



Sensitivity and specificity of the animal fluency test for predicting postoperative delirium

Sensibilité et spécificité du test de fluidité verbale sur les animaux pour la prédiction du délirium postopératoire

Lawrence S. Long, MD · Jed T. Wolpaw, MD ·
Jacqueline M. Leung, MD

Received: 30 May 2014 / Accepted: 16 December 2014 / Published online: 24 December 2014
© Canadian Anesthesiologists' Society 2014

Abstract

Background Preoperative cognitive impairment is a major risk factor for postoperative delirium. We therefore investigated the prognostic significance and feasibility of administering a brief cognitive screen before surgery.

Methods Patients > 65 yr of age undergoing hip, knee, or spine surgery were enrolled. A 60-sec cognitive screen, the animal fluency test (AFT), was administered preoperatively. Postoperative delirium was measured using a chart-based tool previously validated using criteria from the Confusion Assessment Method.

Results Of the 362 patients satisfying the inclusion/exclusion criteria, 199 (55%) underwent the AFT. Among them, 57 patients (29%) had an AFT score < 15, and 38 patients (19%, 95% confidence interval [CI]: 14 to 25%) developed postoperative delirium as measured by chart review. Patients with scores of < 15 were more likely to develop postoperative delirium than those who scored ≥ 15 (54% vs 5%, $P < 0.01$). A multiple logistic

regression, with postoperative delirium as the dependent variable, identified an AFT score of < 15 (odds ratio 20.1, 95% CI: 7.9 to 51.4) and high American Society of Anesthesiologists classification (odds ratio 3.5, 95% CI: 1.3 to 9.2) as independent predictors.

Conclusions The AFT is a potentially useful brief cognitive screen for identifying patients at risk of developing postoperative delirium. Limited participation by eligible participants in this study, however, raises questions about how useful and feasible systematic administration of the test is. Large studies using prospective measurement of postoperative delirium are indicated to validate our results.

Résumé

Contexte Le trouble cognitif préopératoire est un facteur de risque majeur de délirium postopératoire. Nous avons donc étudié la signification pronostique et la faisabilité de l'administration d'un bref test cognitif de dépistage avant une intervention chirurgicale.

Méthodes Des patients âgés de plus de 65 ans subissant une chirurgie de hanche, genou ou colonne vertébrale ont été recrutés. Un test cognitif de dépistage de 60 secondes, le test de fluidité verbale des animaux (AFT: Animal Fluency Test), a été administré en préopératoire. Le délirium postopératoire a été mesuré à l'aide d'un outil déjà validé reposant sur les dossiers et utilisant des critères issus de la méthode d'évaluation de la confusion.

Résultats Sur les 362 patients répondant aux critères d'inclusion/exclusion, 199 (55 %) ont passé l'AFT. Parmi eux, 57 patients (29 %) avaient un score AFT < 15, et 38 patients (19 %, intervalle de confiance [IC] à 95 %: 14 à 25 %) ont développé un délirium postopératoire tel que mesuré par l'analyse du dossier. Les patients ayant des scores < 15 ont été plus susceptibles de développer un

Author contributions Lawrence S. Long contributed to the conception and design, acquisition of data, analysis, and interpretation of data; drafting the article or revising it critically for important intellectual content; and giving final approval of the version to be published. Jed T. Wolpaw contributed to conception and design, acquisition of data, and interpretation of data; drafting the article or revising it critically for important intellectual content; and giving final approval of the version to be published. Jacqueline M. Leung contributed to conception and design, analysis, and interpretation of data; drafting the article or revising it critically for important intellectual content; and giving final approval of the version to be published.

L. S. Long, MD (✉) · J. T. Wolpaw, MD · J. M. Leung, MD
Department of Anesthesia and Perioperative Care, University of California San Francisco, 521 Parnassus Ave, C-455,
San Francisco, CA 94143-0648, USA
e-mail: longsl@anesthesia.ucsf.edu

délirium postopératoire que ceux ayant un score ≥ 15 (54 % contre 5 %, $P < 0,01$). Une analyse en régression logistique multifactorielle utilisant le délirium postopératoire comme variable dépendante a identifié qu'un score AFT < 15 (rapport de cotes: 20,1; IC à 95 %: 7,9 à 51,4) et une cote élevée dans la classification de l'American Society of Anesthesiologists (rapport de cotes: 3,5; IC à 95 %: 1,3 à 9,2) étaient des éléments prédictifs indépendants.

