NICE National Institute for Health and Care Excellence

NICE

2020 exceptional surveillance of delirium: prevention, diagnosis and management (NICE guideline CG103)

Surveillance report

Published: 30 July 2020

www.nice.org.uk

Contents

Surveillance decision

We will update the NICE guideline on delirium: prevention, diagnosis and management.

The update will focus on the risk assessment and diagnosis of delirium.

Reason for the exceptional review

This exceptional review examined any impact of the National Institute for Health Research funded health technology assessment (HTA): development and validation of the 4AT: a new rapid screening tool for delirium on the NICE guideline.

Methods

The exceptional surveillance process consisted of:

- Considering the new evidence that triggered the exceptional review.

- Considering the evidence used to develop the original guideline in 2010.

- Considering the evidence identified in previous surveillance of this guideline.

- Assessing the new evidence against current recommendations to determine whether or not to update sections of the guideline, or the whole guideline.

Full updated literature searches were not needed because as the review progressed, there was sufficient information from the original guideline, routine surveillance, and the HTA publication itself to allow decision making.

For further details about the process and the possible update decisions that are available, see ensuring that published guidelines are current and accurate in developing NICE guidelines: the manual.

Information considered in this exceptional surveillance review

The guideline includes recommendations about the risk factors for delirium, indicators of delirium at presentation and daily observations, and preventing, diagnosing and treating delirium.

The guideline currently recommends a 2-stage diagnostic process for delirium:

- An initial stage based on the identification of warning signs and symptoms including an assessment of cognitive and physical function, perception, and social behaviour (recommendations based on consensus).

- A confirmatory stage based on the use of diagnostic tests (recommendations based on low-moderate quality of evidence). The DSM-IV criteria, the confusion assessment method (short version; short CAM) or CAM-ICU (for use in critical care or in the recovery room after surgery) are the confirmatory tests recommended. The long version of the CAM test (long CAM) was not recommended in the guideline because it was mainly used in research settings.

The 4 'A's test is a screening tool for delirium that can also be used to identify pre-existing cognitive impairment. The tool was developed in 2010 after a review of screening tools available in this area. The 4 'A's test was not available when the guideline was developed.

Development and validation of the 4 'A's test: a new rapid screening tool for delirium

The MacLullich study aimed to assess the usability of the 4 'A's test in clinical practice, its diagnostic accuracy and cost effectiveness.

To assess the diagnostic accuracy, a randomised diagnostic trial was conducted in 3 hospitals in the UK. People 70 years and over attending emergency departments (within 12 hours of arrival) or in acute general medical wards (within 96 hours of admission to the ward), coming by ambulance as an emergency or sent by their GP were included. Initially, the eligibility criteria were assessed using an alphabetical list with the names of the potentially eligible participants. However, a preliminary analysis of the included patients suggested that those with a low risk of delirium were more likely to be recruited. The recruitment strategy of the study changed during the recruitment period and participants

with a higher risk of delirium (elderly patients, likelihood of admission, increased number of comorbidities) were approached first, increasing their chances to be recruited.

Participants were randomised to receive the 4 'A's test or the short CAM test. All the participants underwent a reference standard delirium assessment (maximum of 2 hours interval between tests). The performance of the 4 'A's test individual items was also assessed. People with life-threating conditions, in a coma, non-English speakers, or previously enrolled in other 4 'A's test studies were excluded.

Index test and reference standard

The 4 'A's test assesses 4 items: alertness, abbreviated mental test, attention using the months backwards test, and an assessment of acute changes in mental status (informed by case notes, GP letters or obtaining data from an informant). The performance of the 4 'A's was compared with the short CAM test, which assesses 4 items: acute change and fluctuating course, inattention, disorganised thinking, and altered level of consciousness.

The reference standard was a diagnosis of delirium based on the Diagnostic and Statistical Manual of Mental Disorders IV (DSM-IV) criteria with a final verification by clinical consensus of 3 evaluators. Results from a Delirium Rating Scale-Revised-98 (DRS-R98) test and other neuropsychological assessments were used to inform the DSM-IV criteria. The informant questionnaire for cognitive decline in the elderly was used to assess pre-existing cognitive impairment.

All the evaluators were nurses or clinical research associates, they received specific training in delirium and in how to administer the different tests assessed. They were blinded to the results of the other tests.

The main outcomes evaluated were the diagnostic accuracy of the test, length of hospital stay, new institutionalisation, and mortality at 12 weeks.

Results

A total of 4,928 people were eligible to participate in the study, 843 were randomised and 785 were included in the data analysis (395 in the 4 'A's test group and 390 in the short CAM test group). Fifty-eight participants were not included in the analysis because of missing data or because they were classified as delirium indeterminate as defined by the reference standard. The diagnosis was delirium in 12.1% according to the reference

standard, 14.3% using 4 'A's test only, and 4.8% using the short CAM test only.

