Published in final edited form as:
J Am Geriatr Soc. 2009 April ; 57(4): 697–702. doi:10.1111/j.1532-5415.2009.02185.x.

Reliability and Validity of Prehospital Case Finding for Depression and Cognitive Impairment

Manish N. Shah, MD MPH[a,b,c], **Jurgis Karuza, PhD**[c], **Erik Rueckmann, MD**[a], **Peter Swanson**[a], **Yeates Conwell, MD**[d], and **Paul Katz, MD**[c]

a *Department of Emergency Medicine, University of Rochester Medical Center, Rochester, NY*

b *Department of Community & Preventive Medicine, University of Rochester Medical Center, Rochester, NY*

c *Department of Medicine, Division of Geriatrics and Aging, University of Rochester Medical Center, Rochester, NY*

d *Department of Psychiatry, University of Rochester Medical Center, Rochester, NY*

Abstract

Context—The emergency medical services (EMS) system has the potential for a more global public health role to identify unmet needs of older adults.

Objective—To evaluate the test-retest reliability, the concurrent criterion validity, and the construct validity of prehospital, EMS case finding for depression and cognitive impairment in older adults.

Design—Cross-sectional study.

Setting—Prehospital EMS system and hospital emergency department.

Participants—EMS providers and community-dwelling older adult (age≥60) patients.

Interventions—Case finding instruments for depression (PHQ-2) and cognitive impairment (Six Item Screener.)

Main Outcome Measures—The reliability and validity of these instruments.

Results—We found moderate test-retest reliability for prehospital application of the PHQ-2 (kappa=0.50) and Six Item Screener (kappa=0.52); fair concurrent criterion validity for depression (kappa=0.36) and slight to fair concurrent criterion validity for cognitive impairment (kappa=0.11–0.23). Construct validity was demonstrated through the Multitrait-Multimethod Matrix.

Conclusions—We demonstrated moderate test-retest reliability and construct validity for prehospital case finding by EMS providers for cognitive impairment and depression using these instruments. Slight to fair concurrent criterion validity was found, a result that could be explained by methodological limitations. These findings provide additional support for the concept of using EMS providers to detect older adults at risk for these conditions. Further work is needed to confirm the validity and effectiveness of prehospital screening before such programs are implemented.

Corresponding Author: Manish N. Shah, MD MPH, 601 Elmwood Avenue, Box 655, Rochester, NY 14642, 585-463-2920 (office), 585-463-2966 (fax), E-mail: manish.shah@rochester.edu.
Alternate Corresponding Author: Erik Rueckmann, MD, Department of Emergency Medicine, University of Rochester Medical Center, 601 Elmwood Avenue, Box 655, Rochester, NY 14642, Phone: 585-463-2920, Fax: 585-463-2966, erik_reuckmann@urmc.rochester.edu

This study will be presented in abstract form at the National Association of EMS Physicians, 2009 Annual Meeting in Jacksonville, Florida.

Trial Registration: www.clinicaltrials.gov. NCT00545155

Keywords

Emergency medical services; Prehospital care; Depression; Cognitive Impairment; Geriatrics

INTRODUCTION

Older adults are a large, vulnerable segment of the population and many of these individuals suffer from undetected health problems. Failure of healthcare providers to identify and treat these conditions can lead to unnecessary morbidity and mortality and a decreased quality of life.[1,2]

Depression and cognitive impairment are significant conditions that impact older adults and often go undetected. Up to 20% of community dwelling older adults suffer from clinically significant depressive symptoms, which are associated with increased use of medical services, disability, and death.[3,4] However, estimates suggest that less than one-half of depressed older adults are identified and treated despite the existence of effective therapies.[5] This failure to identify and treat depressed older adults results in many at risk individuals suffering unnecessarily.

Cognitive impairment is also common, affecting 10% of community dwelling seniors and over 20% of older adult emergency department (ED) patients.[6,7] It has been associated with significant morbidity and mortality and increased healthcare use and cost.[2,8,9] Detection, particularly early detection, of cognitive impairment benefits older adults because reversible causes can be treated, existing therapies to slow the progression of the disease can be provided, and families can prepare for the advancement of the disease. However, cognitive impairment is poorly identified and treated in primary care practices.[10,11] This failure to identify and treat cognitively impaired patients results in many at risk individuals suffering unnecessarily.