Conclusions L'AFT est un bref test cognitif de dépistage potentiellement utile pour l'identification des patients à risque de développer un délirium postopératoire. Toutefois, dans cette étude, la participation limitée des patients admissibles soulève des questions sur l'utilité et la faisabilité de l'administration systématique de ce test. De grandes études utilisant des mesures prospectives du délirium postopératoire sont indiquées pour valider nos résultats.

Older surgical patients with preexisting cognitive impairment are at increased risk for developing postoperative delirium.^{1,2} Screening surgical patients for preexisting cognitive impairment, however, has yet to become routine clinical practice, in part because many cognitive screening tests are time-intensive and require special training to administer.

Although abbreviated cognitive screening tests exist, few have been studied in the preoperative setting.³ One of these brief cognitive screens is the animal fluency test (AFT). It requires patients to name as many animals as possible within 60 sec, with one point given for each unique animal. Scores < 15 points have demonstrated high sensitivity and specificity (88% and 96%, respectively) for the presence of Alzheimer dementia in the memory clinic setting.⁴ In addition to its brevity, the AFT has the advantage of requiring limited training to administer it correctly.

We designed a study to investigate the feasibility of administering the AFT before surgery and its effectiveness for predicting postoperative delirium. We hypothesized that patients who scored < 15 on the AFT immediately before surgery would be more likely to develop postoperative delirium than patients who scored ≥ 15 .

Methods

The University of California, San Francisco Human Research Protection Program, Committee on Human Research approved the study in February 2013, and the requirement for written informed consent was waived. The study was conducted between July 1 and December 31, 2012 at a single, large university hospital. Inclusion criteria

were consecutive patients > 65 yr of age who underwent hip, knee, or spine surgery and who required a postoperative stay of more than 24 hr. We chose patients undergoing these types of surgery because prior research has shown that they are typically older and have a high incidence of postoperative delirium.⁵ Exclusion criteria were non-English speakers, hearing or speaking impairment (e.g., preoperative endotracheal intubation), and preexisting neurological impairment (e.g., dementia, altered mental status).

Before study commencement, all anesthesia providers – residents in training, nurse anesthetists, and faculty anesthesiologists – were shown a video that demonstrated how to administer the AFT. The AFT was then administered prospectively by the anesthesia providers during the preoperative interview, which typically took place immediately before surgery. The AFT scores were entered into the electronic medical record. If the AFT was not administered preoperatively, the anesthesia providers were instructed to provide the reason(s) why. They reported clinician oversight or time constraint as reasons.

One investigator (L.S.L.) measured postoperative delirium using a previously validated chart-based tool published by Inouye *et al.*,⁶ after which the results were validated by a second investigator (J.M.L.), who was blinded to L.S.L.'s delirium designation. The two investigators resolved disagreements by reviewing the chart together to achieve consensus. The chart-based tool to measure delirium was previously validated with interviewer ratings using the Confusion Assessment Method in a large cohort of hospitalized older adults. Results indicated that it had a sensitivity of 74% and specificity of 83%. Finally, a third investigator (J.T.W.) determined which patients had taken the AFT preoperatively. Because AFT scores were not abstracted until delirium was diagnosed, investigators who made the diagnosis had been blinded to the AFT scores. The medical records were also abstracted to obtain data on demographics, length of hospital stay, intensive care unit (ICU) admission status, discharge disposition, and in-hospital mortality.

Bivariate analyses were conducted using Chi-square and t-tests. The final model was constructed using multiple logistic regression analysis. Statistical analysis was conducted using Stata[®] 8.0 software (College Station, TX, USA) or Microsoft Excel[®] for Mac 2011.

Results

During the six-month period, 432 patients were examined and 362 satisfied the study inclusion/exclusion criteria. Reasons for excluding patients were communication difficulty ($n = 32/432$, 7%), patient refusal or anxiety ($n = 23$, 5%), and neurological impairment ($n = 15$, 4%).

