Prediction of Prospective Medical Claims and Absenteeism Costs for 1284 Hourly Workers from a Manufacturing Company

Louis Tze-ching Yen, PhD

Dee W. Edington, PhD

Pamela Witting, MS

Multivariate regression models were developed to predict the employer's economic costs (medical claims payments and loss from employees' absenteeism) from selected health-related measures among 1284 hourly employees. Twelve health-related measures, mainly from a Health Risk Appraisal (HRA) completed in 1985, were selected as predictors. Regression models were developed according to the economic measures for the three subgroups (men under 35, men 35 or older, and women). All the regression models significantly predicted economic costs from selected health-related measures with a multiple R^2 range of .23 to .13 ($P < .001$).

The interrelationships between employees' health, absenteeism loss, and medical claims costs have become a major research component in the field of work-site disease prevention and health promotion.[1-4] A number of work-site studies have reported that employees' health is significantly related to their absenteeism[5-10] and medical claims costs.[8-14] The reports suggest that healthier employees (those who have lower health risks, those who practice healthy lifestyles, or those who participate in work-site health promotion programs) have lower absenteeism and less medical claims costs. However, most of those studies are limited to a bivariate relationship that explores a single health-risk effect on medical claims costs and absenteeism.

Several studies have used multivariate analysis techniques to study the relationships of health risks to mortality, morbidity, and medical cost.[15-18] However, there are few studies using the multivariate approach to study the combined economic impact of employees' health on their absenteeism and medical claims costs. In a previous article[19] and with other data (L. Yen and D. W. Edington, unpublished data, May, 1990) we confirmed that there were strong associations between personal health-related meas-

From the Fitness Research Center, University of Michigan, Ann Arbor (Dr Yen, senior research analyst, and Dr Edington, Director) and Steelcase Incorporated, Grand Rapids, Michigan (Ms Witting, Manager, Wellness Program).

Address correspondence to: Dee W. Edington, Director of Fitness Research Center, The University of Michigan, 401 Washtenaw Avenue/3040 CCRB, Ann Arbor, MI 48109-2214.

0096-1736/92/3404-0428$03.00/0

Copyright © by American College of Occupational and Environmental Medicine

ures and prospective medical care costs. We also indicated that a combination of personal health-related measures could significantly predict prospective medical claims costs.

The purpose of this study was to determine the economic impact of employee's health on the employer's prospective cost (both medical claims cost and absenteeism loss).

Methods

Study Sample

The study sample was drawn from 5470 hourly employees who worked for a large manufacturing company. Using a sample selection procedure[19] (and L. Yen and D. W. Edington, unpublished data, May, 1990), the study sample model was limited to the 1284 employees who completed the Health Risk Appraisal (HRA) in 1985, whose absenteeism was recorded by the company from 1986 to 1987, who were covered by the company's traditional medical claims plans for the same 2 consecutive years, and who were less likely to have been covered by the other insurance plans (we excluded those employees with nonsingle marital status and those who had no medical claims for either themselves or for their family members during 1986 and 1987). In addition, the employee membership record for those who joined the employer-sponsored fitness center was merged with the sample's data base. When the employer-sponsored fitness center began operation in 1986, of the 1284 employees 211 paid the membership fee.

Health-Related Measures

The health-related measures of this study mainly came from self-reported HRA data.[19] A health center administrated the HRA process and performed blood pressure and total cholesterol screening for those employees who completed the HRA questionnaire. The HRA data were then analyzed by the Fitness Research Center, University of Michigan. The employer paid both the HRA and screening charges.

The HRA used in this study was a CDC-type HRA (originally developed by the Centers for Disease Control) and modified by the Fitness Research Center for the employees and the health center. There were 37 self-reported questions, three clinical measures, and 15 computer projected scores in the HRA. A strict quality control was incorporated with the HRA process to ensure the reliability of the HRA data.[19]

In addition, because several previous studies indicated that the selection of medical insurance options significantly affects the utilization of medical care,[20-22] employees' self-selected medical insurance options (different plans were available at various premiums) in both 1986 and 1987 were included as a health insurance utilization indicator. It was expected this selection would be highly dependent on the employee's past utilization experiences and could significantly relate to the employee's (or family members') health status.

