ORIGINAL RESEARCH | Annals of Internal Medicine

The CAM-S: Development and Validation of a New Scoring System for Delirium Severity in 2 Cohorts

Sharon K. Inouye, MD, MPH; Cyrus M. Kosar, MA; Douglas Tommet, MS, MPH; Eva M. Schmitt, PhD; Margaret R. Puelle, BS; Jane S. Saczynski, PhD; Edward R. Marcantonio, MD, SM*; and Richard N. Jones, ScD*

Background: Quantifying the severity of delirium is essential to advancing clinical care by improved understanding of delirium effect, prognosis, pathophysiology, and response to treatment.

Objective: To develop and validate a new delirium severity measure (CAM-S) based on the Confusion Assessment Method.

Design: Validation analysis in 2 independent cohorts.

Setting: Three academic medical centers.

Patients: The first cohort included 300 patients aged 70 years or older scheduled for major surgery. The second included 919 medical patients aged 70 years or older.

Measurements: A 4-item short form and a 10-item long form were developed. Association of the maximum CAM-S score during hospitalization with hospital and posthospital outcomes related to delirium was evaluated.

Results: Representative results included adjusted mean length of stay, which increased across levels of short-form severity from 6.5 days (95% CI, 6.2 to 6.9 days) to 12.7 days (CI, 11.2 to 14.3 days) (P for trend < 0.001) and across levels of long-form severity from 5.6 days (CI, 5.1 to 6.1 days) to 11.9 days (CI, 10.8 to 12.9 days) (P for trend < 0.001). Representative results for the composite outcome of adjusted relative risk of death or nursing home residence at 90 days increased progressively across levels of short-form severity from 1.0 (referent) to 2.5 (CI, 1.9 to 3.3) (P for trend < 0.001) and across levels of long-form severity from 1.0 (referent) to 2.5 (CI, 1.6 to 3.7) (P for trend < 0.001).

Limitation: Data on clinical outcomes were measured in an older data set limited to patients aged 70 years or older.

Conclusion: The CAM-S provides a new delirium severity measure with strong psychometric properties and strong associations with important clinical outcomes.

Primary Funding Source: National Institute on Aging.

Ann Intern Med. 2014;160:526-533. www.annals.org

For author affiliations, see end of text.

* Drs. Marcantonio and Jones contributed equally as co–senior authors.

Delirium is a common, serious, costly, and potentially preventable condition for older persons and has been identified as the leading complication of hospitalization for this population (1, 2). Given the associated hospital mortality rates of 25% to 33% (3, 4) and annual health care costs of greater than $182 billion (in 2011 U.S. dollars) (5), delirium has assumed increasing attention as a public health and patient safety priority (2, 6, 7). The Confusion Assessment Method (CAM) is a standardized, validated measure (8) that has gained widespread use in screening for delirium (2, 9, 10). To date, the CAM has been used in more than 4000 original articles and translated into 14 languages. Although previous studies have used an additive score of CAM features as a severity indicator (11, 12), the validity of this approach has not been examined. Extending the CAM as a severity measure would greatly enhance its clinical value.

A CAM-based severity measure would have substantial utility in clinical and research applications. Clinically, a measure to track delirium severity would be useful to follow response to delirium treatment and management interventions across clinical settings. The measure could also be useful to show the graded effect of delirium on health care delivery, such as clinical care staffing and costs. Measurement of delirium severity is essential to understanding its clinical course and recovery. Thus, such a measure could serve as a primary outcome for prognostic studies and treatment trials. The measure would facilitate studies of pathophysiology, in which quantifying the level of delirium and its change over time may clarify mechanistic relationships. Of note, the availability of a standardized measure would facilitate comparison across studies.

Therefore, we developed a new CAM-based scoring system for delirium severity, called the CAM-S. The specific aims were to create the CAM-S scoring system (including a short form based on the 4-item CAM algorithm and a long-form based on the 10-item CAM instrument) (8); evaluate its distribution, test properties, and interrater reliability; examine how it performs in persons with dementia; and determine its association with clinical outcomes likely to be related to delirium severity. We hypothesized that a valid severity measure should be a strong independent predictor of adverse outcomes associated with delirium, including prolonged hospital stay, functional and cognitive decline, nursing home placement, death, and high health care costs.

See also:

Print

Editorial comment. 574

METHODS

Study Samples

The 2 samples were from prospective cohort studies with consecutive sampling, which have been described pre-

viously. The Successful Aging After Elective Surgery (SAGES) study, which is ongoing, provided the first sample (13). Potential participants were consecutive patients scheduled for elective major noncardiac surgery from 10 June 2010 to 29 March 2012 at 2 hospitals affiliated with Harvard Medical School. A total of 951 patients met the initial eligibility criteria (aged ≥70 years; scheduled for major orthopedic, vascular, or general surgical procedures under general or regional anesthesia; anticipated to have a ≥2-day hospital stay; and living within 50 miles). Of these, 446 patients were ineligible because of delirium (n = 2), clinically documented dementia (n = 14), hospitalization within 3 months (n = 78), a terminal condition (n = 12), legal blindness or severe deafness (n = 3), an inability to speak English (n = 80), or other reasons primarily related to being unreachable by telephone or unable to communicate verbally (n = 257). An additional 205 patients declined participation, and their eligibility could not be determined. The final sample included 300 patients, which is a subgroup of the total planned sample.

Project Recovery, described previously (5, 11, 14), provided the second sample. Potential participants were consecutive patients admitted to the medicine service at Yale–New Haven Hospital from 25 March 1995 to 18 March 1998. A total of 2434 patients met eligibility criteria (aged ≥70 years and no delirium on admission but intermediate to high risk for the condition). Of these, 1265 patients were excluded because of an inability to participate in interviews (n = 298, including 154 patients who were nonverbal because of profound dementia, 92 with a language barrier, 38 with severe aphasia, 12 with respiratory isolation, and 2 with intubation); coma or terminal illness (n = 69); a hospital stay less than 48 hours (n = 219); prior enrollment (n = 324); or other reasons (n = 355). An additional 250 patients declined enrollment. The final sample included 919 participants.

Assessment of clinical outcomes was conducted by research assistants (interviewers and chart abstractors) who were blinded to the CAM delirium status of the patients. The CAM ratings were conducted by a separate hospital-based team of research assistants. The screening and recruitment of the participants in each cohort was completed before the CAM was administered. For the SAGES study, informed consent for participation was obtained from all patients according to procedures approved by the institutional review boards at the Harvard-affiliated hospitals. For Project Recovery, informed consent was obtained from the patients or, for those with substantial cognitive impairment, a proxy as approved by the Yale University Institutional Review Board.

Delirium Assessment

In both cohorts, delirium was rated by using the CAM, which was scored daily during hospitalization by experienced research assistants on the basis of observations made during a standardized interview. The interview included a sleep questionnaire and brief cognitive tests, which lasted 10 to 15 minutes. Study interviewers had intensive training and standardization (15).

Context

Delirium is a common condition among hospitalized patients and is associated with poor outcomes. Although the severity of delirium may fluctuate, reliable methods to quantify it are lacking.

Contribution

This study validated a novel scale for delirium severity and showed its association with important clinical outcomes, including length of stay, functional decline, and nursing home placement or death at 90 days.

Implication

A reliable means of rating delirium severity may assist in research and improving treatments.

—The Editors

Development and Scoring of the CAM-S

The CAM-S is intended to be used in addition to the original CAM algorithm; it will not yield a delirium diagnosis but only a means to quantify the intensity of delirium symptoms seen at the bedside. These symptoms can be present in persons with and without delirium. We created a short form and a long-form for the CAM-S scoring system. The short form was based on the 4 features from the CAM diagnostic algorithm (7) that can be rated at bedside: acute onset or symptom fluctuation, inattention, disorganized thinking, and altered level of consciousness. Each symptom of delirium, except fluctuation, was rated 0 (absent), 1 (mild), or 2 (marked). Acute onset or fluctuation was rated 0 (absent) or 1 (present). The sum of these ratings yielded a short-form severity score ranging from 0 to 7, with 7 being the most severe. The long form was based on the 10 features from the full CAM instrument (8): acute onset or symptom fluctuation, inattention, disorganized thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, psychomotor agitation, psychomotor retardation, and sleep–wake cycle disturbance. Each symptom was rated 0 to 2, except acute onset or fluctuation, as previously described. The sum of these ratings yielded a long-form score ranging from 0 to 19, with 19 being the most severe. Features scored as "uncertain" did not contribute to the severity score. Uncertain ratings were present for 1 or more items in only 13 of 1456 (<1%) short-form items and 38 of 1456 (<3%) long-form items.

Interrater Reliability

To assess interrater reliability, a total of 73 paired CAM-S ratings (14 patients with delirium and 59 without) have been conducted on a quarterly basis in the SAGES

study to date, with 2 observers rating each patient simultaneously in a blinded manner.

Evaluation of Convergent Agreement

To show convergent agreement, closely related measures should be highly correlated. We assessed this by examining the correlation of daily CAM-S scores with concurrent measures of confusion and cognitive functioning completed daily during hospitalization. In the SAGES study, convergent agreement was assessed by comparing daily CAM-S scores with a brief cognitive screening test administered to patients (13) and a global rating of confusion (scored 0 to 10, with higher values indicating worse confusion) rated by the interviewers daily. In Project Recovery, convergent agreement was assessed by comparing daily CAM-S scores with the concurrent Mini-Mental State Examination (licensed from Psychological Assessment Resources) administered to patients (scored 0 to 30, with higher values indicating better mental state) (16) and a visual analogue scale for confusion (scored 0 to 100, with higher values indicating worse confusion) rated by the interviewers.

Association With Clinical Outcomes

The association with clinical outcomes was assessed in Project Recovery, where data collection was completed. These clinical outcomes are not yet available in the SAGES study. Hospital outcomes included length of stay and nursing home placement (abstracted from the medical record) and hospital costs (derived from the hospital's billing database). The costs are for hospital services that were submitted on UB-92 hospital billing forms to Medicare Part A, which typically represent about 50% to 60% of hospital charges and do not include professional fees and services (Medicare Part B). We defined functional decline as a partial or complete decline in 1 or more activity on the standard 7-item Activities of Daily Living (ADL) scale between baseline and discharge (17–19). We defined cognitive decline as a decrease of 2 or more points on the 30-point Mini-Mental State Examination score between baseline and discharge (18, 20). Posthospital outcomes comprised death within 90 days, including hospital deaths, determined from medical records, Medicare Part A and Social Security databases, the National Death Index, and death certificates (5, 21). Nursing home residence at 90 days and postdischarge health care costs per day for 90 days were obtained from Medicare Part A data. To avoid bias, the hierarchical outcome of either death or nursing home residence was used because patients who die can no longer be placed in a nursing home. Functional decline at 30 days was defined as a partial or complete decline in 1 or more activity on the standard 7-item ADL scale between baseline and the 30-day follow-up interview. No 90-day interview was conducted.

Statistical Analysis

The **Appendix** (available at www.annals.org) provides further details on validation analyses. For all analyses, we used 1 measure per patient (the highest CAM-S score during hospitalization). The only exception was convergent agreement, in which all observations were used for purposes of daily comparison. We estimated interrater reliability with overall agreement on exact scores and intraclass correlation coefficients. Convergent agreement was estimated by the Pearson correlation coefficient with daily measures of cognitive functioning in the hospital. To rule out bias due to selecting more than 1 observation per person, we repeated our analyses and selected only 1 pair of daily CAM-S and cognitive functioning scores from each person at random. We also verified fulfillment of the linearity assumption (assumed by the correlation coefficient) by comparing total variance with linear and polynomial models.

