Estimating the Cost of Illness: The Case of Diabetes

Gary Olin and Steven R. Machlin and Jeffrey Rhoades

Agency for Healthcare Research and Quality Working Paper No. 08001

January 2008

Suggested citation: Olin G, Machlin SR, Rhoades J. Estimating the Cost of Illness: The Case of Diabetes. Agency for Healthcare Research and Quality Working Paper No. 08001, January 2008, http://gold.ahrq.gov.

Estimating the cost of illness: the case of diabetes
Gary Olin, Steven Machlin, and Jeffrey Rhoades
January 2008

ABSTRACT

We use the Medical Expenditure Panel Survey to estimate the direct medical costs of diabetes. Separate estimates of the cost of diabetes are developed from an attributable risk procedure and a two-part regression model. Costs include expenditures to treat diabetes, chronic complications of the disease, and co-morbidities requiring additional health care resources because the person is diabetic. The cost estimates range from $79.1 billion based on attributable risk to $88.8 billion based on the two-part regression model, but the difference is not statistically significant. These estimates are derived from the actual health care use and expenditure patterns of the disease population. They indicate that 59 to 66 percent of health care spending by diabetics for physician and hospital care and prescription medicines should be attributed to the disease itself.

Gary L. Olin, Ph.D.
Agency for Healthcare Research and Quality
Division of Survey Operations
Center for Financing, Access, and Cost Trends
540 Gaither Road
Rockville, MD 20850
Gary.Olin@ahrq.hhs.gov

Steven R. Machlin, MS
Agency for Healthcare Research and Quality
Division of Research and Survey Methods
Center for Financing, Access, and Cost Trends
540 Gaither Road
Rockville, MD 20850
Steven.Machlin@ahrq.hhs.gov

Jeffrey Rhoades, Ph.D.
Agency for Healthcare Research and Quality
Division of Survey Operations
Center for Financing, Access, and Cost Trends
540 Gaither Road
Rockville, MD 20850
Jeffrey.Rhoades@ahrq.hhs.gov

Estimating the cost of illness: the case of diabetes

Introduction

Health researchers and policymakers use cost of illness (COI) estimates as a measure of the economic burden of a disease on society. However, the estimates are not comparable in many cases, even for the same disease, because of differences in estimation techniques, data sources, and definitions of cost and disease (Kirschstein 2000). The primary data sources for most COI studies have been surveys and reports of the federal government (Songer and Ettaro 1998). They include health expenditure data from the Centers for Medicare and Medicaid Services (CMS), cause of death data from the National Center for Health Statistics (NCHS), and health services use and cost data from CMS, NCHS, and the Agency for Healthcare Research and Quality (AHRQ). The Medical Expenditure Panel Survey (MEPS) sponsored by AHRQ and NCHS is the only one containing all of the information necessary to estimate the direct costs of a disease for a nationally representative sample of the U.S. civilian non-institutionalized population.

In this paper we use the 2002 MEPS to illustrate two estimation methods commonly employed in COI studies. We have chosen diabetes as the example condition to illustrate the methods because it is a relatively prevalent and costly condition, and complications of the disease are well established (American Diabetes Association 1993). The methods consist of an attributable risk procedure and a two-part regression model. We compare cost estimates from the two methods and assess the strengths and weaknesses of using MEPS as the sole data source for the estimates. We also compare

the MEPS estimates to those from a report by the American Diabetic Association (ADA) for the same year (ADA 2003). The ADA report also uses a variant of the attributable risk procedure, but the cost estimates are based on data from several sources, including MEPS and other federal surveys.

Data and Methods

The MEPS is an annual survey of a nationally representative sample of the U.S. civilian non-institutionalized population conducted by the AHRQ and cosponsored by the NCHS. Descriptions of the survey and its methodology can be found in J. Cohen 1997, S. Cohen 2000, and S. Cohen 2003. We use the household component (HC) of the 2002 MEPS in our study. The MEPS HC collects detail on demographic characteristics, health conditions, health status, use of medical care services, charges and payments, access to care, satisfaction with care, health insurance coverage, income, and employment.

