

Validation of the Delirium Diagnostic Tool-Provisional (DDT-Pro) With Medical Inpatients and Comparison With the Confusion Assessment Method Algorithm

José G. Franco, M.D., Ph.D., María V. Ocampo, M.D., Juan D. Velásquez-Tirado, M.D., Daniel R. Zaraza, M.S., Alejandra M. Giraldo, M.D., Paola A. Serna, M.D., Carolina López, M.B., Adolfo Zuluaga, M.B., Esteban Sepúlveda, M.D., Ph.D., Jacob Kean, Ph.D., Paula T. Trzepacz, M.D.

Objective: Delirium remains underdetected as a result of its broad constellation of symptoms and the inadequate neuropsychiatric expertise of most medical-surgical clinicians. Brief, accurate tools are needed to enhance detection.

Methods: The authors extended validation of the Delirium Diagnostic Tool-Provisional (DDT-Pro), originally validated in a study of inpatients with traumatic brain injury for diagnosis of delirium by nonexpert clinicians, for 200 general medical inpatients in Colombia. The three structured, quantitatively rated items in DDT-Pro represent the three core delirium domains.

Results: High interrater reliability between physician and nurse (0.873) administrators, internal consistency (>0.81), and content validity were found. Compared with independent reference standard diagnosis with DSM-5 or the Delirium Rating Scale-Revised-98, the area under the receiver operating characteristic (ROC) curve (global diagnostic accuracy) range was 93.8%–96.3%. ROC analysis revealed the same cutoff score (≤ 6) as that for the original study, with somewhat lower sensitivities of 88.0%–90.0% and

specificities of 85.3%–81.2% (independent expert physician or nurse ratings). Even when rated by a trained expert physician, the original version of the Confusion Assessment Method algorithm (CAM-A) performed moderately, with lower sensitivities (61.8%–70.0%) than the DDT-Pro (88.0%–100%) and somewhat higher specificities (84.8%–95.3% versus 67.4%–86.7%), with values depending on dementia status, reference standard, and rater type. Accuracies for the DDT-Pro and CAM-A were comparable (DDT-Pro: 83.0%–87.5% versus CAM-A: 87.5%–88.5%), although lower in the dementia subgroup, especially for CAM-A. However, these tools were significantly discordant, especially in negative cases, which suggests that they do not detect diagnosis of patients in the same way.

Conclusions: The DDT-Pro had high validity and reliability in provisional delirium diagnosis by physicians and nonexpert clinicians, although further validation is warranted before widespread use can be recommended.

J Neuropsychiatry Clin Neurosci 2020; 32:213–226;
doi: 10.1176/appi.neuropsych.18110255

Delirium has a high point prevalence, occurring in one out of five patients in general hospitals (1). Polypharmacy, organ insufficiency, postsurgical states, and greater medical burden are causal factors (2). Preexisting cognitive impairment and any brain disorder are major risk factors for delirium, as is advanced age (3). Delirium is associated with longer hospital stay, more medical complications, greater difficulty with rehabilitation compliance, psychological sequelae, and increased mortality during and after hospitalization (4). Delirium among elderly persons is associated with long-term cognitive and functional decline in the year following hospitalization, attributed to either

acceleration of preexisting cognitive impairment or new insults as a result of medical conditions, anesthesia, or surgical procedures (5).

Despite many harbingers of poor prognosis, almost two-thirds of delirium cases are undetected in clinical practice (6). The breadth of characteristic neuropsychiatric symptoms of delirium is daunting for clinicians without expertise in psychiatric interviewing and differential diagnosis. Referral to an expert is not guaranteed in emergency and medical-surgical settings (7). Misdiagnosis is common (depression, dementia, etc.), and patients with hypoactive delirium may never be diagnosed (8). Elderly patients are at

Online supplemental materials can be found by selecting "View Options" (located above the Abstract when accessing article on computer or tablet) or "About" (located above the title when accessing on smartphone).

higher risk for delirium but also for other neurocognitive conditions that complicate differential diagnosis and result in delirium being misdiagnosed as dementia and proper management not being sought (9).

Delirium affects many higher cerebral cortical functions, in which inattention is a cardinal feature. It is an acute brain impairment causing an altered state of consciousness but not a lack of consciousness, which defines stupor and coma. Imbalances in neurotransmitter levels and thalamocortical and cortico-cortico neural network dysregulation likely underlie the impairment (10, 11). Oftentimes, core symptoms of delirium are not assessed, while instead nursing management issues or disturbing psychotic symptoms are more likely noticed (12).

A brief and easy-to-administer tool for use by nonexpert physicians and nurses to routinely detect delirium in at-risk patient populations is needed. The challenge is to not compromise validity or standardization of administration, given the breadth and unique differentiating features of delirium symptoms.

The most commonly used brief detection tool is the Confusion Assessment Method algorithm (CAM-A) (13). The CAM-A rates four symptoms of delirium (temporal features, inattention, disorganized thinking, and altered level of consciousness) as being present or absent, although items were not originally standardized for administration or well anchored for item rating. The measure also allows coma or stupor to be considered as delirium. CAM-A performance metrics vary considerably across studies by training and rater expertise in delirium (14, 15). Its delirium case identification differs when more comprehensive approaches are used (16, 17). As a result of its brevity and convenience, it is widely used by nonexpert clinicians for delirium detection. To improve its standardization and accuracy, the CAM-A was adapted with interview techniques, although not all clinicians utilize these techniques.

There have been other efforts to develop brief assessment tools for delirium other than CAM-A; however, these tools are not as widely used or studied (18). The 4A's Test (4AT), an assessment test for cognitive impairment and delirium, is a screening tool with four items that assess the alertness, orientation, attention, and history of a patient's recent change or fluctuation in cognition or other mental function. The global accuracy of the 4AT for delirium diagnosis in accordance with DSM-5 criteria among geriatric inpatients from an urban teaching hospital was 84%, with a sensitivity of 86.7% and a specificity of 69.5% for a cutoff score >4 (19).

Other screening tools are intended for nurses' use, but reviews of such tools have not been prevalent in the literature. These tools usually involve observations of behaviors rather than phenomenological assessment of core delirium characteristics, although one could argue that the items are more accessible to the work focus of nursing staff. Additionally, the use of these tools does not provide a provisional diagnosis. The Single Question in Delirium, a single screening question to family members asking whether the patient is more confused lately, had an 80% sensitivity and a 71% specificity, and its sensitivity outperformed the CAM-A by 40% when administered by untrained staff (20). The

Nursing Delirium Screening Scale does not assess attention, a cardinal feature of delirium (21). The Delirium Observation Screening Scale rates 25 behavioral observations (22), and a 13-item algorithm was developed from it. The Neecham Confusion Scale rates cognition, behaviors, and three non-brain body functions (23). The Delirium-O-Meter rates 12 behaviors (24).

Phenomenological research over the course of nearly two decades has revealed that delirium has three core symptom domains: cognition (attention and vigilance as the cardinal characteristic, plus orientation, memory, and visuospatial ability), higher-order thinking (language, thought process, and comprehension, as well as executive function), and circadian rhythm (sleep-wake cycle and motor activity) (12, 25–29). This work led to the development of a brief tool designed to enhance accuracy by only assessing these core symptom domains: the Delirium Diagnostic Tool-Provisional (DDT-Pro) (30). The DDT-Pro provides a provisional diagnosis of delirium until an expert clinician can verify it. The DDT-Pro scale has three structured items that quantitatively rate three core domains (cognition [vigilance], higher-order thinking [comprehension], and circadian rhythm [sleep-wake]) using observable data, thereby enhancing standardization for nonexpert raters.

The DDT-Pro was first validated among subacute traumatic brain injury (TBI) rehabilitation inpatients rated by nonphysician clinical staff (30). It accurately classified 97% of patients, according to an independently ascertained DSM-IV-TR delirium diagnosis status by experts, and, in receiver operating characteristic (ROC) curve analysis, it performed similarly to the independently rated Delirium Rating Scale-Revised-98 (DRS-R-98). It is a simple, brief tool and, as a result of its content, construction, and administration, may have some advantages over the CAM-A for initial detection of delirium. Before it can be more widely used, however, it needs further validation in other patient populations and languages.

The main objective of the present study was to translate the DDT-Pro into Spanish to validate its performance among patients ≥ 60 years old who were admitted to an internal medicine service. A second aim was to compare performances of the commonly used original CAM-A version and the DDT-Pro against independently ascertained reference standards for delirium diagnosis: DSM-5 and the DRS-R-98.

