REPORTS OF ORIGINAL INVESTIGATIONS

Three-Minute Diagnostic Assessment for Delirium using the Confusion Assessment Method (3D-CAM): French translation and cultural adaptation

Évaluation diagnostique de l'état confusionnel de l'instrument 3D-CAM: traduction française et adaptation culturelle

**John G. Gaudet, MD · Corey Kull, MD · Marc L. Eskenazi, BS · John Diaper, RN ·
Julien Maillard, MD · Florence Mollard, RN · Christophe Marti, MD ·
Edward R. Marcantonio, MD · Delphine S. Courvoisier, PhD · Bernhard Walder, MD**

Abstract

Purpose The routine use of validated diagnostic instruments is key to identifying delirious patients early and expediting care. The 3-Minute Diagnostic Assessment for Delirium using the Confusion Assessment Method (3D-CAM) instrument is a brief, easy to use, sensitive, and specific delirium assessment tool for hospitalized patients. We aimed to translate the original English version into French, and then adapt it to older high-risk patients.

Methods Translation and adaptation of the questionnaire were guided by an expert committee and the 3D-CAM instrument developer. During the translation phase, we achieved semantic and conceptual equivalence of the instrument by conducting forward and backward translations. During the adaptation phase, we assessed the face validity, clarity of wording, and ease of use of the translated questionnaire by administering it to 30 patients and their caregivers in peri-interventional and medical intermediate care units. During both phases, we used qualitative (goal and adequacy of the questionnaire) and quantitative (Sperber score, clarity score) criteria.

Results Translation: four items were judged inadequate and were revised until all reached a Sperber score of < 3/7. Face validity: 91% of patients thought the questionnaire was designed to assess memory, thoughts, or reasoning. Clarity: eight items required adjustments until all scored ≥ 9/10 for clarity. Ease of use: all bedside caregivers reported that the questionnaire was easy to complete after receiving brief instructions.

Conclusions We produced a culturally adapted French version of the 3D-CAM instrument that is well understood and well-received by older high-risk patients and their caregivers.

J. G. Gaudet, MD · C. Kull, MD · M. L. Eskenazi, BS ·
J. Diaper, RN · J. Maillard, MD · F. Mollard, RN ·
B. Walder, MD (✉)
Department of Anesthesiology, Clinical Pharmacology,
Intensive Care and Emergency Medicine, Division of
Anesthesiology, Geneva University Hospital, 4 Rue Gabrielle-
Perret-Gentil, 1205 Geneva, Switzerland
e-mail: bernhard.walder@hcuge.ch

C. Marti, MD
Department of Internal Medicine, Geneva University Hospital,
Geneva, Switzerland

E. R. Marcantonio, MD
Department of Medicine, Beth Israel Deaconess Medical Center,
Boston, MA, USA

D. S. Courvoisier, PhD
Quality of Care Division, Geneva University Hospital, Geneva,
Switzerland

Résumé

Objectif L'administration systématique d'instruments diagnostiques validés est essentielle pour identifier précocement les patients confus. Le questionnaire 3D-CAM (3 Minute Diagnostic Confusion Assessment Method) est un outil d'évaluation bref, facile à administrer en milieu hospitalier, sensible et spécifique pour l'état confusionnel. Notre objectif était de le traduire en français, puis de l'adapter à une population de patients âgés à haut risque.

Méthode La traduction et l'adaptation ont été guidées par un comité d'experts et le développeur de l'instrument. Nous

avons atteint une équivalence sémantique et conceptuelle en menant des traductions antérogrades, puis rétrogrades. Nous avons évalué la validité de contenu, la clarté lexicale, et la facilité d'administration du questionnaire en le soumettant à 30 patients et 30 soignants dans des unités de soins intermédiaires médicaux et péri-interventionnels. Durant les phases de traduction et d'adaptation, nous avons utilisé des critères qualitatifs et quantitatifs.

Résultats *Traduction : quatre questions ont été jugées inadéquates et ont été révisées pour atteindre un score de Sperber < 3/7. Validité de contenu : 91% des patients pensaient que le questionnaire était conçu pour évaluer la mémoire, les pensées, ou le raisonnement. Clarté : huit questions ont dû être modifiées pour atteindre un score de clarté ≥ 9/10. Facilité d'administration : tous les soignants pensaient que le questionnaire était facile à utiliser après une brève formation.*

Conclusions *Nous avons produit une version française du questionnaire 3D-CAM qui est adaptée aux patients âgés à haut risque et aux soignants en milieu de soins aigus.*

Keywords 3D-CAM · adaptation · delirium · diagnosis · French · surveys and questionnaire · translation