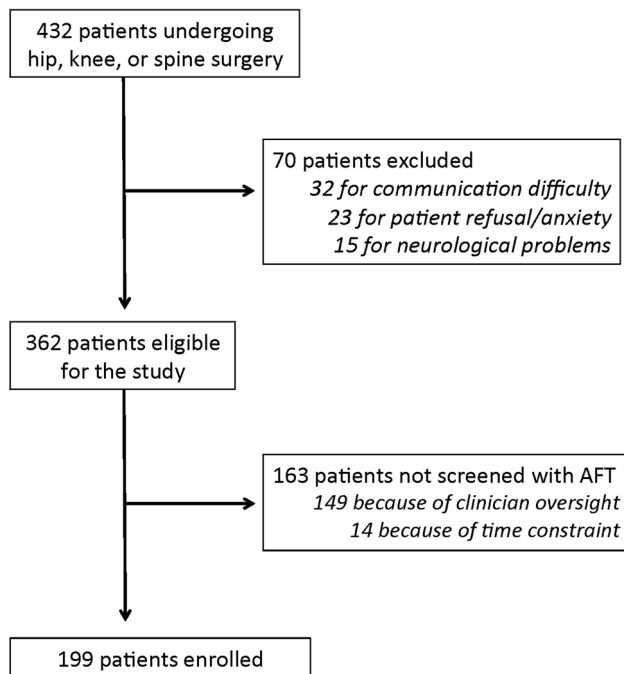


Fig. 1 Enrolment flowchart

A number of patients satisfied the inclusion/exclusion criteria but were not AFT screened because of time constraint ($n = 14/362$, 4%) or clinician oversight ($n = 149$, 41%). The remaining 199 patients (55%) formed the sample for this report (Fig. 1). The demographics of patients included in this report were similar to those who were excluded/not screened with respect to mean age, sex, American Society of Anesthesiologists Physical Status (ASA) classification, emergency case status, rate of ICU admission, and rate of postoperative delirium (see Table 1). These two groups differed only by surgery type.

Of the 199 patients who were screened with the AFT, 57 (29%) scored < 15 . The overall rate of postoperative delirium as measured by chart review was 19% ($n = 38$, 95% confidence interval [CI]: 14 to 25%). Patients with AFT scores < 15 were more likely to develop postoperative delirium than patients with scores ≥ 15 (54% vs 5%,

$P < 0.01$) (Fig. 2). This finding was supported by receiver operating characteristic curve analysis, which yielded an area under the curve of 0.83 (95% CI: 0.76 to 0.90), suggesting that the AFT was able to identify patients at risk for delirium (Fig. 3). At the cutoff of < 15 used in this study, the sensitivity and specificity were 84% and 75%, respectively, and the positive and negative likelihood ratios were 3.39 and 0.21, respectively.

Patients with AFT scores < 15 were more likely to be ≥ 80 yr of age (23% vs 11%, $P = 0.02$) and had a high ASA classification (III or IV) (68% vs 44%, $P < 0.01$). Multiple logistic regression analysis, with postoperative delirium as the dependent variable, identified an AFT score < 15 (crude odds ratio [OR] 19.8, 95% CI: 7.7 to 51.4; adjusted OR 20.1, 95% CI: 7.9 to 51.4) and high ASA classification (crude OR 3.3, 95% CI: 1.2 to 8.8; adjusted OR 3.5, 95% CI: 1.3 to 9.2) as independent predictors. Additional covariates not significant in the final model were age, sex, ICU admission, and emergency case status (see Table 2).

Lastly, patients with postoperative delirium as measured by chart review had worse clinical outcomes than non-delirious patients. They had a higher ICU admission rate (42%, 95% CI: 26 to 59% vs 16%, 95% CI: 10 to 22%, $P < 0.01$), longer length of hospital stay (7.8 days, 95% CI: 5.1 to 6.1 vs 5.6 days, 95% CI: 6.2 to 9.4, $P < 0.01$), higher rate of institutionalization after discharge (92%, 95% CI: 79 to 98% vs 42%, 95% CI: 35 to 50%, $P < 0.01$), and higher mortality rate (5%, 95% CI: 1 to 18% vs 0.6%, 95% CI: 0 to 3%, $P = 0.03$).