METHODS

Design, Ethical Aspects, and Participants

In this cross-sectional prospective validation study, we followed recommendations from the Standards for Reporting of Diagnostic Accuracy guidelines (31), and the study was approved by our institution's ethics committee (Comité de Ética de Investigación en Salud de la Universidad Pontificia Bolivariana, Medellín, Colombia). Written informed consent was obtained from all participants, or from a proxy if the participant was determined to be cognitively impaired by a psychiatrist.

Consecutive patients ages ≥ 60 years who were admitted to the internal medicine ward of the Clínica Universitaria

Bolivariana (an urban teaching hospital from Medellín, Colombia) were evaluated until 200 eligible patients were identified. For inclusion in the study, patients had to be assessed during the first 12–48 hours following hospitalization. Exclusion criteria were refusal to participate, stupor or coma at the assessment period, severe language or hearing difficulty, or inability to speak Spanish. Patients with and without delirium were included.

Study Instruments

We collected relevant demographic and clinical variables using a standard method.

Short-Form Charlson Comorbidity Index. The Short-Form Charlson Comorbidity Index is a measure with prognostic value, used as an indicator of baseline medical severity status. It assesses the history of eight conditions: cerebrovascular accident, diabetes mellitus, chronic obstructive pulmonary disease, congestive heart failure or ischemic cardiopathy, dementia, peripheral arterial disease, chronic renal failure or dialysis, and cancer. For the first six conditions, each score was ranked one point, and for the last two conditions, each score was ranked two points. The score range is 0–10, with scale scores 0–1 indicating no comorbidity, 2 indicating low comorbidity, and ≥ 3 indicating high medical comorbidity (32).

Spanish-Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE). The IQCODE comprises 26 questions to form a structured interview asking patients' caregivers for information about cognitive impairment and functioning during the past 5 years, with a score range of 26–130. We used the validated Spanish version, which indicates likely dementia for scores >85 (33).

DRS-R-98. The DRS-R-98 is a well-validated and widely used instrument with two scales, one for delirium diagnosis (Total scale) and one for severity evaluation (Severity scale), developed on the basis of known delirium characteristics (34). The DRS-R-98 Total scale (16 items) consists of a maximum of 46 points, with 13 symptom items comprising the DRS-R-98 Severity scale. Three diagnostic items are for assessment of temporal onset, fluctuation, and etiology. The DRS-R-98 has a score range from 0 (no alteration) to 3, anchored by detailed phenomenological descriptions of clinical features unique to delirium. The validation study for the DRS-R-98 Spanish version had a cutoff score of >14 for delirium diagnosis in high incidence of delirium samples, and it performed independently of the diagnostic criteria system used (DSM-III-R, DSM-IV, or DSM-5) (35). The rating time frame was the preceding 12–24 hours, with all sources of available information used (caregivers and family members, medical charts, and hospital staff).

CAM-A. To diagnose delirium, CAM-A requires the presence of two items (one for temporal features and one for

inattention), plus either or both of two other items (one for disorganized thinking and the other for altered level of consciousness, in which the latter allows stupor and coma to be considered as delirium). Given that the Spanish version of CAM-A has some modifications from the original version (36), the original CAM-A four-item structured tool was used in the present study (for further details, see the CAM Training Manual and Coding Guide containing the algorithm [37, p 11]). Each item is dichotomously checked as absent (no) or present (yes), and the presence of at least three of the four items indicates delirium, as long as the first two items both are checked as yes. The study rater was trained prior to the study inception. The time frame for rating was 12–24 hours, and the rater had access to any source of information available (caregivers and family members, medical charts, and hospital staff) in addition to clinical assessment of patients for answering the four yes-no CAM-A questions.

DSM-5 diagnostic criteria for delirium. For a DSM-5 diagnosis of delirium, a patient must meet all five criteria established (38). Each criterion from DSM-5 was recorded as being present (yes) or absent (no) for each participant, and the rater had access to any source of information available (caregivers and family members, medical charts, hospital staff) in addition to clinical assessment. The time frame for rating was the previous 12–24 hours.

DDT-Pro. The DDT-Pro was translated from English to Spanish for the present validation study (see the Procedures section). It is a three-item brief, structured scale to provisionally diagnose delirium by nonexpert clinicians (30). It comprises structured and quantitatively scored items for evaluating vigilance and comprehension taken from the Cognitive Test for Delirium (CTD). The DDT-Pro is assessed directly by patient performance (39) and the sleep-wake cycle disturbance item from the DRS-R-98, assessed by patient evaluation or any other source for the preceding 12–24 hours. DRS-R-98 score values were reversed to align with the point scoring for the CTD items.

The three items reflect and represent the core symptom domains of delirium (cognition, higher-order thinking, and circadian rhythm), and to enhance the accuracy, only these three symptoms are rated. The first two items are rated objectively as tests, and the last one has anchored descriptors of observations. Scores range from 0 (worst performance) to 9 (best performance). The best cutoff score for delirium was a DDT-Pro score ≤ 6 (30). It has two equivalent forms, A and B, in which the two CTD vigilance and comprehension items differ to avoid a possible learning effect from repeated administration. These A and B forms were determined by neuropsychological experts to be equivalent. All individuals who administered the DDT-Pro in the present study were trained prior to the study inception. The time frame for rating the sleep-wake cycle disturbance item was the previous 12–24 hours. Each item

was scored from 0 to 3, with CTD items following scoring formulas.

Procedures

DDT-Pro translation and face validity. We followed the Principles of Good Practice for Translation of the original English version to Spanish, which included forward translations; reconciliation of forward translations into one sole version; back translation and back translation review by the tool authors; harmonization according to the original authors' feedback; and, as a final step, "cognitive debriefing" (testing in a small group of relevant people to evaluate face validity) (40).

The forward translations were made by four independent psychiatrists experienced in assessing patients with delirium in two Spanish-speaking countries (Colombia and Spain). The reconciled forward translation was back-translated to English by independent native English translators from a known online company that provides translation of scientific and medical documents.

For cognitive debriefing, we asked four independent psychiatrists from Colombia and Spain to anonymously complete an electronic survey about the face validity of the DDT-Pro that rated each of four questions, from 1 (not understandable at all) to 6 (totally understandable), about the comprehensibility of language and content of the scale and of its instructions. The survey included a space for written suggestions from the experts. (Both the original English and Spanish versions of the DDT-Pro are available upon request from Dr. Trzepacz [ptrzepacz@outlook.com]).

Patient assessment procedures. During the study period, one of two members of the study team went each day (except Sunday) to review the electronic register of the ward for new patients who met study inclusion criteria and sent the patient list with identification codes to the senior researcher on a daily basis. A maximum of four patients per day were entered as potential participants. When there were more than four patients meeting inclusion criteria, the register reviewer used the randomizers app from random.org to randomly select the four patients before reviewing exclusion criteria.

The senior researcher assessed exclusion criteria among potential patients daily and obtained informed consent. This independent evaluator was a psychiatrist experienced in delirium assessment who collected patient demographic and clinical information, evaluated patients using DSM-5 criteria for delirium (from medical charts, hospital staff, patients, and caregivers), and administered the DRS-R-98. Entry into the study occurred immediately after this initial assessment, whereupon the researcher sent a list of identification codes of enrolled study participants to the rest of the team, and therefore they could perform their independent evaluations using study instruments on day 1 of the study.

Two independent researchers who had been trained took turns assessing patients' dementia status using the IQCODE,

with information from family members or home caregivers. Two other independent evaluators administered the DDT-Pro to every patient on day 1 of the study within the same 12-hour shift: the physician researcher administered the A form, and the study nurse without previous experience in delirium research administered the B form, without any prespecified order for their assessments. Another researcher with more than 10 years of experience in consultation-liaison psychiatry had been trained to use the CAM-A and independently rated the CAM-A on day 1 of the study.

To evaluate repeated use of the DDT-Pro for follow-up of clinical improvement of delirium (i.e., sensitivity to clinical change), on day 2, one of the researchers who had rated the IQCODE questions in interviews with family members or home caregivers (but who had not assessed any study participant for delirium or other variables) was told which patients met DSM-5 criteria for delirium; this step was implemented in order to independently administer the DDT-Pro on day 4 or 5. When the patient assessment period ended (day 6 or later), the senior researcher independently reviewed the medical charts of patients with delirium and talked with clinical staff to determine whether delirium symptoms had improved. This information was recorded.

