Concurrent Criterion Validity of the Safe Driving Behavior Measure: A Predictor of On-Road Driving Outcomes

Sherrilene Classen, Yanning Wang, Sandra M. Winter, Craig A. Velozo, Desiree N. Lanford, Michel Bédard

Sherrilene Classen, PhD, MPH, OTR/L, is Director, Institute for Mobility, Activity and Participation, and Associate Professor, Department of Occupational Therapy, College of Public Health and Health Professions, University of Florida, PO Box 100164, Gainesville, FL 32610; sclassen@phhp.ufl.edu

Yanning Wang, MS, is Doctoral Student in Epidemiology, College of Public Health and Health Professions, University of Florida, Gainesville.

Sandra M. Winter, PhD, OTR/L, is Research Coordinator, Institute for Mobility, Activity and Participation, and Postdoctoral Associate, Department of Occupational Therapy, College of Public Health and Health Professions, University of Florida, Gainesville.

Craig A. Velozo, PhD, OTR/L, is Professor, Department of Occupational Therapy, College of Public Health and Health Professions, University of Florida, Gainesville, and Research Health Scientist, North Florida/South Georgia Veterans Health System, Gainesville.

Desiree N. Lanford, MOT, OTR/L, CDRS, is Certified Driving Rehabilitation Specialist, Institute for Mobility, Activity and Participation and Department of Occupational Therapy, University of Florida, Gainesville.

Michel Bédard, PhD, is Professor, Canada Research Chair in Aging and Health, and Director, Center for Research on Safe Driving, Lakehead University, Thunder Bay, Ontario, Canada.

We determined the concurrent criterion validity of the Safe Driving Behavior Measure (SDBM) for on-road outcomes (passing or failing the on-road test as determined by a certified driving rehabilitation specialist) among older drivers and their family members–caregivers. On the basis of ratings from 168 older drivers and 168 family members–caregivers, we calculated receiver operating characteristic curves. The drivers' area under the curve (AUC) was .620 (95% confidence interval [CI] = .514–.725, p = .043). The family members–caregivers' AUC was .726 (95% CI = .622–.829, $p \leq$ .01). Older drivers' ratings showed statistically significant yet poor concurrent criterion validity, but family members–caregivers' ratings showed good concurrent criterion validity for the criterion *on-road driving test*. Continuing research with a more representative sample is being pursued to confirm the SDBM's concurrent criterion validity. This screening tool may be useful for generalist practitioners to use in making decisions regarding driving.

The U.S. population age 65 or older is projected to more than double in the next 30 yr, from 40.2 million in 2010 to 88.5 million by 2050. Although previous researchers have predicted increased crash rates as a result of the rising demographic (Bédard, Stones, Guyatt, & Hirdes, 2001), Cheung and McCartt (2011) reported that fatal crash rates have declined for older drivers in the past decade.

Crashes are a key safety measure for gauging the impact of interventions and policies (National Center for Injury Prevention and Control, 2009), but driving performance as tested via on-road studies is considered the industry standard (Di Stefano & Macdonald, 2005). On-road testing has limitations, however: It is expensive and risky, can be executed validly and reliably only by trained professionals with specialty certifications, and provides limited access to most older drivers, and the process may end in a driver being reported to the licensing authorities if he or she does not do well. To enable older drivers to assess their driving behaviors, researchers and advocacy organizations have developed self-reports and screening tools (AAA Foundation for Traffic Safety, 2010; AARP & Andrus Foundation, 1996; Eby, Molnar, Shope, Vivoda, & Fordyce, 2003; Staplin & Dinh-Zarr, 2006).

Self-reports are criticized for the bias that they may introduce. For example, self-selection bias, recall bias, and rater bias are some of the most common sources of error associated with self-report. If self-reports are to be useful, establishing concurrent validity, predictive validity, or both between the self-report or screening tool and the criterion measure (e.g., on-road performance or crash outcomes) becomes imperative. Self-report or screening tools with criterion validity for driving performance (passing or failing an on-road test) are limited in

the driving literature. In addition to self-reports, proxy or caregiver reports may serve as a useful source of information on driving behaviors among older adults. In response to the limitations of existing tools, we have developed the Safe Driving Behavior Measure (SDBM), and we are testing how this tool may be predictive of on-road outcomes when used by older drivers and their caregivers.

Several driving studies have sought caregiver opinions. For example, Wild and Cotrell (2003) found that caregivers had insight into the driving errors (e.g., managing intersections, managing lane changes) of care recipients with Alzheimer's disease who still drove. However, compared with results on a standardized road test, they underreported some of the care recipients' driving errors. Croston, Meuser, Berg-Weger, Grant, and Carr (2009) reported that family members could provide adequate information on some driving behaviors (e.g., monitoring traffic, maintaining speed) of drivers with dementia (Alzheimer's type). In our previous work, we found that family members and caregivers were more reliable than healthy community-dwelling licensed drivers to report on driving behaviors (e.g., coming to a dead stop or maintaining lane while driving), but they were not as accurate as driving evaluator reports, which were based on standardized on-road tests (Classen et al., 2012b).