We use the expenditure and medical conditions data to estimate the direct cost of diabetes. The expenses represent amounts actually paid for health care services, and they are collected for each hospital stay, outpatient department visit, emergency room visit, office-based doctor visit, month of home health care, dentist visit, and prescription medicine purchase reported in the HC of the survey. Total expenses for an "event" are defined as the sum of direct payments by households, private insurance, Medicare, Medicaid, and other sources to providers of the care. Expenses for all of the event types except dental visits are used to estimate diabetes-attributable expenses. Dental care is excluded from the analysis because the survey respondents are not asked about the reason(s) for their visits to dentists. We also exclude "other medical expenses," a catchall category containing annual expenses for vision products and durable medical equipment.

Persons with diabetes are identified through the responses of survey respondents to questions in the household questionnaire about medical conditions and problems. In the priority conditions section, survey respondents are specifically asked whether they have ever been diagnosed with diabetes. In the conditions enumeration section, survey respondents are asked to list their injuries, physical conditions, and mental or emotional health conditions. In most of the medical events sections, survey respondents are asked about the reasons for their medical provider events and prescription medicine purchases. All of the health problems and medical conditions reported in the conditions enumeration and medical events sections of the questionnaire are recorded verbatim in the interviews and subsequently assigned a code by professional medical coders using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). Each hospital stay, ambulatory care visit, and prescription medicine purchase can have one or more ICD-9-CM codes associated with it. These self-reported conditions may not always conform to diagnoses by physicians, but an AHRQ staff study indicated agreement between household and provider reported conditions in the overwhelming majority of cases at the 3-digit ICD-9-CM level (Krauss and Kass 2000).

Our sample of persons with diabetes (2,164 respondents) contains the maximum number of people in the survey that could be identified as having diabetes. It includes the self-reported diabetics plus all other survey respondents with an event for the treatment of diabetes in 2002. This group represents approximately 15.5 million persons with diabetes, or 5.4 percent of the U.S. civilian non-institutionalized population in 2002 (Table 1). One of the distinguishing characteristics of this group is it has a relatively high proportion of people with renal disease, ophthalmic disease, cardiovascular disease,

endocrine/metabolic disease, peripheral vascular disease, neurological disease, and other chronic complications associated with diabetes. In 2002, for example, about three-quarters of the diabetic population had at least of one these medical conditions compared with slightly more than one-fifth of the non-diabetic population.

Not surprisingly, persons with diabetes have relatively high health care expenses. In 2002, their expenses for hospital stays, ambulatory care, home health care, and prescription medicines were $133.6 billion, or 18.3 percent of all health care spending for these types of services even though persons with diabetes comprised only 5.4 percent of the total population (Table 2). Nearly 90 percent of these expenses were incurred by the three-quarters of the diabetic population with at least one complication of diabetes.

However, not all of the $133.6 billion in health care spending for persons with diabetes should be attributed to diabetes. In 2002, persons with diabetes had expenses of $15.9 billion for medical events where diabetes was the only reason given for the event (Table 3). All of these expenses are counted as part of the cost of diabetes. Another $11.1 billion was for medical events where diabetes was one of the reasons for the event. A large proportion of these expenses should be counted as part of the cost of diabetes. The remaining expenses ($106.5 billion) were for medical events where the reason(s) for the event was something other than the treatment of diabetes. Some proportion of these expenses should be counted as part of the cost of diabetes because they include care for conditions that would not have occurred or would be more easily treated in the absence of diabetes (ADA 1993).

We use two methods—an attributable risk procedure and a two-part regression model—to estimate health care expenses attributable to diabetes. The expenses include

the cost of treating chronic complications of the disease and co-morbidities requiring additional health care resources because the person is diabetic. Diabetics' expenses for dental care, vision products, and durable medical equipment are excluded from the analysis because MEPS does not ask about medical conditions associated with these types of expenses. Further, we do not attempt to estimate indirect costs of diabetes such as lost workdays, permanent disability, and premature mortality.

The Attributable Risk Procedure

An attributable risk procedure can be used to estimate the costs for care associated with chronic complications of diabetes and for additional resources needed to treat a diabetic for the same co-morbidities as a non-diabetic. The procedure uses an etiologic fraction to estimate the relative contribution of a factor such as diabetes to another condition (Rothman 1986). For a sample of known persons with diabetes, the proportion of health care expenses for a condition that is attributable to diabetes is calculated based on the following equation:

(1) $AP_i = (RR_i - 1)) / RR_{i,}$

where AP_i is the proportion of health care for condition "i" that is attributable to diabetes and RR_i is the relative risk of condition "i" for people with diabetes compared with people without diabetes.