Traditionally, screening and case finding have been the responsibility of the primary care provider (PCP). This role has become more difficult as demands on PCPs' time increase, and results in individuals who are at risk for preventable or treatable conditions not being identified.[12,13] One potential solution to this problem involves developing innovative case finding systems in alternative settings, and then communicating those findings to intervention programs. If successful, integrated case finding and intervention programs would improve the health and quality of life of these individuals. Given the prevalence and clinical importance of cognitive impairment and depression and the existence of interventions, older adults may particularly benefit from case finding to identify these conditions.

One alternative setting proposed for case finding is the prehospital, emergency medical services (EMS) setting because older adults frequently use the EMS system and EMS providers spend extended periods of time with patients.[14] During emergency (911) responses, EMS providers may be able to use screening instruments to identify older adults at risk for various conditions, thus helping to ensure that at risk individuals are identified. This information can be then provided to emergency physicians for acute evaluation and intervention, to PCPs for outpatient evaluation and interventions, the patient and/or family members to pursue further evaluation and interventions, or to a single point of entry system that would coordinate evaluation and interventions. Recent work has documented EMS providers' interest in participating in public health activities during emergency responses and the feasibility of prehospital case finding.[15,16,17,18]

However, the reliability and validity of EMS case finding during emergency responses have not been evaluated. This is an essential next step to evaluating the effectiveness of EMS case

finding, which is critical to inform policy changes that make greater public health use of EMS.[19]

The objective of this study was to evaluate the test-retest reliability, the concurrent criterion validity, and the construct validity of EMS case finding for depression and cognitive impairment using previously validated screening tools for depression and cognitive impairment. We hypothesized that: 1) EMS providers can reliably use the Patient Health Questionnaire-2 (PHQ-2) instrument to identify older adults with depression and the Six Item Screener (SIS) to identify older adults with cognitive impairment; 2) the PHQ-2 and SIS instruments have concurrent criterion validity in the prehospital setting; and 3) prehospital case finding for these conditions has construct validity when evaluated using the Multitrait-Multimethod Matrix.

METHODS

Design

We performed a cross-sectional study of older adults (age≥60) cared for by participating EMS providers between June and December 2007. The institution's Research Subjects Review Board approved this study with informed consent.

Setting

This study took place in Monroe County, NY, and involved two EMS agencies, Rural Metro Medical Services and Henrietta Volunteer Ambulance. Both agencies provide basic life support care, such as splinting and lifting, with emergency medical technicians (EMTs) and advanced life support care, such as intravenous medications and intubation, with paramedics.

Protocol

To develop the case finding instrument, we identified tools validated in primary care to identify older adults with depression or cognitive impairment. For depression, we chose the PHQ-2 and for cognitive impairment, we chose the SIS.[20,21,22]

EMS providers from both agencies were invited to participate. Eighteen EMS providers, 12 paramedics and six EMTs consented, completed the training, and participated. The training consisted of the basic life support version of the *Geriatrics Education for EMS* course and reviewing the study procedures and the case finding instruments.

Participating EMS providers were instructed to apply the case finding instrument to all community-dwelling patients aged 60 and older who requested emergency assistance and transport to either of two University EDs. Patients who could not speak English or who refused transport were excluded.

After arriving at the ED, the EMS providers called the study coordinator to provide the screening results. Separately, a trained research assistant, blinded to the EMS results, interviewed consenting patients. Subjects completed the SIS and the PHQ-2 to evaluate the test-retest reliability; completed the Mini-Cog and the CLOX to assess the concurrent criterion validity of the SIS and completed the Patient Health Questionnaire-9 (PHQ-9) to assess the concurrent criterion validity of the PHQ-2.[23,24,25] Finally, the research assistant used the Confusion Assessment Method (shortened version) to evaluate for delirium.[26] The medical records were reviewed for patients' medication list, medical history, and presenting complaint and vital signs.

Two physician investigators (MNS, ER) scored the Mini-Cog and CLOX independently. In the event of a discrepancy between the two scores, the two investigators discussed the results and reached a consensus score.

ED and EMS medical records were also reviewed for those individuals who elected not to participate in the study using a short, structured data abstraction instrument.