Recognizing that caregivers make an important contribution to identifying driving errors or driving behaviors, we have used their input in determining the psychometrics of the SDBM. Family members and caregivers were involved in establishing face and content validity (Classen et al., 2010; Winter et al., 2011), and their ratings were used to determine construct validity (Classen et al., 2012a), rater reliability, and rater effects (leniency vs. severity) among three rater groups (older drivers, family members and caregivers, driving evaluators; Classen et al., 2012b). Our preliminary data (from the studies cited earlier) point to the SDBM's potential usefulness as a screening measure for family members or caregivers to rate the driving behaviors of older drivers, but concurrent criterion validity has not yet been determined.

Measure of Validity Testing: Receiver Operating Characteristic Curves

Receiver operating characteristic (ROC) curves provide a methodology to determine the criterion validity of a screening tool as measured against a gold-standard outcome. Essentially, the ROC curve is a plot of the rate of true positives (true hits or sensitivity) against the rate of false positives (true misses or 1− specificity) resulting from the application of many arbitrarily chosen cutoff

points of the predictor test (Portney & Watkins, 2000). The ROC curve demonstrates the effectiveness of using different cutoff values and reveals the optimal cutoff value for the predictor test. If the area under the curve (AUC), an index of discriminability, is statistically significant and at least .70 in magnitude, then further attention must be paid to the other ROC attributes, such as sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV; Portney & Watkins, 2000).

Sensitivity is the predictor test's ability to obtain a positive test when the condition really exists (a true positive); here, it means that the predictor test would suggest that the participant will fail the on-road test, and the participant would actually fail it. *Specificity* is the predictor test's ability to obtain a negative result when the condition is really absent (a true negative); here, the predictor test would suggest that the participant will pass the on-road test, and the participant would actually pass it (Portney & Watkins, 2000). *PPV* is the probability that the participant will, given a certain cutpoint on the predictor test suggesting a failure on the on-road test, actually fail the on-road test. *NPV* is the probability that the participant will, given a cutpoint on the predictor test suggesting a pass on the on-road test, actually pass the on-road test. Note that the number of *false positives* (those who receive a failing score but pass the road test) and *false negatives* (those who receive a passing score but fail the road test) and, thus, the sensitivity and specificity values change with the cutoff value. Ultimately, one wants the false positives and false negatives to be as close to 0 as possible. For an example of ROC curves using error scores to determine passing or failing on an on-road test, see Shechtman, Classen, Awadzi, and Mann (2009); for using ROC to determine the sensitivity of predictor tests of on-road outcomes, see Classen et al. (2009).

Rationale and Significance

The on-road test is considered the industry gold standard, but because of its characteristics (expensive, time consuming, risky, not accessible to all), efficient screening tests that are predictive of actual on-road outcomes must be developed and tested. The SDBM holds promise for use as a screening tool for family members and caregivers and, potentially, for older drivers, but the criterion validity for the two groups has not yet been established. Thus, the purpose of this study was to determine the SDBM's concurrent criterion validity, as completed by older drivers and their family members and caregivers, against the on-road test conducted by trained driving evaluators.

Method

This study received institutional review board approval from the University of Florida and Lakehead University, and all participants provided informed consent.

Design

This prospective quasi-experimental study used a convenience sample of 168 older drivers and their 168 family members and caregivers from two sites to examine the concurrent criterion validity of the SDBM against the outcome (pass–fail) of standardized on-road tests.

We recruited older drivers and their family members and caregivers from north central Florida and Thunder Bay, Ontario, by flyer distribution in the local community facilities, local newspaper advertisements, and word-of-mouth referrals. Older drivers were included if they were 65–85 yr old, had a valid driver's license, were driving 3 mo before and at the time of recruitment, and had the cognitive and physical ability to complete the SDBM and participate in an on-road driving test. They were excluded if they had medical advice not to drive, had uncontrolled seizures in the past year, or used medications that cause central nervous system impairments. Family members and caregivers were included if they were able to report on the older adult's driving behaviors and excluded if they had physical or mental conditions that impaired their ability to participate.

Measures and Study Variables

Demographics and Health-Related Characteristics. For the drivers, we reported the following demographic variables: age, gender, race (White vs. other), education (high school graduation, some training after high school graduation, and college graduation), and living status (live with others vs. live alone). We also analyzed number of days driving per week and health-related characteristics, such as self-reported number of medications, self-reported health conditions, and comorbidities.

For the family members and caregivers, we reported age, gender, race, education, relationship with driver (family member vs. caregiver), days per week riding with the driver, and lifestyle impact (a self-reported appraisal of how much the caregiver's lifestyle would be affected if the driver stopped driving).