The attributable risk procedure assumes that other causes of conditions attributable to diabetes have equal impact on people with and without diabetes (Last 1995). We estimate attributable fractions for eight categories of conditions including renal, ophthalmic, cardiovascular, cerebrovascular, peripheral vascular, neurological, other chronic complications, and all other care. The first seven categories are included in

the analysis because they are chronic complications of diabetes—conditions that diabetics are at greater risk of having compared with non-diabetics. The last category—all other care—is included in the analysis because more resources are often needed to treat a diabetic for the same condition as a non-diabetic. In 1992, for example, the average hospital stay for a condition unrelated to be diabetes was estimated to be 2.8 days longer for diabetics than it was for non-diabetics (ADA 1993).

The fraction of diabetes-attributable expenditures for each of the eight categories is derived from the odds of being in a specific category for people with diabetes compared with people without diabetes. To identify people in a specific category, we used the MEPS condition codes to create eight person-level flags for the 37,418 survey respondents in the 2002 MEPS HC. Each flag corresponds to one of the eight categories of conditions. A flag has a value of one if the condition in question was reported by the survey respondent in any section of the questionnaire collecting information on medical conditions and a value of zero otherwise.

Attributable fractions are estimated by age and race/ethnicity group for the eight categories of conditions. We stratified the sample by age and race/ethnicity because the prevalence of diabetes (and therefore a complication) increases with age and varies by race and ethnicity (CDC 2005). The age groups are 0-44, 45-65, and 65+, and the race/ethnicity categories are Hispanic of any race, white non-Hispanic, black non-Hispanic black, and other non-Hispanic. The odds ratios used to determine the attributable fraction for each complication of diabetes and demographic group have been estimated using logistic regression equations specified as follows:

Y_i = 1 if the person was in category "i" and 0 otherwise.

X_i = 1 if the person was diabetic and 0 otherwise.

To determine which fraction to apply to expenditures for medical events by persons with diabetes, nine flags were created for each of their events based on the medical conditions associated with the events. The flags show whether the reason for the medical event was diabetes, one of the other eight categories of conditions, or a combination of conditions. The process of determining the share of expenses that should be attributed to diabetes is complicated, however, because the primary condition associated with an event is not identified in MEPS. To resolve this problem, we use the following rule to determine the share of expenses attributable to diabetes. If the sole condition code for an event was for diabetes, 100 percent of the expenses were attributed to diabetes. For other events with one condition code, expenses attributable to diabetes were determined by the attributable fraction for that condition and demographic group. For events with two or more condition codes, the share of expenses attributable to diabetes was calculated as a simple average of the attributable fractions for the conditions and demographic group associated with that event.

The Two-Part Regression Model

Our second method of estimating diabetes-attributable expenses is an application of the Rand Health Insurance Experiment two-part regression model. This model was proposed as a method of estimating health care expenditures characterized by a large number of zero values and a highly skewed distribution of expenditures among people with expenditures (Duan et al. 1983). Part one of the model is a logistic regression in which the dependent variable is set to 1 if the individual had expenses and to 0 otherwise. The independent variables include a dummy variable for diabetics (equal to 1 if the

individual had diabetes and 0 otherwise) and indicator variables capturing the effect of the other control variables (age, race/ethnicity, and insurance coverage). Part two is an ordinary least-squares regression on the sub-sample of persons with expenses. The dependent variable is the natural logarithm of expenses, and the independent variables are the same ones used in part one.

In contrast to the attributable risk procedure, the two-part regression model uses expenses for all sample persons rather than just those for persons with diabetes to estimate the cost of the disease (37,418 respondents in the 2002 MEPS). The model provides an estimate of the difference in expenses for diabetics relative to non-diabetics while controlling for other factors that may affect health care spending. This difference in expenses by the two groups represents the total cost of diabetes including expenses for complications of the disease and additional resources needed to treat other co-morbidities. We control for age (<45, 45-64, 65 and over), race/ethnicity (Hispanic, white non-Hispanic, black non-Hispanic, all other), and insurance coverage for persons under age 65 (private coverage, only public coverage, uninsured) because they represent a basic set of factors associated with both the likelihood of a person having diabetes and the level of health care expenses.