Delirium is a neuropsychological syndrome characterized by a sudden disturbance in attention and awareness. This syndrome typically presents with acute cognitive and/or perceptive alterations that are not best explained by a pre-existing or evolving neurodegenerative disorder.[1] Delirium occurs in up to 50% of older patients who require acute hospital care in the context of major illness or following major surgery.[2,3] Despite being a strong predictor of adverse events[4] and cognitive decline[5] resulting in longer hospital stays,[6] loss of autonomy,[7] and increased mortality,[8] delirium remains under-recognized.[9] The routine use of validated instruments is key to the early identification of delirious patients.[10] Since its validation in 1990,[11] the Confusion Assessment Method (CAM) has become the most widely used diagnostic instrument in clinical practice; however, it is time consuming and requires specialized training.[12] The 3-Minute Diagnostic Assessment for Delirium using the Confusion Assessment Method (3D-CAM) was designed and validated in 2014 to facilitate routine delirium diagnosis by bedside caregivers, irrespective of their level of expertise in the field.[13] The instrument is divided in two sections: ten questions administered to the patient directly, followed by ten to 12 questions to the caregiver. The original English version of the 3D-CAM instrument showed excellent psychometric properties (sensitivity, 95%; specificity, 94%) in a cohort of 201 older patients with and without dementia. Minimally trained clinical staff needed only three minutes on average to complete the assessment. A brief and easy to use questionnaire, 3D-CAM is well suited to postoperative or medical high-risk patients.[14]

Since 2014, the 3D-CAM has been translated and adapted to German, Italian, Danish, Spanish, Portuguese, Malay, Chinese, Japanese, Polish, Thai, and Turkish. To this day, it remains unavailable in French. We translated the 3D-CAM into French and optimized its adaptation (face validity, ease of use, identification of factors associated with lack of clarity) to older high-risk patients and caregivers in perioperative and medical intermediate care units.

Methods

After obtaining authorization to use the 3D-CAM instrument for research purposes from the instrument developer (Dr. Edward R. Marcantonio, Beth Israel Deaconess Medical Center, Boston, MA, USA), the local Research Ethics Board (Cantonal Research Ethics Commission, Geneva, Switzerland) approved the protocol on 11 July 2018 (Swissethics ID 2018-00211).

Basic principles

We used principles of good practice for the translation and cultural adaptation process for patient-reported outcomes measures issued by the International Society for Patient Outcomes Research (ISPOR)[15] to produce a culturally adapted French version of the 3D-CAM instrument.

Translation process

We adopted a multistep multidisciplinary approach involving a centralized review process[16] (Fig. 1, upper half). To achieve semantic and conceptual equivalence of the translated instrument, four bilingual clinicians used to working with older inpatients conducted forward and backward translations. Translators worked independently of each other. They were instructed to favour conceptual rather than literal translations, strive for simplicity and clarity, and use language for a broad audience. All translations were assessed and reconciled by a panel of experts composed of a researcher experienced in instrument development, two acute care physicians, and an acute care nurse.

The instrument developer formally reviewed the backward translation using the Sperber rating scale (1 is best agreement; 7 is worst agreement).[17] Using this scale, any item with a semantic and/or conceptual score ≥ 3/7 was identified as inadequate, revised by the panel of
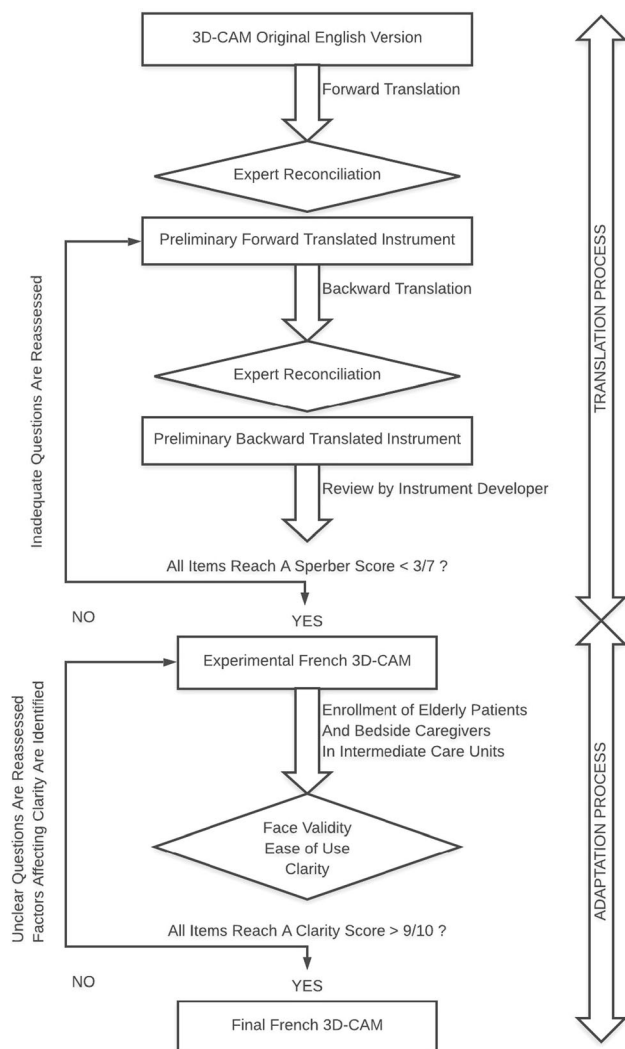
Fig. 1 3-Minute Diagnostic Assessment for Delirium using the Confusion Assessment Method (3D-CAM) French translation and cultural adaptation process. Both the translation (upper half) and adaptation (lower half) processes relied on qualitative (expert panel reconciliation, face validity assessment by patients, ease of use assessment by caregivers) and quantitative (Sperber score, clarity score) criteria

experts, then resubmitted to the instrument developer until all items scored < 3/7.