SDBM. The SDBM is available for drivers, family members and caregivers, and professionals (e.g., driving rehabilitation specialists, driving evaluators, and therapists). The driver SDBM, a 68-item questionnaire to determine the level of difficulty a driver experienced in the past 3 mo when executing driving behavior, has three

sections: Section A, Demographics (gender, race, education level, etc.); Section B, Driving History (days per week of driving, crashes or violation numbers, etc.); and Section C, Driving Behaviors. Difficulty with the driving task was rated on a 5-point adjectival scale ranging from 1 (*cannot do*) to 5 (*not difficult*; Classen et al., 2010, 2012a). The family member and caregiver's SDBM includes only Sections A and C. In this study, we used scores from Section C (interval data derived from Rasch analysis), not the total of the raw scores (ordinal data) as documented in detail in Classen et al. (2012a). We used the SDBM as the independent predictor of on-road outcomes.

Clinical Tests. The validated clinical test battery, with reported psychometrics, included tests of vision, visual cognition, and cognition and motor performance and has been fully documented in previous studies. For the purposes of this study, we include information only on the abilities described in the sections that follow (Stav, Justiss, McCarthy, Mann, & Lanford, 2008).

Vision. Visual acuity and contrast sensitivity were tested using the Optec® 2500 visual analyzer (Stereo Optical Company Inc., Chicago). We categorized the binocular (both eyes open) visual acuity as 20/20–20/40 and 20/50 or poorer (e.g., ≥20/70). We dichotomized contrast sensitivity as *intact* (all five Optec 2500 contrast sensitivity slides were intact) or *impaired* (any of the five contrast sensitivity slides were impaired).

Visual Cognition. We reported the Useful Field of View (UFOV) risk index (1 = *very low risk*, 2 = *low risk*, 3 = *low–moderate risk*, 4 = *moderate–high risk*, 5 = *high risk*) and three UFOV subsets (UFOV 1, visual search and visual processing; UFOV 2, divided attention; UFOV 3, selective attention; Ball & Owsley, 1993; Edwards et al., 2006). The cutpoint for each subtest is 500 ms, meaning that if a person exceeds this score per subtest, he or she will not be able to continue to the next sections and may have impaired visual processing speed.

Cognition. We used the Mini-Mental State Examination (MMSE; maximum score = 30; Folstein, Folstein, & McHugh, 1975) as an indicator of baseline cognitive functioning.

Motor Performance. We used the Rapid Pace Walk (RPW; in seconds) to test the motor performance (gait, postural control, balance, speed of walking) of older drivers. The RPW, when executed for longer than 7 s, is predictive of adverse driving events (accidents, violations, being stopped by the police; Marottoli, Cooney, Wagner, Doucette, & Tinetti, 1994), and this test is statistically significantly correlated with on-road driving performance (Stav et al., 2008).

On-Road Test. The Florida on-road test consisted of driving a standardized road course with demonstrated reliability (intraclass correlation coefficient = .94, $p < .05$) and validity (driving performance score was correlated with the global rating score; $r = .84$, $p < .001$) for older drivers (Justiss, Mann, Stav, & Velozo, 2006; Posse, McCarthy, & Mann, 2006). The Canadian site used a demerit point system consistent with the method used by its licensing authority. The outcome of the road tests included a pass–fail measure of driving: 3 = *pass*, 2 = *pass with restrictions or recommendations*, 1 = *fail with remediation*, 0 = *fail, not remediable*. Both the University of Florida, the primary site, and Lakehead University, the secondary site, used a dichotomized pass–fail outcome.

Procedure

All older drivers and their family members and caregivers gave written informed consent before the study. Older drivers completed the SDBM first and then a brief clinical test battery before completing an on-road test. All aspects of testing were performed by a certified driving rehabilitation specialist (CDRS) at the University of Florida site and by a trained driving evaluator at the Lakehead University site. The evaluators had 100% interrater reliability (Classen et al., 2010). The on-road driving test occurred on the same day, or close to the same day, as the SDBM and clinical test administration, except if rain or adverse weather events interfered with the on-road test; in this situation, the on-road driving test was rescheduled for a different day.

Family members and caregivers completed SDBM Section A (Demographics) to provide information on themselves and their relationship with the driver (e.g., how often they rode with the driver). They also completed Section C (68 items on driving behaviors), based on their observations over the past 3 mo.

Data Management

All the data (SDBM, demographic information, scores on the clinical tests, and on-road test results) of the older drivers and family members and caregivers were entered into the database by trained research assistants. This database was located on a central, secure, and password-protected data repository at the primary site. Data entry was monitored by the principal investigator, and quality control spot checks and corrections were made intermittently during data entry to ensure data completion and accuracy. Missing data were reported to the driving evaluators, obtained from participants by means of phone calls, or reported as missing when data were not available.

Data Analysis

We used PASW Statistics 18 (SPSS Inc., Chicago) and WINSTEPS 3.70.0 (www.winsteps.com/winsteps.htm) to perform the analyses.