Because several previous studies have reported that a work-site health promotion program resulted in economic savings from absenteeism and medical claims costs,[8-10] we also selected the employer-sponsored fitness center membership status (yes or no: as of July 1986) as a health-related measure. This measure partially reflected an employee's early involvement in the employer-sponsored work-site fitness facility and programs (the center was opened in July 1986).

Economic Measurements

The economic measurements in this study came from two data sources: medical claims expenditures provided by the insurance company and the absenteeism records administrated by the employer.

The medical claims cost data, excluding pregnancy cost, were the averages between the 1986 and adjusted 1987 costs.[19] The causes for these claims were illness or accidents. The adjusted 1987 annual medical claim costs were reflected in 1986 dollar values by adjusting for the 1987 medical inflation rate of 6.6%.[23,24]

The absenteeism data were obtained from the company's daily attendance records for hourly workers. Two major absentee categories that would result in an economic loss for the company and relate to personal health (sick leave days and absent days for medical treatments) were selected to measure absenteeism. In addition, tardiness, regardless of reason, was considered as an indicator of job performance and an economic loss for the company and was included as the third category of absenteeism. Both absence days (8 h/d) and tardy occurrences (4 h/occurrence) were converted into dollar amounts by using the average hourly rates of the employees ($17.12/h for 1986 and $17.60/h for 1987). The 1987 annual absence costs were converted to the 1986 dollar value by adjusting for the 1987 general inflation rate of 3.6%.[23,24] The average annual absence costs were then calculated based on the 1986 and adjusted 1987 costs.

A total annual absenteeism cost was calculated by summing the absence costs of sickness, medical leave, and tardiness. Finally, an overall economic measure (total absenteeism plus medical insurance costs) was created by totaling the dollar amounts of these two measures. All six annual economic measures are listed in Table 1.

As shown in Table 1, the normality measures indicated that without data transformation these six economic measures were severely skewed with a sharp peak (kurtosis). However, after the log 10 transformation (a $10 constant was added before the transformation) the economic measures in log 10 form were close-to-normally distributed with reasonable ranges and variations and acceptable skewness and kurtosis values. Log transformation is the most commonly used technique to reduce the severity of skewness on health economic measures.[24-29] In the current study, the purpose of the log 10 transformation is to permit the use of linear multiple regression analysis techniques to

TABLE 1
Distributions of Major Economic Measures* Among 1284 Employees (Before and after the Log10 Transformation†)

Categories	Causes of Cost	Range		Mean	Standard Deviation	Normality Measure	
		Minimum	Maximum			Skewness	Kurtosis
Annual medical claims costs‡	Illness and accidents	$0.0	$23579.0	$637.46	$1740.50	6.53	56.19
	(Log 10 transformation)	(1.0)	(4.4)	(2.07)	(0.83)	(0.24)	(−0.91)
Annual absenteeism§	Illness	$0.0	$7011.7	$382.04	$636.16	3.70	21.51
	(Log 10 transformation)	(1.0)	(3.8)	(2.03)	(0.81)	(−0.11)	(−1.38)
	Leave for medical treatment	$0.0	$9504.3	$318.14	$916.88	4.81	28.80
	(Log 10 transformation)	(1.0)	(4.0)	(1.51)	(0.84)	(1.25)	(−0.01)
	Tardy	$0.0	$1410.5	$125.16	$181.80	2.43	7.64
	(Log 10 transformation)	(1.0)	(3.2)	(1.76)	(0.61)	(0.05)	(−1.23)
	Total absenteeism	$0.0	$9848.8	$825.35	$1296.40	3.26	13.72
	(Log 10 transformation)	(1.0)	(4.0)	(2.41)	(0.80)	(−0.49)	(−0.73)
Total	Absenteeism plus medical claims	$0.0	$27013.0	$1462.80	$2529.00	4.27	25.27
	(Log 10 transformation)	(1.0)	(4.4)	(2.71)	(0.73)	(−0.58)	(0.11)

* Converted to 1986 dollars with a medical inflation index of 6.6% and a general inflation index of 3.6%.

† To include the employees with "0" dollar amounts, the Log 10 transformation was made after a $10 value was added to the costs for each employee.