Statistical Data

We created a SPSS 23.0 database for analyses, unless otherwise specified.

General analyses. Age, which was compared using *t* tests, is reported with means and standard deviations. Education years, short Charlson index scores, and DRS-R-98 Total and Severity scores are also reported using means and standard deviations, although nonnormally distributed to facilitate data interpretation, and their mean ranks were compared with the Mann-Whitney *U* test. Chi-square or Fisher's exact test was used for comparison of all other dichotomous demographic and clinical variables.

DDT-Pro reliability, score description, and concordance of item ratings. The intraclass correlation coefficient was used to assess interrater reliability of the equivalent DDT-Pro forms administered by the physician (form A) and nurse (form B). Median and interquartile range are reported for scores of total DDT-Pro and for each one of the three items assessed (vigilance, comprehension, and sleep-wake cycle disturbance). These three items were compared with the Mann-Whitney *U* test between patients without delirium and patients with delirium, in accordance with DSM-5 criteria. We also compared the concordance of the frequency (percentage) of responses (0, 1, 2, or 3) of each of the three DDT-Pro items by both evaluators (physician and nurse) for patients with and without delirium (Cochran *Q* test).

DDT-Pro internal validity and convergent validity. We assessed internal validity of the DDT-Pro with Cronbach's alpha for the total scale, when each item was eliminated, and

with item-scale correlations. Convergent validity of the DDT-Pro versus the DRS-R-98 Total and Severity scores was evaluated by means of one-tailed Spearman's rho correlations.

Criterion validity and selection of the DDT-Pro cutoff score for delirium. For criteria validation for delirium diagnosis, we used nonparametric ROC analyses to assess accuracy (percentage of patients correctly classified), sensitivity, and specificity of the continuous DDT-Pro scores' span against DSM-5 criteria as a main reference standard and against the DRS-R-98 Total cutoff score of >14 as a second reference standard. We used the Hanley-McNeil test for ROC curves (global accuracies) comparisons in Epidat 3.1 software. We selected the DDT-Pro score with the best balance of sensitivity and specificity as the cutoff, preferring higher sensitivity than specificity.

Accuracy of DDT-Pro cutoff score and comparison with CAM-A. We used Cochran Q test to compare the concordance between the CAM-A and DDT-Pro for delirium diagnosis. Spreadsheet crosstabs were used for calculation of the accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (+LR), and negative LR (-LR) of the DDT-Pro cutoff (dichotomous variable) for DSM-5 criteria and for DRS-R-98 diagnoses. Sensitivity, specificity, PPV, NPV, and LRs for CAM-A are also reported.

Sensitivity of the DDT-Pro to patients' clinical change. Wilcoxon test was used to compare DDT-Pro score from the day 1 study evaluations of patients with delirium with the follow-up evaluation for patients who had improved at day 4 or 5.

RESULTS

Face Validity of the DDT-Pro Spanish Translation

Mean scores from the expert survey about the comprehensibility of language and content of the tool and its instructions, as a final step of the translation process, ranged between 4.0 and 5.5 (score range, 1–6). The scale was slightly modified according to the experts' suggestions for improvement, and the final Spanish version of the DDT-Pro was approved by both the translators and the authors and then used in the assessment of patients.

Characteristics of the Study Sample

All assessments were performed during day shifts, most within the same 8-hour period for a given patient. The patient flow diagram is presented in Figure 1. Demographic and clinical characteristics of the 200 study subjects, 74 (37.0%) of whom had a diagnosis of dementia, are summarized in Table 1. Fifty (25.0%) patients met DSM-5 criteria for delirium; of these, 34/50 (68.0%) had comorbid dementia. As expected, patients with delirium were older, had more comorbid conditions, and had a lower education level and

higher dementia prevalence compared with patients without delirium. When the DRS-R-98 cutoff score of >14 for a diagnosis of delirium was used, 40 (20%) patients from the study sample had a positive diagnosis, 28/40 (70%) of whom had comorbid dementia. In the subgroup of 74 patients with a diagnosis of dementia, 34/74 (45.9%) met DSM-5 criteria for delirium, and 28/74 (37.8%) met DRS-R-98 criteria for delirium.

Description of DDT-Pro Scores, Concordance on Item Ratings, and Reliability

All DDT-Pro median scores (interquartile range) were significantly different between patients who met DSM-5 criteria for delirium and patients without delirium, regardless of who administered the measure or whether comparisons were made in the dementia subgroup (for further details, see Table S1 in the online supplement). Many of the percentages of DDT-Pro item ratings were similar, irrespective of administration by the physician or nurse (for further details, see Table S2 in the online supplement). The item with the highest number of concordances between raters was vigilance, followed by sleep-wake, and comprehension.

The intraclass correlation coefficient between the equivalent A and B forms of the DDT-Pro, as independently administered by a physician or nurse, was 0.873 (95% CI=0.832–0.904). In the dementia subgroup, it was also very good (intraclass correlation coefficient=0.875, 95% CI=0.801–0.921).

Internal Consistency

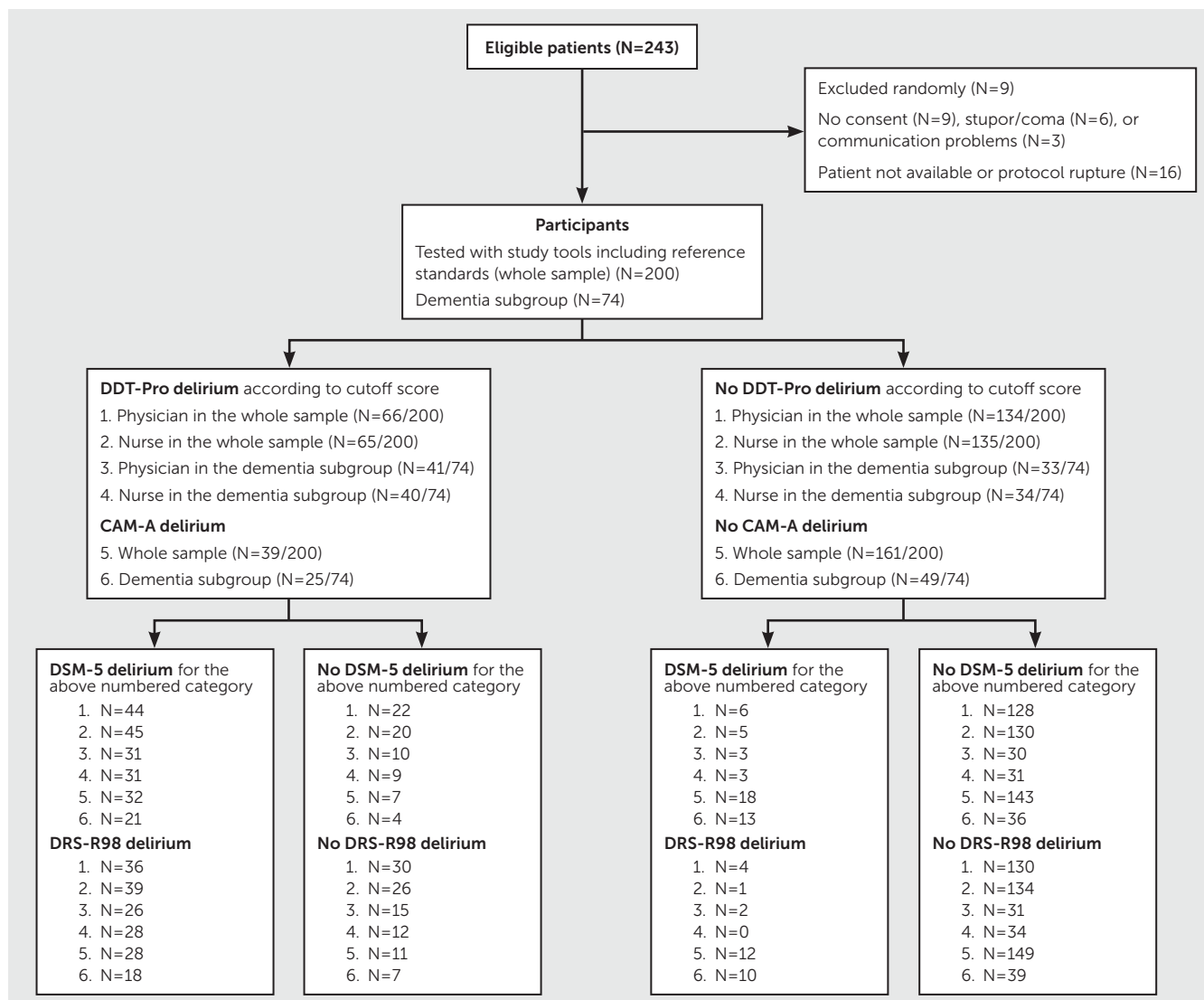
The internal consistency of the DDT-Pro was very good (Cronbach's α for administration by the physician and nurse, 0.809 and 0.816, respectively). All items were important for consistency, because the alpha value lowered or remained almost the same when any item was removed (range, 0.698–0.794 and 0.669–0.823 for administration by the physician and nurse, respectively) and because all item-scale correlations were high (≥ 0.624 and ≥ 0.594 for the physician and nurse, respectively). In addition, the DDT-Pro had a very good Cronbach's alpha in the subgroup of patients with dementia (0.824 and 0.843 for the physician and nurse, respectively).

