CME Article #3

Association Between Wellness Score from a Health Risk Appraisal and Prospective Medical Claims Costs

Louis Yen, PhD
Timothy McDonald, MHSA
David Hirschland, MBA
Dee W Edington, PhD

Learning Objectives

- Explain how the updated health risk appraisal used in this study differs from previous versions, and recall the components of the wellness score.
- Point out the predictive factors included in the multivariate regression model and note how successfully they explained the variance in medical claims costs.
- Recall the potential predictive factors that did - and those that did not - correlate significantly with actual medical claims costs, and those that were most predictive.

Abstract

*This study examines how wellness scores generated from the Health Risk Appraisal are associated with prospective medical claims costs, controlling for age, gender, and disease status. The study was conducted among 19,861 active employees who participated in the Health Risk Appraisal and selected indemnity or PPO medical plans from 1996 to 1998. A multiple regression model based on group averages of age, gender, disease status, and wellness score levels was developed among a randomly selected screening subsample (*n = *10,172) from the study sample. Total medical claim costs of −$56, $88, and $3574 were estimated for one additional point on the wellness score, 1 year of additional age, and an existing major disease, respectively. No significant differences were found between the model predicted and actual medical claims costs for the individuals in both screening and calibration (*n = *9689) subsamples.* (J Occup Environ Med. *2003; 45:1049–1057*)

From the Health Management Research Center, University of Michigan, Ann Arbor, Michigan (Dr Yen, Dr Edington); General Motors Corporation, Detroit, Michigan (Mr McDonald); and International Union, United Automobile, Aerospace and Agricultural, Implement Workers of American, UAW, Detroit, Michigan (Mr Hirschland).

L. Yen has no commercial interest related to this article.

Address correspondence to: Louis Yen, PhD, Health Management Research Center, University of Michigan, 1027 E. Huron Street, Ann Arbor, MI 48104-1688; E-mail address: louisyen@umich.edu.

DOI: 10.1097/01.jom.0000088875.85321.b9

The Health Risk Appraisal (HRA) is one of the most widely used health assessment tools in the field of worksite health promotion and disease prevention.[1–6] Over the past 22 years, using the HRA as a tool to evaluate health status and health risks at the individual or organizational level, the University of Michigan Health Management Research Center (UM-HMRC) has continuously modified the HRA questionnaires and developed profile reports to meet the needs of various corporations. Currently, the UM-HMRC has processed over two million HRAs for those corporations.

Researchers from the UM-HMRC have documented the association between the health risk factors from HRA and average annual medical claims costs among various employee populations in several research publications.[7–13] These findings were in agreement with other worksite studies on the health risk and medical cost relationships.[14–19]

In the 1990s, a newly comprised wellness score was developed by the UM-HMRC, to measure individuals' overall health, as reflected by medical claims costs. This score has been used to replace the traditional HRA measurements of appraised and achievable age in the HRA profile reports.[1–2,4–6] The use of this score was in lieu of the trends of development of various health assessment scales and measurements during the 1980s, as the promotion of health and well-being became an increasingly significant social commitment.[4–6] As a result, a variety of wellness scores or health indexes were developed in different HRA instruments as a measure to reflect individual's overall physical, mental, emotional, social, and spiritual health. Since then, more and more health practitioners and behavioral health researchers have used these newly developed measures as a resource to encourage and motivate people to change and maintain healthy lifestyles. These practices emphasize self-responsibility as a prime component of well-being and cease to measure health strictly from an illness-oriented perspective. The move reflected that well-being as a health perspective had been recognized by both health professional and general public.[4–6] Although the development and validity of the wellness scores and indexes used has been documented, few studies have studied the relationships between the health outcome measures and these newly developed measurements used in HRA programs.[2,4–6]

The wellness score developed by the UM-HMRC is a composed score from a variety of health measures. The theoretical basis for this development was the research findings between personal health measures and medical claims costs from various worksite health studies.[7–19] The practical guide for this scoring system was the quantitative relationships estimated between the selected health measures and health outcomes (ie, mortality, morbidity, and medical claims costs) from the Centers for Disease Control, Carter Center, and UM-HMRC's long-term research, from both published[7–13,20–22] and unpublished data, and the national guidelines in health promotion and preventive services.[23,24]

The purpose of this article is to examine how the wellness score generated from current UM-HMRC's HRA predicts prospective medical claims costs for a large employee population. This is the first study to directly link the wellness score used in an HRA to an important health outcome measure: medical claims costs.

Methods

Sample and Measurement

Study Sample. In May 1996, General Motors (GM) and the International Union, United Automobile, Aerospace and Agricultural, Implement Workers of America (UAW) offered active employees, retired employees, and all adult dependents an opportunity to participate in the LifeSteps HRA.[25,26] Approximately 45,000 active employees participated in the HRA program in the first year.