‡ Dollar amount paid by the employer.

§ Dollar amount lost by the employer because of employees' absence; $17.12 and $17.60 hourly rates were applied for 1986 and 1987, respectively.

study the multivariate relationships between health-related measures and the economic measures.

Selection of Predictors

Two variables (medical insurance plan selection and the employer-sponsored fitness center status) were selected as "fixed" predictors for subsequent regression analyses. The findings of previous studies have indicated they are important predictors.[8-9, 20-22] The medical plan selection accounted for the number of years an employee selected the low-deductible/high-premium health insurance plan. The fitness center membership status reflected the number of years an employee did not join the center.

Other predictors were selected from the HRA data. We identified the potential predictors from HRA by producing bivariate correlation matrices between the six economic measures (after log 10 transformation) and health-related measures. Fourteen variables were selected because they were at least correlated to two of the six economic measures ($P < .05$). In addition, we also selected self-reported physical activity as a predictor although it did not correlate to any of the economic measures. We were specially interested in its predictability in a multivariate model.[30]

Because of the nature of grouping among the HRA variables, we assigned 15 selected HRA variables into five categories: (1) demographic measures: age and sex; (2) lifestyle behaviors: tobacco use, drug/medication use, and physical activity level; (3) psychological perceptions: stress score, physical health, life and job satisfaction, and social support; (4) biological risks: blood pressures and total cholesterol level; and (5) medical and absence history: personal health problems and self-reported absence days in past 12 months (for 1984).

Among these 15 selected measures, blood pressures and cholesterol levels were clinically determined, and personal health problems and stress scores were recorded according to the combinations of several questions from the HRA.[19] Of the two recorded measures, stress score was calculated from six HRA questions: physical health, life satisfaction, sleep, personal loss, violent argument, and marital status.[30] Personal health problems were calculated by counting any personal medical problems: serious medical problem in past 5 years, rectal bleeding, rectal growth, diabetes, chronic bronchitis, emphysema, or abnormal pap smear (women only).

To limit the number of potential predictors, factor analyses were performed among these 15 selected HRA variables. Based on the analyses, only one factor, perception index, was formed. This index replaced five perception items (stress score, life satisfaction, job satisfaction, social support, and physical health) and was used in subsequent correlation and regression analyses. The index score accounted for the number of times an employee was classified in the high risk category. The Cronbach's α coefficient over these five items was 0.60, which indicated that the reliability and validity of the index was in an acceptable range. By using this index and reducing the number of predictors in regression analyses, we eliminated the potential multicollinearity (at least from the perception measures) in our regression models.

We used actual data codes for age and gender. We recorded the other 11

proposed predictors into a risk-related direction: higher health risks would get higher values (codes). Six of these 11 predictors were coded to a two-point scale: risk present or risk absent. The cut points for these six variables were: 140 mm Hg for systolic blood pressure; 90 mm Hg for diastolic blood pressure; 240 mg/dL for total cholesterol; current smoking cigarettes for smoking tobacco; almost every day or sometimes using drugs or medication to affect mood or help to relax for drug/medication use; and little or none for physical activity. The remaining five variables were interval scales including: personal health problems, absence in past 12 months, perception index, the employer-sponsored fitness center membership status, and the choice of medical insurance option. For these five variables, the higher the health risk, the larger the coding value.

Statistical Modeling

When the 13 predictors were selected, we performed an interaction check. The Pearson product-moment correlation matrices were produced to compare the correlations between the economic measures and each of the predictors according to the levels of another predictor. Marginal correlation differences according to sex and age subgroups were found, indicating a sex/age-involved interaction. Thus, according to the variations among different age/sex subgroups and the sample sizes of the groups, we decided to perform regression analyses by three subgroups: men under 35, men 35 or over, and women.

No strong interactions were detected in the remaining 12 variables and stepwise regressions were performed. Significant predictors were selected to build regression formula based on the t tests of partial r ($P<.10$).

We studied the interrelationships among the six economic measures by correlations. When we studied the relationships between the "original" economic measures (before the log 10 transformation), we computed Spearman's correlation coefficients (RHO). RHO is the standard product-moment correlation coefficient between the ranks of the case values of the two variables.[31] The reason for using RHO correlations was because the "original" economic measures clearly were not normally distributed.[31-32] We also computed Pearson product-moment correlations among these measures after the log transformation because the distributions of the measures after the transformation were close to normal.