Descriptive Statistics. For the drivers, we conducted a descriptive analysis and included demographic, driving history, health-related characteristic, clinical test, and on-road test data. The descriptive analysis of family members and caregivers included their demographics, their history as a passenger, and how their lifestyle would be affected if the driver reduced or stopped driving.

Bivariate Analysis. We conducted the χ^2 test to compare the difference between family members and caregivers for *lifestyle impact,* that is, to determine whether their lifestyle would be affected (yes–no) if the driver reduced or stopped driving (Fisher's exact test was used when the 2×2 contingency table contained cells with expected counts of <5). We considered $p \leq .05$ significant.

ROC Curve Analysis. We determined the concurrent criterion validity of the SDBM using the ROC curve. In this study, we viewed an AUC between .7 and .9 as having an acceptable magnitude (Streiner & Cairney, 2007). Most important, for the SDBM to be used as a potential screening tool to accurately classify drivers who fail the on-road test, we wanted sensitivity to be high (>.70). Generally, we wanted to minimize misclassification of drivers, or false positives and false negatives. We generated the ROC curve and AUC estimates with PASW Statistics 18 using measures derived from raw scores on the SDBM by means of Rasch analysis and presented as logits[1] (Bond & Fox, 2007; Classen et al., 2012a). Using the measure (logits), we present the ROC curves demonstrating five of these potential SDBM cutpoint measures. On the basis of the cutpoints, we also calculated the associated specificity, sensitivity, error, PPV, and NPV. The AUC of the ROC curve was based on a 95% confidence interval (CI) and $p \leq .05$ to indicate statistical significance.

Results

Descriptive Statistics

Driver. Table 1 presents the demographics and health-related and driving habits for 168 drivers. The drivers' mean age was 72.96 yr (standard deviation [*SD*] = 5.28, range = 65–85). Most of the drivers were White

---

[1]The procedure for and results of converting the SDBM raw scores to interval measures using Rasch analysis are available from Sherrilene Classen.

Table 1. Driver Characteristics (*N* = 168)

Characteristic	Statistic	
Age, *M* (*SD*)	72.96	(5.28)
Gender, *n* (%)		
Female	77	(45.8)
Male	91	(54.2)
Race, *n* (%)		
White	154	(91.7)
Other	12	(7.1)
Missing	2	(1.2)
Education, *n* (%)		
≤ High school	33	(19.6)
Training after high school	36	(21.4)
≥ College	99	(58.9)
Dichotomized live alone, *n* (%)		
Live alone	44	(26.2)
Live with others	124	(73.8)
Driving days/wk, *M* (*SD*)	5.71	(1.66)
No. of medications, *M* (*SD*)	7.01	(4.54)
Self-reported health condition, *n* (%)		
Having limitation on driving	8	(4.8)
No limitation on driving	160	(95.2)
Comorbidities, *n* (%)		
Heart disease	105	(62.5)
Arthritis	87	(51.8)
Other musculoskeletal	69	(41.1)
Cataracts	44	(26.2)
Urinary disease	38	(22.6)
Cancer	27	(16.1)
Stomach disorders	25	(14.9)
Other glandular problems	24	(14.3)
Sleep disorder	23	(13.7)
Respiratory disease	22	(13.1)
Parkinson's disease	20	(11.9)
Other neurological	19	(11.3)
Glaucoma[a]	16	(9.5)
Diabetes	14	(8.3)
Depression	14	(8.3)
Other vision problems	10	(6.0)
Psychiatric problems	6	(3.6)
Macular degeneration[b]	3	(1.8)
Stroke	2	(1.2)
Dementia	2	(1.2)
Diabetic retinopathy[c]	0	(0)
Visual acuity, *n* (%)		
Vision 20/20–20/40	143	(85.1)
Vision ≤20/50 or not tested	16	(9.5)
Missing	9	(5.4)
Contrast sensitivity, *n* (%)		
Impaired	49	(29.2)
Intact	99	(58.9)
Missing[d]	20	(11.9)
UFOV, ms, *M* (*SD*)		
UFOV1	27.20	(25.49)
UFOV2	118.03	(109.97)
UFOV3	273.27	(120.55)

(Continued)

Table 1. Driver Characteristics (*N* = 168) (*cont.*)

Characteristic	Statistic
Risk index, *M* (*SD*)	
Very low	84 (50.0)
Low	42 (25.0)
Low to moderate	22 (13.1)
Moderate to high	16 (9.5)
High to very high	4 (2.4)
MMSE total, *M* (*SD*)	27.96 (1.82)
Rapid Pace Walk, s, *M* (*SD*)[b]	5.72 (1.53)
On-road driving test, *n* (%)	
Failing	29 (17.3)
Passing	139 (82.7)

Note. *M* = mean; MMSE = Mini-Mental State Examination; *SD* = standard deviation; UFOV = Useful Field of View.
[a]*n* = 164. [b]*n* = 167. [c]*n* = 161. [d]The Canada site did not test the contrast sensitivity.