To enhance interpretability for analyses related to clinical outcomes, we divided the CAM-S scores into 4 categories defined empirically on the basis of score distributions. Poisson regression was used to calculate adjusted relative risks (RRs) for outcomes related to nursing home residence, death, and cognitive and functional decline. For analyses relevant to new nursing home placement, we excluded patients who were deceased or were already living in a nursing home before the relevant time frame. Finally, we modeled hospital length of stay and Medicare costs using log-γ regression (22). Predictive margins (23) were obtained and presented as the adjusted mean length of stay and costs. Because delirium severity was strongly related to death, we calculated costs per day to avoid bias in the posthospital analyses. These costs were not available for 88 patients with missing cost data, including patients who had died or were enrolled in HMOs. Linear trend tests were used to evaluate the exposure–response relationship between delirium severity and clinical outcomes. We adjusted all models for age, sex, race, Acute Physiology and Chronic Health Evaluation II score, Charlson comorbidity index score, and dementia. We also adjusted all models, except the one for functional decline, for baseline impairment in activities of daily living. For all models, we verified the robustness of our parameter estimates to potentially influential (high-leverage) observations by repeating the analyses with these observations excluded. Graphic displays of the raw data on maximum CAM-S scores by clinical outcomes are presented in **Appendix Figures 1** to **9** (available at www.annals.org).

To examine the additional contribution of delirium severity beyond delirium diagnosis, all outcomes were reexamined in analyses stratified by delirium status. We modeled the CAM-S scores as continuous measures and presented them in tertiles based on the within-group distributions to enhance interpretability. The tertiles were intended to have approximately equal-sized groups but were limited by the constraints of integer-based scores. We repeated linear trend tests in each stratum and for the overall sample.

All analyses were conducted with Stata, version 13 (StataCorp, College Station, Texas).

Role of the Funding Source

The funding sources had no role in the design, conduct, or reporting of this study.

RESULTS

Characteristics of the patients are shown in **Table 1**. The SAGES and Project Recovery samples included older adults (mean age, 77 and 80 years, respectively) with a predominance of women (55% and 60%, respectively). However, Project Recovery had higher comorbid conditions (Charlson comorbidity index score, 2.9 vs. 1.0), levels of functional and cognitive impairment, and rates of possible dementia (13% vs. 2%) at baseline. Although clinically documented dementia was an exclusionary criterion in the SAGES study, early dementia was present in 2% of participants according to a clinical consensus panel convened after enrollment. In the SAGES study sample, delirium developed during hospitalization in 68 of 300 participants (23%), with missing CAM ratings in 16 of 1456 observations (1%) on short and long forms. In Project Recovery, delirium developed during hospitalization in 115 of 919 participants (13%), with missing CAM ratings in fewer than 19 of 5202 observations (<0.5%) on short and long forms. No adverse events were associated with administering the CAM.

For the SAGES study sample, the mean CAM-S scores were 0.68 (SD, 1.22) for the short form and 2.00 (SD, 2.50) for the long form. For the Project Recovery sample, the mean scores were 0.33 (SD, 0.94) for the short form and 1.90 (SD, 2.20) for the long form. The distribution of CAM-S scores in delirious and nondelirious patients is shown in the **Figure**. For the short form, a difference of 3 or more points in the median separates the delirious and nondelirious groups in both samples, with no overlap in their interquartile ranges (25th to 75th percentile values). The same relationship is shown for the long form, with a 6-point difference in median values and no overlap in interquartile ranges. In stratified analyses, the CAM-S scores (medians and interquartile ranges) were appropriately separated between the delirious and nondelirious groups in the overall sample and in patients with dementia, which supports construct validity. Because the SAGES study excluded patients with dementia, these analyses were conducted only in the Project Recovery sample. The short form shows a 4-point difference in median values for severity, with no overlap in interquartile ranges between delirious and nondelirious patients with dementia. Similarly, the long form shows a difference of 5 or more points in median values for severity, with no overlap in interquartile ranges.

In 73 paired observations, 19% of whom were delirious patients, the mean CAM-S scores were 1.24 (SD, 1.65) for the short form and 3.00 (SD, 3.55) for the long form. The overall agreement for short-form scoring was 98% with an intraclass correlation coefficient of 0.92, indicating high agreement. For the long form, the overall agreement was 97% with an intraclass correlation coefficient of 0.88, also showing high agreement.

Table 1. **Sample Characteristics***

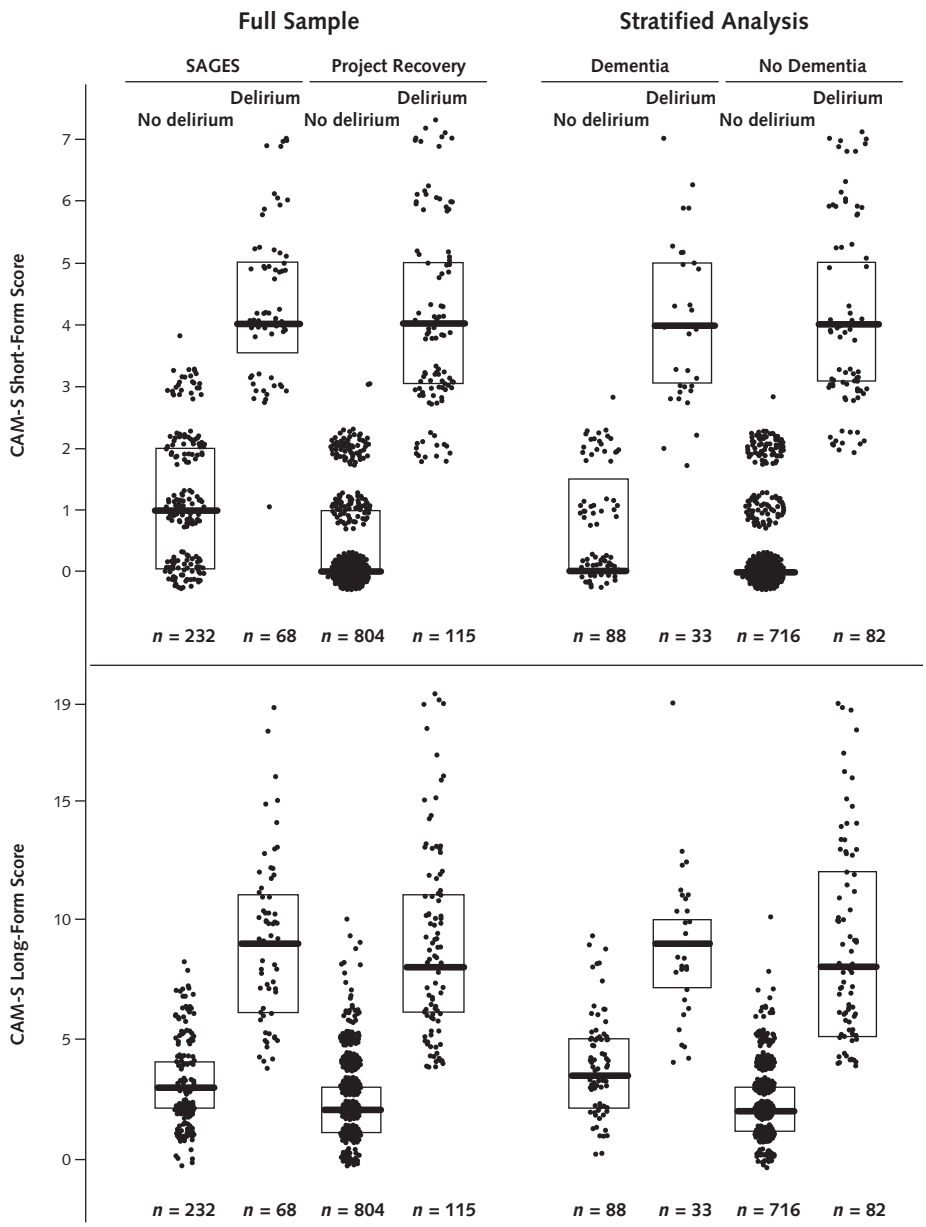
Characteristic	SAGES Study Sample (n = 300)	Project Recovery Sample (n = 919)
Mean age (SD), *y*	76.9 (5.0)	80.0 (6.5)
Male	134 (45)	365 (40)
Nonwhite race	21 (7)	119 (13)
Married	185 (62)	332 (36)
Living alone	85 (28)	371 (40)
Residing in nursing home	0 (0)	56 (6)
Mean education (SD), *y*	15 (2.9)	11.1 (3.5)
Mean Charlson comorbidity index score (SD)	1.0 (1.3)	2.9 (2.2)
Any impairment in basic activities of daily living†	21 (7)	320 (35)
Any impairment in instrumental activities of daily living†	73 (24)	799 (87)
Any cognitive impairment at baseline‡	24 (8)	406 (44)
Dementia at baseline§	5 (2)	121 (13)
Delirium during hospitalization	68 (23)	115 (13)

SAGES = Successful Aging After Elective Surgery.
* Data are numbers (percentages) unless otherwise indicated.
† Defined as any impairment in ≥1 activity.
‡ Defined as Modified Mini-Mental State score <85 in the SAGES study and Mini-Mental State Examination score <24 in Project Recovery.
§ In the SAGES study, although clinically documented dementia was excluded, early dementia was detected by a clinical consensus panel in 2% of patients after enrollment. In Project Recovery, dementia was defined as either a modified Blessed Dementia Rating Scale score >4 or a modified Blessed Dementia Rating Scale score >2, a Mini-Mental State Examination score <20, and a duration of cognitive symptoms of ≥6 mo.

The CAM-S showed moderate to high convergent agreement with other daily cognitive measures. In the SAGES study sample, the CAM-S agreement was high with the daily confusion rating (r = 0.78 and 0.80 with the short and long forms, respectively) and the brief cognitive screen (r = 0.62 and 0.72, respectively). In the Project Recovery sample, the CAM-S agreement was moderate to high with the daily confusion rating (r = 0.45 and 0.64, respectively) and the Mini-Mental State Examination score (r = 0.41 and 0.64, respectively). Similar results were obtained by using repeated random samples of 1 observation per patient and polynomial rather than linear models.

Table 2 shows the association of CAM-S score with hospital outcomes, with significant associations between increasing score and worse hospital outcomes (P for trend < 0.001 in all cases). Length of hospital stay increased across each short form severity category from an adjusted mean of 6.5 days for no delirium symptoms to 12.7 days with high severity. A similar gradient was seen for the long form, which increased across categories from an adjusted mean of 5.6 to 11.9 days. Hospital costs (in 1995 U.S. dollars) increased across CAM-S categories from

Figure. **Distribution of CAM-S scores, by delirium status in total sample and stratified by dementia status.**



The maximum score during each patient's hospitalization was used in all analyses. Boxes around the plots represent the median and 25th and 75th percentiles. If a box is not shown, the median and 25th or 75th percentiles had the same value. The stratified analyses by dementia group were conducted in the Project Recovery sample. SAGES = Successful Aging After Elective Surgery.

an adjusted mean of $5100 for no delirium symptoms to $13 200 for severe delirium. A similar gradient was seen for the long form, which increased from an adjusted mean of $4200 for no delirium symptoms to $11 400 for severe delirium. The adjusted RRs for new nursing home placement were 1.0, 1.4, 2.1, and 2.5 (P for trend < 0.001) across short-form categories and 1.0, 1.4, 2.3, and 3.9 (P for trend < 0.001) across long-form categories. The proportion of patients with functional decline between baseline and discharge increased across short-form (from 36% to 68%) and long-form (from 25% to 61%) categories. Cognitive decline between baseline and discharge also increased across severity categories (from 16% to 65% for the short form and from 10% to 50% for the long form).