Scales

We used two scales, derived from the PRIME-MD, to evaluate depression.[21] We used a version of the PHQ-2 with dichotomous answers for EMS case finding due to the ease of administration and scoring and excellent test characteristics when used in a multistage process.[22] Answering positive to either question was considered positive for depression. The PHQ-9, used for criterion validation, has been validated in the primary care setting. We chose this for the criterion standard because it has excellent test characteristics, with a sensitivity and specificity of 88%.[25] We used the diagnostic algorithm to determine whether patients screened positive for minor or major depression. From one perspective, the comparison of the PHQ-2 to the PHQ-9 could be seen as cross validation of the PHQ-2 because the PHQ-2 items appear in the PHQ-9. However, we feel that applying the term “concurrent criterion validity,” although not ideal, is reasonable given the dissimilarities in the administration and scoring of the version of the PHQ-2 used and the fact that the PHQ-9 has the credibility as the established test in the literature.

For cognitive impairment, we chose the SIS because it is easy to administer and score and has excellent test characteristics.[20] It has been validated in the primary care setting. Greater than two errors on the instrument were considered positive for cognitive impairment. For criterion validation, we chose two instruments. The Mini-Cog has been validated in community samples and it has excellent test characteristics.[24] The CLOX evaluates executive dysfunction and has been shown to be easy to administer and well tolerated by older adults.[23] It includes an unprompted clock-drawing task sensitive to executive control (COLX1) and a prompted version that is not (CLOX2). For both tests, standard scoring schemes were used.

Data Analysis

All data were analyzed using Stata 8.0. Descriptive statistics were calculated to characterize the individuals participating in the study. Those who did not consent were compared to those participating to evaluate for enrollment bias. Continuous data were compared using a t-test and categorical data were compared using the chi-square test.

Test-retest reliability was evaluated by calculating the percent concordance in final diagnoses and the kappa statistic with 95% confidence intervals(CI). Concurrent criterion validity was calculated in two ways. First, the sensitivity and specificity of the EMS applied instrument, as compared to the reference standard instrument, was calculated. Second, the kappa statistic between the two instruments were calculated. Finally, we augmented the concurrent criterion validation by examining the validity of the scales using the construct validity approach. The study design permits the use of the Multitrait-Multimethod approach, which measures multiple traits (e.g., cognitive function and depression) using multiple screening methods and tools to determine both the convergent and discriminant validity of the scales.[27]

RESULTS

Participating EMS providers screened 269 eligible subjects; 187 (70%) consented to participate, while the remaining refused (43, 16%) or could not be consented (39, 14%) due to the absence of decisional capacity and a surrogate. Table 1 demonstrates the characteristics of

older adults who consented to participate in the study and the limited characteristics obtained on older adults who refused to participate. Table 1 also demonstrates the results from the application of the various instruments on the study population. The Mini-Cog and CLOX could not be completed by a large number of subjects (14%, 30% respectively) due to the inability to draw clock faces. Notably elevated levels of cognitive impairment were found among those completing the Mini-Cog and CLOX, compared to the SIS.

Table 2 shows the reliability of performing the SIS and PHQ-2 in the prehospital setting. The kappa's showed moderate test-retest reliability. The concordance between the two applications of the SIS was 87% (161/185; 95% CI: 81%–92%) and the concordance between the two applications of the PHQ-2 was 75% (128/170; 95% CI: 68%–82%). When stratified for presence of delirium no statistically significant differences were found in the reliability of performing either instrument. Furthermore, stratifying by cognitive impairment (SIS applied in the ED) did not find any statistically significant differences in the PHQ-2 reliability (results not shown).

The concurrent criterion validity of the prehospital case finding is also shown in Table 2. The kappa's for depression show fair agreement while the kappa's for cognitive impairment show fair (SIS vs. Mini-Cog) to slight (SIS vs. CLOX1 or 2) agreement. When stratified for the presence of delirium, no statistically significant differences were found for cognitive impairment or depression screening (results not shown).