(91.7%), educated beyond high school (80.3%), and lived with others (73.8%). The self-reported average number of medications was 7.01 (*SD* = 4.54). Only 4.8% of the drivers reported having health conditions that limited their driving abilities. Although the secondary site did not collect data on contrast sensitivity, 33.1% of the drivers (*n* = 49) from the primary site (*N* = 148) had impaired contrast sensitivity, 9.5% (*n* = 16) had binocular visual acuity of ≤20/50 or could not be tested, and 11.9% (*n* = 20) had the UFOV risk index of moderate to high or high to very high. The mean score on the MMSE was 27.96 (range = 22–30; *SD* = 1.82), and the mean time for the RPW was 5.72 s (*SD* = 1.53).

Family Members and Caregivers. One hundred sixty-eight family members and caregivers completed the study. Table 2 shows that the majority of the family members and caregivers were female (72.0%), White (93.5%), and family members of the drivers (79.8%) and received further education after high school graduation (83.9%). They were ages 19–85 with a median age of 67.5 (25th percentile = 56.3, 75th percentile = 74.0) and were the driver's passenger an average of 2.77 days per wk (*SD* = 2.42). Family members were more likely to report that their lifestyle would be affected if the driver reduced or stopped driving than were caregivers (35.1% of the family members vs. 8.8% caregivers, *p* < .05; results are not shown in Table 2).

Receiver Operating Characteristic Curves

Driver. Figure 1 shows the ROC curve and the AUC based on drivers' responses. The AUC based on drivers' responses was .620, 95% CI = (.514, .725), *p* = .043. Five SDBM cutpoints and the associated specificity, sensitivity, error, PPV, and NPV are reported with the ROC curve. As an example, a cutoff point of 4 on the

Table 2. Family Members' and Caregivers' Characteristics (*N* = 168)

Characteristic	Statistic
Age,[a] *M* (*SD*)	63.52 (14.38)
Gender, *n* (%)	
Male	47 (28.0)
Female	121 (72.0)
Race, *n* (%)	
White	157 (93.5)
Other	11 (6.5)
Education, *n* (%)	
≤ High school	27 (16.1)
Training after high school	61 (36.3)
≥ College	80 (47.6)
Relationship with driver, *n* (%)	
Family member	134 (79.8)
Other	34 (20.2)
How many days a week do you ride with the driver, *M* (*SD*)	2.77 (2.42)
Lifestyle impact,[b] *n*[c] (%)	
No	110 (67.9)
Yes	52 (32.1)

Note. M = mean; *SD* = standard deviation.
[a]25th percentile = 56.3, 50th percentile = 67.5, 75th percentile = 74.0. [b]"If the driver reduced or stopped driving, would it significantly impact your current lifestyle?" [c]*n* = 162.

ROC curve, a value of 4.55 logits (converting raw scores to interval measures on the basis of Rasch analysis), yields sensitivity of .79, specificity of .46, error of .75, PPV of .24, and NPV of .91.

Family Members and Caregivers. Figure 2 shows the ROC curve and the AUC based on family members and caregivers' responses, AUC = .726, 95% CI = (0.622, 0.829), *p* ≤ .01. Five SDBM measures and the associated specificity, sensitivity, PPV, and NPV are reported with the ROC curve. The AUC of .726 is above the acceptable AUC level of .7. As an example, a cutoff point of 4 on the ROC curve (a value of 4.57 logits) yields an associated sensitivity .79, specificity of .59, error of .62, PPV of .29, and NPV of .93.

Discussion

We examined the concurrent criterion validity of the SDBM for on-road outcomes (passing or failing the on-road test as determined by a CDRS) among older drivers and their family members and caregivers in Gainesville, Florida, and Thunder Bay, Ontario.

A majority of our drivers were licensed community-dwelling White men and women of a high educational level who drove almost daily and had relatively few self-reported medications. Although the group reported a variety of comorbidities, only about 5% reported that these conditions affected their driving. Their clinical profiles showed that they had adequate visual, visuocognitive, cognitive, and motor performance skills; as a result, we