The estimate of diabetes-attributable expenses is developed from the two-part model as follows. First, we use the coefficients from the model to predict two expenses for each individual in the sample. The first predicted expense assumes the individual has diabetes. It is obtained by setting the diabetes indicator to one and multiplying the predicted probability of having any expenses from the first equation by the predicted average expenses from the second equation in the model. The second predicted expense

assumes the individual does not have diabetes. It is obtained by setting the diabetes indicator to zero and multiplying the predicted probability of having any expenses from the first equation by the predicted average expenses from the second equation in the model. For both sets of predicted expenses, the log dollars are transformed to actual dollars using a smearing factor to correct for transformation bias (Duan 1983). Second, we estimate the predicted average per-person increase in medical expenses attributable to diabetes by taking the difference in predicted expenses for each person and computing the weighted average of the difference across the entire sample. Third, we estimate total expenses attributable to diabetes by summing the weights for all individuals with diabetes in the sample and then multiplying the total by the estimated per-person increase in expenses attributable to diabetes.

Results

Persons with diabetes incurred health care expenses of $133.6 billion in 2002 (Table 4). The expenses include $15.9 billion for the treatment of diabetes, $11.1 billion for diabetes and other conditions, and $106.5 billion for conditions other than diabetes. Using the attributable risk procedure, we estimate that persons with diabetes had expenses of $79.1 billion for diabetes and associated problems in 2002. Approximately thirty percent of the diabetes-attributable expenses ($24.1 billion) were for medical events where diabetes alone ($15.9 billion) or diabetes and other conditions ($8.2 billion) were given as the reason for the event. The other diabetes-attributable expenses ($55.0 billion) were for medical events where diabetes was not identified as a reason for the event (i.e., expenses for chronic complications of diabetes and other co-morbidities in general). These expenses represent about 59 percent of all health care expenses incurred

by persons with diabetes. The other 41 percent of their expenses would have been incurred even if they did not have diabetes.

The two-part model produces an estimate of $88.8 billion in diabetes-attributable expenses, or 66 percent of the total expenses of persons with diabetes. It has been derived by multiplying the weighted average difference in per person expenses of people with and without diabetes ($5,724) by the weighted estimate of persons with diabetes in the 2002 civilian noninstitutionalized population (approximately 15.5 million persons). The $88.8 billion represents the incremental cost of diabetes but, unlike estimates based on attributable risk, it cannot be disaggregated by type of condition being treated. However, the difference between the two estimates of total expenses attributable to diabetes is not statistically significant, and both estimates have small relative standard errors (5.4% for the attributable risk estimate and 7.9% for the two-part model estimate).

We compare our estimates to those in the most recent report by the American Diabetic Association on the cost of diabetes (ADA 2003). The estimates in the ADA report are based on a prevalence-based attributable risk procedure designed for a mixed population of people with and without a disease (Rothman 1986, pp.38-39), and the cost and use data are from a variety of sources. Table 8 of that report (not shown) contains estimates of diabetes-attributable expenses for the entire U.S. population by medical condition and type of service in 2002. If nursing home and hospice expenses are excluded from the $91.9 billion in health care expenses attributable to diabetes, the remaining expenses of $77.4 billion are mostly those of diabetics in the civilian non-institutionalized population. This estimate falls within the same range as ours, but the

precision of the ADA estimate is unknown because standard errors are difficult or impossible to calculate for estimates derived from multiple sources of data.

Discussion

MEPS Strengths and Limitations

MEPS is the only survey containing detailed information on medical conditions, health care use and expenses, and demographic and economic characteristics of a nationally representative sample of the U.S. civilian non-institutionalized population. The person-level information allows researchers to use the survey as a stand alone source to identify a disease population and estimate the proportion of the population's expenses that are attributable to the disease. Assuming the survey respondents do not systematically under-report their health care, these estimates of disease-attributable expenses should be more accurate than those from administrative data and surveys of medical providers because they are based on the actual health care use and expenses of persons with the disease (Szava-Kovats and Johnson 1997). Moreover, the impact of sampling variability on the results can be assessed by calculating standard errors for the estimates. Because the MEPS has a complex survey design, we use SUDAAN (Profession software for SUrvey DAta ANalysis for Multi-stage Sample Designs) to compute the variance of weighted estimates of statistics derived directly from the survey data. For predictions of diabetes-attributable expenses, we use a balanced half-sample replication method to compute the variance of the predicted expenses (Korn and Graubard 1999).