Individual medical claims cost data were available from indemnity insurance accounts and preferred provider organizations (indemnity/PPO) and only those active employees who selected these medical plans during 1996 to 1998 were included in the study. Active employees aged 65 and over in 1998 (the conclusion of this study) were excluded because of enrollment in Medicare. These criteria reduced the sample to 19,861 active employees who completed an HRA in the first program year, worked continuously for GM, were under the age of 65, and enrolled in the indemnity/PPO medical plan for the entire period from 1996 to 1998. The study group (n = 19,681) accounted for approximately 40% HRA participants and 6% of the total eligible active employees. The age and gender distributions of the study sample were similar to the age and gender distribution of the total active employee population of UAW-GM during the study period of 1996 to 1998.

HRA Instrument and Wellness Score. The HRA used in this study was developed by the UM-HMRC for the LifeSteps program of GM and UAW.[25,26] The HRA measures personal health status of HRA participants and provides recommendations for health behavioral and preventive service practices.[23,24] To summarize the health status for each LifeSteps HRA participant, a wellness score was developed as an overall health indicator.

The wellness score is generated from three major components: behavioral health risks; mortality risks; and preventive services usage. Behavioral health risks are weighted the most among the three components in the wellness score and preventive services weighted the least. The behavioral health risks are selected from 10 variables that demonstrate strong associations with future medical claims costs as determined by multiple research studies.[7–12] These variables include smoking status, physical activity, alcohol consumption, safety belt usage, blood pressure, total cholesterol, high-density lipoprotein cholesterol, body weight, illness days, and self-assessment of health. The mortality risks are estimated according to the algorithms developed by CDC and Carter Center[1–3,20] and modified and updated by the UM-HMRC. It is calculated as a function of the rates between achievable and appraised probabilities of the deaths from all causes in the next 10 years according to a HRA participant's age, gender, and health risks. The preventive services selected are based on the findings and recommendations of the US Preventive Services Task Force Guideline according to participants' age and gender.[24]

Disease Status. In the current study, disease status was defined as any existing condition from the self-reported diseases that were significantly associated with high medical claims costs. These diseases included heart disease, past stroke, cancer, diabetes, emphysema, and/or chronic bronchitis.

Medical Claims Costs. Medical claims cost data for the people who selected the indemnity/PPO medical plans were provided to UM-HMRC. The medical care providers' charged amounts were recorded on a per-claim basis using the claim incurred date. Respective individual claims were summed to determine the annual cost for health services per person per year. Similarly, the pharmaceutical charged amounts were recorded on a per-prescription basis using the prescription filled date. Respective individual prescription charges were totaled to calculate the annual cost for drug/medication per person per year. The charges for health services and drug/medication were added to create the total medical claims costs. The charged amounts by the medical care providers reflected the cost of the medical services used by an individual without taking into account the deductible and discount proportions significantly varied by different medical benefit plans and care providers.

Using similar methodology as other studies,[7–13] we studied the medical claims costs based on a 2-year annual average. It was assumed that the average annual costs over a 2-year period would be a more consistent measure than the single-year costs to reflect a person's health related medical utilization pattern.

Using the medical inflation rate of the respective year and present value formula,[27] each annual cost was converted to 1999 dollar value and an average annual cost was then calculated. Therefore, the medical claims costs used in the current study represent the dollar amounts medical providers charged annually for a particular individual from July 1, 1996 to June 30, 1998, expressed in 1999 dollar value.

Statistical Analysis

The current study selected a unique statistical approach to study the associations related to medical claims costs. First, the relationships between selected variables from HRA and medical claims costs were examined through both Median tests and Analysis of Covariance.[28–30]

Second, the study sample was split into two subsamples for cross-validation.[29–31] Each individual was assigned into one of the two subsamples using the last digit of the individual's home zip code. Five numbers were randomly picked from the 10 possible numerical choices to split the sample. As a result, two subsamples were selected with slightly different sample sizes. The group with more people (n = 10,172) was assigned as the screening sample for regression model development.[29–31] The other 9689 individuals were reserved as a calibration sample for cross-validation.[29–31]

Third, among the screening sample (n = 10,172) 96 groups were formed and the group averages were calculated according to each individual's age, gender, disease status, and wellness score levels based on the grouping concept.[32–36] Using the logic of hierarchical linear modeling,[37–39] a parametric regression model was developed on the group data.