To show the associations between the health-related measures and the economic measures, we first computed Pearson correlations to demonstrate the bivariate relationships and then performed multiple regression analyses to examine the multivariate relationships. All the statistical analyses were performed on the Michigan Terminal system using the Michigan Interactive Data Analysis System[31] and the OSIRIS Statistical Analysis and Data Management Software System.[32]

Results

We calculated both RHO coefficients and Pearson correlations, to find the interrelationships among the six economic measures. As shown in Table 2, the RHO correlation coefficients between the total cost measure (absence plus medical claims costs) and its constructed five measures were significant (r ranged from .872 to .451). Because the average annual total absenteeism costs were about 30% higher than the average annual medical claims costs ($825 vs $637, as shown in Table 1) the total cost measure was more strongly correlated to total absenteeism ($r = .87$) than to the medical claims costs ($r = .70$). For the total absenteeism cost measures, illness contributed the most ($r = .83$), followed by medical leave ($r = .66$) and tardy ($r = .59$).

In Table 2, we present a set of Pearson correlations to demonstrate the associations among those economic measures after the log 10 transformation. Indeed, these correlations showed similar trends to the RHO correlations with slight differences in the values for the coefficients. Thus, the comparisons confirmed that the log 10 transformation had not changed the interrelationships among these economic measures.

The bivariate relationships between the six economic measures and the selected 13 predictors for the total employee group are shown in Table 3. The Pearson correlations between the six economic measures and physical activity level and the employer-sponsored fitness center membership status were not significant. Medical plan selection, as expected, was only correlated to medical claim cost, not to absenteeism.

Four of the 13 health-related measures (sex, perception index, self reported 1984 illness absence days, and personal health problems) were significantly related to all the six economic measures. As indicated by the signs of the correlations, female employees, or those employees who negatively perceived their life, job, social support, health, and felt more stress, who had at least one health problem, and who had been absent more in 1984, and who used drug or medication frequently, were more likely to cost more in absenteeism and medical claims expenditures in 1986 and 1987. Age and medical claims were positively correlated, which indicated that older employees were more likely to cost more. In contrast, age and absenteeism were negatively correlated, indicating that older employees were less likely to be absent from work. These opposite age "effects" on the economic measures resulted in the correlation between age and total cost measure becoming insignificant.

Other health-related measures significantly correlated to the absenteeism costs included smoking, drug or medication use, blood pressures, and total cholesterol.

Because the total absenteeism cost was the sum of the three individual absenteeism cost measures, and the correlations were strong (shown in Table 2), we decided to use only the total absenteeism measure in further analyses. In addition, we selected the medical claims costs and total economic

TABLE 2
Correlations among Selected Economic Measures Before and after Log10 Transformation

Correlations	Economic Measures	Correlation Coefficients*				
		Medical Claims Costs	Absenteeism Costs			
			Illness	Medical Leave	Tardy	Total
RHO correlation†	Illness absence	.242				
	Medical leave absence	.480	.369			
	Tardy	.081	.427	.161		
	Total absenteeism	.366	.826	.660	.592	
	Total absence and medical claims	.698	.696	.680	.451	.872
Pearson correlation‡	Illness absence	.239				
	Medical leave absence	.498	.350			
	Tardy	.084	.423	.162		
	Total absenteeism	.359	.814	.611	.623	
	Total absence and medical claims	.704	.671	.624	.475	.845

* All measures were significantly correlated at .005 level.

† Spearman's coefficient of rank correlation were computed before log10 transformation. The correlation (RHO) is the standard product-moment correlation coeffecient between the ranks of the case values of the two variables.

‡ Product-moment correlation coefficients were calculated after log10 transformation. To include employees with "0" claim costs in the analysis, $10 was added to cost for each employee before the log10 transformation.

