Published in final edited form as:
Exp Aging Res. 2013 ; 39(5): 481–492. doi:10.1080/0361073X.2013.839021.

PHYSICAL PERFORMANCE AND SHORT-TERM MORTALITY IN VERY OLD MEXICAN AMERICANS

Lawrence J. Panas,
Preventive Medicine and Community Health, University of Texas Medical Branch, Galveston, Texas, USA

Carlos Siordia,
Center for Aging and Population Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Ronald J. Angel,
Department of Sociology, University of Texas at Austin, Austin, Texas, USA

Karl Eschbach, and
Internal Medicine- Geriatrics, University of Texas Medical Branch, Galveston, Texas, USA

Kyriakos S. Markides
Preventive Medicine and Community Health, University of Texas Medical Branch, Galveston, Texas, USA

Abstract

Background—Physical performance measures have been found to be strong predictors of adverse outcomes in aging populations. Few studies have examined the predictive ability of physical performance measures exclusively within populations of the very old. This study explores the predictive ability of the Short Physical Performance Battery (SPPB) and its three subcomponents - a timed walk, balance test, and timed repeated chair stands - on mortality in a sample of Mexican- Americans aged 75 and older.

Methods—Logistic regression analyses were used with data from the Hispanic Established Population for the Epidemiological Study of the Elderly (Hispanic EPESE), to investigate the relationship between timed walk, balance test, repeated timed chair stands, and the SPPB and mortality over a 2 ½ year period.

Results—We find that being unable to complete the timed walk, the balance test, repeated timed chair stands, or unable to complete any of the SPPB was significantly associated with mortality over 2 ½ years.

Conclusions—These findings indicate that physical performance measures may be less predictive of short term mortality in very old Mexican Americans than previously thought. More research is needed to understand this relationship.

INTRODUCTION

Measures of physical performance have been found to be strong predictors of adverse outcomes in aging populations. Such measures are less intrusive than large batteries of tests, especially when used in studies conducted outside of the clinical setting. Performance measures of lower body function in the elderly have been used in a variety of populations as

Address correspondence to Lawrence Panas, Preventive Medicine and Community Health, University of Texas Medical Branch, Galveston, Texas, 77555. Phone: (409) 772-2551, Fax: (409) 772-5272, lawrence.panas@utmb.edu.

predictors of mortality. These studies have been conducted with non-Hispanic Whites in the United States (Guralnik et al., 1994; Studenski et al., 2011), older Mexican Americans living in the Southwestern United States (Markides et al., 2001; Ostir, Kuo, Berges, Markides, & Ottenbacher, 2007) and with other 65 and over populations (Volpato et al., 2010; Volpato et al., 2008).

Using data from three of the Established Populations for the Epidemiologic Study of the Elderly (EPESE) sites, Guralnik and associates (1994) investigated the predictive ability of the Short Physical Performance Battery (SPPB), a test of lower body function, with respect to several important health outcomes. The battery includes a test of the subjects': (1) ability to rise from a chair five times while only using their legs; (2) timed short walk; and (3) balance. This three-part battery of physical performance strongly predicted short-term mortality, entering a nursing home, and disability. This study was limited to populations from Massachusetts, Iowa, and Connecticut, and did not include substantial Hispanic representation. Markides and colleagues (2001) showed that the SPPB was a strong predictor of short-term mortality in Mexican Americans aged 65 and over using data from the Hispanic EPESE. Several studies examining the individual components of the SPPB, including analyses of the Hispanic EPESE data, have found that the timed walk is as predictive of mortality as the full SPPB (Guralnik et al., 1994; Markides et al., 2001). The timed walk also has the advantage that it provides an additional advantage that it can be administered in most locations with minimal strain on the subject.

Most research on the impact of lower body function has focused on adults aged 65 years and older. Few studies focused on the predictive ability of these physical performance measures among the very old (i.e., those age 75 and above). Physical performance measures may be less predictive of mortality and physical function among the very old, who may exhibit less variation in physical functioning than their younger counterparts. Cesari and associates (2008) found that the full battery, compared to individual tests of walking speed, balance, and chair stands, was a stronger predictor of mortality in a very old sample recruited in Italy (aged 80 and older), though the chair stands were almost as good a predictor as the full performance battery.