The association of CAM-S score with posthospital outcomes is shown in **Table 3**, with significant associations between increasing score and worse posthospital outcomes (P for trend < 0.001 in all cases). The cumulative rates of death within 90 days increased across each short-form cat-

egory from 7% for no delirium symptoms to 27% with high severity (adjusted RR, 3.3). The gradient for the long form ranged from a cumulative death rate of 7% to 22% (adjusted RR, 2.3) from the lowest to highest categories. Finally, rates of death or nursing home residence at 90 days increased across each short-form category from 15% to 51% (adjusted RR, 2.5) and from 13% to 48% (adjusted RR, 2.5) for the long form. Among patients available for the 1-month follow-up interview, functional decline at 30 days increased across severity categories from 29% to 52% (adjusted RR, 1.6) for the short form and from 20% to 52% (adjusted RR, 2.3) for the long form.

Appendix Tables 1 and **2** (available at www.annals.org) show the results of the analyses from **Tables 2** and **3** stratified by delirium diagnosis. These analyses reveal the additional contribution of CAM-S score within strata defined by the presence or absence of delirium. In each subgroup in **Appendix Tables 1** and **2**, poorer clinical outcomes were seen with higher CAM-S scores. These trends are all statistically significant in the nondelirious subgroup. However, statistical significance was found for few outcomes in the delirium subgroups because of small sample sizes.

In **Appendix Figures 1** to **9**, the contribution of the CAM-S score beyond delirium diagnosis is suggested by the spread of points in each subgroup, indicating higher CAM-S scores with worse adverse outcomes regardless of delirium status.

Discussion

This study provides evidence for the usefulness and validity of a new delirium severity measure, the CAM-S. Because this measure is based on the CAM, which is already widely used, it poses distinct advantages for future work. This study shows that the CAM-S has good psychometric properties; high interrater reliability; and, most important, strong association with clinical outcomes related to delirium. The CAM-S short and long forms show strong psychometric properties (**Appendix**), and each has unique advantages. The short form (5-minute completion time), which is based on the CAM diagnostic algorithm alone, is quicker and simpler to rate; however, the long form (10-minute completion time) provides a broader range of severity scores in delirium and nondelirium groups.

Strengths of this study include the rigorous prospective validation of the CAM-S in 2 independent samples. Although the many differences between the study samples might be viewed as a limitation, their disparate nature lends strong support for the potential generalizability of the CAM-S to various populations and conditions. The use of state-of-the-art methods for delirium assessment, high-quality data with relatively few missing values, and the broad range of clinical outcomes for comparison enhanced the validation process. The rich clinical database with long-term follow-up enabled detailed assessment of the effect of the CAM-S on important clinical outcomes.

Table 2. **Association of CAM-S Score With Hospital Outcomes***

Outcome	Adjusted Mean Length of Stay (95% CI), *d* (*n* = 919)	Adjusted Mean Hospital Costs (95% CI), *thousand $* (*n* = 919)	New Nursing Home Placement (*n* = 851)†		Functional Decline (*n* = 908)‡		Cognitive Decline (*n* = 902)§	
			Patients, *n/N (%)*	Adjusted RR (95% CI)	Patients, *n/N (%)*	Adjusted RR (95% CI)	Patients, *n/N (%)*	Adjusted RR (95% CI)
CAM-S short-form severity rating‖								
None (*n* = 598)	6.5 (6.2–6.9)	5.1 (4.8–5.5)	61/571 (11)	Referent	212/589 (36)	Referent	93/589 (16)	Referent
Low (*n* = 91)	8.8 (7.6–9.9)	7.0 (5.8–8.2)	14/85 (16)	1.4 (0.8–2.3)	46/90 (51)	1.4 (1.1–1.8)	23/89 (26)	1.6 (1.1–2.4)
Moderate (*n* = 128)	11.1 (9.9–12.3)	9.7 (8.3–11.0)	29/112 (26)	2.1 (1.4–3.1)	80/128 (63)	1.7 (1.4–2.0)	34/126 (27)	1.6 (1.1–2.3)
High (*n* = 102)	12.7 (11.2–14.3)	13.2 (11.1–15.3)	32/83 (39)	2.5 (1.6–3.7)	69/101 (68)	1.8 (1.5–2.2)	64/98 (65)	3.9 (3.0–5.0)
P value for trend	<0.001	<0.001	–	<0.001	–	<0.001	–	<0.001
CAM-S long-form severity rating¶								
None (*n* = 205)	5.6 (5.1–6.1)	4.2 (3.7–4.7)	13/198 (7)	Referent	50/202 (25)	Referent	20/202 (10)	Referent
Low (*n* = 288)	6.8 (6.4–7.3)	5.4 (4.9–5.9)	25/278 (9)	1.4 (0.7–2.5)	115/284 (40)	1.6 (1.2–2.1)	53/282 (19)	1.9 (1.1–3.0)
Moderate (*n* = 234)	8.8 (8.1–9.5)	7.3 (6.5–8.1)	40/215 (19)	2.3 (1.3–4.2)	126/231 (55)	2.1 (1.6–2.8)	47/230 (20)	2.0 (1.2–3.3)
High (*n* = 192)	11.9 (10.8–12.9)	11.4 (10.0–12.8)	58/160 (36)	3.9 (2.1–7.0)	116/191 (61)	2.3 (1.8–3.0)	94/188 (50)	4.8 (3.0–7.7)
P value for trend	<0.001	<0.001	–	<0.001	–	<0.001	–	<0.001

RR = relative risk.

* Analyses were conducted in the Project Recovery sample. The maximum CAM-S score during each patient's hospitalization was used in all analyses. All models were adjusted for age, sex, race, Acute Physiology and Chronic Health Evaluation II score, Charlson comorbidity index score, and baseline dementia. All models, except the one for functional decline, were also adjusted for baseline impairment in activities of daily living.

† Analyses excluded 54 patients residing in a nursing home at baseline and 14 who died during hospitalization.

‡ Defined as a partial or complete decline in ≥1 activity on the standard 7-item Activities of Daily Living (ADL) scale between baseline and discharge. Analyses excluded 11 patients with missing ADL data at discharge.

§ Defined as a decrease of ≥2 points on the 30-point Mini-Mental State Examination score between baseline and discharge. Analyses excluded 17 patients with missing values at discharge.

‖ 0 (none), 1 (mild), 2 (moderate), or 3–7 (severe) points.

¶ 0–1 (none), 2 (mild), 3–4 (moderate), or 5–19 (severe) points.

Table 3. **Association of CAM-S Score With Posthospital Outcomes***

Outcome	Death Within 90 d (n = 919)†		Adjusted Mean Cost per Day for First 90 d (95% CI), $ (n = 831)‡	Death or Nursing Home Residence at 90 d (n = 844)‡		Functional Decline at 30 d (n = 712)§	
	Patients, *n/N (%)*	Adjusted RR (95% CI)		Patients, *n/N (%)*	Adjusted RR (95% CI)	Patients, *n/N (%)*	Adjusted RR (95% CI)
CAM-S short-form severity rating‖							
None (n = 598)	39/598 (7)	Referent	115.8 (100.2–131.3)	81/544 (15)	Referent	143/497 (29)	Referent
Low (n = 91)	14/91 (15)	2.0 (1.1–3.5)	158.4 (108.0–208.7)	27/82 (33)	1.9 (1.3–2.7)	29/70 (41)	1.4 (1.0–1.9)
Moderate (n = 128)	20/128 (16)	1.8 (1.1–3.2)	175.3 (129.8–220.8)	48/121 (40)	2.1 (1.5–2.9)	41/91 (45)	1.5 (1.2–2.0)
High (n = 102)	28/102 (27)	3.3 (2.1–5.1)	168.2 (116.4–220.0)	49/97 (51)	2.5 (1.9–3.3)	28/54 (52)	1.6 (1.2–2.2)
P value for trend	–	<0.001	<0.001	–	<0.001	–	<0.001
CAM-S long-form severity rating¶							
None (n = 205)	14/205 (7)	Referent	97.6 (76.0–119.2)	24/181 (13)	Referent	35/175 (20)	Referent
Low (n = 288)	20/288 (7)	0.8 (0.4–1.5)	126.7 (103.8–149.7)	37/263 (14)	1.0 (0.6–1.6)	67/238 (28)	1.4 (1.0–2.0)
Moderate (n = 234)	25/234 (11)	1.3 (0.7–2.4)	138.0 (110.2–165.7)	57/217 (26)	1.6 (1.0–2.5)	77/180 (43)	2.0 (1.4–2.8)
High (n = 192)	42/192 (22)	2.3 (1.3–4.1)	177.2 (138.2–216.3)	87/183 (48)	2.5 (1.6–3.7)	62/119 (52)	2.3 (1.6–3.3)
P value for trend	–	<0.001	<0.001	–	<0.001	–	<0.001

RR = relative risk.
* Analyses were conducted in the Project Recovery sample. The maximum CAM-S score during each patient's hospitalization was used in all analyses. All models were adjusted for age, sex, race, Acute Physiology and Chronic Health Evaluation II score, Charlson comorbidity index score, and baseline dementia. All models, except the one for functional decline, were also adjusted for baseline impairment in activities of daily living.
† Includes all in-hospital deaths.
‡ Medicare data were missing for 75 patients (including those receiving care in HMOs). The cost-per-day analyses also excluded 13 patients who died during hospitalization. See text for details.
§ Defined as a partial or complete decline in ≥1 activity on the standard 7-item Activities of Daily Living (ADL) scale between baseline and 30 d. These analyses included all 728 patients who were available for telephone follow-up interviews at 1 mo but excluded 16 with missing ADL data.
‖ 0 (none), 1 (mild), 2 (moderate), or 3–7 (severe) points.
¶ 0–1 (none), 2 (mild), 3–4 (moderate), or 5–19 (severe) points.

Several important caveats of this study are worth mentioning. First, the age of the Project Recovery data may be viewed as a limitation. The relatively long lengths of hospital stays and low hospital costs reflect hospital care from 1995 to 1998; however, we were primarily interested in the comparison of these outcomes among our severity groups. Thus, internal validity of the comparisons was paramount and the absolute values were less important. Second, patients in both cohorts were aged 70 years or older, and the performance of the CAM-S requires future validation in younger adults. Third, there may be inherent dependencies between CAM-S score and adverse outcomes. For example, patients with longer lengths of stay may have had higher CAM-S scores because of more opportunities for measurement. Finally, for accurate rating of the CAM-S, brief but formal cognitive testing of the patients should be conducted, which will require training and standardization of staff and some additional staff time.

