PHYSICAL ACTIVITY AND LABORATORY ASSESSMENT

Steven J. Rigatti, MD
Medical Director
steven.rigatti@crlcorp.com

Michael Fulks, MD
Medical Director
fulksmd@volcano.net

Robert L. Stout, PhD
Chief Science Officer
stoutr@crlcorp.com

Clinical Reference Laboratory
Lenexa, KS

Executive Summary *Activity trackers and smartphones have made it increasingly appealing to consider activity levels in life underwriting. In order to understand the mortality risks associated with activity (or rather inactivity), it is necessary to evaluate longitudinal data which includes activity assessment, as well as the other variables which insurers commonly use to make risk decisions. This article presents an evaluation of NHANES data, which considers activity derived from surveys and from pedometers. Even this fairly rough approximation of activity level shows relatively good mortality risk selection in this data set. Difficulties incorporating activity assessment into an underwriting requirement scheme remain.*

In the life insurance industry there is considerable interest in the use of physical activity as a tool for the assessment of life risk. This interest has been spurred by the recent explosion in the use of wearable activity tracking devices and smartphone apps. It is well established that physical activity has a myriad of benefits, including reduced incidence of hypertension, diabetes, obesity, heart disease, stroke, peripheral vascular disease and depression. It has even been shown to be associated with a lower risk of death in peer-reviewed scientific publications, going all the way back to the 1950s.[1]

Though the benefits of activity are clear, it remains an open question whether physical activity assessment could be useful as a risk selection tool in the life insurance industry. Because exercise may benefit many laboratory and biometric parameters such as cholesterol, hemoglobin A1c and body mass index (BMI), it is reasonable to compare the mortality impact of activity to that of laboratory and biometric measurements.

In order to make this comparison, it is necessary to have, in a single data set, a group of individuals who have been followed for several years after undergoing both an assessment of their activity level and a set of laboratory and physical measurements. The National Health and Nutrition Examination Survey (NHANES) is just such a data set. Individuals selected for participation in the NHANES studies are interviewed, examined, and have blood and urine tests. The interview contains questions about demographics, income, education, medical and family history, as well as detailed questions about work and leisure time physical activity. In 2003 and 2005, the NHANES subjects were also given pedometers to wear for 1 week during waking (non-bathing) hours. NHANES data and the accompanying mortality follow-up file are freely available for download.

In order to make this study more relevant to life insurance, study participants were excluded if they admitted to medical conditions or had laboratory results that would ordinarily exclude them from life insurance. Participants were also excluded if they had very low levels of income or were outside the age window of 20 to 80 years. For each participant, a risk score (CRL SmartScore) was generated from laboratory tests and biometric measurements (BMI and blood pressure)

AURA

YOUR BRIDGE TO THE FUTURE

TRUSTED TECHNOLOGY, PROVEN RESULTS.

AURA is a proven decision management platform powered by RGA's global underwriting experience. In the digital age, the underwriting process is rapidly transitioning to data-driven evidence sources. With increased demands for simplicity and speed, we amped up the power of AURA with new capabilities while maintaining accurate risk assessment. AURA is a modern SaaS-based technology with the advanced business intelligence you'll need for tomorrow's insurance market.

AURA not only adapts to the future, it's your bridge to it.

Benefits of AURA

- Interviews — mobile device capable — online and offline
- SaaS ensures software remains current
- User-friendly configuration
- Supports all distribution channels, customized user journeys
- 10+ languages / 20 countries
- Multivariate decisioning

YOU NEED AURA TODAY.

E-mail **askaura@rgare.com** for a demonstration.

collected at the time of examination. CRL SmartScore is a multivariate, mortality-based, risk model which produces a score roughly equivalent to underwriting debits or credits. After all exclusions, the final data set had 24,536 participants, 1,963 of whom had died after an average of 10.4 years follow-up.