Cultural adaptation process and setting

Once translation was completed, we prospectively assessed the face validity, ease of use, and clarity of wording of the experimental French 3D-CAM in two intermediate care units (Fig. 1, lower half). We also aimed to identify demographic, cultural, or medical factors that may affect clarity.[18] All cultural adaptation assessments were conducted in the accredited, anesthesia-led peri-interventional (perioperative) or medical intermediate

care (high dependency) units at Geneva University Hospital (Geneva, Switzerland) between 9 January 2019 and 10 October 2019.

Patient selection for cultural adaptation

We approached patients aged 65 yr and older who were hospitalized in peri-interventional or medical intermediate care units for at least 24 hr. We excluded patients who (1) could not communicate in French, (2) could not give informed consent, (3) were visually impaired, or (4) were diagnosed with a terminal condition. Since delirium poses significant ethical challenges for the informed consent process,[18] we excluded patients identified as delirious at the time of assessment. Nevertheless, we did not exclude those with cognitive deficits able to give direct or surrogate written consent. All eligible patients were informed that their help was needed to ascertain whether the questionnaire is understandable and acceptable. They were encouraged to make comments during the assessment.

Selection and role of caregivers during cultural adaptation

For questions 11A through to 22 (observations by caregivers), we approached a different intermediate care physician or nurse for each patient enrolled in the study. The data set thus contained as many caregivers as patients. We collected information on their professional experience (years of practice) and first language, and assessed ease of use by asking whether they thought the instrument was adequate in an acute care setting.

Face validity of the instrument

To assess the extent to which the instrument appears to measure what it claims to measure based on face value, we asked patients what they thought the questionnaire was designed to test, then grouped their answers by keywords.

Clarity of the instrument

We asked all participants (patients and caregivers) to grade the wording and phrasing of each item on a ten-point Likert scale (0: the question is not clear at all; 10: the question is very clear). Any time participants did not find a question very clear (score < 10 points), we inquired about problematic words or expressions that may need adjustment. The first round of 20 patients was designed to identify inadequate questions as any item with a mean clarity score < 9/10. Each following round of ten patients was designed to test successive revised versions of the questionnaire until all items scored ≥ 9/10 for clarity. We

anticipated the process of cultural adaptation would require at least one round of revision.

Patients' health status and education level

We assessed physical health, mental health, and level of education to identify factors that may affect clarity. We used an American Society of Anesthesiologists Physical Status classification (ASA) of > 3/5 to identify patients with a severe systemic disease that is a threat to life (unstable condition).[19] We used a Clinical Frailty Scale (CFS) of ≥ 5/9 to identify frail patients.[20] We used a Mini-Cog score of < 3/5 to identify patients more likely to have dementia.[21] We used a Two-Question Depression Screening test (2QDS) score of ≥ 1/2 to identify patients more likely to have a depressive disorder.[22] We used the 2011 International Standard Classification of Education (ISCED) scale to distinguish patients with a low level (< 4/8) or a high level (≥ 4/8) of education.[23] Additionally, we asked all participants (patients and caregivers) which language they felt more comfortable with to distinguish native from non-native French speakers.

Role of the instrument developer

The final translated and culturally adapted French version of the 3D-CAM instrument was submitted to the original instrument developer for approval.

Statistical analyses

Compared with psychometric validation studies that require large sample sizes, 30 to 50 participants are typically necessary to translate and adapt instruments to a different culture.[24–26] We intended to include at least 30 patients and 30 caregivers to maximize the power to detect potential problems during pretest experiments, as previously recommended.[27] We used mean with standard deviation (SD) or median with interquartile range [IQR] to summarize continuous variables. We used count with percentage to describe categorical variables. We used Fisher's exact test to study whether severe systemic disease (ASA > 3/5), frailty (CFS ≥ 5/9), dementia (Mini-Cog < 3/5), depression (2QDS ≥ 1/2), or low educational level (ISCED < 4/8) are associated with lack of clarity (mean clarity score < 9/10). We also examined whether non-native French speakers struggled with the questionnaire, leading to negative association with clarity scores. Unadjusted P values < 0.05 were considered statistically meaningful. We used Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) and R 3.6.2 (R Core Team 2020, R Foundation for Statistical Computing, Vienna, Austria) to store and analyze all data.

Results

We enrolled a total of 30 patients with a mean (SD) age of 76.6 (7.4) yr and 30 caregivers (2/3 were nurses) for the adaptation process (Table 1).

Translation process

The expert panel made several incremental adjustments to the forward and backward translations before preliminary work was submitted to the instrument developer. Four items were judged inadequate (Sperber score ≥ 3/7) by the original instrument developer. They were revised until all reached a score < 3/7 for both language and interpretation (Table 2).