Convergent Validity

As expected, the DDT-Pro and the DRS-R-98 scores were inversely related according to Spearman's rho. In the whole sample, the DDT-Pro administered by the physician correlated with the DRS-R-98 Total score at -0.698 and with the DRS-R-98 Severity score at -0.701 (-0.693 in both cases for the nurse). Correlations were even better in the dementia subgroup (-0.829 and -0.848 for the physician; -0.820 and -0.821 for the nurse).

When we assessed correlations only in the DSM-5 delirium cases (N=50), DDT-Pro scores correlated well with DRS-R-98 Total scores (-0.775 and -0.687 for administration by the physician and nurse, respectively) and DRS-R-98 Severity scores (-0.775 and -0.681 for

FIGURE 1. Flow diagram of the study participants^a



^a CAM-A=Confusion Assessment Method-Algorithm, DDT-Pro=Delirium Diagnostic Tool-Provisional; DRS-R-98=Delirium Rating Scale-Revised-98.

administration by the physician and nurse, respectively). In the subgroup of patients with DSM-5 delirium and comorbid dementia (N=34), Spearman’s rho for DDT-Pro with the DRS-R-98 Total was -0.798 for administration by the physician and -0.615 for administration by the nurse, and Spearman’s rho for DDT-Pro with DRS-R-98 Severity was -0.803 for the physician and -0.608 for the nurse.

Criterion Validity and Selection of the DDT-Pro Cutoff Score for Delirium

The area under the curve (AUC) for the ROC analysis of the DDT-Pro for accuracy of the diagnosis of delirium per DSM-5 criteria as administered by the physician was 94.1% (95% CI=90.6%–97.5%), and as administered by the nurse it was 93.8% (95% CI=90.0%–97.6%). Accuracy of the DDT-Pro by both evaluators in the subgroup of patients with dementia was >92.3%. AUC of the DDT-Pro for DRS-R-98 delirium as

administered by the physician was 94.9% (95% CI=91.6%–98.1%), and as administered by the nurse it was 96.3% (95% CI=93.9%–98.7%). Accuracy for both evaluators in the dementia subgroup was always >93.3%. Figure 2 shows that the AUCs for the physician administration were not significantly different from those of the nurse, nor were the AUCs for the physician or nurse in the whole sample different from those in the dementia subgroup.

ROC analysis sensitivity and specificity values for DDT-Pro diagnosis of delirium compared with either DSM-5 delirium or DRS-R-98 delirium are presented in Table 2. A gradient for ascending sensitivity and descending specificity was found. DDT-Pro scores ≤ 5 and ≤ 6 both had balanced sensitivity and specificity, although scores ≤ 6 had higher sensitivity, whereas scores ≤ 5 had higher specificity. Therefore, we selected a score of ≤ 6 as the DDT-Pro cutoff for delirium diagnosis.

Comparisons and Performance of the DDT-Pro and CAM-A

Concordance crosstabs for delirium determined by DDT-Pro (≤ 6 cutoff score) compared with delirium determined by CAM-A are presented in Table 3. CAM-A was administered only by the expert physician. The results were discordant as a consequence of classification differences for one or the other test, regardless of the DDT-Pro administration by the physician (Cochran $Q=19.703$, $p<0.001$) or by the nurse (Cochran $Q=21.125$, $p<0.001$). The tests were also discordant in the dementia subsample when the DDT-Pro was administered by the physician (Cochran $Q=12.800$, $p<0.001$) or the nurse (Cochran $Q=13.235$, $p<0.001$).

Accuracy of the DDT-Pro at the cutoff score of ≤ 6 was comparable to the CAM-A, when compared with the reference standards of DSM-5 criteria and with the DRS-R-98 Total scale cutoff score of >14 (Table 4). Almost all accuracies were within the 80%–90% range. However, accuracies were lower in the dementia subgroup than in the whole sample for both tools, although more so for the CAM-A (>10 percentage points decline versus 6 percentage points for DDT-Pro).

Sensitivity of the DDT-Pro at the cutoff score of ≤ 6 was very high, ranging between 88.0% and 100%, considering all scenarios, whereas the CAM-A values were much lower, in the 61.8%–70.0% range. Further DDT-Pro sensitivity values were higher in the dementia subgroup (91.2%–100%) than in the whole sample, whereas CAM-A values were even lower (61.8%–64.3%). Specificity values tended to be higher for the CAM-A (84.8%–95.3%) than for the DDT-Pro (67.4%–86.7%).

Indicators assessing the likelihood of a patient actually having delirium when results from the DDT-Pro or the CAM-A were positive for the disorder (PPVs, +LRs), or not having it when the tools revealed negative results (NPVs, -LRs), are in accordance with indicators assessing internal properties of these instruments (sensitivity, specificity). At the cutoff score of ≤ 6 , all DDT-Pro NPVs were $>90\%$, whereas the PPVs were lower. Conversely, CAM-A NPVs ranged from 73.5% to 92.5%, with PPVs a little better than those of the DDT-Pro. The +LRs of the CAM-A tended to

TABLE 1. Demographic and clinical characteristics of 200 inpatients from internal medicine wards, with or without delirium, per DSM-5 criteria^a

Characteristic	Whole sample (N=200)		Delirium (N=50)		No delirium (N=150)	
	Mean	SD	Mean	SD	Mean	SD
Age (years)	76.5	9.0	81.6 ^b	7.5	74.8 ^b	8.8
Education (years)	5.5	4.0	4.2 ^c	2.9	5.9 ^c	4.2
Short Charlson Comorbidity Index	2.1	1.4	2.8 ^c	1.6	1.9 ^c	1.3
DRS-R-98 Severity scale	6.6	7.7	17.8 ^c	7.6	2.8 ^c	2.2
DRS-R-98 Total scale	8.1	9.7	22.5 ^c	8.4	3.3 ^c	3.0
	N	%	N	%	N	%
Female	122	61.0	30	60.0	92	61.3
Retired	166	83.0	46	92.0	120	80.0
Dementia ^d	74	37.0	34 ^e	68.0	40 ^e	26.7
Most frequent admission diagnosis						
Systemic infection	87	43.5	25	50.0	62	41.3
Organ insufficiency	46	23.0	10	20.0	36	24.0
Metabolic/endocrine	24	12.0	7	14.0	17	11.3
Neoplasia or related complications	11	5.5	2	4.0	9	6.0
Two or more active medical diagnoses	142	71.0	42 ^e	84.0	100 ^e	66.7

^a Comparisons are between patients with delirium and patients without delirium. DRS-R-98=Delirium Rating Scale-Revised-98.

^b Data from the t test are shown ($p<0.01$).

^c Data from the Mann-Whitney U test are shown ($p\leq 0.025$).

^d Dementia diagnoses were in accordance with the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) score >85 , except two cases (1.0%) that were diagnosed by psychiatrist clinical impression using all available sources as a result of absence of any familiar or home caregiver for IQCODE assessment.

^e Data from the chi-square test are shown ($p<0.02$).

be higher than those of the DDT-Pro (i.e., better), and the -LRs of DDT-Pro tended to be lower than those of the CAM-A (i.e., better).