Another strength of MEPS is that estimates of disease-attributable expenses can be compared to the total expenses of the disease population to paint a more complete

picture of the health care burden of people with the disease. In addition, although our illustrations focus on estimates of total expenses attributable to a disease, detailed estimates of the expenses can be made in many instances. Depending on the disease, it may be possible to estimate expenses by subpopulation, type of care or complication, or source of payment. All of this information can be used by health services researchers to highlight areas with the greatest potential for savings or by policymakers weighing decisions to fund research and other programs related to a disease.

MEPS does have limitations. It excludes nursing home residents, a group with significant expenses for some diseases including diabetes. Furthermore, sample size may be an issue in estimating expenses of the civilian non-institutionalized population for less common diseases. Cost estimates will have relatively large standard errors if the sample contains few people with a disease or the disease population has few medical events. However, the annual files can be combined to get more precise estimates of the cost of a disease if sample size is a problem.

Another consideration in using MEPS for COI work is that medical conditions are reported by the survey respondents. Household-reported conditions should produce better estimates of the prevalence and cost of a disease compared with diagnostic information from medical providers because identification of the disease population is not dependent on its seeking health care. Nonetheless, the information is household-reported and must be converted from verbatim text into ICD-9-CM codes before it can be used to estimate the prevalence and cost of a disease. This process of collecting and coding household information on medical conditions is subject to a variety of errors including lack of knowledge or recall problems regarding sample persons' conditions,

deliberate failure to report some conditions due to social stigma, inadequate information to code a condition, incomplete information on the number of conditions, and error due to proxy reporting (Johnson and Sanchez 1993).

Comparison of Cost Methods

Either method—the attributable risk procedure or the two-part model—can be used to estimate the cost of a disease such as diabetes. Both methods use a control group, although in very different ways, to estimate the cost of a disease. The attributable risk procedure uses odds ratios or estimates of relative risk to determine the fraction of expenses for a particular condition that should be attributed to a disease. Then the fraction is used to estimate expenses for the excess prevalence of the condition in the disease population. In contrast, the regression model is used to predict average expenses for people with and without a disease while controlling for other factors which can influence health care expenses. To estimate disease-attributable expenses, the difference in predicted average spending by the two groups is applied to the disease population.

A general caveat in using attributable risk is that the fraction of expenses for a condition caused by a disease may be overstated if the influence of confounding factors is not accounted for in the procedure (Last 1995). For example, obesity may have a confounding effect on the proportion of services for cardiovascular disease that should be attributed to diabetes because it is associated with both diabetes and cardiovascular disease. In addition, estimates of the relative risk of a complication for people with and without a disease may be biased when based on household survey data. If diabetics, for instance, are monitored more closely by their physicians for complications of the disease,

they are more likely than non-diabetics to be aware of the complications and report them to interviewers (Huse 1989).

Caveats notwithstanding, one strength of the attributable risk procedure is that it yields information on the extent to which each complication of a disease contributes to medical expenses of people with the disease (not shown in this paper). This information is valuable because complications can add substantially to the total cost of a disease. However, the complications must fall within international classification of disease categories, and the sample must be large enough to provide reliable estimates of the relative risk of each complication for people with and without the disease.

Another consideration in choosing to use the attributable risk procedure is that disease-specific attributable risks cannot always be estimated for every type of care. In MEPS, for example, medical conditions are not linked to dental care or purchases of vision products and durable medical equipment. This problem does not affect the two-part model because it relies on total expenses for diabetics and non-diabetics rather than on condition-specific expenses. However, we deliberately excluded dental care and other medical expenses from the two-part model estimates of the cost of diabetes in order to compare estimates from the two methods.