Face validity of the instrument

At the end of the questionnaire, 21/23 (91%) patients thought the scale was designed to assess "memory", "thoughts", or "reasoning". Two patients thought the scale was designed for other, noncognitive purposes. Seven patients opted out of summarizing what the instrument is designed to assess, mostly because of fatigue.

Ease of use of the instrument

All caregivers (100%) thought the instrument was adequate in an acute care setting.

Clarity of the instrument

After 20 questionnaires were completed (first round), five items (Q4, Q8, Q10, Q11B, and Q14) were judged inadequate (clarity score < 9/10). Additionally, three items (Q5, Q9, and Q12) included terms or expressions that were judged ambiguous. Using comments made by participants, we tested a revised version of the questionnaire in a separate group of ten patients and ten caregivers (second round). All questionnaire items reached clarity scores ≥ 9/10 (Table 3).

Identification of patient factors that may affect clarity

We focused on questions 4, 8, and 10, which several patients found less clear in the first round. Neither medical nor sociodemographic characteristics were associated with

Table 1 Patients and caregivers' characteristics during the cultural adaptation phase

	First round N = 20	Second round N = 10
Patient data		
Time of assessment (days), median [IQR]	2 [1-3]	3 [1-3]
Length of hospitalization (days), median [IQR]	13.5 [8.5–26.5]	13.0 [11.0–22.0]
Female sex, *n*/total *N* (%)	7/20 (35%)	6/10 (60%)
Age (yr), mean (SD)	76.2 (6.3)	77.4 (9.5)
ASA physical status score > 3/5, *n*/total *N* (%)	7/20 (38.9%)	3/10 (30.0%)
CFS score ≥ 5/9, *n*/total *N* (%)	4/20 (20.0%)	3/10 (30%)
Mini-Cog score < 3/5, *n*/total *N* (%)	14/20 (70.0%)	5/10 (50.0%)
2QDS score ≥ 1/2, *n*/total *N* (%)	9/20 (45.0%)	7/10 (70.0%)
2011 ISCED level < 4/8, *n*/total *N* (%)	6/20 (30.0%)	3/10 (30.0%)
French as a first language, *n*/total *N* (%)	14/20 (70.0%)	8/10 (80.0%)
Caregiver data		
Nurse caregiver, *n*/total *N* (%)	15/20 (75.0%)	5/10 (50.0%)
Professional experience (years), median [IQR]	13.0 [10.0–22.0]	10.0 [7.0–16.0]
French as a first language, *n*/total *N* (%)	18/20 (90.0%)	10/10 (100%)
Medical intermediate care unit, *n*/total *N* (%)	6/20 (30.0%)	2/10 (20.0%)

2QDS: Two-Question Depression Screen; ASA = American Society of Anesthesiologists; CFS = Clinical Frailty Scale; IQR = interquartile range; ISCED: International Standard Classification of Education; SD = standard deviation

clarity judgement. Nevertheless, 91% of demented patients found question 8 unclear (score < 9/10), compared with 50% of nondemented patients ($P = 0.04$) (Fig. 2).

Approval by the Instrument Developer and access to the instrument

The final French version of the 3D-CAM instrument was approved by the original instrument developer. To access the final French version and all updated 3D-CAM material freely and without registration, go to https://americandeliriumsociety.org/cam-and-help-tools/. For further information, please contact info@americandeliriumsociety.org.

Discussion

We produced a French version of the 3D-CAM instrument that is culturally adapted to older patients and professionals in peri-interventional and medical settings. We complied with ISPOR good practice standards[15] and used both qualitative and quantitative methods to identify and address several issues during the translation and cultural adaptation phases. We found that technical and conceptual words are harder to translate. We showed that the French 3D-CAM is easy to use and has good face validity. We also showed that

cognitively impaired patients hospitalized in an acute care setting are more likely to struggle answering questions about changes over time.

With approximately 270 million speakers, French was the seventh most spoken language worldwide in 2020.[28] Once validated, implementation of the French 3D-CAM will facilitate the identification of delirious patients hospitalized in French-speaking areas.[29,30] Our results are particularly relevant in acute monitored care settings, such as intermediate care units, where more than one in five patients develop delirium on average across medical and surgical specialties.[31,32] Compared with the standard CAM, which is available in French, the 3D-CAM is much shorter to administer at the bedside. All 30 caregivers involved in our study found the 3D-CAM to be adapted to their setting. Compared with the CAM-ICU or the ICDSC, the 3D-CAM was designed to assess noncritically ill patients[33] and is therefore adapted to a broader population of patients. Contrary to screening tools such as the Nu-DESC or the 4AT that typically require further testing when positive, the 3D-CAM is a diagnostic instrument that identifies delirious patients who should receive adequate care promptly.[34]

We used robust, ISPOR-compliant methods to achieve crosscultural adaptation, which is an essential step towards fully validating any questionnaire in a foreign language.[26] Unfortunately, the quality and methodological approaches of crosscultural studies have been shown to vary greatly,