Sensitivity of the DDT-Pro to Patients' Clinical Change

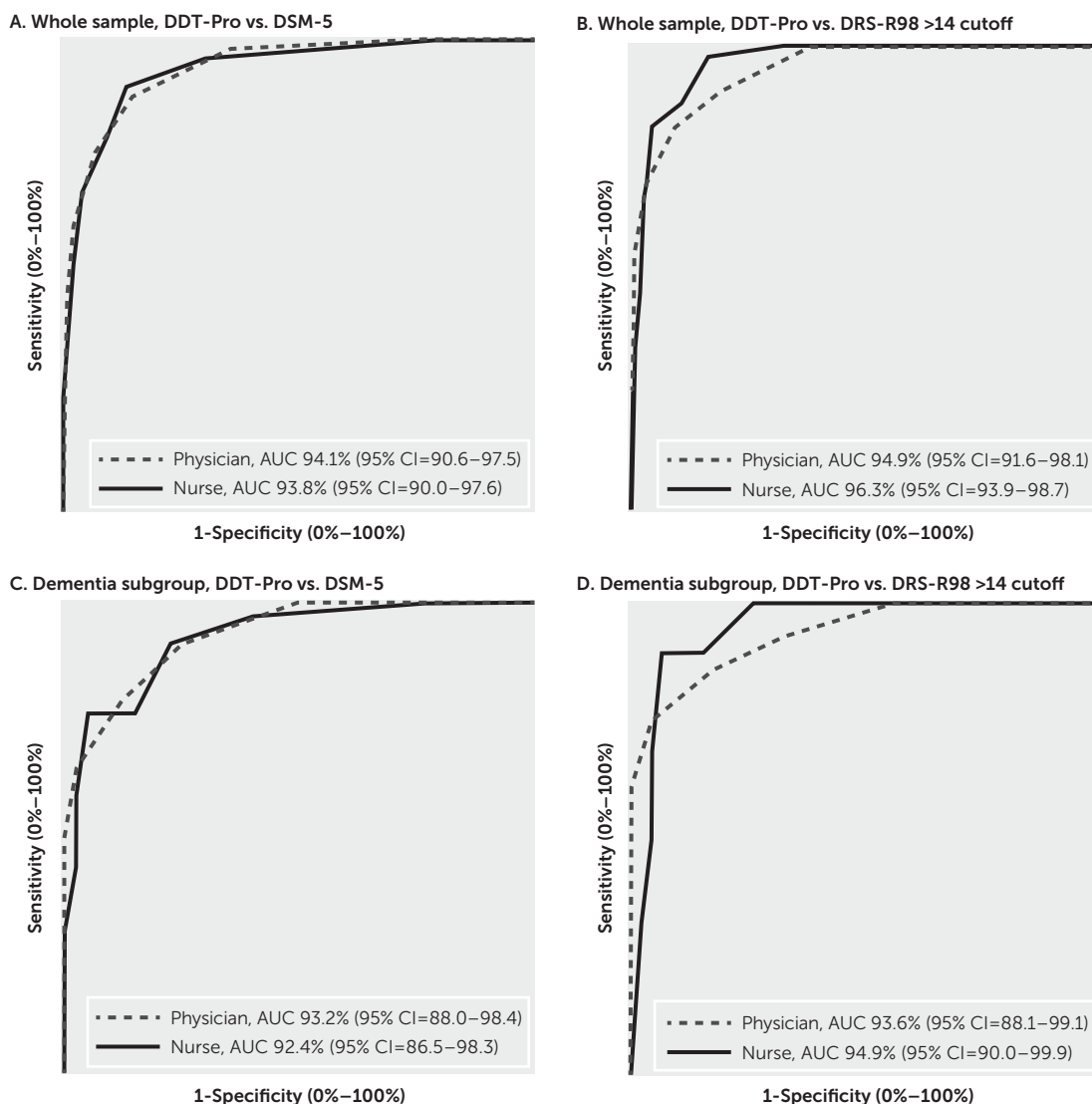
Delirium clinical improvement at day 4 or 5 of the study could be documented in 18 patients' charts, 13 of whom had comorbid dementia. The median DDT-Pro score at follow-up was 7.0 (interquartile range, 6.0–8.0), which implies a three-point difference from the initial score by the physician, who had a median of 4.0 (interquartile range, 1.7–5.0; Z score of Wilcoxon test= -3.741 , $p<0.001$), and the same regarding the nurse, who had a median of 4.0 (interquartile range, 1.7–6.0; Z score of Wilcoxon test= -3.312 , $p=0.001$).

DDT-Pro was also sensitive to clinical change in the dementia subgroup. Its median of 7.0 (interquartile range, 4.5–7.5) at follow-up implied a four-point difference from that of the initial physician score of 3.0 (interquartile range, 1.0–5.0; Z score of Wilcoxon test= -3.195 , $p=0.001$) and from that of the nurse score of 3.0 (interquartile range, 0.5–6.0; Z score of Wilcoxon test= -2.819 , $p=0.005$).

DISCUSSION

There remains a clinical need for a briefly administered, highly accurate, and reliable diagnostic tool for delirium to enhance delirium detection in clinical settings until expert

FIGURE 2. Area under the receiver operating characteristic curves for the Delirium Diagnostic Tool-Provisional (DDT-Pro) for delirium diagnosis^a



^a Area under the ROC curves are shown for the DDT-Pro administered by a physician (dotted lines) or nurse (solid lines) for 200 patients from an internal medicine ward by using either DSM-5 delirium criteria or a Delirium Rating Scale-Revised-98 (DRS-R-98) Total scale cutoff score >14 for delirium as the diagnostic reference standards (panels A and B). The ROC curve was created by plotting the true-positive rate against the false-positive rate at various threshold settings. Analyses are shown for the subgroup of 74 patients with dementia (panels C and D). Physician versus nurse areas for graphs of the whole sample or the dementia subgroup and areas for each rater type between the whole sample and the dementia subgroup are similar (all p values for Hanley and McNeil test for comparison of areas >0.050). Axes show sensitivity versus probability of false alarm (1-specificity). AUC=area under the curve.

physicians can confirm a diagnosis or when they are not available for all patients who might benefit from an evaluation. Such a tool should be accurate enough to be used as a provisional or “working” diagnosis to begin delirium evaluation until or in lieu of a delirium expert physician assessment or confirmation. The DDT-Pro was developed to provisionally diagnose delirium by assessing each of the three core domains of delirium (cognition [vigilance], higher-order thinking [comprehension], and circadian rhythm [sleep-wake]) and to be quantitative and highly structured, as a potential alternative to the widely used CAM-A screening tool. Therefore, we extended the validation

of the DDT-Pro, which was originally studied in 36 subacute inpatients with TBI.

When the DDT-Pro interrater reliability was compared between a physician rater who had delirium expertise and a nurse who did not, we found high comparability between their ratings on the equivalent forms of the tool, which suggests that it could be useful in routine clinical care settings by clinical staff who are not physicians or experts in delirium. This may be attributable to its content, which is focused only on the key features of the syndrome, as well as its structured ratings, which enable better standardization and interrater reliability. Furthermore, it comprises three

items from well-validated delirium tools, which enhance its validity and reliability. The latter is also revealed by the high rate of concordance between the physician and nurse for each of the score ratings of the DDT-Pro items.

Internal consistency and item-scale correlational analysis showed that all items were important for the performance of the DDT-Pro. Moreover, all individual item scores were statistically different between patients with delirium and control subjects (without delirium). This is consistent with the three items representing the three core domains of delirium. The delirium groups had high correlations of the DDT-Pro with the DRS-R-98 Total and Severity scores, which supports a high degree of construct validity (41).

Unlike in CAM-A, the DDT-Pro items are not dichotomously rated as present or absent without regard to severity. Its quantitative nature allows for ROC analyses of the DDT-Pro against delirium reference standards.