Errors in the reporting of medical conditions can affect either cost method, but they may have more impact on the attributable risk procedure. It relies on household-reported conditions to identify people with and without specific conditions as well as all of the medical events with disease-attributable expenses. The two-part model, on the other hand, only uses the household-reported conditions to differentiate between people with and without a disease. Relatively complex statistical and modeling issues must be

addressed, however, in a using a regression model to estimate the cost of disease. The two-part model used in this paper, for example, is a popular least squares method of modeling health care utilization and expenditures data characterized by a large number of zero values and a highly skewed distribution for the nonzero observations. Predictions from this model must be adjusted to remove the bias that results from retransforming log dollars to actual dollars—a straightforward process when the variance of the error term is constant across all observations (homoskedastic). The adjustment for retransformation bias is more complicated, though, when the error term is related to one or more of the explanatory variables in the model (heteroscedastic), and the two-part model does not necessarily produce estimates as precise as those from generalized linear models even when corrected for bias. As a consequence, the best choice of models can vary depending on the problem and data set (Manning and Mullahy 2001; Buntin and Zaslavsky 2005).

Conclusion

We have used the 2002 MEPS to illustrate two popular methods of estimating the cost of disease. Our estimates of the cost of diabetes range from $79.1 billion based on attributable risk to $88.8 billion based on the two-part model. These estimates, which are not statistically different, indicate that a large proportion of the $133.6 billion in total health care spending by persons with diabetes should be attributed to the disease. Moreover, the two estimates are comparable to those published by the ADA ($77.4 billion) , which used a prevalence-based attributable risk procedure and a variety of data sources in its report.

MEPS sample size limitations and concerns about the accuracy of household-reported medical events and conditions may limit its use in estimating the cost of some diseases. Nonetheless, the survey seems well suited for analyses of highly prevalent diseases such as diabetes because it contains a wealth of detail on individuals and their health care use and expenses. The demographic and economic information can be used to create appropriate categories of people for cost estimation, and the detail on their use and expense patterns should produce cost estimates superior to those from administrative data sets and surveys of medical providers. In addition, disease-attributable expenses can be compared to total expenses of the disease population. Our point estimates of the cost of diabetes, for example, range from 59 to 66 percent of all expenses incurred by persons with diabetes. This information helps to paint a more complete picture of the impact of a disease on total expenses of the disease population, but it usually not presented in studies using other sources to estimate disease-attributable costs.

The choice of estimation methods depends on the disease and the requirements of the analyst. Attributable risk provides more detail on expenses for disease-related conditions, but it also requires more information on the conditions of people with and without the disease. The two-part model requires less information on conditions affecting the disease population, but its estimates of the incremental cost of a disease can hide other information of interest to researchers and policymakers. Nonetheless, regardless of choice, we would argue that MEPS-based estimates of the cost of a disease can offer more insight about the economic burden of a disease than those from many other sources, including information on the impact of sampling variability on the cost

estimates. The variance estimates from MEPS may be particularly important because they provide critical information on the precision and reliability of the cost estimates.

References

American Diabetes Association. 1993. *Direct and Indirect Costs of Diabetes in the United States in 1992*. Alexandria, VA.

American Diabetes Association. 2003. "Economic Costs of Diabetes in the U.S. in 2002." *Diabetes Care*, 26 (3): 917-932.

Buntin, M. B. and A. M.. Zaslavsky. 2005. "Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures." *Journal of Health Economics*, 23(3): 525-542.

Centers for Disease Control and Prevention. 2005. "National Diabetes Fact Sheet." National Center for Chronic Disease Prevention and Health Promotion. [accessed on November 1, 2005]. Available at: http://www.cdc.gov/diabetes.

Cohen, J. W. 1997. "Design and Methods of the Medical Expenditure Panel Survey Household Component." Rockville (MD): Agency for Health Care Policy and Research. MEPS Methodology Report No. 1. AHCPR Pub. No. 97-0026.

Cohen, S. B. 2000. "Sample Design of the 1997 Medical Expenditure Panel Survey Household Component." Rockville (MD): Agency for Healthcare Research and Quality. MEPS Methodology Report No. 11. AHRQ Pub No. 01-0001.

Cohen, S. B. 2003. "Design Strategies and Innovations in the Medical Expenditure Panel Survey." *Medical Care* 41(7): 5-12.

Duan, N. 1983. "Smearing Estimate: A Nonparametric Retransformation Method." *Journal of the American Statistical Association,* 78(383):605-610.

Duan, N., W. G. Manning, C. N. Morris, and J. P. Newhouse. 1983. "A Comparison of Alternative Models for the Demand for Medical Care." *Journal of Business & Economic Statistics*, 1(2):115-126.