Table 2 Quantitative assessment of translation versions

Item	Preliminary version Sperber scores*			Amended version Sperber scores*	
	Lg	Ip	Problematic word or concept	Lg	Ip
1	2	1	None	-	-
2	2	1	None	-	-
3	2	1	None	-	-
4	2	1	None	-	-
5	1	1	None	-	-
6	2	1	None	-	-
7	2	1	None	-	-
8	2	2	None	-	-
9	2	1	None	-	-
10	2	2	None	-	-
11A	2	2	None	-	-
11B	2	3	Inadequate distinction between different levels of consciousness	2	1
12	3	2	The word "hypervigilance" should be retained	1	2
13	2	2	None	-	-
14	4	3	The translation does not capture the scope of the question	2	2
15	2	2	None	-	-
16	2	2	None	-	-
17	2	2	None	-	-
18	2	2	None	-	-
19	2	3	Unclear coding instructions (mistakes *vs* lack of attention)	2	1
20	2	2	None	-	-
21	2	2	None	-	-
22	1	1	None	-	-

*The Sperber rating scale is used to quantify the quality of translation.

Using this scale, any item with a score $\geq$ 3/7 is identified as inadequate

Ip = interpretation (conceptual score); Lg = language (semantic score)

and often depart from recommendations.[35] Although poor cultural adaptation work leads to poor validation studies and biased psychometric results,[36] translating and culturally adapting research instruments is often treated as an unimportant step of study protocols.[17]

All four questions that required additional translation work also proved more challenging to adapt culturally. Translators and several healthcare professionals struggled with the terminology used to describe levels of consciousness (questions 11B and 12) or quality of communication (question 14). Words used to describe disorders of consciousness fail to capture the complexity of the underlying neurobiology.[37] Perhaps unsurprisingly, they are often confusing, if not controversial.[38] Qualitative assessment of communication can be affected by multiple cultural and medical factors.[39] We found that replacing technical terms such as "stuporous", "hypervigilance", "verbose", and "tangential" with more

common, though less precise, wording such as "sleepy", "state of heightened alertness", "speaks a lot", or "lacks focus" achieved both translational and cultural goals. In addition to improving clarity, shorter and simpler questions are generally perceived to be better as they tend to reduce the number of missing answers.[40] Nevertheless, oversimplifications carry the risk of inducing measurement errors.[41] Such issues highlight the importance of appropriate training prior to administration of the questionnaire, as well as the availability of a clear training manual in French.[42]

We found that cognitively impaired patients (identified using the Mini-Cog test) were more likely to find question 8 unclear. Declines in cognitive function interfere with the question-answer process.[43] Question 8 asks patients to identify a change over time ("during the past day"). Although cognitively impaired patients often present with compromised ability to project themselves in time[44],

Table 3 Clarity scores during the cultural adaptation phase

Question	First round N = 20	Second round N = 10	Adjustment
1, mean (SD)	9.80 (0.62)	9.80 (0.42)	None
2, mean (SD)	9.28 (1.53)	9.20 (2.20)	None
3, mean (SD)	9.61 (1.24)	9.60 (0.84)	None
4, mean (SD)	8.61 (2.12)	9.30 (0.95)	We stressed the importance of reading the digit sequence at a rate of one per second
5, mean (SD)	9.06 (1.95)	9.50 (0.97)	We stressed the importance of reading the digit sequence at a rate of one per second
6, mean (SD)	9.56 (0.92)	9.70 (0.67)	None
7, mean (SD)	9.35 (1.06)	9.60 (0.70)	None
8, mean (SD)	8.12 (1.93)	9.80 (0.63)	1. We simplified the time reference by replacing "durant ces dernières 24 heures" with "entre hier et aujourd'hui". 2. We shortened the question by making the clarification optional
9, mean (SD)	9.24 (1.52)	9.50 (1.27)	We simplified the time reference by replacing "durant ces dernières 24 heures" with "entre hier et aujourd'hui"
10, mean (SD)	8.31 (2.24)	9.30 (1.16)	1. We simplified the time reference by replacing "durant ces dernières 24 heures" with "entre hier et aujourd'hui" 2. We avoided referring to reality and replaced "pas réelles" with "bizarres"
11A, mean (SD)	10.00 (0)	9.80 (0.63)	None
11B, mean (SD)	7.25 (2.97)	9.70 (0.67)	We replaced the word "stuporeux" with "endormi"
12, mean (SD)	9.15 (1.53)	9.40 (0.84)	We replaced the word "hypervigilance" with "état d'alerte"
13, mean (SD)	9.75 (0.79)	9.30 (1.16)	None
14, mean (SD)	7.90 (1.59)	9.90 (0.32)	We replaced the word "prolixe" with "parlait-il beaucoup"
15, mean (SD)	9.70 (0.98)	9.60 (0.97)	None
16, mean (SD)	10.00 (0)	10.00 (0)	None
17, mean (SD)	10.00 (0)	9.90 (0.32)	None
18, mean (SD)	10.00 (0)	9.70 (0.95)	None
19, mean (SD)	8.45 (1.93)	9.70 (0.95)	We simplified the coding instruction distinguishing wrong answers from inattention
20, mean (SD)	9.80 (0.62)	9.70 (0.95)	None
Instruction for optional questions 21 and 22, mean (SD)	8.10 (2.32)	9.70 (0.95)	We shortened the instruction by suggesting questions 21 and 22 should be answered whenever column 1 (time pattern) is empty, irrespective of other columns
21, mean (SD)	9.85 (0.49)	9.10 (1.10)	None
22, mean (SD)	8.00 (2.32)	9.70 (0.95)	We replaced the expression "toute évaluation positive" with "une réponse / une observation différente"