The DDT-Pro performed very well with regard to criterion validity, regardless of administration by a physician or nurse, against diagnoses made independently by a senior research physician using DSM-5 criteria or DRS-R-98 Total score cutoff for delirium (Figure 2). The DDT-Pro global diagnostic accuracy was high (>94.0%) and little affected by dementia, even though our population had a high comorbidity prevalence (nearly 40%). Accuracies were even higher against a DRS-R-98 delirium diagnosis (96.3% for

TABLE 2. Sensitivity and specificity analysis using receiver operating characteristic for the Delirium Diagnostic Tool-Provisional (DDT-Pro) for delirium diagnosis in accordance with DSM-5 and DRS-R-98 criteria^a

Cutoff	Whole sample (N=200)		Dementia subgroup (N=74)	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
DDT-Pro by a physician compared with DSM-5 delirium				
0	6.0	100.0	5.9	100.0
1	18.0	100.0	20.6	100.0
2	32.0	99.3	41.2	100.0
3	44.0	99.3	50.0	100.0
4	60.0	98.0	64.7	97.5
5	76.0	93.3	79.4	87.5
6	88.0	85.3	91.2	75.0
7	98.0	64.7	100.0	50.0
8	100.0	23.3	100.0	25.5
9	100.0	0.0	100.0	0.0
DDT-Pro by a nurse compared with DSM-5 delirium				
0	22.0	100.0	29.4	100.0
1	34.0	99.3	44.1	97.5
2	40.0	98.7	50.0	97.5
3	52.0	98.0	58.8	97.5
4	68.0	96.0	76.5	50.0
5	78.0	91.3	76.5	85.0
6	90.0	86.7	91.2	77.5
7	96.0	70.0	97.1	60.0
8	100.0	20.0	100.0	20.0
9	100.0	0.0	100.0	0.0
DDT-Pro by a physician compared with DRS-R-98 delirium				
0	7.5	100.0	7.1	100.0
1	22.5	100.0	25.0	100.0
2	40.0	99.4	50.0	100.0
3	55.0	99.4	60.7	100.0
4	70.0	96.9	75.0	95.7
5	82.5	90.6	85.7	82.6
6	90.0	81.2	92.9	67.4
7	100.0	61.2	100.0	43.5
8	100.0	21.9	100.0	19.6
9	100.0	0.0	100.0	0.0
DDT-Pro by a nurse compared with DRS-R-98 delirium				
0	25.0	99.4	32.1	97.8
1	40.0	98.7	50.0	95.7
2	47.5	98.1	57.1	95.7
3	62.5	97.5	67.9	95.7
4	82.5	95.6	89.3	93.5
5	87.5	89.4	89.3	84.8
6	97.5	83.7	100.0	73.9
7	100.0	66.9	100.0	54.3
8	100.0	18.7	100.0	17.4
9	100.0	0.0	100.0	0.0

^a The data indicate results for the DDT-Pro as administered by an expert physician and a nurse to patients from an internal medicine ward. The cutoff score ≤ 6 was chosen because it offered balance sensitivity and specificity, with relatively higher sensitivity. DRS-R-98=Delirium Rating Scale-Revised-98.

TABLE 3. Concordance crosstabs for the Confusion Assessment Method-Algorithm (CAM-A) and the Delirium Diagnostic Tool-Provisional (DDT-Pro) according to a cutoff score ≤ 6 for delirium for 200 patients in an internal medicine ward (dementia, N=74)^a

Crosstab	CAM-A by a consultation-liaison psychiatrist					
	Positive		Negative		Total	
	N	%	N	%	N	%
DDT-Pro by a physician, whole sample						
Positive	34	87.2	32	19.9	66	33.0
Negative	5	12.8	129	80.1	134	67.0
Total	39	100	161	100	200	100
DDT-Pro by a nurse, whole sample						
Positive	36	92.3	29	18.0	65	32.5
Negative	3	7.7	132	82.0	135	67.5
Total	39	100	161	100	200	100
DDT-Pro by a physician, dementia subgroup						
Positive	23	92.0	18	36.7	41	55.4
Negative	2	8.0	31	63.3	33	44.6
Total	25	100	49	100	74	100
DDT-Pro by a nurse, dementia subgroup						
Positive	24	96.0	16	32.6	40	54.0
Negative	1	4.0	33	67.3	34	45.9
Total	25	100	49	100	74	100

^a Data in each one of the four crosstabs are discordant (all Cochran Q p values <0.050).

administration by a nurse), with values >93% in the dementia subgroup. Beyond determining a single provisional diagnostic cutoff for the DDT-Pro, its excellent global accuracy warrants further study of its performance using other cutoffs for broader purposes, such as delirium confirmation in clinical settings or delirium diagnosis in research.

ROC analysis sensitivity and specificity of DDT-Pro scores enabled selection among cutoff scores in accordance with clinical interest. The cutoff score of ≤ 6 on the DDT-Pro had high sensitivity with balanced specificity and was the same cutoff score derived in the original TBI study (30), in which the mean DRS-R-98 score in the delirium group was nearly the same as the mean in the present cohort (22.3). This cutoff had a higher sensitivity than specificity in both studies, which is preferred for case detection, whereas those more interested, for example, in higher specificity may prefer the ≤ 5 cutoff. The dementia subgroup had the lowest specificity values, irrespective of rater type or reference standard. However, clinically speaking, it is better to overdiagnose possible delirium in an at-risk elderly population, for whom dementia may be a confounding factor, than to miss it. Additionally, when comorbid, delirium overshadows dementia symptoms (9).

Convergent validity was high for the DDT-Pro regardless of rater type. Correlations to the DRS-R-98 Total and Severity scores were both around -0.70 and even better in the

dementia subgroup (-0.82 to -0.85). In the delirium-only group, there was a discrepancy between rater types, such that the nurse's values were somewhat lower than the physician's. This difference was most apparent in the comorbid delirium-dementia subgroup, although this was small (N=34).

We preferred to use the original CAM-A, which is not anchored by tests with the original descriptors (for further details, see the training manual [37]). We used the original version, because most routine clinical evaluators would likely prefer this version for its ease of use. Even though we chose a delirium expert and consultation-liaison psychiatrist to administer the CAM-A, its sensitivity was considerably lower (64.0%) against DSM-5 criteria than for the nonexpert nurse's administration of the DDT-Pro (90%), which we attribute to differences in scale design and not to inadequate delirium training.

Concordance statistical analysis for the diagnostic performance of the DDT-Pro cutoff and the CAM-A for positive and negative diagnoses revealed notable differences when one was used over the other (Table 3). Even though they demonstrated high concordance for positive delirium (most scenarios >90%), their concordance was poorer in negative cases (<83.0%), especially in the dementia subgroup (as low as 63.3%). This degree of discordance implies that the performance of the tools was different for negative cases. This was corroborated by the DDT-Pro's very high sensitivity (90.0%–100.0%), revealing a low percentage of false negative diagnoses, versus much lower sensitivity for the CAM-A (61.8%–70.0%), suggesting more false negatives among the true positives. It is possible that scale structure (CAM-A items 3 and 4 have an either-or option) and the fact that stupor and coma are allowed to be rated as delirium were contributors to the lower sensitivity when tested against reference standards that do not involve those potential confounds. It is preferable to have higher sensitivity than specificity for a diagnostic tool when the condition is a medical urgency, as with delirium.

There were also important differences in the performance characteristics between these tools when compared with the two reference diagnostic standards (Table 4). Sensitivity values were higher for DDT-Pro, whereas specificity values tended to be higher for CAM-A, although CAM-A values were more adversely affected by dementia. This was true irrespective of whether the DDT-Pro was administered by the expert physician or nonexpert nurse, even though CAM-A was always administered by a delirium expert physician. The better specificity than sensitivity for CAM-A in this study sample could be a result of our use of the expert physician, who would know that delirium is not consonant with stupor or coma (even though CAM-A allows the rating of stupor and coma as delirium). Additionally, an expert in delirium using the unanchored CAM-A would apply full clinical knowledge of the syndrome, thereby enhancing diagnostic specificity.