Construct validity is demonstrated in Table 3. We found that the monotrait-monomethod correlations, which reflect the relationship between measures of the same trait, using the same method are among the greatest (0.48, 0.53), followed by the monotrait-heteromethod correlations, which reflect the relationship between measures of the same trait, using different measures, (range from 0.12 to 0.58), followed by heterotrait-heteromethod, which reflect the relationship between measures that neither share the same trait or the same method (range from −0.12 to 0.013). This pattern, where the cognitive case finding results were more highly intercorrelated among themselves than with the depression screening results, and the depression screening results were more highly intercorrelated among themselves than the with the cognitive screening results, provides evidence for the convergent and discriminant validity of the two screening tools. This is the order expected for a set of measures with construct validity.

DISCUSSION

In this study, we found that the test-retest reliability of prehospital, EMS provider application of the PHQ-2 was moderate (kappa=0.50). Furthermore, we found that the test-retest reliability of prehospital application of the SIS was also moderate (kappa=0.52). This level of reliability is very good and supports prehospital case finding for depression with the PHQ-2 and cognitive impairment with the SIS.

These results must be tempered by three factors that may have altered the reliability. First, our evaluation was not ideal because the reference standard assessment was conducted in the ED, not the ambulance, and a short delay occurred between the two assessments. However, practical and ethical considerations limited our ability to perform the ideal evaluation. This change threatens the external validity and potentially reduces the level of reliability identified. Second, a Hawthorne effect may have existed, potentially improving the results. We had a highly motivated group of EMS providers who were aware they were being evaluated. If universally implemented, this program may not be as successful. Third, prior to this study, EMS providers had not regularly evaluated their patients for depression or cognitive impairment, either as part of case finding or clinical care. As the EMS providers become more comfortable with these

concepts and instruments, their skills and their results may improve, resulting in better reliability. Nonetheless, we believe that the reliability testing results support continued research evaluating prehospital case finding for depression and cognitive impairment with the PHQ-2 and SIS.

The concurrent criterion validity testing for depression showed only fair performance of the prehospital PHQ-2. In particular, the prehospital PHQ-2 suffered from low specificity, which may have led to this unanticipated performance level. The reasons for this result were not specifically explored by this study, but the low specificity may stem from this version of PHQ-2 used, which itself only has a specificity of 57%.[22] This theory is supported by the fact that trained study staff using the same instrument found similarly elevated levels of individuals screening positive (40% vs. 48%) even though they found only 22% of individuals screen positive with the PHQ-9. It is possible that using the PHQ-2 with the interval scoring system rather than the dichotomous scoring system may result in better test performance, but the added difficulty of use may serve as a barrier to use.[28] Future studies need to evaluate the performance of the PHQ-2 with the interval scoring system to better determine the reason for this fair performance and to identify ways to improve the validity of prehospital case finding for depression.

The concurrent criterion validity testing for cognitive impairment showed only slight to fair performance of the SIS, with kappas ranging from 0.11–0.23. However, evidence points to difficulties related to the chosen reference standards. In applying the Mini-Cog and CLOX, it became evident that these tests are inappropriate for the emergency setting. Almost a quarter of subjects could not draw the clock faces due to limitations such as limited arm mobility or visual difficulties, thus potentially biasing the results. Additionally, for subjects who completed the test, the proportion with cognitive impairment was atypically high, lacking face validity. This study was not structured to evaluate this surprising finding, but we hypothesize that distractions in the ED may have impacted subjects' ability to concentrate on the clock drawing, resulting in worse performance on the two instruments. Despite these challenges, the results are encouraging. Future studies need to evaluate the validity of the prehospital application of the SIS using instruments that are appropriate to the ED.

The construct validity, which found the correlation coefficients of similar items being greater than different items, supports our hypotheses. One additional benefit of this approach is that it helps evaluate method variance. Especially problematic in the use of screening tools in such unorthodox environments is the confound of method variance—that the results reflect more of the variance due to the administration method of rather than the actual trait being measured. That the intercorrelations between the measures of cognitive function and depression within the ED and within the EMS setting were in the low range, strongly suggests the scores were not an artifact of the location and method of administration.