Fourth, the regression equation from the group data of the screening sample was applied to each individual in both the screening and the calibration sub-samples.[29–31] Predicted costs for each individual were estimated according to the equation.[29–30]

Finally, a cross-validation was made by performing the paired t test comparisons and Spearman's rank correlations (r_s) between the actual costs and predicted costs in both screening and calibration samples.[28–31]

The details of the above five steps and its background were presented in the Appendix. All statistical analyses were performed using SAS.[40,41]

Results

Associations Between Selected Variables and Medical Claims Costs

The age and gender of the study sample are presented in Table 1. There were over 75% males. The average age in 1998 for the study sample was 46 years, with 46.3% aged 45 to 54 and 10.4% over age 54. As stated in previous section, these distribution were similar to the age and gender distribution of the approximate 300,000 active employee population of UAW-GM during the study period of 1996 to 1998.

Table 1 describes the average annual medical claims costs according to the four selected variables: age, gender, disease status, and wellness core range. Although medical claims costs were severely skewed, the unadjusted means shown in Table 1 indicated that there were significant mean cost differences (all $P < 0.001$) according to the categories within each of the four selected variables without controlling for the other variables' effects. The descriptive statistics show that being female, older, with existing disease, or at a lower wellness score level, is associated with higher average medical claims costs than being male, younger, without disease, or having higher wellness scores.

As shown in Table 1, using median tests to examine the relationships with the severely skewed medical claims costs,[28] we confirmed that there were statistically significant differences in medical claims costs by gender, disease status, age group, and wellness score levels (all $P < 0.001$). The analysis of covariance further confirmed that there were significant group differences among the adjusted means according to the levels of one of the four selected variables when the effects of the other three variables were adjusted (all $P < 0.001$).[29,30,40]

Developing Multivariate Regression Models to Predict Costs

As described in the methods session, the study sample was split into two groups. The group averages were calculated among the screening group of 10,172 individuals to develop a regression model.[29–31] Based on the data from the screening subsample, a new sample of 96 groups was formed according to their age, gender, disease status, and wellness score. In absence of the variances

TABLE 1
Association Between the Select Variables and Medical Claim Costs*

Groups Described	*N* (%)	Measures of Medical Claim Costs*		
		Mean** (unadj.)	Median**	Mean*** (adjusted)
Overall Study Sample	19,861 (100.%)	$3,237	$1,251	N/A
Grouped by Gender				
Male	14,940 (75.2%)	$2,955	$1,027	$2,931
Female	4,921 (24.8%)	4,092	2,091	4,165
Grouped by Age				
Younger than 45	8,605 (43.3%)	$2,617	$ 914	$2,728
45 to 54 years old	9,186 (46.3%)	3,522	1,437	3,468
55 years or older	2,070 (10.4%)	4,544	1,947	4,322
Grouped by Disease Status***				
No self-reported disease	17,597 (88.6%)	$2,779	$1,066	$2,804
With existing disease****	2,264 (11.4%)	6,792	3,556	6,595
Grouped by Wellness Score				
<60.3	2,464 (12.4%)	$3,237	$1,599	$4,120
60.3–69.5	2,481 (12.5%)	3,611	1,557	3,753
69.6–76.2	2,486 (12.5%)	3,540	1,444	3,643
76.3–81.2	2,417 (12.2%)	3,404	1,356	3,408
81.3–85.4	2,533 (12.7%)	3,253	1,212	3,241
85.5–89.2	2,491 (12.5%)	2,906	1,089	2,832
89.3–92.9	2,446 (12.3%)	2,710	1,056	2,596
>92.9	2,543 (12.8%)	2,544	867	2,324

* The medical claims costs are the average yearly charges for medical care of 1996 and 1997 adjusted to 1999 dollar value.
** There were significant differences in mean medical claims costs by Analysis of Variance ($p < .001$) and median by Median Test ($p < .01$) for all four variables.
*** There were significant differences in adjusted mean medical claims costs by Analysis of Covariance, controlling for the other three predictors' effects ($p < .001$).
****Those HRA participants who reported having heart disease, past stroke, cancer, diabetes, emphysema and/or chronic bronchitis.

TABLE 2
Multivariate Models: Predicted Average Group Claims Costs From A Group Sample ($N = 96$)

Results from Group Averages	Regression Based on Group Average from 96 Groups*					
	Total Medical		Health Services		Drug/Medication	
	Estimate	P-value	Estimate	P-value	Estimate	P-value
Multiple R-Square	0.595	0.0001	0.526	0.0001	0.611	0.0001
Predictors						
Gender (female)	$ 438	0.2312	$ 274	0.4430	$165	0.0033
Year of Age (one-year older)	$ 88	0.0007	$ 72	0.0039	$ 16	0.0001
Disease Status (with disease)	$3,574	0.0001	$3,011	0.0001	$563	0.0001
Wellness Score Point (one-point higher)	−$ 56	0.0001	−$ 52	0.0001	−$ 5	0.0181
Intercept	$4,175	0.0052	$4,020	0.0058	$155	0.4811

* Multiple regression was performed on the second layer sample ($N = 96$), representing 10,172 individual HRA responders.

among the individuals within the homogenous group, the distributions of medical claims costs among the 96 groups became less skewed (skewness reduced from 11.47 to 2.86 and kurtosis reduced from 385 to 12). A regression model, using wellness score, disease status, age and gender to predict medical claims costs, was then developed for the new sample. The results are presented in Table 2.