For a diagnosis of delirium, the sensitivity of the 4 'A' test (score greater than 3) was 75.5% (95% confidence interval [CI] 61.1% to 86.6%), and the specificity was 96.4% (95% CI 93.8% to 98.1%). The area under the curve (AUC) was 0.89. The AUC is a global assessment of the test performance. An AUC of 0.5 suggests no discrimination (same as random) and an AUC of 1.0 excellent discrimination. The sensitivity of the short CAM test was 40% (95% CI 27% to 57%), and the specificity was 100% (95% CI 98 to 100). The individual items included in the 4 'A's test had sensitivities between 31% and 96%, and specificities between 79% and 99%; the results varied depending on the thresholds used in some of the items assessed. For severity of delirium, there was a moderate correlation between the results of the 4 'A's and the DRS-R98 total score.

People with a 4 'A's test score of more than 3 had an increased mortality risk compared with those with a score of 3 or less, but this was not statistically significant (odds ratio 2.00; 95% CI 0.85 to 4.70; p value 0.0261). They also had a greater length of hospital stay at 12 weeks compared with those with a 4 'A's test of 3 or less (hazard ratio 0.64; 95% CI 0.46 to 0.88; p value 0.0009). All these results were adjusted by age, gender and dementia status. New incidences of institutionalisation appeared to be very low in the study (high number of missing data), so no analyses were performed.

The study has some limitations, the main one related to the risk of selection bias that could affect the representativeness of the target population in the study. Only 17% of those eligible consented to participate, and the recruitment strategy changed during the recruitment period. The final number of participants included in the study was lower than initially calculated (targeted 900, participants included in the analysis 785), with low representation of minority ethnic groups. The test was available only in English language format, so applicability for non-English speaking minority groups was not assessed. The prevalence of delirium was in the lower range of what was expected (12%, expected between 10% and 20%). Differences in mortality and length of stay between the tests were not assessed, so it is unclear if the differences in the diagnostic accuracy of the tests had an impact on the relevant outcomes evaluated.

Cost-effectiveness analysis

A cost-effectiveness analysis comparing 4 'A's test with the short CAM test was conducted. The differences in terms of benefits and costs at 12 weeks were small. Sensitivity analyses suggested that the 4 'A's test is cost effective compared with the

short CAM test, but there was a high degree of uncertainty around the estimates. The authors stated that the cost-effectiveness analysis has several limitations that could affect the confidence in the results obtained. These limitations included the lack of knowledge of the delirium clinical pathway and a lack of data to inform the different model parameters, so an important part of the data used in the analysis was formally elicited by expert opinion (data on mortality, quality of life and cost associated).

Information considered when developing the guideline

The committee developing the original guideline in 2010 was interested in examining the diagnostic accuracy of different tests that could be used as screening tools to identify delirium in hospital (including intensive care units) and long-term care settings. They were looking for tests with high sensitivity that would 'rule in' patients with delirium. Most of the studies identified at the time were cross-sectional studies that assessed different tests including the abbreviated mental test, long CAM test, short CAM test, confusion assessment method, clock-drawing test, mini mental state examination (MMSE), delirium index, DRS-R-98, and chart assessment. Studies comparing tests with DSM-IV criteria, ICD-10, DSM-III-R, DSM-III, CAM and clinical review or consensus diagnosis as reference standards were considered. The mean age of the participants in most of the studies identified was above 65 years and had varying proportions of non-English speakers and people with dementia. The quality assessment of the individual studies highlighted different types of bias, including spectrum bias, disease progression bias, partial verification bias, review bias, and incorporation bias, among others.

Information considered in previous surveillance of this guideline

A subsequent evidence update (published in 2012) and 2 surveillance reviews (published in 2014 and 2018) identified evidence assessing the diagnostic accuracy of diagnostic tests for delirium including the PRE-DELIRIC tool (prediction of delirium in ICU patients), delirium scale, CAM test, CAM for intensive care units, intensive care delirium screening checklist, months backwards test, MMSE, the observation scale of level of arousal, the Richmond agitation sedation scale for postoperative delirium, the post anaesthetic recovery score, and the nursing delirium screening scale. However, it was considered that the evidence identified on individual tools was limited and unlikely to have an impact on

the current guideline recommendations.

Equalities

Minority ethnic groups were not well represented in the study. The 4 'A's test was only available in English limiting the generalisability of the results to non-English speaking minority groups. Although this issue is not specific to the 4 'A's test, it is an important equality consideration if the guideline is to be updated.

Overall decision

The guideline currently recommends the use of clinical criteria as a first step to identify people with a higher risk of delirium, and the use of DSM-IV criteria or the short CAM test to diagnose delirium. The evidence identified when developing the guideline was limited in terms of the quality and came mostly from observational studies. Evidence identified in previous surveillance of the guideline on other tests was considered unlikely to have an impact on the current guideline recommendations.

The MacLullich study demonstrated that the 4 'A's test is a useful tool for diagnosis of delirium in acute medical settings through a more robust study design. The results of the cost-effectiveness analysis showed that the 4 'A's test is a cost-effective option compared with the short CAM test, but there is a high degree of uncertainty around the economic estimates.

Considering the new evidence that shows the 4 'A's test could be used as a screening tool for delirium, alongside the evidence from the previous routine surveillance review, these data for the 4 'A's test may have an impact on the current recommendations.

ISBN: 978-1-4731-3728-8