Discussion

Our study shows that older patients who scored < 15 on the AFT before surgery were more likely to become delirious postoperatively, and that patients who became delirious had worse clinical outcomes overall. Our results are novel in demonstrating an association between a 60-sec preoperative cognitive screening test and the development of postoperative delirium as measured by chart review.

Table 1 Characteristics of patients included in this report vs those who were excluded or not screened with the AFT

| Characteristic | Patients included (n = 199) | Patients excluded / not screened (n = 233) | P-values |
|------------------------|-----------------------------|--|-----------|
| Age (yr) | 73.4 (5.5) | 74.0 (6.6) | 0.30 |
| Female sex | 53% (n = 106) | 55% (n = 128) | 0.77 |
| ASA III/IV | 51% (n = 101) | 44% (n = 103) | 0.18 |
| Emergency cases | 11% (n = 22) | 13% (n = 30) | 0.55 |
| Spine surgery | 71% (n = 142) | 52% (n = 121) | < 0.001 |
| ICU admission | 20% (n = 40) | 16% (n = 37) | 0.26 |
| Postoperative delirium | 19% (n = 38) | 19% (n = 44) | 0.90 |

AFT = animal fluency test; ASA = American Society of Anesthesiologists physical status classification; ICU = intensive care unit

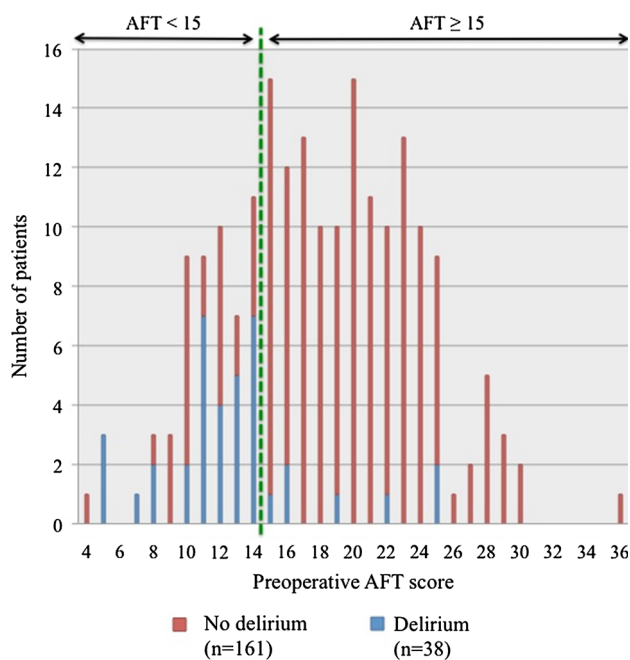


Fig. 2 Number of patients with and without postoperative delirium at any given animal fluency test (AFT) score. The AFT scores (range 4–36) are shown on the horizontal axis. Numbers of patients with those AFT scores are shown on the vertical axis. Blue bars = patients who developed postoperative delirium; red bars = patients who did not develop delirium; bars to the left of the green dotted line = patients whose AFT scored was < 15. It is evident that patients with scores < 15 were more likely to develop delirium, as seen by the predominance of blue bars to the left of the green dotted line and red bars to the right of it

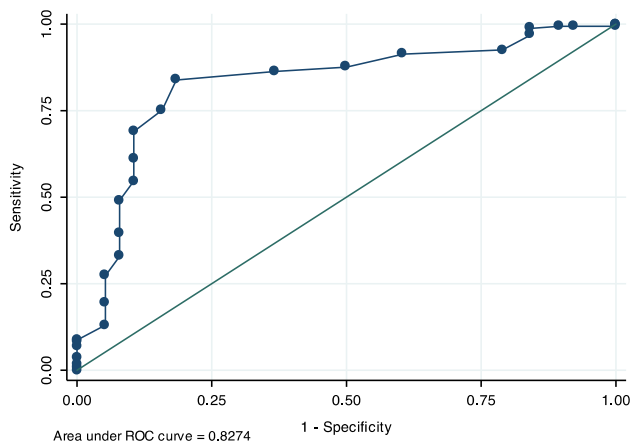


Fig. 3 A receiver operating characteristic curve was generated to determine the sensitivity and specificity of the AFT for predicting postoperative delirium