As shown in Table 2, with the four predictors, the models explain 60%, 53%, and 61% of the variance in health services, drug/medication, and overall medical claims costs, respectively. This finding indicates that the group based regression model from the selected four predictors can successfully explain over 50% of the cost variations among the 96 groups in the screening subsample.

The model shows that for a given homogenous group (as reflected by a unique group of individuals), an existing disease condition (heart diseases, cancer, past stroke, diabetes, emphysema, and/or chronic bronchitis), an additional year of age, and one-point improvement in wellness score equates to $3574, $88, and −$56 of the total medical claims per year, respectively. Gender, however,

TABLE 3
Comparisons of Average Annual Medical Claims Costs Between Actual Charges and Predicted Charges

	Comparison between Actual and Predicted Costs Results					
	Annual Average		T-Test Result			
Measurements	Actual	Predicted	Diff	T	r_s**	Prob >[T]
Results from Screenings Sub-sample: N = 10,172						
Annual Medical Claims (Total)	$3,306	$3,289	$17	0.26	.7965	.281
Health Service Charges	$2.780	$2,772	$ 8	0.12	.9047	.227
Drug/Medication Charges	$ 526	$ 516	$10	1.12	.2622	.359
Results from Calibration Sub-sample: *N = 9,689						
Annual Medical Claims (Total)	$3,164	$3,259	−$95	−1.54	.1230	.267
Health Service Charges	$2,661	$2,748	−$87	−1.49	.1355	.217
Drug/Medication Charges	$ 503	$ 511	−$ 8	−0.88	.3783	.342

* The calibration group (N = 9,689) were reserved for the verification of the prediction model and excluded from the model development as shown in Table 2.

** The Spearman Correlation Coefficients between actual and predicted charges ($p < .001$).

was not significantly related to health services and total medical claims costs ($P > 0.10$). However, females had $165 more drug/medication charges than that of males.

Cross-Validation

To verify the models' predictability based on the group data, the predicted costs in health services, drug/medication, and overall medical claims costs were calculated for each of the 19,861 individuals in the study sample according to the regression coefficients (as shown in Table 2) and each individual's age, gender, disease status, and wellness score. The comparisons between the predicted and actual medical claims costs among the screening (n = 10,172) and the calibration subsamples (n = 9689), are shown in Table 3.

The average predicted annual costs in the total medical claims for the screening and calibration subsamples are $3289 and $3259, respectively. The predicted costs are $17 or 0.51% lower than the actual average annual charges of $3306 for the screening sample and $95 or 3% higher than the actual average annual charges of $3164 for the calibration sample, respectively. These −$17 and +$95 differences between the predicted and actual total costs were not statistically significant as determined by the paired t tests for the respective subsamples. The r_s between the predicted and actual charges for both groups is statistically related ($P < 0.001$) at 0.28 and 0.27, respectively. Similarly, the differences between predicted and actual costs for both health service and drug/medication claims are not significant ($P > 0.10$), and are within 3% of the actual average costs.

Discussion

This study is the first study to show that the wellness score (developed by the UM-HMRC), used by an HRA to measure an individuals' overall health, is significantly associated with the HRA respondent's prospective short-term medical claims costs. The comparisons between the predicted and actual medical claims costs among both screening and calibration groups confirm that wellness scores, age, gender, and disease status can be used to estimate the incurred medical claims charges within 3% of the actual charges.[29–31] This shows that the wellness score can be used to estimate future short-term medical claims costs among the HRA respondents. This estimation would be more accurate than the traditional estimation by age, gender, and smoking behavior, and show the significant association of health and disease status to the future costs.

Controlling for disease status, age, and gender, each additional point on the wellness score equals −$52, −$5, and −$56 of health services, drug/medication, and total annual medical claims costs, respectively. Although there have been numerous studies using HRA data to estimate a risk-cost relationship,[7–19] the current study is the first to establish and verify an inverse quantifiable relationship between the wellness score and average annual medical claims costs within a 2-year period.