In the current study, we investigate the ability of the SPPB, and each of the individual physical performance components: timed walk, balance test, and repeated time chair stands, to predict short-term mortality in Mexican Americans age 75 and over using data from the Hispanic EPESE. We examine the association of physical performance measures while considering several important covariates including age, gender, and chronic conditions. Based on the work by Cesari and colleagues (2008), we might expect that the full battery will be a strong predictor of mortality, but evidence from studies of the Hispanic EPESE suggests that walking speed, rather than chair stands, is the best individual measure for older Hispanics (Markides et al., 2001; Ostir, Markides, Black, & Goodwin, 1998; Ostir et al., 2007). Given the absence of much direct evidence, specifically pertaining to populations after age 75 and among Hispanics, we begin the study without strong and specific expectations. As detailed above, there is strong evidence that physical performance measures are very useful in research on mortality. Here we ask: Do all physical performance measures retain their ability to predict short-term mortality among very old Mexican Americans?

METHODS

Sample

The Hispanic EPESE is an ongoing cohort study of older Mexican Americans living in the Southwestern United States (Arizona, California, Colorado, New Mexico, and Texas) begun in 1993-1994, when a representative sample of 3,050 older Mexican Americans were interviewed in their homes. To study the predictive power of the performance battery and walking speed in a population of very old Mexican Americans, this analysis focuses on Wave 5 and Wave 6 of the Hispanic EPESE. Wave 5, baseline for this study, was conducted in 2004-2005 and Wave 6 was completed in 2007, approximately two and a half years later. At Wave 5, an additional representative sample of 902 Mexican American adults from the same region aged 75 years and over were added to supplement the surviving cohort of 1167 aged 75 years and over for a combined sample of 2069 Mexican Americans aged 75 years and older (Beard, Al Ghatrif, Samper-Ternent, Gerst, & Markides, 2009; Markides, Ray, Angel, & Espino, 2009).

Our analysis excludes respondents whose vital statuses during Wave 6 were unknown (n=259), respondents interviewed via a proxy at baseline (n=280), those with missing information on chronic conditions, and the performance battery (n=164). This left a final analytic sample size of 1,366 respondents.

Measures

Lower body performance was measured with a timed 8-foot walk, timed repeated chair stands, and standing balance. Walking speed was measured by timing the respondent's best of two attempts to walk 2.8 meters as quickly as they were able to from a standing start. Those in the highest quartile with speeds of 5.6 or more seconds were assigned a score of 1; those in the second quartile with speeds of 4.1-5.5 seconds were assigned a score of 2; those in the third quartile with speed of 3.2-4.0 seconds were assigned a score of 3; and those with speeds of 3.1 seconds of less were assigned a score of 4. Those unable to complete the timed walk were assigned a score of zero.

The chair stands were attempted from a sitting position with arms crossed. An initial test was conducted to determine if the respondent could successfully complete the task. If the first chair stand was completed, respondents were asked to attempt the same task five more times. The chair stands were also measured in quartiles based on seconds to complete: more than 16.7 seconds was scored 1; 13.7-16.6 seconds was scored 2; 11.2-13.6 was scored 3; and 11.1 or less was scored 4. Again, those unable to complete the task were assigned a score of 0.

The item for balance included four balance types: full tandem (score of 4), tandem (score of 3), semi-tandem (score of 2), and side-by-side (score of 1). Respondents were asked to maintain each balance task for about 10 seconds. Inability to complete the balance tasks resulted in a score of 0.

Scores on the three tasks were combined to create an overall performance score (0-12). The total lower body function score is then categorized into five groups to mirror the categorization of other measures: 0=unable to complete any physical component; 1= combined score of 1-3; 2= combined score of 4-6; 3= combined score of 7-9; and 4= combined score of 10-12.