TABLE 3
Correlations between Selected Economic Measures and Health-Related Measures

Health Related-Measures*	Pearson Correlation with Economic Measures†					
	Medical Claim Costs	Absenteeism Costs				Total Absence and Medical Claims
		Illness	Medical Leave	Tardy	Total	
Medical plan selection	.139	NS‡	NS	NS	NS	NS
Membership status	NS	NS	NS	NS	NS	NS
Sex	.088	.299	.197	.166	.258	.224
Age	.206	−.235	NS	−.286	−.192	NS
Smoking status	NS	.132	.115	.153	.144	.127
Drug/medication use	.150	.122	.175	NS	.151	.180
Physical activity	NS	NS	NS	NS	NS	NS
Perception index	.149	.198	.146	.180	.220	.225
Systolic blood pressure	NS	NS	NS	.111	.097	.088
Diastolic blood pressure	NS	.090	NS	.164	.101	NS
Total cholesterol	NS	.103	NS	.100	.088	NS
Personal health problem	.246	.129	.187	.082	.163	.221
Absence in past 12 months (self-reported for 1984)	.219	.480	.242	.241	.414	.390

* Scoring methods for health-related measures are detailed in text under Methods.

† Only correlation coefficients significant at .05 level are presented.

‡ NS, not significant.

costs as dependent variables. Thus, nine regression analyses were performed (three economic measures with each of three age/sex subgroups). The findings of the analyses are presented in Table 4. Each of the nine models significantly predicts the three economic measures. The multiple R^2 ranged from .125 to .229, which indicates that, depending on the age/sex subgroups and the economic measures (absenteeism or medical claims costs or both), the health-related measures collected in 1985 explains about 12.5% to 22.9% of the variance of the employer's 1986 and 1987 economic costs.

The medical plan selection did significantly predict employees' medical claims costs regardless of age and sex, which is consistent with previous findings.[19,29] However, the medical option selection was not a significant predictor for the costs related to absenteeism or to the total cost.

Self-reported 1984 absence days was a significant predictor in eight of

TABLE 4
Multiple Regression Models

Selected Health-Related Measures*	Coefficients to Predict Economic Measures								
	Medical Claims Costs			Absenteeism			Medical Claims and Absenteeism		
	Men < 35 (N = 455)	Men ≥ 35 (N = 671)	Women All (N = 158)	Men < 35 (N = 455)	Men ≥ 35 (N = 671)	Women All (N = 158)	Men < 35 (N = 455)	Men ≥ 35 (N = 671)	Women All (N = 158)
Constant	.935	NS†	1.222	2.124	1.278	2.252	1.482	1.245	2.180
Medical plan selection	.164	.139	.271	NS	NS	NS	NS	NS	NS
Personal medical problem	.391	.248	.302	.209	.155	NS	.234	.212	NS
Reported 1984 absent days	.080	.084	NS	.256	.158	.135	.205	.132	.080
Age	NS	.025	NS	−.031	NS	NS	−.020	.011	NS
Perception index	NS	.094	NS	NS	.091	.094	.068	.088	.091
Smoking	NS	NS	NS	.150	.233	NS	NS	.218	NS
Drug/medication usage	.247	.224	NS	NS	.173	NS	NS	.213	NS
Physical activity	NS	NS	NS	NS	NS	NS	NS	NS	NS
Membership status	.127	NS	NS	NS	NS	NS	.092	NS	NS
Systolic blood pressure	NS	NS	NS	NS	NS	NS	NS	NS	NS
Diastolic blood pressure	NS	NS	NS	NS	NS	NS	NS	NS	.182
Cholesterol	NS	NS	NS	NS	NS	−.101	NS	NS	−.081
Multiple-R^2‡	.125	.151	.178	.229	.153	.176	.223	.180	.170

* The significant level of t tests for each of the predictors was set at .05. The coefficients presented are predictors that significantly contribute to the predictions of economic measures in the multiple regression models ($P < .05$).

† NS, not significant (partial r). The variable was not selected as a predictor for the regression model.

‡ The F tests for all multiple regression models were significant at .0001 level.

the nine regression models, and self-reported personal medical problems in 1985 could significantly predict 1986 and 1987 economic measures in seven of nine models. These findings indicate the significance of employees' self-reported absence history and current medical problem to the employer's future costs.