Two unique statistical approaches were used in the current study. First, the regression model used to predict health-related individual medical claims costs was based on group data. To use group data to describe a relationship has been a general practice in statistical science and other fields,[32–36] where group averages were calculated to smooth the dependent variables' curves according to different time intervals. Within a given the time interval, the larger the measurement unit, the smaller the variance of the dependent variable, and the better the fit of the regression formula. Similarly, in the current study, when the study sample units was changed from individuals (n = 10,172) to homogeny groups (n =

96) according to health and demographic characteristics, the skewness measures of the total average annual medical claims cost were reduced from 11.47 by individuals to 2.86 by groups and kurtosis reduced from 385 to 12. Focus on counting the variances "between" groups and discounting the variances "within" group, the group-based regression explained over 50% of the "between" group variances of the medical claims costs, a level of explanation that has not been reached by individual-based analyses in other studies.[7–19,42–44] In fact, focusing on the between-group variances, also agrees with the logic of hierarchical linear modeling.[37–39]

Second, the current study using cross validation revealed the regression modeling and verified the quantifiable relationship established between the wellness score and average annual medical claims costs. To split a study sample into a screening and a calibration subsample is a general procedure used for cross validation for the establishment of statistical relationships.[29–31] As a result, no significant differences ($P > 0.10$) were found between the predicted and actual medical claims among the screening subsample ($n = 10{,}172$) validated that the group-based regression model can be successfully applied to the individual. Furthermore, no significant differences ($P >$ 0.10) between actual and predicted medical claims costs were found among the individuals in the calibration sub-sample ($n = 9689$).[29–31] The findings crossly validated the statistical relationships between age, gender, wellness score, disease status and medical claims costs estimated from a screening subsample that can be re-verified from a reserved calibration subsample.

The current study quantifies the relationships of age, disease status, and wellness score to medical claims costs. These qualified relationships are useful references for the strategic planning for worksite health management programs. According to these qualified relationships, average medical claims costs for HRA participants can also be more precisely estimated than only using age and gender. Thus, the HRA can be considered as a valid measurement tool to estimate prospective medical claims costs with additional health information. Since the differences between predicted and actual costs are not statistically significant after cross validation, health care industries may consider adopting HRA as one of their prediction tools to estimate future medical claims costs.

The regression models also shows the dominant effects of existing diseases on medical claims costs. However, it also shows for those with an existing disease, higher wellness scores were associated with lower medical claims costs. This finding suggests that it is still possible for persons with existing disease to contain their medical claims costs. In agreement with a previous UM-HMRC study,[13] the current study provides evidence on potential economic incentives for a disease management program. For medical cost containment purposes, it would be economically beneficial to integrate wellness and disease management programs.

As with most worksite health management studies,[7–19] sampling is a main limitation for the current study. The current study focused on the 19,861 active employees who completed a LifeSteps HRA, were under age 65, and enrolled in the Indemnity/PPO medical plan for the entire period of 1996 to 1998. First, the findings cannot be applied to the dependents under 65 years of age in the same plan because there were significant cost differences between the employees and their dependents with same age. Second, the findings cannot be applied to people 65 or older, since the group was not included in the current study and likely would be enrolled in Medicare. Third, the findings from the current study cannot be directly applied to the active employees under the age of 65 who were not in the Indemnity/PPO medical plan for the entire period of 1996 to 1998, since the medical claims costs information were not available. The HMO subscribers are administrated at a pre-paid base. All of the studies related to medical economics and worksite health economics[7–19,42–44] are focused on the indemnity/PPO medical plan subscribers. Therefore, to apply the relationships of health, age, gender, and disease status to medical claim costs from the current study to other populations, such as dependents, HMO subscribers, retirees, and older adults aged 65 or older, should be done with caution. Further research is needed to expand the study to other populations.

The study sample uses self-reported HRA data from either mailed HRA or onsite biometric screening.[25,26] The validity and reliability of HRAs has been addressed in our previous publications.[2,21,22] In a previous study of the same UAW-GM active employees,[25] we concluded that the risk measures from the HRA were reliable. For the current study, we consider self-reported HRAs adequately reflect personal health.

In conclusion, this study shows that a wellness score, generated from an HRA, is significantly associated with prospective medical claims costs among active employees under age 65. It showed a lower cost of $56 per year for a one-point wellness score gain and a higher cost of $88 for 1 year of additional age. Further study is needed to examine the longitudinal associations between changes in wellness scores and age and the changes in medical claims costs from repeat HRA participants.

Appendix

Background

As with most studies involved with medical claims costs,[7–19,42–44] the average annual medical claims costs is the most commonly used statistic in health service research.