Despite the challenges encountered in this study, the results were encouraging. The moderate reliability between the prehospital and study staff applied PHQ-2 and SIS and the construct validity identified supports the continued development of the prehospital use of these instruments for case finding. The research agenda can move to validity testing, taking into account the lessons learned from the concurrent criterion validity testing that we performed. A diagnostic standard evaluation such as detailed neuropsychiatric testing for cognitive impairment and semi-structured diagnostic interviews for depression to evaluate the validity of the EMS case finding activity would be appropriate. If those results prove to be accurate, then formal program development and effectiveness trials can be performed. In addition, case finding for other conditions can be considered, including, for instance, vaccination status, fall risk, abuse, and medication management strategies.

A number of limitations were identified. First, the generalizability of these results is unknown, as this study took place with a small group of EMS providers who were motivated to participate. However, the goal was to evaluate the reliability and criterion validity of prehospital implementation of these instruments, which relates primarily to the instruments. Second, the research staff in the ED performed both the screening tests and the criterion standard measurements. Thus, administration of one measure was done by individuals not blinded to another, potentially improving the concurrent criterion validity results. Third, a number of subjects did not consent or could not consent to participate in the study. Based on the limited data collected on these individuals, we do not expect that the bias would impact the reliability and validity results although it may impact the proportion of older adults found to have needs. Finally, we did not evaluate anxiety, which is interrelated with depression. Future studies will also need to evaluate anxiety.

CONCLUSION

We demonstrated moderate test-retest reliability and construct validity for prehospital case finding by EMS providers for cognitive impairment using the Six Item Screener and depression using the PHQ-2. Slight to fair concurrent criterion validity was found, a result that could be explained by methodological limitations. These findings provide additional support for the concept of using EMS providers to detect older adults at risk for these conditions. Further work is needed to confirm the validity and effectiveness of prehospital screening using these instruments before such programs are implemented.

Acknowledgements

Dr. Shah is supported by the Paul B. Beeson Career Development Award (NIA 1K23AG028942). We would like to thank the EMS providers from Henrietta Volunteer Ambulance and Rural Metro Medical Services, Rochester, NY, for their assistance with this study.

References

1. Stuck AE, Walthert JM, Nikolaus T, et al. Risk factors for functional status decline in community-living elderly people: a systematic literature review. Soc Sci Med 1999;48:445–469. [PubMed: 10075171]
2. Blazer DG, Sachs-Ericsson N, Hybels CF. Perception of unmet basic needs as a predictor of mortality among community-dwelling older adults. Am J Public Health 2005;95:299–304. [PubMed: 15671468]
3. Unutzer J, Patrick DL, Simon G, et al. Depressive symptoms and the cost of health services in HMO patients aged 65 years and older. JAMA 1997;277:1618–1623. [PubMed: 9168292]
4. Callahan CM, Hui SL, Nienaber NA, et al. Longitudinal study of depression and health services use among elderly primary care patients. J Am Geriatr Soc 1994;42:833–838. [PubMed: 8046192]
5. Unutzer J, Katon W, Callahan CM, et al. Depression in a sample of 1801 depressed older adults in primary care. J Am Geriatr Soc 2003;51:505–514. [PubMed: 12657070]
6. Hustey FM, Meldon SW. The prevalence and documentation of impaired mental status in elderly emergency department patients. Ann Emerg Med 2002;39:248–253. [PubMed: 11867976]
7. Graham JE, Rockwood K, Beattie BL, et al. Prevalence and severity of cognitive impairment with and without dementia in an elderly population. Lancet 1997;349:1793–1796. [PubMed: 9269213]
8. Stump TE, Callahan CM, Hendrie HC. Cognitive impairment and mortality in older primary care patients. J Am Geriatr Soc 2001;49:934–940. [PubMed: 11527485]
9. Zhao Y, Kuo TC, Weir S, et al. Healthcare costs and utilization for Medicare beneficiaries with Alzheimer's. BMC Health Serv Res 2008;8:108.10.1186/1472-6963-8-108 [PubMed: 18498638]
10. Callahan CM, Hendrie HC, Tierney WM. Documentation and evaluation of cognitive impairment in elderly primary care patients. Ann Intern Med 1995;122:422–429. [PubMed: 7856990]
11. Boise L, Neal BM, Kaye J. Dementia assessment in primary care: Results from a study in three managed care systems. J Gerontol A Biol Sci Med Sci 2004;59:M621–M626. [PubMed: 15215282]