Our methodological approach of assigning zeros to those unable to walk, stand from a chair, or balance is of crucial importance. Previous research has highlighted the importance of including those unable to perform on these measures instead of treating inability to perform as a missing category (Eeles, White, & Bayer, 2009; Rockwood, Jones, Wang, Carver, & Mitnitski, 2007). We echo the actions of others (Guralnik et al., 1994) and argue that individuals who are unable to participate in performance measures should not be excluded from analysis. The fact that they are unable to perform the task is theoretically significant—as it may indicate their low functional ability.

Other measures included in our models are age (continuous), gender (male =1), and measures of chronic conditions including diabetes, heart attack, hypertension, hip fracture, cancer, and stroke. For assessment of chronic conditions, respondents were asked if they had ever been informed by a doctor that they had any of the conditions.

Analytic strategy

We conduct several nested logistic regression analyses with physical performance measures collected at Wave 5 predicting mortality at Wave 6, an average of 2 ½ years follow-up. We then performed more detailed analysis of the full physical performance battery, including analyzing interaction terms for health conditions and performing goodness-of-fit tests, Receiver Operating Curves (ROC), and sensitivity and specificity calculations to determine the model that provided the best and most parsimonious fit. We used SAS software, version 9.2 (SAS Institute, Inc., 2008) and Stata software, version 12.1 (StataCorp, 2011) to conduct our analyses.

RESULTS

Table 1 shows that our analytic sample at baseline (2004-05) had a mean age of 81.4 (+ 4.7), 37.7% were male, 17.5% were unable to complete the full physical performance battery, 18.7% were unable to complete the short walk, 35.1% were unable to complete the timed repeated chair stands, and 26.7% were unable to complete the balance test. Slightly over one-third of the subjects reported having diabetes, and 67.1% of the reported having hypertension.

Of the 1,366 respondents who were alive at baseline and not missing information for physical function or other relevant factors, 207 (15.2%) were determined to have died by the follow-up interview in 2007.

Table 2 shows the results of logistic regression models predicting mortality using the full total lower body function score. Model 2 shows that, controlling for age and gender, that subjects unable to complete the test had significantly higher odds of mortality (OR 4.5; 95% CI 2.9-6.8) than those in the highest performing category (scores 10-12). After controlling for demographics and health conditions (Model 3), only those unable to complete the full physical battery show a significant risk of mortality (OR 3.9 95% CI 2.5-6.0). We also tested for possible interaction effects in our model, but no interactions were significant and they did not alter our main effects significantly when included in the model, so they are not shown.

Table 3 shows the results of logistic regression models predicting mortality from the short walk measure. In Model 1 only those unable to complete the short walk were at higher risk (OR 3.8; 95% CI 2.2-6.4) of mortality compared to the best walk score (score of 4), when controlling for demographics. This association decreased only slightly and remained significant after including other relevant factors in Model 2 (OR 3.2; 95% CI 1.8-5.4).

Table 4 shows the results of logistic regression models predicting mortality with the balance test. Only those unable to complete the balance test were at higher risk (OR 3.0; 95% CI 2.0-4.4) of mortality compared to completion of the full Tandem, when controlling for age and gender. This association decreased somewhat but remained significant after including other relevant factors in Model 2 (OR 2.7; 95% CI 1.8-4.0). Table 5 shows results from the analysis of the Repeated Timed Chair Stands test. The analysis controlling for demographics (Model 1) showed that only respondents unable to complete the test were at higher risk (OR 2.7; 95% CI 1.8-4.0) of mortality compared to the best performers. This association decreased but remained significant after including other relevant factors in Model 2 (OR 2.4; 95% CI 1.6-3.6).

SENSITIVITY AND MODEL FIT

To further test our full performance battery model, we calculated sensitivity and fit statistics to determine how well the SPPB predicted mortality when considering health conditions. For a model with demographics only (not shown) sensitivity was 53.6% with a goodness-of-fit (GoF) of .49 and an AUC of .65. Table 2 shows that adding the SPPB (Model 1) increases sensitivity (61.8%) and improves the AUC (.72) but appears to overfit the model (GoF =.72). Adding in the health conditions in Model 2 reduces GoF (.29), improves the AUC (.74) and increases sensitivity (62.8%). Adding in interaction terms (not shown) increased sensitivity (63.8), but reduced GoF (.24) and did not change the AUC (.74). These analyses indicate that the best, and most parsimonious, model for predicting mortality in the current analytic sample is the full performance battery controlling for demographics and select chronic conditions.