The first 10 items (clear area) are administered to the patient. The last 10–12 items (shaded area) are administered to the caregiver. Items with a mean clarity score < 9/10 and/or ambiguous terms or expressions during the first round were revised and tested in further rounds until all items reached clarity scores ≥ 9/10 and no ambiguity remained

SD = standard deviation

Marcantonio *et al.* did not encounter similar clarity issues in their original validation study.[13] Overall, they found that the 3D-CAM performs well in patients both with and without dementia (sensitivity, 96% *vs* 93%; specificity, 86% *vs* 96%). Contrary to their study that included older patients in general medicine or geriatric medicine services, we studied high risk patients admitted to intermediate care units, where symptoms of dementia may be aggravated.[45]

More specifically, noise and frequent interventions disrupt circadian rhythms and exacerbate otherwise minor timekeeping issues.[46] Our finding further reinforces the importance of multicomponent strategies to promote "chronofitness" in acute care settings.[47]

Our work has two significant limitations. First, although we used robust methods to translate the 3D-CAM instrument, all French speakers around the world do not
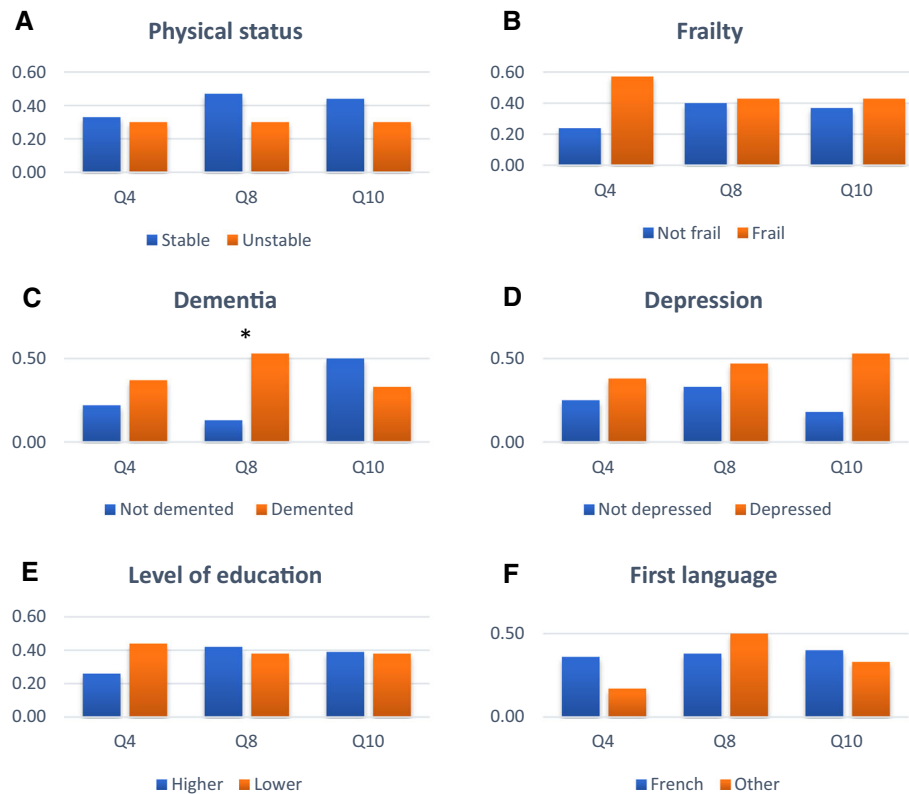
Fig. 2 Impact of language skills, health status, and education level on the proportion of patients reporting lack of clarity (y axis, higher value = lower clarity) for questions 4, 8, and 10 (x axis). Those three questions were selected because they had mean clarity scores < 9/10 during the first round of cultural adaptation. For each question, we compared patients with *vs* without (A) an unstable physical status (defined as an American Society of Anesthesiologists Physical Status classification > 3/5), (B) frailty (defined as a Clinical Frailty Score ≥ 5/9), (C) dementia (defined as a Mini-Cog score < 3/5), (D) depression (defined as Two-Question Depression Screen score ≥ 1/2), (E) a lower level of education (defined as an International Standard Classification of Education level < 4/8), and (F) French as their first language

share the same cultural background. The meaning of words can slightly fluctuate from one French speaking region to another, and local semantic or conceptual variations may affect the psychometric properties of the questionnaire.[48] The French version presented here, being produced in a single Swiss hospital with surgical and medical patients, may not be fully adapted to all 29 countries where French is listed as an official language. We encourage clinicians working in different French-speaking areas to verify the clarity of the instrument locally. Second, we produced a culturally adapted, French version of the 3D-CAM instrument but did not aim to validate it by investigating its psychometric equivalence for delirium diagnosis with the original English instrument. Our work now calls for further investigations to confirm the validity and reliability of the French 3D-CAM. Such work will require a larger sample size and the inclusion of patients with and without delirium.

