	0.76
Rheumatism	0.31	0.48	0.65
Enuresis	0.29	0.55	0.65
Stroke	0.28	0.46	0.72
Abdominal pains	0.26	0.42	0.52

3.3 Agreement in Reporting Medical Conditions as a Function of Demographic and Socioeconomic Factors

There are several demographic and socio-economic characteristics that are potentially related to health status and to the ability to report medical condition. The characteristics that we have identified are:

o Demographic: age, race, gender, place of residence (region), and poverty status

411

o Health Insurance status: covered by either private insurance, public insurance, or covered by either medicare or medicaid

o Billing for health care provided: did the household receive a bill with the recorded set of medical conditions?

The results of the various examinations showed that no socio-demographic characteristics explain the observed rates of agreement on the medical condition reports associated with the event. The rates of agreement on condition as reported by hispanics or by white people is 57 percent with the most extensive collapse (20-groups). The rate of agreement on medical reporting for blacks is not much different, at 55 percent. The uniformity in the reporting of medical condition was the same for the other socio-demographic characteristics mentioned above.

The only characteristic that needs to be investigated is the relationship between the level of education and the rate of agreement on condition reported.

3.4 Effects of Household Proxy Reports on the Probability of Agreement between Household and Medical Provider Condition Data

NMES is a household survey which collects information about the family unit and its individual family members. In multi-person households, a designated family respondent will answer questions about him/herself, and also act as a proxy respondent for questions about other family members. We want to explore whether the probability of agreement between the medical provider data and the household-supplied information is influenced by the degree of first-hand knowledge that family respondents may have about the medical encounters of other family members for whom they are reporting.

All things being equal, self reports are likely to be more accurate than reports provided by a proxy; however, there are qualifications to this assumption. Anyone who has direct responsibility for the care and welfare of another person is likely to have intimate knowledge about the affairs of that other person. Thus, the mother of a young child should have more direct knowledge about the child's medical conditions than about the husband's.

Who reports for whom in the family, and how much detail the reporter is likely to know about the person who is reported about, may have an effect on the agreement levels between data supplied by family respondents and data obtained from medical providers about medical conditions for the same visits. To test for this potential effect, each event was classified as to the person who provided the information (self/proxy) and the age of the sample person if someone other than self gave the event information in NMES.

Table 7. Probability of Agreement by Type of Household Reporter (Self/Proxy) and Age of Sample Person (SP) whose Information was Supplied by a Proxy

Reporter	Agreement Probability (Any-to-Any Variable)	#Events
Self	0.402	18,797
Proxy; SP<17	0.406	6,029
Proxy; SP 17+	0.373	7,380
Unknown	0.394	1,308
TOTAL	0.396	33,514

The resulting variable and distribution of events is as follows: person reported about self (56.1%); proxy reported about sample person 16 years old or younger (18.0%); proxy

reported about sample person older than 16 (22.0%); type of reporter is unknown (3.9%).

Agreement levels on medical condition reporting between household and medical provider (any-to-any version) were computed for each reporting class, and a significant difference detected, as predicted, for the class of events involving proxy reports for family members older than 17 compared to self (Z=4.35) and compared to proxy reporting for persons under the age of 17 (Z=3.90).

4.0 Summary

The National Medical Expenditure Survey provides a unique opportunity to assess the reliability of medical conditions reported by household respondents. In this study, data were collected about the same medical visits from two sources-- the household respondent and the medical provider who cared for the person in the sampled household.

The results of this study reinforce the need to analyze the concordance in reporting medical condition between household respondents and medical providers in terms of: (1) recall bias; (2) questionnaire design; (3) proxy reporting; and, (4) medical coding of conditions.

The analysis shows that demographic characteristics, and the socio-economic status of the household respondent do not affect the level of congruence between respondents and providers in reporting medical conditions. Harlow and Linet (1987) reached the same conclusion when summarizing the limited number of studies addressing this issue.

Coditz et al. (1986) suggested that accurate reporting, and consequently agreement between medical provider reports and household reports, is more likely for diseases that have clear and unambiguous diagnostic profiles. The results of our study indicate that the levels of agreement range from 38 to 82 percent for various diagnoses using the most collapsed recode of conditions and the least restrictive definition of agreement; the possible exceptions are the agreement levels of pregnancy related visits. Once the levels of agreement are sorted by magnitude, chronic diseases such as diabetes and hypertension seem to have the highest level of agreement.

The rates and the patterns of missing data for household and for medical provider reports in 1987 NMES data are reassuring. With proper training and instrumentation, interviewers were able to obtain codeable information. However, the limited number of conditions elicited from household respondents may reflect the need to revise the training and the probing process of interviews. There does not seem to be reluctance to report medical conditions by either the household or the medical reporters.

Although agreement levels reach the 90 percent mark infrequently in 1987 NMES, it is nonetheless the case that this study achieved significantly higher agreement rates than those in 1977 NMCES, the precursor of NMES. Using major groupings of condition codes, NMCES reports an average agreement value of 40 percent, whereas in NMES the comparable agreement is 68.2. This improvement is partly the result of the decision to code data from the two 1987 NMES sources (household and provider) according to the same modified ICD-9 coding scheme.

Future research should focus on analysis of the levels of agreement in the context of health status measures, medication use, and length of recall period associated with the health events as control measures that may help identify subgroups with significantly higher agreement rates.

References available from authors upon request.