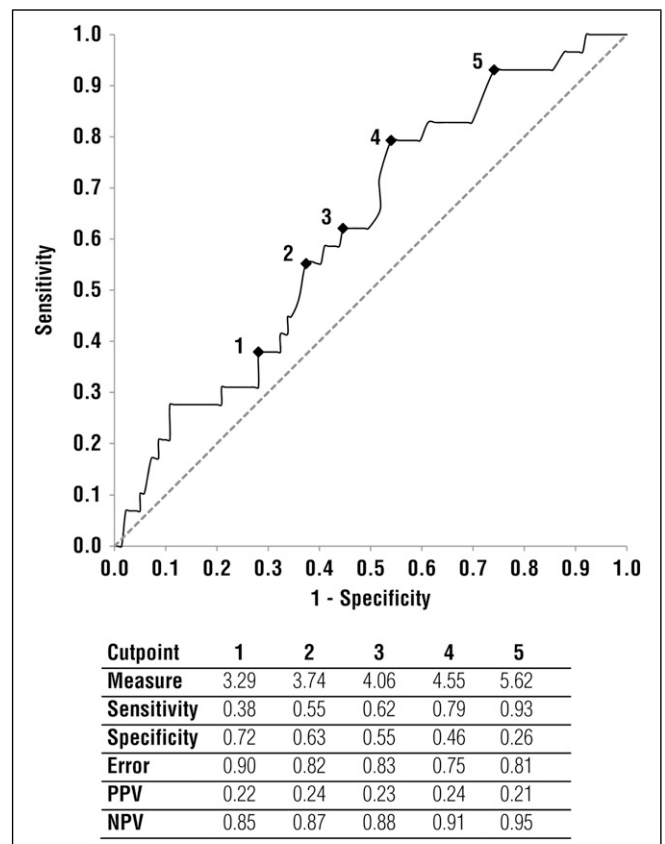| Cutpoint	1	2	3	4	5
Measure	3.29	3.74	4.06	4.55	5.62
Sensitivity	0.38	0.55	0.62	0.79	0.93
Specificity	0.72	0.63	0.55	0.46	0.26
Error	0.90	0.82	0.83	0.75	0.81
PPV	0.22	0.24	0.23	0.24	0.21
NPV	0.85	0.87	0.88	0.91	0.95

Figure 1. Receiver operating characteristic curve with cutpoints based on the drivers' ratings. Area under the curve = .620, 95% confidence interval = (.514, .725), *p* = .04.

Note. NPV = negative predictive value; PPV = positive predictive value.

surmise that they could be considered a relatively healthy group of older drivers. This group is not representative of the general spectrum of older adults, because our sample had low representation of minorities, people of low educational status, and those with poor health status. Generalizations can be made only to drivers who fit the profile described earlier.

Most of the family members and caregivers were community-dwelling White women with education beyond high school. About 80% of the group were family members of the drivers. Thirty percent of the group reported that they would be affected if the driver reduced or stopped his or her driving. In terms of the general U.S. demographics for caregivers of older adults, our group showed similarities in that they were also mainly female caregivers. However, a study of the U.S. general population found that most care recipients were women; in our study, most of the drivers (i.e., care recipients) were men (National Alliance for Caregiving & AARP, 2009). The same study found that 40% of those in the U.S. study lived alone, whereas only 26% of our group lived alone (National Alliance for Caregiving & AARP, 2009). Generalizations can only be made to family members and caregivers who fit the profile described earlier.
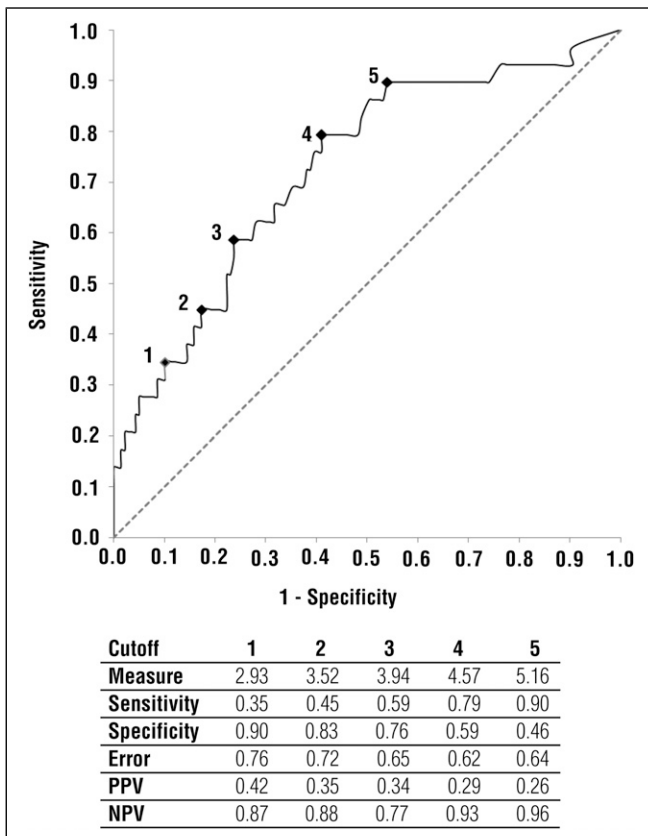
Figure 2. Receiver operating characteristic curve with cutpoints based on the family members' and caregivers' ratings. Area under the curve = 0.726, 95% confidence interval = (0.622, 0.829), $p \le .01$.
Note. NPV = negative predictive value; PPV = positive predictive value.

The AUC of the older drivers' self-assessment based on SDBM, although statistically significant, yielded low accuracy in predicting the on-road driving test results. We therefore conclude that the SDBM, when used by drivers, is not an accurate self-report screening tool to make determinations regarding on-road outcomes. That being said, drivers' ratings may still be used by occupational therapists in discussing differences between drivers' self-ratings and those of family members and caregivers to increase self-awareness of driving behaviors. Likewise, the driver report may also be used, in combination with the caregiver's report, to start a conversation about future driving interventions, driving alternatives, or driving cessation.