Because the CAM is well-known and widely used, the CAM-S may pose distinct logistic advantages over existing delirium severity measures and fills an important gap in the applicability of the CAM. Unlike the Delirium Rating Scale (24, 25), a clinician rater is not required for the CAM-S. Instead, well-trained research assistants can reliably conduct the assessments. Compared with the Memorial Delirium Assessment Scale (26), the CAM-S short form is quicker and simpler to use. In addition, the short form has the important advantage of being relatively unbiased with respect to rating the severity of hyperactive and hypoactive forms of delirium because the features rated are present in both forms. All delirium severity measures to date have been limited by the overrepresentation of hyperactive compared with hypoactive features (that is, hyperactive features, such as agitation and hallucinations, contribute more to the total severity score than hypoactive features, such as psychomotor retardation). Although not directly examined in the present study, the severity of hypoactive delirium, which is the more common type among older persons, may not be captured adequately with these instruments. This imbalance has important implications for clinical trials targeting the management of delirium. If a treatment converts patients from hyperactive to hypoactive delirium, it may be rated as efficacious if the outcome measure does not adequately capture the severity of the hypoactive delirium. The results of many clinical trials for delirium have shown discrepant results (2, 7, 27), with many showing no benefits for (and even worsening of) clinical outcomes. This discrepancy may be a direct consequence of these measurement limitations. Delirium severity is a complex and multifaceted construct, and weighing the relative contributions of different symptom categories (for example, cognitive vs. behavioral and hyperactive vs. hypoactive) can create unique challenges. Separate severity ratings for these symptom subgroups may be needed. Finally, examining the relative and combined contributions of delirium severity, duration, and recurrence to outcomes is essential to better define the clinical impact of delirium. Although beyond the scope of the present study, these are important areas for future investigation.

The CAM-S provides a new scoring method with strong psychometric properties to add to the armamentarium of delirium measures. This measure may serve as a primary outcome for clinical trials and studies of the pathophysiology or prognosis of delirium. We hope that the availability of this measure will serve to facilitate critically needed studies of delirium and its outcomes and ultimately lead to improved quality of life for older persons and their families.

From Beth Israel Deaconess Medical Center, Harvard Medical School, and Hebrew SeniorLife, Boston, Massachusetts; University of Massachusetts Medical School, Worcester, Massachusetts; and Warren Alpert Medical School of Brown University, Providence, Rhode Island.

Note: The CAM algorithm and instrument are copyrighted to the Hospital Elder Life Program, LLC (www.hospitalelderlifeprogram.org). Instructions and training manual are available at the Web site.

Acknowledgment: The authors thank the patients, families, physicians, and research staff who participated in the SAGES and Project Recovery studies and made this study possible. This work is dedicated to the memory of Joshua Bryan Inouye Helfand and Bradley Yoshio Inouye.

Grant Support: By National Institute on Aging grants P01AG031720 and K07AG041835 (Dr. Inouye), R01AG030618 and K24AG035075 (Dr. Marcantonio), and K01AG033643 (Dr. Saczynski). Dr. Inouye holds the Milton and Shirley F. Levy Family Chair.

Disclosures: None. Forms can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M13-1927.

Reproducible Research Statement: *Study protocol, statistical code, and data set*: Available on request from Dr. Inouye (e-mail, AgingBrainCenter@hsl.harvard.edu).

Requests for Single Reprints: Sharon K. Inouye, MD, MPH, Aging Brain Center, Institute for Aging Research, Hebrew SeniorLife, 1200 Centre Street, Boston, MA 02459; e-mail, AgingBrainCenter@hsl.harvard.edu.

Current author addresses and author contributions are available at www.annals.org.

References

1. **Inouye SK, Westendorp RG, Saczynski JS.** Delirium in elderly people. Lancet. 2013. [PMID: 23992774]
2. **National Institute for Health and Care Excellence.** Delirium: diagnosis, prevention and management. NICE clinical guideline 103. London: National Institute for Health and Care Excellence; 2010. Accessed at www.nice.org.uk/nicemedia/live/13060/49909/49909.pdf on 5 February 2014.
3. **Inouye SK.** Delirium in older persons. N Engl J Med. 2006;354:1157-65. [PMID: 16540616]
4. **Witlox J, Eurelings LS, de Jonghe JF, Kalisvaart KJ, Eikelenboom P, van Gool WA.** Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: a meta-analysis. JAMA. 2010;304:443-51. [PMID: 20664045]
5. **Leslie DL, Marcantonio ER, Zhang Y, Leo-Summers L, Inouye SK.** One-year health care costs associated with delirium in the elderly population. Arch Intern Med. 2008;168:27-32. [PMID: 18195192]
6. **Wachter RM.** Understanding Patient Safety. 2nd ed. New York: McGraw-Hill Medical; 2012. Accessed at http://langetextbooks.com/0071765786/downloads/0071765786_Wachter.pdf on 5 February 2014.
7. **Greer N, Rossom R, Anderson P, MacDonald R, Tacklind J, Rutks I, et al.** Delirium: Screening, Prevention, and Diangosis–A Systematic Review of the Evidence. VA Evidence-based Synthesis Program Reports. 2011. [PMID: 22206108]
8. **Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI.** Clarifying confusion: the Confusion Assessment Method. A new method for detection of delirium. Ann Intern Med. 1990;113:941-8. [PMID: 2240918]
9. **Wei LA, Fearing MA, Sternberg EJ, Inouye SK.** The Confusion Assessment Method: a systematic review of current usage. J Am Geriatr Soc. 2008;56:823-30. [PMID: 18384586]
10. **Wong CL, Holroyd-Leduc J, Simel DL, Straus SE.** Does this patient have delirium?: value of bedside instruments. JAMA. 2010;304:779-86. [PMID: 20716741]
11. **Inouye SK, Bogardus ST Jr, Charpentier PA, Leo-Summers L, Acampora D, Holford TR, et al.** A multicomponent intervention to prevent delirium in hospitalized older patients. N Engl J Med. 1999;340:669-76. [PMID: 10053175]
12. **Milisen K, Foreman MD, Abraham IL, De Geest S, Godderis J, Vandermeulen E, et al.** A nurse-led interdisciplinary intervention program for delirium in elderly hip-fracture patients. J Am Geriatr Soc. 2001;49:523-32. [PMID: 11380743]
13. **Schmitt EM, Marcantonio ER, Alsop DC, Jones RN, Rogers SO Jr, Fong TG, et al; SAGES Study Group.** Novel risk markers and long-term outcomes of delirium: the Successful Aging After Elective Surgery (SAGES) study design and methods. J Am Med Dir Assoc. 2012;13:818.e1-10. [PMID: 22999782]
14. **Inouye SK, Leo-Summers L, Zhang Y, Bogardus ST Jr, Leslie DL, Agostini JV.** A chart-based method for identification of delirium: validation compared with interviewer ratings using the Confusion Assessment Method. J Am Geriatr Soc. 2005;53:312-8. [PMID: 15673358]
15. **Inouye SK.** The Confusion Assessment Method (CAM): Training Manual and Coding Guide. New Haven: Yale University School of Medicine; 2003. Accessed at www.hospitalelderlifeprogram.org/pdf/TheConfusionAssessmentMethodTrainingManual.pdf on 29 January 2014.
16. **Folstein MF, Folstein SE, McHugh PR.** "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12:189-98. [PMID: 1202204]
17. **Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW.** Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. JAMA. 1963;185:914-9. [PMID: 14044222]
18. **Inouye SK, Bogardus ST Jr, Baker DI, Leo-Summers L, Cooney LM Jr.** The Hospital Elder Life Program: a model of care to prevent cognitive and functional decline in older hospitalized patients. Hospital Elder Life Program. J Am Geriatr Soc. 2000;48:1697-706. [PMID: 11129764]
19. **Covinsky KE, Palmer RM, Fortinsky RH, Counsell SR, Stewart AL, Kresevic D, et al.** Loss of independence in activities of daily living in older adults hospitalized with medical illnesses: increased vulnerability with age. J Am Geriatr Soc. 2003;51:451-8. [PMID: 12657063]
20. **Tombaugh TN, McIntyre NJ.** The Mini-Mental State Examination: a comprehensive review. J Am Geriatr Soc. 1992;40:922-35. [PMID: 1512391]
21. **Leslie DL, Zhang Y, Bogardus ST, Holford TR, Leo-Summers LS, Inouye SK.** Consequences of preventing delirium in hospitalized older adults on nursing home costs. J Am Geriatr Soc. 2005;53:405-9. [PMID: 15743281]
22. **Manning WG, Mullahy J.** Estimating log models: to transform or not to transform? J Health Econ. 2001;20:461-94. [PMID: 11469231]
23. **Graubard BI, Korn EL.** Predictive margins with survey data. Biometrics. 1999;55:652-9. [PMID: 11318229]
24. **Trzepacz PT, Baker RW, Greenhouse J.** A symptom rating scale for delirium. Psychiatry Res. 1988;23:89-97. [PMID: 3363018]
25. **Trzepacz PT, Mittal D, Torres R, Kanary K, Norton J, Jimerson N.** Validation of the Delirium Rating Scale-Revised-98: comparison with the Delirium Rating Scale and the Cognitive Test for Delirium. J Neuropsychiatry Clin Neurosci. 2001;13:229-42. [PMID: 11449030]
26. **Breitbart W, Rosenfeld B, Roth A, Smith MJ, Cohen K, Passik S.** The Memorial Delirium Assessment Scale. J Pain Symptom Manage. 1997;13:128-37. [PMID: 9114631]
27. **Lacasse H, Perreault MM, Williamson DR.** Systematic review of antipsychotics for the treatment of hospital-associated delirium in medically or surgically ill patients. Ann Pharmacother. 2006;40:1966-73. [PMID: 17047137]

Current Author Addresses: Drs. Inouye and Schmitt, Mr. Kosar, and Ms. Puelle: Aging Brain Center, Institute for Aging Research, Hebrew SeniorLife, 1200 Centre Street, Boston, MA 02131.
Mr. Tommet: Butler Hospital, Duncan Building, 700 Butler Drive, Providence, RI 02912.
Dr. Saczynski: University of Massachusetts Medical School, Division of Geriatric Medicine, 377 Plantation Street, Biotech 4, Suite 315, Worcester, MA 01655.
Dr. Marcantonio: Harvard Medical School, Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215.
Dr. Jones: Butler Hospital, 345 Blackstone Boulevard, Providence, RI 02906.

Author Contributions: Conception and design: S.K. Inouye, E.R. Marcantonio, R.N. Jones.
Analysis and interpretation of the data: S.K. Inouye, C.M. Kosar, D. Tommet, E.R. Marcantonio, R.N. Jones.
Drafting of the article: S.K. Inouye.
Critical revision of the article for important intellectual content: S.K. Inouye, C.M. Kosar, E.M. Schmitt, M.R. Puelle, J.S. Saczynski, E.R. Marcantonio, R.N. Jones.
Final approval of the article: S.K. Inouye, C.M. Kosar, E.M. Schmitt, M.R. Puelle, J.S. Saczynski, E.R. Marcantonio, R.N. Jones.
Provision of study materials or patients: S.K. Inouye, E.R. Marcantonio.
Statistical expertise: S.K. Inouye, C.M. Kosar, D. Tommet, E.R. Marcantonio, R.N. Jones.
Obtaining of funding: S.K. Inouye, E.R. Marcantonio.
Administrative, technical, or logistic support: S.K. Inouye, C.M. Kosar, E.M. Schmitt, M.R. Puelle.
Collection and assembly of data: S.K. Inouye, C.M. Kosar, D. Tommet, E.M. Schmitt, E.R. Marcantonio.