DISCUSSION

The current analysis found a smaller and less consistent influence of the performance measures on short term mortality than previous analyses. We find that only those who were unable to complete the full performance battery, or each of the individual component tasks, were more likely to have died within a few years after their measurement was taken. The current investigation supports the previous conclusions drawn by Rockwood and colleagues (2007) and provides incentive for further examination of the health of those unable to perform on these tasks.

As with most investigations, there are limitations in our study. Methodologically, we do not have a contrast group. It could be that performance measures work with other minority (e.g., non-Latino-black) or non-minority (i.e., non-Latino-white) populations. Theoretically, it could be that the ability to predict mortality with performance measures is a limitation created by the “width of time” in our short-term mortality outcome. Perhaps performance measures are most predictive with long-term mortality in the very old. Future research should explore these issues.

Notwithstanding these limitations, we feel this project provides substantive insight to the relationship between performance measures and short term mortality. Mortality may become less predictable in older ages due to a selection effect in socioeconomically challenged Mexican Americans who survive into their very old ages. It is also possible that the measurement of total body lower functioning in those aged 75 and above may not be as reliable as with younger survey participants. For example, it could be that “measurement error” arises if interviews experience greater difficulty collecting performance measures in the very old. There may be factors influencing these older respondents’ desire to complete the tasks required.

REFERENCES

Arias E. United States life tables by Hispanic origin. Vital and health statistics. Series 2, Data evaluation and methods research. 2010; (152):1.

Beard HA, Al Ghatrif M, Samper-Ternent R, Gerst K, Markides KS. Trends in Diabetes Prevalence and Diabetes-Related Complications in Older Mexican Americans From 1993–1994 to 2004–2005. Diabetes Care. 2009; 32(12):2212–2217. doi:10.2337/dc09-0938. [PubMed: 19755626]

Cesari M, Onder G, Zamboni V, Manini T, Shorr R, Russo A, Bernabei R, et al. Physical function and self-rated health status as predictors of mortality: results from longitudinal analysis in the ilSIRENTE study. BMC Geriatrics. 2008; 8(1):34. doi:10.1186/1471-2318-8-34. [PubMed: 19102751]

Eeles, EMP.; White, S.; Bayer, T. Inability to complete cognitive assessment is a marker of adverse risk. 2009. Age and ageing. Retrieved from http://ageing.oxfordjournals.org/content/early/2009/07/23/ageing.afp115.full

Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, Scherr PA, et al. A Short Physical Performance Battery Assessing Lower Extremity Function: Association With Self-Reported Disability and Prediction of Mortality and Nursing Home Admission. Journal of Gerontology. 1994; 49(2):M85–M94. doi:10.1093/geronj/49.2.M85. [PubMed: 8126356]

Markides KS, Black SA, Ostir GV, Angel RJ, Guralnik JM, Lichtenstein M. Lower Body Function and Mortality in Mexican American Elderly People. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2001; 56(4):M243–M247. doi:10.1093/gerona/56.4.M243.

Markides, KS.; Ray, LA.; Angel, R.; Espino, DV. Hispanic Established Populations for the Epidemiologic Study of the Elderly (HEPESE) Wave 5, 2004-2005 [Arizona, California, Colorado, New Mexico, and Texas].. Inter-university Consortium for Political and Social Research (ICPSR) [distributor]. Sep 23. 2009 Retrieved from http://dx.doi.org/10.3886/ICPSR25041

Ostir GV, Markides KS, Black SA, Goodwin JS. Lower Body Functioning as a Predictor of Subsequent Disability Among Older Mexican Americans. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 1998; 53A:M491–M495. doi:10.1093/gerona/53A.6.M491.