The nonparametric tests on cost status or parametric tests on the log or cube root transformed cost averages are the most commonly used techniques to statistically test the relationships between medical claims costs and other variables.[7,13–19]

The current study used a statistical approach to study the associations related to medical claims costs. Because of only 5% of the employees during the 2-year period had zero claims costs in the current study, the complex models, such as two-part modeling,[14–16,42–44] would not significantly improve the predictability. With a large study sample in the current study, the individuals with similar demographic and health characteristics were grouped together and a regression model based on group data was developed. To validate this approach, the study sample was also split into screening and calibration sub-samples for cross-validation.[29–31] Four variables, age, gender, disease status, and wellness score, were selected as potential predictors for the regression model based on the results from previous studies.[7–12]

Because both age and wellness scores are continuous variables, we recoded the two variables into categorical variables according to the age range and wellness score levels. The age ranges were determined according to a generally used age grouping method to get the midpoint of every 10-year age range. The eight levels for the wellness score were based on the percentile distributions of the study sample. Each wellness score level represented approximate one-eighth of the study sample. The eight divisions were the appropriate levels where there would be significant differences in medical claims costs among each level of wellness scores.

Both parametric and nonparametric statistic tests were performed to show the significance of the associations between the four selected variables and medical claims cost measures.[34,35] The analyses provided statistical evidence for the selection of predictors in the regression model.

Split Study Sample

The study sample of 19,861 active employees was split into two subsamples: a screening sample and a calibration sample.[29–31] Each individual was assigned into the two sub-samples using the last digit of the individual's home zip code. Five numbers were randomly picked from the 10 possible numerical choices. Any employees whose last digit of the home zip code was in one of these five numbers were selected into a subsample. Employees not selected formed the second subsample.

As a result, two subsamples were selected with different sample sizes. The group with more people (n = 10,172) was assigned as the screening sample to develop a regression model. The other 9689 individuals were reserved as a calibration sample for cross-validation. The difference in group size was related to the number of individuals residing within the last digit of five-digit zip code area.

Regression Model Development from Group Data

Based on the group average concept,[32–36] 96 groups were formed from the 10,172 individuals in the screening sample according to the age-ranges (<45; 45–54; 55+), gender (male or female), disease status (exist or none), and wellness score ranges (eight groups according to the values of wellness score). The averages, including medical claims costs, age, and wellness scores, were calculated for each of the 96 groups. These group averages were used to establish a multivariate regression model. By using the grouping data, the problem of the severely skewed cost measures at individual levels was significantly reduced.[32–36]

Using the logic of hierarchical linear modeling,[37,38] a parametric regression model was then developed to describe the relationships between wellness scores and medical claims costs based on the grouping data.

Calculation of Predicted Medical Claims Costs for Individuals

Using the data (n = 96) drawn from the screening subsample, the multiple regression models were developed. In the regression formula, the coefficients for age, gender, disease status, and wellness score were estimated. The predicted medical claims costs for each individual in the study sample (n = 19,861) were then calculated based on these regression coefficients from the group data and the person's age, gender, disease status, and wellness score.

Cross-Validation

To examine whether the mean differences for the individuals between predicted medical claims costs calculated according to the regression equation and actual medical claims costs were significantly different from zero, t tests were performed for both the screening (n = 10,172) and calibration (n = 9689) sub-samples.[29–31]

The correlations between the predicted and actual medical claim costs were also calculated. Because the actual medical claims costs were not normally distributed, Spearman's correlation coefficients (r_s)–the standard product-moment correlation coefficient,[28,29] between the ranks of individual values of the predicted and actual cost measures were computed to show the associations between these two measures.

Discussion

The unique statistical approaches used group averages to develop a regression model and to use split sub-samples to cross valid the predicting results at the individual level.

The skewness and variance of the cost data was significantly reduced when the measurement of unit changed from individual to group. For the total average annual medical claims costs, the skewness measure was reduced from 11.47 at the individual level to 2.86 at the group level and kurtosis was reduced from 385 to 12. The average and standard deviation was $3306 and $6815 at the individual level (n = 10,172), but $5117 and $3242 at the group level (n = 96), respectively. The curve of cost distributions was changed to close to normal distribution. Thus, the group based modeling resulted in less violations of the necessary statistical assumptions for regression analyses.