The family members' and caregivers' AUC yielded acceptable accuracy for using the SDBM measure to predict outcomes of the on-road driving test. Several previous studies have used caregivers to provide a proxy report on older drivers' driving errors (Wild & Cotrell, 2003) and behaviors (Croston et al., 2009). Similarly, in our previous work we have shown that family members' and caregivers' ratings on the SDBM are reliably correlated with driving evaluators' SDBM ratings (Classen et al., 2012b). We propose that these finding have implications for both research and clinical practice.

Implications for Future Research

The implication for future research is that even though the family members' and caregivers' ROC findings illustrate acceptable AUC, using a cutoff point to achieve good sensitivity results in a large number of false positives. For example, using a cutoff point of 5 yields a sensitivity of .79 and a specificity of .59. To improve the SDBM's accuracy, we are testing the efficacy of a caregiver training program to enhance its accuracy in identifying driving difficulties in older drivers. Although preliminary findings are promising, this approach will have to be tested in multisite, multicenter settings with representative samples to make population-based generalizations.

Implications for Occupational Therapy Practice

The SDBM is one of few screening tools for use by family members and caregivers to rate older drivers' behaviors. To our knowledge, this screening tool is the first showing concurrent criterion validity for family members' and caregivers' reports in classifying older drivers who fail an on-road test. As such, occupational therapists may use this screening tool (completed by family members and caregivers) to form a picture of the driver's driving behaviors. This screening tool may also be used to facilitate a conversation about difficulty with driving (from the caregiver's perspective, client's perspective, or both) and help in identifying driving problems, which may in turn lay the foundation for intervention planning by a CDRS or evaluator. Moreover, the SDBM operationalizes driving by means of 68 behavioral items. Thus, it gives the practitioner, perhaps a generalist who is not extensively familiar with all the underlying driving-related issues, a concrete description of driving abilities that can be viewed as difficult to perform and provides an entry point for clinical decision making, intervention, adaptation (e.g., suggesting safer strategies, such as not driving on the interstate), or referral to a driving rehabilitation specialist.

Limitations beyond those already mentioned (e.g., race) pertain to the error associated with the family members' and caregivers' SDBM ratings, as well as the less-than-desirable specificity and low PPV. Only two sites were involved in the testing of participants. A Web-based tool (in development) may enhance our chances of involving more sites in continued research.

Conclusion

Family members or caregivers may be a group providing valid and reliable ratings of older adults' driving behaviors. This study established that the SDBM, when used by family members and caregivers to rate the driving behaviors of older drivers, has achieved concurrent criterion validity for on-road outcomes but requires further validation (a larger research study with a more representative sample). Clinically, this screening tool may be useful for occupational therapy practitioners to make decisions regarding intervention or referral or to start conversations about driving cessation. Future developments for Web-based completion, receiving outputs, and formulating action-oriented recommendations are under way. ▲

Acknowledgments

References

AAA Foundation for Traffic Safety. (2010). *Drivers 65 plus: Check your own performance: A self-rating form of questions, facts and suggestions for safe driving.* Washington, DC: Author.

AARP, & Andrus Foundation. (1996). *Development of an assessment instrument for the clinical evaluation of elderly drivers.* Washington, DC: Transportation Research Board of the National Academies. Retrieved October 1, 2011, from http://trid.trb.org/view.aspx?id=649248

Ball, K. K., & Owsley, C. (1993). The Useful Field of View test: A new technique for evaluating age-related declines in visual function. *Journal of the American Optometric Association, 64,* 71–79.

Bédard, M., Stones, M. J., Guyatt, G. H., & Hirdes, J. P. (2001). Traffic-related fatalities among older drivers and passengers: Past and future trends. *Gerontologist, 41,* 751–756. http://dx.doi.org/10.1093/geront/41.6.751

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Erlbaum.

Cheung, I., & McCartt, A. T. (2011). Declines in fatal crashes of older drivers: Changes in crash risk and survivability. *Accident Analysis and Prevention, 43,* 666–674. http://dx.doi.org/10.1016/j.aap.2010.10.010

Classen, S., McCarthy, D. P., Shechtman, O., Awadzi, K. D., Lanford, D. N., Okun, M. S., . . . Fernandez, H. H. (2009). Useful Field of View as a reliable screening measure of driving performance in people with Parkinson's disease: Results of a pilot study. *Traffic Injury Prevention, 10,* 593–598. http://dx.doi.org/10.1080/15389580903179901

Classen, S., Wen, P., Velozo, C., Bédard, M., Winter, S. M., Brumback, B., & Lanford, D. N. (2012a). Psychometrics of the self-report Safe Driving Behavior Measure for Older Adults. *American Journal of Occupational Therapy, 66,* 233–241. http://dx.doi.org/10.5014/ajot.2012.001834