Web-Only References

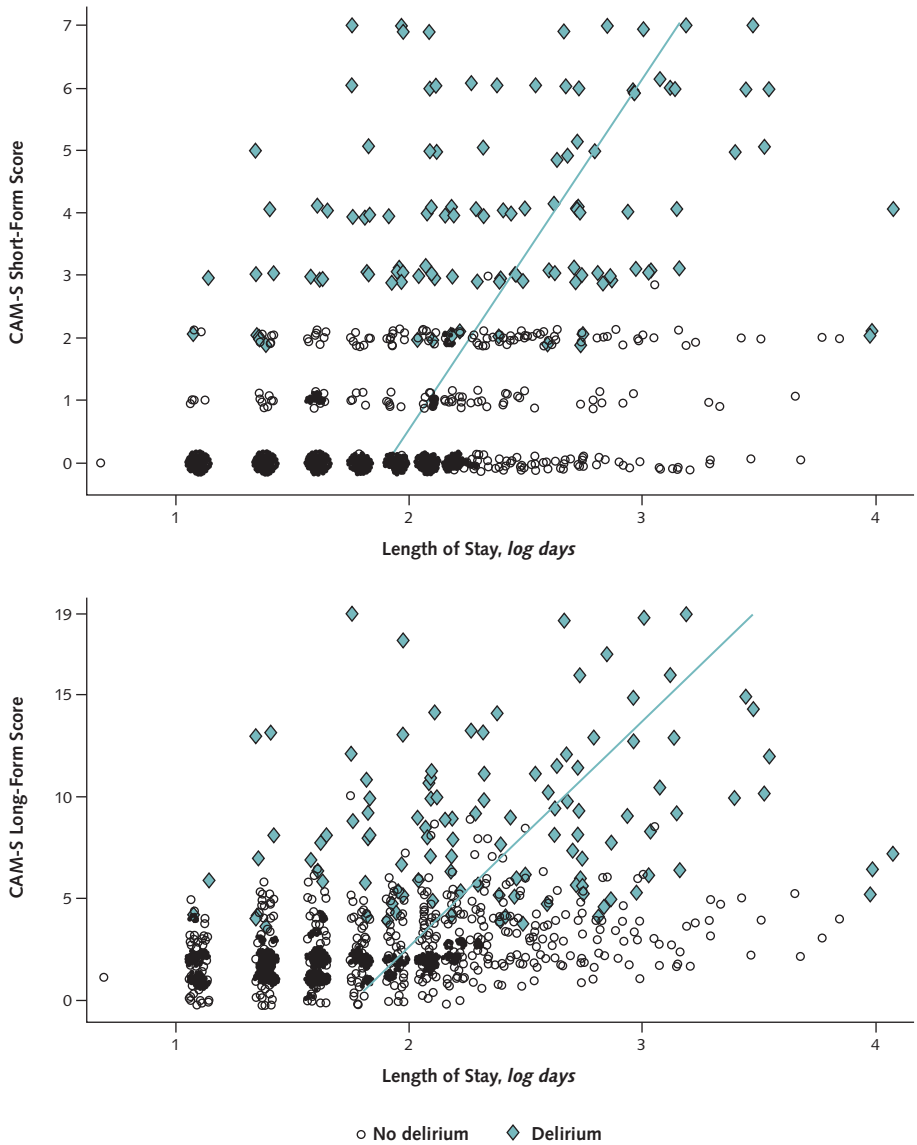
28. **Buja A, Eyuboglu N.** Remarks on parallel analysis. Multivariate Behav. 1992; 27:509-40.
29. **McDonald RP.** Test Theory: A Unified Treatment. Mahway, NJ: Psychology Pr; 1999.

Appendix: Validation Analyses: Factor Analysis

Unidimensionality and model fit were assessed with permuted parallel analysis (28) and confirmatory factor analysis in the SAGES study sample. Internal reliability was assessed with the Cronbach α and McDonald ω statistics (29).

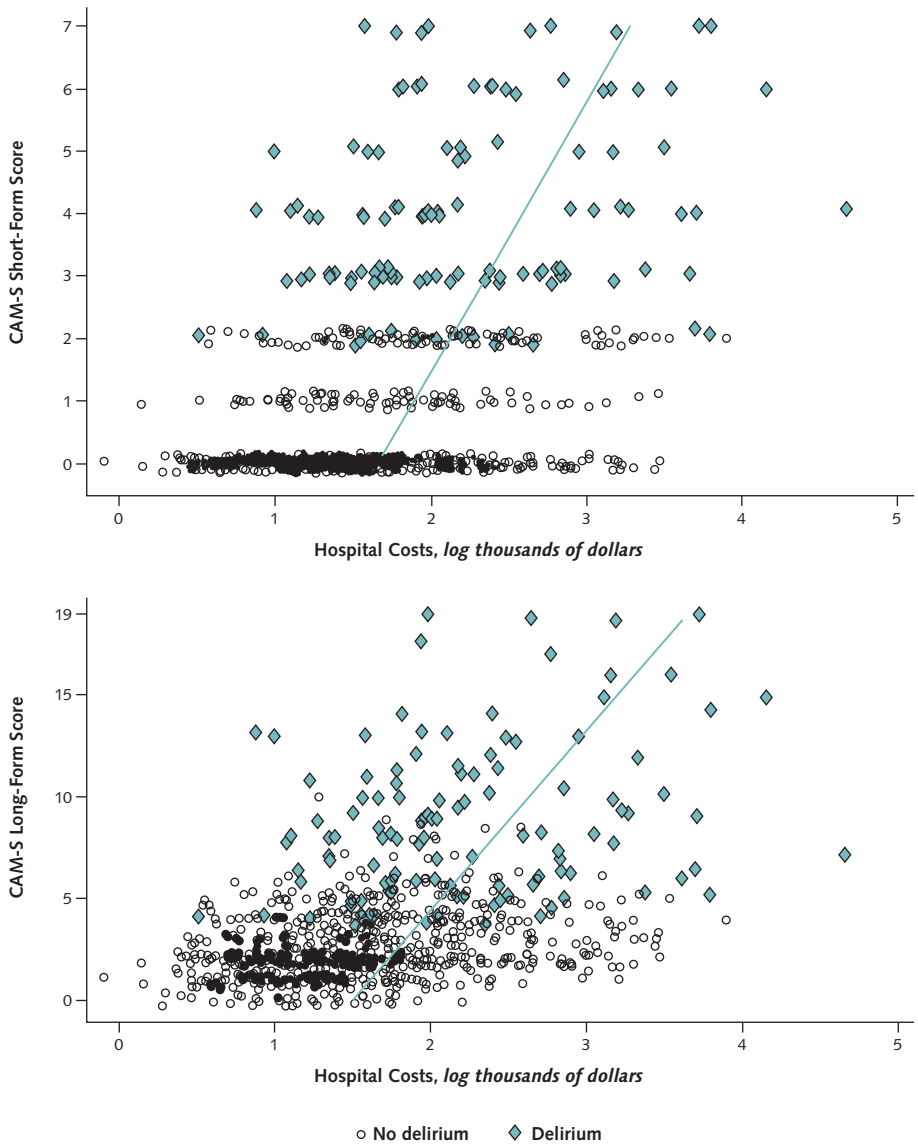
Parallel and confirmatory factor analyses show that the CAM-S is a unidimensional measure. The single-factor model shows good fit (confirmatory fit index, 0.99; root mean squared error of approximation, 0.04). Good internal reliability of the CAM-S is shown by the Cronbach α (0.73 for the short and long forms) and McDonald ω (0.92 for the short form and 0.90 for the long form).

Appendix Figure 1. **CAM-S scores, by length of stay.**



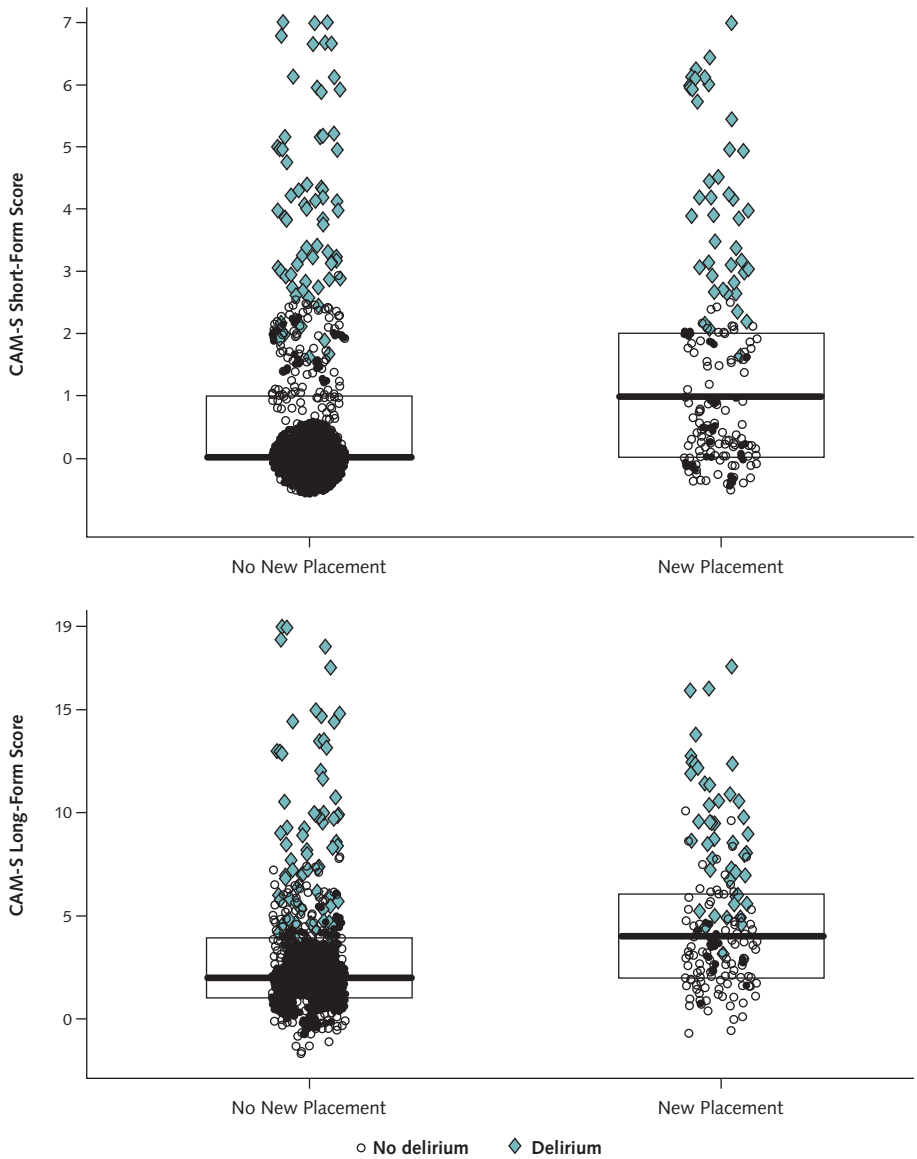
Plots are of maximum scores per patient by length of hospital stay. The green line runs through fitted values derived from log-γ regression.