In addition, using group data also weighted the sample by cost. The group sample (n = 96) was generated with an average of 106 individuals per group based on their similar characteristics in gender, disease status, age-range, and wellness score level. Although the numbers of individuals in each group varied, each group was treated the same in the grouping sample. As discussed above, there was in an increase in population average from $3306 from the individual data to $5117 from the group data, which indicated the individuals with high-costs were over represented in the group sample. As a result, the group sample produced a better regression fit than the individual-based sample and explained as much as 50% of variance in medical claims costs between the groups with unique age, gender, wellness score level, and disease status. In contrast, using individual based data from the screening subsample (n = 10,172), the regression model could only explain 5% of the variance on the actual costs and 10% of the variances on the log transformed cost measures. Thus, the explanation power was significantly improved from the grouping sample and was at least doubled in the explanation power from previous studies that used individual-based medical claims costs data.[7–19,42–44]

The findings validated the statistical approach used in the current study. The paired t tests showed no significant statistical differences between predicted and actual costs measures, it indicated that the relationships estimated between age, gender, wellness score, and disease status and future medical claims costs through a group-based model among the screening subsample can be successfully applied to the individuals in both the screening and calibration subsamples. Furthermore, through computing the r_s,[28,29] the significant associations were found between the ranks of the predicted and actual costs at the individual levels for both screening and calibration subsamples.

References

1. Robbins LC, Hall J. *How to Practice Prospective Medicine.* Indianapolis, IN: Methodist Hospital of Indiana; 1970.
2. Edington DW, Yen L, Braunstein A. The reliability and validity of HRAs. In: Hyner G, Peterson K, Travis J, Dewet J, Foerster J, Framer E. eds. *SPM Handbook of Health Assessment Tools.* 3rd ed. Pittsburgh, PA: The Society of Prospective Medicine and Institute for Health and Productivity Management; 1999:135–141.
3. Lasco R, Moriarty D, Nelson CF. *CDC Health Risk Appraisal User Manual.* Atlanta, GA: Centers for Disease Control, US Government Printing Office; 1984.
4. Terry P, Anderson DR, Serxner S. Health Assessment at Worksite. In: Hyner G, Peterson K, Travis J, Dewet J, Foerster J, Framer E, eds. *SPM Handbook of Health Assessment Tools.* 3rd ed. Pittsburgh, PA: The Society of Prospective Medicine and Institute for Health and Productivity Management; 1999:207–216.
5. Wetzler H. The status of Health Status Assessment. In: Hyner G, Peterson K, Travis J, Dewet J, Foerster J, Framer E, eds. *SPM Handbook of Health Assessment Tools.* 3rd ed. Pittsburgh, PA: The Society of Prospective Medicine and Institute for Health and Productivity Management; 1999:17–24.
6. Frankish CG, Herbert C, Milsum JH, Peters HF. Measurement of positive health and well-being. In: Hyner G, Peterson K, Travis J, Dewet J, Foerster J, Framer E, eds. *SPM Handbook of Health Assessment Tools.* 3rd ed. Pittsburgh, PA: The Society of Prospective Medicine and Institute for Health and Productivity Management; 1999:41–48.
7. Yen L, Edington DW, Witting P. Associations between health risk appraisal scores and employee medical claims costs in a manufacturing company. *Am J Health Promot.* 1991;61:46–54.
8. Edington DW, Yen L. Is it possible to simultaneously reduce risk factors and excess health care cost? *Am J Health Promot.* 1992;64:403–409.
9. Yen L, Edington DW, Witting P. Prediction of prospective medical claim costs and absenteeism costs for 1284 hourly workers from a manufacturing company. *J Occup Med.* 1992;344:428–435.
10. Yen L, Edington DW, Witting P. Corporate medical claim cost distributions and factors associated with high-cost status. *J Occup Med.* 1994;365:505–515.
11. Edington DW, Yen L, Witting P. The financial impact of changes in personal health practices. *J Occup Environ Med.* 1997;3911:1037–1046.
12. Edington DW. Emerging research: a view from one research center. *Am J Health Promot.* 2001;155:341–349.
13. Musich S, McDonald T, Hirchland D, Edington DW. Excess healthcare costs associated with health risks in diseased and non-disease health risk appraisal participants. *Dis Manage Health Outcomes.* 2002;104:251–258.
14. Goetzel R, Anderson D, Whitmer W, et al. The relationship between behavioral health risks and health care expenditures: an analysis of the multi-employer HERO health risk and cost database. *J Occup Environ Med.* 1998;4010:843–857.
15. Anderson DR, Whitmer RW, Goetzel RZ, et al. The relationship between behavioral health risks and group-level health care expenditures. *Am J Health Promot.* 2000;151:45–52.
16. Fries J, Bloch D, Harrington H, Richardson N, Beck R. Two-year results of a randomized controlled trail of a health promotion program in a retiree population: the Bank of America Study. *Am J Med.* 1993;94:455–462.
17. Golaszewski T, Lynch W, Clearie A, Vickey D. The relationship between retrospective health insurance claims and a health risk appraisal-generated measure of health status. *J Occup Med.* 1989;313:262–264.
18. Pelletier K. A review and analysis of the health and cost-effective outcome studies of comprehensive health promotion and disease prevention programs at the worksite: 1991–1993 update. *Am J Health Promot.* 1993;81:50–62.
19. Pelletier K. Clinical and cost outcomes of multifactorial, cardiovascular risk management interventions in worksite: a comprehensive review and analysis. *J Occup Environ Med.* 1997;3912:1154–1167.
20. Hutchins EB, Amler RW, Evans RC,

Jacobs TA, Sceigaj M, Sligh D, eds. *Healthier People: Version 4.0.* Atlanta, GA: The Carter Center of Emory University, Health Risk Appraisal Program; 1991.