A previous small study was conducted on a group of patients aged 70 yr or older who were undergoing proximal femoral fracture repair and were free of known dementia and delirium preoperatively. Results showed that a brief cognitive screen – which included a three-word recall test, simplified clock drawing, and spatial orientation – predicted

the risk of postoperative delirium in a univariate analysis and after adjustment for age.⁷ Our results are in agreement with the conclusion of that study: that a limited cognitive screen such as a one-minute fluency test can be prognostically useful for identifying patients at risk of developing postoperative delirium. One potential advantage of the AFT – in contrast to other cognitive screens, such as the Mini Cog⁸ – is that it can be administered in person or over the telephone. Tests that include a clock drawing, for example, require in-person testing, and the results may be difficult to store and retrieve.

Our study was not powered to examine the relative importance of all known risk predictors of postoperative delirium. In addition to AFT scores < 15, however, a high ASA physical status independently predicted the development of postoperative delirium, which was demonstrated previously.¹ Although advanced age is also an established risk factor for postoperative delirium, we did not find an independent association in our analysis. This discrepancy may be explained by the selection bias of including only older patients, thereby truncating the age distribution and attenuating the effect of age on delirium.

A key aspect of this study was to examine the potential barriers of adding a cognitive screen to the routine preoperative evaluation. With a screening rate of 55%, our study suggests that the completion rate could be augmented in future, larger studies. Given that clinician oversight was the most common reason cited for not administering the AFT, education and reinforcement of this practice as part of a routine preoperative evaluation may improve its success. More importantly, investigators should seek better understanding of the barriers to administering a preoperative cognitive screen so strategies can be developed to overcome them.

The overall 19% rate of delirium reported in our study is comparable to that in previous studies of older surgical patients undergoing major surgery.^{1,9–11} The incidence of delirium has ranged from 10–62%, with higher rates reported in higher-risk patients, such as those with a hip fracture.¹² In contrast, studies on more homogeneous patient populations, such as those undergoing elective orthopedic arthroplasty, have reported delirium rates of approximately 10%.^{10,13}

There are limitations to our study. First, this one-minute cognitive screen may not be sensitive enough to identify mild cognitive impairment. Second, we investigated only the cutoff value of < 15. We did not consider other values, which need to be investigated in future larger studies. Third, the educational background and/or anxiety state of patients may influence test performance. Fourth, we included only patients undergoing spine surgery or arthroplasty, so the results may not be generalizable to patients undergoing other types of surgery. Fifth, a large

Table 2 Crude and adjusted odds ratios for predictor variables of postoperative delirium

| Predictor variables | Crude model | | Adjusted model | |
|---------------------|-------------|-------------|----------------|-------------|
| | Odds ratio | 95% CI | Odds ratio | 95% CI |
| AFT < 15 | 19.8 | 7.7 to 51.4 | 20.1 | 7.9 to 51.4 |
| ASA III/IV | 3.3 | 1.2 to 8.8 | 3.5 | 1.3 to 9.2 |
| Age \geq 80 yr | 1.4 | 0.5 to 4.6 | | |
| Female sex | 1.3 | 0.5 to 3.1 | | |
| ICU admission | 0.8 | 0.2 to 2.8 | | |
| Emergency case | 0.7 | 0.2 to 3.0 | | |

AFT = animal fluency test; ASA = American Society of Anesthesiologists classification; ICU = intensive care unit

number of patients were not screened with the AFT, mostly because of clinician oversight, which may have introduced a selection bias. Sixth, because we found that patients who developed delirium were more likely to be admitted to the ICU and had a longer length of hospital stay, it is possible their mental status was assessed more frequently, resulting in an observation bias. Finally, the choice of the chart-based tool to screen for incident delirium may result in misclassification. For example, a recent study demonstrated that a chart-based tool was more likely to miss patients with hypoactive delirium.¹⁴ However, the patients in our study who were deemed delirious as measured by chart review had worse postoperative outcomes than non-delirious patients, supporting the validity of the chart-based tool used for measuring delirium. Future studies with adequate sample size should include a standardized, validated tool to measure delirium actively and prospectively. It should be noted that this was an initial investigation into the feasibility of administering the AFT and its effectiveness for predicting delirium. Accordingly, we did not measure a large number of covariates for postoperative delirium. A future study with adequate sample size is needed that includes consideration of additional covariates for delirium development.