Conclusion

We used robust qualitative and quantitative ISPOR-compliant methods to produce a French version of the 3D-CAM instrument that is adapted to the acute monitored care setting. Prior to clinical implementation in French-speaking areas, the French instrument must be culturally validated by showing its psychometric equivalence for delirium identification with the original English instrument.

References

1. *European Delirium Association; American Delirium Society.* The DSM-5 criteria, level of arousal and delirium diagnosis: inclusiveness is safer. BMC Med 2014; https://doi.org/10.1186/s12916-014-0141-2.

2. *Marcantonio ER.* Delirium in hospitalized older adults. N Engl J Med 2017; 377: 1456-66.

3. *Demeure MJ, Fain MJ.* The elderly surgical patient and postoperative delirium. J Am Coll Surg 2006; 203: 752-7.

4. *Raats JW, van Eijsden WA, Crolla RM, Steyerberg EW, van der Laan L.* Risk factors and outcomes for postoperative delirium after major surgery in elderly patients. PLoS One 2015; https://doi.org/10.1371/journal.pone.0136071.

5. *Austin CA, O'Gorman T, Stern E, et al.* Association between postoperative delirium and long-term cognitive function after major nonemergent surgery. JAMA Surg 2019; 154: 328-34.

6. *Gleason LJ, Schmitt EM, Kosar CM, et al.* Effect of delirium and other major complications on outcomes after elective surgery in older adults. JAMA Surg 2015; 150: 1134-40.

7. *Shi Z, Mei X, Li C, et al.* Postoperative delirium is associated with long-term decline in activities of daily living. Anesthesiology 2019; 131: 492-500.

8. *Thein MZ, Pereira JV, Nitchingham A, Caplan GA.* A call to action for delirium research: meta-analysis and regression of delirium associated mortality. BMC Geriatr 2020; https://doi.org/10.1186/s12877-020-01723-4.

9. *Ritter SR, Cardoso AF, Lins MM, Zoccoli TL, Freitas MP, Camargos EF.* Underdiagnosis of delirium in the elderly in acute care hospital settings: lessons not learned. Psychogeriatrics 2018; 18: 268-75.

10. *Tamune H, Yasugi D.* How can we identify patients with delirium in the emergency department?: A review of available screening and diagnostic tools. Am J Emerg Med 2017; 35: 1332-4.

11. *Inouye SK, Dyck CH van, 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.

12. *Wong EK, Lee JY, Surendran AS, et al.* Nursing perspectives on the confusion assessment method: a qualitative focus group study. Age Ageing 2018; 47: 880-6.

13. *Marcantonio ER, Ngo LH, O'Connor M, et al.* 3D-CAM: derivation and validation of a 3-minute diagnostic interview for CAM-defined delirium: a cross-sectional diagnostic test study. Ann Intern Med 2014; 161: 554-61.

14. *Olbert M, Eckert S, Mörgeli R, Kruppa J, Spies CD.* Validation of 3-minute diagnostic interview for CAM-defined delirium to detect postoperative delirium in the recovery room. A prospective diagnostic study. Eur J Anaesthesiol 2019; 36: 683-7.

15. *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 Aadaptation. Value Health 2005; 8: 94-104.

16. *Chávez LM, Canino G.* Toolkit on Translating and Adapting Instruments. The Evaluation HSRI Center. MA: Cambridge; 2005: 1-58. Available from URL: https://www.hsri.org/files/uploads/publications/PN54_Translating_and_Adapting.pdf (accessed December 2021).

17. *Sperber AD.* Translation and validation of study instruments for cross-cultural research. Gastroenterology 2004; 126(1 Suppl 1): S124-8.

18. *Auerswald KB, Charpentier PA, Inouye SK.* The informed consent process in older patients who developed delirium a clinical epidemiologic study. Am J Med 1997; 103: 410-8.

19. *Hackett NJ, De Oliveira GS, Jain UK, Kim JY.* ASA class is a reliable independent predictor of medical complications and mortality following surgery. Int J Surg 2015; 18: 184-90.

20. *Rockwood K, Song X, MacKnight C, et al.* A global clinical measure of fitness and frailty in elderly people. CMAJ 2005; 173: 489-95.

21. *Borson S, Scanlan JM, Chen P, Ganguli M.* The Mini-Cog as a screen for dementia: validation in a population-based sample. J Am Geriatr Soc 2003; 51: 1451-4.

22. *Tsoi KK, Chan JY, Hirai HW, Wong SY.* Comparison of diagnostic performance of two-question screen and 15 depression screening instruments for older adults: systematic review and meta-analysis. Br J Psychiatry 2017; 210: 255-60.