While there are many different ways to quantify activity using pedometers, for this study, average counts per minute (CPM) were used. A pedometer count is not the same thing as a "step" measured by modern activity trackers. A count includes a measure of intensity, thus ensuring a step while sprinting is counted as a higher level of activity than a step while walking. Because activity declines with age, the average CPM was corrected for age. Since the pedometer was only used in two of the NHANES cohorts, the data is somewhat limited (5,786 participants, 429 deaths, 10.6 years average follow-up). Therefore, activity was also quantified on the basis of the questionnaires, which have been given to all participants since 1999. The questionnaires ask about the frequency and intensity of work- and leisure-related activities, and can be quantified using METs (metabolic equivalents), combined with average hours per day, yielding MET-hours per day (MHD).

Setting up the comparison

In order to make a comparison between methods of mortality selection, a few different underwriting classification schemes were developed. First a typical knockout system was created, using build, blood pressure, total cholesterol and cholesterol:HDL ratio (see Table 1). Thresholds were chosen to resemble common knockout systems used in the industry. Anyone failing to meet the full set of class thresholds would be limited to the next best class in the usual manner. This knockout (KO) system was used to set the class distribution for the other systems (34% best class, 15% second best, 21% third best, 31% residual standard). Note the residual standard class would also include substandard and declined risks.

A SmartScore class was then generated. The 34% of people with the best (lowest) SmartScores received best class, and so on, based on the class distribution of the knockout system. Then, an activity-based class was created using MHD. This entire procedure was repeated using the smaller pedometer data set, substituting counts per minute for MHD.

Finally, for both the smaller and larger data sets, an underwriting class system was created using SmartScore after adjustment for activity-based mortality. The numerical adjustment was based on previously conducted mortality research which has been submitted to the *Journal of Insurance Medicine* for publication. Essentially, hazard ratios from Cox models were translated into debits and added or subtracted from the SmartScore.

Naturally, these various underwriting systems did not choose equivalent age and sex distributions. Therefore, each underwriting class was compared to the expected US population mortality matched by age, sex and calendar year. The final statistic for evaluation is the mortality ratio for each class in each system. Better risk selection is implied by those systems which have escalating MRs as one goes from best class to residual standard, those which have the lowest MR in best class, and the highest in standard class.

As shown in Figure 1 (next page), for the larger data set using questionnaires to assess activity, a knockout system performs similarly to an activity-based system. This is fairly impressive when one considers activity is only one variable doing the work of the three variables included in the KO system. Perhaps this is not unexpected when one considers that physical activity has been demonstrated to improve BMI, blood pressure and cholesterol levels. The SmartScore-based system outperforms the others, likely because it is based on mortality risk directly, rather than on intermediate outcomes, and considers many more variables.

Table 1

KO Class	BMI	Blood Pressure	Cholesterol:HDL Ratio	Total Cholesterol
Best Class	≤29 (age <60) ≤33 (age 60+)	≤130/80 (age <50) ≤135/85 (age 50+)	≤4.5	<220
Second Best	≤29 (age <60) ≤33 (age 60+)	≤135/85 (age <50) ≤140/90 (age 50+)	≤5.5	<250
Third Best	≤33 (age <60) ≤37 (age 60+)	≤140/90 (age <50) ≤145/90 (age 50+)	≤6.5	<270
Standard	>33 (age <60) >37 (age 60+)	>140/90(age <50) >145/90(age 50+)	>6.5	≥270