Ostir, Glenn V.; Kuo, Y-F.; Berges, IM.; Markides, KS.; Ottenbacher, KJ. Measures of Lower Body Function and Risk of Mortality over 7 Years of Follow-up. American Journal of Epidemiology. 2007; 166(5):599–605. doi:10.1093/aje/kwm121. [PubMed: 17566063]

Rockwood K, Jones D, Wang Y, Carver D, Mitnitski A. Failure to complete performance-based measures is associated with poor health status and an increased risk of death. Age and ageing. 2007; 36(2):225–228. Retrieved from http://ageing.oxfordjournals.org/content/36/2/225.short. [PubMed: 17259639]

SAS Institute, Inc.. SAS software, Version 9.2. SAS Institute, Inc.; Cary, NC, USA: 2008.

StataCorp. Stata Statistical Software: Release 12. StataCorp LP; College Station, TX: 2011.

Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, Brach J, et al. Gait Speed and Survival in Older Adults. JAMA: The Journal of the American Medical Association. 2011; 305(1):50–58. doi:10.1001/jama.2010.1923. [PubMed: 21205966]

Volpato S, Cavalieri M, Sioulis F, Guerra G, Maraldi C, Zuliani G, Fellin R, et al. Predictive Value of the Short Physical Performance Battery Following Hospitalization in Older Patients. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2010; 66A(1):89–96. doi:10.1093/gerona/glq167.

Volpato, Stefano; Cavalieri, M.; Guerra, G.; Sioulis, F.; Ranzini, M.; Maraldi, C.; Fellin, R., et al. Performance-Based Functional Assessment in Older Hospitalized Patients: Feasibility and Clinical Correlates. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2008; 63(12):1393–1398. Retrieved from http://biomedgerontology.oxfordjournals.org/content/63/12/1393.abstract.

Table 1

Descriptive statistics of baseline sample (2004-2005)

	n	% or m(sd)
Age	1,366	81.4 (±4.7)
Gender		
Men	515	37.7
Women	851	62.3
SPPB		
0 (unable)	239	17.5
1-3	129	9.4
4-6	248	18.2
7-9	447	32.7
10-12 (Best)	303	22.2
Timed Walk		
0 (Unable)	256	18.7
1	146	10.7
2	323	23.7
3	439	32.1
4 (Best)	202	14.8
Timed Chair Stands		
Unable (0)	364	26.7
Poor	238	17.4
Moderate	208	15.2
Good	210	15.4
Best (4)	346	25.3
Timed Balance		
Unable (0)	480	35.1
Side-by-Side	104	7.6
Semi-Tandem	133	9.7
Tandem	188	13.8
Full Tandem (Best)	461	33.8
Medical Conditions		
Hypertension	917	67.1
Heart Attack	221	16.2
Cancer	93	6.8
Diabetes	462	33.8
Hip Fracture	80	5.9
Stroke	168	12.3
Deceased at Time 6	207	15.2

SPPB – Short Physical Performance Battery

Table 2

Logistic Regression Models Predicting Mortality at Time 6 (2007) from SPPB, Age, Gender, Chronic Conditions, and Categorical Interactions at Baseline (2004-2005)

	(n=1,366)	
	Model 1	**Model 2**
SPPB (vs 10-12)		
7-9	0.78 (0.47-1.29)	0.75 (0.45-1.25)
4-6	1.12 (0.65-1.92)	1.06 (0.61-1.83)
1-3	1.78 (0.99-3.19)	1.64(0.90-2.98)
Unable	4.27 (2.66-6.85)	3.69 (2.26-6.02)
Demographics		
Age	1.09 (1.05-1.12)	1.09 (1.06-1.13)
Gender (male)	1.65 (1.19-2.28)	1.64 (1.18-2.28)
Chronic Conditions		
Hypertension		1.15 (0.81-1.63)
Heart Attack		1.63 (1.10-2.42)
Cancer		0.93 (0.50-1.72)
Diabetes		1.49 (1.06-2.08)
Hip Fracture		0.84 (0.44-1.61)
Stroke		1.06 (0.67-1.67)
Hosmer-Lemeshow	.72	.29
Sensitivity	61.8%	62.8%
Specificity	72.7%	71.1%
AUC	.72	.74

Model 1 includes age, gender, and the full physical performance battery (10-12 is the best score range).