12. Yarnall KSH, Pollack KI, Ostbye T, et al. Primary care: Is there enough time for prevention? Am J Public Health 2003;93:635–641. [PubMed: 12660210]
13. Frame PS. Health maintenance in clinical practice: strategies and barriers. Am Fam Phys 1992;45:1192–1200.
14. Shah MN, Bazarian JJ, Lerner EB, Fairbanks RJ, Barker W, Auinger P, Friedman B. The epidemiology of emergency medical services use by older adults: an analysis of the National Hospital Ambulatory Medical Care Survey. Acad Emerg Med 2007;14:441–447. [PubMed: 17456555]
15. Lerner EB, Shah MN, Fernandez AR. Do EMS providers think they should participate in disease prevention? Prehosp Emerg Care 2008;12(1):108–109.
16. Shah MN, Clarkson L, Lerner EB, et al. An emergency services program to promote the health of older adults. J Am Geriatr Soc 2006;54:956–962. [PubMed: 16776792]
17. Gerson LW, Schelble DT, Wilson JE. Using paramedics to identify at-risk elderly. Ann Emerg Med 1992;21(6):688–691. [PubMed: 1590608]
18. Shah MN, Lerner EB, Chiumento S, et al. An evaluation of paramedics' ability to screen older adults during emergency responses. Prehosp Emerg Care 2004;8:298–308. [PubMed: 15295732]
19. Delbridge TR, Bailey B, Chew JL Jr, et al. EMS agenda for the future: where we are … where we want to be. EMS Agenda for the Future Steering Committee. Ann Emerg Med 1998;31(2):251–263. [PubMed: 9472190]
20. Callahan CM, Wnverzagt FW, Hui SL, et al. Six-item screener to identify cognitive impairment among potential subjects for clinical research. Med Care 2002;40:771–781. [PubMed: 12218768]
21. Spitzer RL, Williams JB, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. JAMA 1994;272:1749–56. [PubMed: 7966923]
22. Whooley MA, Avins AL, Miranda J, et al. Case finding instruments for depression. Two questions are as good as many. Gen Intern Med 1997;12:439–45.
23. Royall DR, Cordes JA, Polk M. CLOX: An executive clock drawing task. J of Neurol, Neurosurg, Psychiatry 1998;64:588–594. [PubMed: 9598672]
24. Borson S, Scanlan JM, Chen P, et al. The Mini-Cog as a screen for dementia: validation in a population-based sample. J Am Geriatr Soc 2003;51:1451–1454. [PubMed: 14511167]
25. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med 2001;16:606–613. [PubMed: 11556941]
26. Inouye SK, van Dyck CH, Alessi CA, et al. Clarifying confusion: the Confusion Assessment Method. Ann Intern Med 1990;113:9410948.
27. Campbell D, Fiske D. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull 1959;56:81–105. [PubMed: 13634291]
28. Li C, Friedman B, Conwell Y, et al. Validity of the Patient Health Questionnaire 2 (PHQ-2) in identifying major depression in older people. J Am Geriatr Soc 2007;55:596–602. [PubMed: 17397440]

Table 1

Characteristics of Older Adults

Characteristic	Consenting (N=187)	Not Consenting (N=82)
Age, mean (standard deviation)	76 (9.2)	76 (9.9)
Gender, Female	120 (64%)	43 (53%)
Race		
White	128 (68%)	61 (74%)
Black	55 (29%)	18 (22%)
Other/Unknown	4 (2%)	3 (4%)
Ethnicity		
Not Hispanic	181 (97%)	NA
Hispanic	1 (0.5%)	NA
No Answer	5 (2.7%)	NA
Education		
< High School	47 (25%)	NA
High School	71 (38%)	NA
College or more	58 (31%)	NA
No Answer	11 (5.9%)	
Marital Status		
Single	22 (12%)	NA
Married	59 (32%)	NA
Widowed	61 (33%)	NA
Separated/Divorced	38 (20%)	NA
No Answer	7 (3.7%)	NA
Accessed ED For Care in Past 6 Months	78 (42%)	NA
Hospitalized in Past 6 Months	43 (23%)	NA
Chief Complaint*		
Trauma/Hemorrhage	40 (21%)	17 (21%)
Cardiac	33 (18%)	4 (5%)
GI/GU	22 (12%)	9 (11%)
Syncope/Dizzy	23 (12%)	6 (7%)
Respiratory	25 (13%)	4 (5%)
Pain, not traumatic	12 (6%)	9 (11%)
Weakness/General Illness	10 (5%)	9 (11%)
Neurological (CVA, Seizure)	4 (2%)	7 (9%)
Endocrine	8 (4%)	1 (1%)
Alcohol/Drug Overdose	1 (1%)	8 (10%)
Confusion	2 (1%)	3 (4%)
Psychiatric	1 (1%)	4 (5%)
Other	6 (3%)	0 (0%)
Length of Illness Symptoms		
<4 hours	53 (28%)	NA