Age did not significantly predict economic measures for the female employee group and had a different impact on the economic measures between the two age groups within male employees. As shown by the regression analyses, for male employees under 35, older employees cost less in absenteeism and the total costs than did the younger but not in medical claim costs. In contrast, for men 35 or over, older employees cost more in medical claims and total costs than did the younger but not in absenteeism. This finding clearly demonstrates the interactive effects of age and other predictors on the economic measures and provides supporting evidence for us to divide male employees into two age groups in the model development.

The perception index was a significant predictor in six of the nine models. This finding shows that employees' perceptions about their health, stress, social support, job, and life in 1985 predicts their medical claims cost and absenteeism for 1986 and 1987.

The limited predictability of seven lifestyle-related variables in the nine models was unexpected. Two of these seven predictors, self-reported physical activity level and clinically measured systolic blood pressure, did not significantly contribute to any of the multiple regression models. Diastolic blood pressure was a significant predictor only of total economic costs, and total cholesterol level was negatively associated with absenteeism and total costs for the female employee group, after adjusting for the effects of the other predictors. Smoking tobacco was a significant predictor of absenteeism costs and total costs for the male employees 35 or older, and only for absenteeism for the male employees under 35. Use of drug/medication significantly predicted all the three economic costs for the older male employee group but only predicted medical insurance costs for the younger males. In addition, the employer-sponsored fitness center membership status was a significant predictor of the medical claims costs for the male employees under 35 (members cost less). These findings demonstrate the impact of employees' lifestyle and biological risks on their absenteeism and medical claims may be limited, but they are highly specific to the employees' age and sex subgroups.

Conclusion and Discussion

The health-related measures collected in 1985 significantly predicted employer's economic costs of both medical claims expenditures and absenteeism loss, as shown by the nine regression models presented in Table 4. Six of the 12 selected health-related

measures (age, perception index, personal health problem, self-reported 1984 absences, smoking and drug/medication) contributed significantly to the multiple regressions. Another predictor, medical plan selection significantly predicted medical insurance costs but was not a significant predictor for absenteeism-related costs.

This study demonstrates that health effects on costs are age-sex specific. This finding provides support for health promotion programming and its marketing. For example, because in the male-under-35-years-old employee group, fitness center members cost significantly less in medical claims and total costs (as shown in Table 4), the fitness center may focus on this target group for a cost-containment purpose. In contrast, target marketing for smoking cessation programs may focus on the male employees aged 35 or over, because the employer was more likely to pay more in both absenteeism and medical claims for the smokers in this subgroup. In addition, the finding regarding the age predictability in the different models also suggests that the employer may focus on younger men for absenteeism reduction and older men for medical claim control.

One argument that challenges our current model is related to the selections of the predictors. We used self-reported 1984 absence days to predict the absence days in 1986 and 1987, which were accurately recorded by the employer. We also used self-reported current or past personal health problems (reported in 1985) to predict future medical claims and/or absenteeism costs (1986 and 1987). To clarify the potential conflicts of these predictors to predict similar components, we calculated the partial multiple R^2 for these two predictors versus other predictors from the regression models. The partial multiple R^2 of these two predictors ranged from .03 to .06, compared with partial multiple R^2 of .11 to .21 for other predictors. These findings suggest that the significance of the health-cost relationships was not overly affected by one or two apparently "dominant" health-related measures.

Another concern relates to the findings regarding predictability of the perception index. This index is a composed scale selected from available HRA data. It reflects a person's overview about life in general and his/her environment. In today's highly stressed and diversified society, it should not be surprising that a person's positive perception in life, job, personal health, and stress would result in better job performance, less absenteeism and better mental and spiritual health, less medical claim costs. Thus, the fact that the index significantly predicted costs in six of the nine models indicates the importance of spiritual, mental, and social health to employees' future job performance and medical claims. Therefore, to promote economic benefits, a work-site health promotion program also should emphasize improved employees' spiritual, mental, and social health.

Another major consideration to question with our current model is related to lack of association between the measures of lifestyle and biological risk and the costs, especially medical claim costs. As shown in Table 4, smoking, physical activity, blood pressure, and cholesterol were all nonsignificant predictors of medical claim costs and only were selected significant predictors for absenteeism or total costs. Several factors may contribute to this lack of predictability.