Huse, D., G. Oster, G., A. Killen, M. Lacey, and G. Colditz. 1989. "The Economic Costs of Non-Insulin Dependent Diabetes Mellitus." *Journal of the American Medical Association* 262(19): 2708-2713.

Johnson, A. and M. Sanchez. 1993. "Household and Medical Reports on Medical Conditions: National Medical Expenditure Survey." *Journal of Economic and Social Measurement*, 19(2): 199-223.

Kirschstein, R. February 2000. "Disease-Specific Estimates of Direct and Indirect costs of Illness." Department of Health and Human Services, National Institutes of Health.

Korn, E.and B. Graubard. 1999. *Analysis of Health Surveys*. New York: John Wiley & Sons, Inc.

Krauss, N. and B. Kass. August 2000. "Comparison of Household and Medical Provider Reports of Medical Conditions." Paper presented at the Joint Statistical Meetings, Indianapolis, Indiana.

Last, J. (ed). 1995. *A Dictionary of Epidemiology*. Oxford: Oxford University Press.

Manning, W. G. and J. Mullahy. 2001. "Estimating log models: to transform or not to transform?" *Journal of Health Economics* 20(4): 461-494.

Rothman, K. 1986. *Modern Epidemiology*. Boston/Toronto: Little, Brown and Company.

Songer, T., L.Ettaro, and the Economics of Diabetes Project Panel. June 1998. "Studies on the Cost of Diabetes." Prepared for the Division of Diabetes Translation, Centers for Disease Control and Prevention, Atlanta GA.

Szava-Kovata, G. and J. Johnson. 1997. "A Review of Diabetes Cost of Illness Studies." Institute of Pharmaco-Economics Working Paper 97-3, Edmonton, Canada.

Table 1. U.S. civilian non-institutionalized population, by diabetic status, 2002

Diabetic Status	No. of persons (in millions)	Percent of total population	Percent by diabetic and complication status
All diabetics	15.5	5.4	100.0
No complications*	3.7	1.3	(24.1)
At least 1 complication*	11.8	4.1	(75.9)
All non-diabetics	272.7	94.7	100.0
No complications*	213.2	74.0	(78.2)
At least 1 complication*	59.5	20.7	(21.8)
All persons	288.1	100.0	

*Complications include renal disease, ophthalmic disease, cardiovascular disease, endocrine/metabolic disease, peripheral vascular disease, neurological disease, and other chronic complications.

Table 2. Distribution of health care expenses* of the U.S. civilian non-institutionalized population, by diabetic status, 2002

Diabetic Status	Median**	Mean**	Total (in millions)	Percent of total
All diabetics	3,937	8,612	133,561	18.3
No complications	1,528	4,303	(15,870)	(2.2)
At least 1 complication	4,814	9,957	(117,691)	(16.1)
All non-diabetics	363	2,181	594,729	81.7
No complications	197	1,260	(268,574)	(36.9)
At least 1 complication	2,099	5,483	(326,155)	(44.8)
All persons	421	2,527	728,290	100.0

*Does not include expenses of diabetics for dental care, vision products, or durable medical equipment.
** Includes a small number of persons with no expenses.

Table 3. Total health care expenses* of diabetics for diabetes and other conditions, 2002

Condition	Diabetes only	Diabetes and other conditions	Other conditions Only	All conditions
Expense (in millions)	$15,927	$11,126	$106,507	$133,561
(Standard error)	(1,202)	(1,143)	(6,315)	(7,127)

*Does not include expenses of diabetics for dental care, vision products, or durable medical equipment.

Table 4. Total and diabetes-attributable expenses* of diabetics by type of condition, 2002

Type of Condition	Diabetes only	Diabetes and other conditions	Other conditions only	Total expenses for diabetics	Total expenses attributable to diabetes
Total expenses of diabetics (in millions)					
Total	$15,927	$11,126	$106,507	$133,561	--
(Standard error)	(1,202)	(1,143)	(6,315)	(7,127)	
Diabetes-attributable expenses (in millions)					
AR estimates	$15,927	$8,211	$55,006	--	$79,144
(Standard error)	(1,089)	(1,165)	(5,208)		(4,266)
Two-part estimate	--	--	--	--	$88,800
(Standard error)					(7,020)

*Does not include expenses of diabetics for dental care, vision products, or durable medical equipment.