TABLE 4. Accuracy of the Delirium Diagnostic Tool-Provisional (DDT-Pro) cutoff score ≤ 6 and of the Confusion Assessment Method-Algorithm (CAM-A) for delirium diagnosis according to DSM-5 criteria and to the DRS-R-98 Total scale cutoff score >14 among 200 patients from an internal medicine ward (dementia, N = 74)^a

Variable	Accuracy		Sensitivity		Specificity		Positive predictive value		Negative predictive value		Positive likelihood ratio		Negative likelihood ratio	
	%	95% CI	%	95% CI	%	95% CI	%	95% CI	%	95% CI	Likelihood ratio	95% CI	Likelihood ratio	95% CI
DDT-Pro by a physician or by a nurse versus DSM-5 delirium														
Whole sample (physician)	86.0	80.2, 90.3	88.0	75.0, 95.0	85.3	78.4, 90.4	66.7	53.9, 77.5	95.5	90.1, 98.2	6.00	4.02, 8.95	0.14	0.07, 0.30
Whole sample (nurse)	87.5	81.9, 91.6	90.0	77.4, 96.3	86.7	79.9, 91.5	69.2	56.4, 79.8	96.3	91.1, 98.6	6.75	4.44, 10.26	0.12	0.05, 0.27
Dementia subgroup (physician)	82.4	71.5, 90.0	91.2	75.2, 97.7	75.0	58.5, 86.8	75.6	59.4, 87.1	90.9	74.5, 97.6	3.65	2.11, 6.30	0.12	0.04, 0.35
Dementia subgroup (nurse)	83.8	73.0, 91.0	91.2	75.2, 97.7	77.5	61.1, 88.6	77.5	61.1, 88.6	91.2	75.2, 97.7	4.05	2.26, 7.27	0.11	0.04, 0.34
DDT-Pro by a physician or by a nurse versus DRS-R-98 delirium														
Whole sample (physician)	83.0	76.9, 87.8	90.0	75.4, 96.7	81.2	74.1, 86.8	54.5	41.9, 66.7	97.0	92.1, 99.0	4.80	3.42, 6.73	0.12	0.05, 0.31
Whole sample (nurse)	86.5	80.8, 90.8	97.5	85.3, 99.9	83.7	76.9, 88.9	60.0	47.1, 71.7	99.3	95.3, 100	6.00	4.21, 8.56	0.03	0.00, 0.21
Dementia subgroup (physician)	77.0	65.5, 85.7	92.9	75.0, 98.7	67.4	51.9, 80.0	63.4	46.9, 77.4	93.9	78.4, 98.9	2.85	1.86, 4.37	0.11	0.03, 0.41
Dementia subgroup (nurse)	83.8	73.0, 91.0	100	85.0, 100	73.9	58.6, 85.2	70.0	53.3, 82.9	100	87.4, 100	3.83	2.36, 6.23	0.00	— ^a
CAM-A by a consultation-liaison psychiatrist versus DSM-5 delirium														
Whole sample	87.5	81.9, 91.6	64.0	49.1, 76.7	95.3	90.2, 97.9	82.0	65.9, 91.9	88.8	82.7, 93.1	13.71	6.46, 29.11	0.38	0.26, 0.55
Dementia subgroup	77.0	65.5, 85.7	61.8	43.6, 77.3	90.0	75.4, 96.7	84.0	63.1, 94.7	73.5	58.7, 84.6	6.18	2.35, 16.24	0.42	0.27, 0.66
CAM-A by a consultation-liaison psychiatrist versus DRS-R-98 delirium														
Whole sample	88.5	83.0, 92.4	70.0	53.3, 82.9	93.1	87.5, 96.3	71.8	54.9, 84.4	92.5	87.0, 95.9	10.18	5.56, 18.65	0.32	0.20, 0.52
Dementia subgroup	77.0	65.5, 85.7	64.3	44.1, 80.7	84.8	70.5, 93.2	72.0	50.4, 87.1	79.6	65.2, 89.3	4.22	2.02, 8.82	0.42	0.25, 0.70

^a No confidence interval is reported as a result of the 100% sensitivity.

DDT-Pro's performance data for PPVs, NPVs, and LRs were consistent with its design to maximize a provisional delirium diagnosis among patients who actually had delirium. The low number of false negative test results (with high percentage of sensitivity) also increased the likelihood of excluding patients without delirium with negative test results (true negatives), where its NPV was >90%, and the -LR was close to 0.1. In contrast to the DDT-Pro, NPVs of the CAM-A were lower, and the -LR was around 0.3–0.4 (as a rule of thumb for the understanding of LRs, +LRs >5 and -LRs <0.2 are considered good) (42).

Overall, our data suggest that the DDT-Pro administered by a physician or nurse would be more useful than the CAM-A to provisionally diagnose delirium, as a result of its performance characteristics. Furthermore, it has advantages over the CAM-A for use in elderly populations, among whom comorbid dementia is common. More research is needed before we can better determine the DDT-Pro performance specifications as a provisional diagnostic tool in clinical settings other than internal medicine wards and acute brain injury rehabilitation services, where it has been validated. Very brief screening tools, such as the 4AT or a single question, may be useful for triage in acute medical settings, where delirium prevalence is high and false positives from other psychiatric disorders is less likely, or in emergency departments, where time is of the essence, but such tools should not be relied on for provisional delirium diagnosis. A variety of nurse screening tools have not been as well studied as the 4AT, and their content is less focused on core delirium characteristics.

As the follow-up DDT-Pro scores revealed for 18 patients with delirium or with comorbid dementia, its continuous value scoring approach allows its use in repeated evaluation of patients. Additionally, another possible advantage of quantitative scoring could be for evaluation of subsyndromal delirium, in which core domain symptoms are present but at subthreshold intensity (43). Future studies might delineate different cutoff values for subsyndromal delirium using the DDT-Pro.

This study has several limitations. The low mean years of education of the study sample imposes restrictions with regard to the generalization of our data. However, our cutoff score was the same as that in the Kean et al. (30) study, in which U.S. patients had more than twice as many years of education, which suggests that the DDT-Pro is suitable across a wide range of education and in different cultures. IQCODE assessments were performed by two different researchers, which would have affected the prevalence of dementia; however, the IQCODE has shown good interrater reliability (44). Because of the natural fluctuation of delirium symptom severity throughout a 12- to 24-hour period, this contributes variability among measurements made, even though they occurred during day 1 of the study. Patient fatigue could also have influenced some results.

Because delirium severity can fluctuate over a 24-hour period, it is recommended that for tools such as the DRS-R-98,

symptoms used for ratings should include the preceding 12–24 hours in addition to the information gleaned at the interview. The DDT-Pro sleep-wake cycle item uses information for the preceding 24 hours, whereas its other two items rate the contemporaneous performance during the interview. Therefore, a potential limitation is that the measure might insufficiently capture the most severe symptoms for two of three items if administered when symptom severity is milder, thereby potentially producing a lower score that does not meet the cutoff score for delirium diagnosis (a false negative). However, our ROC analyses revealed a high sensitivity for the DDT-Pro versus reference standards, both of which assessed symptoms for up to 24 hours, consistent with false negatives not being a serious limitation.

In conclusion, we found excellent reliability and validity for the DDT-Pro in this large, medical inpatient study sample, which was less homogeneous than the original TBI cohort. High accuracies and performance characteristics may relate to the tool's design: anchored item ratings and objective scoring on the basis of previously validated delirium scales, straightforward administration, no need for training or interview expertise on delirium phenomenological characteristics, brevity, and items representing the three core domains of delirium. Encouraging are the confirmation of the original cutoff score and its high performance for non-physician clinician assessors, which is its intended purpose for ease and accuracy of routine delirium detection and provisional diagnosis. Additionally, the DDT-Pro's high sensitivity among dementia patients suggests utility in real-life clinical settings with older patients, an advantage over existing tools. Its highly structured, quantitative items may enhance its performance over less-structured brief assessment tools, such as the CAM-A. Further validation in other settings is warranted, such as in emergency departments, nursing homes, and units for critically ill and postsurgical patients, where the severity of the clinical state of the patient, the resource availability, and the environment are different than that of the present study.

AUTHOR AND ARTICLE INFORMATION

Grupo de Investigación en Psiquiatría de Enlace (GIPE), Facultad de Medicina, Universidad Pontificia Bolivariana, Medellín, Colombia (Franco, Ocampo, Velásquez-Tirado, Giraldo, Serna, López, Zuluaga); Grupo de Investigación en Cuidado, Facultad de Enfermería, Universidad Pontificia Bolivariana, Medellín, Colombia (Zaraza); Hospital Psiquiátrico Universitario Institut Pere Mata, IISPV, Universitat Rovira i Virgili, Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Reus, Tarragona, Spain (Sepúlveda); the Departments of Population Health Sciences and Communication Sciences and Disorders, VA Salt Lake City Health Care System, and Health System Innovation and Research, University of Utah, Salt Lake City (Kean); and the Department of Psychiatry, Indiana University School of Medicine, Indianapolis (Trzepacz).

Send correspondence to Dr. Franco (josefranco11@hotmail.com).

Supported by the Centro de Investigación para El Desarrollo y la Innovación from the Universidad Pontificia Bolivariana, Medellín, Colombia.

Drs. Kean and Trzepacz are copyright holders for the Delirium Diagnostic Tool-Provisional. Dr. Trzepacz holds the copyright for the Delirium Rating

Scale-Revised-98, and she is President and CEO of conXiome, LLC, a privately held biotech consulting company that is without direct relationship to this study. All other authors report no financial relationships with commercial interests.