First, the above variables are the key variables in health promotion and disease prevention, but improvement in these measures may not necessarily directly reduce medical claim costs. As a nation, during the past three decades, there has been significant public health improvement in smoking and to a lesser extent in blood pressure and cholesterol control. This improvement has resulted in a significant reduction in mortality from lifestyle-related diseases, i.e., heart diseases and stroke. However, during the same period there has occurred more than a twentyfold increase in national health care expenditure.[33] Thus, from the national trend, we surmise that health-risk reduction may not of itself lead to a medical cost reduction.

Second, these key variables are the major precursors of cardiovascular disease, which accounts for more than half the deaths in the U.S. adult population. However, these diseases account for only about 12% of the national direct medical costs among the 18 to 64 year old population.[34] In our study sample, cardiovascular diseases share a similar portion of total medical costs, accounting for 13.6% of total costs from 1984 through 1987. Although these precursors significantly contribute to the onset of cardiovascular diseases, they may not significantly contribute to total medical costs because of the limited proportion of the costs caused by these diseases.

Third, factors affecting medical costs are very complex and the supply-demand factor plays a major contributing role. This may be one of the main reasons the annual changeable medical plan selection option significantly predicted medical claim costs. This selection, by the employees, is partially determined by the previous year medical claim demand or the coming year anticipated demand. From a previous study,[19] we found that the total number of risk factors were more significantly associated with medical claim costs than any other single risk factor alone. The perception index, in fact, is a sum of risk factors related to perception measures. We were able to use this index in the regression modeling because its reliability and validity for index composition were in an acceptable range. In contrast, we were unable to form an index for lifestyles, for biological risks, or for overall health because of the independence of their underlying measures showed in factor analyses. Thus, by using single behavior or biological risk variables in the models, we may have discounted the effect of these key variables on medical claim costs.

Fourth, the 2-year short-term cost data and limited sample size may affect the predictability of these key

variables. In addition, the effect of the employer-sponsored work-site health promotion program, which was focused on risk reduction in these key measures, may have influenced workers in our study to make significant changes from their 1985 reported lifestyle behaviors and biological risks.

Although there are weaknesses, partially due to the limited data source, this study demonstrates that multiple regression models based on HRA information significantly predict future medical claims costs, absenteeism, or both. Indeed, the findings of this study provide evidence that the health-cost associations are significant and age-sex specific. It indicates that promoting employees' health should be an important management strategy for reducing the business costs from both absenteeism and medical claims for an employee population. It also suggests that to focus on only cardiovascular disease risk factors may not be comprehensive enough for cost containment purposes.

Acknowledgments

Preparation of this article was assisted by a research contract from Steelcase Incorporated and was partially funded by the Fitness Research Center, the University of Michigan, Ann Arbor.