23. *UNESCO Institute for Statistics.* International Standard Classification of Education; 2011. Available from URL: http://uis.unesco.org/en/topic/international-standard-classification-education-isced (accessed December 2021).

24. *Sousa VD, Rojjanasrirat W.* Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. J Eval Clin Pract 2011; 17: 268-74.

25. *Hui CH, Triandis HC.* Measurement in cross-cultural psychology: a review and comparison of strategies. J Cross Cult Psychol 1985;

26. *Tsang S, Royse CF, Terkawi AS.* Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi J Anaesth 2017; 11(Suppl 1): S80-9.

27. *Perneger TV, Courvoisier DS, Hudelson PM, Gayet-Ageron A.* Sample size for pre-tests of questionnaires. Qual Life Res 2015; 24: 147-51.

28. *Wikipedia.* French language. Available from URL: https://en.wikipedia.org/wiki/French_language (accessed December 2021).

29. *Travers C, Henderson A, Graham F, Beattie E.* CogChamps: impact of a project to educate nurses about delirium and improve the quality of care for hospitalized patients with cognitive

impairment. BMC Health Serv Res 2018; https://doi.org/10.1186/s12913-018-3286-4.

30. Hasemann W, Tolson D, Godwin J, Spirig R, Frei IA, Kressig RW. A before and after study of a nurse led comprehensive delirium management programme (DemDel) for older acute care inpatients with cognitive impairment. Int J Nurs Stud 2016; 53: 27-38.

31. Schubert M, Schürch R, Boettger S, et al. A hospital-wide evaluation of delirium prevalence and outcomes in acute care patients - a cohort study. BMC Health Serv Res 2018; https://doi.org/10.1186/s12913-018-3345-x.

32. Ryan DJ, O'Regan NA, Caoimh RÓ, et al. Delirium in an adult acute hospital population: predictors, prevalence and detection. BMJ Open 2013; https://doi.org/10.1136/bmjopen-2012-001772.

33. Gusmao-Flores D, Salluh JIF, Chalhub RA, Quarantini LC. The confusion assessment method for the intensive care unit (CAM-ICU) and intensive care delirium screening checklist (ICDSC) for the diagnosis of delirium: a systematic review and meta-analysis of clinical studies. Crit Care 2012; https://doi.org/10.1186/cc11407.

34. American Psychological Association. Distinguishing Between Screening and Assessment for Mental and Behavioral Health Problems. Available from URL: https://www.apaservices.org/practice/reimbursement/billing/assessment-screening (accessed December 2021).

35. Maneesriwongul W, Dixon JK. Instrument translation process: a methods review. J Adv Nurs 2004; 48: 175-86.

36. Berkanovic E. The effect of inadequate language translation on Hispanics' responses to health surveys. Am J Public Health 1980; 70: 1273-6.

37. Zasler ND, Aloisi M, Contrada M, Formisano R. Disorders of consciousness terminology: history, evolution and future directions. Brain Injury 2019; 33: 13-4.

38. Bayne T, Hohwy J, Owen AM. Reforming the taxonomy in disorders of consciousness. Ann Neurol 2017; 82: 866-72.

39. Insua-Summerhays B, Hart A, Plummer E, Priebe S, Barnicot K. Staff and patient perspectives on therapeutic engagement during one-to-one observation. J Psychiatr Ment Health Nurs 2018; 25: 546-57.

40. Schaeffer NC, Dykema J. Questions for surveys: current trends and future directions. Public Opin Q 2011; 75: 909-61.

41. Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. Trials 2010; https://doi.org/10.1186/1745-6215-11-2.

42. Blevins CS, DeGennaro R. Educational intervention to improve delirium recognition by nurses. Am J Crit Care 2018; 27: 270-8.

43. Kutschar P, Weichbold M, Osterbrink J. Effects of age and cognitive function on data quality of standardized surveys in nursing home populations. BMC Geriatr 2019; https://doi.org/10.1186/s12877-019-1258-0.

44. Haj ME, Kapogiannis D. Time distortions in Alzheimer's disease: a systematic review and theoretical integration. Npj Aging Mech Dis 2016; https://doi.org/10.1038/npjamd.2016.16.

45. Cunningham C, Archibald C. Supporting people with dementia in acute hospital settings. Nurs Stand 2006; 20: 51-5.

46. Chan MC, Spieth PM, Quinn K, Parotto M, Zhang H, Slutsky AS. Circadian rhythms: from basic mechanisms to the intensive care unit. Crit Care Med 2012; 40: 246-53.

47. McKenna H, van der Horst GT, Reiss I, Martin D. Clinical chronobiology: a timely consideration in critical care medicine. Crit Care 2018; https://doi.org/10.1186/s13054-018-2041-x.

48. Solano-Flores G. Language, dialect, and register: sociolinguistics and the estimation of measurement error in the testing of English language learners. Teachers College Record 2006; 108: 2354-79. Available from URL: https://www.tcrecord.org/books/Content.asp?ContentID=12808 (accessed December 2021).