Figure 1

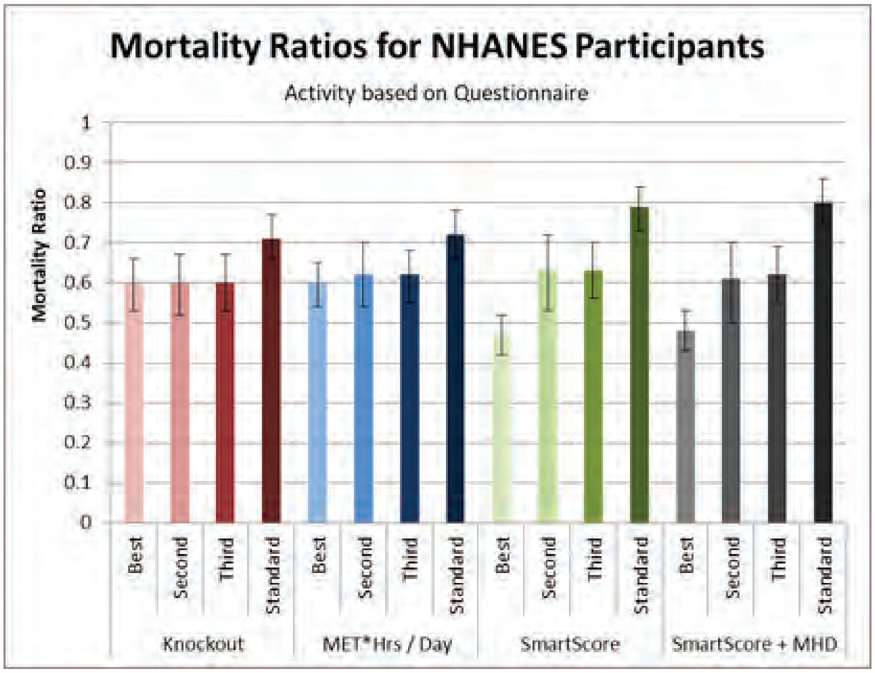
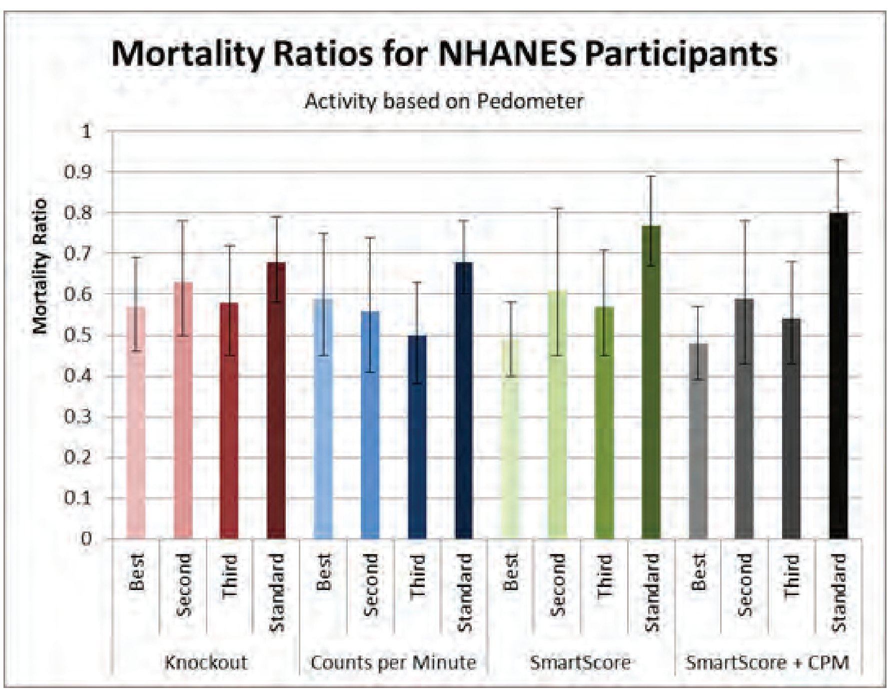


Figure 2 displays the results for the smaller data set where the pedometer is used to assess activity. As can be seen, a similar trend holds. However, the error bars are larger due to smaller number of deaths; therefore, conclusions are more tentative. In both data sets, adding an activity adjustment to the SmartScore did slightly improve classification – with the MR for the best class moving lower and the MR for the residual standard class moving higher vs. the unadjusted SmartScore.

Figure 2



Discussion

This research demonstrates that, within the confines of the NHANES data, a knockout system and a system based purely on physical activity, whether assessed by questionnaire or pedometer, perform similarly. Also, a system based on a multivariate mortality risk score based on labs and biometrics performs better than either of these systems, and is improved further by the incorporation of an activity assessment.