In conclusion, a simple, brief preoperative cognitive screening test, such as the AFT, is potentially useful when applied prior to surgery. Based on our data, the AFT was able to identify older surgical patients at risk for developing postoperative delirium as measured by chart review. Although limited participation by eligible participants in this study raises questions about the usefulness and feasibility of systematic administration of the AFT, our study adds to the growing “call for action” for anesthesiologists to incorporate cognitive assessment in the preoperative evaluation of older surgical patients.¹⁵

Acknowledgement The authors would like to thank John Feiner for his assistance with the ROC analysis.

Conflicts of interests None declared.

Funding This project was supported in part by the National Institute of Aging, National Institutes of Health, Bethesda, MD, Grant # NIH 1RO1AG031795-05 (Leung).

References

1. Dasgupta M, Dumbrell AC. Preoperative risk assessment for delirium after noncardiac surgery: a systematic review. *J Am Geriatr Soc* 2006; 54: 1578-89.
2. Youngblom E, DePalma G, Sands L, Leung J. The temporal relationship between early postoperative delirium and postoperative cognitive dysfunction in older patients: a prospective cohort study. *Can J Anesth* 2014; 61: 1084-92.
3. Long LS, Shapiro WA, Leung JM. A brief review of practical preoperative cognitive screening tools. *Can J Anesth* 2012; 59: 798-804.
4. Canning SJ, Leach L, Stuss D, Ngo L, Black SE. Diagnostic utility of abbreviated fluency measures in Alzheimer disease and vascular dementia. *Neurology* 2004; 62: 556-62.
5. Leung JM, Sands LP, Lim E, Tsai TL, Kinjo S. Does preoperative risk for delirium moderate the effects of postoperative pain and opiate use on postoperative delirium? *Am J Geriatr Psychiatry* 2013; 10: 946-56.
6. 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.
7. Meziere A, Paillaud E, Belmin J, et al. Delirium in older people after proximal femoral fracture repair: role of a preoperative screening cognitive test. *Ann Fr Anesth Reanim* 2013; 32: e91-6.
8. Borson S, Scanlan JM, Watanabe J, Tu SP, Lessig M. Simplifying detection of cognitive impairment: comparison of the Mini-Cog and Mini-Mental State Examination in a multiethnic sample. *J Am Geriatr Soc* 2005; 53: 871-4.
9. Bohner H, Hummel TC, Habel U, et al. Predicting delirium after vascular surgery: a model based on pre- and intraoperative data. *Ann Surg* 2003; 238: 149-56.
10. Kalisvaart KJ, Vreeswijk R, de Jonghe JF, van der Ploeg T, van Gool WA, Eikelenboom P. Risk factors and prediction of postoperative delirium in elderly hip-surgery patients: implementation and validation of a medical risk factor model. *J Am Geriatr Soc* 2006; 54: 817-22.
11. Robinson TN, Raeburn CD, Tran ZV, Angles EM, Brenner LA, Moss M. Postoperative delirium in the elderly: risk factors and outcomes. *Ann Surg* 2009; 249: 173-8.

12. *Bitsch M, Foss N, Kristensen B, Kehlet H.* Pathogenesis of and management strategies for postoperative delirium after hip fracture: a review. *Acta Orthop Scand* 2004; 75: 378-89.
13. *Jankowski CJ, Trenerry MR, Cook DJ, et al.* Cognitive and functional predictors and sequelae of postoperative delirium in elderly patients undergoing elective joint arthroplasty. *Anes Analg* 2011; 112: 1186-93.
14. *Saczynski JS, Kosar CM, Xu G, et al.* A tale of two methods: chart and interview methods for identifying delirium. *J Am Geriatr Soc* 2014; 62: 518-24.
15. *Crosby G, Culley DJ, Hyman BT.* Preoperative cognitive assessment of the elderly surgical patient: a call for action. *Anesthesiology* 2011; 114: 1265-8.