Appendix Figure 2. **CAM-S scores, by hospital costs.**



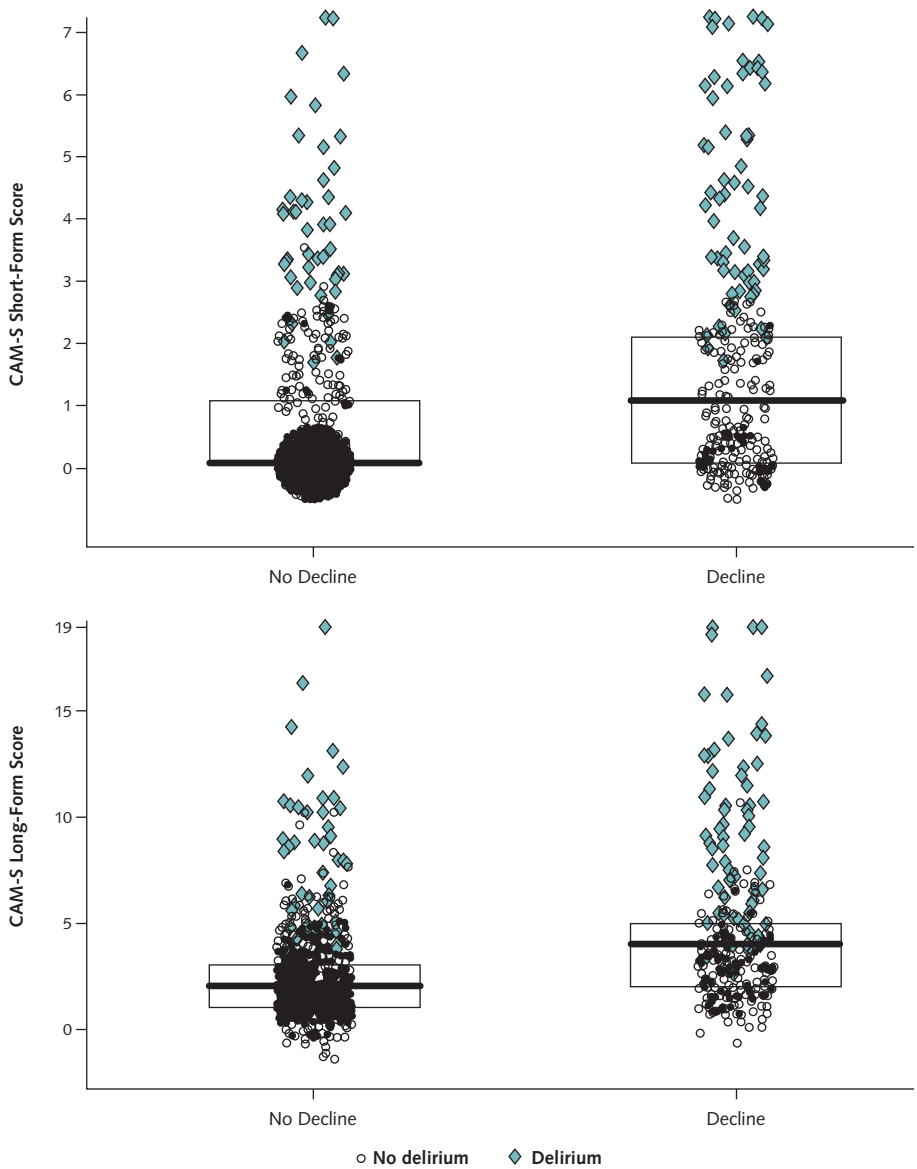
Plots are of maximum scores per patient by hospital costs. The green line runs through fitted values derived from log-γ regression.

Appendix Figure 3. **CAM-S scores, by new nursing home placement.**



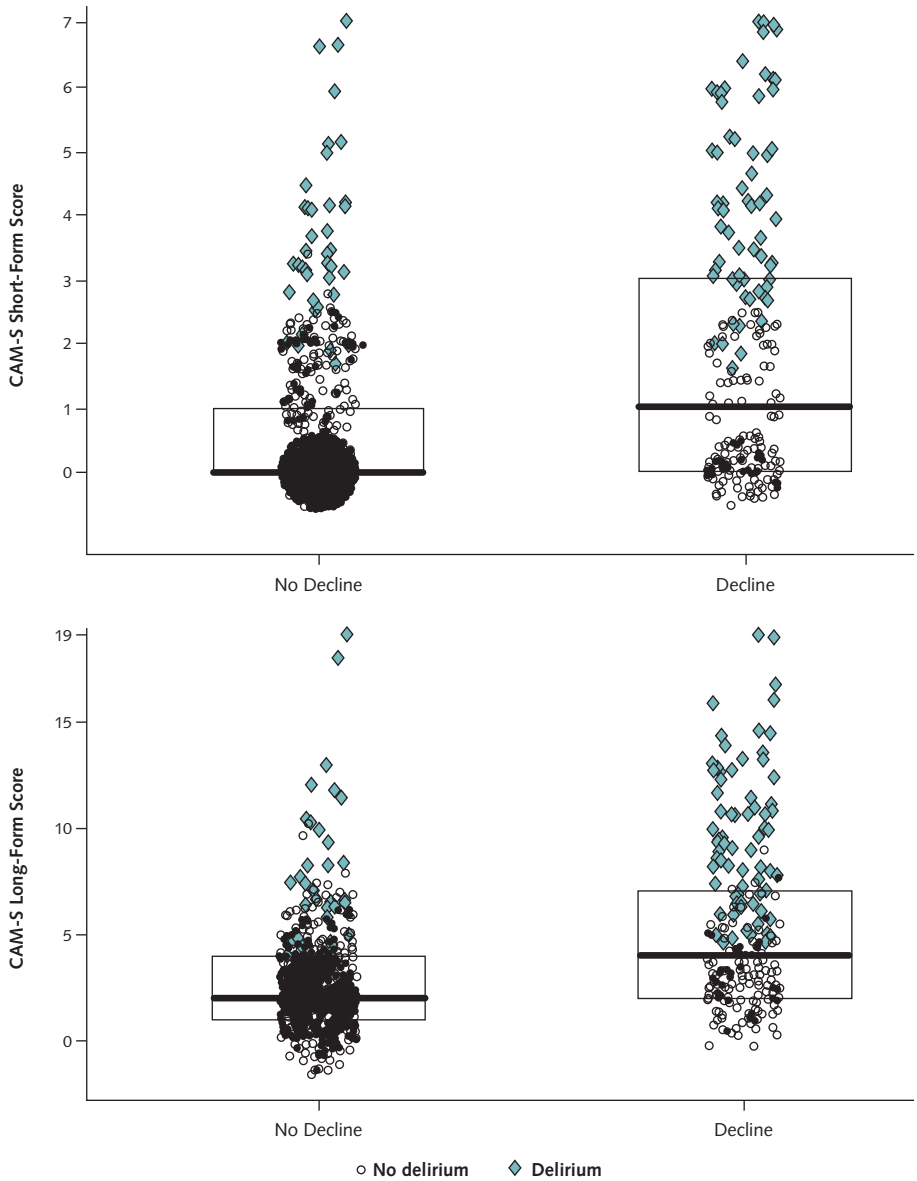
The maximum score during each patient's hospitalization was used in all analyses. Boxes around the plots represent the median and 25th and 75th percentiles. If a box is not shown, the median and 25th or 75th percentiles had the same value.

Appendix Figure 4. **CAM-S scores, by functional decline.**



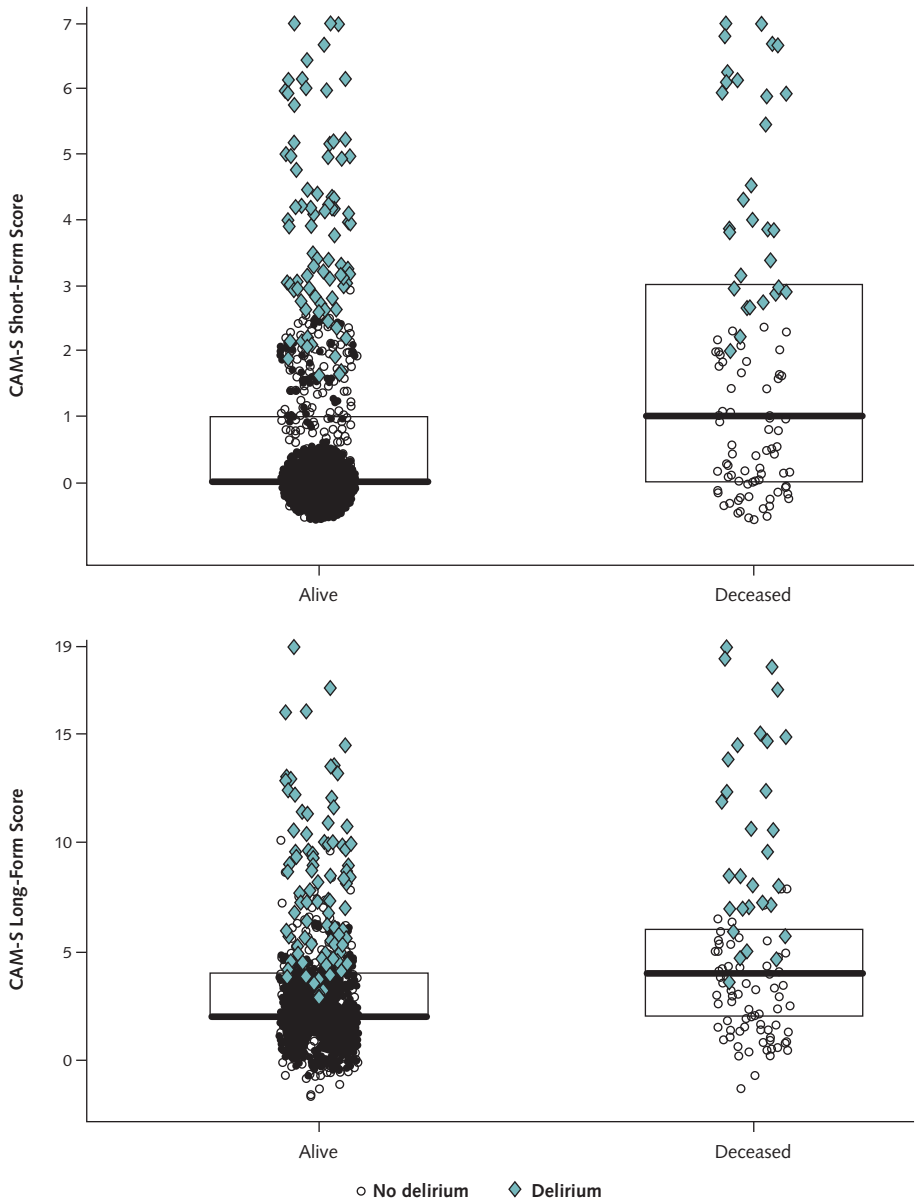
The maximum score during each patient's hospitalization was used in all analyses. Functional decline was defined as a partial or complete decline in ≥1 activity on the standard 7-item Activities of Daily Living scale between baseline and discharge (see text for details). Boxes around the plots represent the median and 25th and 75th percentiles. If a box is not shown, the median and 25th or 75th percentiles had the same value.