Classen, S., Wen, P., Velozo, C. A., Bédard, M., Winter, S. M., Brumback, B., . . . Lanford, D. N. (2012b). Rater reliability and rater effects of the Safe Driving Behavior Measure. *American Journal of Occupational Therapy, 66,* 69–77. http://dx.doi.org/10.5014/ajot.2012.002261

Classen, S., Winter, S. M., Velozo, C. A., Bédard, M., Lanford, D. N., & Lutz, B. J. (2010). Item development and validity testing for a self- and proxy report: The Safe Driving Behavior Measure. *American Journal of Occupational Therapy, 64,* 296–305. http://dx.doi.org/10.5014/ajot.64.2.296

Croston, J., Meuser, T. M., Berg-Weger, M., Grant, E. A., & Carr, D. B. (2009). Driving retirement in older adults with dementia. *Topics in Geriatric Rehabilitation, 25,* 154–162.

Di Stefano, M., & Macdonald, W. (2005). On-the-road evaluation of driving performance. In J. M. Pellerito (Ed.), *Driver rehabilitation and community principles and practice* (pp. 255–274). St. Louis, MO: Elsevier-Mosby.

Eby, D. W., Molnar, L. J., Shope, J. T., Vivoda, J. M., & Fordyce, T. A. (2003). Improving older driver knowledge and self-awareness through self-assessment: The driving decisions workbook. *Journal of Safety Research, 34,* 371–381. http://dx.doi.org/10.1016/j.jsr.2003.09.006

Edwards, J. D., Ross, L. A., Wadley, V. G., Clay, O. J., Crowe, M., Roenker, D. L, & Ball, K. K. (2006). The Useful Field of View Test: Normative data for older adults. *Archives of Clinical Neuropsychology, 21,* 275–286. http://dx.doi.org/10.1016/j.acn.2006.03.001

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198. http://dx.doi.org/10.1016/0022-3956(75)90026-6

Justiss, M. D., Mann, W. C., Stav, W. B., & Velozo, C. A. (2006). Development of a behind-the-wheel driving performance assessment for older adults. *Topics in Geriatric Rehabilitation, 22,* 121–128.

Marottoli, R. A., Cooney, L. M., Jr., Wagner, R., Doucette, J., & Tinetti, M. E. (1994). Predictors of automobile crashes and moving violations among elderly drivers. *Annals of Internal Medicine, 121,* 842–846.

National Alliance for Caregiving, & AARP. (2009). *Caregiving in the U.S.: A focused look at those caring for the 50+.* Bethesda, MD: National Alliance for Caregiving and MetLife. Retrieved October 1, 2011, from http://assets.aarp.org/rgcenter/il/caregiving_09.pdf

National Center for Injury Prevention and Control. (2009). *CDC injury research agenda, 2009–2018.* Atlanta, GA: Centers for Disease Control and Prevention. Retrieved October 1, 2011, from www.cdc.gov/injury/ResearchAgenda/CDC_Injury_Research_Agenda-a.pdf

Portney, L., & Watkins, M. P. (2000). *Foundations of clinical research: Applications to practice* (2nd ed.). Upper Saddle River, NJ: Prentice Hall Health.

Posse, C., McCarthy, D. P., & Mann, W. C. (2006). A pilot study of interrater reliability of the Assessment of Driving-Related Skills: Older driver screening tool. *Topics in Geriatric Rehabilitation, 22,* 113–120.

Shechtman, O., Classen, S., Awadzi, K. D., & Mann, W. (2009). Comparison of driving errors between on-the-road and simulated driving assessment: A validation study. *Traffic Injury Prevention, 10,* 379–385. http://dx.doi.org10.1080/15389580902894989

Staplin, L., & Dinh-Zarr, T. B. (2006). Promoting rehabilitation of safe driving abilities through computer-based clinical and personal screening techniques. *Topics in Geriatric Rehabilitation, 22,* 129–138.

Stav, W. B., Justiss, M. D., McCarthy, D. P., Mann, W. C., & Lanford, D. N. (2008). Predictability of clinical assessments for driving performance. *Journal of Safety Research, 39,* 1–7. http://dx.doi.org/10.1016/j.jsr.2007.10.004

Streiner, D. L., & Cairney, J. (2007). What's under the ROC? An introduction to receiver operating characteristics curves. *Canadian Journal of Psychiatry, 52,* 121–128.

Wild, K., & Cotrell, V. (2003). Identifying driving impairment in Alzheimer disease: A comparison of self and observer reports versus driving evaluation. *Alzheimer Disease and Associated Disorders, 17,* 27–34. http://dx.doi.org/10.1097/00002093-200301000-00004

Winter, S. M., Classen, S., Bédard, M., Lutz, B., Velozo, C. A., Lanford, D. N., & Brumback, B. A. (2011). Focus group findings for the self-report Safe Driving Behaviour Measure. *Canadian Journal of Occupational Therapy, 78,* 72–79. http://dx.doi.org/10.2182/cjot.2011.78.2.2