Challenges to the incorporation of physical activity assessment remain. In the NHANES data, participants were asked detailed questions about both leisure time and work-related activity. In that context, participants are likely to be fairly honest, as they have nothing to gain by embellishment. In the context of a life insurance transaction, prospective customers would likely be aware a question about their exercise habits may translate into better or worse underwriting classifications and embellishment becomes more likely. Collection of data from wearable devices may still be falsified, but it likely improves the objectivity of activity assessment. The variety of devices and apps introduces a high degree of variability in the types of data collected. Step count monitors may miss activities like weight training, cycling or swimming, while GPS-enabled devices capture this information but may create privacy concerns. Heart rate monitors introduce yet another potential level of assessment, offering access to tantalizing and important measurements, such as heart rate variability and response to exercise. However, some devices, especially those worn on the wrist and relying on photoplethysmography (LED light used to detect pulsations), may be inaccurate.[2]

When evaluating public data sets for important medical outcomes like mortality, it is helpful to validate research findings against those already present in the peer-reviewed medical literature. While non-industry studies would not consider SmartScore and would not necessarily exclude subjects in a manner similar to underwriting, it is still valuable to examine the work of others. In a 2015 study, Schmid et al.[3] evaluated the first NHANES cycle to include pedometer data. Follow-up was fairly short (35 months) and included only 112 deaths. Lower levels of moderate/vigorous activity were associated with a tripling of mortality risk. Higher levels of sedentary time produced a doubling of mortality risk. These associations were largely unaffected by exclusions for various medical issues. A 2012 study by Koster et al.[4] evaluated the NHANES data. Follow-up was short (2.8 years) and included only 145 deaths. Very high mortality ratios were seen for the highest quartiles of sedentary time, but there may have been an element of reverse causation (illness causing lack of activity, rather than the other way around). In 2016, Matthews et al.[5] published a similar study, which included 700 deaths after an average of 6.6 years of follow-up. In addition to confirming the risks of high levels of sedentary time and low levels of light and moderate/vigorous activity, this study demonstrated highly active adults (those who were active at any level for more than 5.8 hours per day) were protected from the mortality effects of high levels of sedentary time. In other words, if you exercise once you stand up, it is OK to sit. Finally and most recently, Lee et al.[6] studied data from 18,289 American women over the age of 65 (average age 72) who were given activity monitors to wear for 1 week. Follow-up averaged 4.3 years and included 504 deaths. The authors concluded as few as 4,400 steps per day was associated with lower mortality, and that mortality continued to improve with higher levels until approximately 7,500 steps per day.

Conclusion

This paper demonstrates that activity is an important component of mortality risk. The results should be used cautiously, if wearable-based physical activity is to be integrated into an underwriting scheme. While the direction of the mortality effect is likely to remain, the numerical relationship is likely to be different and difficult to quantify.

Notes

1. Morris JN, Heady JA, Raffle PA, Roberts CG, Parks JW. Coronary heart-disease and physical activity of work. *Lancet* 265:1111-1120.
2. Wang R, Blackburn G, Desai M et al. Accuracy of Wrist-Worn Heart Rate Monitors. *JAMA* Cardiol. 2017;2(1):104-106.
3. Schmid D, Ricci C, Leitzmann MF. Associations of Objectively Assessed Physical Activity and Sedentary Time with All-Cause Mortality in US Adults: The NHANES Study. *PLOS One*. 10(3):e0119591.
4. Koster A, Caserotti P, Patel KV et al. Association of Sedentary Time With Mortality Independent of Moderate to Vigorous Physical Activity. *PLOS One*. 7(6): e37696.
5. Matthews CE, Keadle SK, Troiano RP. Accelerometer-measured dose-response for physical activity, sedentary time, and mortality in US adults. *Am J Clin Nutr*. 2016;104:1424-32.
6. Lee IM, Shiroma EJ, Kamada M. Association of Step Volume and Intensity With All-Cause Mortality in Older Women. *JAMA*. Published online May 29, 2019.