Appendix Figure 5. **CAM-S scores, by cognitive decline.**



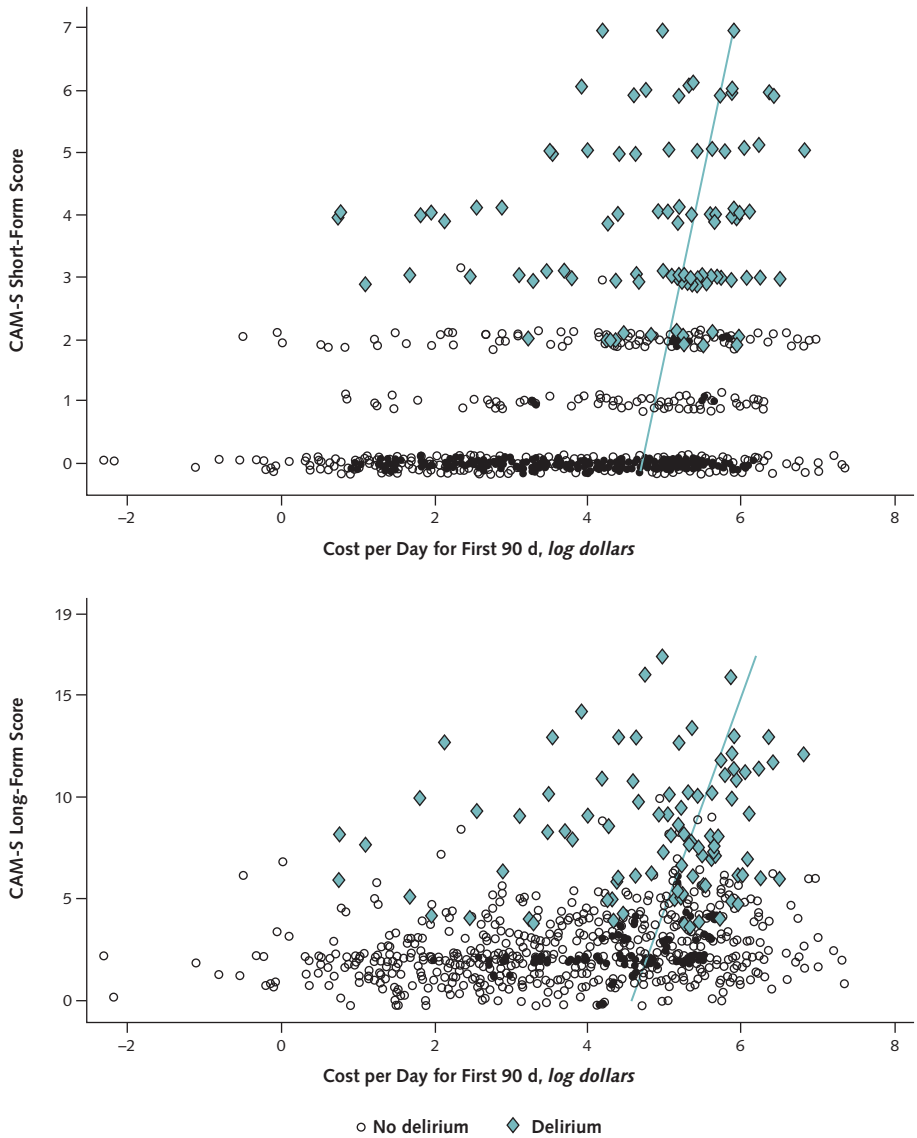
The maximum score during each patient's hospitalization was used in all analyses. Cognitive decline was defined as a decrease of ≥2 points in the Mini-Mental State Examination score between baseline and discharge. Boxes around the plots represent the median and 25th and 75th percentiles. If a box is not shown, the median and 25th or 75th percentiles had the same value.

Appendix Figure 6. **CAM-S scores, by death within 90 d.**



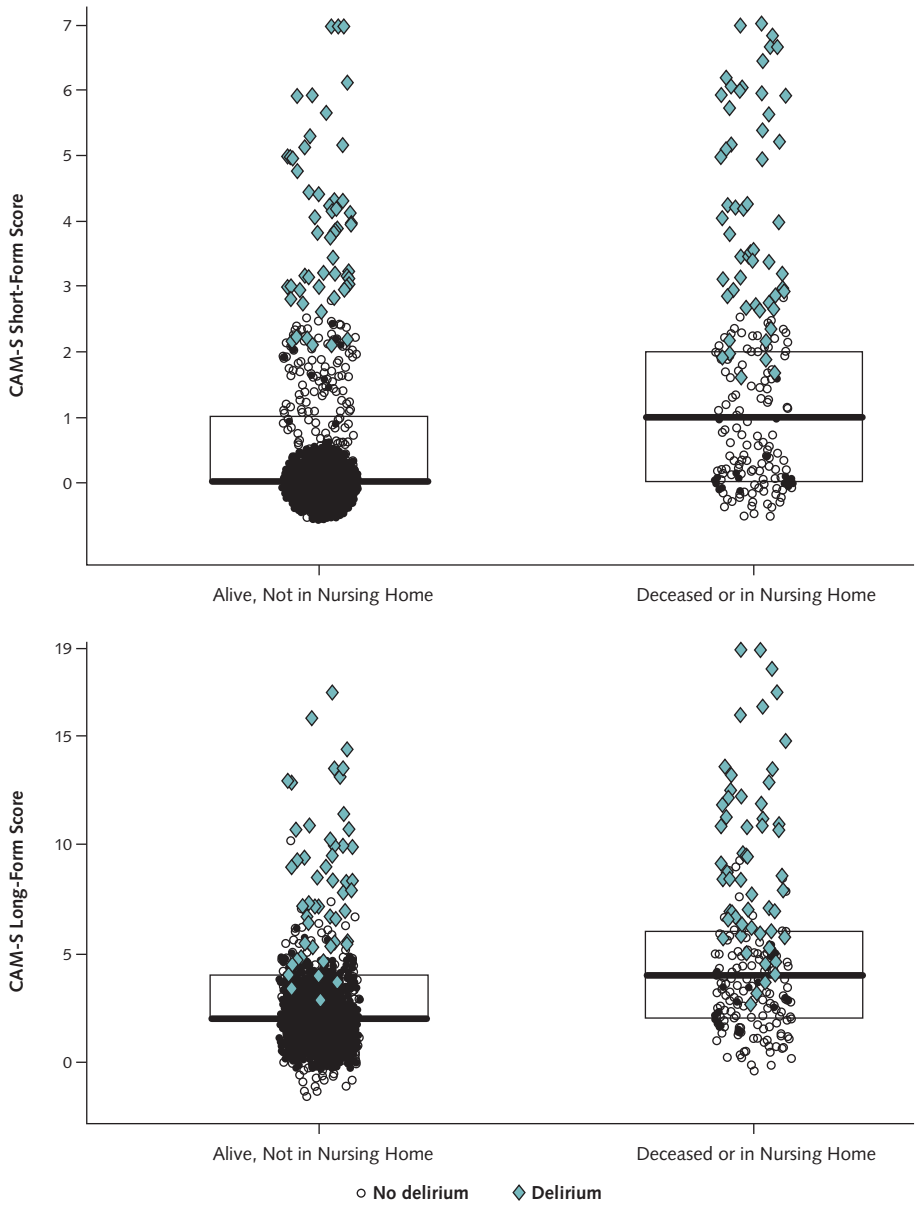
The maximum score during each patient's hospitalization was used in all analyses. Boxes around the plots represent the median and 25th and 75th percentiles. If a box is not shown, the median and 25th or 75th percentiles had the same value.

Appendix Figure 7. **CAM-S scores, by cost per day for the first 90 d.**



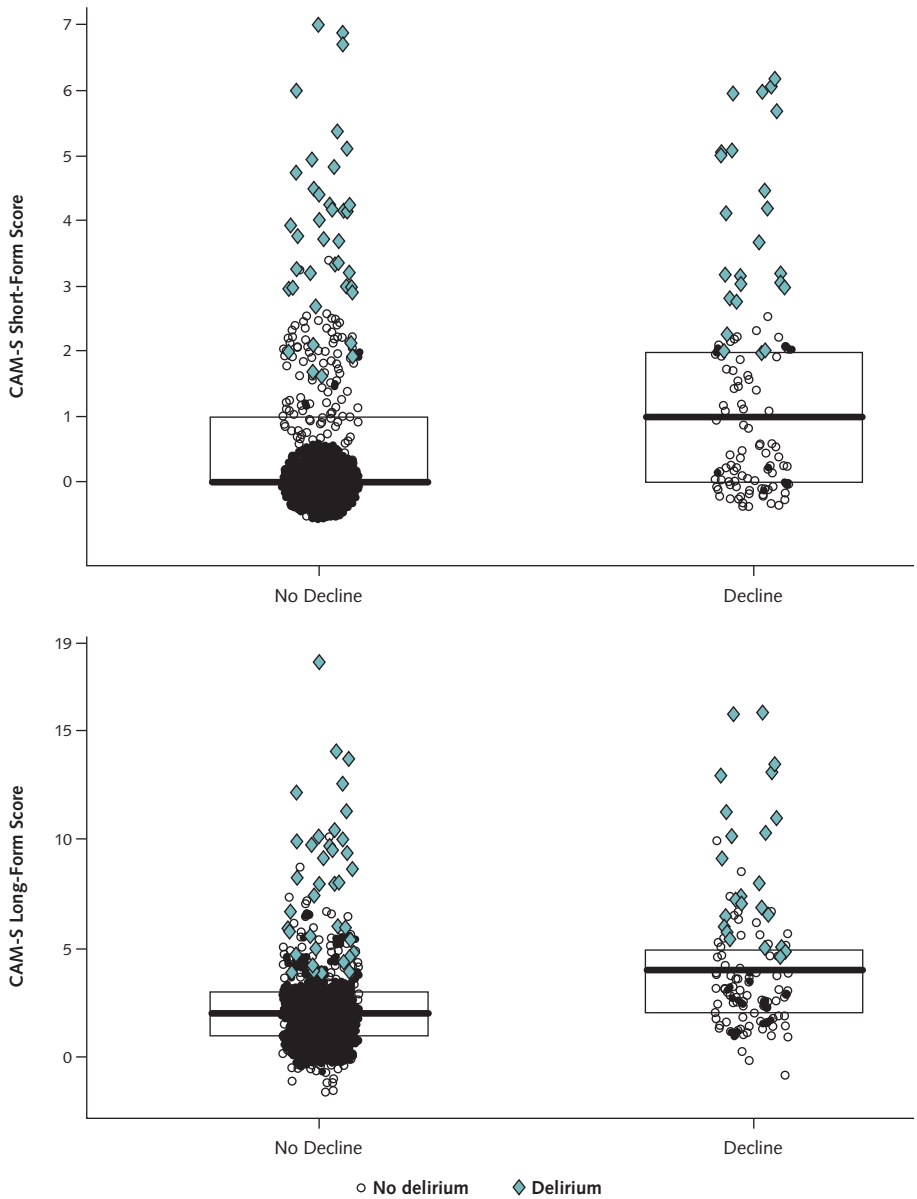
Plots are of maximum scores per patient by cost per day. The green line runs through fitted values derived from log-γ regression. For 14 patients, the observed values were <0 log dollars. The values were 0.10 to 0.98 dollars per day (−2.30 to −0.018 log dollars per day).