References

1. Edington DW, Edington M, Yen L. The formula for proving your program's worth. *Employee Services Management.* 1988–89;31:12–17.
2. Wolfe RA, Ulrich DO. Employee health management programs: review, critique, and research agenda. *J Management.* 1987;13:603–615.
3. Warner KE. Selling health promotion to corporate America: uses and abuses of the economic argument. *Health Educ Q.* 1987;14:39–55.
4. Fielding JE. Health promotion and disease prevention at the worksite. *Ann Rev Public Health.* 1984;5:237–265.
5. Jones RC, Bly JL, Richardson JE. A study of a work site health promotion program and absenteeism. *J Occup Med.* 1990;32:95–99.
6. Blair SN, Smith M, Collingwood TR, et al. Health promotion for educators: impact on absenteeism. *Prev Med.* 1986;15:166–175.
7. Bell BC, Blanker D. The effects of a worksite fitness program on employee absenteeism. *Health Values.* 1989;13:3–11.
8. Baun W, Bernacki EJ, Tsai SP. A preliminary investigation: effects of a corporate fitness program on absenteeism and health care cost. *J Occup Med.* 1986; 28:18–22.
9. Reed RW, Mulvaney DE, Bellingham RE, et al. *Health Promotion Service Evaluation and Impact Study.* The Blue Cross and Blue Shield of Indiana, Benchmark Press; 1986.
10. Warner KE, Wickizer TM, Wolfe RA, et al. Economic Implications of Workplace Health Promotion Program: Review of the Literature. *J Occup Med.* 1988; 30:106–112.
11. Golaszewski T, Lynch W, Clearie A, et al. The relationship between retrospective health insurance claims and a health risk appraisal-generated measure of health status. *J Occup Med.* 1989; 31:262–264.
12. Bly JL, Jones RC, Richardson JE. Impact of worksite health promotion on health care costs and utilization. *JAMA* 1986;256:3235–3240.
13. Brink SD. *Health Risks and Behavior: The Impact on Medical Costs.* A preliminary study by Milliman & Robertson, Inc. and Control Data Corp; 1986.
14. Jose WS, Anderson DR, Haight SA. The stay well strategy for health care cost containment. In: Opatz JP, ed. *Health Promotion Evaluation.* Steven Point, Wis: National Wellness Association Publication; April 1987.
15. Gordan T, Kannal WB. Multiple risk functions for predicting coronary heart disease: the concept, accuracy, and application. *Am Heart J.* 1982;103:1031–1039.
16. Leaverton PE, Sorlie PD, Kleinman JC, et al. Representativeness of the Framingham risk model for coronary heart disease mortality: a comparison with a national cohort study. *J Chron Dis.* 1987;40:775–784.
17. Buring JE, Evans DA, Fiore M, et al. Occupation and risk of death from coronary heart disease. *JAMA* 1987; 258:791–792.
18. Weistein MC, Coxson PG, Williams LW, et al. Forecasting coronary heart disease incidence, mortality, and cost: the coronary heart disease policy model. *Am J Public Health.* 1987;77:1417–1426.
19. Yen L, Edington DW, Witting P. Associations between HRA scores and prospective medical claim costs in a manufacturing company. *Am J Health Prom.* 1991;6:23–30. The University of Michigan; 1981.
20. Newhouse JP, Manning WG, Morris CN, et al. Some interim results from a controlled trial of cost sharing in health insurance. *N Engl J Med.* 1981;305: 1501–1507.
21. Valdez RB, Brook RH, Rogers WH, et al. Consequences of cost-sharing for children's health. *Pediatrics.* 1985;75:952–961.
22. Leibowitz A, Manning WG, Keeler EB, et al. Effect of cost-sharing on the use of medical services by children: interim results from a randomized controlled trial. *Pediatrics.* 1985;75:942–951.
23. Gramlich EM. *Benefit-Cost Analysis for Government Program.* Englewood Cliffs, NJ: Prentice-Hall; 1981.
24. Metropolitan Life and Affiliated Companies. Trends in medical care costs. *Stat Bull Metrop Insur Co.* 1988;68(1):2–8.
25. Box GEP, Cox DR. An analysis of transformations. *J R Stat Soc.* 1964;Ser B 24:211–252.
26. Duan N: Smearing estimate: a nonparametric retransformation method. *J Am Stat Assoc.* 1983;78:605–610.
27. Manning WG, Duan N, Rogers WH. Monte Carlo evidence on the chronic between sample selection and two-part models. *J Econometrics.* 1987;35:1–24.
28. Duan N, Manning WG, Morris CN, et al. A comparison of alternative models for the demand for medical care. *J Business Econ Stat.* 1983;1:115–126.
29. Connell FA, Diehr P, Hart LG. The use of large data bases in health care studies. *Annu Rev Public Health.* 1987;8:51–74.
30. Edington MP, Yen L. The association between CDC/HRA S-Scale and stress related measures. *Proc 22nd Annu Meet Soc Prospect Med.* 1986;22:45–56.
31. Fox DJ, Guire KE. *Documentation for MIDAS.* Ann Arbor, Mich: Statistical Research Laboratory, The University of Michigan; 1976.
32. Institute for Social Research. *OSIRIS IV: Statistical Analysis and Data Management System.* Ann Arbor, Mich: DATE.
33. Levit KR, Lazenby HC, Letsch SW, et al. National health care spending, 1989. *Health Affairs.* 1991;Spring:118–130.
34. US Department of Health and Human Services. *Costs of Illness United States, 1980.* 1986;Ser C, Analytic Report 10.