Model 2 adds hypertension, heart attack, cancer, diabetes, hip fracture and stroke.

SPPB = Short Physical Performance Battery

Table 3

Logistic Regression Models Predicting Mortality at Time 6 (2007) from Timed Walk Test, Age, Gender, and Chronic Conditions at Baseline (2004-2005)

	(n=1,366)	
	Model 1	**Model 1**
Timed Walk (vs 4)		
3	0.75 (0.43-1.30)	0.70 (0.40-1.22)
2	0.97 (0.55-1.7 1)	0.89 (0.50-1.58)
1	1.46 (0.78-2.74)	1.35 (0.71-2.56)
Unable	3.76 (2.22-6.38)	3.15 (1.83-5.43)
Demographics		
Age	1.09 (1.06-1.12)	1.10 (1.06-1.13)
Gender (male)	1.65 (1.19-2.28)	1.65 (1.19-2.30)
Chronic Conditions		
Hypertension		1.15 (0.82-1.63)
Heart Attack		1.65 (1.12-2.45)
Cancer		0.94 (0.51-1.74)
Diabetes		1.56 (1.12-2.18)
Hip Fracture		0.88 (0.46-1.69)
Stroke		1.06 (0.67-1.68)

Model 1 includes age, gender, and the short walk measure (4 is the best score).

Model 2 adds hypertension, heart attack, cancer, diabetes, hip fracture and stroke.

4 = Top Quartile of performance

Table 4

Logistic Regression Models Predicting Mortality at Time 6 (2007) from Balance Test, Age, Gender, and Chronic Conditions at Baseline (2004-2005)

	(n=1,366)	
	Model 1	**Model 2**
Timed Balance (vs Full Tandem)		
Tandem	0.84 (0.43-1.54)	0.83 (0.44-1.58)
Semi-Tandem	1.37 (0.75-2.52)	1.33 (0.72-2.46)
Side-by-Side	1.57 (0.83-2.97)	1.48 (0.78-2.81)
Unable	2.97 (2.00-4.41)	2.68 (1.78-4.02)
Demographics		
Age	1.09 (1.05-1.12)	1.09 (1.06-1.13)
Gender (male)	1.50 (1.10-2.05)	1.51 (1.09-2.07)
Chronic Conditions		
Hypertension		1.18 (0.84-1.66)
Heart Attack		1.73 (1.17-2.54)
Cancer		0.92 (0.50-1.70)
Diabetes		1.51 (1.08-2.09)
Hip Fracture		0.90 (0.47-1.70)
Stroke		1.16 (0.74-1.81)

Model 1 includes age, gender, and the balance test (Full Tandem is the best category).

Model 2 adds hypertension, heart attack, cancer, diabetes, hip fracture and stroke.

Table 5

Logistic Regression Models Predicting Mortality at Time 6 (2007) from Repeated Timed Chair Stands, Age, Gender, and Chronic Conditions at Baseline (2004-2005)

	(n=1,366)	
	Model 1	**Model 2**
Timed Chair Stands (vs Best)		
Good	0.65 (0.36-1.17)	0.63 (0.35-1.15)
Moderate	0.74 (0.42-1.31)	0.72 (0.40-1.29)
Poor	0.68 (0.39-1.19)	0.65 (0.37-1.15)
Unable	2.68 (1.80-4.00)	2.38 (1.58-3.58)
Demographics		
Age	1.08 (1.05-1.12)	1.09 (1.06-1.12)
Gender (male)	1.56 (1.14-2.15)	1.57 (1.13-2.16)
Chronic Conditions		
Hypertension		1.16 (0.82-1.64)
Heart Attack		1.74 (1.18-2.58)
Cancer		0.91 (0.49-1.68)
Diabetes		1.52 (1.09-2.12)
Hip Fracture		0.87 (0.45-1.66)
Stroke		1.10 (0.70-1.73)

Model 1 includes age, gender, and the repeated timed chair stands.

Model 2 adds mobility disability, hypertension, heart attack, cancer, diabetes, hip fracture and stroke.