Appendix Figure 8. **CAM-S scores, by death or nursing home residence at 90 d.**



The maximum score during each patient's hospitalization was used in all analyses. Boxes around the plots represent the median and 25th and 75th percentiles. If a box is not shown, the median and 25th or 75th percentiles had the same value.

Appendix Figure 9. **CAM-S scores, by functional decline at 30 d.**



The maximum score during each patient's hospitalization was used in all analyses. Functional decline was defined as a partial or complete decline in ≥1 activity on the standard 7-item Activities of Daily Living scale between baseline and 1-mo follow-up (see text for details). Boxes around the plots represent the median and 25th and 75th percentiles. If a box is not shown, the median and 25th or 75th percentiles had the same value.

Appendix Table 1. **Association of CAM-S Score With Hospital Outcomes, Stratified by Delirium Diagnosis***

Outcome	Adjusted Mean Length of Stay (95% CI), *d* (*n* = 919)†	Adjusted Mean Hospital Costs (95% CI), *thousand $* (*n* = 919)†	New Nursing Home Placement (*n* = 851)‡		Functional Decline (*n* = 908)§		Cognitive Decline (*n* = 902)‖	
			Patients, *n/N (%)*	Adjusted Mean Probability (95% CI)¶	Patients, *n/N (%)*	Adjusted Mean Probability (95% CI)¶	Patients, *n/N (%)*	Adjusted Mean Probability (95% CI)¶
CAM-S short-form severity rating								
Nondelirious group								
0 points (*n* = 598)	6.5 (6.2–6.8)	5.1 (4.8–5.4)	61/571 (11)	10.6 (8.3–13.0)	212/589 (36)	36.3 (32.6–40.1)	93/589 (16)	16.2 (13.3–19.1)
1 point (*n* = 91)	8.4 (8.0–8.9)	6.8 (6.3–7.3)	14/85 (16)	17.2 (13.7–20.6)	46/90 (51)	48.2 (43.6–52.7)	23/89 (26)	21.1 (17.3–24.8)
2–3 points (*n* = 115)	11.3 (10.1–12.4)	9.6 (8.3–10.8)	26/101 (26)	25.4 (18.1–32.6)	69/115 (60)	60.6 (52.5–68.6)	27/115 (23)	25.1 (17.8–32.4)
P value for trend	<0.001	<0.001	–	<0.001	–	<0.001	–	0.020
Delirious group								
2 points (*n* = 15)**	11.8 (9.0–14.6)	10.3 (7.0–13.6)	3/12 (25)	26.3 (11.2–41.5)	11/15 (73)	62.6 (46.3–79.0)	7/13 (54)	50.9 (34.2–67.6)
3–4 points (*n* = 64)	12.4 (10.6–14.2)	12.0 (9.7–14.4)	21/57 (37)	35.4 (25.9–44.8)	38/63 (60)	65.9 (56.5–75.3)	36/61 (59)	59.9 (50.0–69.8)
5–7 points (*n* = 36)	15.1 (11.8–18.4)	16.4 (11.5–21.3)	11/25 (44)	46.7 (30.1–63.4)	31/36 (86)	80.8 (69.5–92.0)	28/35 (80)	79.6 (68.0–91.2)
P value for trend	0.26	0.033	–	0.22	–	0.084	–	0.024
CAM-S long-form severity rating								
Nondelirious group								
0–1 point (*n* = 205)	5.8 (5.4–6.2)	4.5 (4.2–4.9)	13/198 (7)	6.3 (4.2–8.4)	50/202 (25)	30.2 (25.6–34.8)	20/202 (10)	13.5 (10.2–16.9)
2 points (*n* = 288)	6.9 (6.6–7.2)	5.6 (5.2–5.9)	25/278 (9)	9.4 (7.3–11.6)	115/284 (40)	38.2 (34.6–41.8)	53/282 (19)	16.4 (13.6–19.2)
3–10 points (*n* = 311)	8.9 (8.3–9.5)	7.2 (6.6–7.9)	63/281 (22)	22.2 (18.2–26.2)	162/308 (53)	51.1 (46.3–56.0)	70/309 (23)	22.4 (18.3–26.6)
P value for trend	<0.001	<0.001	–	<0.001	–	<0.001	–	<0.001
Delirious group								
4–6 points (*n* = 41)	11.7 (9.5–13.9)	10.8 (8.0–13.6)	9/35 (26)	27.0 (15.0–39.0)	27/40 (68)	67.0 (54.8–79.2)	18/38 (47)	50.8 (37.4–64.1)
7–10 points (*n* = 42)	13.0 (11.2–14.8)	12.6 (10.2–15.0)	16/38 (42)	42.7 (33.3–52.2)	27/42 (64)	66.7 (57.6–75.9)	29/40 (73)	66.6 (57.7–75.6)
11–19 points (*n* = 32)	15.3 (11.7–18.8)	17.2 (11.6–22.8)	10/21 (48)	44.4 (27.3–61.6)	26/32 (81)	78.7 (67.0–90.3)	24/31 (77)	80.8 (69.5–92.1)
P value for trend	0.29	<0.047	–	0.134	–	0.189	–	0.005

* Analyses were conducted in the Project Recovery sample. The maximum CAM-S score during each patient's hospitalization was used in all analyses. All models were adjusted for age, sex, race, Acute Physiology and Chronic Health Evaluation II score, Charlson comorbidity index score, and baseline dementia. All models, except the one for functional decline, were also adjusted for baseline impairment in activities of daily living.

† Analyses included 804 and 115 participants in the nondelirious and delirious groups, respectively. Predictive margins were obtained from log-γ regression.

‡ Analyses included 757 and 94 participants in the nondelirious and delirious groups, respectively. Fifty-four patients residing in a nursing home at baseline and 14 who died during hospitalization were excluded.

§ Defined as a partial or complete decline in ≥1 activity on the standard 7-item Activities of Daily Living (ADL) scale between baseline and discharge. Eleven patients with missing ADL data at discharge were excluded. Analyses included 794 and 114 participants in the nondelirious and delirious groups, respectively.

‖ Defined as a decrease of ≥2 points on the 30-point Mini-Mental State Examination score between baseline and discharge. Seventeen patients with missing values at discharge were excluded. Analyses included 793 and 109 participants in the nondelirious and delirious groups, respectively.

¶ Predictive margins obtained from logistic regression.

** Based on clinical adjudication and chart review. Sensitivity analyses that excluded this group revealed similar results.

Appendix Table 2. **Association of CAM-S Score With Posthospital Outcomes, Stratified by Delirium Diagnosis***

Outcome	Death Within 90 d (n = 919)†		Adjusted Mean Cost per Day for First 90 d (95% CI), $ (n = 831)‡§	Nursing Home Residence or Death at 90 d (n = 844)‡‖		Functional Decline at 30 d (n = 712)¶	
	Patients, *n/N (%)*	Adjusted Mean Probability (95% CI)**		Patients, *n/N (%)*	Adjusted Mean Probability (95% CI)**	Patients, *n/N (%)*	Adjusted Mean Probability (95% CI)**
CAM-S short-form severity rating							
Nondelirious group							
0 points (n = 598)	39/598 (7)	6.9 (5.0–8.8)	109.8 (95.6–124.1)	81/544 (15)	15.3 (12.4–18.2)	143/497 (29)	29.1 (25.3–32.9)
1 point (n = 91)	14/91 (15)	11.4 (8.9–14.0)	151.0 (126.6–175.4)	27/82 (33)	27.4 (23.2–31.6)	29/70 (41)	38.1 (33.2–43.0)
2–3 points (n = 115)	18/115 (16)	17.0 (11.0–23.0)	204.2 (147.3–261.0)	41/109 (38)	39.6 (31.5–47.7)	36/83 (43)	44.1 (34.6–53.6)
P value for trend	–	<0.001	<0.001	–	<0.001	–	0.004
Delirious group							
2 points (n = 15)††	2/15 (13)	14.5 (3.6–25.3)	174.2 (112.4–235.9)	8/14 (57)	41.1 (27.0–55.1)	5/10 (50)	46.6 (25.2–67.9)
3–4 points (n = 64)	15/64 (23)	21.9 (14.1–29.8)	182.3 (148.0–216.6)	27/61 (44)	49.8 (41.0–58.6)	17/35 (49)	49.8 (36.7–62.8)
5–7 points (n = 36)	13/36 (36)	38.3 (24.3–52.3)	226.8 (160.6–292.9)	21/34 (62)	58.5 (44.6–72.3)	11/17 (65)	64.3 (44.9–83.7)
P value for trend	–	0.023	0.21	–	0.32	–	0.32
CAM-S long-form severity rating							
Nondelirious group							
0–1 point (n = 205)	14/205 (7)	5.5 (3.5–7.4)	91.5 (76.1–107.0)	24/181 (13)	11.2 (8.2–14.2)	35/175 (20)	22.8 (18.2–27.3)
2 points (n = 288)	20/288 (7)	8.3 (6.3–10.4)	115.1 (100.6–129.7)	37/263 (14)	16.2 (13.3–19.1)	67/238 (28)	27.7 (24.0–31.4)
3–10 points (n = 311)	37/311 (12)	11.5 (8.5–14.5)	164.7 (133.6–195.7)	88/291 (30)	29.6 (25.1–34.1)	106/237 (45)	43.1 (37.7–48.6)
P value for trend	–	0.003	<0.001	–	<0.001	–	<0.001
Delirious group							
4–6 points (n = 41)	8/41 (20)	16.2 (7.4–24.9)	161.7 (118.8–204.6)	16/38 (42)	37.9 (26.6–49.1)	12/26 (46)	46.9 (30.1–63.6)
7–10 points (n = 42)	8/42 (19)	23.8 (16.1–31.5)	208.7 (166.3–251.1)	20/40 (50)	56.3 (48.2–64.5)	12/21 (57)	53.1 (40.4–65.8)
11–19 points (n = 32)	14/32 (44)	41.8 (27.5–56.1)	216.1 (151.3–281.0)	20/31 (65)	61.5 (47.7–75.4)	9/15 (60)	64.4 (44.8–84.1)
P value for trend	–	0.007	0.182	–	0.058	–	0.22

* Analyses were conducted in the Project Recovery sample. The maximum CAM-S score during each patient's hospitalization was used in all analyses. All models were adjusted for age, sex, race, Acute Physiology and Chronic Health Evaluation II score, Charlson comorbidity index score, and baseline dementia. All models, except the one for functional decline, were also adjusted for baseline impairment in activities of daily living.

† Analyses included 804 and 115 participants in the nondelirious and delirious groups, respectively. Deaths within 90 d include all in-hospital deaths.

‡ Medicare data were missing for 75 patients (including those receiving care in HMOs). The cost-per-day analyses also excluded 13 patients who died during hospitalization. See text for details.

§ Analyses included 732 and 99 participants in the nondelirious and delirious groups, respectively. Predictive margins were obtained from log-γ regression.

‖ Analyses included 735 and 109 participants in the nondelirious and delirious groups, respectively.

¶ Among those who were available for telephone follow-up interviews at 1 mo (n = 712), functional decline was defined as a partial or complete decline in ≥1 activity on the standard 7-item Activities of Daily Living scale between baseline and 30 d. Analyses included 650 and 62 participants in the nondelirious and delirious groups, respectively.

** Predictive margins obtained from logistic regression.

†† Based on clinical adjudication and chart review. Sensitivity analyses that excluded this group revealed similar results.