Received November 10, 2018; revision received February 11, 2019; accepted March 5, 2019; published online Oct. 30, 2019.

REFERENCES

- Ryan DJ, O'Regan NA, Caoimh RO, et al: Delirium in an adult acute hospital population: predictors, prevalence and detection. *BMJ Open* 2013; 3:e001772
- Ahmed S, Leurent B, Sampson EL: Risk factors for incident delirium among older people in acute hospital medical units: a systematic review and meta-analysis. *Age Ageing* 2014; 43:326-333
- Ferro JM, Caeiro L, Verdelho A: Delirium in acute stroke. *Curr Opin Neurol* 2002; 15:51-55
- Trzepacz P, Meagher DJ, Leonard M: Delirium, in *American Psychiatric Publishing Textbook of Psychosomatic Medicine*. Edited by Levenson J. Washington, DC, American Psychiatric Publishing, 2011
- Fong TG, Jones RN, Shi P, et al: Delirium accelerates cognitive decline in Alzheimer disease. *Neurology* 2009; 72:1570-1575
- Ritter SRF, Cardoso AF, Lins MMP, et al: Underdiagnosis of delirium in the elderly in acute care hospital settings: lessons not learned. *Psychogeriatrics* 2018; 18:268-275
- Hustey FM, Meldron SW, Smith MD, et al: The effect of mental status screening on the care of elderly emergency department patients. *Ann Emerg Med* 2003; 41:678-684
- Farrell KR, Ganzini L: Misdiagnosing delirium as depression in medically ill elderly patients. *Arch Intern Med* 1995; 155:2459-2464
- Sepulveda E, Leonard M, Franco JG, et al: Subsyndromal delirium compared with delirium, dementia, and subjects without delirium or dementia in elderly general hospital admissions and nursing home residents. *Alzheimers Dement (Amst)* 2016; 7:1-10
- Trzepacz PT: Update on the neuropathogenesis of delirium. *Dement Geriatr Cogn Disord* 1999; 10:330-334
- Choi S-H, Lee H, Chung T-S, et al: Neural network functional connectivity during and after an episode of delirium. *Am J Psychiatry* 2012; 169:498-507
- Leonard M, Donnelly S, Conroy M, et al: Phenomenological and neuropsychological profile across motor variants of delirium in a palliative-care unit. *J Neuropsychiatry Clin Neurosci* 2011; 23:180-188
- Inouye SK, van Dyck CH, Alessi CA, et al: Clarifying confusion: the Confusion Assessment Method: a new method for detection of delirium. *Ann Intern Med* 1990; 113:941-948
- Wei LA, Fearing MA, Sternberg EJ, et al: The Confusion Assessment Method: a systematic review of current usage. *J Am Geriatr Soc* 2008; 56:823-830
- Rolfson DB, McElhane JE, Jhangri GS, et al: Validity of the Confusion Assessment Method in detecting postoperative delirium in the elderly. *Int Psychogeriatr* 1999; 11:431-438
- Adamis D, Rooney S, Meagher D, et al: A comparison of delirium diagnosis in elderly medical inpatients using the CAM, DRS-R98, DSM-IV and DSM-5 criteria. *Int Psychogeriatr* 2015; 27:883-889
- Velásquez Gaviria LM: Instrumentos para el diagnóstico de delirium en hispanohablantes: artículo de revisión. *Med UPB* 2016; 35:100-110
- Hall RJ, Meagher DJ, MacLulich AM: Delirium detection and monitoring outside the ICU. *Best Pract Res Clin Anaesthesiol* 2012; 26:367-383
- Hendry K, Quinn TJ, Evans J, et al: Evaluation of delirium screening tools in geriatric medical inpatients: a diagnostic test accuracy study. *Age Ageing* 2016; 45:832-837
- Sands MB, Dantoc BP, Hartshorn A, et al: Single Question in Delirium (SQiD): testing its efficacy against psychiatrist interview, the Confusion Assessment Method and the Memorial Delirium Assessment Scale. *Palliat Med* 2010; 24:561-565
- Gaudreau JD, Gagnon P, Harel F, et al: Fast, systematic, and continuous delirium assessment in hospitalized patients: the Nursing Delirium Screening Scale. *J Pain Symptom Manage* 2005; 29:368-375
- Schuermans MJ, Shortridge-Baggett LM, Duursma SA: The Delirium Observation Screening Scale: a screening instrument for delirium. *Res Theory Nurs Pract* 2003; 17:31-50
- Gemert van LA, Schuurmans MJ: The Neecham Confusion Scale and the Delirium Observation Screening Scale: capacity to discriminate and ease of use in clinical practice. *BMC Nurs* 2007; 6:3
- de Jonghe JF, Kalisvaart KJ, Timmers JF, et al: Delirium-O-Meter: a nurses' rating scale for monitoring delirium severity in geriatric patients. *Int J Geriatr Psychiatry* 2005; 20:1158-1166
- Meagher DJ, Moran M, Raju B, et al: Phenomenology of delirium: assessment of 100 adult cases using standardised measures. *Br J Psychiatry* 2007; 190:135-141
- Meagher D, Adamis D, Trzepacz P, et al: Features of subsyndromal and persistent delirium. *Br J Psychiatry* 2012; 200:37-44
- Thurber S, Kishi Y, Trzepacz PT, et al: Confirmatory factor analysis of the Delirium Rating Scale Revised-98 (DRS-R98). *J Neuropsychiatry Clin Neurosci* 2015; 27:e122-e127
- Trzepacz PT, Meagher DJ, Franco JG: Comparison of diagnostic classification systems for delirium with new research criteria that incorporate the three core domains. *J Psychosom Res* 2016; 84:60-68
- Trzepacz PT, Franco JG, Meagher DJ, et al: Delirium phenotype by age and sex in a pooled data set of adult patients. *J Neuropsychiatry Clin Neurosci* 2018; 30:294-301
- Kean J, Trzepacz PT, Murray LL, et al: Initial validation of a brief provisional diagnostic scale for delirium. *Brain Inj* 2010; 24:1222-1230
- Cohen JF, Korevaar DA, Altman DG, et al: STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016; 6:e012799
- Berkman LF, Leo-Summers L, Horwitz RI: Emotional support and survival after myocardial infarction: a prospective, population-based study of the elderly. *Ann Intern Med* 1992; 117:1003-1009
- Morales González JM, González-Montalvo JI, Del Ser Quijano T, et al: Validation of the S-IQCODE: the Spanish version of the Informant Questionnaire on Cognitive Decline in the Elderly. *Arch Neurobiol (Madr)* 1992; 55:262-266
- Trzepacz PT, Mittal D, Torres R, et al: 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; 13:229-242
- Sepulveda E, Franco JG, Trzepacz PT, et al: Performance of the Delirium Rating Scale-Revised-98 against different delirium diagnostic criteria in a population with a high prevalence of dementia. *Psychosomatics* 2015; 56:530-541
- González M, de Pablo J, Fuente E, et al: Instrument for detection of delirium in general hospitals: adaptation of the Confusion Assessment Method. *Psychosomatics* 2004; 45:426-431
- Inouye S: *The Confusion Assessment Method (CAM): training manual and coding guide*. Boston, Hospital Elder Life Program, 2003. <https://www.hospitalelderlifeprogram.org/delirium-instrumentsbook>
- American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. Washington, DC, American Psychiatric Publishing, 2013
- Hart RP, Levenson JL, Sessler CN, et al: Validation of a cognitive test for delirium in medical ICU patients. *Psychosomatics* 1996; 37:533-546
- Wild D, Grove A, Martin M, et al: Principles of good practice for the translation and cultural adaptation process for patient-reported

- outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005; 8:94–104
41. Mukaka MM: Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012; 24:69–71
 42. Hayden SR, Brown MD: Likelihood ratio: a powerful tool for incorporating the results of a diagnostic test into clinical decision-making. *Ann Emerg Med* 1999; 33:575–580
 43. Trzepacz PT, Franco JG, Meagher DJ, et al: Phenotype of subsyndromal delirium using pooled multicultural Delirium Rating Scale–Revised-98 data. *J Psychosom Res* 2012; 73:10–17
 44. Jorm AF: The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): a review. *Int Psychogeriatr* 2004; 16:275–293