21. Foxman B, Edington DW. The accuracy of health risk appraisal in predicting mortality. *Am J Publ Health.* 1987;778:971–974.
22. Gazmararian JA, Foxman B, Yen L, Morgenstern H, Edington DW. Comparing the predictive accuracy of health risk appraisal: the Center for Diseases Control versus Carter Center program. *Am J Publ Hlth.* 1991;8110:1296–1301.
23. US Public Health Services. *Health People 2000: National Health Promotion and Disease Prevention Objectives.* PHS 90–50212. Washington, DC: US Department of Health and Human Services; 1991.
24. US Department of Health and Human Services, Office of Public Health and Science, Office of Disease Prevention and Health Promotion. *Guide to Clinical Preventive Services.* 2nd ed. Washington, DC: Government Printing Office; 1996.
25. Yen L, Edington MP, Lu C, Edington DW. Early survey evaluation of the United Auto Workers and General Motors LifeSteps Health Promotion Program. *AWHP's Worksite Health.* 1999; 9–14.
26. Yen L, Edington MP, McDonald T, Hirschland D, Edington DW. Changes in health risks among the participants in the UAW-GM LifeSteps Health Promotion Program. *Am J Health Promot.* 2001;161: 7–15.
27. Metropolitan Life and Affiliated Companies. Medical care costs update. *Statistical Bull.* 1999;80:2–8.
28. Conover WJ. *Practical Nonparametric Statistics.* 2nd ed. New York: Jon Wiley & Sons, Inc.; 1980.
29. Stevens J. *Applied Multivariate Statistics for the Social Science.* 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 1997.
30. Pedhazur EJ. *Multiple Regression in Behavioral Research: Explanation and Prediction.* 3rd ed. Fort Worth, TX: Harcourt Brace College Publishers; 1997.
31. Qin J, Zhang B. A goodness-of-fit test for logistic regression models based on case-control data. *Biometrika.* 1997;843:609–618.
32. Ware JH. The fitting of straight lines when both variables are subject to error and the ranks of the means are known. *J Am Statistical Assoc.* 1972;67:891–897.
33. SAS Institute Inc. *Stock Market Analysis–Using SAS System: Technical Analysis.* 1st ed. Cary, NC: SAS Institute Inc.; 1995.
34. Achelis, SB. *Technical Analysis from A to Z.* 2nd ed. New York: McGraw-Hill Trade; 2000.
35. Smith SW. Moving average filters. In *The Scientist and Engineer's Guide to Digital Signal Processing.* 2nd ed. San Diego, CA: California Technical Publishing; 1999.
36. Branas K, de Gooijer J G, Terasvirta T. Testing linearity against nonlinear moving average models. *Communi Stat, Theory and Methods.* 1998;278:2025–2035.
37. Bryk AS, Raudenbush SW. *Hierarchical Linear Models: Applications and Data Analysis Methods.* 1st ed. Newbury Park, CA: Sage; 1992.
38. Littell RC, Milliken GA, Stroup WW, Wolfinger RD. *SAS System for Mixed Models.* 1st ed. Cary, NC: SAS Institute Inc.; 1996.
39. Fox J. *Linear Statistical Models and Related Methods.* 1st ed. New York: Jon Wiley & Sons, Inc.; 1984.
40. SAS Institute Inc. *SAS Procedures Guide.* 3rd ed. Cary, NC: SAS Institute Inc.; 1990.
41. SAS Institute Inc. *SAS/STAT User's Guide.* 4th ed. Cary, NC: SAS Institute Inc.; 1989.
42. Duan N, Manning WG, Morris CN, Newhouse JP. A comparison of alternative models for the demand for medical care. *J Business Econ Stat.* 1983;1:115–126.
43. Manning WG, Newhouse JP, Duan N, et al. Health insurance the demand for the demand for medical care: evidence from a randomized experiment. *Am Econ Rev.* 1987;77:251–277.
44. van Vilet RCJA. Predictability of individual health care expenditures. *J Risk Insur.* 1992;59:443–465.