Characteristic	Consenting (N=187)	Not Consenting (N=82)
>=4 hours	134 (72%)	NA
Past Medical History		
Cancer	46 (25%)	NA
CHF	25 (13%)	NA
Dementia	16 (8.6%)	NA
Diabetes	71 (38%)	NA
Heart Disease	68 (36%)	NA
Hypertension	133 (71%)	NA
Kidney Disease	29 (16%)	NA
Liver Disease	8 (4.3%)	NA
Lung Disease	45 (24%)	NA
Stroke	28 (15%)	NA
Glasgow Coma Scale		
Score=15	178 (95%)	72 (88%)
Cognitive Impairment Evaluation		
Six Item Screener by EMS, impaired	31/186 (17%)	NA
Six Item Screener by Study Staff, impaired	29/186 (16%)	NA
Mini-Cog by Study Staff, impaired	88/161 (55%)	NA
CLOX 1 by Study Staff, impaired	83/138 (60%)	NA
CLOX 2 by Study Staff, impaired	66/130 (51%)	NA
Depression Evaluation		NA
PHQ-2 by EMS, depressed	86/178 (48%)	NA
PHQ-2 by Study Staff, depressed	71/179 (40%)	NA
PHQ-9, by Study Staff, minor or major depression	40/180 (22%)	NA
Delirium Evaluation		
CAM, by Study Staff, impaired	16/175(9.1%)	NA

* $p<0.05$

Table 2

Reliability and Concurrent Criterion Validity of Depression and Cognitive Impairment Case Finding

Evaluation	Domain	Instrument	Reference Standard	Sensitivity	Specificity	Kappa
				N % 95% CI*	**N % 95% CI***	**Kappa 95% CI***
Reliability Testing	Cognitive Impairment	Six Item Screener	N/A	N/A	N/A	0.52 (95% CI: 0.36–0.69)
	Depression	PHQ-2**	N/A	N/A	N/A	0.50 (95% CI: 0.37–0.63)
Concurrent Criterion Validity Testing	Cognitive Impairment	EMS Six Item Screener	ED Mini-Cog	25/87 29% 20–39%	70/73 96% 88–99%	0.23 0.13–0.34
		EMS Six Item Screener	ED CLOX 1	17/82 21% 13–31%	51/55 93% 82–98%	0.11 0.017–0.21
		EMS Six Item Screener	ED CLOX 2	15/65 23% 14–35%	59/64 92% 83–97%	0.15 0.029–0.27
	Depression	EMS PHQ-2**	ED PHQ-9 (minor or major depression)	33/38 87% 72–96%	85/133 64% 55–72%	0.36 0.24–0.48

* 95% Confidence Interval

** Dichotomous scoring

Table 3

Multitrait-Multimethod Correlation Matrix Validation N=119

		Cognitive Impairment					Depression		
		EMS SIS	**ED SIS**	**MiniC og**	**CLOX 1**	**CLOX 2**	**EMS PHQ-2***	**ED PHQ-2***	**PH Q9**
Cognitive Impairment	**EMS SIS**								
	ED SIS	0.48							
	Mini Cog	0.30	0.20						
	CLOX1	0.18	0.12	0.56					
	CLOX2	0.21	0.19	0.35	0.35				
Depression	**EMS PHQ-2***	−0.12	−0.060	−0.090	−0.12	−0.044			
	ED PHQ-2*	−0.067	−0.0091	−0.081	−0.069	−0.050	0.53		
	PHQ-9 (minor or major depression	−0.029	0.013	−0.083	−0.045	−0.053	0.41	0.58	

*Dichotomous scoring